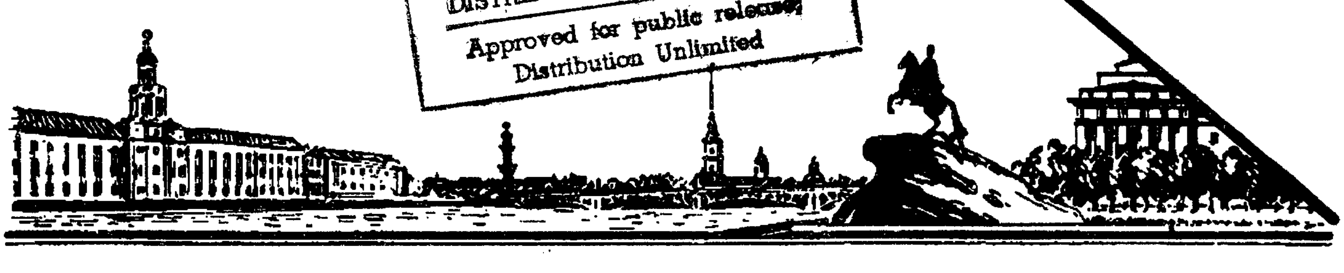
IC I&C'97

INTERNATIONAL CONFERENCE ON INFORMATICS AND CONTROL

PROCEEDINGS

June 9-13, 1997
St. Petersburg
RUSSIA



ISBN 5-85733-077-7

PUBLISHED BY ST. PETERSBURG INSTITUTE FOR INFORMATICS AND AUTOMATION OF THE RUSSIAN ACADEMY OF SCIENCES (SPIIRAS)
39, 14 Linia, VO
St. Petersburg, 199178, Russia
spiiran@iias.spb.su

COVER DESIGNED BY NIKOLAY D. GORSKY

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 12 June 1997	3. REPORT TYPE AND DATES COVERED Conference Proceedings

4. TITLE AND SUBTITLE International Conference on Informatics and Control (ICI&C'97) Proceedings Vol 1	5. FUNDING NUMBERS F6170897W0117
6. AUTHOR(S) Conference Committee	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) St.Petersburg Institute For Informatics & Automation 39, 14 Line, V.O. St.Petersburg 199178 Russia	8. PERFORMING ORGANIZATION REPORT NUMBER N/A
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 BOX 14 FPO 09499-0200	10. SPONSORING/MONITORING AGENCY REPORT NUMBER CSP 97-1025
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE A

13. ABSTRACT (Maximum 200 words)

The Final Proceedings for International Conference on Informatics and Control, 9 June 1997 - 13 June 1997

The Topics covered include: Intelligent control systems and information processing; Nueroinformatics and nueronet control; Robots' control and computer integrated technologies; Qualitative techniques for control and information processing tasks; Stochastic systems for information processing in motion control tasks; Safety control in complex systems, Transport Control (vehicles, vessels, aircraft, etc.); Satellite remote sensing in control and control related applications; Telecommunications: basics.

14. SUBJECT TERMS Control, Control System, Advanced Computing, Computers, Electronic Devices, Instrumentation, Optimization, Signal Processing, Human Factors	15. NUMBER OF PAGES 418 16. PRICE CODE N/A

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19, SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

International Conference on Informatics and Control
(ICI&C'97)

PROCEEDINGS

VOLUME 1 OF 3

St. Petersburg, Russia

International Conference on Informatics and Control (ICI&C'97),
June 9 -13, 1997,
St.Petersburg, Russia

CONFERENCE ORGANIZATION

Sponsored by:

United States Air Force European Office of Aerospace Research and Development,
United States Army research development and Standardization Group (UK),
United States Navy Office of Naval Research, Europe
Ford Forschungszentrum, Aachen, Ford Motor Company,
Russian State Committee for Science and Technologies,
Department of Informatics, Computer Engineering and Automation of the Russian Academy of Sciences
St.Petersburg Scientific Center of the Russian Academy of Sciences

Co-sponsored by:

IMACS,
Object Management Group (OMG),
IFIP (WG 5.3)

Hosted by:

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)

Organized with participation of:

Committee for Informatization Policy under the auspices of the Russian President, Department of Informatics, Computer Engineering and Automation of RAS,
Department of Problems in Mechanical Engineering, Mechanics and Control Processes of RAS,
Institute for Problems of Information Transmission of RAS,
Institute for Problems of Mechanical Engineering of RAS,
Institute of Transport Problems of RAS,
Institute of Control Sciences of RAS,
Central Research Institute "GRANIT",
Research and Production Company "AURORA",
Admiral Kuznetsov Naval Academy,
Moscow State University,
St. Petersburg State University,
St. Petersburg State Technical University,
St. Petersburg State Electrotechnical University,
St. Petersburg Institute of Fine Mechanics and Optics,
St. Petersburg Academy of Aerospace Instrumentation

Conference General Chair:

Acad. Emeljyanov S. V. (Rus)

Program co-chairmen:

Acad. Kuznetsov N. A. (Rus)
Prof. Yusupov R. M. (Rus)

International Advisory Committee:

Academician Zh. I. Alferov (Rus)
Prof. L.M. Camarinha-Matos (Portugal)
Prof. J.-P. Crestin (France)
Prof. T. Kjellberg (Sweden)
Prof. D. Kochan (Ger)
Prof. R. C. Michelini (It)
Prof. G. J. Olling (USA)
Prof. V.M. Ponomaryov (Rus)
Prof. O. I. Semenkov (Byelorus)
Prof. L. Nemes (Aust)

Members of the International Program Committee

Prof. Bely O. V. (Rus)
Dr. Bojko V. V. (Rus)
Prof. Bulatov V. P. (Rus)
Prof. Chernorutsky I. G. (Rus)
Acad. Chernousko F. L. (Rus)
Prof. Groumpos P. (Greece)
Prof. Fang Ting-Jian (China)
Acad. Fedosov E. A. (Rus)
Prof. Fradkov A. L. (Rus)
Prof. Gorodetsky V. I. (Rus)
Prof. Ivannikov V. P. (Rus)
Dr. Kortnev A. V. (Rus)
Prof. Koshevoi A. A. (Ua)
Prof. Kozlov V. N. (Rus)
Dr. Krasilnikov I. A. (Rus)
Prof. Krasnoselsky M. A. (Rus)
Acad. Krasovsky A. A. (Rus)
Prof. Krotov V. F. (Rus)
Prof. Kryzhanovsky G. A. (Rus)
Prof. Kurdyumov S. P. (Rus)
Prof. Lopota V. A. (Rus)
Prof. Lototsky V. A. (Rus)
Prof. V. A. Lyubetsky (Rus)

Prof. Madani K.	(Fr.)
Prof. Miller B. M.	(Rus)
Prof. Miroshnik I. V.	(Rus)
Prof. Mizin I. A.	(Rus)
Acad. Okhotsimsky D. E.	(Rus)
Prof. Nikoltsev V. A.	(Rus)
Prof. Pavlov B. V.	(Rus)
Prof. Perona G.	(It)
Prof. Pervozvansky A. A.	(Rus)
Prof. Polonnikov R. I.	(Rus)
Acad. Prangishvili I. V.	(Rus)
Acad. Pugachov V. S.	(Rus)
Prof. Puzankov D. V.	(Rus)
Prof. Pyatnitsky E. S.	(Rus)
Prof. Ryabinin I. A.	(Rus)
Acad. Salyga V. I.	(Rus)
Acad. Samarsky A. A.	(Rus)
Acad. Sazonov A. E.	(Rus)
Prof. Sgurev V.	(Bul)
Prof. Smirnov A. V.	(Rus)
Dr. Sobolewski M	(USA)
Prof. Solunin V. L.	(Rus)
Acad. Teryaev E. D.	(Rus)
Prof. Timofeev A. V.	(Rus)
Acad. Tsypkin Ya. Z.	(Rus)
Prof. Varzhapetyan A. G.	(Rus)
Prof. Vasiljev V. N.	(Rus)
Prof. Venets V. I.	(Rus)
Prof. Vishnevsky V. M.	(Rus)
Dr. Voitetsky V. V.	(Rus)
Acad. Vukobratovich M.	(Serb)
Prof. Yakovlev B. V.	(Rus)
Academician Zhuravlev Yu. I.(Rus)	

Conference Coordinator and Organizer:

Irina P. Podnozova	(Rus)

WELCOME ADDRESS FROM THE CONFERENCE HOST

On behalf of organizers of the International Conference on Informatics and Control (ICI&C'97) and as a director of the hosting institution: St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS) I am pleased and proud to extend a warm welcome to all participants and contributors to ICI&C'97, including the Russian Academy of Sciences (RAS), the Russian State Committee for Science and Technologies, the Russian Foundation for Basic Research, Government of the City of St. Petersburg, Committee for Informatization Policy under the auspices of the Russian President, Department of Informatics, Computer Engineering and Automation of RAS, Department of Problems in Mechanical Engineering, Mechanics and Control Processes of RAS, Institute for Problems of Information Transmission of RAS, St. Petersburg Scientific Center of RAS, St.Petersburg Institute for Informatics and Automation of RAS.

The conference consists of a large number of broad based high quality refereed papers in the areas of adaptive control, nonlinear systems of automatic control, complex systems modeling, control of moving objects, control and information technologies in ecology, nueroinformatics and nueronet control, robots control, quality of information control systems, safety control in complex systems, telecommunications and computer networks, intelligent manufacturing processes and systems, informatics and control in biology and medicine, simulation of ocean processes.

The authors come from 15 countries around the world, and represent research, industry, university and government. Through the published proceedings, you will have an opportunity to refer to the content of these papers long after the conference concluded.

As many of you might remember, in 1960 the First IFAC Congress was held in our country in Moscow, and since then for long years an attention of professional peers was continuously drawn to biannual All Union Meetings on Automatic Control. Unfortunately this tradition ceased existing. In the above regard our ability to organize and host the current Conference is a meaningful to us as a milestone of restoring a long ago established tradition to conduct professional forums in Russia in the area of control science, and of adding an informatics and computer science issues to the conventional problem scope. This Conference fosters an establishing of strong programs in research, industry, university and government cooperation as well.

We also would like to take this opportunity and express our thanks to the following for their contribution to the success of this conference. Specifically I want to appreciate Conference sponsors:

United States Air Force European Office of Aerospace Research and Development,
United States Army Research Development and Standardization Group (UK),
United States Navy Office of Naval Research, Europe
Ford Forschungszentrum, Aachen, Ford Motor Company
Russian State Committee for Science and Technologies,
Department of Informatics, Computer Engineering and Automation of the Russian Academy of Sciences
St.Petersburg Scientific Center of the Russian Academy of Sciences

for an outstanding job and sincere understanding of our current problems. Also we would like to further thank the professional organizations of IMACS, Object Management Group (OMG), IFIP and particularly IFIP WG 5.3, who have given their endorsement for this conference, the members of the program committee, advisory committee and organizing committee who have helped with the formulation of and conduct of the meeting, and extend special thanks to session chairs and organizers who contributed significant amount of their time and effort in ensuring a high quality program, and finally to Ms. Irina Podnozova whose countless hours and attention to detail made this event happen.

Welcome then again, to what we hope will be the first international conference in the renewed raw of informatics and control events in Russia. Take the opportunity to listen to the papers and to interact with your colleagues. Please feel free to make this your "professional home" for the next three days.

Welcome, learn, and enjoy.

Sincerely,

Rafael M. Yusupov
Professor, Director of SPIIRAS

TABLE OF CONTENTS

CHAPTER VII. NEUROINFOMATICS AND NEURONET CONTROL 535

INTRODUCTORY CHAPTER

THIS CHAPTER INCLUDES PAPERS

PRESENTED AT THE CONFERENCE:

PLENARY SESSION

Organized by: ***Acad. Stanislav V. Emeljanov***

Acad. Nikolai A. Kuznetsov

Prof. Rafael M. Yusupov

INFORMATION, SCIENCE of MODEL FREE EXPERT CONTROL - PATH of PRESERVATION of the STATE

Alexander A. Krasovskii
Zhukovskii Air Force Engineering Academy
Leningradskii prospect, 40, Moscow, Russia

Abstract. The problem of reliability and perativeness of feed-backs of centralized control of organizational - economic systems, including state, are being considered. The solution may be achieved with the help of telecommunication nets with the special software and data bases.

Keywords. Organizational-economic control, telecommunication nets, simulation and prediction.

Introduction. In a modern socio economic and geopolitical circumstances engaging the newest reachings of technological progress and intelligent potential of exact sciences for a solution of problems of organizational - economic control at all levels, including state is necessary.

Egosteadness problem. The information, telecommunication nets allow to transmit and to treat vast information contents, but themselves the problems are not decided in view of a lack of an egosteadness. An egosteadness is understood as protection of action of a personal and party egoism, corruption, bureaucratism, organized criminality, hostile to antistate of activity [1]. Simulation and prediction - powerful, but not egosteady links of information process engineerings of control. A bright development of a lack of an egosteadness in conventional and new contours of organizational - economic control is delibarate distortion of datas, eye-wash, prospering at the moment. Conventional checks, commission, the inspections are frequently ineffective owing to corruption, controlling structures being criminal.

Addressing to concrete branch of state control - carrier, namely air carrier, we shall mark, that USA, meeting with delibarate distortions of an information, yet in 1975 have created a system ("program") " of the voluntary messages ". For a today's day such "systems" of informers are created in 15 developed countries, including Russia (on the public beginnings, 1992). However it is clear, that similar "systems" or "programs" cannot considerably decide a problem of reliability and perativeness of feed-backs of centralized control. This problem can be solved by means of telecommunication nets (TKN) with the special software and data bases.

Telecommunication nets. The creation (acquisition) hardware and common programm parts TKN by the non-state enterprises and airlines is recommended to be made by a condition of their legal existence. Development centralized, immediately updated a data base and special software, also continuously updated, expediently to assign to state centers with engaging of a fundamental science. This security can be under construction on principles of artificial intelligence and comparison of datas, obtained in TKN from many of the same type or interconnected enterprises, firms, airlines.

The reliability of treated datas not yet guaranteeigs high effectiveness of centralized control even only in a sense of realization of the operating spefications. The high management efficiency is inevitably connected to a solution many-dimensional, extremum problems of optimization. It is functions of direct channels TKN in contours of control.

Many-dimensional extremum problem. The solution of many-dimensional extremum problems, even there is a single-criteria nature usually requires high computing expenditures and is hampered at operative control in real time, necessary in modern conditions. We and independently by other authors recently develop a so-called selectively-averaging method (SAM) solution of extremum problems [2,3], suitable for operative shaping of controlling actions of many-dimensional plants.

However information of many criterions and spefications (restrictions) to one criterion with different parameters for different plants (principles the Pareto) deprives, generally speaking, information process engineering of an egosteadness.

Experts. Therefore development and introduction of a process engineering of expert-automated control is recommended, in which the experts of state and regional control centers fulfill functions of terms jury.

Each expert - term jury on an information showed him by TKN estimates outcomes of control on each rather short interval of time in numbers irrespective of other terms jury. Searching the best variant for the next stage of functioning is carried out automatically by means of SAM or other rather effective method, in which the role of a sole criterion fulfills a flowing summarized number jury. The terms jury are subjected to a professional selection and monitoring.

The controlling actions of the model free expert TKN can as actuate volumes of state budget financing as concerns to staff, labour discipline, training and tutoring of personnel and etc.

Conclusions. The similar information and control can prevent a further disintegration of transport, thermal, electropower and other state systems.

References

1. Krasovskii A.A. (1996) Some actual problems of a science of control. The theory and control systems, 6, 8-16.
2. Krasovskii A.A. (1992) A selectively-averaging method of a solution of extremum problems. Automatica and Telemechanica, 9,
3. Naumov A.I., Shelkovenkov I.E. (1996) A selectively-averaging method of searching of a global extremum in problems of optimization of driving of dynamic plants. The theory and control systems, 6, 152-158.

Mobile Robotic Manipulation

Oussama Khatib
Stanford University
Stanford, CA, 94305, USA
khatib@cs.stanford.edu

Abstract. Mobile manipulation capabilities are key to many new applications of robotics in space, underwater, construction, and service environments. In these applications, consideration of vehicle/arm dynamics is essential for robot coordination and control. This article discusses the inertial properties of holonomic mobile manipulation systems and presents the basic strategies developed for their dynamic coordination and control. These strategies are based on extensions of the *operational space formulation*, which provides the mathematical models for the description, analysis, and control of robot dynamics with respect to the task behavior.

Keywords. Operational Space, Redundancy, Dynamic Consistency, Vehicle/Arm Coordination, Multi-Robot Cooperation.

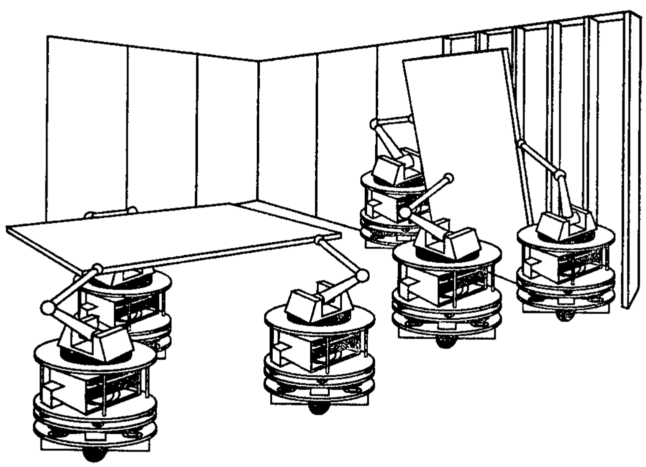


Figure 1: Robotics in Construction: Drywall

Introduction. A central issue in the development of mobile manipulation systems is vehicle/arm coordination [1,2]. This area of research is relatively new. There is, however, a large body of work that has been devoted to the study of motion coordination in the context of kinematic redundancy. In recent years, these two areas have begun to merge [3], and algorithms developed for redundant manipulators are being extended to mobile manipulation systems. Typical approaches to motion coordination of redundant systems rely on the use of pseudo- or generalized inverses to solve an under-constrained or degenerate system of linear equations, while optimizing some given criterion. These algorithms are essentially driven by kinematic considerations and the dynamic interaction between the end-effector and the manipulator's internal motions are ignored.

Our approach to controlling redundant systems is based on two models: an *end-effector dynamic model* obtained by projecting the mechanism dynamics into the operational space, and a *dynamically consistent force/torque relationship* that provides decoupled control of joint motions in the null space associated with the redundant mechanism. These two models are the basis for the dynamic coordination strategy we are implementing for the mobile platform.

Another important issue in mobile manipulation concerns cooperative operations between multiple vehicle/arm systems. An example of cooperative operations involving multiple vehicle/arm systems in construction tasks, is illustrated in Figure 1. Our study of the dynamics of parallel, multi robot structures reveals an important additive property. The effective mass and inertia of a multi-robot system at some operational point are shown to be given by the sum of the effective masses and inertias associated with the object and each robot. Using this property, the multi-robot system can be treated as a single *augmented object* [5] and controlled by the total operational forces applied by the robots. The control of internal forces is based on the *virtual linkage* [6] which characterizes internal forces.

Operational Space Dynamics. The joint space dynamics of a manipulator are described by

$$A(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \mathbf{\Gamma}; \tag{1}$$

where $\mathbf{q}$ is the n joint coordinates and $A(\mathbf{q})$ is the $n \times n$ kinetic energy matrix. $\mathbf{b}(\mathbf{q}, \dot{\mathbf{q}})$ is the vector of centrifugal and Coriolis joint-forces and $\mathbf{g}(\mathbf{q})$ is the gravity joint-force vector. $\mathbf{\Gamma}$ is the vector of generalized joint-forces. The operational space equations of motion of a manipulator are [4]

$$\Lambda(\mathbf{x})\ddot{\mathbf{x}} + \mu(\mathbf{x}, \dot{\mathbf{x}}) + \mathbf{p}(\mathbf{x}) = \mathbf{F}; \tag{2}$$

where $\mathbf{x}$, is the vector of the m operational coordinates describing the position and orientation of the effector, $\Lambda(\mathbf{x})$ is the $m \times m$ kinetic energy matrix associated with the operational space. $\mu(\mathbf{x}, \dot{\mathbf{x}})$, $\mathbf{p}(\mathbf{x})$, and $\mathbf{F}$ are respectively the centrifugal and Coriolis force vector, gravity force vector, and generalized force vector acting in operational space.

Redundancy. The operational space equations of motion describe the dynamic response of a manipulator to the application of an operational force $\mathbf{F}$ at the end effector. For non-redundant manipulators, the relationship between operational forces, $\mathbf{F}$, and joint forces, $\mathbf{\Gamma}$ is

$$\mathbf{\Gamma} = J^T(\mathbf{q})\mathbf{F}; \tag{3}$$

where $J(\mathbf{q})$ is the Jacobian matrix. However, this relationship becomes incomplete for redundant systems. We have shown that the relationship between joint torques and operational forces is

$$\mathbf{\Gamma} = J^T(\mathbf{q})\mathbf{F} + \left[I - J^T(\mathbf{q})\overline{J}^T(\mathbf{q})\right]\mathbf{\Gamma}_0; \tag{4}$$

with

$$\overline{J}(\mathbf{q}) = A^{-1}(\mathbf{q})J^T(\mathbf{q})\Lambda(\mathbf{q}); \tag{5}$$

where $\overline{J}(\mathbf{q})$ is the *dynamically consistent generalized inverse* [5] This relationship provides a decomposition of joint forces into two dynamically decoupled control vectors: joint forces corresponding to forces acting at the end effector ($J^T\mathbf{F}$); and joint forces that only affect internal motions, $\left([I - J^T(\mathbf{q})\overline{J}^T(\mathbf{q})]\mathbf{\Gamma}_0\right)$.

Using this decomposition, the end effector can be controlled by operational forces, whereas internal motions can be independently controlled by joint forces that are guaranteed not to alter the end effector's dynamic behavior. This relationship is the basis for implementing the dynamic coordination strategy for a vehicle/arm system. The end-effector equations of motion for a redundant manipulator are obtained by the projection of the joint-space equations of motion (1), by the *dynamically consistent* generalized inverse $\overline{J}^T(\mathbf{q})$,

$$\overline{J}^T(\mathbf{q})\,[A(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q},\dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \Gamma] \quad \Longrightarrow \quad \Lambda(\mathbf{q})\ddot{\mathbf{x}} + \mu(\mathbf{q},\dot{\mathbf{q}}) + \mathbf{p}(\mathbf{q}) = \mathbf{F}; \qquad (6)$$

The above property also applies to non-redundant manipulators, where the matrix $\overline{J}^T(\mathbf{q})$ reduces to $J^{-T}(\mathbf{q})$.

Vehicle/Arm Coordination. In our approach, a mobile manipulator system is viewed as the mechanism resulting from the serial combination of two sub-systems: a "macro" mechanism with coarse, slow, dynamic responses (the mobile base), and a relatively fast and accurate "mini" device (the manipulator). The mobile base referred to as the *macro structure* is assumed to be holonomic. Let Λ be the *pseudo kinetic energy matrix* associated with the combined macro/mini structures and Λ_{mini} the operational space *kinetic energy matrix* associated with the mini structure alone. The magnitude of the inertial properties of macro/mini structure in a direction represented by a unit vector $\mathbf{w}$ in the m-dimensional space are described by the scalar [5]

$$\sigma_{\text{w}}(\Lambda) = \frac{1}{(\mathbf{w}^T\Lambda^{-1}\mathbf{w})};$$

which represents the effective inertial properties in the direction $\mathbf{w}$.

Our study has shown [5] that, *in any direction* $\mathbf{w}$*, the inertial properties of a macro/mini-manipulator system* (see Figure 2) *are smaller than or equal to the inertial properties associated with the mini-manipulator in that direction:*

$$\sigma_{\text{w}}(\Lambda) \le \sigma_{\text{w}}(\Lambda_{\text{mini}}). \qquad (7)$$

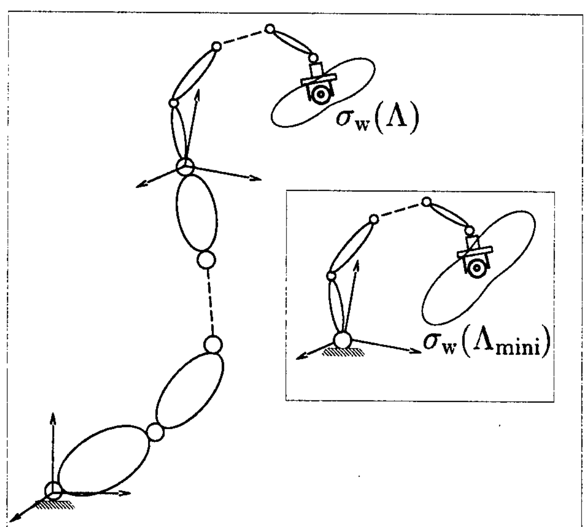


Figure 2: Inertial properties of a macro/mini-manipulator

A more general statement of this *reduced effective inertial* property is that the inertial properties of a redundant system are bounded above by the inertial properties of the structure formed by the smallest distal set of degrees of freedom that span the operational space.

The reduced effective inertial property shows that the dynamic performance of a combined macro/mini system can be made comparable to (and, in some cases, better than) that of the lightweight mini manipulator. The idea behind our approach for the coordination of macro and mini structures is to treat them as a single redundant system.

The *dynamic coordination* we propose is based on combining the operational space control, $\mathbf{\Gamma} = J^T\mathbf{F}$, with a minimization of some criterion associated with the macro/mini relative posture, e.g. the deviation from the midrange joint positions of the mini-manipulator. This minimization is implemented with a joint torque vector, $\mathbf{\Gamma}_{\text{Coordination}}$, selected as the gradient of a potential function constructed for this minimization. To eliminate any coupling effect on the end-effector, these torques are projected in *dynamically consistent* null space of equation (4). The total joint torques are

$$\mathbf{\Gamma} = J^T(\mathbf{q})\mathbf{F} + \left[I - J^T(\mathbf{q})\overline{J}^T(\mathbf{q})\right]\mathbf{\Gamma}_{\text{Coordination}}. \tag{8}$$

Cooperative Manipulation. Our research in cooperative manipulation has produced a number of results which provide the basis for the control strategies we are developing for mobile manipulation platforms. Our approach is based on the integration of two basic concepts: The *augmented object* [5] and the *virtual linkage* [6]. The *virtual linkage* characterizes internal forces, while the *augmented object* describes the system's closed-chain dynamics. These models have been successfully used in cooperative manipulation for various compliant motion tasks performed by two and three PUMA 560 manipulators [7].

Augmented Object. The *augmented object* model provides a description of the dynamics at the operational point for a multi-arm robot system. The simplicity of these equations is the result of an additive property that allows us to obtain the system equations of motion from the equations of motion of the individual mobile manipulators. The *augmented object* model is

$$\Lambda_{\oplus}(\mathbf{x})\ddot{\mathbf{x}} + \mu_{\oplus}(\mathbf{x},\dot{\mathbf{x}}) + \mathbf{p}_{\oplus}(\mathbf{x}) = \mathbf{F}_{\oplus}; \tag{9}$$

with

$$\Lambda_{\oplus}(\mathbf{x}) = \Lambda_{\mathcal{L}}(\mathbf{x}) + \sum_{i=1}^{N}\Lambda_i(\mathbf{x}); \tag{10}$$

where $\Lambda_{\mathcal{L}}(\mathbf{x})$ and $\Lambda_i(\mathbf{x})$ are the kinetic energy matrices associated with the object and the i^{th} effector, respectively. The vectors, $\mu_{\oplus}(\mathbf{x},\dot{\mathbf{x}})$ and $\mathbf{p}_{\oplus}(\mathbf{x})$ also have the additive property. The generalized operational forces $\mathbf{F}_{\oplus}$ are the resultant of the forces produced by each of the N effectors at the operational point.

$$\mathbf{F}_{\oplus} = \sum_{i=1}^{N}\mathbf{F}_i. \tag{11}$$

The dynamic decoupling and motion control of the augmented object in operational space is achieved by selecting a control structure similar to that of a single manipulator. The dynamic behavior of the augmented object of equation (9) is controlled by the net force $\mathbf{F}_{\oplus}$. Due to the actuator redundancy of multi-effector systems, there is an infinity of joint-torque vectors that correspond to this force.

Virtual Linkage. Object manipulation requires accurate control of internal forces. Recently, we have proposed the *virtual linkage* [7] as a model of internal forces associated with multi-grasp manipulation. In this model, grasp points are connected by a closed, non-intersecting set of virtual links, as illustrated in Figure 3 for a three-grasp task.

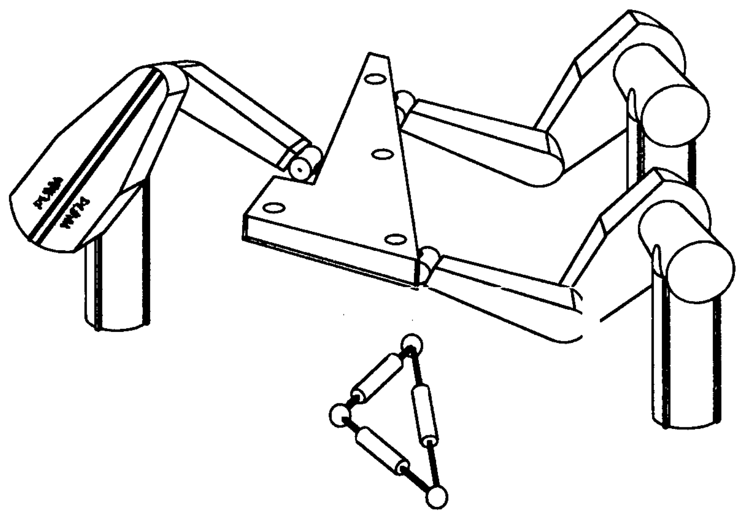


Figure 3: The Virtual Linkage

In the case of an N-grasp manipulation task, a *virtual linkage* model is a $6(N-1)$ degree of freedom mechanism that has $3(N-2)$ linearly actuated members and N spherically actuated joints. Forces and moments applied at the grasp points of this linkage will cause forces and torques at its joints. We can independently specify internal forces in the $3(N-2)$ members, along with $3N$ internal moments at the spherical joints. Internal forces in the object are then characterized by these forces and torques in a physically meaningful way. The relationship between applied forces, their resultant and internal forces is

$$\begin{bmatrix} \mathbf{F}_{res} \\ \mathbf{F}_{int} \end{bmatrix} = \mathbf{G} \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_N \end{bmatrix}; \tag{12}$$

where $\mathbf{F}_{res}$ represents the resultant forces at the operational point, $\mathbf{F}_{int}$ the internal forces and $\mathbf{f}_i$ the forces applied at the grasp point i. $\mathbf{G}$ is called the grasp description matrix, and relates forces applied at each grasp to the resultant and internal forces in the object.

Decentralized Cooperation. For fixed base manipulation, the *augmented object* and *virtual linkage* have been implemented in a multiprocessor system using a centralized control structure. This type of control is not suited for autonomous mobile manipulation platforms.

In a multiple mobile robot system, each robot has real-time access only to its own state information and can only infer information about the other robots' grasp forces through their combined action on the object. Recently, we have developed a new control structure for decentralized cooperative mobile manipulation [8]. In this structure, the object level

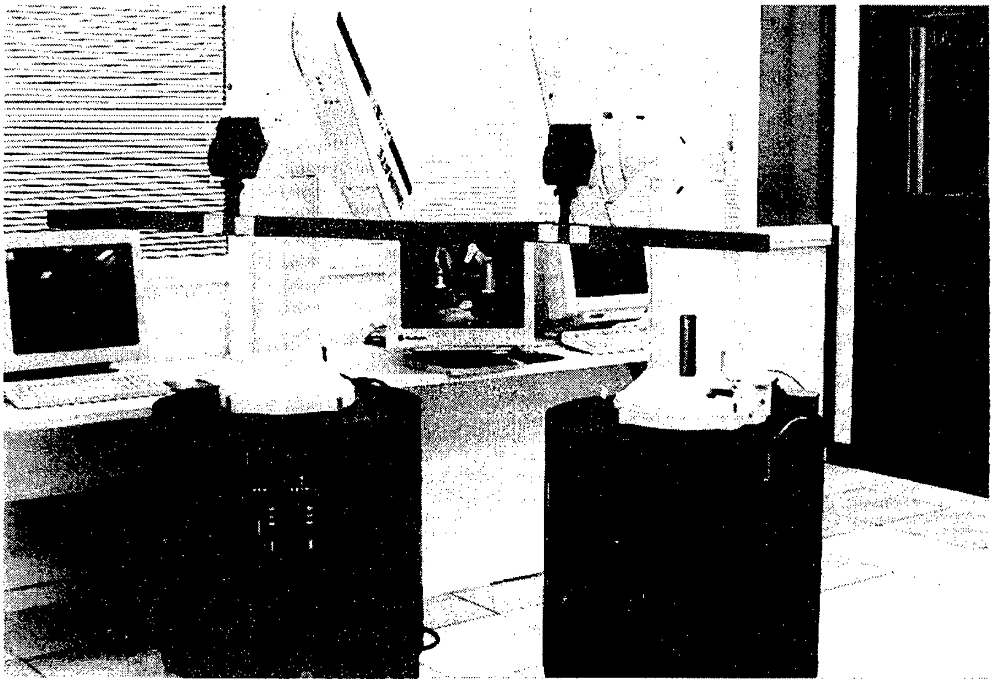


Figure 4: The Stanford Mobile Platforms

specifications of the task are transformed into individual tasks for each of the cooperative robots. Local feedback control loops are then developed at each grasp point. The task transformation and the design of the local controllers are accomplished in consistency with the *augmented object* and *virtual linkage* models.

Experimental Mobile Platforms. In collaboration with Oak Ridge National Laboratories and Nomadic Technologies, we have completed the design and construction of two holonomic mobile platforms (see Figure 4). Each platform is equipped with a PUMA 560 arm, various sensors, a multi-processor computer system, a multi-axis controller, and sufficient battery power to allow for autonomous operation. The base consists of three "lateral" orthogonal universal-wheel assemblies which allow the base to translate and rotate holonomically in relatively flat office-like environments [9].

The control strategies discussed above have been implemented on these two platforms. Erasing a whiteboard, cooperating in carrying a basket, and sweeping a desk are examples of tasks demonstrated with the Stanford Mobile Platforms [10]. The dynamic coordination strategy has allowed full use of the relatively high bandwidth of the PUMA. Object motion and force control performance with the Stanford mobile platforms are comparable with the results obtained with fixed base PUMA manipulators.

Conclusion. We have presented extensions of various operational space methodologies for fixed-base manipulators to mobile manipulation systems. A vehicle/arm platform is treated as a macro/mini structure. This redundant system is controlled using a dynamic coordination strategy, which allows the mini structure's high bandwidth to be fully utilized. Cooperative operations between multiple platforms rely on the integration of the *augmented object*, which describes the system's closed-chain dynamics, and the *virtual* *linkage*, which characterizes internal forces. These models are the basis for the decentralized control structure presented in [8]. Vehicle/arm coordination and cooperative operations have been implemented on the two mobile manipulator platforms developed at Stanford University.

Acknowledgments. The financial support of Boeing, General Motors, Hitachi Construction Machinery, and NSF (grants IRI-9320017 and CAD-9320419) is gratefully acknowledged. Many thanks to Alain Bowling, Oliver Brock, Arancha Casal, Kyong-Sok Chang, Robert Holmberg, Francois Pin, Diego Ruspini, David Williams, James Slater, John Slater, Stef Sonck, and Kazuhito Yokoi for their valuable contributions to the design and construction of the Stanford Mobile Platforms.

References

1 Ullman, M., Cannon, R. (1989). Experiments in Global Navigation and Control of a Free-Flying Space Robot. Proc. Winter Annual Meeting, Vol. 15, pp. 37-43.

2 Umetani, Y., and Yoshida, K. (1989). Experimental Study on Two-Dimensional Free-Flying Robot Satellite Model. Proc. NASA Conf. Space Telerobotics.

3 Papadopoulos, E., Dubowsky, S. (1991). Coordinated Manipulator/Spacecraft Motion Control for Space Robotic Systems. Proc. IEEE Int. Conf. Robotics and Automation, pp. 1696-1701.

4 Khatib, O. (1987). A Unified Approach to Motion and Force Control of Robot Manipulators: The Operational Space Formulation. IEEE J. Robotics and Automation, vol. 3, no. 1, pp. 43-53.

5 Khatib, O. (1995). Inertial Properties in Robotics Manipulation: An Object-Level Framework, *Int. J. Robotics Research,* vol. 14, no. 1, pp. 19-36.

6 Williams, D. and Khatib, O. (1993). The Virtual Linkage: A Model for Internal Forces in Multi-Grasp Manipulation. Proc. IEEE Int. Conf. Robotics and Automation, pp. 1025-1030.

7 Williams, D. and Khatib, O. (1995). Multi-Grasp Manipulation," IEEE Int. Conf. Robotics and Automation Video Proceedings.

8 Khatib, O. Yokoi, K., Chang,K., Ruspini, D., Holmberg, R. Casal, A., Baader., A. (1996). Force Strategies for Cooperative Tasks in Multiple Mobile Manipulation Systems, *Robotics Research 7, The Seventh International Symposium,* G. Giralt and G. Hirzinger, eds., Springer, pp. 333-342.

9 Pin, F. G. and S. M. Killough (1994). A New Family of Omnidirectional and Holonomic Wheeled Platforms for Mobile Robots, IEEE Trans. on Robotics and Automation, Vol. 10, No. 4, pp. 480-489.

10 Khatib, O., K. Yokoi, K. Chang, D. Ruspini, R. Holmberg, A. Casal, and A. Baader (1996). The Robotic Assistant, *IEEE Int. Conf. Robotics and Automation Video Proceedings*, 1996.

DEVELOPMENT OF CONTROL AND INFORMATICS IN ST.PETERSBURG

Valentin Ponomaryov and Rafael Yusupov
St.Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences
39, 14th Line, St.Petersburg, 199178, Russia
e-mail: spiiran@iias.spb.su

Abstract. Brief outline of the history of the control theory and informatics in St.Petersburg is given.

Keywords. Control theory, control systems, informatics.

In XIXth century St.Petersburg became the main industrial and scientific center of Russia. So the first fundamental results in the control theory were obtained by prominent Russian scientists in St.Petersburg.`

The industrial revolution brought a wide application of the steam engines with regulator. The stability and quality of the controlled process in the steam engine became of the great importance and several attempts, were made to develop the necessary mathematical approach for solving the problem.

The first outstanding result in this field was obtained by I.A.Vischnegradskii. His work [1] may be considered as a beginning of the classical automatic control theory and had a great impact in regulator design. The main features of Vischnegradskii approach were:

- mathematical model of the steam engine with regulator,
- linearization of the differential equation describing the controlled process,
- stability criteria,
- parameter plane stability analysis,
- connection between parameters of the regulator and the quality of the controlled process.

The next important stage of the automatic control theory progress is connected with famous work of A.M.Lyapunov [2], in which he developed a stability theory of dynamical systems described by nonlinear timevarying ordinary differential equations. He was the first who formulated the definition of a stability in mathematical terms.

The first Method of Lyapunov permits to deduce the stability of nonlinear system from the stability properties of the linearized system, and Lyapunov's Direct Methods represent the general necessary condition of the stability of equilibrium state of any nonlinear system. Until now it is one of the most important foundations of the control theory.

The following development in the control theory in our city is represented in works of I.N.Voznesenskii [3], where he formulated the principles of autonomous regulation and obtained new results concerning the problems of the energetic machines control.

The control education in St.Petersburg began during the first years of XXth century. The first control textbook was written by N.Gukovskii and published in 1909. In 30th the chair of automation and remote control was founded at the Leningrad Electrical Engineering Institute and appropriate specialization was established.

In the post-war II time the most remarkable results in control theory were obtained by researchers belonging to groups concentrated around several outstanding scientists.

A.I.Lurie [4] formulated the absolute stability problem as the development of Lyapunov's approach. His followers investigated the problems of theory of the multidimensional discrete control systems (V.Y.Katcovnik, K.A.Poluektov), automata theory (V.I.Varshavskii), stohastic control processes (A.A.Pervozvanskii), theory of discrete time control of continuous dynamic systems (E.N.Rosenvasser), optimal control (V.A.Troizkii) and others.This group or school was formed in the Leningrad Polytechnical Institute.

The big and productive group of specialists in control theory was formed by E.P.Popov, now academician of the Russian Academy of Sciences. His widely known textbook [5] is of great educational value until now. He developed harmonic linearization approach [6] which became of great use in control systems engineering. His followers solved a number of problems connected with control systems analysis and synthesis. Among them were problems of control systems analysis (D.A.Bashkyrov, V.T.Kochetkov, V.A.Bessecerskii, V.N.Kalynin), nonlinear control, systems analysis (I.P.Paltov, R.A.Nelepin, E.I.Cheypalo, V.I.Tchernetzkii), discrete control systems analysis (S.M.Fedorov), optimal control systems (V.M.Ponomaryov, L.A.Mayboroda, V.I.Gorodetzkii), reliability of the control systems (A.M.Polovko), sensitivity of the control systems (R.M.Yusupov) and others. This group worked mostly in Military the Engineering Academy.

Lyapunov's approach was developed as well in works of V.I.Zubov, now corresponding member of the Russian Academy of Sciences [7]. He founded the department of the applied mathematics control processes in the Leningrad State University and became the leader of the group of researchers working on control theory problems such as control system synthesis (R.I.Truchaev, V.V.Khomenyuk), optimal control systems (V.F.Demjanov), robust control systems (V.U.Kharitonov) and others.

Important results were obtained in the general control theory (A.A.Voronov, D.V.Vasyliev, V.B.Yakovlev, A.V.Basharin, A.A.Vavilov), theory of stability of nonlinear systems (V.N.Fomin, G.A.Leonov, A.K.Gelig) the adaptive control theory (V.A.Yakubovich, A.L.Fradkov., A.V.Timofeev).

It was typical for the research in control theory in the 50 th -60 th that the most difficult problems came from the important practical necessity to build automatic control systems for such objtcts as aircrafts, rockets, space vehicles etc. In 70th the main attention of specialists in control theory was turned to the problem, of control of processes in other kind of systems. First of all it were manufacturing systems, economic systems, ecological systems, military systems and such like. It was typical that forinstance the manufacturing system as a whole included such complicated enough objects as robots, automatic machine tools, flexible manufacturing cells etc. At that time a computer became an essential part of a control system, because it was not possible to control such processes using less powerful means.

In many cases the control included decision making based on results,of the processing of big set, of dataand other kinds of information.

It became clear that industrial automation problems in Leningrad have some unique features. It happened so that great part of industry in the city produced nonserial or small serial products like ships, electrical machines, big turbines, big machine tools etc. In this case the only way to increase the productivity was to build the computer integrated production systems based on computer aided planning, computer-aidid research, computer-aided engineering and flexible manufacturing systems. So a number of new theoretical problems appeared.

In the very end of the 50th S.P.Mytrophanov formulated the principles of group technology which gave the possibility to increase productivity of the nonserial or small serial manufacturing processes. The only drawback of this approach was a long time required to reorganize the technological process for every new group of products. When computers began to be used for this purpose the difficulty disappeared and first flexible manufacturing systems in Japan were build using the group technology approach.

The various parts of the computer integrated manufacturing problem were investigated by Leningrad specialists in control theory and automation. Important theoretical and application results were obtained in computer-aided research (V.M.Akhutin, V.V.Alexandrov, V.V.Ivanitshev), computer-aided design (A.Y.Zvonitskii, A.A.Leskin, A.E.Bor-Ramenskii), flexible manufacturing (V.Y.Sovetov, M.B.Ignatijev, C.A.Mayorov, G.V.Orlovskii, V.G.Kolosov, A.I.Fedotov, Y.M.Smirnov), control robots (E.I.Yurevich, F.M.Kulakov) etc. The Institute of Informatics and Automation of the Academy of Sciences became the most active theoretical center in these fields.

The 70th were the remarkable period in history of the control theory. It was a period when the large-scale and comlex control systems began to be the main subject of the control theory and computer began to be the essential part of such systems. But the success in development of these new systems with required quality and reliability depended on solving a number of new problems.

New controlled system consisted of the subsystems which were large scale and complex also. So the complexity of mathematical models of such systems increased substancially.

The analytical simulation of the control processes became not practical and gave way to computational experiment.It brought the special problems, connected with the computer properties and pecularities of algorithms and computational processes.

In old continuous control systems an optimal control was a function of state variables and time. The optimal parameters of the control system were obtained using various methods of system synthesis when the system was designed. In the large-scale and complex system the quality criteria may be not formalized completely and may be changed during the control process. In many cases the quality of control is to be evaluated by the man participating in the control process and making decisions about control and disturbances. The final decision may be made as a result of alternative control strategies analysis. As to the decision support process it can be the problem belonging to the area of the Artificial Intelligence.

The amount of information required to create the large scale and complex control system and to form the appropriate control commands increased substantially. It includes the initial information about the object to be controlled, which is necessary to model the control process, the general requirements to the characteristics of the control process, the state variable at every moment used when forming the control commands, predicting the possible results and, if it is necessary correcting the model of the control process. It is evident that the model of the control process in large scale and complex system is to be complex enough. So it became necessary to create the computer-aided research system which is capable to build the controlled process of programming which results in the required model corresponding with the initial description of the system.

The information technologies became the important part of the control systems. But the control theory is not less important to control the computational process in proper way. And that situation created the close connection between the control theory and the theory of information processes and systems or the informatics. First time the term "informatics" was proposed by Prof. F.E.Temnicov in his paper published in 1963. He proposed to use this term for the new science consisting of the theory of information components, the theory of information processes and the theory of information systems. It is very close to the present understanding of Informatics in Russia inspite of the fact that the science is progressing very fast. Now it includes the theory of information, theory of computers and computer systems, the theory of algorithms, the theory of programming, the artificial intelligence etc.

Historically the mathematical foundations of the informatics were developed in St.Petersburg by such prominent mathematicians as P.L.Tshebyshev, V.A.Steklov, A.N.Krylov (numerical calculation and approximation theory), A.V.Kantorovich (linear programming), A.A.Marñov, N.A.Shanin, Y.V.Matiasevich (mathematical logic) and others.

The electronic components, computers and telecommunication equipment were developed and produced in our city in a number of research institutes, beginning from 60 th. At the same time research and education in programming was initiated in Leningrad State University (L.V.Kantorovich, S.S.Lavrov) and a number of big higher school institutes (S.A.Mayorov, V.B.Smolov, T.N.Sokolov and others). The big amount of the system and application software was developed in the institutes of the Academy of Sciences, higher school institutes, and big industrial research institutes. In 1978 Leningrad Research Computing Center of the Academy of Sciences was founded (from 1985 the Institute of Informatics and Automation).

Joint efforts of the specialists in control theory and informatics, working in St.Petersburg, (in many cases that is the same person) brought good results such as new complex control systems and new types of computer systems. As the most important impact of informatics in control theory a number of computer-aided software engineering systems and tools for control systems development must be mentioned. Among them are system for computer-aided algoritmic model development, computer-aided decision support systems development, computer-aided system optimization tools development, computer-aided optimal control algorithms for continuous and discrete time control systems development and others. As to the control theory impact in informatics there must be considered new results in control of the distributed computer systems and control of the distributed computer networks.

Until now St.Petersburg is one of the biggest world scientific, educational and industrial centers where more than 1600 institutions and companies are working on control and informatics problems. And we hope to keep this position in future.

Limited volume of this presentation does not to describe more completely the impact of all St.Petersburg researchers in control and informatics. Authors apologize in advance but hope to prepare later more detailed review.

References

1. Vischnegradskii I.F. (1877) Uber direktwirkende Regulatoren, civilingenieur.
2. Lyapunov A.M. (1893). The General Problem of Stability of Motion, Com.Soc.Math., Kharkow, in Russian.
3. Voznesenskii I.N. (1922). On Regulators of Non-direct Action, Proc. Leningr.Technol.Inst., Leningrad, in Russian.
4. Lurie A.I., Postnikov V.N.(1944). On the Theory of Stability of Control Systems, (1944) Appl.Math.Mech., 8, Moscow, in Russian.
5. Popov E.P. (1954). Dynamics of Automatic Regulation Systems, Gostechizdat, Moscow, in Russian.
6. Popov E.P., Paltov I.P. (1960). Approximate Methods for Analysis of Nonlinear Automatic Systems. Gostechizdat, Moscow, in Russian.
7. Zubov V.I. (1964). Methods of A.M.Lyapunov and their Applications, P.Noordhoff Ltd., Groningen.

DECOMPOSITION AND SYNTHESIS OF CONTROL IN A NONLINEAR DYNAMIC SYSTEM

Felix L. Chernousko, Sergey A. Reshmin
Institute for Problems in Mechanics of the Russian Academy of Sciences
pr. Vernadskogo 101-1, Moscow 117526, Russia
e-mail: chern@ipmnet.ru

Abstract. A nonlinear controlled dynamic system with many degrees of freedom is considered that is described by the Lagrangian equations. The system is subjected to bounded control forces and has a restricted area of motion. Matrix of the kinetic energy of the system is assumed to be close to a constant diagonal matrix with positive diagonal elements. This condition is satisfied for robots driven by actuators with high gear ratios. A feedback control law is proposed that satisfies the imposed constraints and drives the system to the prescribed terminal state in finite time. The design of feedback control is based on the decomposition of the given system into subsystems with one degree of freedom [1]. Under certain assumptions, the nonlinearities are estimated. Using these estimates, a feedback control is designed counteracting the influence of the nonlinearities for each subsystem. Therefore, the decomposition of the original system into subsystems is accomplished. As a result, explicit formulae representing a feedback control for the original nonlinear system are obtained. Applications of the described control to robots are presented. The paper follows the works [1–4], but differs in the assumptions made. Another approach to control design for dynamic systems, based on the principle of decomposition, is proposed in [5,6].

Keywords. Dynamic system, feedback control, nonlinear system, decomposition, robots.

Problem Statement. Consider a nonlinear controlled dynamic system with n degrees of freedom governed by Lagrangian equations

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = U_i + Q_i, \qquad i = 1, ..., n \tag{1}$$

Here, t is time; $q = (q_1, ..., q_n)$ is the n-vector of generalized coordinates, $q \in D \subset R^n$; U_i are the control generalized forces that are to found; Q_i are all the other external and internal forces; T is the kinetic energy of the system given by

$$T(q, \dot{q}) = \frac{1}{2}(A(q)\dot{q}, \dot{q}) = \frac{1}{2}\sum_{j,k=1}^{n} A_{jk}(q)\dot{q}_j\dot{q}_k \tag{2}$$

where $A_{jk}(q)$ are elements of the symmetric positive definite $n \times n$-matrix $A(q)$. The domain of motions D in q-space is bounded and given by

$$D = \{q : \ q_i^- \le q_i \le q_i^+, \quad i = 1, ..., n\} \tag{3}$$

where q_i^-, q_i^+ are constants. The control forces are subjected to the constraints

$$|U_i| \le U_i^0, \qquad i = 1, ..., n \tag{4}$$

Here, U_i^0 are positive constants.

Suppose the kinetic energy T defined by (2) and the generalized forces Q_i from (1) satisfy the following conditions. Let the matrix $A(q)$ consists of a constant diagonal matrix with positive diagonal elements and a remaining position-dependent part

$$A(q) = I + \tilde{A}(q), \qquad I = \mathrm{diag}(I_1, ..., I_n), \quad I_i > 0 \tag{5}$$

We assume that

$$|\tilde{A}(q)z| \le \mu|z|, \qquad \forall z \in R^n \tag{6}$$

$$|\partial \tilde{A}_{jk}/\partial q_i| \le c, \qquad i, j, k = 1, ..., n \tag{7}$$

for all $q \in D$. Here, $\mu > 0$ is a sufficiently small parameter (its possible value will be specified below), and $c > 0$ is a constant.

Let generalized forces Q_i consist of two terms: the bounded forces G_i and the forces F_i that are sufficiently small, if the velocities $\dot{q}_i$ are small

$$Q_i = G_i + F_i \tag{8}$$

$$|G_i| \le G_i^0 < U_i^0, \qquad i = 1, ..., n \tag{9}$$

$$|F_i(q, \dot{q}, t)| \le a|\dot{q}| + b|\dot{q}|^2, \qquad i = 1, ..., n \tag{10}$$

Here, G_i^0, a, b are positive constants. The forces $G_i(q, \dot{q}, t)$, $F_i(q, \dot{q}, t)$ in (8) may be uncertain, with only bounds on them (9), (10) being essential. We consider the following control problem.

Problem 1. Find a feedback control $U_i(q_i, \dot{q}_i)$, $i = 1, ..., n$ that satisfies (4) and steers the system (1) from any given initial state

$$q(0) = q^0, \qquad \dot{q}(0) = \dot{q}^0, \qquad q^0 \in D \tag{11}$$

to the prescribed terminal state with zero velocities

$$q(\tau) = q^1, \qquad \dot{q}(\tau) = 0, \qquad q^1 \in D \tag{12}$$

Here, the initial instant of time is taken equal to zero without loss of generality, and the terminal time is denoted by $\tau > 0$. Time of the control process τ is finite and free.

Decomposition. In order to construct the desired control in Problem 1, we can use the method of decomposition [1]. Substituting (2), (8) into (1), we obtain the equations of motion in the form

$$A(q)\ddot{q} = U + G + S(q, \dot{q}, t) \tag{13}$$

where the following notation is introduced

$$U = (U_1, ..., U_n), \quad G = (G_1, ..., G_n), \quad S = (S_1, ..., S_n)$$

$$S_i(q, \dot{q}, t) = F_i(q, \dot{q}, t) + \sum_{j,k=1}^{n} \left(\frac{1}{2} \frac{\partial A_{jk}}{\partial q_i} - \frac{\partial A_{ij}}{\partial q_k} \right) \dot{q}_j \dot{q}_k \tag{14}$$

Note that $S_i \to 0$ as $|\dot{q}| \to 0$. Let us multiply both sides of equation (13) by IA^{-1} (the matrix I is introduced in (5)). Thus we obtain

$$I_i\ddot{q}_i = U_i + V_i, \qquad i = 1, ..., n \tag{15}$$

$$V_i = G_i + S_i - [\tilde{A}A^{-1}(U + G + S)]_i \tag{16}$$

The system (15), (16) is equivalent to the initial equation (13).
Suppose the following inequalities hold

$$|V_i| \leq \rho_i U_i^0, \qquad \rho_i < 1, \qquad i = 1, ..., n \tag{17}$$

where ρ_i are some constants. We shall regard the terms V_i in equations (15) as independent uncertain disturbances. Hence, we can treat (15) as a set of n separate subsystems with one degree of freedom each that are subjected to bounded disturbances. Below, we propose a feedback control law for each of these subsystems and give conditions that ensure the assumption (17) for the fixed dynamic system (15).

The Control of a Subsystem. A scalar feedback control $U_i(q_i, \dot{q}_i)$ which satisfies (4) and brings the corresponding subsystem (15) from any initial state $(q_i^0, \dot{q}_i^0)$ to the terminal state $(q_i^1, 0)$ in finite time for any admissible disturbance V_i, satisfying (17), is given in [1] as

$$U_i(q_i, \dot{q}_i) = U_i^0 \text{sign } \psi_i(q_i, \dot{q}_i), \qquad \psi_i \neq 0$$

$$U_i(q_i, \dot{q}_i) = -U_i^0 \text{sign } \dot{q}_i, \qquad \psi_i = 0$$

$$\psi_i(q_i, \dot{q}_i) = q_i^1 - q_i - \dot{q}_i\,|\dot{q}_i|/2X_i \tag{18}$$

Here, X_i is the positive parameter of control, which is connected with ρ_i by the expression

$$X_i = U_i^0(1 - \rho_i)/I_i \tag{19}$$

Note that the parameter X_i is unknown yet, because the constant ρ_i is not determined.

The presented control is obtained as the time optimal control for the differential game described by equations (15), where U_i and V_i are considered as the bounded controls of two players [7]. This control is of the bang-bang type: $U_i = \pm U_i^0$. The switching curve $\psi_i(q_i, \dot{q}_i) = 0$ consists of two parabolic arcs which are symmetric with respect to the point $(q_i^1, 0)$.

Let us introduce the set Ω_i in the $(q_i, \dot{q}_i)$-space of the ith subsystem

$$\Omega_i = \{(q_i, \dot{q}_i) : \; q_i^- \leq q_i \leq q_i^+, \quad f_i^- \leq \dot{q}_i \leq f_i^+\} \tag{20}$$

$$f_i^-(q_i) = -[2X_i(q_i - q_i^-)]^{1/2}, \qquad f_i^+(q_i) = [2X_i(q_i^+ - q_i)]^{1/2}$$

Let us describe the motion of the ith subsystem (15) under the control (18), (19), assuming that $|V_i| \leq \rho_i U_i^0$, $\rho_i < 1$ and the initial state $(q_i^0, \dot{q}_i^0)$ belongs to Ω_i

$$(q_i^0, \dot{q}_i^0) \in \Omega_i \tag{21}$$

The process of control of the subsystem (15) is divided into two basic stages. At the first stage, the subsystem moves with a constant control until its phase point falls on the switching curve. During this stage, according to (15) and (17) – (19), we have (to be definite, we assume that $\psi_i(q_i^0, \dot{q}_i^0) < 0$)

$$\ddot{q}_i \leq -X_i \tag{22}$$

From (20) – (22) we have the following relations

$$\frac{d\dot{q}_i}{dq_i} \le -\frac{X_i}{f_i^+(q_i)} = \frac{df_i^+(q_i)}{dq_i}, \quad \dot{q}_i > 0; \qquad \frac{d\dot{q}_i}{dq_i} > 0, \quad \dot{q}_i < 0$$

Therefore, the phase trajectory, having started in the domain Ω_i, cannot leave it and falls on the switching curve. For $\psi_i(q_i^0, \dot{q}_i^0) > 0$, this fact is proved similarly.

Suppose the ith subsystem (15) reaches the switching curve. After that, the phase point moves along the switching curve to the terminal state. The parabolic branches of this curve coincide with the phase trajectories of the considered subsystem, if U_i is chosen according to (18), (19), and $V_i = -\rho_i U_i$. If $V_i \ne -\rho_i U_i$, then the phase point also follows the switching curve, but in a chattering regime: the control U_i switches between $+U_i^0$ and $-U_i^0$ infinitely often, so that in the average we have $\ddot{q}_i = X_i$ and $\ddot{q}_i = -X_i$ along the corresponding branches of the switching curve.

So, if at the initial instant of time the conditions (20), (21) are satisfied for all $i = 1, ..., n$, then all the considered motions of the subsystems (15) lie within the corresponding domains Ω_i. Therefore, the restrictions (3) and

$$|\dot{q}_i| \le (2d_i X_i)^{1/2}, \qquad d_i = q_i^+ - q_i^- \tag{23}$$

hold true.

As shown in [1], the time of motion for the ith subsystem (15) is maximal under the "worst" disturbance $V_i = -\rho_i U_i$ and is determined by the expression

$$\tau_i^*(q_i^0, \dot{q}_i^0) = X_i^{-1}\{2[(\dot{q}_i^0)^2/2 - X_i(q_i^0 - q_i^1)\gamma_i]^{1/2} - \dot{q}_i^0\gamma_i\} \tag{24}$$

$$\gamma_i = \text{sign } \psi_i(q_i^0, \dot{q}_i^0), \qquad \psi_i \ne 0$$

$$\gamma_i = \pm 1, \qquad \psi_i = 0$$

The total time of motion τ does not exceed the greatest of times calculated for each degree of freedom. According to (24), we have

$$\tau \le \tau^* = \max_i(\tau_i^*), \qquad i = 1, ..., n \tag{25}$$

Choice of Parameters. The use of control (18) is possible only if conditions (17) hold during the motion. Let us show that there exist parameters X_i entering the control law (18) such that relations (17) are really fulfilled.

First, let us estimate the absolute values of the functions V_i. Due to (4) – (6), (9), and (16), we have

$$|V_i| \le G_i^0 + \left(1 + \frac{\mu n^{1/2}}{I_{\min} - \mu}\right) S^0 + \frac{\mu}{I_{\min} - \mu}\left[\sum_{j=1}^{n}(U_j^0 + G_j^0)^2\right]^{1/2}, \quad \mu < I_{\min} \tag{26}$$

Here $I_{\min} = \min_i(I_i)$, $i = 1, ..., n$; S^0 is the estimate on the absolute values of the functions $S_i(q, \dot{q}, t)$ from (14) under the control (18). Using the constraints (7), (10), and (23), we have

$$S^0(X) = a\left(2\sum_{j=1}^{n} d_j X_j\right)^{1/2} + 2b\sum_{j=1}^{n} d_j X_j + 3c\left[\sum_{j=1}^{n}(d_j X_j)^{1/2}\right]^2, \; X = (X_1, ..., X_n) \tag{27}$$

Let us replace $|V_i|$ by their estimates (26) in inequalities (17) and substitute ρ_i from (19) into the obtained inequalities. We get

$$I_i X_i + \left(1 + \frac{\mu n^{1/2}}{I_{\min} - \mu}\right) S^0(X) \leq U_i^0 - G_i^0 - \frac{\mu}{I_{\min} - \mu}\left[\sum_{j=1}^{n}(U_j^0 + G_j^0)^2\right]^{1/2}, \qquad (28)$$

$$i = 1, ..., n$$

Suppose the parameter μ is sufficiently small so that the following inequality is fulfilled

$$\mu < \frac{\min_i(U_i^0 - G_i^0) I_{\min}}{\min_i(U_i^0 - G_i^0) + [\sum_{j=1}^{n}(U_j^0 + G_j^0)^2]^{1/2}}, \qquad i = 1, ..., n \qquad (29)$$

By virtue of (29), the expressions in the right-hand sides of inequalities (28) are positive. Since, according to (27), $S^0(X) \to 0$ as $X_i \to 0$, there always exist positive parameters X_i that satisfy inequalities (28) and, thus, inequalities (17).

The obtained results are summarized in the following theorem.

Theorem 1. Let conditions (5) – (10) and (29) hold. The feedback control $U_i(q_i, \dot{q}_i)$ solving Problem 1 is given by equations (18), where the parameters X_i should be chosen so as to satisfy inequalities (28). This bounded control brings the system (1) from the initial state (11) to the terminal state (12), if the initial velocities $\dot{q}_i^0$ satisfy the constraints $f_i^-(q_i^0) \leq \dot{q}_i^0 \leq f_i^+(q_i^0)$, $i = 1, ..., n$. The system does not leave the domain D defined by (3), and the time of motion τ does not exceed τ^* given by (24), (25).

Let us point out how admissible values of X_i can be chosen. We search for them in the form

$$X_i = Y^2 d_i, \qquad i = 1, ..., n \qquad (30)$$

where Y is unknown. Using (30), we rewrite inequalities (28) as follows

$$Y^2 + 2g_i Y \leq h_i, \qquad i = 1, ..., n \qquad (31)$$

where g_i, h_i are positive coefficients whose specific values can be found from (27) and (28). Then the solution of (31) is determined by

$$Y \leq \min_i[(g_i^2 + h_i)^{1/2} - g_i], \qquad i = 1, ..., n \qquad (32)$$

Now, we find the maximal value Y which satisfies (32), and then determine the parameters X_i from (30).

Applications to Robots. We consider a manipulation robot which consists of n rigid links connected consecutively by revolute or prismatic joints. Angles of relative rotation of links and their relative linear displacements for revolute and prismatic joints, respectively, are denoted through $q = (q_1, ..., q_n)$. Equations of motion for the manipulator can be presented in the form (1) with the kinetic energy T given by (2). The generalized forces here are the torques for revolute joints and the forces for prismatic joints. The terms U_i in (1) are the control torques and the forces created by the actuators in revolute and prismatic joints, respectively. The terms Q_i include all external and internal forces except the controls, namely, the weight, resistance, friction, etc. We assume that the forces Q_i satisfy (8) – (10) and that joints are driven by independent electric DC (direct current) actuators placed at the robot joints.

The kinetic energy T consists of the kinetic energy of the robot links $T^1(q,\dot{q})$ and the kinetic energy $T^2(q,\dot{q},N)$ of the actuator rotors, here $N = (N_1, ..., N_n)$ are gear ratios of the reduction gears. According to König's theorem, the kinetic energy T_i^2 of the ith rotor is equal to the sum of the kinetic energies of its progressive motion and rotation in relation to the center of mass

$$T_i^2(q,\dot{q},N_i) = T_i^v(q,\dot{q}) + T_i^\omega(q,\dot{q},N_i)$$

Let J_i and J_i' are the moments of inertia of the ith rotor with respect to the axis of the ith joint and the perpendicular axis passing through the center of mass of the rotor, respectively. Denote by ω_i and ω_i' the projections of the angular velocity of the stator of the ith motor on the axis of the rotor and on the perpendicular direction, respectively. Then we have

$$T_i^\omega(q,\dot{q},N_i) = \frac{1}{2}\left[J_i(N_i\dot{q}_i + \omega_i)^2 + J_i'\omega_i'^2\right]$$

The components ω_i and ω_i' are linear combinations of the generalized velocities $\dot{q}_1, ..., \dot{q}_n$ with coefficients that depend of q. Hence, the kinetic energy of the robot can be presented in the form

$$T = \frac{1}{2}\sum_{j=1}^{n} J_j(N_j\dot{q}_j)^2 + \frac{1}{2}N_{\max}(B\dot{q},\dot{q}) \tag{33}$$

where $B(q.N)$ is a $n \times n$-matrix such that

$$|B(q,N)z| \leq \lambda|z|, \qquad \lambda = \text{const}, \qquad \forall z \in R^n \tag{34}$$

for all $q \in D$ and all $N_i \geq 1$, $i = 1, ..., n$. Here and below, $N_{\max} = \max_i(N_i)$, $N_{\min} = \min_i(N_i)$, $i = 1, ..., n$.

We substitute expression (33) into Lagrange's equations (1). As a result, we obtain the following equations of motion

$$N_i^2 J_i\ddot{q}_i + N_{\max}[B(q,N)\ddot{q}]_i = U_i + G_i + S_i(q,\dot{q},t,N), \qquad i = 1, ..., n \tag{35}$$

Let us introduce a new variable p

$$q = Hp, \qquad H = \text{diag}(N_1^{-1}, ..., N_n^{-1}) \tag{36}$$

After the change of variables (36), equations (35) can be written in the form

$$(J + \tilde{B})\ddot{p} = M + G^* + S^* \tag{37}$$

$$J = \text{diag}(J_1, ..., J_n), \qquad \tilde{B} = N_{\max}HBH \tag{38}$$

$$M = HU, \qquad G^* = HG, \qquad S^* = HS$$

Here, M_i, $i = 1, ..., n$ are the electromagnetic torques created by the actuators. Hence, under the notations (36) and (38), the equations of motion (37) are presented in the form (5), (13). According to (34) and (38), we have the following inequality

$$|\tilde{B}z| \leq \mu^*|z|, \quad \mu^* = N_{\max}N_{\min}^{-2}\lambda \tag{39}$$

which is similar to the restriction (6). The initial and terminal conditions for the vector p from (36) are similar to (11), (12).

Let us consider two versions of Problem 1 for robots.

1) Suppose the control electromagnetic torques M_i, $i = 1, ..., n$, can vary arbitrarily within the bounds

$$|M_i| \leq M_i^0, \qquad i = 1, ..., n \tag{40}$$

Then the feedback control can be obtained by means of Theorem 1. Rewriting the restriction (29) in terms of the system (37), we can obtain the inequality that specifies admissible values of the parameter μ^*. Substituting in this inequality the expression (39) for μ^*, we obtain a restriction on the gear ratios of the reduction gears

$$\frac{N_{\min}^2}{N_{\max}} > \frac{\lambda}{J_{\min}} \left(1 + \frac{[\sum_{j=1}^n (M_j^0 + G_j^{*0})^2]^{1/2}}{\min_i(M_i^0 - G_i^{*0})}\right), \qquad i = 1, ..., n \tag{41}$$

Here $J_{\min} = \min_i(J_i)$; $G_i^{*0} > 0$ is a constant, which restricts the absolute values of the function G_i^*, $i = 1, ..., n$.

2) A more detailed analysis is required, if the input voltages of the motors are subjected to constraints. In this case, there is a need to extend the considered system (37). The equation of balance of voltages in the circuit of the ith actuator has the form

$$L_i \frac{dj_i}{dt} + R_i j_i + k_i^E \dot{p}_i = u_i, \qquad i = 1, ..., n \tag{42}$$

Here, L_i is the inductance, R_i is the electrical resistance, k_i^E is the constant coefficient, and u_i is the electric voltage in the circuit of the ith actuator. The torque M_i is proportional to the electric current j_i in the ith actuator

$$M_i = k_i^M j_i, \qquad k_i^M = \text{const} > 0 \tag{43}$$

The first term in (42) is usually small in comparison with other terms and, hence, can be omitted. Then we obtain from (42) and (43)

$$M_i = k_i^M R_i^{-1}(u_i - k_i^E \dot{p}_i) \tag{44}$$

Substituting M_i from (44) into equation (37), we obtain

$$(J + \mu^* \tilde{B})\ddot{p} = U^* + G^* + S^{**} \tag{45}$$

$$S^{**} = S^* - \Lambda \dot{p}, \quad \Lambda = \operatorname{diag}(k_1^M k_1^E R_1^{-1}, ..., k_n^M k_n^E R_n^{-1})$$

$$U^* = (k_1^M R_1^{-1} u_1, ..., k_n^M R_n^{-1} u_n)$$

The voltages u_i of the actuators are usually restricted by the constraints

$$|u_i| \leq u_i^0, \qquad i = 1, ..., n \tag{46}$$

where u_i^0 are constants. The constraints (46) are transformed into constraints imposed on the components U_i^* of the vector U^* from (45)

$$|U_i^*| \leq U_i^{*0} = k_i^M R_i^{-1} u_i^0, \qquad i = 1, ..., n \tag{47}$$

The equations of motion (45) are again presented in the form (5), (13). The constraints (47) have the same form as in (4). Clearly, in this case, we can consider Problem 1 for the system (45) under constraints (47). Now Theorem 1 can be applied, and the feedback control voltages can be obtained in the explicit form. According to the Theorem 1, we obtain a restriction which is similar to (41)

$$\frac{\Lambda_{\min}^2}{\Lambda_{\max}} > \frac{\lambda}{J_{\min}} \left(1 + \frac{[\sum_{j=1}^{n}(k_j^M R_j^{-1} u_j^0 + G_j^{*0})^2]^{1/2}}{\min_i(k_i^M R_i^{-1} u_i^0 - G_i^{*0})}\right), \qquad i = 1, ..., n \tag{48}$$

Conclusions. The proposed feedback control law brings the nonlinear dynamical system to the prescribed terminal state with zero velocities in finite time. For robots, this control is applicable, if the gear ratios and the parameters of robots satisfy the inequalities (41) and (48). The control is robust: it can cope with additional small disturbances and parameter variations. To ensure the robustness, we should decrease the parameters X_i, creating thus a sufficient margin in the control possibilities *.

References

1. Chernousko, F.L. (1990) Decomposition and suboptimal control in dynamic systems. J.Applied Mathematics and Mechanics (PMM), 54, 6, 727–734.
2. Chernousko, F.L. (1992) Feedback control for a nonlinear dynamic system. J.Applied Mathematics and Mechanics (PMM), 56, 2, 157–166.
3. Chernousko, F.L. (1993) The decomposition of controlled dynamic systems. In A. B. Kurzhanski (ed), Advances in Nonlinear Dynamics and Control: A Report from Russia, Birkhäuser, Boston, pp. 1–40.
4. Chernousko, F.L. (1995) Decomposition and synthesis of control in nonlinear dynamic systems. Proc. of Steklov Mathematical Institute, 211, 457–472.
5. Piatnitskii, E.S. (1987) Feedback control of manipulation robots. Izvestiya of USSR Acad. Sci., Tekhnicheskaya Kibernetika, 3, 92–99.
6. Piatnitskii, E.S. (1988) The decomposition principle in the control of mechanical systems. Doklady of USSR Acad. Sci., 300, 2, 300–303.
7. Krasovskii, N.N. (1970) Game Problems of the Encounter of Motions. Nauka, Moscow (in Russian).

* This research is conducted within the research grant No. 96–01–01137 from the Russian Foundation for Basic Research

FORMATION, DEVELOPMENT AND FUTURE PROSPECTS OF INFORMATION PROCESSING AND CONTROL ON SHIPS AND VESSELS

Voitetsky V.V.
Scientific&Production Association "Aurora" of the Ministry of Defence Industry,
15, Karbysheva Street, 194021, St. Petersburg, Russia
Telephone: +7(812)247-22-50, Fax: +7(812)247-80-61.

Abstract. The author presents retrospective analysis and characteristics of the main stages of ships and vessels` complex automation development, the most important scientific and technical achievements in theory and practice of naval automation control systems. The study gives a concept of creation and principles of structural formation of control systems for ships and vessels of the first decade of the XXI century.

Keywords. Ships` complex automation, control systems` structural formation, hierarchy of control systems, decentralisation of organisational, functional-topological and technical structures of control systems, reliability, survivability, safety of control systems, intellectualisation and self-organisation of control systems.

Introduction. As before, automation of information processing and control processes is one of the vital tasks for rising efficiency of naval ships (efficiency and accuracy of combat tasks achievement, level of combat readiness), as well as the increase of fail-safe operation, safety and survivability of naval ships alongside with minimisation of expenses for design, shipbuilding and maintenance. The global tasks for automation - the principal increase of naval ships and vessels` effectiveness (hereinafter ships) and labour productivity alongside with lower costs of public labour (live and materialised), and also the economy of live labour, with restriction of work in a hazardous and dangerous environment. It can be achieved only under conditions of complex automation, which assumes fully- centralised control over all shipborne functional complexes of facilities, (FCF) when the level and grade of its automation is justified economically (on the corresponding development stage of science, engineering and technology in the country). The retrospective analysis of development of formation principles and technical implementation of naval information and control systems is based on the example of complex control systems for technical facilities (CCS TF), developed by SPA"Aurora" during its 40 year history of naval ships` complex automation development [1] and analysis of theoretical & methodological basis of CCS TF formation [2]. It is possible to single out four stages of quality development for such large control systems as CCS TF and incorporated highly developed control systems of naval technical facilities functional complexes (CS TFFC). The CCS TF-based characteristic of these stages, which is mentioned below, relates mainly to nuclear submarines (NS). It is an example of the most developed large-scale control systems. The study also presents a formulated and scientifically-based concept of a prospective unified automated control system (UACS) of a ship and all its FCF. It is based on the home experience of development and maintenance of naval information and control systems, results of fundamental scientific research of Russian scientists in the sphere of large-scale control systems and achievements of foreign companies in utilisation of up-dated microprocessor controlled software-hardware complexes (SHC) in control systems type ACS TP.

Formation of ships complex automation. The first, initial stage of ships` complex automation is often linked with the deployed studies in the 60's, which focused on the development of small-crew, highly automated NS and partially with the development of CCS TF new type". Despite the accumulated experience in solving a number of basic problems of automatic regulation, design and putting into production enough large-scale naval automation systems [1], the development of new large-scale systems with extreme control centralisation of all technical facilities, a minimum number of operators and absence of continuous watch, moreover without local control posts (LCP) in peripheral compartments, required other, than already tested, methods of system organisation, its design and co-ordination of multiple departments and co-production enterprises. When the attempt was made to look at the problem of control for all TF as a structural-complicated system with an enormous number of components, requiring co-ordination of its interaction in the interests of the whole ship, the flaws in the theory of control had been detected.

Among the basic problems, which demanded radical theory development we can note the most acute:
1. Provision of the possibility of control from the central post (CP) of all TF of the ship (including its cruise), by the minimal amount of the operators with the absence of posts and watches in peripheral compartments during the duration of the endurance cruise. The control over nuclear-propulsion plant (NPP) is in charge of one operator during the entire routine, end and emergency modes. This is the main precondition and limiting term when looking for both basically new technical decisions concerning algorithmization of information processing and control processes and structural forming of the system and its hardware implementation. Previously, there was no such a precedent in home and foreign practice.
2. Provision of high levels of control reliability and effectiveness, workability and TF fail-safe operation in conditions of physical elements, inevitably increase. The problem of reliability then becomes a key one and determines the "to be or not to be" of complex automation.
3. Creation of a control panel, of accessible inboard dimensions, for its placement with a simultaneous true increase of information access to the operator and inevitable increase of controls due to the elimination of local posts and watches in peripheral compartments, which come into contradiction with human psychophysiological abilities.
4. Provision of operational functional and diagnostical control over CS TF and its integral parts' good working order during the cruise preparation period and during the cruise itself, with fast detection of damages and recovery of system and its components' workability.

At this period, there had been an "explosion" of dynamic research of complicated naval control objects, and simulation technology on analogue and digital computers had been worked out, including objects with distributed parameters. Mathematically complicated tasks of synthesis of optimal algorithms for nuclear-propulsion plants' control and nuclear submarine manoeuvres, with the use of updated variation methods had been solved. The principles of control with forecasting, adaptation and self-trimming, methods of forming fail-safe devices for information logical processing had been developed. The basic problems for achieving the appropriate levels of reliability for such scale control systems and ergonomic provision of the operator's activities had been solved. Special attention was paid for matching the features and characteristics of control systems with real human abilities during work, as a central link of the control system for the multi-assembly complex of technical facilities and their aggregates. The last mentioned point required the necessity to conduct a number of theoretical and experimental works not only to select the type of control system, which is vital for utilisation of human capabilities, but also to match control panel designs with the requirements and recommendations of engineer psychology. Ideas of a hierarchical method for information processing and display were suggested, and had been implemented. Semantic principles of its presentation to the operator had been developed, which enabled the operator to quickly and simply diagnose the state of the object. Experimental methods for the evaluation of the operator's functional state and analysis of his activities had been developed and practically used. These methods are based on usage of visualograms and their analysis (analysis of the operator eyes` trajectory during an information readout from the control panel), which reflects the path of control and diagnostic processes in the object being solved by a human being, and determines the essential peculiarities of thinking in emergency situations. A scientific basis, methods of forming information presentation systems, with related features of a control object, its control functions, psychological peculiarities of information perception by a human being had been created. In the process of researching the development problems of large man-machine control systems it became evident, that the operator must be in a calm state of mind, because he is constantly waiting for an emergency or damage. Technical decisions were found to ease tension, which provide systematic forecasting of possible failures, which normally mature gradually. The key scientific and practical result was elaboration and implementation on the first stage, of a complex automation of functional-hierarchical principle of control and arrangement of control systems, which became the basis of forming all successive generations of CCS TF. Grounding and elaboration of hierarchical structures of naval control systems refer back to the beginning of the 60's. The main developments have author certificates [2]. Similar structures later were published in the US by M. Mesarovich.

The development of ships` complex automation. The second stage (the 70's) is directly connected with the creation of CCS TF for new nuclear submarines, which took into account the higher requirements for ship survivability, complexes of its technical facilities and thus to survivability of technical facilities control systems [3]. It required further development of methods and principles of control systems` structural arrangement.

The decisive influence on the development of principles of CCS TF structural arrangement was exerted by requirements for failure safety [3], which characterise the feature of control systems not to lead the controlled TF equipment to the state of emergency, and do not cause emergency situations on board the ship under inevitable conditions, during long-term working cycle faults of the components, short circuits and breaks in communication lines and circuitry, various disturbances in electric power supply systems of electrical, electronic and hydro equipment of CS TF . Some influence on structural arrangement of control systems was exerted by higher requirements to failure safety and fault-free functioning of CCS TF, caused, in particular, by a drastic increase of a continuous operation working period (time intervals of endurance cruise), requirements to repair suitability and endurance (overhaul period and service life) of CCS TF devices.

In order to meet the requirements of survivability of CCS TF, the principles of functional and territorial decentralisation of control systems had been designed and scientifically grounded, gave in detail and implemented. They led to the creation of complex control systems with basically distributed topological (or functional-topological) and organisational structures [3]. In this systems there was generated transition to automation of decision-making processes on the upper, commander level of ship control.

The necessity of the absolute provision of CS TF failure safety features, high levels of fail-safe and durable continuous functioning required the development of methods and improvement of technology, related to structural, in particular majority redundancy of discrete and continuous (analogue) control devices, usage of specific structural decisions, which provide protection for CS TF in emergency modes of operation, against false actuation and faults in operation of discrete control channels, usage of two-pole commutation of discrete armature actuation mechanism electromagnet windings (of actuation bodies), developed functional and diagnostical control of equipment.

The principle system decision in the ideology of system arrangement, was practical implementation of territorial decentralisation of CCS TF functional-topological structure principle, in order to provide an increase of the vitality level of a ship and its automated TF survivability levels, including crew structure survival, provision of fail-safe and safety of automated complexes, nuclear safety in particular. It can be stressed, that the principle of territorial decentralisation is added to the functional-hierarchy principle and provides centralised control over all TF complexes with basically higher levels of control processes vitality, compared with the preceding systems.

The third stage (the beginning of 80's) of complex automation is connected with the creation of CCS TF for a new generation of nuclear submarines and large surface ships. This systems are using well-shaped, during the previous stages, fundamental principles of control systems` arrangement and experience of its operation. At the same time, they represent a quality breakthrough in ships` automation process. From the point of view of system organisation and principles of information processing and control, the principle distinguishing features are:

- usage in contrast to 1 and 2 stages of complex automation the home-made microprocessor technology (LSI, VLSI), and displays for information mapping,
- usage of new principles of control and information processing, algorithmically developed methods of functional and diagnostical control over TF equipment and CCS TF devices,
- high increase of control processes automation level in the systems of all levels of ship's hierarchy, in particular in emergency situations, which are caused by multiple faults and damages of the main equipment and during damage control,
- improvement of man-machine interface, utilisation of new, more advanced methods and means of data presentation for operators in order to rise the effectiveness of their activities, also in non-standard situations, extreme modes and during damage control,
- more elaborated functional and territorial decentralisation of control systems with dispersion of functions of centralised control of technological parameters (STSK) between control systems of single types of technical facilities,
- a basically new stage in rising the grade of ship TF control processes automation had been implemented, i.e. automation of decision-making processes on the highest level of control over the whole aggregate of TF functional complexes in all normal and emergency modes of operation and during damage control.

The structure of CCS TF incorporates a new co-ordinating control system of Technical Facilities with a specially arranged control panel. This system characterises transition to a quality new stage of ship`s complex automation. This system is considered in accordance with [3] generalisation of information, concerning the state of TF and the internal situation on the ship through their control systems in all modes of operation, localisation of emergency situations, which cannot be solved within control systems of separate TF types. It also provides recommendations for the operators how to control technical facilities during damage control tasks solving. An increase of efficiency and accuracy (quality) control decision-making during damage control, organisation of effective co-ordinated interaction of personnel at all levels of control and provision of co-ordinated functioning of the whole aggregate of automated TF, is one of the most important aims in co-ordinating system creation. It is important to stress, that this system, together with other information-control systems of the same (commander) level of hierarchy, in particular with the combat information-control system (CICS), in the future provides organisation of ship unified automated control system (ship UACS), as an aggregate of large-scale control systems, interacting within regulated hierarchical structures.

Future prospects of naval ships and vessels complex automation. At different periods, the authors paid much attention to the forecasting of naval complex automation systems development for the future 20-30 years (e.g. see [4]). Among it we can note the forecast of US expert M. Murphy]. Murphy predicted, that in the year 2000 the ship will be completely automated and controlled from the coastal control centres - initially at the departure port, then - at the port of destination, without crew on board the ship. But despite substantial success in automation, which has been achieved by the end of the second millennium, we can not say, that this forecast is realised. Is there any reason for that at all? Because a large ship can become damaged due to failure, which is hard to foresee. But mainly, it is not expedient to have control centres in each port to manage hundreds of ships [5]. But, from the technical point of view, taking into account the achievements in the sphere of control systems, accumulated database, expert diagnostic and real-time control systems, the system of global communication, it can be implemented.

In this work, we connect the future prospects of ship automated control systems` and its functional complexes development with the forth stage of complex automation, which started in the middle of the 90's for a ship which will be built after the year 2000. The changes of the social-economic system in the country, and liberalisation of external economic relations with western countries left a serious imprint on the further development of the complex automation process. One of the decisive tasks in the future is to eliminate the lagging in development of microprocessor-operated complexes and software design, to build automated control systems for potentially hazardous objects as nuclear ships, power units of nuclear power stations and other objects which have hard working conditions for the equipment located there [5].

The characteristic features of a new generation of automated control systems (ACS) of future ships and incorporated technical facilities, ACS of radio-electronic equipment (functional complexes of navigation, mapping of ambient situations, external communication), ACS of target assignment complexes, automated commander control complexes, will be:
- usage of network architecture during realisation of distributed control and information systems with unconditional implementation of the functional-hierarchical principle of its building, principles of functional and territorial decentralisation,
- provision of higher than previous levels of survivability, fail-free safety on the basis of sound scientific-technical experience in the theory of reliability, survivability, safety and efficiency of large-scale systems, also on the basis of developed principles of decentralised control over reconfiguration of structures, decentralised control of information exchange between information processing centres, dynamic re-distribution of tasks and functions among information processing levels during faults of information processing centres (network centres) of any assigned rate and, finally, the principle of cannibalism,

- intellectualisation of control systems of all ACS hierarchy levels, including realisation of, developed and tested in STA, "Aurora" technology for the creation of control systems with database, which implement the functions of intellectual, calculating more sophisticated types of information support for decision-making, made by operational personnel and leading officers of the ship in non-standard, unpredictable emergency situations and during damage control,
- integration of ACS and onboard intelligible (intellectual) simulators for self-training (without instructor) of operational and leading officers with built-in expert system-technologist.

These features will provide the ship's ACS with basically new qualities, supported by the appropriate selection of its structure and functioning, "implantation" in the system of automatic mechanisms of "self-preservation" and "self-adaptation", i.e. homeostasis mechanisms. At the same time, practically tested fundamental principles of structural organisation and complex systems, real necessity to select categories of local control systems (LCS) will be unchanged. LCS are, in accordance with [3], the components of separate ship-based equipment (unit, set of units, etc.) and provide the very possibility of physical processes in this or that unit, fail-safe operation, diagnosis of the unit state, carry out, when necessary local control, and also receiving and processing of commands from the control system of a higher level and presentation of generalised information about the state of controlled equipment. The selection of categories of local control systems, which are not incorporated in ACS, is intended to test the TF equipment in factory conditions (before delivery to the ship) and to conduct complex tests. Besides, the development of local control systems as an integral part of ship-based equipment, is expedient during development of multi-purpose equipment, designed for utilisation within TF of ships of different types and assignments (e.g. automated refrigerator unit, automated turbo-generator or diesel-generator, etc.).

The results of research and technical studies show the possibility to implement the task of centralised control over all the facilities of the ship, by 2-5 operators, depending on ship type and assignment, without continuous watch in peripheral compartments and control posts.

Conclusions. The main scientific and technical achievements of home industry in the sphere of ships' complex automation, first of all are connected with the successful solution during the last three decades of the most complicated problems of ship control processes and its large functional-control systems automation. Large-scale man-machine control systems were developed for such multi-unit complexes as nuclear-propulsion plant, electrical power system, aggregate of all ship-based equipment, devices and mechanisms, complex of propulsion means and manoeuvring. The volume, which has been achieved, (coverage) level and grade of automation of control enabled the following :

1. considerable reduction in the number of personnel, which took part in information processing and TF control (3-5 times, compared with known foreign analogues);
2. the increase of the quality of control processes, furthermore the quality of TF functioning;
3. to increase operating reliability and safety of TF complexes, considerably reduce losses, caused by emergency situations in TF;
4. to raise the survivability level of the ship and its TF complexes, cruise safety during assignment of combat, maintenance and other damages.

Practical realisation and implementation of TF complex automation designs became a reality, due to the successful solution of a number of basic scientific and technical problems, connected with the provision of high level of automation and quality of information processing and control, under high levels of workability and functioning reliability of large-scale man-machine control systems, covering multi-functional complexes of TF, which are interconnected by common tasks of functioning or unity of physical processes. Among the most important and widely tested principles of ACS building and incorporated large-scale control systems, it is necessary to distinguish the functional-hierarchy principle of system structure organisation, decentralisation principles of functional, topological and organisational structures, specific principles for provision of control systems` survivability and safety. Successive realisation of tested principles enabled the following :

1. centralised or co-ordinated control of all aggregates of TF complexes (complexes of energy provision, cruise provision, provision of stabilisation and manoeuvrability, provision of inhabitancy, provision of unsinkability and ship`s combat ruggidness), co-ordination of an enormous number of processes of interconnected units, realised within hierarchical type of control with the help of a organised hierarchy of functional sub-systems (systems) of control ;
2. the principle possibility of one-man control over such multi-unit complexes as a nuclear-propulsion plant, even in emergency situations, caused by failures of some components of automatic control;
3. possibility of a parallel solution of huge amount of tasks of control and control under limited conditions (capacity and memory size, etc.) for means of information processing;
4. reduction of information flows and volumes, passing from controlled objects to higher levels of hierarchy due to generalisation of information within functional sub-systems of lower levels of control;
5. possibility to find compromise variants of control systems structures, which meet the requirements` contradicting demands of high levels of control processes quality (dynamic features), on the other hand, high levels of reliability, aimed at provision of fail-safe functioning of TF and cruise safety, reduction in "price" of components failures, which are inevitable during long-time operation;
6. possibility to organise parallel development, production, and to put functional sub-systems into action, due to relatively small number of information or functional links between them;
7. possibility of independent updating of separate sub-systems, without radical re-arrangement of the whole system, including the possibility to change sub-systems, which produces its service life.

References.

1. Voitetsky V.V. Formation and development of the Association and the main results of its scientific and production activities. - In the jubilee scientific-technical Collection, devoted to the 25-th anniversary of SPA"Aurora", St. Petersburg, 1995, page 3-20. (in Russian).
2. Voitetsky V.V., Simakov I.P. Development of methodology, theory and principles of naval ships CCS TF and control systems type ACS TP, technical complexes with operation of a high level of risk. See the same, page 46-86. (in Russian).
3. Control systems of ship technical facilities. Terms and definitions. - State standard GOST 19176-85. (in Russian).
4. Jung V.N. Automation of ship-based technical facilities at the beginning of the XXI century. - Shipbuilding, 1980, № 10, pages 25-30. (in Russian).
5. Prangishvili I.V. Modern state and prospects of automated control systems for complicated, potentially hazardous energy objects, chemistry and petro-chemistry. Devices and control systems, 1993, № 7, pages 1 - 5. (in Russian).

BASIC GENERAL SYSTEM LAWS OF CONTROL OF COMPLEX SYSTEMS OF VARIOUS NATURE IN CRISES

Academician Iveri Prangishvili,
Institute of Control Sciences,
Russian Academy of Sciences,
65, Profsoyuznaya St., Moscow , Russia

Abstract. A scientific approach to system-based control of investments and other managerial processes in citical transient period of making a market economy is presented. Consideration is given to the main system-generating attributes and eight main universal system laws governing the operation of complex systems of any nrature.

Keywords. General system laws, economics, complex systems, analysis, attributes.

A scientific approach to system-based control of investments and other managerial processes in citical transient period of making a market economy is presented. Consideration is given to the main system-generating attributes and eight main universal system laws governing the operation of complex systems of any nrature. Their application to economics enables one to determine the most efficient points of time for investing minimal resources and to find mechanisms for merging several opposites into stable systems or, vice versa, for decomposing systems into several opposites with emission of the disintegration energy. The law of cause-effect relationships is described in form of signed graphs or cognitive maps. It is proved that when resources in a system are moved upward to the next level of the pyramid, the system loses stability. A new system law of manifestation of system instability or crises through coherent (coordinated) interaction with the background of the observed system is proposed. It enables us to detect and recognize the desired objects or processes by analyzing the coherent system background. Changes in the nature of interaction between the system elements are shown to cause substantial changes in the resulting system potential. The main laws of the general systems theory and control theory can be used to determine the possible degree of sovereignty, evolution, and fate of systems and individuals.

System-generating attributes and eight main system laws

From the point of view of control and the general systems theory, any complex (engineering, economic, ecological, social, etc.) system features system-generating attributes of which major ones are as follows: (1) **primary** system elements, (2) **relations** between system elements, (3) laws of **composition** of system relations, and (4) system **background. Primary** elements can be industries, economic regions, administrative unities. The elements of an economic system are **related** by competition, cooperation, neutrality, etc. **Laws** of composition of relations comprise formulas, empirical regularities, and laws of operation of complex systems. System **background** corresponds to internal stable oscillatory processes supporting composition of system relations.

Let us consider the **eight laws of composition** governing operation of any complex system - be it engineering, economic, social, ecological, managerial, etc, system. The eight fundamental system laws define largely the processes of system control - **what, when, and how** to control. They enable one to determine **when and what** should be invested, the outcome of incorrect investing, and how to manage complex systems more efficiently. It is namely the system insight into engineering, economic, social, and other processes that enables one to tackle complicated managerial problems [1, 2, 3].

First system law is that of "driving the system from one qualitative state to another by a minimal action in the critical point of system phase change." To understand it, we note that the crisis state of any system is a kind of phase change from one state to another with its own critical point where transition is caused by a minor action upon the system. It is well known from the secondary-school textbooks on physical chemistry that there exist phase changes of substances from one qualitative state to another and critical points or instants of this change which are often referred to as points of crystallization (melting, boiling, evaporation, transition to the gaseous phase, or bifurcation). Here, a slightest action by the "crystallization nucleus" suffices to start a snowballing transition of the substance from one qualitative state to another.

Studies have demonstrated that the economic crisis of branch, industry, region, or country also is a kind of phase change with its critical (bifurcation) point and that if at this point and time one introduces a crystallization nucleus, to which small investments correspond, then minor effort could result in appreciable economic effect. Judging from the general systems theory and theory of control, this analogy between systems of different nature and types is justified.

In our crisis economy, an active process of **disrupting the existing established connections and eliminating the old structures and forming new connections and structures** corresponds to phase change. The point of time when the **old structural connections** are eliminated or very weakened and the **new ones have not yet formed or gained a foothold** corresponds to the critical point (time) of phase change. It is precisely here that the economic system of a branch, industry, or region can be driven to a new qualitative state by smallest investments playing the role of the crystallization nucleus, because then the rest will be finished by the system itself.

Second general system law is the **"system-genetic law or law of evolution"** saying that within individual evolution any system briefly repeats the evolutionary path including all stages (similar to the evolution of the human foetus in the mother's womb). Neither in economics, nor ecology, nor social spheres one can skip any stage by making the Great Leap forward. Independently of whether we want it or not, all evolutionary stages must be passed through. Their duration can be artificially retarded or accelerated, but they cannot be skipped. Therefore, to create a stable market economy, the stage of primitive accumulation of capital must be passed despite all its negative phenomena [2].

Third system law is the "law of pyramid" relying on the first and second laws of thermodynamics saying that the efficiency of any physical system cannot reach 100% because a part of its energy is dissipated into the environment. In economic systems, the external initial energy moves from level to level and concentrates in the product. In industry, for example, first raw materials are produced, then parts, then products. At each production stage the initial system energy tapers down like a "pyramid" because, similar to the law of energy dissipation in thermodynamics, it is used to interact with the environment. It turns out that, according to the law of pyramid, an industrial or economic system starts to **"vibrate" and loses its stability** when appreciable amounts of money, resources, or energy are fed into the next level of pyramid [1, 2].

Fourth system law, which is called the **"insular effect,"** establishes the dependence of the degree of sovereignty or isolation of an island or system (for example, sovereign state) of its population, area, and of quality and quantity of its natural resources. This law also rules out the possibility of longtime existence of a powerful socioeconomic or ecological system in a primitive environment - for example, in the environment of petty separatists. The lovers of now-so-popular separatism must be aware of this law.

Fifth system law, which is applicable to complex systems of any nature that are formed by merging ("pasting") opposites (antagonists - local subsystems with opposing objectives) or, on the contrary, by dividing ("splitting") a complex system into opposites (or antagonists), is the **law of unity and conflict of opposites**. It defines the possibility and conditions of merging several opposite systems into more powerful systems or splitting a system into several subordinated opposites (antagonists).

This law is universally applicable to complex systems of various nature and can explain many (homoeostatic, as a rule) processes in complex systems that feature interaction of antagonists and more than one feedback loop. Here, we are interested in considering the homoeostatic mechanisms of control as a unity of opposites with controllable internal contradiction playing in such systems the role of "spinning reserve." Notably, "pasting" of antagonists (opposites) into stable systems and "splitting" of such systems with emanation of the "degradation energy" allow us to identify some interesting properties of complex systems.

If the necessary and sufficient conditions of "pasting" are satisfied, then two potential antagonists, both structurally stable and unstable, and their combinations (one stable, one unstable) can be merged into a structurally stable system. The necessary condition of "pasting" two antagonistsis is that they be mirror symmetrical, the opponent of an antagonist playing the role of feedback. This results in a stable system featuring as if double feedback and eight possible combinations of such "pastings."

From the standpoint of the control theory and systems theory, disintegration of a complex chemical compound, enriched uranium, state (the USSR, for instance), or economics are system "fissions" accompanied by emanation of a good deal of **"fission energy."** Sometimes, this process is followed by a tendency to "secondary" splitting of the separated antagonists. For example, the separation of some states in the course of disintegration of the USSR was accompanied by emanation of a great amount of "splitting" energy in places where processes of secondary disintegration develop (for example, Moldavia, Georgia, and Azerbaidjan). This energy plus low stability of multinational states such as Moldavia, Georgia, or Azerbaidjan cause processes leading to their possible secondary disintegration. Of course, other, stabilizing processes superimpose on them, and it is still difficult to forecast their outcome, but the processes observed today in the post-Soviet states could have been predicted from this simplistic behavior of splitting homoeostatic system.

We note that, because of lost system stability or compulsory splitting, "recession" of opposites takes place accompanied by emanation of energy that is in power dependence of the contradictions between the antagonists and their instability. In doing so, the emanated energy can be either absorbed by the environment, or by the separated opposites, causing their secondary disintegration. Yet, we emphasize here that it is possible to prevent this secondary disintegration by timely "pasting" of the opposites [3].

It follows from the properties of the "pasted" systems that if the control structure of an antagonist deteriorates (for example, a feedback is disrupted), then the nature of their interrelations changes and, instead of parity, subordination appears, the antagonists with weakened control system becoming satellites of other antagonists. To illustrate this statement, the worse is management in the Russian Federation, the more this country becomes an American satellite and the more the USA become the leader. This follows from the properties of systems consisting of two "pasted" antagonists.

If internal contradictions in a system that was constructed by "pasting" antagonists reach a value near the stability limit, then the system can be rescued from disintegrating by applying to its inputs contradictions of an opposite sign (the authorities abide to this method to prevent social explosions) [3].

Sixth system law is the law of cause-effect relations. To gain insight into a complex system and analyze its behavior, often the block diagram of its cause-effect relations is constructed. Let us explain the gist of this law. If the element A is the cause and the element B effect, then they are related by a cause-effect relation representable as two vertices A and B connected by an oriented positive or negative arc (positive or negative cause-effect relation). The relation between A and B is called the positive cause-effect relation if an increase in A leads to an increase (gain) in B, and vice versa. Otherwise, the cause-effect relation and the sign over the arc connecting A and B are negative. Such graphs, which are called the oriented signed graphs or cognitive maps, are most suitable for understanding and analyzing behavior of complex engineering, socioeconomic, ecological, managerial, and other systems [1].

The cognitive map or oriented signed graph enable one to readily analyze strange human behavior, as well as views and opinions of decision makers. The apparatus of signed graphs is especially suitable for system-oriented thinking and analysis of practically unmeasurable factors [1]. The relationship between human thinking and the situation into which persons get involves a cognitive function of understanding the situation and control function for applying their conclusions to a real situation [1]. In view of some experts, the cause-effect law helps to forecast not only the behavior of complex systems, but also human fate - if some person does something essentially evil (good) in the beginning (cause), then he/she gets evil (good) at the end (effect). This means that human malfeasances return to them in one or another form as a just retaliation (cf. "As a man sows, so shall he reap.") [4].

Seventh system law seems to be that of "manifestation of instabilities or system crises in coherent interaction with the background of the system." By system background are meant stable coherent oscillatory processes within the system supporting composition of system relations. This general system law makes use of the background principle of detecting a moving object from its coherent interaction with the background [5]. Coherent interaction with background, for its part, consists in forming within the observation (control) system of oscillatory processes acting in coordination with the background of the observed (controlled) system and in tracking this coordination (correlation). Detection of violation of this coordination (correlation) suggests the beginning of instability or crisis, which is the counterpart of the invisible moving object in the principle of background [5]. The seventh law (in form of the background principle) has been realized in the detection of objects in radio location, security systems, biology, psychology (human ecology), and protection of objects against technogenic accidents and catastrophes. It is planned to use further this principle in socioeconomic systems, ecology, etc.

The reason for using the background principle is based on the fact that the objects or processes we are interested in (for instance, an imminent catastrophe, beginning of war, etc.) are practically unobservable, whereas the environmental information (background) is quite observable and enables us to detect the imminent danger and to take preventive measures. The background principle is actively used by the intelligence services that learn about the oncoming processes from nontypical perturbations of the environment. The background principle enabled the advent of a basically new and highly efficient nonlocational method of detecting weakly dissipating nonreflecting (absorbing and transparent) objects based on the "stealth" technology or nonreflecting materials. The signals about an object are obtained from its background, rather than from the object itself. Since each object leaves its trace in the background or environment, the coherent component of background carries information about the object [5].

Eighth general system law is the law of "essential dependence of the system potential on changes in the nature of interaction between the system elements."

The potential E of system A depends substantially on the nature of interactions between the autonomous primary elements of the system having each the potential [6]. For example, if systems are represented by branches, industries, or countries, then we deal with respective potentials.

If interactions of the primary elements of systems (subsystems) are goal-oriented and coordinated (for example, the cells of an organism), then the potential E of system A exceeds many times the total potential of its constituent elements:

$$P(A)>[p(a_1)+p(a_2)+p(a_3)+...+p(a_n)].$$

If interactions of the primary elements of a system is uncontrollable ("random"), the potential E of A corresponds to the potential of its individual averaged element a:

$$P(A)=[p(a_1), p(a_2), p(a_3),..., p(a_n)].$$

If interactions of the primary elements of a system are totally antagonistic, that is, each is at war with all others, the system potential is less than that of the weakest element:

$$P(A)< \min [p(a_1), p(a_2), p(a_3),..., p(a_n)].$$

It follows from above that insignificant changes in the nature of interactions between system elements produce striking differences in the resulting system potential.

Conclusion

We can conclude that the above eight general system laws, which are known from the general system theory and control theory, enable us to carry out analysis of and gain insight into the basic processes occurring in crises in any complex system independently of its type or nature. This is an attempt to approach from scientific standpoint analysis and solution of specific problems of control in complex systems of engineering, socioeconomic, and other nature. The first and third system laws enable us to establish the most advantageous points of time for making investments into a system and the conditions for investing resources to retain system's stability. The second law indicates to the evolutionary path that must be passed by any normal system within its individual development. The fourth law establishes the dependence of system's sovereignty and isolation vs. its major indices. The fifth law defines the conditions for merging several opposites into a more powerful system or, vice versa, for decomposing one powerful system into several antagonists with emanation of the "splitting" energy. The sixth law provides an insight into behavior of complex systems or humans, views and opinions of decision makers. The seventh law enables one to detect moving objects "invisible" to radar and hydroacoustic systems, create new security devices and systems protecting against technogenic accidents and catastrophes, to design detection and control systems for ecology and socio-economic systems. The eighth law establishes the dependence of system potential on the nature of interaction of its constituent subsystems and demonstrates that their goal-oriented coordinated interaction could result in system potential many time exceeding the sum of subsystem potentials.

References

1. I.V. Prangishvili, "Features of managing national development in predictable and stable manner," Prib. Sis. Upravlen. No.12 (1996).
2. A.S. Martinov, V.V. Artyukhov, V.G. Vinogradov, and others, Russia: A Strategy of Investments in the Period of Crisis [in Russian], PAIMS (1994).
3. Yu. Gorskii and V. Lavshuk, "Life and death of a civilization (model, forecasts, role of intelligence and information)" [in Russian], Express-Informatization, Int. Acad. of Informatization, Center of Strategic Studies, LISNYa, Irkutsk (1995)
4. Ya. Zubtsova and I. Abramenko, "Thoughts with comfortable envelope," AiF (June 1996), p. 16.
5. I.V. Prangishvili and A.N. Anuashvili, "Background principle of detecting moving objects," Autom. Remote Control, 58, No.4 (1997).
6. A.A. Bogdanov, Tectology: A General Science of Organization [in Russian], Moscow (1929).} □

CONTROL OF NONLINEAR OSCILLATIONS AND CHAOS

A.L.Fradkov
Inst. for Problems of Mech. Engg.
61, Bolshoy, V.O. 199178, St. Petersburg, Russia
alf@ccs.ipme.ru

Abstract. A survey of the class of control problems is given related to oscillatory and chaotic systems. The classification of the problems is presented in terms of control goals and potential applications.

1. Control goals. The importance of oscillatory processes and models has been recognized in many fields of science and technology. In many applications it is necessary to modify the behavior of the system, e.g. to change the amplitude or the period of oscillations, to transform chaotic (irregular) motions into periodic (regular) ones and vice versa and so on. Such problems can be understood as those of control. Recently a lot of interest has been attracted to the control problems for chaotic systems.

Traditionally the two main kinds of control goals are considered in the control theory: *regulation* and *tracking.*

Goal 1: Regulation (or stabilization, or positioning) is understood as approaching the variable $x(t)$ of the control system to some equilibrium (fixed point) x_*. Sometimes only regulation of the output variable $y(t)$ is significant. Formally we can write the goal 1 as a limit relation

$$\lim_{t\to\infty} x(t) = x_* \tag{1}$$

or

$$\lim_{t\to\infty} y(t) = y_* \tag{2}$$

Goal 2: Tracking means approaching the variable $x(t)$ to the *desired trajectory* $x_*(t)$, i.e.

$$\lim_{t\to\infty}[x(t) - x_*(t)] = 0 \tag{3}$$

or

$$\lim_{t\to\infty}[y(t) - y_*(t)] = 0 \tag{4}$$

The desired trajectory is often referred to as *reference trajectory* or *command signal.* Sometimes the reference signal is defined explicitly as a function of time, otherwise it may be defined as a solution of another auxillary system called the *reference model*, or the goal model. In the latter case the problem of finding the controller ensuring the goal (3), or (4) is called the *model reference control problem.*

Turning to the control of oscillatory behavior we can see at least two more kinds of the control goals:

Goal 3: Synchronization is understood as coincidence or coming close the variables of two or several systems; or concurring change of some quantitative characteristics of the systems. The synchronization problem differs from that of model reference control problem because it allows for coincidence of different variables taken at different time instances, i.e. with time shifts, which may be constant, or tending to constants (asymptotic phases). Also in the typical synchronization problems the bidirectional connections between systems are assumed. It means that the desired limit solution can not be prespecified. On the other hand synchronization with unidirectional connections (master-slave or drive-response synchronization) may be of interest too.

The synchronization phenomenon in systems without control (self-synchronization)is well known and well studied [2, 3, 38]. However synchronization as a control goal was paid much less attention. Its investigation started only recently [51, 48, 16]. The general view of synchronization problem is presented in [5].

Goal 4: Modification of asymptotic behavior of the system. This class of the goals includes various special cases as follows:

- change of the type of equilibrium (e.g. turn unstable fixed point ito stable one and *vice versa*);
- change of the type of limit set (e.g. turn limit cycle into chaotic attractor and *vice versa*);
- change of the bifurcation point and/or the type of the bifurcation point in the parameter space;
- create oscillations posessing the desired properties.

The important special case of the last subclass is represented by the so called *swinging/unwinding* which means excitation of the oscillations having the desired energy (or frequency or another parameter of intensity. Related problem is the modification of fractal dimension of the invariant compact set (attractor). Some general approaches to the swinging problem are described in [17, 19].

An important distinctive feature of the control of oscillations and chaos is problems that the small level (or low power, or cheap) control is required. The reason is that the life time of an oscillatory system covers large number of cycles and the average power of control action per one cycle should be small to remain the total energy of the control bounded. Fortunately, in many cases the "small control" requirement can be achieved, see below.

Chaotic systems from the control theorist's point of view are just particular class of nonlinear dynamical systems having irregularly oscillating solutions. Since oscillations can be either periodic or nonperiodic, control of chaos can be considered as a special case of control of nonlinear oscillatory systems.

However problems of control of chaos have distinctive features which were first pointed out by E. Ott, C. Grebogi and J. Yorke (1990) [49]. Namely chaotic trajectory may be turned into periodic one by means of arbitrary small control stabilizing inherent periodic orbits.

To be rigorous, mathematicians were aware of such a possibility from the so called "closing lemma" or Pugh lemma [55] which holds even for wider class of recurrent trajectories. However Ott, Grebogi and Yorke studied this property as that of control and interpreted it as a possibility to either create or suppress chaos by means of tiny corrections. It opened new perspectives in both natural sciences and technology and initiated an avalance of publications.

In [18] an adaptive control algorithm motivated by the method of Ott, Grebogi and Yorke (OGY) has been presented. The conditions of the control goal achievement for recurrent motions which are more general than chaotic ones have been presented.

Paradoxally it is not necessary to know what the chaos is to be able to control it. However this does not surprise control theorists and engineers who know that good control of the plant does not necessary require good model and good prediction. Indeed, feedback can suppress the errors. Moreover controlling a plant we do not need to know the uncontrolled plant behavior. Who knows, may be it will be possible to control weather earlier than to predict it.

And now let us list the applications of the control of oscillations and chaos.

2. Applications. As it was already mentioned one of the reasons of emerging interest to control of chaos was a wide range of its potential applications. Therefore it is important to give an idea of possible applications to the reader. Again we will be talking about control of general oscillations rather than about control of chaotic oscillations. The range of applications is indeed so wide that each heading in the list covers the whole branch of science and technology.

2.1. Mechanics and mechanical engineering. A variety of swinging and synchronization problems for mechanical systems arises in the design of vibration equipment: vibrotransporters, mills, crashers, etc., (see [2, 3]). The simpliest one is swinging the one degree-of-freedom (pendulum-like) systems. The solution of this problem may help to launch vibroactuators with low power motors, to overturn large vehicle or ship in case of emergency, or even to draw the car out of the hole.

One more example is known from robotics where the so called "brachiation robot" was developed which is a kind of mobile robot that moves using its arms like gibbon moving from branch to branch [60]. To design a locomotion control for such a robot some swinging amplitude control algorithms were suggested in [60].

Synchronization problem also has numerous engineering applications, see [2, 3, 4]. In the most interesting applications it is required to achieve synchronization of subsystems by introducing small interconnections between them. If the desired synchronization mode is not achieved without external action then the problem of controlled synchronization arises, see [5]. Creation of stable synchronous oscillatory and rotatory motions is also important for advanced vibrational technology. As to chaotic motions they turned out to be useful, e.g. for drilling because chaotic rotation of the drill has been proved to make more smooth surface of the hole than periodic rotation makes.

One more example is the problem of control of a ship rolling subject to lateral ocean waves [12]. The ship's rolling dynamics can exhibit chaos even for waves purely periodic in time. Thus the problem is to decrease the amplitude of chaotic oscillations under the presence of disturbances. Again only small level of control is admissible.

Finally, methods of creating or suppressing chaos can be applied in washing machines. It is known that washing may be accelerated if the angular velocity of the rotor (activator) is oscillating. Moreover the desirable kind of oscillations is chaotic because it provides better mixing of loundry and detergent.

2.2. Electrical engineering and telecommunications. The interest in oscillatory behavior and synchronization in electrical and electronical engineering is rapidly growing together with the growth of these branches of technology. Till recent years it was concerned mainly with periodic oscillations [38]. Even for this case the nonlinear control theory may be of practical use, e.g. for enlargening the capture range of a phase-locked loop in the state space [6]. Application of adaptive control allows to broaden the capture range also in the parameter space [33]. For many years chaotic modes were either overlooked or considered as undesirable ones. The classical example can be found in the paper by Van der Pol and Van der Mark (1927) [68], who just made a brief remark "Often an irregular noise is heard in the telephone receiver before the frequency jumps" and never discussed this issue further [44].

During the last 15 years, however, the interest in chaotic oscillations in circuits theory emerged dramatically. The most popular applications of chaos lie in the field of telecommunications.

<u>Telecommunications</u>. A few ways of using chaos for signals storage and transmission were reported. The most popular ones suggest to use the chaotic carrier instead of the periodic one. To use advantages of chaotic carrier electrical circuits of the transmitter and receiver must be synchronized which gives a possible application for various schemes of synchronization of nonlinear dynamical systems. The problem is familiar to the control enginieers because similar tasks arise in the problems of observation or model reference control. Nevertheless it turned out that traditional solutions do not encompass all required synchronization schemes. For example conventional observation problem is posed as reconstruction of the whole state vector based on the on-line output measurements. In terms of synchronization problem it allows only for unidirectional connection between the systems, while the general synchronization problem statement allows for biderectional connections. Some possible applications of synchronization of oscillatory systems particularly, chaotic ones can be found in cryptography (see e.g. [11, 34]).

<u>Control of chaos in lasers</u>. One of the first experiments on control of chaotic oscillations was devoted to control of lasers, see [58] where the system consisting of neodymium-yttrium-alluminium-garnate laser and a frequency-doubling crystal was considered. It is known that at high levels of input power the intensity of the laser's output power fluctuates chaotically. Using simple feedback control algorithm of Hunt [31] the laser was driven into periodic mode. The final result was that the range of steady output was increased from about 20 percent to more than 300 percent above the lasing threshold. The other succesful experiments with gas CO_2 lasers [64, 25], solid-state lasers [26, 50, 9] and semiconductor lasers [24, 36] were also reported.

Control of oscillations in power systems. An important application field is a synchronization of electrical generators in power systems. Since electrical power systems are becoming heavily loaded and distributed, some possible sudden faults such as lightening, equipment failure, etc. may result in instability of the overall system. Therefore the problem of the broading the attraction region of normal operating mode is of utmost importance (see e.g. [56]). The whole stabilization problem can be splitted into two ones: a) transient stabilization which should be achieved immediately after the fault and b) local stabilization near the desired operating point. The goal of transient stabilization is to recover the synchronism of generators in the system. The key point is to turn the irregular oscillations into the synchronous ones by means of small control efforts. One possible solution of the problem of transient stabilization was presented in [54].

2.3. Chemistry and chemical engineering. Studies of oscillatory dynamics in chemical reactions have dramatic history. First example of oscillatory chemical reaction was the famous Belousov-Zhabotinsky reaction which was discovered in 50s but recognized much later. First mathematical model of the hypothetic chemical reaction admitting oscillatory behavior was suggested by Lotka still in 1925 but later was used in ecology rather than in chemistry. The significance of oscillatory behavior in chemistry was recognized only in 70s, see [47]. Then the simple models of oscillatory chemical kinetics have become of wide use, e.g. the brusselator model which was in fuct suggested by Turing in 1952 [67] for other purpose. Chaotic oscillations were discovered in 70s numerically by simulation [57] and later experimentally in the forced brusselator model [66], coupled brusselators [62] and Belousov-Zhabotinsky reaction [61].

2.4. Biology, biochemistry, medicine. In biology in contrast to chemistry the oscillatory behavior was undoubtfully recognized long ago, see e.g. excellent book by Glass and Mackey [28]. It is even emphasized that rhythmic behavior is a basic property of living organisms because it is a sign of regulation processes caused by natural selection. Therefore rhythmic behavior facilitates the survival and evolution of the organisms [29]. Oscillations indeed occur at all levels of biological organization, with periods ranging fromm milliseconds (neurons) to seconds (cardiac cells), minutes (oscillatory enzymes), hours (pulsatile hormone secretion), days (circadian rhythms), weeks (ovarian cycle) and years (circannual rhythms, epidemiological processes and predator-prey interactions in ecology). Chaotic rhythms are not exceptional. Besides, biological systems are pierced by positive and negative feedbacks which makes them well prepared for control purposes. For example a population of chaotic amebae was subjected to a small-amplitude periodic forcing which appeared to be sufficient to transform chaotic behavior into periodic [29].

During the last decade much attention in biochemistry and molecular biology was attracted to modelling and simulation of molecular dynamics at different levels: from local molecular bonding to proteins, DNA, chromosomes and larger units, see e.g. [45]. The models of molecular (e.g. biomolecular) dynamics are built much like to those of mechanical systems using Hamiltonian form of description. Therefore methods of control of Hamiltonian systems [17] can be applied to create oscillations in biomolecular structures, synchronization of oscillations in different parts of the structure, etc. To implement these methods electromagnetic or laser radiation can be used as a control action while measurements of secondary field induced by the control can serve as a sensor information to organize the appropriate feedback.

2.5. Economics and finance. Existence and importance of oscillations in economical activities has been widely recognized since the last century. One of the most well established economic theories is that of business cycles where the classical models of nonlinear dynamics were suggested by Kaldor [32], Metzler [43], Goodwin [30]. A large number of monographs and surveys were published focused on oscillations in particular markets or the entire economy. Among those devoted to nonlinear dynamics and chaos one may mention [40, 42, 10]. In the majority of publications various economical mechanisms and explanations of oscillatory behavior were suggested. Much less attention was paid to prediction of cycles and almost no approaches to control based on models of nonlinear dynamics were reported.

In [1] we suggest the statement and solution to the problem of control of business-cycles based on Metzler model which admits chaotic behavior. As a control goal we choose proximity of the system behavior to the periodic limit cycle modelled by the reference model. Such a goal ensures better predictability of buisiness-cycles which makes natural economical sense.

Quite recently an interest in models of nonlinear dynamics and chaos arised in financial studies previously based mainly on stochastic diffusion models. By means of rescale-range analysis it has been shown that dynamics of many financial time series is better described by chaotic models than by conventional ones of Brownian motion [52, 53]. The interest in nonlinear dynamics was also heated by achievements of artificial intelligence suggesting new analysis and prediction methods based on neural networks, genetic algorithms, fuzzy models. Computer-aided trading systems create new hopes and expectations for financial professionals. "Chaos theory" and "neural networks" have become new centers of attraction in financial industry and hot topics in the journals "Wall-Street Technology", "Artificial Intellegence in Finance", "Neurowest", etc. Summarizing, we anticipate an increasing role of economics and finance as a new market for control of oscillations and chaos.

3. Conclusions. The abovementioned publicatons (see also the bibliography [8]) demonstrate both growing interest in control of nonlinear oscillatory and chaotic systems and great variety of its potential applications. Some solutions to the posed problems were surveyed in [7, 21]. Also recently a unified approach to adaptive and nonadaptive control of periodic and chaotic oscillations in nonlinear continuous-time systems was suggested [16, 17, 20, 19] . The approach is based on the so called speed-gradient method [15](changing the control parameters proportionally to the gradient of the derivative of the given objective functional along the controlled system trajectories). The speed-gradient method is extended to achieve both conventional control objectives (regulation and tracking) and specific goals like excitation (swinging) and syncronization of oscillations (see [17, 19]).

Another approach deals with the extension of the so called OGY method [49] of controlling chaos. The rigorous analysis is based on the new concept "controlled Poincare map" and on the method of recursive goal inequalities suggested by V.Yakubovich in 1966 [70]. The obtained control algorithms solve the tracking problem for the recurrent trajectories under bounded disturbances of sufficiently small level [18].

More information on this emerging and exciting field can be found in the book [22].

References

1. Andrievsky B.R., Dymkin E.G., Fradkov A.L. (1997) "Adaptive control of nonlinear business-cycle models" (this volume).
2. Blekhman, I.I. (1971) *Synchronization of dynamic systems*, (Moscow-Nauka), in Russian.
3. Blekhman, I.I. (1988) *Synchronization in science and technology*, (ASME Press, New York).
4. Blekhman I.I., (1994) *Vibrational Mechanics.* Moscow: Fizmatlit, 1994.
5. Blekhman, I.I., A.L. Fradkov, H. Nijmeijer and A.Yu. Pogromsky, (1997), "Self-synchronization and controlled synchronization of dynamical systems", Proceedings of 4th Europ. Contr. Conf., Brussels, July 1-4, 1997.
6. Bradley E., (1993), "Using chaos to broaden the capture range of a phase-locked loop," *IEEE Trans. Circ. Sys. I*, 40, pp. 808-818.
7. Chen, G. and X. Dong, (1993) "From chaos to order – perspectives and methodologies in controlling chaotic nonlinear systems", *Int. J. Bifurcation Chaos*, vol. **3**(6), pp. 1363-1409.
8. Chen, G. "Control and synchronization of chaotic systems", (a bibliography), ECE Dept., Univ. of Houston, TX, available from
ftp: ftp.egr.uh.edu/pub/TeX/chaos.tex (anonymous login).
9. Colet, P., R. Roy and K. Wiesenfeld, (1994), "Controlling hyperchaos in a multimode laser model," *Phys. Rev. E*, 50, pp. 3453-3457.
10. Creedy, J. and V.L. Martin, (eds.), (1994), *Chaos and non-linear models in economics. Theory and applications*, (Edward Elgar).
11. Cuomo, K.M., A.V. Oppenheim and S.H. Strogatz (1993) "Synchronization of Lorenz-based chaotic circuits with applications to communications", *IEEE Trans. Circ. Sys, Part II*, vol. **40**, pp.626-633.
12. Ding, M., E. Ott and C. Grebogi, (1994) "Controlling chaos in a temporally irregular environment," *Physica D*, 74, pp. 386-394.
13. Ditto, W.L. and L. M. Pecora, (1993), "Mastering chaos," *Scientific American*, pp. 78-84, Aug.
14. Field, R.J. and L. Gyorgyi (eds.), (1992), *Chaos in chemistry and biochemistry*, (World Scientific).
15. Fradkov, A.L., (1990), *Adaptive control in complex systems*, (Moscow-Nauka), in Russian.
16. Fradkov, A.L., (1995), "Adaptive synchronization of hyper-minimum-phase systems with nonlinearities," *Proc. of 3rd IEEE Mediterranean Sympo. on New Directions in Contr. Auto.*, Limassol, July 1995, pp.272-277.
17. A.L.Fradkov, Swinging control of nonlinear oscillations. International J. Control, v.64, No 6, 1996, pp.1189-1202.
18. A.L. Fradkov, P.Yu.Guzenko (1997), "Adaptive control of oscillatory and chaotic systems based on linearization of Poincare map" Proceedings of 4th Europ. Contr. Conf., Brussels, July 1-4, 1997.

19. A.L. Fradkov, I.A.Makarov, A.S.Shiriaev, O.P.Tomchina, (1997), "Control of oscillations in Hamiltonian systems". Proceedings of 4th Europ. Contr. Conf., Brussels, July 1-4, 1997.
20. A.L. Fradkov, A.Yu. Pogromsky, Speed-gradient control of chaotic continuous-time systems. IEEE Trans.Circuits and Systems, part I, v.43, No11 pp.907-913.
21. Fradkov, A.L., Pogromsky A.Yu.(1996), "Methods of nonlinear and adaptive control of chaotic systems," *Proc. of 13th IFAC World Congress*, San-Francisco, July 1996, v.E, pp.185-190.
22. A.L. Fradkov, A.Yu. Pogromsky, *Introduction to Control of Oscillations and Chaos* Singapore: World Scientific, 1997.
23. Gandolfo, G., (1983), *Economic Dynamics: methods and models.* 2nd ed., Amsterdam, North Holland.
24. Gavrielides, A., V. Kovanis and P. M. Alsing, (1993), "Controlling chaos in semiconductor laser devices," *Proc. of the SPIE - The Int'l Society for Optical Engr.*, 2039, pp. 250-262.
25. Genesio, R., M.Basso, M.Stanghini, A. Tesi, M.Ciofini, R.Meucci, (1995), "Limit cycle bifurcation analysis and its application for controlling complex dynamics in a CO_2 laser," *Proc. Nonlin. Dynamics of Elec. Circ. Conf. (NDES'95)*, Dublin, July 1995.
26. Gills, Z., C. Iwata, R. Roy, I. Schwartz and I. Triandaf, (1992), "Tracking unstable phenomena in chaotic laser experiments: Extending the stability regime of a multimode laser," *Phys. Rev. Lett.*, 69, pp. 3169-3172.
27. Glanz, J., (1994), "Do chaos-control techniques offer hope for epilespy?" *Science*, August, p.1174.
28. Glass, L., M.C.Mackey, (1988), *From Clocks to Chaos: The Rhythms of Life*, Princeton,NJ: Princeton Univ.Press.
29. Goldbeter, A., (1993), "From periodic behavior to chaos in biochemical systems," in*Chaos in Chemistry and Biochemistry*, R. J. Field and L. Györgyi (eds.), World Sci. Pub., pp. 249-283.
30. Goodwin, R.M., (1951), "The nonlinear accelerator and the persistence of business cycles," *Econometrica*, v.19, pp.1-17.
31. Hunt, E. R. and G. Johnson, (1993), "Keeping chaos at bay," *IEEE Spectrum*, Nov., pp. 32-36.
32. Kaldor, N., (1940), "A model of the trade cycle" *Economic Journal*, 1940, vol. **50**, pp.78-92.
33. Kazakov, L.N., Yu.V.Shirokov, (1992), "The design of the broadband adaptive phase-locking circuits," In: *Intern. Seminar "Nonlinear Circuits and Systems"*, Moscow: IEEE Circuits and Systems Society - Russian Society for Radio, Electronics and Communications, vol. **2**, pp.253-262.
34. Dedieu H., Kennedy M.P., M.Hasler (1993), "Chaos shift keying: modulation and demodulation of a chaotic carrier using using self-synchronizing Chua's circuits ". IEEE Trans. Circ.Syst.,II v.40, pp.634-642.
35. Landa, P.S., (1980) *Self-Oscillations in Systems with Lumped Parameters*, (Nauka), Moscow. (in Russian)
36. Liu, Y., N. Kikuchi and J. Ohtsubo, (1995), "Controlling dynamical behavior of a semiconductor laser with external optical feedback," *Phys. Rev. E*, 51, pp. 2697-2700.
37. Li T., J.Yorke, "Period three implies chaos". Amer.Math.Monthly, 1975, v.82, pp.985-992.

38. Lindsey, W., (1972), *Synchronization systems in communications and control*, NJ, Prentice-Hall.
39. Lorenz, E.N., (1963), "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. **20**, pp. 130-141.
40. Lorenz, H.-W., (1989), *Nonlinear Dynamical Economics and Chaotic Motion*, Berlin-Heidelberg-New York: Springer-Verlag.
41. Lorenz, H.-W. "Complexity in deterministic, nonlinear business-cycle models - foundations, empirical evidence and predictability," in *Nonlinear dynamics in economics and social sciences*, Lecture notes in economics and mathematical systems, v.**399**, pp. 17–52, Springer-Verlag.
42. Medio, A., (1991), Continuous-time models of chaos in economics. *Journal of Economic Behavior and Organization*, vol. **16**, pp.115-151.
43. Metzler, L.A., (1941), "The nature and stability of inventory cycles," *Review of Economic Studies*, vol.**23**, pp.113-129.
44. Moon, F., (1992), *Chaotic and fractal dynamics: an introduction for applied scientists and engineers*. (Wiley).
45. Mosekilde, E. and L. Mosekilde (eds), (1991), *Complexity, chaos, and biological evolution*, NATO ASI Series, Series B: Physics Vol. 270, Plenum Press, New York, London.
46. Neimark, Yu. I. and P.S. Landa, (1987), *Stochastic and chaotic oscillations*, (Moscow-Nauka), (in Russian).
47. Nicolis, G. and I.R. Prigogine, (1977), *Self-organization in non-equilibrium systems*, (New York: Wiley).
48. Ogorzalek, M. J. (1993) "Taming chaos - part I: synchronization; part II: control", *IEEE Trans. Circ. Syst. Part I*, vol. **40**, pp. 639-706.
49. Ott, E., C. Grebogi, and J. Yorke, (1990), "Controlling chaos", *Phys. Rev. Lett.*, vol. **64**(11), pp. 1196-1199.
50. Parisi, J., R. Badii, E. Brun, L. Flepp, C. Reyl, R. Stoop, O. E. Rossler, A. Kittel and R. Richter, (1993), "Explicit realization of chaos control in an NMR-laser experiment," *J. of Zeitschrift fur Naturforschung, Teil A*, 48A, pp. 627-628.
51. Pecora L.M., Carroll T.M. (1990) "Synchronization in chaotic systems" Phys.Rev.Letters, v.64, pp.821-823.
52. Peters, E., (1991), *Chaos and Order on capital markets*, New York: John Wiley& Sons.
53. Peters, E., (1994), *Fractal market analysis*, New York: John Wiley& Sons.
54. Pogromsky, A.Yu., A.L. Fradkov and D.J. Hill, (1996), "Passivity based damping of power system oscillations", *35th IEEE Conf. on Decision and Control*, v.4, pp.3876-3881.
55. Pugh C.C. (1967) "The closing lemma", Amer. J. Math., v.89, pp.956-1009.
56. Qu Z., Dorsey J., Bond J., McCalley J.,(1992) "Application of robust control to sustained oscillations in power systems", IEEE Trans. Circ.Syst., I, v.39, pp.470-476.
57. Rossler, O.E., (1976), "An equation for continuous chaos," *Phys. Lett. A*, vol.**57**, pp.397-399.
58. Roy, R., T. W. Murphy, T. D. Maier, Z. Gills and E. R. Hunt, (1992), "Dynamical control of a chaotic laser: Experimental stabilization of a globally coupled system," *Phys. Rev. Lett.*, vol. **68**, pp. 1259-1262.
59. Ruelle D., F.Takens (1971), "On the nature of turbulence", Comm. Math. Physics, 1971, v.20, No2, pp.167-192.

60. Saito, F., T. Fukuda and F. Arai, (1994), "Swing and locomotion control for a two-link brachiation robot," *IEEE Control Systems*, Feb.1994, No1, pp.5-12.

61. Schmitz, R.A., K.R.Graziani and J.L.Hudson, (1977), *J. Chemical Physics*, vol. **67**, pp.3040.

62. Schreiber, I., M.Kubicek, M.Marek, (1980), "On coupled cells", In: *New Approaches to Nonlinear Problems in Dynamics*, Ed. by P.Holmes Philadelphia: SIAM, pp.496-508.

63. Simin, S.G., A. L. Fradkov and S. A. Shadchin, (1995), "Chaotic oscillations emission by NDC semiconductor devices", *Proc. of ISDRS-95*, Sharlotteswille, Wa, USA, pp. 211-224.

64. Sugawara, T., M.Tachikawa, T.Tsukamoto, T.Shimuzu, (1994), "Observation of synchronization in laser chaos" *Physical Review Letters*, vol. **72**(22), pp.3502-3505.

65. Takens, F., (1981), "Detecting strange attractors in fluid turbulence," In *Dynamical Systems and Turbulence. Ed. by D.Rand and L.-S.Young*, Berlin: Springer-Verlag.

66. Tomita, K. and T.Kai, (1979), *J. Stat. Phys.*, vol. **21**, p.65.

67. Turing, A.M., (1952), "The chemical basis of morphogenesis", *Philos. Trans. Royal Soc.*, vol. **B237**, N641, pp. 37–72.

68. Van der Pol, B. and J. Van der Mark, (1927), "Frequency demultiplication" *Nature*, vol. **120**(3019), pp.363-364.

69. Wu, C.W. and L.O. Chua (1993) "A simple way to synchronize chaotic systems with applications to secure communications," *Int. Journal of Bifurcation and Chaos*, vol. **3**, pp. 1619-1627.

70. Yakubovich, V.A., (1996), "Finitely-convergent recursive algorithms for the solution of the systems of inequalities," *Soviet Math. Dokl.*, vol.**166**.

RELIABILITY OF AVIONICS AND ITS "HISTORY OF ABUSE" A PROGNOSTIC TECHNIQUE

Victor A. Skormin and Leonard J. Popyack
Binghamton University, Binghamton, NY, 13902, USA
Rome Laboratory US Air Force, Rome, NY, 13441-5700, USA
e-mail: vskormin@binghamton.edu

Abstract. Wide application of real-time monitoring and recording devices results in vast amounts of operational and environmental data. This data can be utilized for compiling a history of adverse exposures of flight-critical avionics components and relating this history to their probability of failure. A statistical technique utilizing cluster analysis is proposed for the "off-line" analysis of the impact of various adverse conditions on avionics failures. Bayes' technique is suggested for "on-line" reassessment of failure probabilities of particular components thus facilitating the "service when needed".

Key Words. Failure, avionics, statistical analysis, probability of failure, environmental exposures

Introduction. High reliability of avionics is crucial for safe and efficient operation of the entire aircraft. While in operation, avionics components are exposed to mechanical vibration, excessive temperatures, humidity, radiation, etc. These adverse conditions, acting individually and in combination, are known to have cumulative effects leading to performance degradation and failures. Monitoring of environmental conditions and operational status of avionics provides important information which can be viewed as a "history of abuse" of individual modules. Implemented on a large scale, this monitoring practice would result in a growing data base suitable for correlating the "history of abuse" to probability of failure of particular devices. This will facilitate periodic re-evaluation of probability of failure of flight-critical avionics modules on the basis of the cumulative exposure time to various adverse conditions. As the probability of failure exceeds some marginal value, the module could be replaced *before* the occurrence of a failure thus facilitating the maintenance policy known as "service when needed". It can be seen that this policy would result in significant enhancement of the reliability of the entire aircraft.

The suggested concept is not new. Personnel of nuclear power plants are subjected to continuous monitoring of exposure to radiation. Individualized exposure data is periodically analyzed and is used for changing the individual's work assignments "when needed". Continuous increase of complexity and cost of avionics, although incomparable with cost of human life but already exceeding the cost of aircraft, provides a sufficient argument supporting the adoption of the described approach. Recently, this approach became not only justifiable but feasible due to the development and implementation of high capacity data/event monitoring devices. For example, TSMD [1] is capable of recording and accumulating significant amounts of statistical data representing the environmental/operational conditions of particular hardware units.

Availability of TSMD data facilitates formulation and solution of important on-line and off-line reliability-related problems. These problems include investigation of the role of various environmental factors in the development of particular failures, investigation of combined effects of several factors, re-evaluation of probability of failure on the basis of known exposure to particular adverse conditions, as well as development of special types of mathematical models and model-based diagnostics techniques [2].

Off-Line Analyses. Failure development processes can be viewed as Continuous Processes with Discrete Event Output (CPDEO). The word "continuous" reflects the nature of the recorded environmental variables such as intensity of vibration, humidity, radiation, excessive temperature, etc., and exposure time to these factors. The "discrete event output" refers to the occurrence or non-occurrence of various types of failures as the result of the environmental exposures and operational conditions.

Methods of failure analysis are always based on mathematical models. The traditional methods of mathematical description are focused primarily on the investigation of the physical nature of the failure development process, or on the development of statistics-based regression models. The first approach deals with particular theories, explaining physical phenomena behind failure development, therefore, it is more suitable for designers and manufacturers of avionics. Application of the second approach is limited due to the large number of factors and complex interactions between various factors, responsible for failures. Both approaches result in unnecessary complex analytical procedures, often too complex and unreliable for practical implementation.

Cluster analysis, developed in the beginning of our century, presents a technique for detection and formalization of common features specific to realizations of one of two processes subjected to the analysis. It is, therefore, ideally suitable for modeling a CPDEO with two distinct outcomes. Application of cluster analysis results in the selection of a so-called group of informative factors responsible for the process outcome. "Common features", defined in terms of informative factors, result in a mathematical model of a CPDEO, often defined in the form of a set of separation rules. Once established, such a model can be successfully used for predicting outcomes of the CPDEO based on recorded realization of its continuous variables (factors). Cluster analysis has the potential for providing valuable methodology for reliability-related problems.

Consider a particular record of environmental exposures and performance status of an avionics module, generated by a TSMD. Such a record would be comprised of a group of continuous exposure variables $x_1(t),x_2(t),...,x_M(t)$ and a binary status variable Q(t). Variables $x_j(t)$ represent cumulative exposure to particular adverse environmental factors, j=1,2,...,M, in hours; variable Q(t) defines the performance status of an electronic module on the "pass-fail" basis, and t is the record index. Since variable Q(t) has numerical value of 1 or 0, it can be seen that $\{x_1(t),x_2(t),...,x_M(t),Q(t)\}$ constitutes one realization of a CPDEO representing the entire history of "environmental abuse" of a particular module and the outcome of this "abuse", i.e. "failure" or "no failure".

It is understood that outcomes Q(t) are defined by the particular combinations of numerical values of variables $x_1(t),x_2(t),...,x_M(t)$. Theoretically, the value of variable Q can be associated with any feasible combination of numerical values of the environmental variables, i.e. can be assigned to any particular location in the factor space $\mathbf{X}=[x_1,x_2,...,x_M]$. Although the relationship Q=Q(X) may not be defined analytically, particular locations in the factor space X, marked as 1 or 0, form clusters according to this relationship. Analysis of these clusters provides an important insight to the failure development process and results in a specific form of a mathematical model, representing the unknown relationship Q=Q(X).

Analysis of clusters defined in a multidimensional space cannot be done by visual inspection because of the obvious reason. However, inspection of the projections of these clusters on particular subspaces (planes) can be very fruitful. It can facilitate the detection of the groups of dominant variables, responsible for the outcome Q, as well as variables which can be labeled as non-informative. This analysis can lead to the detection of particular environmental factors which acting independently or in combination are responsible for specific failures. Analysis of the projections of clusters allows for the definition of specific regions in the factor space **X** where the value of variable Q most likely indicates a failure. Defined mathematically, these regions are viewed as a cluster model and present a convenient way to describe a CPDEO. The presented problem is common in pattern recognition: a pattern, described by a large number of continuous variables (factors), can be recognized as a pattern of a particular class by analyzing only a small group of dominant factors. These groups, associated with particular classes, can be selected by processing available statistical data.

Implementation of TSMD-based monitoring systems for flight-critical avionics and appropriate compilation of obtained records, t=1,2,...,N, will result in a data base sufficient for a comprehensive statistical study of the impact of various environmental factors on avionics failures. Application of cluster analysis will facilitate solution of the following problems:

- ranking of particular environmental factors as having effect on failure based on the "history of abuse" of failed modules,
- detection of particular groups of environmental factors, and characterization of their combined effects on a general failure or a particular failure mode,
- development of a cluster model representing the "average exposure to particular combination of environmental factors" of avionics modules which had failed or exhibited a particular failure mode,
- definition of the "probability of a failed module to be exposed to particular combination of environmental factors over a certain time period".

Successful solution of these problems would provide valuable information to designers and manufacturers of avionics, pointing at specific failure mechanisms. At the same time, it will justify development of special protective measures and devices shielding avionics from adverse environmental conditions. Finally, some recommendations on avoiding certain combinations of adverse conditions could be formulated.

Application of cluster analysis for mathematical description, analysis and control of CPDEO is suggested in publications [3,4]. It implies that statistical data array downloaded from dedicated TSMD of flight-critical avionics components, $\{x_1(t),x_2(t),...,x_M(t),Q(t)\}$, t=1,2,3,...,N, be subjected to

1. Definition of particular subspaces (planes) $\mathbf{X}_{ij} \in \mathbf{X}$, typically comprised of two exposure variables, and selection of the projections of the original data array on these subspaces thus forming arrays $\{x_i(t),x_j(t),Q(t)\}$, $t=1,2,3,...,N$, $\forall$ $i,j=1,...,M$, $i \neq j$

2. Evaluation of a clustering criterion $F_{ij}=F\{x_i(t),x_j(t),Q(t)\}$ at each subspace representing the extent of cluster separation. One of the simplest expressions for criterion $F\{.\}$ is given as,

$$F_{ij}= (\sigma_i \sigma_j N^1 N^0)^{-1} \sum_{m=1}^{N^1} \sum_{n=1}^{N^0} \{[x_i^1(m)-x_i^0(n)]^2+[x_j^1(m)-x_j^0(n)]^2\}^{1/2} \tag{1}$$

where

N^1 and N^0 are total numbers of realizations $\{x_i(i),x_j(i),1\}$ and $\{x_i(i),x_j(i),0\}$,

$x^1_i(m)$ and $x^1_j(m)$ represent coordinates of a particular location in the plane $\mathbf{X}_{ij}$ where $Q[x^1_i(m),x^1_j(m)]=1$, similarly $Q[x^0_i(n),x^0_j(n)]=0$, and

σ_i and σ_j are standard deviations of variables x_i and x_j .

3. Selection of a "set of informative subspaces" $\Re$ comprised of subspaces $\mathbf{X}_{ij}$ with largest values of criterion $F\{.\}$

4. Definition of separation rules in terms of particular subspaces $\mathbf{X}_{ij}$,$ij \in \Re$, which could be accomplished by establishing functions (predicates) $\Phi_{ij}[x_i ,x_j]$ and $\Psi_{ij}[x_i ,x_j]$ such that

$$\begin{aligned} &\Phi_{ij}[x^1_i(k),x^1_j(k)]=1 \;\; \forall\, k=1,...,N^{11} \;\text{ and }\; \Phi_{ij}[x^0_i(k),x^0_j(k)]=1 \;\forall\, k=1,...,N^{1E} \\ &\Psi_{ij}[x^0_i(k),x^0_j(k)]=1 \;\; \forall\, k=1,...,N^{00} \;\text{ and }\; \Psi_{ij}[x^1_i(k),x^1_j(k)]=1 \;\forall\, k=1,...,N^{0E} \end{aligned} \tag{2}$$

where

$N^{11}>>N^{1E}$, $N^{00}>>N^{0E}$, $N^{11}+N^{0E}=N^1$, and $N^{00}+N^{1E}=N^0$

It could be seen that a system of predicates $\Phi_{ij}(x_i , x_j)$ and $\Psi_{ij}(x_i , x_j)$ established for informative subspaces constitutes a cluster model of a failure process. Such a model can be used for failure prediction on the basis of the known vector of adverse environmental exposures $\mathbf{X}(t^*)=[x_1(t^*),x_2(t^*),...,x_M(t^*)]$. Indeed, if $Q=1$ corresponds to failure as the outcome of the CPDEO, then condition

$$\Phi_{ij}[x_i(t^*), x_j(t^*)]=1 \;\forall\; ij \in \Re \tag{3}$$

undoubtedly indicates that the avionics component subjected to $x_1(t^*)$, $x_2(t^*)$, ..., $x_M(t^*)$ hours of exposure to adverse conditions has a high probability of failure. Analysis of the set of informative subspaces can be used for assessing the combined effect of several environmental factors on the failure development process.

However, the most useful information which can be extracted from the cluster model of the failure development process is the probability of a failed avionics module to have a certain cumulative exposure to adverse conditions,

$$\begin{aligned} &P_{ij}(X/1)=P\{[x^1_i(k),x^1_j(k)] \;/\; Q[x^1_i(k),x^1_j(k)]=1\} \\ &P_{ij}(X/0)=P\{[x^0_i(k),x^0_j(k)] \;/\; Q[x^1_i(k),x^1_j(k)]=1\} \end{aligned} \tag{4}$$

These probabilities are to be defined for particular subspaces as follows,

$$P_{ij}(X/1)=N^{11}/N^{1} \quad \text{and} \quad P_{ij}(X/0)=N^{1E}/N^{0} \quad \forall\, ij\in\Re \qquad (5)$$

It will be shown that probabilities (4), (5) can be utilized for updating failure probability of particular modules on the basis of their individual cumulative exposures to adverse conditions recorded by their dedicated TSMD.

On-Line Analysis. While the above problems are aimed at the detection of general trends "hidden" in the environmental and operational data, reliability characterization of a particular avionics module constitutes an important problem. This problem implies that general trends representing failure statistics of many modules of the same type will be utilized for computing probability failure of a particular module. Successfully solved, this problem would facilitate "service when needed" and result in the enhanced reliability of the entire aircraft.

Denote initial probability of failure of an avionics module as P_0. Assume that after taking into account the most recent batch of TSMD data, vector $\mathbf{X}^*=[x_1^*,x_2^*,...,x_M^*]$ representing the cumulative exposure of this module to adverse environmental conditions, has been defined. Then, based upon the established cluster model, the probability of failure of the module can be updated by computing the posterior failure probability P_1. As suggested in [3,4], this problem can be solved by application of Bayes' technique utilizing the developed cluster model.

Analysis of realization $\mathbf{X}^*$ may result in the occurrence of one of the following mutually exclusive events,

$$\text{event } A_L\text{:} \quad \Phi_{ij}[x_i(t^*), x_j(t^*)]=1 \;\forall\, ij\in\Re_L\,,\; L=1,2,...\,,\; \Re_L\in\Re \qquad (6)$$

Then posterior probability P_1 can be defined as follows,

$$P_1= P_0^1P(\Re_L/1) \,/\, [P_0^1P(\Re_L/1)+P_0^0P(\Re_L/0)] \qquad (7)$$

where

$P_0^1=P_0$ and $P_0^0=1-P_0$ are prior probabilities of failure and no failure,

$P(\Re_L/1)$ is the probability of having condition $\Phi_{ij}[x_i(t), x_j(t)]=1$ satisfied for those realizations of vector $\mathbf{X}$ that $Q[\mathbf{X}(t)]=1 \;\forall\, ij\in\Re_L$,

$P(\Re_L/0)$ is the probability of having condition $\Phi_{ij}[x_i(t), x_j(t)]=1$ satisfied for those realizations of vector $\mathbf{X}$ that $Q[\mathbf{X}(t)]=0 \;\forall\, ij\in\Re_L$.

These results, however, are strictly application-driven and require rigorous mathematical justification.

Numerical Example. For the purpose of illustration consider the following simulated example featuring data comprised of status of a particular avionics modules represented by a binary variable and their exposure time to such adverse environmental conditions as a high level of vibration (x_1), high humidity (x_2), high radiation (x_3), abnormally high voltage of the power source (x_4), abnormally low voltage of the power source (x_5), abnormally high temperature (x_6), abnormally low temperature (x_7). It is assumed that prior probability of failure P_0=.0001

An example of computer printout of one of the subspaces is given in Figure 1 below. The set of informative subspaces included the following combinations of variables, x_1&x_6, x_1&x_7, and x_6&x_7. The predicates were defined as follows,

$$\Phi_{16}=1 \text{ if } [(x_1-1.)^2+(x_6-5.8)^2-25.8 \leq 0]$$
$$\Phi_{17}=1 \text{ if } [x_1-21.x_7 - 7.5 \leq 0]$$
$$\Phi_{67}=1 \text{ if } [-x_6 -.25x_7 + 5.7 \leq 0]$$

Results of data analysis are summarized in the tables below

	failure (Q=1)	no failure (Q=0)
total number of cases	27	56
$\Phi_{16}=1$ & $\Phi_{17}=1$ & $\Phi_{67}=1$ ($\Re_1$)	15	1
$\Phi_{16}=1$ & $\Phi_{17}=1$ & $\Phi_{67}\neq1$ ($\Re_2$)	5	2
$\Phi_{17}=1$ & $\Phi_{67}=1$ & $\Phi_{16}\neq1$ ($\Re_3$)	2	3
$\Phi_{16}=1$ & $\Phi_{67}=1$ & $\Phi_{17}\neq1$ ($\Re_4$)	2	4
$\Phi_{16}=1$ & $\Phi_{17}\neq1$ & $\Phi_{67}\neq1$ ($\Re_5$)	1	5
$\Phi_{17}=1$ & $\Phi_{16}\neq1$ & $\Phi_{67}\neq1$ ($\Re_6$)	1	5
$\Phi_{67}=1$ & $\Phi_{17}\neq1$ & $\Phi_{16}\neq1$ ($\Re_7$)	1	6

probabilities	$\Re_1$	$\Re_2$	$\Re_3$	$\Re_4$	$\Re_5$	$\Re_6$	$\Re_7$
P_0^1	.556	.185	.074	.074	.037	.037	.037
P_0^0	.018	.036	.054	.071	.089	.089	.107

Now assume that a particular data set $\mathbf{X}(t^*)=[x_1(t^*),x_2(t^*),x_3(t^*),x_4(t^*),x_5(t^*),x_6(t^*),x_7(t^*)]$ representing cumulative exposure of a module to adverse environmental conditions is such that

$$[x_1(t^*)-1.]^2+[x_6(t^*)-5.8]^2-25.8 \leq 0$$
$$x_1(t^*)-21.x_7(t^*) - 7.5 \leq 0$$
$$-x_6(t^*) -.25x_7(t^*) + 5.7 > 0$$

It could be seen that event A_2 takes place and therefore probability of failure can be updated as follows,

$$P_1= .0001\times.185 / [.0001\times.185+.9999\times.036] = .0005$$

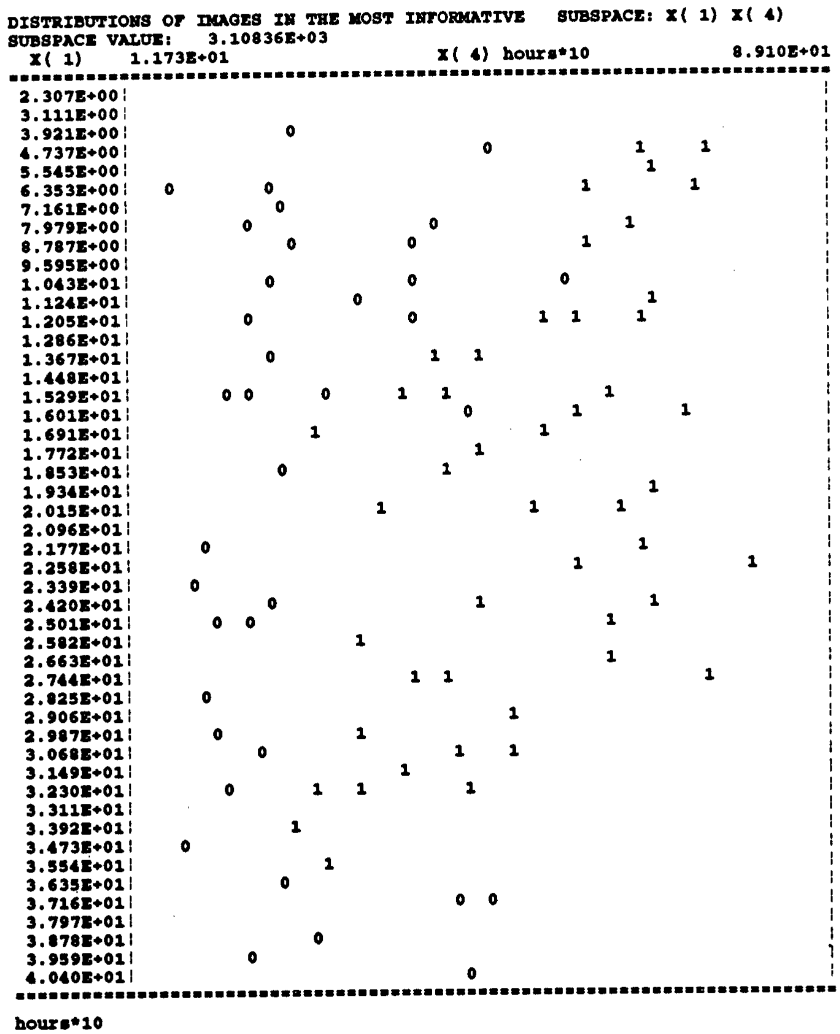


Figure 1. Example of a printout of the cluster analysis program

References

1. McCallum, M., Popyack, L., Collins, J. (1990) Environmental Measurement and Recording Techniques Utilizing a Time Stress Measurement Device (TSMD), RADC, Griffiss AFB, Proceedings - Institute of Environmental Sciences

2. Skormin, V., Popyack, L., (1995) Reliability of Electronic Modules and Their History of Abuse, Proceedings of the National Aerospace Electronics Conference NAECON'95, Dayton OH, May 95.

3. Skormin, V., Siciliano, R., (1991) A Novel Approach to Quality Analysis and Control of the Film Coating Process, Journal of Coating Technology, January 1991.

4. Skormin, V., Herman, C., Popyack, L. (1995) Application of Statistical Clustering for Process Control of Screen Printing, ASME Electronic Packaging Journal, September, 1995

THEORY AND APPLICATIONS OF INTELLIGENT CONTROL*

Stanislav Vassilyev
Irkutsk Computing Centre
of Siberian Branch of Russian Academy of Science
134, Lermontov street, Irkutsk, 664033, RU
e-mail: vas@icc.ru

Abstract. New logical language and calculus for real-time control are described. The main idea is in alternating the admissible controls and in logical derivation of consequences in order to evaluate these controls and to select more preferable ones.

As application we consider the problem of intelligent control of a group of passenger elevators, which shows that our approach is quite practical and efficient one.

Keywords. Intelligent control, automated reasoning, deduction.

Introduction. The term *intelligent control* is now becomming common in the field of computerized control. However, we are agree with [1] that this term has been overused (and often misused) in the scientific and engineering literature, and that there is a considerable lack of unanimity on the definition and structure of intelligent control systems. We favor also the following definition from [2]:

An intelligent control system has the ability to comprehend, reason, and learn about processes, disturbances and operating conditions.

We are agree with [2] that even if the word *intelligence* is interpreted in a very restrictive sense, it appears that current control systems have a long way to go before they can qualify this name.

In this paper we shall consider only the problem of increasing the capability of control systems to reason. We shall start from the question: why we can not be satisfied with the level of intelligence of today's knowledge based systems with respect to their ability to reason?

A well known and essential form of reasoning is deduction. The most powerful systems of automated deduction are automatic theorem proving techniques in mathematical domains. In contrast to this, real-world problems of control and learning have not experienced a similar success of using such well developed automatic reasoning capabilities. The existing knowledge based (KB) systems of intelligent control, especially those very popular production rule-based systems are created mostly for a restricted class of real-world problems, namely those that can be described by some set of rules which represent a priori knowledge of the designer about purposeful actions required to change the state of the controlled system in the direction of a control goal. A control system designer has to provide heuristically the applicability of some of these rules to any new state. Thus a plan for achieving the goal is not constructed or reconstructed before or during the control process and a very high responsibility and artistic skill (as opposed to scientific) of the designer are necessary in complicated control problems to guarantee the achievability of the control goal for different initial states and admissible disturbances of the operating environment. The way to free the designer from this difficult task is to increase the level of intelligence of the control system. For this purpose, it is required to extend the expressiveness of the language of rules such as to describe logically some properties of a new state of the system which is obtained after any given control action. These properties are necessary for planning the controlled process at least for some nearest future.

The knowledge based systems of action planning which have also been developed in AI, operate in comparatively simpler worlds (e.g., in programming environment). These worlds

*This research is conducted within the research grants No 96-01-00144, 95-07-20058 from RFBR.

allow knowledge representation either in very simple languages such as in intuitionistic proposition calculuses with constructive semantics, or in more expressive languages as PROLOG (in which descriptive and constructive semantics coincide); but in the latter case the deduction problem becomes semidecidable. Very important modern formalizations in the style of temporal, nonmonotonic, relevant and other logics are not sufficiently oriented on the problem of semidecidability.

Instead of automatic theorem proving we propose to derive logically only consequences of alternative controls in order to evalute them and to select more preferable ones.

We describe new method of deduction which is based on a new logical language with large-size structural elements like typical quantifiers [3]. The latter are similar to corresponding very stable linguistic constructs in natural language. A new logical calculus has developed with complete strategies of automatic theorem proving.

As application of the method we consider the problem of intelligent control of a group of passenger elevators.

A language of positively constructed formulas. We denote by Con a set of all conjucts and assume that *a conjunct* is either a finite set of atoms of the 1-st order language or $\mathbf{F}$, where $\mathbf{F}$ satisfies the property $A \subset \mathbf{F}$ for each $A \in Con$. The empty conjunct is denoted by $\mathbf{T}$. As some extension of [4] language L of *positively constructed formulas* (PCFs) is defined as follows: 1) if $A \in Con$, then $\mathbf{K}X{:}A$ is a $\mathbf{K}$-formula, where X is a set of individual variables, $\mathbf{K} \in \{\forall, \exists\}$; 2) if $B \in Con$, Φ is a set of $\mathbf{K}$-formulas, then $\mathbf{M}Y{:}B\ \Phi$ is a $\mathbf{M}$-formula, where $\mathbf{M} \in \{\forall, \exists\}, \mathbf{M} \neq \mathbf{K}$, Y is a set of variables; 3) each PCF is either a $\forall$-formula or a $\exists$-formula and vice versa. We call the nodes of tree-structure of PCF as *positive quantifiers* (PQs) and without loss of generality we assume that I) a root of PCF is $\forall{:}\mathbf{T}$ (the corresponding sets of variables and atoms are empty ones), II) each leaf is $\exists$-node, i.e. contains the sign $\exists$.

Any node $\exists X{:}A$ that follows the root of PCF immediately is called as *a base* of the PCF. A subformula which root is a base of the whole PCF is called as *a basic subformula.* We call any of immediate successors $\forall Y{:}B$ of a base $\exists X{:}A$ as *a question* to $\exists X{:}A$. It is conveniently to suppose that III) in each base $\exists X{:}\mathbf{T}$ the set X is nonempty.

A semantics of PCF $\mathcal{F}$ is defined by a common semantics of a corresponding formula in 1-st order predicate calculus $(\mathcal{F})^*$: 1) if $A \in Con, A \notin \{\mathbf{F}, \mathbf{T}\}$, then $A^{\&} = \&\{\alpha{:}\alpha \in A\}, \mathbf{F}^{\&} = False, \mathbf{T}^{\&} = True$ (propositional constants); 2) $(\exists X{:}A\ \Phi)^* = \exists x_1 \ldots \exists x_m (A^{\&} \& (\Phi)^*), (\forall X{:}A\ \Psi)^* = \forall x_1 \ldots \forall x_m (A^{\&} \rightarrow (\Psi)^*)$, where $\{x_1 \ldots x_m\} = X$, $(\Phi)^* = \&\{(\alpha)^* : \alpha \in \Phi\}, (\Psi)^* = \vee\{(\alpha)^*{:}\alpha \in \Psi\}$.

The common concepts of logical satisfiability, validy, inconsistency, equivalence, free and bound occurrences of variables and so on for PCF $\mathcal{F}$ are understood as for $(\mathcal{F})^*$. We will assume that IV) inside of each basic subformula any variable cannot be free and bound simultaneously as well as cannot be bounded by different PQs simultaneously.

Proposition 1. The language L of PCFs is complete w.r.t. expressibility in the 1-st order predicate calculus.

The question answering method of ATP. A question $\forall Y{:}B$ to a base $\exists X{:}A$ has *an answer* Θ iff Θ is a mapping (substitution) $Y \rightarrow TermA$ and $B\Theta \subseteq A$, where $TermA$ consists of all terms from A if $A \neq \mathbf{T}$ and $TermA = X$ otherwise.

If PCF $\mathcal{F}$ has the structure $\forall{:}\mathbf{T}\{\Psi, \exists X{:}A\ \ \Phi\}$, where Ψ is a list of other PCFs as subformulas (subtrees) of $\mathcal{F}$, and Φ contains a subformula $\forall Y{:}B\ \{\exists Z_i{:}C_i\ \Psi_i\}_{i=\overline{1,k}}$, then the result $\omega\mathcal{F}$ of application of *the inference rule* ω to the question $\forall Y{:}B$ with the answer $\Theta{:}Y \rightarrow TermA$ is the formula $\omega\mathcal{F} = \forall{:}\mathbf{T}\{\Psi, \{\exists X \cup Z_i{:}A \cup C_i\Theta\ \Phi \cup \Psi_i\Theta\}_{i=\overline{1,k}}\}$. After appropriate renaming some of bound variables inside of each subformula the expression $\omega\mathcal{F}$ will satisfy all the requirements for PCF's. Such renaming we will imply always during application of ω as well as the following simplifying substitutions: a) $\exists X{:}\mathbf{F}\ \ \Phi\ /\exists{:}\mathbf{F}$, i.e. $\exists X{:}\mathbf{F}\ \Phi$ will be replaced by $\exists{:}\mathbf{F}$, b) $\forall{:}\mathbf{T}\ \{\Psi, \exists\mathbf{F}\}/\forall{:}\mathbf{T}\ \Psi$ if $\Psi \neq \emptyset$.

Theorem 1. For any PCF $\mathcal{F}$ $\vdash (\mathcal{F})^* \leftrightarrow (\omega\mathcal{F})^*$.

Any finite sequence of PCFs $\mathcal{F}, \omega\mathcal{F}, \omega^2\mathcal{F}, \ldots, \omega^n\mathcal{F}$, where $\omega^s\mathcal{F} = \omega(\omega^{s-1}\mathcal{F}), \omega^1 = \omega, \omega^n\mathcal{F} = \forall{:}\mathbf{T}\ \exists : \mathbf{F}$, is called *a derivation* of $\mathcal{F}$ in the calculus $J = < \forall{:}\mathbf{T}\ \exists : \mathbf{F},\ \omega >$. Accordingly to Th. 1 the calculus J is correct: if $\vdash_J \mathcal{F}$, then $\vdash \neg(\mathcal{F})^*$.

We will assume that a search strategy tests the questions in consecutive order without ommisions (with repeating only after the whole cycle of bypass) and does not use reapplication of ω to a question with the same Θ (QA-method, i.e. question answering method of ATP).

Theorem 2. The calculus J is complete, i.e. for any PCF $\mathcal{F}$ if $\vdash \neg(\mathcal{F})^*$, then $\vdash_J \mathcal{F}$.

In comparison with [4] the conjuncts from PCF can have functional symbols (and constants) and any subformula of PCF (as subtree) is PCF due to explicit indication of sets X in the nodes of PCF.

Example 1. Let us consider one example from [5]: "Some patients love all doctors. None of patients loves any quack. Therefore, none of doctors is a quack." The formalization of this text is $A_1 \& A_2 \to B$, where $A_1 = \exists x(P(x) \& \forall y(D(y) \to L(x,y)))$, $A_2 = \forall x(P(x) \to \forall y(Q(y) \to \neg L(x,y)))$, $B = \forall x(D(x) \to \neg Q(x))$.

At first we consider the ATP-problem: how to prove the theorem $A = (A_1 \& A_2 \to B)$ of 1-st order predicate calculus by the refutation procedure of calculus J. The negation of the formula A is $A_1 \& A_2 \& \neg B$ and $A_1, A_2, \neg B$ can be represented as PCFs $A_1 = \forall{:}\mathbf{T}\ \exists x{:}\ P(x)\ \forall y{:}D(y)\ \exists L(x,y)$, $A_2 = \forall x{:}P(x), Q(y), L(x,y)\ \exists{:}\mathbf{F}$, $\neg B = \forall{:}\mathbf{T}\ \exists x{:}D(x), Q(x)$. Then for $\mathcal{F} = \neg A = \forall : \mathbf{T}\ \ \exists u : \mathbf{T}\ \ \{\forall : \mathbf{T}\ \ \exists x : P(x)\ \ \forall D : (y)\ \ \exists : L(x,y),\ \forall x_1 y_1 : P(x_1), Q(y_1), L(x_1, y_1)\ \ \exists : \mathbf{F},\ \forall : \mathbf{T}\ \ \exists x_2 : D(x_2), Q(x_2)\}$. We obtain an example of derivation of $\mathcal{F}$ in J :

$\mathcal{F}$;

$\forall{:}\mathbf{T}\ \exists ux{:}P(x)\ \{\forall y{:}D(y)\ \exists{:}L(x,y),\ \forall x_1 y_1{:}P(x_1), Q(y_1), L(x_1,y_1)\ \exists{:}\mathbf{F},$
$\quad \forall{:}\mathbf{T}\ \exists x_2{:}D(x_2), Q(x_2)\}$;

$\forall{:}\mathbf{T}\ \exists u x x_2{:}P(x), D(x_2), Q(x_2)\ \{\forall y{:}D(y) \exists{:}L(x,y),$
$\quad \forall x_1 y_1{:}P(x_1), Q(y_1), L(x_1,y_1)\ \exists{:}\mathbf{F}\}$;

$\forall{:}\mathbf{T}\ \exists u x x_2{:}P(x), D(x_2), Q(x_2), L(x,x_2)\{\forall y{:}D(y)\ \exists{:}L(x,y),$
$\quad \forall x_1 y_1{:}P(x_1), Q(y_1), L(x_1,y_1)\ \exists{:}\mathbf{F}\}$;

$\omega^4\mathcal{F} = \forall{:}\mathbf{T}\ \exists{:}\mathbf{F}$.

Corollary 1. If PCF is propositional and is not refutable, then finite number of applications of ω gives a PCF which is different from the axiom $\forall{:}\mathbf{T}\ \exists{:}\mathbf{F}$ and for which the rule ω is not more applicable.

Representation of knowledge on real − time systems (RTS) and real − time control on the basis of the calculus J. The 1-st part of the complete logical model of RTS must directly describe the operating of the RTS and be used to construct a tree of possible variants of system's operation in global and inexact terms. Each of the variants may possibly be a solution of our problem. Simultaneously, by using a knowledge in the language L we must maximally reduce the number of acceptable variants under consideration. We can describe heuristic control principles (2-nd part) intended to reduce the number of admissible variants of control generated in the 1-st part. The final selection from the reduced set of solutions (i.e. irrefuted disjunctive branches of the formula describing the state of a controlled object a few cycles in advance) may be realized with the help of special logical axioms (3-rd part) or a quality functional (in latter case some numerical methods of vector optimization are used).

Since the control information received at the output of the derivation subsystem is quantized (cycled) w.r.t. time, for simplicity of consideration we now assume the hypothesis that events observed within one cycle can be considered to be simultaneous. For example, when we have to control of a group of cabins of passenger elevator is some building as a cycle we may consider a time interval when elevator cabin passes a distance between neighboring floors (assuming that the speed of motion to be constant). Decision making has to realize within the time of one cycle.

The 1-st part is composed of the description of states and the description of operating of system.

The 2-nd part can be based on zoning the floors (i.e. determination of some groups of floors as zones). Being based on the introduced zones, one of the heuristics may require that in each zone there were no more than one free cabin. Another heuristic principle requires that situations be ignored that have statistical expectation smaller than a small threshold value. The efficiency and the priorities may be estimated in experiments.

We do not give the logical model for elevator in complete form, since it would have taken much space. Only two subformulas will now be given, which illustrate the mechanism of search and decision making. We assume that a building has 12 floors and 4 cabins of elevator. We have 5 alternative controls:

$$\forall e\alpha t' t : \mathrm{Call}(e,d,t',t), N(t,t_1)\ \exists \alpha_1\alpha_2\alpha_3\alpha_4 : \mathrm{Dist}(1,e,d,t,\alpha_1),\ldots,\mathrm{Dist}(4,e,d,t,\alpha_4)$$

$$\{\exists :\ \mathrm{Conn}(1,e,d,t),\ \ldots,\ \exists :\ \mathrm{Conn}(4,e,d,t_1),\ \exists :\ \mathrm{Call}(e,d,t'+1,t_1)\} \qquad (1)$$

(we use the notation $\forall x : A\ \ \exists Y : B\{\exists : B_1,\ldots,\exists : B_n\}$ as a shortening for $\forall x : A\ \ \{\exists Y : B, B_1,\ldots,\exists Y : B, B_n\}$).

Here, Call (e,d,t',t) means "there is a call from the floor e in the direction $d \in$ {upwards, downwards}, and at the time moment t this call has the duration t' of expectation of beginning of its service". $N(t,t_1)$ iff "t_1 is the time moment next to t". This predicate in the axiom of the next time moment existence is used. $\mathrm{Dist}(i,e,d,t,\alpha_i)$ iff "the distance between the cabin i at the time moment t and the call (e,d) is equal to α_i". The distance in our understanding is an average expected number of cycles, during which a cabin i can begin to service the call (e,d), i.e. can come to the floor e and after that move in the direction d, but not in contradiction with commands given to the cabin i by the passengers, which are in the cabin. We use incorporated procedure of computing the distance. This distance as a rule is larger than the distance (in floor numbers) of the cabin from the floor e. $\mathrm{Conn}(i,e,d,t)$ iff "the cabin i is assigned, i.e. at the time moment t it is directed to service the call (e,d)". This means that the floor e is included in the route of the cabin i as a stop.

The formula (1) "says" that a nonassigned call can either be bound up in time moment t with one of cabins or remain nonassigned until the next time moment $t_1 = t+1$. The task is to restrict the number of variants of possible reactions of the system to incoming calls. The following formula represents one of instruments of such a restriction:

$$\forall ied\alpha_i j\alpha_j t : \mathrm{Dist}(i,e,d,\alpha_i),\ \mathrm{Dist}(j,e,d,\alpha_j),\ \ \alpha_i \geq \alpha_j + m,\ldots,\ \ \mathrm{Conn}(i,e,d,t)\ \exists : \mathbf{F},$$

where m is some definite number, for example, 8 (for 12-floor building with 4 cabins of elevator). If some α_i is larger than or equal to $\alpha_j + m$, then the variant of assigning a call to cabin i is unacceptable, and along the corresponding branch of PCF we come to the contradiction. We have to remark that introducing next time moments has to be restricted by special heuristic limitations of seach strategy in QA-method.

Thus, here we use the logical derivation of corollaries of single, origin PCF formula instead of complete proving the given formula. These corollaries describe the preferable variants of control of cabins. It has been stated that this intelligent control has many advantages in comparison with the traditional one based on finite-state automata (sequential machines) and vector optimization technique. Particularly, we have to say about the high flexibility of the logical model for accounting many peculiarities of concrete application domain (the type of building, the season of year, the day of week, the time of day and so on).

Conclusion. A new logical language and calculus (with using some analogs of type quantifiers) to real-time and other problems have been developed. The main point of real-time application consists of the outstripping modelling trajectories for different admissible controls (on the time interval between two corrections of control) with eliminating nonrational trajectories. Our approach means the continuous synthesis of theorems on trajectory properties in some 1-st order calculus for estimating and selecting the most preferable controls. In comparison with temporal logics of propositional type (e.g.[6]) we use only axiom of the next time moment existence and some relevant derivation strategy. We do not eliminate the formula structure as in [7], and therefore our derivation technique is well compatible with heuristics.

References

1. Sinha, N. K., M. M. Gupta (1996) Toward intelligent machines. In M. Gupta, N. Sinha (ed.), Intelligent Control Systems. Theory and Applications, IEEE Press, New York, pp. 804-807.
2. Åström K. J., T. J. McAvoy (1992) Intelligent control: an overview and evaluation. In D. White, D. Sofge (ed.), Handbook on Intelligent Control, Van Nostrand Reinhold, New York, pp. 3-34.
3. Vassilyev, S. N. (1990) Machine synthesis of mathematical theorems. J. Logic Programming, 9, *N*2&3, pp. 235-266.
4. Vassilyev, S. N., A. K. Zherlov (1991) Logical modelling and control in real time. Proc. USSR All-Union Conf. "Intelligent Systems in Technical Engineering", III, Samara, pp. 33-38.
5. Chang, C.-L., R. C.-T. Lee (1973) Symbolic logic and mechanical theorem proving, Academic Press, New York, San Francisco, London.
6. Gabbay, D., R. Owens (1991) Temporal logics for real-time systems. Proc. of IMACS-IFAC Symposium on Modelling and Control of Technological Systems, 2.
7. Robinson, J.A. (1965) A machine-oriented logic based on the resolution principle. J. of ACM, 12, N 1.

CHAPTER I:

CONTROL OF MOVING OBJECTS

THIS CHAPTER INCLUDES PAPERS

PRESENTED AT THE CONFERENCE SESSION:

CONTROL OF MOVING OBJECTS

Organized by: ***Acad. Anatoly E. Sazonov.,***

Prof. Vitoljd V. Vojtetskij,

Dr. Gennady S. Osipov,

Dr. Igor P. Simakov.

POWERTRAIN CONTROL SYSTEM DEVELOPMENT FOR A GLOBAL MARKET

Frode Maaseidvaag and Paul Moraal
Ford Forschungszentrum Aachen GmbH
Technologiezentrum am Europaplatz, Dennewartstr. 25
D-52068 Aachen, Germany
e-mail: pmoraal@ford.com

Jing Sun and Jeffrey Cook
Ford Research Laboratory
PO Box 2053 Mail Drop 2036 SRL
Dearborn, MI 48121-2053 USA
email: jcook2@ford.com

Abstract. Powertrain control system development for a global market necessitates a systems approach to minimize time-to-market and capture diverse requirements. Off -line tuning and on-line adaptation are required to minimize calibration effort and provide robustness to parameter variations.

Keywords. Systems engineering, sensitivity points, adaptive, powertrain control.

Introduction. The challenge of being competitive in a global marketplace is to provide innovative products which meet the needs of customers worldwide, and to bring those products to market rapidly and efficiently. In the automotive industry, customer requirements for high value, reliable personal mobility must be met without neglecting societal needs reflected by governmental regulations on emissions, fuel economy and safety. These issues are of particular importance to the automotive powertrain control system designer to whom these customer and societal needs translate into system constraints and performance requirements that must be attained with minimal cost. Specifically, a powertrain control system must be developed rapidly by establishing functional requirements, and using them to define and validate plant models, control laws, and, ultimately, the embedded software. This "systems approach" to control development reduces *in-situ* calibration and improves time-to-market. A key requirement is that the system must be robust with respect to variations in parameters associated with different markets, employing automatic tuning off-line, and adaptation on-line. This paper is organized as follows: first, the diverse requirements of a global marketplace are described. Secondly, two control designs are discussed: one which uses sensitivity techniques to provide an automatic method of tuning the gains of an idle speed controller, and another which uses adaptive methods to adjust transient fuel control parameters on line to account for changes in fuel quality. Finally, a systems engineering process is described which is requirements driven and incorporates the elements of modeling, simulation, rapid prototype implementation and target microprocessor code generation.

Global Requirements and Constraints. The requirements of automotive customers throughout the world are different. Of particular importance to the powertrain control designer are the often conflicting requirements of fuel economy and performance. In Europe, diesel engines are desired to avoid the high taxes imposed on gasoline, and small engines are required for fuel economy and avoidance of taxes leveled on displacement. Nonetheless, the capability for high speed operation on the autobahn or motorway is absolutely essential. The fuel quality, the environment in which a vehicle operates, and special attributes associated with specific markets (particularly fuel economy and emissions regulations) must also be taken into account in the design of the powertrain controller. To satisfy time-to-market imperatives and contain cost, control system software must be reused, not redesigned for these different powertrains. Consequently, the system must be immune to parametric variations caused by age or operating environment, and easily calibrated to meet the needs of particular markets. Some of these requirements are encapsulated in Table 1. Additionally, stringent emissions regulations in Europe and the United States pose significant constraints on the powertrain control system.

Table 1 Typical Powertrain Operation Requirements

Country	Gasoline				Vehicle Usage		Environment	
	Price US $/litre	Quality	Availability	Octane	Distance traveled per Year (Km)	Typical Highway Speed (Km/hr)	Avg. Low Temp. (Winter / Summer deg. C)	Avg. High Temp. (Winter / Summer deg. C)
Germany	0.96	good	good	95-98	13000	130	-15/10	5/35
UK	0.95	good	good	87-98	14000	120	0/15	10/22
India	0.61	poor	limited	87-93	7500	75	10/38	15/45
Russia	0.33	poor	regionally variable	76-98	15000	90	-30/5	30/45

Off-line Tuning of Idle Speed Control. The quality of idle speed control (ISC) affects almost every aspect of vehicle performance including fuel economy, emissions, and driveability. The primary objective for ISC is to maintain the engine speed at a desired setpoint in the presence of load disturbances including air conditioning, power steering, and other accessory loads. The control variables for ISC are air flow (regulated by the throttle, or an air-bypass valve) and ignition timing. Typically, only engine speed is used for ISC feedback. A schematic control diagram is shown in Fig. 1. The ISC problem lends itself to the application of various control design methods ranging from classical (PID) to modern (LQG, H_∞) and nonconventional (neural networks, fuzzy logic) [1]. A typical ISC feedback strategy incorporates a PID loop for the air bypass valve and simple proportional feedback on ignition timing. This control configuration is dictated by the requirements that spark should return to its nominal value at steady state independent of load disturbances, and zero steady state error must be achieved. A method of rapidly calibrating an ISC system for different powertrains is described here.

The difficulty in searching for optimal PID gains is due primarily to two factors. First, the correlation between the controller parameters and the dynamic behavior of the controlled system is complicated and, in most cases cannot be explained by known steady state relationships. Secondly, the three gains affect the system in an interactive fashion and the contribution from each function cannot be easily isolated. A systematic PID tuning procedure based on dynamic control theory was developed in [2] using a validated engine model structure described in [3]. A block diagram of the engine mathematical model is shown in Fig. 2. The model includes the air-bypass valve (ABV) characteristics, intake manifold filling dynamics, engine pumping characteristics, intake-to-power stroke delay, torque characteristics and engine rotational dynamics. The differential equations describing the system are given by:

$$MAF = f_a(u)$$
$$\dot{P} = K_m(MAF - \dot{m})$$
$$\dot{m} = cyl(N, P)$$

$J_e\dot{N} = T_q - T_L$

$T_q(t) = f_T(\dot{m}(t-\sigma), N(t), r(t-\sigma), \delta(t))$

where

u = duty cycle for the air by-pass valve

r = air-fuel ratio, A/F

δ = ignition timing

T_L = load torque

J_e = engine rotational inertia

K_m = constant

σ = 180 degree crank angle delay

N = engine speed

The method of parameter tuning that has been successfully applied to the ISC system is based on the sensitivity functions of the engine speed with respect to the controller parameters. Let K be a (possibly) vector valued, generic controller parameter, ΔN be the engine speed error, and $J(\Delta N)$ be a performance cost function we wish to minimize by adjusting K. Then

$$J(K+\Delta K) \approx J(K) + 2\Delta N \frac{\partial N}{\partial K}\Delta K + (\Delta K)^T \left(\frac{\partial N}{\partial K}\right)^T \frac{\partial N}{\partial K}\Delta K$$

The ΔK which minimizes J is given by

$$\Delta K = -\left[\left(\frac{\partial N}{\partial K}\right)^T \frac{\partial N}{\partial K}\right]^{-1} \left(\frac{\partial N}{\partial K}\right)^T \Delta N$$

By measuring the sensitivity function $\frac{\partial N}{\partial K}$, a simple gradient method can be used to minimize the cost function and adjust the controller gains for the air and spark loops simultaneously. The sensitivity functions with respect to the controller parameters are easily measurable and shown in Fig. 3. It should be pointed out that this off-line tuning method can be used to develop an on-line adaptive PID control scheme (referred to as the MIT rule [4]). The sensitivity function method can also be used to investigate the robustness of the ISC system with respect to key plant parameters by evaluating $\frac{\partial N}{\partial K}$ where K is the vector of plant parameters. Other, more empirical, methods of PID tuning are also available including the well known Ziegler-Nichols technique [5].

Adaptive Compensation of Fuel Dynamics. Powertrain controllers employ feedforward compensation to eliminate the effects of fuel puddling dynamics on air-fuel ratio. Uncompensated (or improperly compensated) dynamics cause significant A/F excursions during transient operation, resulting in increased emissions and poor drive quality. In [6], an adaptive technique is described that assures proper compensation regardless of vehicle-to-vehicle variability, changing dynamics due to aging, or different fuel properties.

The system describing the air and fuel dynamics is nonlinear with a variable time delay. The fuel dynamics depend on operating condition (including ambient temperature) and may change over time. Importantly, the only feedback information available is from an exhaust gas oxygen (EGO) sensor located in the exhaust system upstream of the catalytic converter. This sensor is essentially a switch, indicating only that the A/F is either rich or lean of stoichiometry, but not by how much. The main characteristics of the fuel dynamics are illustrated in Fig. 4 and described by the following transfer function:

$$T_{fuel}(s;\tau,X) = \frac{X}{\tau s+1} + (1-X)$$

where X represents the fraction of fuel contributing to the formation of a liquid fuel film on the intake manifold wall and τ is an evaporative time constant. A typical approach to feedforward compensation is to identify the parameters X and τ over the operating range of the engine and use them to invert the fuel dynamics. To provide robustness, it is desirable to adapt these parameters on-line. This requires a means of monitoring the exhaust A/F during transient operation. As previously described, the binary valued EGO sensor is not immediately amenable to this task. Extracting A/F magnitude information from the switching EGO sensor may be accomplished by imposing an A/F modulation and deducing the unmodulated A/F based on the magnitude of the imposed signal and the resultant switching frequency of the EGO sensor. This is illustrated in Fig. 5 which shows (1) the modulated and unmodulated A/F, (2) the EGO sensor signal, and (3) the reconstructed A/F. The adaptation problem remains to determine how to update both transient fuel parameters from information contained in a single signal. An examination of the transfer function describing the fuel dynamics reveals a time-scale difference between parameters that can be used to advantage. That is, the feedthrough term, X, will affect the initial part of the transient, whereas the influence of the time constant, τ, will be observed in the latter part of an A/F transient. The adaptation algorithm then, consists of the following steps:

1. If there is no air flow transient, do not adapt parameters
2. If there is no A/F excursion, do not adapt parameters
3. If a large air flow transient causes a small A/F excursion, modify the parameters by a small amount
4. If a small air flow transient causes a large A/F excursion, modify the parameters by a large amount

It is clear that this algorithm is most easily implemented as a rule-based, "fuzzy logic" control. An example of the input-output map for the adaptation algorithm is illustrated in Fig. 6. The correction factors X and τ provided by the algorithm are stored in a table indexed by engine operating temperature and air flow. Fig. 7 shows experimental results for a vehicle driven over a specified cycle (part of the US Federal test procedure) with changes in fuel composition. First, the compensator was calibrated to provide good A/F behavior in a vehicle with high quality "certification" fuel. Testing was done using a special "hesitation" fuel (so called due to its effect on driving quality) with very low volatility. It can be seen that the algorithm has adapted to the new fuel parameters, reducing the A/F excursions dramatically.

A Systems Engineering Process for Powertrain Control Development. Often, the powertrain control development process proceeds as illustrated in the top portion of Fig. 8. A strategy designer is given, essentially, a fully specified system. That is, the base engine configuration has been determined, and most often, the subsystems have been specified: the set of sensors and actuators, their locations, and performance characteristics such as bandwidth and accuracy. The engineer, generally working from experience and judgment, prepares a software specification which is eventually hand coded, perhaps in assembly language, for the embedded application. The first real test of the control design is often in the vehicle during calibration. It is evident from this process model that not enough effort is put into requirements specification and systems analysis early in the process, resulting in entirely too much effort being expended in calibration and error correction late in the process. Furthermore, the software generation procedure is labor intensive and error prone. Finally, because the hardware development and control analysis are not concurrent, it is very difficult for the control designer to influence the subsystem requirements, even though the selection of appropriate sensors and actuators is key to both the performance and cost of the control system. It is essential to redefine this development process on the basis of sound systems engineering principles to remain competitive in a global market.

The new process vision for powertrain control is illustrated in the bottom half of Fig. 8. The process is driven by requirements, beginning at the system level (which are different for different markets), and ends with high confidence that the first vehicle implementation of the control strategy will be of the highest quality, with minimal calibration required to meet the system goals. The process vision has an analytical foundation incorporating the elements of modeling and functional analysis with verification at each step: first by validation of the modeled system, then by simulation of the system and controller. Rapid prototype implementation of the controller prior to software specification and validation of the embedded software prior to vehicle integration completes the process. Implicit in Fig. 8 are the concepts of concurrence and cross-functionality with powertrain hardware designers. Communication and refinement in both the "hardware" and "software" arenas is critical to the development of a systems solution that satisfies public, regulatory and business imperatives. Key components and important challenges of the process are described below:

Requirements Capture. The systems engineering approach to powertrain development is driven by requirements. Fundamental systems engineering research in this area, based on optimization and sensitivity techniques, is aimed at providing a mathematically rigorous approach of cascading system level objectives (fuel economy, tailpipe emissions) to subsystem requirements (feedgas emissions, component efficiency, accuracy), and providing quantitative robustness metrics for high confidence in the system performance prior to prototype hardware [7].

Control Law Development. Control models are nominally first-principle models with identified dynamics and empirically derived characteristics obtained by testing prototype hardware. As mentioned previously, by the time hardware is available, it is often too late for the control engineer to influence subsystem specifications such as sensor and actuator locations and bandwidth. The challenge is to develop control oriented models from solid modeling, kinematic, and thermodynamic design data to facilitate concurrent engineering.

Only about 20% of all the software in an embedded powertrain control system is devoted to control algorithms: most embedded "control software" is mode switching and other logic. Consequently, the development of techniques to analyze such hybrid systems is another important theoretical area [8]. Besides the hybrid nature of the control implementation, advanced technology engines are generally distinguished by the fact that they are fundamentally multivariable, high order, nonlinear systems. Research into advanced control techniques is required to attain the performance available from control intensive advanced technology hardware [9].

Rapid prototyping refers to the process of implementing a control system in a high performance computing system, usually by automatic means, directly from the graphical representation of the system in appropriate computer aided control system design (CACSD) and computer aided software engineering (CASE) tools. Rapid prototyping provides design verification prior to target microprocessor implementation, but it must flow seamlessly from the CASE/CACSD design tools.

Software Development. Software development is time consuming, often error prone, and on the critical path to product development. Software, *per se*, is not a particularly valuable commodity: it is the algorithms which underlie the software realization that distinguish the product and provide a competitive advantage. Substantial gains in efficiency can be anticipated with the capability of automatically generating embedded code from the verified software specification contained in the CASE/CACSD representation of the system. Current research is dedicated to that goal.

References

1. Hrovat, D. and Powers, W.F. (1990), Modeling and Control of Automotive Powertrains, in Control and Dynamic Systems Vol. 37, Academic Press, NY 33-64.
2. Cook, J.A., Grizzle, J.W., and Sun, J. (1995), Engine Control, in W.S. Levine (ed.) The Control Handbook, CRC Press 1270-1274.
3. Powell, B.K., and Cook, J.A. (1987), Nonlinear Low Frequency Phenomonological Engine Modeling and Analysis, in Proc. 1987 Am. Control Conf., Minneapolis, MN 332-340.
4. Ioannou, P.A., and Sun, J. (1996) Robust Adaptive Control, Prentice-Hall, Inc., NJ
5. Astrom, C., and Wittenmark, B. (1990), Computer Controlled Systems, Theory and Design, Prentice-Hall, Inc., NJ
6. Moraal, P.E. (1995), Adaptive Compensation of Fuel Dynamics in an SI Engine using a Switching EGO Sensor, in Proc. 1995 IEEE Conf. on Decision and Control, New Orleans, LA
7. Sun, J. and Sivashankar, N., (1997), An Application of Optimization Methods to the Automotive Emissions Control Problem, in Proc. 1997 American Control Conference, Albuquerque, NM
8. Butts, K.R., Kolmanovsky, I.V., Sivashankar, N., and Sun, J. (1996), Hybrid Systems in Automotive Control Applications, in A.S. Morse (ed.) Lecture Notes in Control and Information Sciences 222, Control Using Logic-Based Switching, Springer-Verlag, London, 173-189.
9. Stefanopoulou, A.G. et.al. (1995), Consequences of Modular Controller Development for Automotive Powertrains: A Case Study, in Proc. 1995 IEEE Conf. on Decision and Control, New Orleans, LA

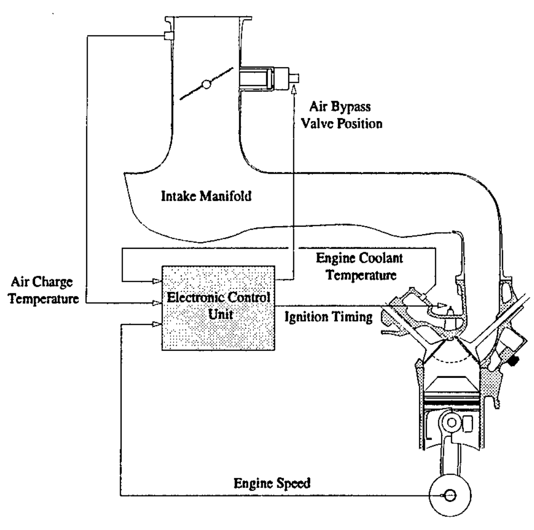


Fig. 1. Idle speed control system

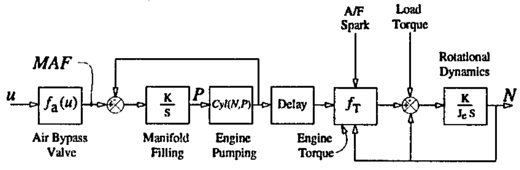


Fig. 2. Engine block diagram for idle speed control

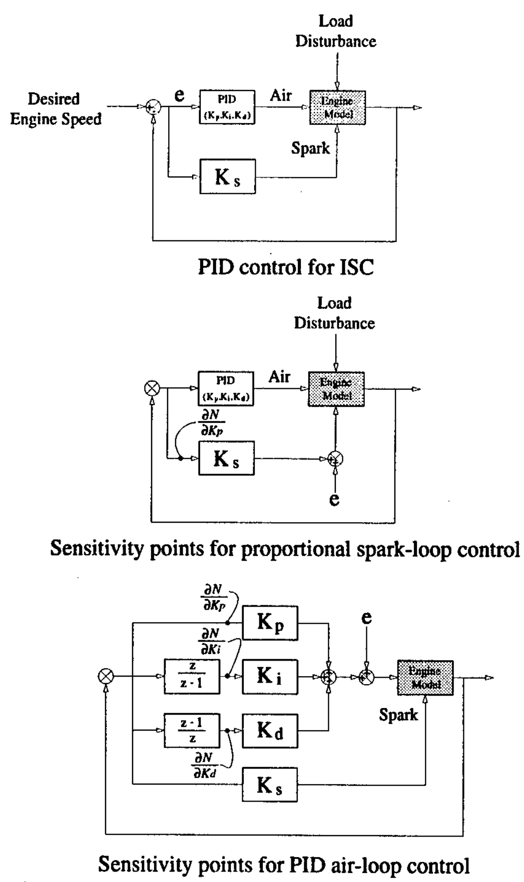


Fig. 3. Idle speed control sensitivity functions

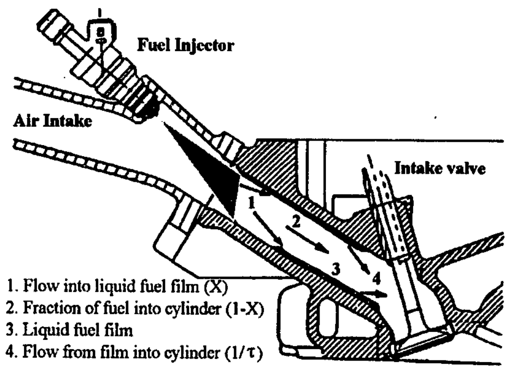


Fig. 4. Wall-wetting phenomenon

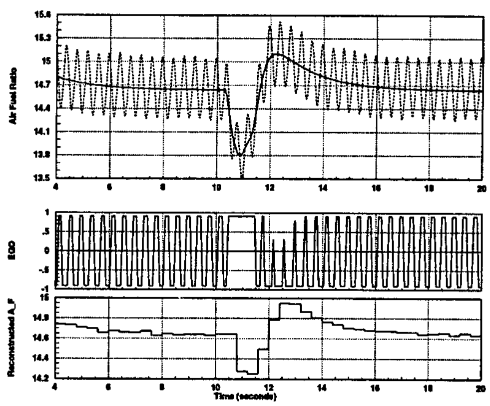


Fig. 5. Example of reconstructed A/F trace from EGO signal and modulation signal

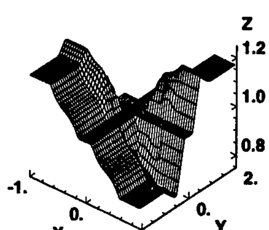


Fig. 6. Input-output map for the adaptation algorithm. X-axis: A/F transient; Y-axis: air charge transient; Z-axis: correction factor.

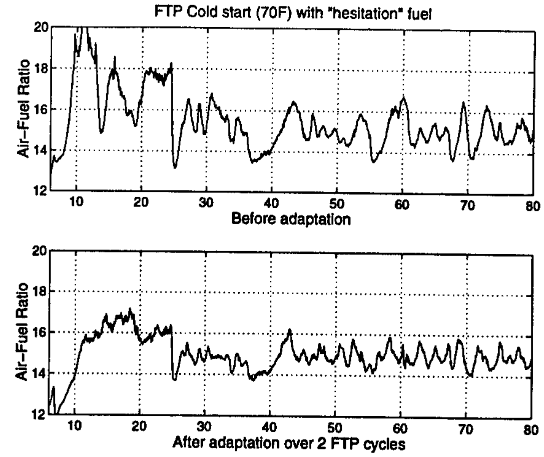


Fig. 7. Compensator optimized for "regular" fuel, tests are run with low volatility fuel. Plots show before (Top) and after (Bottom) the adaptive algorithm was active.

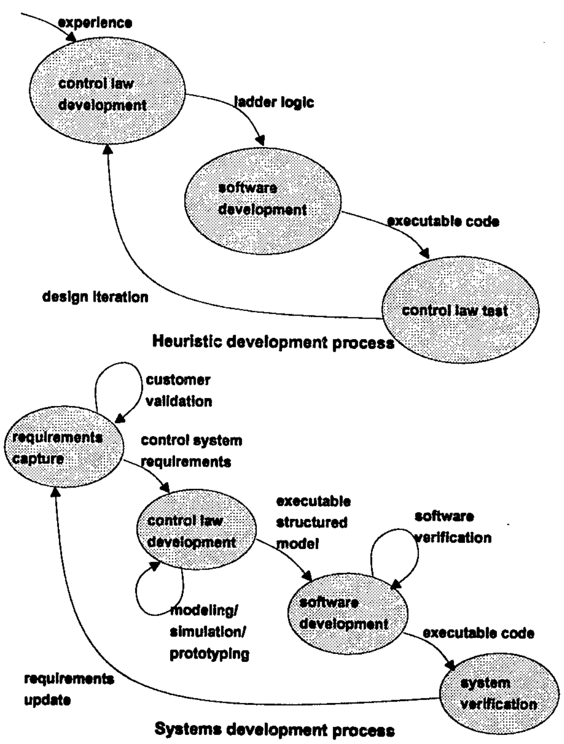


Fig. 8. Current process model (top) and proposed process model (bottom) for powertrain control development

ADAPTIVE STATISTICAL MODELS IN FUEL ECONOMY SYSTEMS FOR SEA VESSELS

Anatoly Sazonov, Sergey Smolentsev
Institute Transport's Problems
of Russian Academy of Science
V.O. 12 Line #13 199178
St.Petersburg, Russia
e-mail: wasily@transas.com

Abstract. This article touches upon the problem of statistical models of ship speed and route fuel consumption and their dependence on ship power, draught and trim as well as statistical models adaptation algorithms under real sailing conditions. The algorithms proposed can find application in fuel economy systems.

Keywords. Statistical models, adaptation, fuel economy systems.

Introduction. The main idea of statistical models usage in fuel economy systems is as follows [1]. The ship speed is determined by the movement power and medium resistance the latter being dependent, to a large extent, on the ship's draught and trim. Thus in the case of a fixed power resulting in a fixed fuel consumption, the change of the ship 's draught and trim must bring ahead the change of the ship speed (relative to water) as well as the route fuel consumption. The route fuel consumption here can be defined as the fuel consumption per 1 mile of the route covered, relative to water.

The experiments of 1990-91 made on ships belonging to Baltic Shipping Company confirmed the existence of some functional dependence between the speed, fuel consumption, power, draught and the trim of the ship. However, the general universal dependence (even for a ship taken separately) hasn't been deduced yet. It has to do with a number of so far ignored unmeasured factors functioning simultaneously on ships so under different conditions the abovementioned dependence would be different [2].

Let's introduce the following indices:

- V - ship speed relative to water
- R - route fuel consumption
- P - power
- T - draught
- D - trim.

Let's designate the functional dependence of V - speed (the route fuel consumption R) on P , T , D parameters as the speed model (fuel consumption model). Then according to the experimental data, the speed-route fuel consumption - dependence is of linear character, which makes it possible to investigate only the ship speed model. The fuel consumption model can be easily obtained with that model.

Ship Speed Statistical Model : Modeling and Optimization. The ship speed model can be represented as

$$V = f(P, T, D) \qquad (1)$$

The Statistical Model is build as a polynomial with the coefficients determined by means of natural measurements. The structure of models (parameters as well as their order in a model included) was determined with the help of experimental results obtained on ships of Baltic Shipping Company. The model structure is shown as :

$$V = a_0 + a_1 \cdot P^{1/3} + a_2 \cdot T + a_3 \cdot D + a_4 \cdot D^2 + a_5 \cdot D^3 \qquad (2)$$

The route fuel consumption dependence model can be built by analogy with the same parameters. However, consumption value can be also obtained through the speed values with the help of the route fuel consumption model correlating to the ship speed (linear equation):

$$R = b_0 + b_1 \cdot V, \qquad b_1 < 0; \qquad (3)$$

Model (2) is a regressive model of the 3d D-order, which allows to base the optimization problem on this model. Optimal value of D_{opt} trim can be found:

$$\begin{aligned} &a_3 + 2 \cdot a_4 \cdot D + 3 \cdot a_5 \cdot D^2 = 0; \\ &2 \cdot a_4 + 6 \cdot a_3 \cdot D < 0; \end{aligned} \qquad (4)$$

The ship's movement with D_{opt} trim must provide the minimum movement resistance, the other conditions being equal, resulting in its turn in the maximum speed (relative to water), whereas the engine revolutions power being fixed and steady.

Ship Speed Statistical Model : Adaptation. As mentioned above Statistical Models usage makes it possible to solve the problem of ship fuel consumption when in motion. However, the Statistical Model build on the basis of a priori data includes only few measurable parameters out of the whole number of factors affecting ship movement resistance and consequently the ship speed as well as fuel consumption. Permanent change of uncontrolled parameters leads to inadequacy of the a priori model build for definite sailing conditions in case of different conditions. Taking into account (to say nothing of measuring) all the factors is impossible, therefore it is necessary to solve the problem of model adaptation under real sailing conditions.

Adaptation algorithm can be built on the basis of information about the sequence of differences between forecast speed value according to $\overline{V}_t = f(P_t, T_t, D_t)$ and actual speed measurements for the moment $t - V_t$: $v_t = V_t - \overline{V}_t$. The absence of a systematic shift alongside with non-correlation of sequence $\{v_t\}$ testifies the correspondence of the model with real conditions while the dispersion value allows to make a conclusion as to the transducer errors and model noise.

The peculiarity of the adaptation problem is that there is some global adaptation of the model within the range limits of its existence rather than local one which necessity results from the purpose to use the model not for controlling (where the correspondence within the short range of momentary parameters values is enough) but for optimization, i.e. for solving the problem within the whole range of the model existence. In this context the model existence range is defined as the range of possible model parameters variation.

Thus to use usual methods of adaptation aimed at solving problems of local character seems to be troublesome for the case. For instance, building of the regressive polynomial by the method of least squares with short range parameter values changing measurements leads to serious errors of the model, out of the limits of this range. The method of the regular net as a global adaptation method can be proposed which uses the information about the sequence of differences $\{v_t\}$ for the local region of momentary parameters, while out of this region it deals with the primary measurements [3]. Geometrically it can be represented as shifting hipersurfices according to $V = f(P, T, D)$ in the coordinate system $< V, P, T, D >$ to measurements of momentary parameters values in the local region. As far as it concerns the model structure the surface remains stable within the whole model existence range.

The method is as follows.
Let's introduce:

f_t - model at t-moment

r_t - parameters vector at t-moment : $r_t = [V_t, P_t, T_t, D_t]^T$

$\Re_M$ - set of vectors r_t: $|\Re_M| = \mathbf{M}$

$\mathbf{L}_N[*]$ - model building operator out of N measurements set (for instance the procedure of least square method)

$\mathbf{L}_M^{-1}[*]$ - "model decomposition" inverse operator - model substitution by equivalent N-vectors set (i.e. f - surface substitution by set of vectors on the net with M knots):

$$\mathbf{L}_M \circ \mathbf{L}_M^{-1}[f] = f; \qquad \mathbf{L}_M^{-1} \circ \mathbf{L}_M[\Re] \neq \Re;$$

Then the equation for the model at the moment $t+1$ with vector of measurements $r_{t+1} = [V_{t+1}, P_{t+1}, T_{t+1}, D_{t+1}]^T$ can be represented as:

$$f_{t+1} = \mathbf{L}_{M+1}[\tilde{\Re}_M \cup r_{t+1}];$$

where $\tilde{\Re}_M = \mathbf{L}_M^{-1}[f_t]$ - equivalent representation of the model at t-moment;

M - the primary model weight;

and finally:

$$f_{t+1} = \mathbf{L}_{M+1}[\mathbf{L}_M^{-1}[f_t] \cup r_{t+1}]; \qquad f_0 = \mathbf{L}_N[\Re_N]. \qquad (5)$$

The adaptation algorithm (5) stipulates the existence of some a priori model built through a N dimensional measurements set. The M - weight variation makes it possible to change the velocity of adaptation.

Clear generalization of algorithm (5) is possible though not for the case of single vectors r_t but for some set $\mathfrak{R}_{P,t}$ (then it is necessary to use the operator $\mathbf{L}_{M+P}[*]$), using non-regular net for building equivalent set or carrying out adaptation in case of observing some conditions affecting the consequence characteristics $\{v_t\}$ etc.

The usage of the proposed adaptation algorithm, as it was mentioned above, allows to locally correct the apriori model $V = f(P, T, D)$ within the primary data scheme (as far as the model structure goes) out of the momentary parameters values region.

Conclusions. The algorithms of building and adaptation of statistical models of ship speed and route fuel consumption can find application in the creation of ship fuel economy systems. the model proposed allows to find the optimal trim value minimizing route fuel consumption as well as possible fuel overconsumption forecasting for the non-optimal trim ship sailing case. The algorithm suggested provides, for the statistical model, adaptation to real sailing conditions. However, adaptation algorithm is universal enough and can be used for adjusting a large group of statistical models.

References

1. Sazonov, A.E., N.S. Solomenko, O.V. Beliy (1989) Use of Statistic Modelling for Optimum Ship Control and Diagnostics. IV Int.Conf. of Hydrodynamics. Bulgaria. pp.62-63.
2. Smolentsev, S.V. (1992) The Adaptation of the Statistical Models. Institute Transport's Problems: Technical report. St.Petersburg. pp.5-40.
3. Sazonov, A.E., S.V. Smolentsev (1993) The Applications of the Statistical Model in Ship Automatic Control. - Transport: scientific, technical and control, 3, pp.18-22.

THEORY AND MODERN PRINCIPLES OF ARRANGING HIGHLY AUTOMATED CONTROL SYSTEMS IN RELATION TO SHIP'S NAVIGATION AND TECHNICAL FACILITIES

Igor P. Simakov, Jury S. Nemokaev,
Scientific&Production Association "Aurora"
of the Ministry of Defence Industry of the Russian Federation,
15, Karbysheva Street, St. Petersburg, 194021, Russia.
Telephone: +7(812)535-97-23 (home) E-mail: ppnovod@dux.ru

Vasily V. Antipov
The First Central Scientific and Research Institute of the Defence Ministry
(Navy Shipbuilding), St. Petersburg, Russia

Dmitry A. Skorokhodov
Transportation Problems Institute of Russian Academy of Sciences,
13, 12 line of Vasilyevsky Ostrov, St. Petersburg, Russia
Telephone: +7(812)213-09-17

Alexandr N. Novodjilov
Naval Academy named after Kuznetsov N.G., St. Petersburg
73/1, Vyborgskaya Embankment, St. Petersburg, Russia
E-mail: ppnovod@dux.ru

Abstract. The authors give a theoretical and methodological basis for modern, distributed, multifunctional control systems' formation in regard to technical facilities and ships' navigation. They present a system of criteria for a scientifically based selection of technical decisions, related to structural organisation. Also, the authors formulate fundamental principles for arranging different types of structure. Moreover, they present new ideas both related to algorithmical and structural methods, whose realisation will provide higher levels of efficiency, safety and survivability of shipborne systems, based on its self-algorithmization and self-organisation.

Keywords. synthesis of a control system structural arrangement, functional and hierarchical principle, principles of functional and territorial decentralisation, network architecture, systems with knowledgebase, real time expert systems, self-algorithmization, self-arrangement, elegant degradation, cannibalisation, criteria's system, survivability, safety, efficiency.

Among a variety of problems, concerning substantiation of design decisions related to the development of modern control systems (CS) for technical facilities (TF) and ship, it is necessary to choose problems, connected with the selection (synthesis) of its structural arrangement, i.e. with substantiation of formation ideology and principles of functioning. The acuteness of the problem, dealing with structural synthesis has been stipulated by the fact, that both the ship and its functional control systems of TF, NPP in particular, relate to the class of emergency hazardous objects with higher risk of operation. The analysis of emergency accidents gives evidence, that the core of the problems, connected with provision of such vital features as safety, fail-safe operation and survivability, related to the ship and its technical facilities, lies in the field of the processes' organisation. These processes deal with the control and control systems' structure, which provide a high level of the above-mentioned features. The exact control systems carry out developed functions of emergency protection in regard to separate units and components, the functions of control over specific systems of safety and multi-unit complex' anti-emergency control, taking it as an integrity. Besides, the control systems provide purposeful functioning of the complex by the methods of co-ordinated or concerted alteration of units' operation modes and re-arrangement (reconfiguration) of its structure. At the same time, the efficiency of CS, regarded as an active means of TF survivability provision, is determined not only by the principles of control, the range of the tasks being solved, control functions being carried out, information processing algorithms or evaluation of control influences, but its structural arrangement and therefore its own survivability, failure safety and reliability.

The control systems are artificial products, so they are vulnerable to failures, malfunctioning, damages. They can lead to serious after-effects, if the structure is not properly arranged. It is evident, that the quality of CS structural arrangement is manifested, first of all, during failures in its operation and must be assessed by results of "TF + CS" interaction, which is considered as an integrity. The research, which has been conducted in relation to the problem of grounding the structural arrangement of a large CS, enabled the development of a methodical basis for its solution [1, 2, 5-8].

Among the most important and widely practically tested principles of CCS TF and incorporated large control system for controlling such functional complexes as TF, NPP, EPS, the complex of stabilisation and manoeuvrability, first of all, it is necessary to distinguish the functional and hierarchical principle for control systems' formation. This principle is a fundamental methodological basis for development of large-scale control systems.

The process of controlling the ship and all its facilities, according to its nature relates to the category of centralised processes and must be carried out from one centre in the frame of unified automated control system (ACS). The main problems, connected with the control system arrangement, which provides the ship's centralised control, deal with:
- key difficulties of formalised definition and solution of the common task, concerning control influences on a large amount of structural elements, which need co-ordinated actions for achieving common tasks functioning,
- limited opportunities of a common centre (control service and information processing TF personnel) for collecting and processing large information flows in a dynamically changing environment and inboard situation in real time mode,
- necessity of an unconditional provision of safety and survivability in relation to control processes, which are the most important means of achieving the tasks of the ship's operation in inevitable conditions during long time operation failures of TF and CS elements, and possible damages of TF main equipment, devices, communication lines under the impact of an emergency environment,
- determination of criteria for co-ordinated control over the main tasks.

The principle feasibility to realise centralised control in practice is provided due to ship's natural structural peculiarities as a controlled object, presence of exact structural formations or groups of elements, (aggregates of interconnected units, mechanisms and so on). The result of elements' joint functioning in each group is characterised by one or several physical values - output aggregated variables (co-ordinates of state), which reflect the role and the rate of participation of elements aggregate, being considered, in the processes of higher level objects' functioning. These variables, in essence, are the control influences for objects of a higher level. This natural structuring of shipborne objects of control leads to the feasibility of making a multi-level distribution of joint control tasks and split them into a number of local and co-ordinating tasks. The essence

of the synthesis procedure for a hierarchical system of interconnected tasks dealing with controlling such large objects as a ship is as follows [2, 5, 6]:
- structural elements (mechanisms, devices, units) are completed (grouped) in some sub-multitudes according to the attribute of closeness of physical (technological processes which flow in these elements, and (or) the unity of typical, universal tasks (executed episodes) of functions in regard to the assigned tasks for the ship ("horizontal "decomposition),
- common (global) tasks related to the ship and the aggregate of its functional complexes of facilities are being "split "into a multi-level system of interconnected local and co-ordinating tasks, which a common solution enables, a suboptimal solution of global tasks ("vertical "decomposition) to be found.

The most large-scale structural formations in the ship, which are considered to be a controlled system, are functional complexes of facilities (FCF), each of them providing execution of clearly determined typical function, which is universal in relation to the variety of ship's actions or assigned tasks. The typical functional complexes of technical facilities (FCTF) include the complexes of: power supply, navigation and manoeuvrability, provision of inhabitancy, provision of survivability. Other typical FCF of the ship include the complexes of: navigation, mapping of the environmental situation, external communication. The last three functional complexes are related to radio-electronic facilities (REF) or radio-electronic armament (REA). The complexes of ship's target assignment are not typical and are not discussed in this study. We can only note, that the most completed ship's target complex includes the complex of strategic missile weaponry (SMW), the complex of anti-ship missile weaponry (ASMW), complex of torpedo-missile weaponry and means of self-defence (TMW and SD). Selection of FCTF as the most large-scale (after the ship's level) controlled objects, at least leads to consideration of two-level structure of control tasks, which are being solved by the aggregate of TF: the tasks of control for each FCTF; tasks of co-ordinated control over this or that aggregate of FCTF when the ship is solving different assignments (actions), and also in emergency situations, which are not considered by CS of separate FCTF. Depending on FCTF complexity, the physical nature of processes, which flow in the complex and objective necessity in co-ordinated functioning of its elements, (units, mechanisms) the control tasks of separate FCTF are also split into a corresponding hierarchy of local and co-ordinating tasks. Thus, the control tasks related to these multi-unit complexes as NPP, are split into three-level structure of local and co-ordinating tasks.

The functional and hierarchical principle in control systems in relation to the sphere of controlled dynamic systems, generalises the basic principle of a systematic approach to the task of complicated systems` design, ship including [9, 10]. This principle means a "reasonable" decomposition of the system into sub-systems (tabulation of the optimisation task to 2- or more level scheme) and establishing such local criteria for sub-system optimisation, which help to attain an optimal or close to an optimal solution of the common task by solving the combination of practically independent tasks for the optimisation in each sub-system.

Here is necessary to stress the following basic difference between design tasks, related to large-scale systems of control and systematic tasks' type [9, 10]. When designing such control systems, the method of system decomposition into sub-systems, which is traditional for systematic analysis is insufficient. There is the necessity to use additional methods of decomposition, which makes a separate depiction of aspects, related to the CS arrangement and functioning. We have to say, that the notion "structural arrangement" [7] is to some extent "voluminous". Enough complete description of

structural arrangement and principles of functioning, can be made with the help of several diversified simulators or types of structures [7]: i.e. functional (what the system "does"), algorithmical (how the system "does" it), topological (where and what the system "does"), technical (with the help of what the system executes its functions), organisational (what the Centres of decision-making are responsible for - how operators interact). These structures essentially characterise the ideology of the system formation and principles of operation. The definitions of the above-mentioned types of structures are given in [7]. Such specific decomposition of control systems makes it possible to "split" the problem of design into a number of successively interconnected or interationally completed tasks of synthesis for separate types of structures.
Use of microprocessor means and systems with network architecture expanded the range of adopted decisions, related to "ideology". Thus, it is necessary to make a choice:
- network topology (star, ring, main circuit and so on)- on access to common channels of information communication (accidental, marker, mixed, marking ring, marking bus, multiple access with carrier control),
- protocols of communication (TCP/IP, MAP, DECnet, and so on) - way of information communication (wide band, single-band, and so on),
- strategy of controlling information exchange in the network (decentralised, centralised),
- distributed real-time operational systems - systems for database control and other common software and so on.
One of the most important problems is the evaluation of scientifically grounded criteria for the selection of design solutions and principles of arrangement, irrespective of who develops the arrangement and what software is used [1]. The designer needs criteria, which the systems must satisfy. At least, these criteria must be split into two groups:
a) criteria, which must be unconditionally executed irrespective of the ship's assignment at any (practical) expenses and irrespective of software which is planned to be used; examples of such criteria can be the criteria of the complete safety of NPP and ship navigation (cruise) safety at any, at least single, sudden malfunctioning of the object, control system or related support systems;
b) criteria, which execution is necessary to strive for, but the level of its execution (or its weight factors) must be determined in dependence of ship assignment and availability of different resources.

It is important, that these criteria and values take into account the influence of the controlled technical facilities, their structural and dynamic features, "fitness" for control automation. It is evident, that criteria of the second group need to consider the missions assigned to the ship, tactical and technical, technical and economical, military and economical values, the values of production efficiency in regard to control systems and so on.

Problem solving for the provision of CS structural arrangement's high quality levels leads to the necessity of a thorough analysis, concerning possible damages' after-effects on TF functioning quality level, and on the other hand, on losses' level, caused by accidents and risk level during safety conditions' violation. Such comprehensive analysis can be accomplished only on the basis of research, dealing with simulators of "object + control system (CS + TF)" complex operation. This complex is considered as an integrity at any environmental and inner disturbances of a different nature.

Thus, there is a necessity to widely use the fundamental principle of theory of control during analysis and synthesis related to CS TF structure. This principle presupposes the necessity to consider and investigate the "object + control system" complex as an integrity, using an appropriate set of simulators of different nature and assignment (dynamic, structural, probability, others).

In that way, the objective necessity to arrange and analyse simulators, which describe joint functioning of the object and control system, appears not only in the process of research, concerning classic problems of stability, dynamics and solution of optimisation tasks and processes algorithmization, but in selection of structures for control systems. The following hierarchical system of criteria was used in the basis of developed methodology for selection of technical decisions on CS structural arrangement (organisation): [2, 5, 6]

A. Determined criteria of structural stability, which are to be unconditionally executed (not taking into account the expenses of any resources):
A1. Criteria of "complete" safety (failure safety) of CS at any single malfunctioning, caused by the manifestation of natural unreliability, including possible accumulation of hidden failures in safety and emergency protection systems.
A2. Criteria of CS survivability during equipment and cables' partial damaging, when any single emergency environmental impact from standard multitude of incompatible likelihood hypothesises is realised.
Input criteria must be assessed and accomplished not statistically (on probability or on average for multitude of systems), but in a determined way, for each exact system and be acknowledged by object straight simulation, with failures of assigned multiplicity' imitation, but not in the form of resource tests.
Additional criteria in this group are the criteria of failure stability (e.g., the criterion of single failure in nuclear power engineering), the multiplicity of which is determined in regard to allocated volume of resources.

B. Probability criteria and values of functional efficiency and system reliability of the whole CS and also:
* probability (arbitrary and absolute) risk values of safety conditions' violation (for example, nuclear),
* losses from unreliability and total expenses for system development and operation.

It is important, that these values organically consider the characteristics of "TF + CS "complex' heterogeneous features (dynamic, reliability, systems of technical maintenance). Simultaneously, the selection of reasonable versions from the multitude of alternative ones is carried out. It is made in compliance with Pareto criteria in axises' type "efficiency - full expenditures" or "efficiency - safety (risk)".

C: The criteria of multi-level optimisation type "the effectiveness of expenditure". These criteria provide the solution of the task for optimum distribution of limited resources among systems of two co-ordinated ranks. C. The criteria's type "time of development - live labour expenses", which provide the correct correlation of requirements to CS TF with the state of automation equipment, possibilities of design departments and equipment-producing plants.
The criteria of the A and B group determine the structural arrangement of CS TF. They imply the execution of analysis and quantitative assessments of determined and probability values of safety, values of CS survivability and efficiency and also systematic values of reliability through the results of object functioning.

In order to accomplish the following analysis it is necessary to use the simulators of a different nature, level and assignment:

- imitation simulators of physical processes dynamics, aimed at information (knowledge, facts) output about faults after-effects,
- structural simulators in the form of logical equations' system or Boolean differential equations in order to receive structural functions of reliability, survivability and safety, and also a system degradation graph, which is the initial point for calculating both determined and probability values.

The calculation of probability values is made on the basis of simulators and programs for PC, which are developed on the basis of logical and probability methods, simulators of Mark processes of accelerated statistical simulation. The features of safety, survivability, efficiency and system reliability are the development of traditional notions, related to the theory of automatic control, i.e. the notions of movement stability and control processes quality for a wider grade of disturbances, which cause structural malfunctions in "Object of control (OC)+ Control system (CS)" complex, considered as an integrity.

The noted features of CS structural stability are provided during usage of the following principles related to structures arrangement [2]:
* the principle of hierarchical arrangement of functional structure;
* the principle of territorial decentralisation of organisational and topological structures, which provides crew structural stability, as an organisational structure with the retention of ship centralised control;
* the principle of functional decentralisation related to the technical structure with expedient selection of information processing devices number and initial distribution of functions among them;
* the principle of decentralised control over information exchange processes between information processing devices;
* the principle of connective stability related to closed control systems, at least in relation to part of the co-ordinates, emergency hazardous ones;
* the principle of dynamic task redistribution between data processing devices in the network architecture of control systems, taking into account the current state of workability and calculation load, task characteristics and task priority (applying to the system features of "elegant degradation" or features of cannibalism");
* the principle of reconfiguration of interacting duplex structures of programmable controllers with flexible utilisation of higher trustworthiness, reliability and productivity modes;
* the principle of maximum dispersal of electric power sources and distribution of the electric power supply system.

On the basis of these principles, the authors present architectural, design and algorithmical solutions [3, 4] which enable, under the conditions of object, equipment and cables (lines of communication and power supply circuits) partial damages, standard hypothesises to be made, related to damaging environmental impacts (including failure of equipment, device and cable of any single room (compartment), including central command post (CCP- GK., PEG, BSI and so on at the maximal accident level, envisaged by the design) to provide:
* prevention of the cascade development (transfer) of damages (breakage) of technologically linked equipment, which had not suffered direct damage from environmental impacts (prevention of "domino" effect);
* information exchange between devices placed in all the rooms, of the object, excluding (maybe) emergency devices;
* continuation of electric power supply for devices of all rooms, excluding (maybe) emergency devices;

* data access about the state of the object and its technical facilities, excluding (maybe) the equipment of the emergency compartment, with data concentration and mapping in the central command post or in an emergency (reserve) post (EPIC);
* preservation (realisation) of centralised control from CCP or EPIC over all technical facilities, excluding (maybe) equipment in the emergency compartment, providing co-ordination of actions, for example, of a ship, integrity of FCTF for maintaining the ship's navigation, controllability, navigation safety and other extremely important characteristics of the ship, and also
* full processing of information related to the solution of the ship and its technical facilities' survivability tasks.

In case of failures of separate modules and devices, including malfunctions and errors in functional programs, the control system continues to exercise its functions in full scale and without fail. The achievement of more than a single multiplicity of "bearable" failures are determined during the design period, both by limitations on weight and dimensional characteristics of the equipment and the efficiency of means for provision of failure stability at minimum use of structural redundancy.

Employment of network architecture, related to interacting integrity of microprocessor devices for data processing and control makes it possible to provide an "outgrow" or the development of a failure stability feature into a "strong" feature of elegant degradation (the principle of cannibalism), at which even with the absence of recovery (replacement) of inoperative modules or devices by operative systems (from a spare parts and accessories set), it reconfigures its own architecture, proceeds the continuous execution of control, at least as it vitally relates to the ship's operation equipment during the entire period of its continuous use or operation.

In the case of malfunctioning, caused by accumulation (or appearance) of a number of failures and its ability to surpass the minimum level determined for failure stable functioning, and also by failures of information sources and actuation mechanisms, ruptures and short circuits in lines of communication or power supply circuitry, discontinuities in supply of working fluids and other reasons, CS TF provides:
* prevention of issuing erroneous commands to IM (IOU) for unsanctioned changes of the commutation state of shipborne systems and equipment operation modes;
* prevention of an unsanctioned change of data in RAM database;
* fixation of IOU positions at breaks or loss of electric power supply and working fluids ("conservatism");
* in case of the impossibility to prevent "erroneous" commands or incorrect impacts on the controlled objects in regard to IM (IOU) (violation of functioning logic), CS TF provides:
- retention of emergency hazardous co-ordinates within permissible safety conditions limits (the principle of "connective" stability) with an objects transfer into a safety mode of operation with possible reduction of functioning quality (up to temporary disabling);
- prevention in the course of the transient process of disturbances' propagation in technologically linked controlled units for the elimination of its damages or its disabling by the means of local emergency protection.

The suggested integrity of interconnected technical solutions imparts to the ship's ACS basically new qualities, which are provided by an appropriate selection of its structure and operation (functioning), "implantation" of automatic "self-preservation" and "self-fitness" methods, which are practically homeostasis mechanisms.

References

1. Demchenko O.P. The problems of arranging automated control systems of shipborne technical facilities. Shipbuilding, (1976), № 8. (in Russian).

2. Voitetsky V.V., Simakov I.P. Development of methodology, theory and principles for the arrangement of ships' (vessels) control systems type ACS TP, technical complexes with higher risk of operation - In Jubilee Scientific and technical Collection of SPA "Aurora ". St. Petersburg, (1995), page 46-86. (in Russian).

3. Voitetsky V.V., Dmitriev A.N., Ipatov I.N., Korchanov V.M., Travkin N.A. The main principles of ACS TF related to advanced submarines - the same Collection, page 21-24. (in Russian).

4. Voitetsky V.V., Korneev V.N., Nemokaev U.S., Obuhovsky S.A. A new generation of complex and local systems for ship technical facilities.- The same Collection, page 24-34. (in Russian).

5. Simakov I.P. The system of criteria to make an assessment of technical solutions on the early stages of design, related to the control systems of ship's technical facilities - Collection of Scientific and technical Conference "problems of shipborne technical facilities' complex automation"- Shipbuilding, (1982), page 82-83. (in Russian).

6. Simakov I.P. Criteria of safety, survivability and control efficiency, successive synthesis of control systems arrangement. Collection: 1X All-Union Conference on Control problems. Nauka, 1983. (in Russian).

7. GOST 19176. The control system of a ship's technical facilities. Terms and definitions. Issue (1980) and (1985). (Developer SPA "Aurora ")(in Russian).

8. Simakov I.P. Program complex for probability analysis of safety, related to automated shipborne NPP and NPS during its design, operation and direct control. In Jubilee scientific and technical Collection of "Aurora" SPA, St. Petersburg, 1995. (in Russian).

9. Pashin V.M. Criteria for co-ordinated optimisation of ship's sub-systems - Shipbuilding, (1976). (in Russian).

10. Zaharov I.G. The theory of compromise in decision-making during ship design. Shipbuilding, (1987). (in Russian).

11. Simakov I.P. The metodological basis of synthesis concerning control systems' structure on the basis of safeny, reliability and survivability criteria. In Collection: X All - Union Conference on control problems. Moscow: Nauka (1986) (in Russian).

12. Simakov I.P. Methods of control systems structure selection, on the basis of survivability criteria at different level of the information availability, concerning environmental emergency impacts. In Collection: XI All-Union Conference on control problems. Moscow: Nauka, (1989) (in Russian).

GROUPE SCHNEIDER: ELECTRICAL AND AUTOMATION EQUIPMENT FOR MARINE TECHNICAL SYSTEMS

Helge Malmberg, Vladimir Khokhlovski, Vadim Petrossov
Schneider Electric, St.-Petersburg
3 Zvenigorodskaya Street, St.-Petersburg, 191126, Russia
Telephone +7 (812) 112-4174. Fax +7 (812) 314-7805
E-mail 101771.577@compuserve.com

Abstract. ***GROUPE SCHNEIDER*** is a leading specialist manufacturer of electrical and industrial control equipment. The company provides customers with recognized expertise and service, supported by the experience of its main brands: Modicon, Telemecanique, Merlin Gerin, Square D. The possibilities of the ***GROUPE SCHNEIDER*** equipment and tools for marine and offshore technical systems are presented. The electrical and automation equipment is suggested for use in the following fields: high and low voltage distribution, contacting, automation and safety systems (PLCs and local communication networks), uninterruptible power supplies. A number of applications are considered as well.

Keywords: High and low voltage distribution, automation and safety marine systems, programmable logic controllers (PLCs), onboard communication networks, uninterruptible power supplies, contracting and offer sales service.

Introduction. Presented in more than 100 countries, the ***GROUPE SCHNEIDER*** , one of the top French industrial groups includes around 100 000 people in 4 major industrial companies: Merlin Gerin, Modicon, Telemecanique, and Square D. The main attention in the presentation is paid to products with Merlin Gerin and Modicon brands.

Created in 1920, Merlin Gerin, part of the ***GROUPE SCHNEIDER***, operates in the field of electrical and electronic equipment and contacting, from generation and transmission to the use of electricity.

Merlin Gerin has equipped more than 2000 ships and offshore platforms, in compliance with the main classification societies.

Modicon was the first firm who created and introduced modular programming controllers in 1965. As a part of the ***GROUPE SCHNEIDER,*** Modicon deals now with PLCs, man-machine interface, communication networks and software tools to promote high quality service to customers in different fields of automation including marine applications.

Marine activity. This activity is handled exclusively by local specialists in France or in subsidiaries close to shipyards and shipping companies: 500 people, 70% of whom are engineers and technicians are involved in the following fields: high and low voltage distribution, contacting, ready-made automation and safety systems for marine applications, uninterruptible power supplies, services, PLCs and networks as a basis for building onboard control and monitoring systems.

Low voltage equipment. **Masterbloc Marine system** is designed in response to shipbuilder's and operator's safety needs in terms of reliability and maintainability in order to guarantee the security and availability of electrical installations. Switchboards built using Masterbloc marine system meet the requirements of the IEC 92 standards and those of the main classification organizations such as BV, LRS, ABS, GL, DNV, RINA, USSR RS, KRS, ZC.

In addition Masterbloc marine system fulfils the IEC standard 439-1 § 3 requirements for " type testing assemblies (TTA)". This means that the performance of each switchboard is guaranteed by actual equipment type testing and switchboard compliance with test samples. Thanks to the combined functions of the Masterbloc marine system, each specific switchboard is covered by the type testing certificates.
Masterbloc marine cubicles, designed to improve installation and personnel security criteria, are partitioned according to forms 2 (minimum) to 4 (cf IEC std 439-1).
Merlin Gerin adds to the strict application of standards its own security criteria, validated by testing, installation of auxiliaries in a specific compartment, natural HP (high performance ventilation, standardized insulating (epoxy) coating of internal structures.

High voltage equipment. **Fluair Marine System** switchboards are used in high voltage distribution from 3 to 17.5 kV on board ships and offshore platforms. They use metalclad enclosures housing fully withdrawable units. All assemblies use SF6 as the arc interruption and insulation medium: Fluarc circuit breakers and Rollarc contactors.
The Fluair Marine System cubicles have successfully undergone the following tests in accordance with IEC standard 298: dielectric tests, temperature rise tests, mechanical endurance tests, electrodynamic and thermal withstand short-circuit current tests, internal arcing protection tests.
Fluair Marine System switchboards comply with IEC international standards and most national standards. They also take into account the recommendations of the main classification societies. All Flair Marine switch-boards already delivered were approved by DNV or LRS or BV.

Ready-made automation system. **Maestro** is an automation system made to be integrated on ships in order to provide:
- greater security and operating comfort,
- higher performance and installation availability,
- reduction of maintenance costs,
- possibility of designing ship to be run by fewer crew members.

All Maestro system hardware and software components are designed, implemented and marketed by Merlin Gerin.
This is the shipping companies' and shipyards' assurance of the companies' sole responsibility and of the systems durability.
The system has a distributed type architecture built around a high performing communications network. It permits greater continuity of service than a centralised architecture would.
The Maestro system is fully modular (hardware and software); there are fewer components thanks to the effect of belonging to a range (e.g. the same I/O cards regardless the type of function); the maintenance procedures are very simple and few spareparts are required.
Apart from these assets derived from the technological choices mentioned above, the following advantages are also taken into account starting at the design phase: ergonomy of the man-machine interface, opening towards other systems and the management level, tolerance of internal failures, possibility of "fall-back" operating mode.
Maestro software modules enable the following functions to be performed: alarm and measurement monitoring, power management, main engines monitoring and control, generator sets monitoring and control, auxiliaries management, supervision.
All of these functions can be connected to same network.

The various functions required for its use can be controlled from different points on the ship, in particular from the wheelhouse. Quick, ergonomic control and display are provided by multipurpose workstations (VDU, keyboard, trackerball) suited to marine constraints.
The system is customised to meet the client's needs by parametering-setting in the plant or aboard the ship.

Services. Merlin Gerin acts as well as a world servicing network including in particular teams of marine specialists settled close to the main shipping routes.The maintenance services of the marine activity of Merlin Gerin includes:
- commissioning and sea trials,
- preventive maintenance with contract upon demand,
- troubleshooting (24h-a-day service if necessary),
- renovation,
- spare parts sales,
- crew training,

and has been entrusted by more than 50 shipping companies for many years.

PLCs and other automation equipment. The following range of PLCs can be utilized by system integrators, contractors, and ship builders (in the order of increasing functionality): Modicon TSX Nano, Modicon 984 Micro, Modicon TSX Compact, Modicon TSX Micro, Modicon TSX Premium, Modicon TSX Quantum. The PLC technical characteristics are listed in the presentation.
For communication from a controller to a controller, a controller to man-machine interface, or a controller to a host computer, two main industrial networking strategies are provided: Modicon *Modbus* and Modicon *Modbus Plus*. The former is one of the most reliable, economical and popular industrial networks, which may be used for long distance connectivity with different types of modems, e.g. telephone and radio modems. The latter combines high speed communication and easy installation to simplify application implementation and reduce installation costs in local area networks. Combinations of these two networks can be used to provide simple high performance architectures, which meet the goals of integrated automation solutions.

Modbus network. Modbus is a *master/slave* network. All communications are initiated by a master device. Up to 247 nodes can be addressed from it. The master device may be a controller, a host computer, operator interface, or a programmable panel. The host computer performs programming, data transfer, upload/download, and other host operations with a number of connected controllers. Modbus offers flexible communications. Communications may be a *query/response*, where the master addresses only one slave, or a *broadcast/no response,* where the master simultaneously addresses all the slaves. Baud rate (up to 19200) and parity are selectable. A local system may be configured by directly connecting the master to the slave device via a standard RS-232C connection. This is a point-to-point configuration, which is used typically for connectivity to a programming panel for local operation interface. One can also connect the controllers using RS-232 modems over common carrier (phone line, radio, fiber-optic line, etc.).

Modbus Plus network. Modbus Plus network offers three very important advantages over Modbus: the higher speed, peer-to-peer communication, and a token bus principle of operation. The first is 1 Mbaud vs. 19.2 Kbaud. The second permits any node on the link to initiate transactions with any other. The third, the token bus media access control provides democratic, deterministic access of nodes to the medium. Thus, nodes on a Modbus Plus network function as a peer members, gaining access to the network upon receipt of a token frame. When a node holds a token, it can initiate message transactions with selected destinations. The standard Modbus Plus network supports up to 32 nodes at distances up to 500m. For applications requiring more nodes or greater distance, a single Modicon RR85 Modbus Plus Repeater permits addressing of up to 64 nodes and up to 1000m. Up to three repeaters are permitted to extend the network to 2000m and 64 nodes. For applications requiring access to more than 64 nodes, the Modicon BP85 Modbus Plus Bridge Plus permits to link two Modbus Plus networks. This is especially useful when one want to link many small networks for optimal performance.

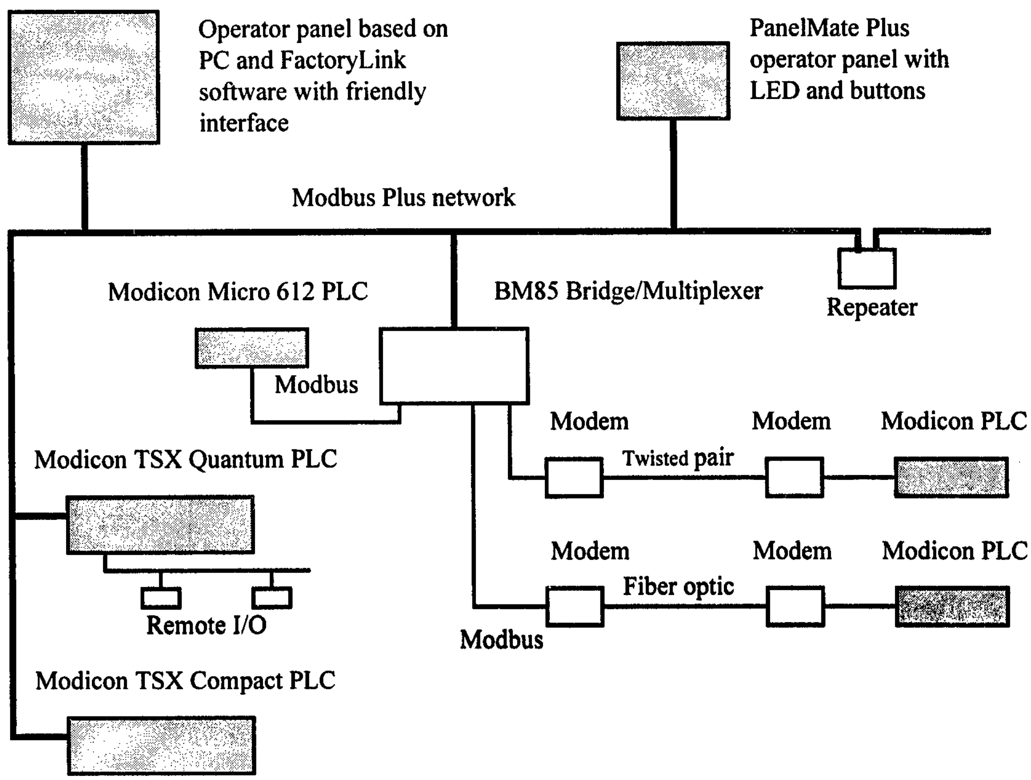


Fig. 1. An example of the Modicon network

In applications where standard Modbus devices require access to Modbus Plus network data, a Modicon BM85 Modbus Plus Bridge/Multiplexer offers four Modbus-compatible RS-232 serial ports which permit Modbus master or Modbus slave devices to link into Modbus Plus network.

Modicon offers a wide range of man-machine interface products to work with a control system. The family of products includes cost-effective replacements for applications requiring traditional discrete push-button and message display units, industrial control workstations, as well as software for production and supervisory control: 8, 16, 32 element operator panels, ASCII operator keypad, PanelMate Plus video control panel series, FactoryMate Plus industrial control workstations series with open architecture. An example of a network is shown on the Fig. 1. Grey blocks are active nodes.

The Modicon compatibility strategy extends beyond its own products. The popularity of Ethernet has made it a requirement that Modbus Plus be connected to the more popular implementations, including DECnet and TCP/IP. Application programs running on Ethernet benefit from all the properties of Ethernet and Digital's Pathwork, including global connection to Ethernet services that permit virtual disk access, program archiving, file transfer/sharing, and many others. Modicon provides also terminal I/O modules for INTERBUS-S and PROFIBUS-DP industrial communications networks.

Modicon programming software enables users to perform off-line program development using standard IEC languages, on-line program maintenance, generating user documentation, I/O module health and status monitoring, process/machine diagnostics, specialized functions for communications.

Thus, Modicon networks of various topologies for ship automation can be implemented to meet user requirements. For instance, a distributed network for measuring critical technical parameters and monitoring can be built with remote inputs, local preprocessing controllers, and an advanced operator panel for data analysis and graphics. Communication among up to hundreds of nodes may be provided onboard. On the other hand a system for local control can be based on powerful controllers and simple operator panels for man-machine interface interconnected with Modbus Plus local area network of high performance. A number of typical network topologies and applications are considered in the presentation.

WWW servers available. More information on electrical and automation equipment, software products and services one can browse through the Internet at web sites in USA and Europe:

http:\\ www.schneiderelectric.com
www.schneider.co.uk
www.modicon.com
www.ups.merlin-gerin.com

SAILING SAFETY AND COORDINATED INTERACTION OF TRANSPORT SHIPS FOR ARCTIC OIL - GAZ DEPOSITS

Victor K.Bolkhovitinov, Yuri V. Bagljuk

(Science & Production Corporation "Aurora"),

Vladimir S. Vovk (RAS "Gazprom")

Russia

Abstract: The automatic control of ship handling process at following of vessels to a place of operations and control procedures means at execution of operations in interaction with other vessels and extractive platforms is based on application of allindustrial hardware-software tools of computer facilities, integrated of tools of navigation and electronic charts, local control systems of ship executive mechanisms and devices which are conforming international standards and rules of navigation safety.

Keywords: automatic control system, shore navigation, discrepancy safety, coordinated interaction of vessels, electronic chart, maneuvering of vessels in group, dispatching and training systems.

Introduction: The international requirements of sailing safety are put in to the transport vessels of technical and serving fleet at their maintenance in regions of marine Deposits on a shelf of the Arctic seas of Russian Federation in full size as in modes of passages to a place of operations or from the port to the port, and at execution of technological operations in water area of a platform or terminal. These problem questions are complicated and have interdependence at decision of the safety tasks in use of marine oil-gaz deposits. This safety of transport ships we will name as " technical sailing safety" of the vessels during the use of these deposits.

1. The common part

The practical implementation of the requirements IMO, Russian Maritime Register of Shipping, Det Norske Veritas to technical sailing safety is connected with motion control systems of transport (technical or serving) vessels (CS MRC), interacting with satellite and coast systems of navigation, hydrometeorological and ice decor in region of vessel motion and in its destination. In particular, the tasks of vessel discrepancy safety, passing narrow waters, humble water areas and waterways are decided at automatic control of vessel motion on specific trajectory at monitoring of speed, heading and course in conditions of weather and state of the sea.. The use of dispatching and training systems in the aim of raising of operators professional skills is directed to technical safety too.

In addition to this, in / 1 / said: "It has been realized that offshore ('technical") safety should be treated much more seriously, as has been done previously in the offshore process industry, and aviation and nuclear power. The assurance of a satisfactory level of safety is now seen as necessary component of a successful offshore proejct."

Accordingly, CS MRC decides the following tasks of vessel sailing safety and its coordinated interaction at share of vessel sailing in technical and serving fleet, for example, in a structure of group " the ice breaker - oil ship - platform ", " an oil ship - terminal - ice breaker " etc.:

- Support of required speed and heading set;
- Maneuvering at a mooring or at pass narrow waters;
- Creation and improvement of recommended parameters of discrepancy with marine objects;
- Imaging a situation on motion route on an electronic chart in a format of the international standard;
- Creation of the messages about parameters of current location, speed, heading, roll and trim of a vessel etc.

In the report the outcomes of executed researches in use CS MRC in specific modes of motion and vessels coordinated maneuvering of technical (serving) fleet are represented.

2. The requirements to operators of CS MRC and the training system

The successful using of complicated integrated automated motion control systems and positioning of serving fleet vessels for oil-gas marine complexes shows high requirements to professional qualification of operators-navigates and technologists.
The professional skills of vessel control procedures and technological equipment's operators are especially important in unexpected, unnominal situations, including emergencies, when it is necessary to select the most effective solution by optimal using of involved and again connected ship technical means, equipment and systems, and also to accept solution on organization of serving fleet vessels interaction for elimination of given situations.
Indicated problems: ship control systems operators and operators of dispatching control systems operators preparation and learning dictate the requirements on creation of appropriate imitative and training systems and complexes. These systems and the complexes should ensure in real time scale simulation of probable situations of vessel motion and operation of a production equipment in external weather conditions, at probable refusals of system units etc., which in an actual decor of maintenance of systems and vessels or require sizable expenditures of time and tools, or create undesirables consequences of accidents with large damage, including the vessel loss.

The offered engineering solutions on the imitative computer system of researches of changes oil-transportation and freight traffic, ensuring required modes of operation of the transport system Prirazlomnoe petroleum deposit, and also systems of similar assignment, and training complex adequate on learning operators of serving fleet vessels are considered too.

3. The example of practical realization of technical safety

The support of construction surface and underwater operations, and also execution of some cargo operations and operations on underwater pipe laying with creation of marine platforms for drilling, terminals and complexes is connected to use multifunctional universal crane-ship of large rated load capacity arranged with an automated control system it with means (ACS TM).
As these means, first of all, the rotary columns (RC), sliding rotary columns (SRC), tunnel thrusts (TT) and anchor cord winches (AW) are considered which at simultaneous operation on signals of ACS TM ensure controlled introduced motion of a vessel on specific trajectory of motion at laying of underwater pipelines, exact positioning of the rather building object or platforms, and also maneuvering of a vessel with simultaneous cargo motion to a specific point.
The interconnected structure of ACS TM realizes required modes of operation of the anchor system of positioning (ASP) both dynamic system of positioning and maneuvering of a vessel (DSP) at their separate and share operation.

ASP in automatic and remote operational modes controls a tension of twelve anchor cord lines, which layout varies concerning ship hull in the process of anchors transfer with towboats controlled by an operator - technologist at underwater of pipe laying, and also ensures processes separate and combined with DSP positioning of a vessel above a point of operations.
DSP in automatic and remote modes controls motion of a vessel at passages (with use the RC) and ensures high-precision maneuvering and positioning of a vessel at support of operations in a specific point.

The high reliability of operation of ACS TM in whole is ensured with an automatic choice of necessary parameters SRC, RC and TT in case of a breakaway of an anchor line ASP at operation of heavily external effects of a wind, sea-way and current on ship hull.
In the report the outcomes of researches of controlled processes of motion and positioning crane-ship, projected CDB "Corall" under the job RAS "Gazprom" are represented, and also the main characteristics and parameters ASP, DSP and ACS TM in whole are reduced.

The specific modes of crane-ship intended for execution of technological operations on building marine extractive complexes and support of cargo operations at civil and erection works, are ensured with the automated control system of crane-ship (ACS CS).
In an organizational structure of ACS CS the central place is taken by combined system of crane-ship positioning consisting of two interacting control systems of hardware components: the anchor system of positioning (ASP) for actuation of anchor lines at handle of their tension depending on a level of external revolting effect and system of dynamic positioning (DPS), ensuring coordinated control of tunnel thrusters (TT), rotary columns (RC) and sliding rotary columns (SRC) of a vessel.

On a design stage of the combined system of positioning the structure is developed and the composition of hardware-software tools ASP and DPS is offered, as which the standard instruments and control panels from a package of tools of automation (PTA) "Aurolog", approved by Russian Maritime Register of Shipping software of the modern operational QNX system with real-time mode Rtime, and also hardware-software tools of a network process engineering Ethernet ensuring required parameters of reception - data transfer and controlling signals ASP, DPS and adjacent ship systems are applied.
The software ASP, DPS with the purposes of rise of a reliability have stand-by basis channels of input-output, processings, representation and imaging of the information, use standard interfaces and are constructed on functional tag of assignment of used procedures and components.

The algorithmic support of systems is executed in view of the requirements of selfdiagnosis of program complexes both possibility of testing and debugging at regulations and set-up of the specialized software - functional programs. Thus the requirements to hardware-software tools on their operation in a mode of call by an operator of programs and algorithms simulating of probable situation of controlled positioning of a vessel and learning of operators at improvement of control skills, including in unnominal situations also are taken.

4. The safety at interaction of ships with platform (terminal)

Main mode of interaction of an ensuring vessel with producing platform or marine terminal at accepting - transmission of cargoes and support of technological operations is the confinement of a preset position (distance and bearing) in conditions of operation on a vessel of external perturbations from a wind, sea-way and current.

Dynamic positioning and controlled vessel maneuvering at confinement of preset distance and bearing concerning a platform or terminal are ensured with a ship automated control system interacting on the information channel to a dispatching control system of a platform (terminal). Main criterion of effective and safe support of technological operations in an indicated mode is observance the compromise requirement of necessary productivity of the process of reception - transmission of cargoes in conditions of a marine bad weather at a support of ecological safety. In the report possible practical modifications of rational positions of mutual layout of a vessel concerning a platform for variable conditions of the external marine environment, state of a navigational decor in region. mode of operation of a platform are represented and examined.

On the basis of made analysis of positioning variants the additional requirements to interacting control systems are stated, and also the modes of their practical use in training complexes to vessel control skills improvement are determined.

Conclusions: Thus accumulating knowledge, statistics of real situations with interacting control systems, sofware and professional skills of operators of mentioned systems, the questions of technical safety for vessels of serving fleet may be decided successfully. Their decision has to have place during the design work of concrete ship automated control systems, dispatching and training systems that have to be used on marine oil-gaz Deposits in region of Arctic seas in Russian Federation.

References:

1. Dr. Alexis N. Chernoplekov, Dr.Tony Cox, Offshore Safety: Engineering and Regulation, The Second International Conference on Development of the Russian Arctic Offshore. September , 1995 , Saint-Petersburg, Russia

SWITCH-TECHNOLOGY.
ALGORITHMIC AND PROGRAMMING METHODS IN SOLUTION THE LOGIC CONTROL PROBLEMS OF SHIPPING EQUIPMENT

Yuri.V.Bagljuk, Anatoly.A.Shalyto
(Science and production
corporation "Aurora",
Saint-Petersburg, Russia)

Abstract. On the basis of reseaches and practical results of the author's new technology of algorithmic and programming of logic control problems, including the complex technological processes,is offered.

Keywords. Logic control, state, state diagram, transmission graft, finite state machine, automation, automata, algorithmization, programming.

Introduction. A choice of graphic language- state diagrams (transition graphs)- of the specifications allows, from the one hand, reasonably, mathematical strictly, completly and nonconflictly to describe problems of this class in the form, clear to the Customer, Technologist, Developer, Programmer, User and Controller, and with other- formality and isomorphicly to build the programming language, which are well understood, structured, observable and manageable.

Environmental Impact. Methods of formal synthesis of functtional programms for three classes of programming languages are stated: algorithmic languages of a high level; algorithmic languages of a low level; specialized languages. These languages correspond to three classes of computing devices, now in use in practice of designing a control system: to industrial computers; tomicrocontpollers; to programming logic controllers (PLC).

With reference to the last class of devices the special methods are stated for the representatives of all classes of languages according to the international standard IEC 1131-3 (languages of the instructions, ladder and function diagrams, sequential functional charts). When carring out the development we aimed to turn the programming of the problems of this class form art into the science.

The following cognitive properties of technologies based on atomata structures can be considered important for the stages of design and application: convenience of mutual understanding of professionals from different areas.

A hierarchy of knowledge representation is discussed at two levels: more generalized, algorithmic level and more detailed or higher resolution programming level.The semantic of the programming should maximize correspondance to the meaning of the algorithmic level but disclose more details. The meaning should be understandable for both the designer and later for the programmer.

Authors have introduced the mechanism of automata interaction which allows for performing the transformation of description from thr lower to the higher level of resolution during the top-down design, and reverse transformation while the realization of the description during the analysis is performed bottom-up. We present this mechanism for a class systems for control of the complex technological objects which we call systems of logical control (where variables have two values: 0 and 1). For these systems their behavior can be crisply described by the state diagrams.The concept of state is fundamental in these class system.There are three classes of descriptions of processes in algorithmic form which can be called: linguistic, tabular (or matrix) and graphical. Authors demonstrates that frumens understand better the descriptions presented in a graphical form - state diagrams. Different alternatives (for example, sequential functional charts) are discussed. However, for important and responsible objects of control (for example, in the case whenthe nuclear objects are involved) thr formal transformation form the algorithm to the programm can be insuffisient because for cecking the correctness of such transfer especially when programming in done manually, the inspectiny individual has to have an opportunity to verify and check the proof of the correctness of using simulation to provide and opportunity to extract new errors otherwise it is impossible to be sure that the program is error free.

The structure of programms developed permits to modify them easily causing minimum errors. These programms also allow for easy observation which simplifies considerably their debugging. There are specifications which are proposed to be served also as tests.

The possibility of composition and decomposition to the complex algorithms on the interconnected modules is also provided. The methods can be also used for the control objects modeling, which permits to develop on PC the suitable model of the complex system "of the interconnected controlling automatic units".

There was also developed the permitting to reach better visualisation of the process (to transfer from the potential observation to real one) due to development of lager scale systems realised with C-programs. This program allows to modify the input data and to visualize on PC output and varying data but also, and this is the most inportant thing, the signatures of each components of the complex system.

This program can be also used for the algorithms modeling using C-language before the automatic or manual relization on the basis of the low-level languages.

The advantages of the approach proposed in comarison with the other methods of solution of the said problems include also those using GRAPHSET language (Telemecanique Company,France) are proved.

Conclusion. The offered technology allows to divide work and main responsibility between the Customer, Developer and Programmer, that is especialy important, when they represent different organizations and furthermore the countries, otherwise arise essential language and, finally,economic problems.

The proposed technology was realised impemented in:

development of the system of control on the diesel-generator of the timber-carring vessels of the project 15760 with the help of SELMA system (ABB stroberg Drives, Finland);

development of the system of control on the diesel-generator of beam-carring vessel of the project 15640 with the help of DATACHIEF-7 system (Norcontrol, Norway);

development of the vessel control systems with the help PLC AUTOLOG (FF-Automation, Finland) and complex of agregated means AVROLOG (Aurora, Russia).

References

1. Shalyto, A.A. Cognitive Properties of Hierarchical Representations of Complex Logical Structures. Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems. 10th IEEE International Symposium on Intellegent Control 1995, Monterey, California.

2. Shalyto A.A.(1996).The Use of Graph Diagrams and Transition Graphs in Program Implementation of Logistic Control Problems.Automation and Remote Control, №6,7.

3. Shalyto, A.A., V.V. Antipov(1996). Algorithmic and Programming Methods in Solution of the Logic Control Problems of Technical Devices. Morintex, St-Petersburg.

4. Shalyto A.A.(1997).Switch-Technology. Algorithmic and Programming Methods in Solution of the Logic Control Problems. Nauka(Science), St-Petersburg.

Principles of Creating and Experience in Introducing a Management Information System for a Motor Transport Agency

Ronald A. Nelepin, Alexander A. Solovyev
St.-Petersburg State University
Faculty of Applied Mathematics & Processes of Control
e-mail: sandro@ptn.spb.ru

Abstract. The paper is concerned with problems in automating processes and operations at motor transport agencies (MTAs). The issue is examined in the context of developments in computer technology and changes in the economic conditions in which MTAs now operate. Basic principles of building an MTA management information system (MIS) are presented, task-doing concepts are suggested, and an example of introducing an MIS is given.

Keywords. Workstation, motor transport agency, computing center, artificial intelligence, personal computer, data base management systems.

Introduction. Automating processes and operations provides a means for enhancing productivity and the efficiency of motor transport agencies (MTAs), which is due to the development of computer technology, a decrease in computer equipment cost, and the enhancement of computer capabilities. In the recent decade, the speed of PCs has increased dozens of times. Internal memory is now measured in gigabytes rather than megabytes. Huge information networks are being set up which join together producers, consumers, banks, governmental entities, etc. This enables production management problems to be solved in an expedient and optimum manner using PC capabilities.

Company management is known to be made up of components, viz. management proper and its information support. All actions encompassed by a "decision-making system" [3] are assigned to the former whereas a process assuring the acquisition, processing and analysis of information required by a company's executives, to the latter. In 1980s, production automation meant specifically computerizing the latter management component. Data bases (DBs) were being created and data base management systems (DMBSs) were developing vigorously by virtue of which speed considerably increased and the quality of processing data files improved. However, the stress has shifted lately to the automation of the "decision-making system" itself. Problems of creating artificial intelligence (AI) and expert systems (ESs) are being solved successfully.

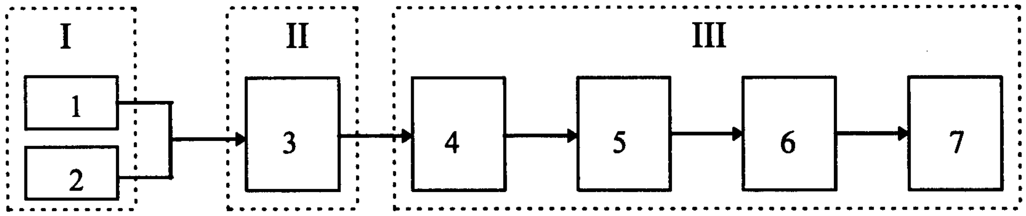
This paper deals with the problem of automating and computerizing a motor transport agency (MTA) [1-3] taking into account today's trends in the development of both MTAs themselves and computer technology.

While the document flow-chart for an MTA's primary transport documents (PTDs) in the past might be represented by Fig.1 [1,2], the industry de-centralization process now has resulted in every company's striving for a data processing system closed within the company and constituting a commercial secret.

Task-doing concepts. The computerization task can be done in different ways such as:

- a comprehensive agreement with a company specializing in management information systems,
- enlisting the services of a company's own computing center,
- a combination of both.

Fig.1 MTA's primary transportation document flow-chart



1. MTA: DDT preparation
2. Customer: WB preparation
3. Customer, MTA, Driver: marking engineering and operational details of PTD
4. MTA: PTD primary check and run
5. Computing center (MTA): PTD primary processing (getting totals)
6. Computing center (MTA): getting report runs
7. MTA (computing center): getting MTA reports.

Herein:

I: preparatory stage, II: working stage, III: check / computational stage

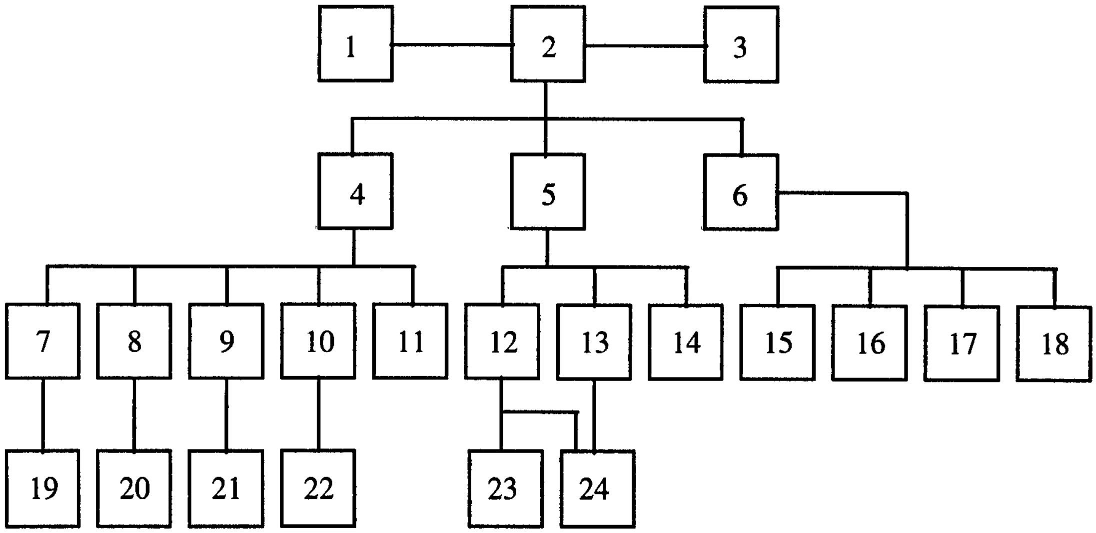
Computing center serves simultaneously several MTAs. Depending on the operation scheme adopted, performs most of the check / computational jobs.

DTT: driver's trip ticket

WB: waybill

DTTS and WBs are principal primary transportation documents since they contain records of actually all of the operations performed by drivers, the MTA and customers.

Fig.2. Sample MTA structural chart



1. Personnel department
2. Manager
3. Administrative housekeeping department
4. Chief engineer
5. Manager, operations department
6. Finance controller
7. Procurement department
8. Production manager
9. Technical department
10. Mechanical manager
11. Electrical manager
12. Operations service
13. Traffic controller
14. Cct
15. Accounting department
16. Legal adviser
17. Foreign relations department
18. Planning department
19. Warehouse
20. Maintenance area
21. Fuel and oil
22. Telephone exchange
23. Traffic & occupational safety
24. Vehicle fleets (branches)

The first way presupposes writing a comprehensive request for proposal defining the company's tasks and structures, and software projected. Here, all particulars and details of the system to be created must be provided for in advance. Then a soft- and hardware platforms for implementing the system are chosen and purchased, a demonstration version is created which constitutes the shell of the system to be created, a final program product is accepted, and personnel is trained.

This option has the important advantage that everybody is engaged in his own business. However, it has equally important drawbacks such as:

- a high cost of creating such a system,
- the impossibility of expedient changes in and repair to the system in emergency because of strong dependence on the contractor's experts,
- breaking off relations with the contractor (because of his ruin, ceasing his activities in this field, etc.) deadens the system, not allowing for further development, which inevitably results in new expenses for updating or even creating a new program product,
- difficulties in making changes in the system,
- the impossibility of providing all details at the initial stage.

The second option features much smaller expenditures and their time span, a step-by-step problem-solving, the possibility of expedient changes in the system, and virtual independence from systems developers since the original program texts constitute the company's property. However, there are also drawbacks in this option:

- a long time required to create the system,
- certain inevitable flaws,
- the necessity of keeping track of changes in legislation on one's own and making appropriate changes in programs.

The third option is a skillful combination of the first and the second automation options through the use of the advantages of both of them. E.g. it is worth while purchasing a legal reference book program and getting the computing center's experts develop a system for processing drivers' trip tickets (DTTs) and waybills (WBs).

Let us dwell now upon the choice of a hard- and software platforms.

The most widespread computers in Russia now are IBM compatible PCs so that it is no problem to find software for them. E.g., a 486DX2-66 IBM PC with a 360 MB winchester and a 8 MB internal memory should suit a company which needs computers to process text and numerical data. However, if the use of the MS Word 7.0 text processor or the MS ACCESS 7.0 system for data base design and maintenance is provided, then a 586DX-100/650/16 IBM PC would be required as a minimum for normal operation.

Since the very statement of the problem presupposes the necessity of an intensive joint use of DBs, setting up a local network is suggested. This raises the question of whether a server is required. It should be noted that the server is a powerful high-performance computer (IBM PC 486DX4-100/1.2Gb/32 as a minimum) which serves:

- to store DBs, programs, etc. shared by different users (a DB or file server),
- to organize and control a joint printer (a printer server),
- as a fax/modem server, etc.

The server processes users' requests, while being invisible since it cannot address them by itself.

The customer is a user's workstation and accordingly serves:

- to provide a user's interface,
- to create requests and send them to the server,
- to process request results received from the server and supply them to the user.

Let us now address ourselves to the software platform. A family of DB management systems (DBMSs) based on the dBASE language [5,7,8] have received wide acceptance in Russia. A total absence of means assuring data integrity is a major shortcoming of such DBMSs. However, there are quite a few other software platforms in which this problem has been solved. The MS ACCESS 7.0 is one of them.

Principles of setting up a MTA management information system. In creating a MTA management information system (MTA MIS), the following basic principles should be adhered to:

1. The system must simultaneously meet several quality criteria, the most important of which are maximum reliability, minimum cost, attaining a maximum profit as the result of the introduction of the system, convenient use, the capability of up-grading the system, serviceability, etc. To take into account all of the requirements to the system, a target graph and vector optimization criterium are constructed. For system optimization on the basis of a vector quality criterium, Pareto's optimum system theory is used with AI elements involved.
2. It is our opinion that the basic subsystems of a MTA MIS should be constituted by the calculation / check team's workstation, the accounting workstation, the traffic control workstation and the manager's workstation whose purposes are presented in what follows.

Elements to be introduced. Figure 2 shows an MTA's structural chart by way of illustration. The MTA incorporates a check / calculation team (CCT).
Prior to management automation at MTAs, this team was responsible for calculating the portion of driver's salary which is directly determined by the information contained in DDTs and WBs.
This particular MTA gave the task of setting up a workstation for the CCT.

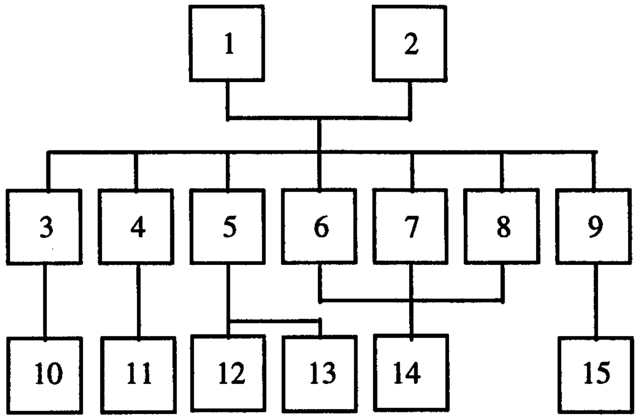
Task-doing method. This job can be done by a purely programming method. That is why the principal problem is to define a correct and clear data structure. It was decided to write a program in FoxPro [5,7].
The initial task in the process of creating a CCT workstation eventually became but one of the items of a program submenu. This was due to the fact that many groups of the MTA took DDT and WB data in parallel. It made sense to process these information sources on a computer and in one group only since one department only was being automated initially.
Data were structured as follows: DDT, WB, PERSONNEL, FLEET, FUEL, CUSTOMERS, BANKS, TIMESHEET, other ancillary DBs.
It was decided to make up the CCT workstation of three program blocks:

- ADMINISTRATOR,
- DATA ENTRY,
- WORK WITH REFERENCE BOOKS.

DATA ENTRY is responsible for entering data into a buffer DB and pre-checking for operator errors.
ADMINISTRATOR is responsible for data transfer from the buffer DB to the accumulation and restricting an access to data for unauthorized persons (password check).
WORK WITH REFERENCE BOOKS is responsible for data processing and analysis, and issue of reports and other documentation.
Figure 3 shows the structure of the output data of the CCT workstation.

Results. The CCT workstation thus created has reduced substantially the time required to process DDTs and WBs not only at the department comprising the CCT but also at all MTA

Fig.3. Structure of the check / computational workstation's output data



1. WB processing
2. DTT processing
3. Fuel consumption accounting
4. Maintenance data
5. Vehicle runs
6. Vehicle depreciation
7. Drivers' salary
8. Invoicing customers
9. Timekeeping
10. Fuel and oil
11. Production department
12. Technical department
13. Planning department
14. Accounting department
15. Personnel department

departments which use the data of these documents, enhanced the calculation accuracy, precluded processing, in parallel, the same information, reduced significantly the routine paper work, and increased the expediency of making decisions.

Prospects for further automating the management process at the MTA depend on the issues presented below:

The development of an ACCOUNTING workstation [4,6] which is to handle most of the financial problems. Provided specifically, within its framework, is the availability of modern communication with the bank of the MTA, which will enable payment documents, invoice reports and currency exchange information to be sent or received without leaving the agency.

It it necessary to determine an optimum route for the vehicle starting from data on roads, parkings, service and filling stations, season, binding delivery time, etc. Solving this problem, necessitates the construction of a strict mathematical model, the use of graph theory, the parameter space section method, and optimum control theory.

The job ends in setting up a MANAGER workstation, which presupposes writing, in one of the object-oriented programming languages, e.g. C + + or SMALLTALK, an expert system using the data of all of the MTA's groups. This program will bring the manager up to date on all of the current affairs of any department, provide for an optimum solution to business problems, enhance the expediency of managing the company.

Conclusions. The results of the job done can be extended to Russia's urban and rural motor transport agencies of up to five hundred vehicles.

References.

1. D.P.Brunstein, V.I.Podolsky, and B.P.Savitsky. Automation of Motor Transport Reporting and Calculation Work. - Moscow: Transport Publishers, 1986. - 247pp. (in Russian)
2. D.P.Brunstein. Computing Centers in Motor Transport Data Control Systems. - Moscow: Transport Publishers, 1988. - 175pp. (in Russian)
3. Yu.A.Kaftanyouk. Computer Technology in Motor Transport. - Moscow: Transport Publishers, 1985. - 183pp. (in Russian)
4. T.A.Krayeva. Accounting Methods and Organization in Automation Enviroment. - Moscow: Finance and Statistics Publishers, 1992.- 160 pp. (in Russian)
5. L.Pinter. FoxPro 2.0 Applications Programming /translated into Russian). - Moscow: Edel, 1994. - 384 pp.
6. L.M.Polkovsky, S.A.Zaydman, and M.Ye.Berkovich. Accounting Automation on the Basis of PCs. - Moscow: Finance and Statisitics Publishers, 1991. - 192pp. (in Russian)
7. I.B.Rogachev. Programming in FoxPro DBMs: The User's Guide. - Nizhni Novgorod, 1992. - 116pp. (in Russian)
8. S.R.Clipper. Programming Guide /translated into Russian/ - Minsk: Tivali, 1994. - 480pp.

CONTROL SYSTEMS WITH KNOWLEDGEBASE: MATHEMATICS, DEVELOPMENT TECHNOLOGY AND USAGE OF AUTOMATED COMPLEXES IN THE SYSTEMS OF OPERATORS` INTELLECTUAL SUPPORT SYSTEMS.

Igor P. Simakov
Scientific and Production Association "Aurora"
of the Ministry of Defence Industry of the Russian Federation,
15, Karbysheva Street, St. Petersburg, 194021, Russia.
Telephone: +7(812)535-97-23 (home) E-mail: ppnovod@dux.ru

Michael A. Zuenkov, Alexey G. Poletykin
The Institute of Control Sciences, Russian Academy of Science, Moscow.
65, Profsojuznaja Street, Moscow, 117806, Russia.
Telephone(Fax): +7(095)334-75-71 (works)

Alexandr N. Novodjilov, Igor I. Simakov
Naval Academy named after Kuznetsov N.G., St. Petersburg
73/1, Vyborgskaya Embankment, St. Petersburg, Russia
E-mail: ppnovod@dux.ru

Abstract. The authors present the principles of formation, mathematical instrument and development technology related to a new class of control systems - control systems with knowledgebase, designed for making operational decisions, concerning control over potentially hazardous multi-unit technical complexes with high risk of operation in unpredictable emergency situations. The tasks, being solved are: operational diagnostics of technological processes under conditions of multiple impacts, failures, damages; evaluation of control responses to prevent emergency cascade development (the domino effect) and redamaging; evaluation of decisions for control object structure reconfiguration at arbitrary combinations of failures and component damages. The experience, related to development of this class system is analysed. The analysis concerns the tasks of control process automation in regard to nuclear-powered ships power systems and nuclear power stations power units, the possibility to use it in training programs and intelligible simulators for operational and leading staff self-training.

Keywords. Control systems with knowledgebase, expert real-time control systems, multi-unit technical complexes, unpredictable emergency situations, multiple failures and damages, safety, fail-safe operation, survivability, qualitative simulation, structural simulators, decision-making support.

The nature of the problem, research and development objectives. One of the main problems related to the theory of control is the problem of operational decisions' evaluation, concerning control over objects in non-standard emergency situations, at any unpredictable failures, combination of failures and damages, irregular emergency modes and emergencies, not considered by the project, i.e. which can not be considered by data processing and control algorithms. These algorithms are incorporated in traditional control systems (TF CS, TF CSC, systems of emergency protection (SEP), safety control systems (SCS), ACS TP). For a very long time, the existing problem restrained the realisation of a qualitative breakthrough in the development of automated control systems, related to large potentially hazardous technical complexes, normally, multi-unit and distributed on the territory complexes with a higher environmental risk of operation. These objects include nuclear-powered ships and vessels, nuclear power plants of different assignments and many other objects in the other fields of industry and production.

During the last 2-3 decades, many institutions of the Academy of Science, scientific and research institutes of the Ministry of Defence carried out prospecting, which characterised by inclination to use the ideology of an artificial intelligence system, in order to solve the control problems of such a class.

In order to solve the problem, related to the development of automated control systems for controlling nuclear-powered ships` survivability and provision of their nuclear safety, and under the technical guidance and direct participation of this study's authors there have been developed :
a) theory, mathematical instrument, simulators and methods, principles and means of controlling decisions` evaluation,
b) industrial technology in regard to the formation of intellectualised programming systems` complexes. First of all, these complexes are formed to develop the systems of information support (SIS) and to make decisions during the control of:
1. Functional complex of technical facilities (FCTF) related to the power provision of nuclear-powered ship or vessel (NPP, hydraulic system, lubrication system, equipment cooling system, others, considered integral),
2. Nuclear power station (NPS) power unit (PU) in non-standard emergency situations, in situations, which are not considered by the project. These situations include events, caused by an unpredictable environmental damaging effect - fires, floods, shock vibrations, explosions, real failures of systems with high energy parameters of working fluids and mediums, other premeditated and unpremeditated damaging impacts, such as acts of terrorism, enemy influence and so on.

The authors have developed a theory, mathematical instrument and information technologies of industrial development for a new class of control systems. The control systems have knowledgebase, lifting the "mysterious curtain" from the so- called artificial intellect systems. They provide a formation of systems, which under appropriate selection of ACS TP structural organisation, in particular their functional-topological, technical and other structures, impart to the whole system high levels of survivability, safety and efficiency. It has become possible through input of "mechanisms" of self-organisation and self-algorithmization, i.e. "homeostasis" mechanisms. These research efforts are unique and have no analogues in the world.

Research efforts` fields of application. Technical facilities complex control systems (TF CCS) of nuclear-powered ships, including :
- Automated control system of ship's survivability control, technical facilities control and provision of nuclear safety (incorporated in TF CCS).
- Intellectualised systems of decision-making information control, regarding the CCP personnel, and/or main power unit control panel staff during survivability control and provision of nuclear safety (re-equipment of current TF CCS).
- Automated control systems of technological processes (ACS TP) inside power units of :
a. Nuclear power stations,
b. Nuclear heat supply stations,
c. Nuclear-powered ships and vessels,
d. Float nuclear power stations.
- Training programs and simulator complexes for operational personnel training, which provide control over nuclear power units and installations.

- ACS TP in multi-unit technical complexes with high risk of operation, related to different fields of industry and production.

Ways of employment.
- Additional equipment of currently used ACS TP (TF CCS and TF CS) with intellectualised systems of information support,
- Additional equipment of NPS power units` control panels with diagnostical expert systems,
- Realisation of a high level for newly developed ACS TP (TF CCS and TF CS),
- Additional equipment of full-scale simulators with expert instructor-systems (technologists),
- Development of independent and ACS TP incorporated (TF CCS and TF CS) intelligible (intellectual) simulators.

Assignment of the systems being developed.
1. Drastic increase of fail-safe operation and safety levels in regard to operation of nuclear power units. It is achieved by solving monitoring and control tasks in non-standard emergency situations, unpredictable combinations of failures and equipment damages and extraordinary environmental impacts.
2. Considerable increase of ACS TP control processes automation level and grade, regarded as large man-machine complexes with appropriate reduction of operational personnel.
3. Prevention of faulty or ineffective actions (inactions) of operators in emergency situations, initialisation of correct and timely actions.
4. Prevention, when possible, of emergency situations` occurrence, considerable reduction of psychophysiological load on the operators.
5. Increase of operational personnel readiness for decision-making in non-standard situations, due to execution of pre-watch anti-emergency training with the use of an expert instructor-system (technologist).
6. Realisation of self-training modes on full-scale simulators, connected with an expert instructor-system (technologist).
7. Development of complicated training scenarios and the search for non-traditional (original) ways of control, in the existence of non-standard emergency situations.

The tasks being solved and functions being executed.
1. Operational diagnostics of technological processes in the power unit, including :
- early detection of faults in processes prior to the surpassing of parameters (state co-ordinates) over the precautionary and, moreover, emergency settings,
- revealing of non-standard emergency situations with multiple failures in the power unit equipment and control systems, which are not considered by the system to be built-in algorithms,
- the control of safety critical functions` hierarchical system,
- assessment of current values related to probable risk of a nuclear and radiation safety condition violation.
2. Analysis of events' succession and disturbances propagation in power unit sub-systems, its prediction with the assessment of final consequences at any unpredictable malfunction and failure combination, answering the question "why?", prevention of faulty actions made by the operator, initialisation of correct actions.
3. Evaluation of controlling influences (or recommendations) in order to promptly eliminate the imbalance, in regard to energy and substance flow between NPP components; its coming to an "equilibrium" condition with the maximum possible level of available capacity. Simultaneously, these controlling influences prevent the cascade character of damage development or pickups of local protection systems and retain the

technological processes' co-ordinates of state within permissible levels, under the conditions of integrity preservation and physical barriers of safety.
4. Evaluation of controlling influences (or recommendations) for reconfiguration of a power unit and related energy systems (own needs power supply (ONPS), hydraulic systems, equipment cooling systems, oil systems, etc.) (basically - solution of the task, concerning the problem of how to arrange the structure from operative components, including "preventive" reconfiguration) in compliance with the criteria of unconditional execution of nuclear safety conditions. Another criteria are : preservation of maximum possible levels of available or generated output at any unpredictable combination of malfunctioning and damages, which appear as a result of fires, floods, shock vibrations and other extraordinary impacts.

Along with this, the system executes the following functions :
5. Data input from standard control systems, dealing with current parameter values, values of equipment physical state, armature and automatic switches positioning, protection circuits' actuations, actions of operational personnel, also functions of RAM and data archives set up. There is a feasibility to connect systems of noise, vibration and other types of equipment diagnostics by physical methods.
6. Man-machine multi-window graphic and speech interfaces, which provide data mapping by practically all modern means of technological graphics, instant perception of current, future and recommended structural states of the power system by its operators, arrangement of effective dialogue between control personnel and the system.
7. System-integrated intelligible (intellectual) self-training simulator (for self-training) with expert instructor-system in order to develop in-depth understanding and a strengthening of knowledge in physics and process dynamics. These processes proceed in a multi-unit technical complex, revealing in the course of training genetic relations of units and power unit sub-systems. The training is aimed at receiving skills of quick perception when evaluating the situation and making operational decisions for controlling the installation in any unpredictable emergency situation, during survivability control and provision of nuclear and radiation safety.

Specific system features and the main "know-how". The specific system features are characterised by the use of :
a) Technology of knowledgebase with automatic generation of large technological and diagnostical principles (knowledgebase) with the use of a special program. The technology is based on the processing of purposeful data of experiments, with verified mathematical simulators in regard to controlled physical and technological processes' dynamics, which are occurring in a power unit. Simultaneously, the theory and modern methods of training, concerning trial data processing are being used. They are used for determining objective conformities and genetic relations.
b) The technologies of so called qualitative simulation, regarding power unit dynamics, with the use of project and service documentation, providing automation of operator' "thinking simulator" in an accelerated time scale. Qualitative simulation is widely used in real time expert systems.
c) Technologies of structural simulation (logical equations systems, systems of differential Boolean equations, graphodynamics simulators, theories of character graphs, automatic simulators), which provide forecasting and determination of final after-effects caused by any (not studied beforehand) combinations of failures and damages. It is necessary to consider the feasibility of its cascade development and work out control decisions to reconfigure the structure of the multiplicity of all interconnected TF complexes, also in an accelerated time scale, with a reasoning system connection and impossibility to meet many criteria.

d) New systematic (nonconventional and nonuniversal) numerical methods of differential equation integration, (including so-called "tough" systems) which enable the achievement of a considerable increase of simulation efficiency in an essentially accelerated time scale of controlled physical and technological processes. Besides this, it carries out a smooth time scaling, when it is necessary to slow down or to accelerate processes in the course of intellectual training, and during analysis of training sessions.
e) Technologies of expert systems for the realisation of an intelligible simulator with an expert technologist-system in order to carry out anti-emergency pre-watch intellectual training.
f) Advanced technologies for programming and industrial standards, supported by leading producers of computers and common (basic) software, and also modern instrumental means of design automation, related to large high level applied program systems.

Technology of programming. The adopted technology of programming uses :
- network multi-task, multi-user real time operational systems type UNIX, including QNX meeting the requirements of POSIX standard,
- programming systems, which meet the requirements of C ANCI standard,
- graphic systems of X-Windows standard,
- means of graphic annex evaluation of Motif 1.1 / Xlib Rel 5.0 standard,
- network logical record of data exchange TCP/IP between computers and other networks type FDDI, Ethernet, Token Ring and so on,
- "Operator" instrumental system (study made by Institute of Control Sciences, Russian Academy of Science) for designing control systems type ACS TP with functions of operators` intellectual support, real time expert systems, training systems and simulator complexes.

Employment of the above mentioned standards provides:
- system functioning as a programming product of high level both on the platform of IBM PC compatible computers, (IBM PC 486, Pentium) and on the other (optional) computer platforms, e.g. HP 9000 700/800 series, HP 700I, HP X-terminals, VAX/VMS, DECstation, RS 6000, SUN SPARC, Selicon Graphics with UNIX-like operational systems and common software, meeting the requirements of POSIX, C ANSI, X-Windows, Motif 1.1 /Xlib Rel 5.0, TCP/IP standards.
- feasibility of easy expansion and development in the course of operation due to employment of open module architecture and widely spread industrial standards,
- feasibility of joint operation with a full-scale simulator,
- feasibility of operation in a common calculation network, which provides data input, received from the systems of power unit (s) centralised control.

Possible special occasions of expert systems realisation. Improved mathematical means and information technologies enable us within short time periods to develop both complex intellectualised systems of operators` information support and a high level for newly developed ACS TP of NPS power units and nuclear-powered ships and vessels with intelligible simulators, which are integrated in control systems, and also special systems to solve separate intellectual tasks of diagnostics and control, or its some multiplicity, in particular :
1. Expert systems of anomalies in power units' technological processes. It is made on the base of data, received from the systems of centralised control, with data output on safety panels; specific display stations (monitors) for operators and station shift chief officer (SSCO) (chief executive).
2. Expert systems of monitoring and probable risk forecasting related to violations of safety conditions, connected with information about the equipment's current state of

operation, unit operation mode, technological parameters and possible reconfigurations of NPS system structure, which are presupposed by SSCO.
3. The system for information collection and processing, concerning equipment running time and failures with the assessment of technical reliability real values of some systems and equipment (fail-safe operation, fitness for repair, durability), with the analysis of systematic values related to reliability, efficiency and unit operation safety, including assessment and forecasting of systems, ACS TP equipment and devices' remaining resources.
4. Expert systems of operational diagnostics, related to technological processes and working out recommendations for control service personnel, how to co-ordinate the control over multiplicity of NPS systems and units, in order to eliminate the appearance of an imbalance of energy and/or substance flows between units (systems). The purpose is to transfer the power unit for the maximum possible level of available capacity without violation of minimum functions of safety.
5. Expert systems for the recommendations evaluation, related to NPS reconfiguration at any unpredictable combinations of failures and damages, caused by extraordinary environmental impacts (fires, floods and so on), also possible acts of terrorism).
6. Training systems with an intelligible simulator for pre-watch anti-emergency training of operators and SSCO at non-standard emergency situations with different degrees of severity, failures, which are not considered by the project and extraordinary environmental impacts.
7. Expert systems incorporated in full-scale simulators and designed for supporting the process of training and training for decision-making in unplanned emergency situations, working out complicated training scenarios with preliminary "interpretation" of "untraditional" decisions on control, which are suggested by the expert system and "self-investigation" of incidents by the command personnel on the basis of emergency processes, fixed by a control system.

Experience of using developed information technology.
1. Development of intellectualised system for the information support of control service personnel, working with nuclear power installation. This system is incorporated in the pilot model of automated control systems for controlling nuclear submarine survivability.
2. Development of the system for continuous monitoring, related to nuclear-powered ships' power plants' nuclear safety conditions' risk of violation (demonstration prototype).
3. Development of simulators with an expert instructor-system for steam-gas electric power station operators' self-training. Technology transfer. Ansaldo S.P.A. (Genoa, Italy), (1993).
4. Creation of real time expert systems for operational diagnostics of technological processes in steam-gas electric power stations. Technology transfer Ansaldo S.P.A. (Genoa, Italy), (1994).
5. Development of ACS TP "RACURS" high level of PIK reactor installation - research reactor with 200 MW capacity (expert system of operational diagnostics related to technological processes and operator support). Institute of nuclear physics (Gatchina, Russia), (1995). "Turnkey" design.
6. Development of ACS TP high level, related to VVER-1000 reactor installation (demonstration prototype) safety stand. ENITS (Electrogorsk, Russia), (1994).,"Turnkey"design, study of the Institute of Control Sciences, Russian Academy of Science.
7. Development of advanced ACS TP high level for NPS with VVER-1000 reactor installation (demonstration prototype). ENITS (Electrogorsk, Russia), (1994)., study of Institute of Control Sciences, Russian Academy of Science.

8. Development of unified software for a new generation of ACS TP high level, This software is designed for NPS with VVER-1000 reactor installations. BNIIA (Moscow, Russia), (1995). Joint study with the Institute of Control Sciences, Russian Academy of science.

Reference

1. Summary scientific and technical report about scientific and research work "Research and development of an automated control system for controlling nuclear submarine survivability with the use of the concept of expert systems". Authors : Simakov I.P. (Scientific instructor and Chief Designer), Zuenkov M.A., Merzliakov V.A., Ponomarev A.A., others. - inventorial no. 101/2-123 - SPA "Aurora", St. Petersburg, (1992). (in Russian).
2. Zuenkov M.A., Simakov I.P., Skorokhodov D.A. Systems of intellectual support for NPP operators. - In the Collection, The First International Conference on naval intellectual technologies. - MORINTEH-95, St. Petersburg, (1995). (in Russian).
3. Simakov I.P. Control systems with knowledgebase: mathematics, technologies for the creation and use on nuclear-powered ships and NPS. - In Jubilee scientific and technical Collection of SPA "Aurora", St.Petersburg, (1995). (in Russian).
4. Naumov M.V., Simakov I.P., Skorokhodov D.A. Automation of naval ships' survivability control. - In the Collection of reports on the International Conference "Navy and ship-building in modern conditions" (NSN), commemorating the 300-th Anniversary of the Russian Navy, St.Petersburg, (1996). (in Russian).

THE INFORMATION ESTIMATIONS OF SEARCH TIME THE SURVIVOR AT SEA

Anatoly E.Sazonov, Ivan P.Boukanov
Admiral Makarov State Maritime Academy
Kosaya linia, 15-a, St.-Petersburg, Russia

Abstract. In the paper, a information approach has been developed to obtain evaluations (of the upper and the lower bounds) of the average time of spotting those distressed on the sea. For the special case of the fixed target search by a single search unit, a theorem has been formulated and proved, which connect the average search time with the differential entropy of the probability density function of the searched object's location. On the basis of the entropy equality condition, a notion of an entropy area has been introduced. It has been shown that a search on the entropy area requires maximum average time, if the distribution of the object's location is taken uniform on it.

Keywords. Entropy, survivor, search and rescue, search unit.

Introduction. The practice of improvement of technical means of execution of search and rescue operations on the sea, in particular, radio engineering ones, has required to make a mathematical formalization of some elements of the search procedure. In this report, a theoretical and information approach has been used, which has allowed to obtain evaluations of average search time of those distressed on the sea (the target).

Let the target be fixed. The probability density function of the target location on the plane has the form $p(x,y)$ and is unimodal. The search is carried out by a single search unit. It has taken that the search unit moves from the initial search point (x_0,y_0) with constant linear velocity v along a certain curve γ, which is called a trajectory. The search unit being at the point (x,y) makes search efforts so that if the target is located on the line passing through the point (x,y) of the trajectory and the curvature centre of the curve at this point at the distance not more than D at both sides of the trajectory, then it will be spotted with the unity probability. So the spotting region on the plane, which is formed by the movement of the search unit along the trajectory γ, is a strip S with the width equal to $2D$. It has been taken that the strip S performs packing of the plane, on which the search operation is being carried out.

Let $\gamma(x_i,y_i,t_i)$ mean that at the moment t_i the search unit is located at the point (x_i,y_i). On the basis of the Kadane [1] rule, it can be noted that the trajectory γ should be such that, with the search unit movement for any two moments of time t_i and t_k, ($t_k > t_i$), the following condition be fulfilled: $p(x_i,y_i). \geq p(x_k,y_k)$. In this case, the initial search point coincides with the maximum value of the unimodal density $p(x,y)$. Let us call the trajectory of the search unit movement so determined A-trajectory. It is required to evaluate the average time T necessary for spotting the target.

Theorem. Let the search unit movement take place along an A-trajectory, and the probability density function of the target spotting be $p(t)$. Then the average time T of spotting the target will satisfy the inequality

$$\frac{1}{e}e^{H_0 \ln 2} \le T \le \frac{1}{2}e^{H_0 \ln 2}, \tag{1}$$

where $H_o = -\int\limits_t p(t)\log_2 p(t)\,d\,t$ – the differential entropy.

Proof. Along the known curve of the A-trajectory, a functional transformation of the probability density function *p(x,y)* into the probability density function *p(l)* of the target location is being performed at the length *l* of the trajectory. As *l=vt*, then there is a single-valued transformation *p(l)* into the probability density function *p(t)* of the target spotting as a function of the search time. Then the average time *T* of the search

$$T = \int\limits_t tp(t)\,d\,t. \tag{2}$$

Let us find the probability density function *p(t)*, which provides the minimal average search time value. To do this, let us solve a variation problem: find a function *p(t)*, which provides the minimum of the functional $T = \int\limits_0^\infty tp(t)dt$ under the extra conditions $\int\limits_0^\infty p(t)\,d\,t = 1;\ -\int\limits_0^\infty p(t)\log_2 p(t)\,d\,t = H_0$.

The solution of the corresponding differential Euler-Lagrange equation $\frac{d}{d\,p(t)}\left[tp(t) + \lambda_1 p(t) - \lambda_2 p(t)\log_2 p(t)\right] = 0$ leads to the power distribution law density function $p(t) = a\,e^{-at}, a = e^{1-H_0 \ln 2}$. The direct calculation according to formula (2) gives the average time value $T = \frac{1}{a} = \frac{1}{e}e^{H_0 \ln 2}$, which proves the left-hand side of inequality (1).
In order to prove the right part of inequality (1), let us use the fact that the probability density function, which provides the maximal average time value under fixed entropy, is equivalent to the density function providing the minimal entropy value under fixed average time. So, a variation problem is being formulated to minimize the functional

$$-\int\limits_0^b p(t)\log_2 p(t)\,d\,t \tag{3}$$

under the extra conditions

$$\int\limits_0^b p(t)\,d\,t = 1; \quad \int\limits_0^b tp(t)\,d\,t = T. \tag{4}$$

According to the A-trajectory definition, a variation problem solution should be in the class of non-increasing probability density functions. Using the known inequality $\ln x \le (x-1)$ under $x>0$, let us reduce the variation problem (3) to a problem of providing maximum of the functional $\int\limits_0^b p^2(t)\,d\,t$ under the previous extra conditions (4). The solution of the corresponding differential Euler-Lagrange equation $\frac{d}{d\,p(t)}\left[p^2(t) + \lambda_1 p(t) + \lambda_2 tp(t)\right] = 0$ has the form

$$p(t) = \frac{4}{b} - \frac{6}{b^2}T_m + \left(-\frac{6}{b^2} + \frac{12}{b^3}T_m\right)t, \quad \left(-\frac{6}{b^2} + \frac{12}{b^3}T_m\right) \le 0.$$

Taking $b = 2T_m$, we have the uniform distribution law function

$$p(t) = \begin{cases} \dfrac{1}{2T_m}, & 0 \le t \le 2T_m; \\ 0, & t < 0, t > 2T_m \end{cases}, \tag{5}$$

for which the average time value $T_m = \dfrac{1}{2} e^{H_0 \ln 2}$ is derived from the known equation for entropy $H_0 = \log_2 2T_m$. So the right-hand side of inequality (1) has been proved, and, consequently, the theorem has also been proved. In Table 1, the calculations of the average time T of spotting the target for various distribution functions $p(t)$ under the constant entropy value H_0..

Table 1 The average time of spotting the target for various distribution functions

N	$p(t)$	T	H_0	T/T_o
1	$\lambda\, e^{-\lambda t}$	$T_0 = \dfrac{1}{\lambda}$	$\log_2 \dfrac{e}{\lambda}$	1.000
2	$\begin{cases} 1/\Delta, & 0 \le t \le \Delta; \\ 0, & t < 0, t > \Delta; \end{cases}$	$\dfrac{1}{\Delta}$	$\log_2 \Delta$	1.365
3	$\dfrac{\mu^2}{2} e^{-\mu\sqrt{t}}$	$\dfrac{6}{\mu^2}$	$\log_2 \dfrac{2e^2}{\mu^2}$	1.098
4	$\dfrac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}}$	$\sigma\sqrt{\dfrac{\pi}{2}}$	$\log_2 \dfrac{\sigma}{\sqrt{2}} e^{\frac{c}{2}+1}$ $c = 0.5772\ldots$	1.325
5	$\dfrac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{t^2}{2\sigma^2}}$	$\sigma\sqrt{\dfrac{2}{\pi}}$	$\log_2 \sigma\sqrt{\dfrac{\pi e}{2}}$	1.052
6	$\dfrac{4t^2}{\sqrt{\pi}(2\sigma^2)^{\frac{3}{2}}} e^{-\frac{t^2}{2\sigma^2}}$	$\sigma\sqrt{\dfrac{8}{\pi}}$	$\dfrac{1}{2}\log_2 \dfrac{2\pi\sigma^2}{e} + c \log_2 e$	1.016
7	$\begin{cases} \left(\dfrac{4}{b} - \dfrac{6}{b^2}T\right) + \left(-\dfrac{6}{b^2} + \dfrac{12}{b^3}T\right) t, & 0 \le t \le b; \\ 0, & t < 0, t > b; \end{cases}$ $b = 3T$	$\dfrac{b}{3}$	$\log_2 \dfrac{\sqrt{e}}{2} b$	1.102

The consequence from the theorem, which is worth to be noted, is the fact that the average search time on the trajectory, along which the probability density function of the target location is distributed according to the uniform law (5), has the maximal value. Irrespective of concrete form of the A-trajectory curve, the search area equals to

$$S_e = 2vD2^{H_0},\qquad (6)$$

which, in order to be definite, we call an entropy one. As the probability density function *p(t)* and the probability density function *p(ρ)* of the radius-vector ρ under the formulation of *p(x,y)* in the polar coordinate system are connected with each other by the relation $p(t) = p(\rho)\left|J\left(\frac{\rho}{t}\right)\right|$, where $\left|J\left(\frac{\rho}{t}\right)\right|$ is the modulus of the transformation Jacobian depending upon the form of the curve γ of the trajectory, then [2] $H_0 = H(\rho) - \int p(\rho)\log_2\left|J\left(\frac{\rho}{t}\right)\right| \mathrm{d}\,\rho$, where $H(\rho) = -\int\limits_t p(\rho)\log_2 p(\rho)\,\mathrm{d}\,\rho$. Consequently, the expression (6) will have the form

$$S_e = 2vD2^{H(\rho)-\int p(\rho)\log_2\left|J\left(\frac{\rho}{t}\right)\right|\mathrm{d}\rho}.\qquad (7)$$

Let us study, in more details, the class of round distributions *p(x,y)*. For the class of distributions mentioned, an A-trajectory curve is an Archimedes spiral with the step equal to *2D* and the centre in the initial search point. It can be shown that the distance *l*, which has been passed by the search unit under the displacement from the initial search point (*t*=0) by the moment of time *t*, is the corresponding arc length of the Archimedes spiral and equals to $l = vt = \frac{\pi\rho^2}{2D}$. In this case, $\rho = \sqrt{\frac{2vDt}{\pi}}$, and the modulus of the transformation Jacobian will have the form

$$\left|J\left(\frac{\rho}{t}\right)\right| = \frac{1}{2}\sqrt{\frac{2vD}{\pi t}}.\qquad (8)$$

By inserting expression (8) into formula (7), we derive $S_e = \pi 2^{H(\rho)+\int p(\rho)\log_2 \rho\,\mathrm{d}\,\rho+1}$. In Table 2, the results of the entropy area S_e calculations have been shown for several round distributions *p(x,y)*.

Consequence. Let *p(x,y)* be a density function of the target location on the plane with the differential entropy equal to $H(x,y) = -\iint p(x,y)\log_2 p(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y$. Then for the entropy area, the following equality is valid

$$S_e = 2^{H(x,y)}.\qquad (9)$$

Without presenting a strict proof of the consequence, we shall only note that the validity of formula (9) is proved by of the differential entropy properties [2]. Really, the presentation of the search region of arbitrary form as a strip *S* will lead to a re-distribution of the density function *p(x,y)* under preserving the set of differential probabilities. The transformation in question does not change the scale of variables, and consequently, the magnitude of the differential entropy retains. As a comparative example, in Table 2 some entropy *H(x,y)* formulae have been presented for several distribution functions *p(x,y)*.

Conclusions. The practical significance of the theorem and the consequence presented herein is that the evaluation of the upper bound of the average time of spotting the target located on the plane with an arbitrary density function *p(x,y)* can be derived by arranging a search on the entropy area where the probability density function of the target location is taken to be uniform. On the basis of the differential entropy property being already mentioned, the form of the search region and, consequently, the form of the A-trajectory curve can be selected as arbitrary and different from those used under the search by the density function *p(x,y)*. In particular, in order to study the influence of the radio engineering provision upon the efficiency of search and rescue operations on the sea, it is expedient to take a straight line as an A-trajectory curve.

Table 2 The results of the entropy area calculations for several round distributions

N	*p(x,y)*	*H(x,y)*	H_0	S_e
1	$\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$	$\log_2 2\pi\sigma^2 e$	$\log_2 \frac{\pi\sigma^2 e}{vD}$	$2\pi\sigma^2 e$
2	$\frac{\lambda^2}{2\pi} e^{-\lambda\sqrt{x^2+y^2}}$	$\log_2 \frac{2\pi e^2}{\lambda^2}$	$\log_2 \frac{\pi e^2}{\lambda^2 vD}$	$\frac{2\pi e^2}{\lambda^2}$
3	$\begin{cases} \frac{3}{\pi R^3}\left(R-\sqrt{x^2+y^2}\right), & \sqrt{x^2+y^2} \le R; \\ 0, & \sqrt{x^2+y^2} > R; \end{cases}$	$\log_2 \frac{1}{3} e^{\frac{5}{6}} \pi R^2$	$\log_2 \frac{e^{\frac{5}{6}} \pi R^2}{6vD}$	$\frac{1}{3} e^{\frac{5}{6}} \pi R^2$
4	$\begin{cases} \frac{1}{\pi R^2}, & x^2+y^2 \le R^2; \\ 0, & x^2+y^2 > R^2; \end{cases}$	$\log_2 \pi R^2$	$\log_2 \frac{\pi R^2}{2vD}$	πR^2

References

1.J.B.Kadane (1968) Discrete Search and the Neyman-Peearson Lemma. Journal of Mathematical Analysis and Applications, v.22, N1, pp.156-171.

2.C.E.Shannon (1948) A mathematical theory of communication. Bell System Techn. J., v.27, N3, pp.379-423.

SOUNDNESS OF AUTOMATED SYSTEMS' STRUCTURE RELATED TO SUBMARINE COMBAT CONTROL

Zavalishin A.A., Kashtankin V.V., Luschik V.L., Fedorov L.E.
"Aurora" Scientific-Production Association of
Ministry of Defence Industry of Russian Federation
15, Karbysheva Street, 194021, St. Petersburg, Russia
Telephone: +7(812)316-82-95

Abstract. The authors consider the problems concerning the synthesis of ASCC as an integral part (subsystem) of integrated SM ACS. At the same time, they are analysing the main problems of its creation, at the early stages of design works. The authors have formulated requirements for ASCC information, program and technical provision and present alternative methods for its realisation.

Keywords. submarine (SM), control system, combat control.

The automated system of combat control (ASCC), is one of the most important elements of submarine ACS, and considerably determines the realisation of ship combat potential, the success of submarine combat operations. ASCC is a complicated system, thus during its grounding and development it is necessary to follow the principles of a systematic approach, taking into consideration the whole variety of tactical, technical and organisational factors, which determine the assignment, composition, structure and employment of this system.

Besides traditional common problems concerning the creation of a complicated system, the development of ASCC has problems directly related to the soundness of the system's structure. They include:

- complication of criteria selection necessary for the whole system and its main sub-systems' soundness,
- difficulty in employment of strict mathematical methods for system structure optimisation, the necessity to provide compatibility of all sub-systems and ASCC as a whole with the control sub-system of submarine technical facilities, environmental (interacting) complexes and systems (part of these systems was developed without considering the requirements for compatibility with the developing ASCC),
- considerable indefiniteness of initial data at the early stages of design, first of all, it implies the composition of specific mathematical support (SMS) and the qualitative characteristics of its elements, characteristics of information exchange, between ASCC sub-systems and with external systems.

At the same time, the general appearance of the system is shaped at the early stages of development, which later (in the process of technical design and producing of pilot samples) is hard to correct.

The solution of these problems is achieved through the realisation of a complete complex of organisational and technical measures and specifically developed methodology for grounding the system structure at the early stages of its development.

The basis of methodology for grounding ASCC structure is a systematic approach, which in relation to the problem being considered, can be determined as a realisation of the following main aspects:

- provision of ASCC conformity with the organisational and technical structure of a submarine control system, its armament, weaponry and technical facilities,
- provision of ASCC functioning as an independent element and simultaneously, as the sub-system of the hierarchy of submarine ASC, possessing information compatibility with corresponding systems and complexes,
- consideration of the whole system, interconnection of all its structural and functional sub-systems,
- consideration of the system in connection with environmental factors, which influence its functioning organisation and efficiency, and taking into consideration the feasibility of enemy impact on the submarine, military and geographical conditions, environmental influence, specific features of its introduction and service, feasibilities of research efforts (design and production), related to system's software and hardware within a targeted period,
- system design in compliance with targets (the tree of targets), for its creation, i.e. according to the scheme from top to bottom with successive decomposition of realised requirements between sub-systems,
- revealing of system functioning quality dependence from the features of its sub-systems,
- finding an optimal (reasonable) version of system formation, providing the best possible values of its efficiency in presupposed conditions of its functioning, and under practical existing limitations of allocated resources and limitations of development and introduction periods.

Realisation of the suggested approach is linked with the necessity to carry out systematic analysis, including multiple (interational) assessment of efficiency, both of the system and its main sub-systems.

The process of systematic analysis is a completely creative action, which can not be absolutely formalised. Nevertheless, the experts, which possess experience in the field of such works, have developed recommendations for its accomplishment and appropriate methodical support. The shaping of the requirements for system information, program and technical support can be considered one of the main problems for ASCC development.

ASCC information support (IS) includes the integrity of input and output data, classification and a standardised basis, and realised solutions, related to information volume, arrangement and existing forms in the system during its operation. The composition, structure and content of the information support must enable the development and effective use of functional software (SW) of ASCC sub-systems and separate workstations. The structure of information support and the arrangement of its components in ASCC hardware must provide an effective data exchange within system.

Information support must correspond with the users' ideas (interests), reflected in the task assignment for functional software, and in combination with the preservation of maximum independence from algorithms and functional SW programs. It must be provided an adaptation of information support and its separate elements within limits of adopted organisational and hardware solutions, in relation to accessible changes of subject sphere and information demands of functional SW, without modification of data processing algorithms. It must provide information and linguistic compatibility with ASCC interacting systems (complexes), between sub-systems and ASCC workstations. This compatibility is expressed in the unambiguity of description, interpretation and identification (encoding) of the data, which are being used in the data exchange.

System software is a component of a hardware-software platform and includes programming facilities for the organisation of the computer process - operating system (OS), compatible elements of network software, programming facilities of functional control.

OS is one of the most important computer (computing complex) elements, which to a large extent determines the output and capabilities of the whole system. The main requirements, which should be taken into account in OS selections are as follows:
- provision of computer real-time operation,
- provision of multi-task mode execution with support of parallel processes,
- compressibility of OS core (< 2 MB),
- high level of OS functioning reliability,
- availability of a graphic shell and support of mülti-window mode of operation,
- support of standard networks,
- support of wide grade drive units,
- compatibility with CSCA, which is used in ASCC,
- the cost and terms of delivery and acceptance.

Functional software (FSW) is designed for control processes automation, and is aimed at data processing and mapping, evaluation of recommendations and data, simulation of situation in the interests of the submarine assigned tasks. FSW must provide "through" realisation of the submarine automated control processes in compliance with the whole ASCC and its sub-systems' functions. It must also meet the technical and economical requirements, put forward to ASCC. The FSW structure must include all necessary software elements, provide the hierarchy of its ability and solution, interaction, arrangement on ASCC hardware, with considering execution of system functions. It is necessary to envisage automatic realisation of program components' determined aggregate, which provide ASCC accomplishment of all necessary procedures in a real tactical situation.

ASCC hardware is the means of computing devices, which provide realisation of control functions on different hierarchical levels of control. The composition of hardware can be divided (but enough conditionally) into two levels: "upper level"- means of processing and control, information storing and mapping, and "lower level"- devices of initial processing, control, management and connection with transducers and actuating devices.

The analysis of tasks, which are being solved by the means of control automation, shows that its number is constantly increasing. The main reasons for this and higher requirements, which are put forward to them are as follows:
- the necessity to raise the level of automation,
- the demand in services, which are required by new types of armament, ways of its employment and protection against enemy's weaponry,
- improvement of "external" systems and complexes, which interact with ASCC, their tactical and technical parameters.

In order to provide the solution of the increasing number of FSW tasks in advanced systems, in time and with needed accuracy, the hardware of "upper level", as it comes from the forecasts, will require the acceleration of data processing speed by several orders. The increase of "lower level" hardware parameters must be achieved through transfer of initial processing functions on the level of transducers and actuation devices and a wider unification of information sources.

SHIP INTEGRATED CONTROL SYSTEM. WAYS AND METHODS OF TECHNICAL APPLICATION

Fedorov L.E., Kashtankin V.V., Luschik V.L., Vorobjev M.N.
"Aurora" Scientific-Production Association of
Ministry of Defence Industry of Russian Federation
15, Karbysheva Street, 194021, St. Petersburg, Russia
Telephone: +7(812)316-82-95

Abstract. The authors have studied the basis of ship integrated ASC arrangement. They have analysed the situation related to this problem, in home and foreign shipbuilding industry, have determined functions and structure of this system, have discussed approaches for assessing efficiency of similar systems and have considered questions of organisational support, dealing with its creation and operation.

Keywords. submarine (SM), control system, combat control.

The development of armaments and weaponry, improvement of acquisition systems, systems of counteraction and destruction lead to the increase of the information volume, which has to be processed on board the ship in a real time scale. The data input is expanding both from inboard facilities and systems (CAS, radar, navigation complex, means of radio and radiotechnical reconnaissance and others) and external sources. The technical rigging of the entire ship has been considerably improved. Because of these reasons, it is necessary to optimise ship control; its combat capabilities depend directly on the ability of its system of combat control to accept, map, store, process and analyse information and proceed on the basis of received results.

The creation and acceptance of ASC must provide maximum employment of ship combat potential due to a complex analysis of situation, elaboration of recommendations for optimum (reasonable) actions in exact situations and its execution with required efficiency and accuracy. It can be achieved by dynamically arranging combat control outlines, which correspond to real situations.

At the same time, in the process of ACS development it is necessary to minimise equipment physical dimensions, input of electric power, the price of the system's design, production and operation.

The given task assignment, in principle, corresponds with common statements of systematic approach towards the development of such complicated systems as ship ACS. It also corresponds with methodology for its grounding and designing.

The achievement of the above-mentioned tasks must be accomplished on the basis of priority adoption and realisation of the following principles, related to system development and arrangement.

1. Arrangement of systems in compliance with the C3I concept (Command, Control, Communication and Intelligence) in relation to an automated object - combat ship.
2. Dynamic formation of all ASC necessary control elements, depending on a real situation and in accordance with commands, which are issued by a higher level of a hierarchical control system - command control sub-system.

3. Provision (with maximum available level) of the system's "openness", i.e. its correspondence to the co-ordination simulator of open systems - international system of standards ISO/OSI.
4. Provision of co-ordination within all ACS sub-systems and with "external" systems on the basis of unified or several interconnected local networks for data communication (if possible, also audio and video information).
5. Realisation of new information technologies (developed user's interface, multimedia - technology, methods of artificial intelligence).
6. Provision of required reliability and survivability of the system. For that purpose, it is necessary to use highly reliable technical facilities, and provide the realisation of redundancy (including functional) in regard to equipment and a developed information protection system, during its communication and processing.
7. Employment of modern hardware and software related to digital computing devices: high output processors, controllers, high capacity, high-speed memory, graphic systems of data mapping, multi-objective real-time operating systems and compatible network software. At the same time, it is necessary to minimise the design period and preserve the system's competitiveness at the beginning of its operation. For this purpose, modern software and hardware must be used as one or several interconnected hardware-software platforms. It is named the aggregate of hardware and software systems.
8. ACS integration of shipborne weapons, armament and technical facilities' systems.

The processes of integration are one of the most evident tendencies in development of advanced naval automated control systems. It is no coincidence, that the names of similar systems, designed in different countries include the word "integrate" - unified, common. Besides, the problem of submarine ACS integration is the most acute. The assessment of integration process' depth can be accomplished by the use of two parameters - degree and level of integration.

The degree of ACS integration - is a qualitative parameter of the system, which answers the question, which types of weaponry, armament and technical facilities of the ship are integrated with information. Information integration means the presence of links, which carry out data input-output, without participation or by only one operator (in other words - without voice communication between operators).

This parameter can be converted into a quantitative one, in case of submarine decomposition into separate elements - systems of weaponry, armament, technical facilities (echo-ranging and direct-listening hydroacoustic means of scanning, means of hydroacoustic signals detection, means of radar signals detection, active radar equipment, means of communication, means for storing and acquisition of SB co-ordinates, torpedo launchers, torpedo defence launchers, missile launchers, shipborne systems and facilities, main propulsion plant). The number of decomposition elements can differ and reach 20 -30 units, depending on SM class and design. In completely integrated ACS (grade of integration 100%), all above-mentioned elements of SM are system integrated. In the case, when some elements are not integrated in the system, the grade of integration reduces (for example, if SM is decomposed into 15 elements and only 9 elements are integrated in ACS, the grade of integration is 60%).
Analysis of development and ACS, introduced into the inventory of the Navy submarines of developed capitalist countries shows, that the grade of integration of these systems is from 50% to 100%.

The grade of ACS integration is a qualitative parameter of an automated control system, which answers the question, on which level of data processing was the integration of weaponry systems made. We can select three levels of ACS integration:

1. Integration on the level of completed (final) information, e.g. information, which is shaped for direct employment. In this case, ACS integrates completed complexes, stations, etc.
2. Integration on the level of preliminary information, i.e. information, which needs further processing before its employment. In this case, ACS integrates complex and stations control systems, etc.
3. Integration on the level of initial information, i.e. information, which enters ACS from aerials and transducers in the form of signals. In this case, ACS integrates initial transducers of information.

The majority of foreign ACS installed on SM are characterised by the first level of integration; the second level of integration is more rare, the third one even rarer.

Considerable changes to the shipborne ACS outline need a different approach towards their creation and development. First of all it refers to changing the existing arrangement: ship's customer - ship's designer and customers of separate complexes - developers of individual complexes and systems.

The creation of ACS (from an arranging development and successive deliveries point of view) is suggested to be accomplished as a whole system: one designer, who unifies the technical assignment, one chief developer, who carries out development, production, testing and system escort (if necessary, an appropriate contractor can be involved).

The ship designer must be a customer for this system. The leading developer - scientific and production association "Aurora ".

FEASIBILITY OF THE FUNCTIONING MODES OF THE COMPLEX TECHNICAL SYSTEMS.

Alexandr N. Novodjilov
Naval Academy named after Kuznetsov N.G., St. Petersburg
73/1, Vyborgskaya Embankment, St. Petersburg, Russia
E-mail: ppnovod@dux.ru

Abstract: .In the given article the construction fundamentals of a system of information support of the operator in sphere of the feasibility of the functioning modes are considered. The functioning mode is understood as a set of states and connections of system elements, which fulfils a defined functions. Use of such system increase safety of systems maintenance and reduce number of the attendants errors. It is reached by the calculations the optimal modes from a non-failure point of view, which realise functions, determined for a system.

Keywords: Safety, disturbing factors, simulation modelling, pre-set system functions, state of the element.

Nowadays the task to provide the exploitation safety for the nuclear plants becomes very vital. Generally the safety is understood as the aggregation of the technical system to resist the influence of the disturbing factors. The main disturbing factors in nuclear engineering are failures and malfunctions of the equipment, wrong actions of the personnel in ordinary and emergency situations as well as in navigation's accidents (in the Navy).

The most effective way to solve the problem of ensuring the safety is to build the complex control systems. This complex system is to define the mostly safe mode of nuclear power plant operation, while fulfilling the assigned functions in the exact situation, and carry it out minimising the possibility of faulty actions of the personnel. In this case, the mode of operation is understood as set of condition relations of the system elements that carry out their pre-set function (or functions). The mode of safe operation is also a set of conditions and relations between the elements that provides the most effective resistance to the disturbing factors.

To solve the problem of feasibility of the safe functioning modes of the complex technical systems operation we offer the method of feasibility, based on the simulation modelling. The method proposed offers the following:

1. Receiving the information on the technical means status from the existing monitoring systems and utilising the complex diagnosing algorithms to increase the degree of data reliability;
2. Evaluation and forecast for the future status of the technical means in a form of calculation of the carried out functions of the systems (subsystems) at various moments;
3. Feasibility of the safe functioning modes of the technical means on the basis of the pre-set system functions.

Receiving the information on the technical systems status is provided as follow: the signals, coming from different sensors or from the operator's control console, transferred to the diagnostic and later to the simulating model by means of which the forecast is calculated. The evaluating results of the current status and the forecast are input into the database and are used in future to evaluate the condition and to issue the recommendations on the functioning modes.

The simulation model is understood as set of the system elements (subsystems) and relations between them in their interaction in space and time during the realisation of some process. By use of such model the algorithm reproduces process of functioning of system during the time, and elementary process's compound phenomena are simulated with preservation of its logic structure and execution sequence. It allows, using the initial data (information, received from sensors, diagnostic system and from operator) to get the information about the technical means status in determined moment of time, giving an opportunity to determine a technical status of system.

For construction of such model the system is subdivided into subsystems (elements). The degree of detailing elaboration is determined by problems of modelling, namely, the number of system characteristics and necessary accuracy, which must be obtained during modelling. Subsystems (the elements) are united by logic model. For them conditions of functioning are written down, various types of connections with other elements of system are defined.

Each element (the subsystem) depending on a organisation, peculiarity of processes proceeding in it, has describing continuously - determined, discretely - determined, discretely - stochastic, continuously - stochastic model or their combination. The accuracy of the elements description is defined by problems of modelling, exact mathematical models (the differential equations and etc.), regressive models, simply mathematical expectations of some event can be used.

For the element of the simulation model in a common case is possible to define subsets of variables, which describes functioning process, as follows:

1. External.
   1.1. Set of the control influence. It is such influence, which can change the state of the element from the point of view of its functioning. Can be applied only to active elements, which are the channels of management.

$$x_i \subseteq X, \quad i=1,n_x; \qquad (1)$$

   1.2. Set of revolting disturbance. It is disturbance of external environments, which can change the state of the element from the point of view of its serviceability. Can be applied to elements of any type.

$$v_i \subseteq V, \quad i=1,n_v; \qquad (2)$$

2. Internal.
   2.1. State of the element from the point of view of its functioning (f - condition).

$$f_j \subseteq F, \quad j=1,n_f; \qquad (3)$$

   2.1.1. State of the element from the point of view of its serviceability (i - condition).

$$i_j \subseteq J, \quad j=1,n_j; \qquad (4)$$

   2.2. State of the element's communications.

$$c_j \subseteq C, \quad j=1,n_c; \qquad (5)$$

In view of determined thus set of variables, the logic of change the state of an element of technical means can be described by system of the equations:

$$\left.\begin{array}{l} f_k(t_j)=F(x_k(t_j),y_k(t_j),f_z(t_{j-1}),i_z(t_{j-1}),c_z(t_{j-1})); \\ i_k(t_j)=I(x_k(t_j),y_k(t_j),f_z(t_j),i_z(t_{j-1}),c_z(t_{j-1})); \\ c_z(t_j)=C(f_z(t_j),i_z(t_j),c_z(t_{j-1})); \end{array}\right\} \qquad (6)$$

where:

k - index of an element;

j - moment of time of beginning of event;

$z \subseteq Z_k$ set of elements, connected to the element k;

f - state of an element from the point of view of functioning;

i - state of an element from the point of view of serviceability;

c - state of communications of an element;

x - controlled influence;

y - revolting disturbance.

For elements, dynamics of transients of which at change of the state it is impossible to neglect and for which dynamic models of transients of the "low" level are started, the variables $f_{j-1}, i_{j-1}, c_{j-1}$ determine initial meanings of parameters of transient, and the variables f_j, i_j, c_j determine a state, into which the element will be switched.

In each moment of time in system there is no more than one event. At realisation of the given model it is offered to go over from a principle " delta t " or " delta z " to the direct generation of events in models of separate elements and detection of these events by other models, including logic. For the first generated and fixed event (control or disturbance influence) the element is determined, on which this event has affected, his new state and state of communications. The logic model is solved with new values of states, new events in system are expected. The process will proceed until the models of elements will stop to cause change of a states of elements and communications, determined for logic model of the "top" level.

Soon after the simulation model is calculating the new state, into which technical system will be passed in result of control or disturbance influence, there are starting algorithm, which determine, which elements in system, not being in an emergency state, can be switched into the functioning state.

The work of this algorithm is similar to algorithm of simulation model work. The distinction is that all active elements of system the algorithm tries to switch into a functioning state and check, that they will not switched in emergency.

Sources of substance (energy), defined for each group (complex, system) are first of all checked. From sources are first of all checked this, which have not entrances of substance (energy) or at which condition of functioning already are carried out. Further for an elements, state of which is changed on functioning, its connections are checked, names of elements, which are connected to this, are entering into buffers. After these elements are taken from buffers in special succession, conditions of their functioning are checking, and, if they are carried out, the element is switching into a functioning state and for him connections are checking, names of elements, which are connected to this, are entering into buffer.

The extraction procedure of elements from buffers and work with them are stopping, when buffers of static and dynamic elements become empty. After this the next source of substance (energy) is checking. The checking and elements switching into a functioning state is stopping, when for one cycle of checking of sources will not be switched into a functioning state of any element. Into a database about a physical state of means, the information, received on the basis of the given algorithm, is writing as forecast. Simultaneously the formation of the connection's matrixes of functional groups for elements, which can be switched into a functioning state, is executing.

After fulfilment the tasks of reception of the information and the forecast of the technical state of system with estimation of an opportunity of switching the elements into a functioning state, on the basis of data, written into database , the task of a substantiation of modes of functioning of means is starting.

During reception of the information from a monitoring system and the realisations of algorithm the simulation model functioning, into a database are been write the data about a state of each element of technical system from the point of view of functioning and serviceability, and also about weight nominated to an element. Weight of an element - numerical size, is equal to probability of a element refusal, and is calculated at the formula:

$$Q(t) = 1 - e^{-\int_0^t \lambda(x)dx} \quad ; \qquad (7)$$

where t - time,

λ - failure rate of an element, calculated on the formula:

$$\bar{\lambda}(t) = \frac{N(t) - N(t + \delta t)}{N(t)\delta t}; \qquad (8)$$

$$\lambda(t) = \bar{\lambda}(t)|_{\delta t \to 0, N(t) \to \infty}; \qquad (9)$$

where:

N(t) - number of elements, was not refused up to a moment of the time t,

N(t+δt) - number of elements, was not refused up to a moment of the time t+δt.

The probability of a refusal of each element is calculating during functioning the technical means and the simulation model.

State of elements of technical system, the connections and weight of elements are initial data for the solving a problem of a feasibility of modes of functioning of technical means.

The substantiation of a mode of functioning of means is understood as the decision of the following two problems:

1. Calculation of functions (target purposes) system, which can be realised at the given technical state;
2. Calculation of optimum ways between sources and consumers, realising accessible functions (target purposes).

Functions (target purposes) system (functional complex, the groups) are assigned by a set of sources and consumers of energy (substance and etc.), i.e. by the set of tops on the graph.

As the mode of functioning of means is understood as set of a states of elements of technical system and connections between them, ensuring fulfilment determined for system (complex, group) functions, the task of a feasibility of a mode is formulated as follows: For functions (target purposes), determined in system to find the minimum ways between sources and consumers of energy, substance and etc.

The criterion function for a feasibility of a mode looks as follows:

$$F = 1 - \prod_{z_i \subset Z} (1 - Q_i) \; ; \qquad (10)$$

Physical sense of the given criterion function - probability of a refusal of fulfilment of function at given realisation the way between a source and consumer of energy (substance etc.). The way between sources and consumers is searched on all set of elements, which can be switched into a functioning state and are not in a condition of failure or dependent refusal.

Thus, the task of a mode's substantiation it is possible to write down:

$$\left.\begin{array}{l} F_n = 1 - \prod_{z_i \subset Zn} (1 - Q_i) \to \min \; ; \\ 0 < Q_i < 1; \\ i_i \not\subset i_{avar}; \\ i_i \not\subset i_{zavis.\ avar}; \end{array}\right\} \qquad (11)$$

Provided that such ways exist.

n - amount of functions (target purposes), determined for system;

Zn - set of elements, realising function with an index n on which are searched the ways, ;

Qi - probability of a refusal of an i element;

i_i - state of an i element from the point of view of serviceability;

By other words, for technical system it is necessary to find such ways between sources of energy (substance) and consumers, the probability of "break" of which would be minimum from possible.

The given problem is decided through algorithm, in a basis of which Floyd's algorithm and Uorsholl's procedure lays. The Floyd's algorithm finds the shortest ways between all n (n-1) ordered pairs of tops in oriented graph with n tops. We shall consider it more in detail.

Let tops of the graph are designated in figures $1,2,\ldots, n$. As at assignment the weights positive numbers were used only, there is not oriented cycle with negative length in the graph. Let $W=[w_{ij}]$ - $n \times n$ - matrix of lengths of graph edges G, i.e. w_{ij} - length of an oriented edge (i,j) in graph G. We believe $w_{ij} = \infty$, if in the graph there is no edge (i,j), oriented from i in j. We believe also, that $w_{ii} = 0$ for all i.

Since a matrix $W^{(0)}$ the Floyd's algorithm builds a sequence $W^{(1)}, W^{(2)}, \ldots, W^{(n)}$ such $n \times n$ of matrixes, that an element $w_{ij}^{(n)}$ the matrix $W^{(n)}$ is distance between i and j in graph G. The matrix $W^{(k)} = [w_{ij}^{(k)}]$ is determined on a matrix $W^{(k-1)} = [w_{ij}^{(k-1)}]$ according to a following rule:

$$w_{ij}^{(k)} = \min\{w_{ij}^{(k-1)}, w_{ik}^{(k-1)} + w_{kj}^{(k-1)}\} \qquad (12)$$

Let $P_{ij}^{(k)}$ way of minimum length among all oriented i - j of ways, which use as internal tops only tops from set $\{1,2,..., k\}$. For him there is the theorem, proving the true of Floyd's algorithm.

Except the shortest lengths it is necessary to receive also ways with such lengths. It is reached at realisation of the Floyd's algorithm as follows: in accordance with construction of a sequence $W^{(0)}, W^{(1)}, \dots, W^{(n)}$ other sequence of matrixes $Z^{(0)}, Z^{(1)}, \dots, Z^{(n)}$, is simultaneously built, such, that an element $z_{ij}^{(k)}$ matrix $Z^{(k)}$ specifies top, which directly follows top i in $P_{ij}^{(k)}$. It is clear, that in this case:

$$z_{ij}^{(0)} = \begin{cases} j, \text{ if } w_{ij} \neq \infty; \\ 0, \text{ if } w_{ij} = \infty; \end{cases} \qquad (13)$$

$Z^{(k)} = [z_{ij}^{(k)}]$ turns out from $Z^{(k-1)} = [z_{ij}^{(k-1)}]$ according to a following rule: let $M = \min\{w_{ij}^{(k-1)}, w_{ik}^{(k-1)} + w_{kj}^{(k-1)}\}$, then:

$$z_{ij}^{(k)} = \begin{cases} z_{ij}^{(k-1)}, \text{ if } M = w_{ij}^{(k-1)}; \\ z_{ik}^{(k-1)}, \text{ if } M < w_{ij}^{(k-1)}; \end{cases} \qquad (14)$$

The validity of this rule is defined following. If $M = w_{ij}^{(k-1)}$, length $P_{ij}^{(k)}$ is equal to length $P_{ij}^{(k-1)}$. Therefore $z_{ij}^{(k)}$ coincides with $z_{ij}^{(k-1)}$. On the other hand, if $M < w_{ij}^{(k-1)}$, $P_{ij}^{(k)}$ - association of ways $P_{ik}^{(k-1)}$ and $P_{kj}^{(k-1)}$. Therefore $z_{ij}^{(k)} = z_{ik}^{(k-1)}$. It is obvious, that shortest i-j the way is defined by a sequence of tops $i, i_1, i_2, \dots, i_p, j$, where:

$$i_1 = z_{ij}^{(n)};\quad i_2 = z_{i1j}^{(n)};\quad i_3 = z_{i2j}^{(n)}; \dots ; j = z_{ipj}^{(n)}; \qquad (15)$$

Updating the given algorithm for the decision of a problem of a substantiation of modes of functioning of means consists in replacement of a rule (12) on following:

$$w_{ij}^{(k)} = \min\{w_{ij}^{(k-1)}, 1-[(1-w_{ik}^{(k-1)})(1-w_{kj}^{(k-1)})]\} \qquad (16)$$

It is caused by that it is necessary to us to find the minimum probability of a refusal of a way, and not minimum sum of lengths. Weight of arches are nominated equal weights of elements, for which this arch is outgoing.

The adapted Floyd's algorithm looks as follows:

1. $W = [w_{ij}]$ - $n \times n$ - matrix of lengths of edges in given oriented graph G. Let's put $w_{ij} = \infty$ for all $i = 1,2, \dots ,n.$. $Z = [z_{ij}]$ - $n \times n$ - matrix, in which:

$$z_{ij} = \begin{cases} j, \text{ if } w_{ij} \neq \infty; \\ 0, \text{ if } w_{ij} = \infty; \end{cases}$$

2. Let $k = 0$;
3. Let $k=k+1$. For all such $i \neq k$, that $w_{ik} \neq \infty$, and all such $j \neq k$, that $w_{kj} = \infty$, carry out the following actions:

   3.1. To put $M = \min\{w_{ij}, 1-[(1-w_{ik})(1-w_{kj})]\}$;

3.2. If $M < w_{ij}$, put $z_{ij} = z_{ik}$ and $w_{ij} = M$;

4. Actions are being carried out:

4.1. If some element of a matrix $w_{ii} < 0$, the top i belongs to some oriented cycle with negative length. A stop of the algorithm realisation;

4.2. If all $w_{ii} \geq 0$ and $k = n$, $[w_{ij}]$ - length of all shortest ways, and z_{ij} - first top after i in shortest i-j of a way. A stop of realisation of algorithm;

4.3. If all $w_{ii} \geq 0$, but $k<n$, go to a step 3.

On ending fulfilment of Floyd's adapted algorithm we receive a matrix Z, in which following information is contained: z_{ij} - first top after i in shortest i-j way. That is, if the way between a source and consumer exists (function is carry out), $z_{ij} \neq 0$. Such way is optimum from the point of view of non-failure operation.

The information on carrying-out functions (target purposes) is deduced to the operator as the hierarchical menu. The top level - function of system as a whole, the level is lower - functional complexes, functional groups is much lower. At each level the list of functions, grouped on systems, is displayed. Depending on that the function is carried out completely, partially or not carried out, the name of each function is highlighted by the appropriate colour. As in result of accounts on Floyd's modified algorithm is arrived the matrix Z, being a matrix of optimum ways in system, there is opportunity to deduce on the screen the slide of the appropriate group (complex), on which the recommended optimum ways of realisation of functions will or be highlighted, or are allocated in another way (3-d, blinking, etc.).
The thus processed and submitted information is initial for acceptance of the decision the operator.

The method described is oriented to realise it in the local computer network in multi-purpose real-time operating systems. The simulation in the Ethernet network under the operating system QNXv4.22 control showed:

1. The possibility of the increased safe use of the technical system by means of proving the mode and elimination of the wrong actions of the personnel;
2. Expediency of this system use under designing and as a training means.

References:

1. Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений. М.: Мир 1984. (in Russian).
2. Клир Дж. Системология. Автоматизация решения системных задач: Пер. с англ. М.: Радио и Связь, 1990. (in Russian).
3. Новожилов А. Н. "Имитационное моделирование в системах управления движением". XXII Всероссийская конференция по управлению движением морскими судами и специальными аппаратами Тез. доклада г. Сочи 1995г. (in Russian).
4. Советов Б. А., Яковлев С.А. Моделирование систем. М., 1985. (in Russian).
5. Дж. Ульман. Основы систем баз данных. М.: Финансы и статистика,1983.(in Russian).
6. Шенон Р. Имитационное моделирование систем -искусство и наука.М.,1978.(in Russian).
7. A. Novodjilov "Substantiation of modes of safe functioning complex technical systems". International Conference "Instrumentation in ecology and human safety" ISA (International society for measurement and control) St. Petersburg Russian Section. Russia, St. Petersburg 1996 p.130-132. (in Russian).

PRINCIPLES OF AUTOMATIC PROCESSING OF MESSAGES ON THE SAFETY AT SEA

Sergey Smolentsev
Institute Transport's Problems of Russian Academy of Science
V.O. 12 Line #13 199178 St.Petersburg, Russia
e-mail: wasily@transas.com

Abstract. This report considers the basic principles and procedure of solving a problem of automatic normalization of the text information related to the safety of navigation. Among other things the issues of formalizing the messages in NAVTEX system are considered.

Keywords. normalization of the text, NAVTEX messages, safety of navigation

System of messages on the safety of navigation. The safety of navigation is one of major problems of sea transport. To solve this problem IMO has established a World Service of the Navigational Warnings. NAVTEX is its component part. It includes a network of ground transmitting stations and NAVTEX receivers on the vessels. NAVTEX service provides the mariners with the navigational and meteorological warnings and other urgent information. These messages are transferred by the ground stations at 518 kHz frequency and are automatically received by the specialized shipboard receiver. Under GMDSS, NAVTEX receiver is mandatory part of the vessel's equipment.

Though NAVTEX messages are received in the automatic mode, the text of the message is not processed in the receiver, and is given to the navigator in an unpacked form. The navigator either takes this information "into consideration", or displays it manually on the navigation chart. NAVTEX information, however, can be used more efficiently, especially in the combination with an electronic chart (ECDIS). For this purpose it is necessary to create a system of automatic message processing.

The system for the automatic processing of NAVTEX messages should solve the following problems:

- to receive data from the NAVTEX receiver as it arrives;
- to classify the messages and assess the quality of the reception;
- to maintain the database using the received messages;
- to process (formalize) the text of the messages;
- to prepare information for the display on the electronic chart.

The problem of retrieving messages from the NAVTEX receiver consists in creating an interface between the receiver and the computer and is of technical character. The problems of classifying the messages (by the types of information) and assessing the received message quality, as well as of creating and maintaining the database up-to-date with the received messages are quite important, but are easily solved by standard methods. The problem of preparing information for the display in the ECDIS consists in creating a converter from some formalized structure to a format appropriate for ECDIS. The most difficult problem is a problem of formalizing the text of the message. It is the consideration of this problem which the present paper is devoted.

The Structure of a NAVTEX message. The format of a NAVTEX message is determined in "the Manual on NAVTEX Service" and has a structure, shown in Fig. 1. The structure of the messages considerably facilitates their initial analysis and classification. Thus signatures ZCZC and NNNN (at the beginning and the end of the message) allow the received information flow too be easily divided into a number of messages. The mandatory fields in the heading: "symbol of the station", "type of information" and "number of the message" can be used for automatic sorting of the received messages. The line containing the time of drawing up the message can also be easily decoded and is used in the system for registration of the received messages.

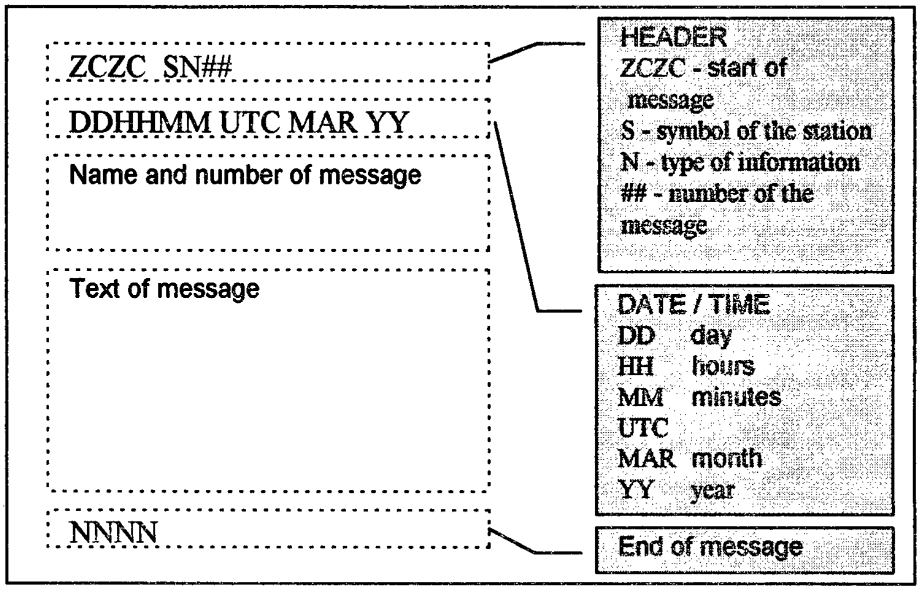


Fig. 1 Format of a NAVTEX message

Nevertheless, the structure of a "text" field is not set and is composed in an arbitrary form. It is this field which contains all the semantic information, transmitted in the message. Thus, the problem of formalizing the text of the message consists in creating a special formal structure, reflecting the meaning of the specified textual field. This formal structure can be used for

- the automatic display of information on an electronic chart;
- the automatic creation of survey charts (for example : charts of weather forecasts, ice conditions, gale regions);
- transmission of data to the vessel control system.

Problem of processing (formalizing) the text of the message. The problem of automatic normalization of the text information consists in translating the texts written in a natural (limited) language and concerning some subject area, into a previously created formal language, and obtaining formal data structures corresponding to these texts, and suitable for the further automatic processing. It means that it is necessary to identify a particular situation, described in the message and to create a formal data structure appropriate to it. Various levels of understanding of the text of the messages are possible:

absence of understanding - referencing the entire text of the message to the region of responsibility of the transmitting station, determined from the heading of the message.

elementary understanding - extracting geographical names and coordinates from the text and referencing the whole text of the message (or appropriate fragments) to the areas , determined by these geographical names or coordinates.

understanding at the level of key words - extracting from the text, along with the geographical names and coordinates, the key words, which are used by the receiving party for getting an approximate idea on the character of a situation.

understanding of a described situation - creating a formal structure describing a situation - frame on the basis the text. The possibility of executing this frame is provided- i.e., implementation of actions corresponding to the described situation (display on the chart, distribution of messages, generation of alarms etc.)

The procedure for solving the normalization problem. The procedure of the text analysis (complete understanding of a situation) can be divided into a number of stages:

- extracting and recognizing individual items of the text (with the use of patterns determined beforehand): signs, words, numbers, coordinates, spoiled words (containing *symbols), special words (of m/s type), etc.;
- forming names marked with " symbols (which can consist of several words);
- recognizing geographical names (using the database of the geographical names and navigation area);
- extracting the idioms and collocations (using the dictionary of idioms);
- recognizing words (taking into account their possible changes:- plural, verb forms etc.) using the dictionary and adding any characteristics (part of speech, tense, etc.) - morphological analysis;
- extracting separate sentences and their grammar analysis with using the language rules database;
- formation the preliminary descriptions of situations (preliminary frames);

search for the appropriate frames in the database, selecting the most suitable option (options) and filling in its fields' values.

Implementation of the message formalizing system. For the implementation of the message formalizing system based on the above principles, it is necessary to define:

- a special work structure of the data for the text of the message;
- the mechanism of this structure transformation.

The order of transformation of the working data structure (see Fig. 2) will be described in the current section in more detail.

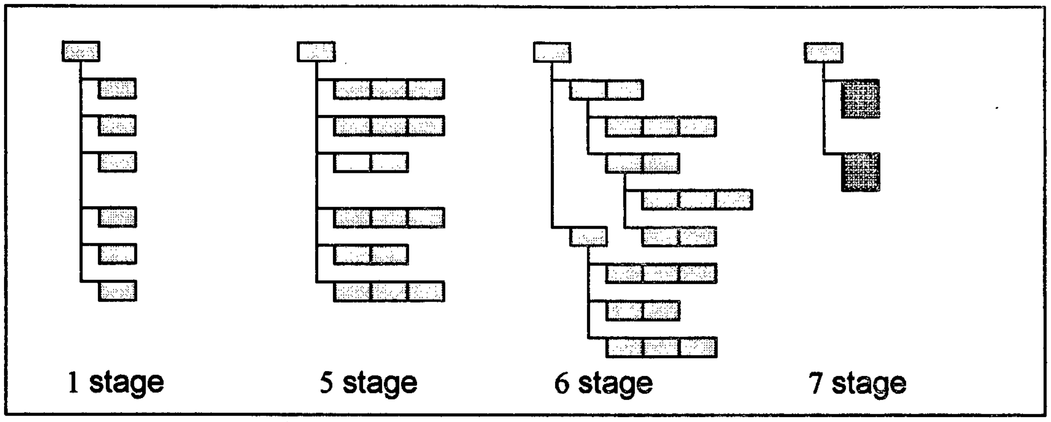


Fig. 2 Evolution of the working structure

Originally the working text structure is b a linear list of the words extracted at the first stage, with a unique attribute - a type of item (word, number, special type, not understood, etc.). At the second, third and fourth stages of processing this structure remains linear. As a result of fulfilling these stages some segments of this list can be combined in a single element (merged in a single item), and the type of the given element can be modified (for example - a certain geographic name defined). Processing of the given linear structure at these stages will be carried out with the use of predetermined pattern lists and databases (geographical terms, dictionary of idioms, abbreviations, etc.).

At the fifth stage the morphological processing of the working structure is implemented. The dictionary of terms and a special algorithm (final automata) for the recognition of the endings (plural, verb forms, participle, etc.). As a result the working structure is not altered(remains a linear list of elements), but some additional grammar attributes (part of speech, number, tense modality, etc.) are added to each element. Several words can be merged in a single item (not necessarily arranged one after another, as in the compound tenses of a verb). It should be noted that at this stage the definition of several versions for a single item is possible. Here, if it is preset in the dictionary, the attribute "meaning" is determined for the elements.

At the sixth stage the grammar processing of the working structure is carried out. A production database of the language grammar rules and the algorithm of unification should be used for this purpose. The working structure is transformed into a tree- like structure and it corresponds to the grammar structure of the text of the message. At this stage a multivalent in grammar attributes is partially removed in individual items. On the basis of the obtained structure, with the use of an additional production database on the subject area the transformation of "meaning" attributes in separate items is performed, and the meanings of these attributes for the higher level structure elements are formed (word collocations, sentences, paragraphs of the text).

At the seventh stage only the "meaning "attributes are handled in the working structure. All structures at the sentence level will be transformed to the framelike structures, reflecting the meaning of these sentences. The working structure is folded to form a linear list of the corresponding framelike elements.

At the last stage the obtained framelike structures are compared to the frames, stored in knowledge base. These frames represent the knowledge of possible situations in the given subject area. In case of finding a match, the element of working structure is replaced with a frame with specified values of its parameters. Otherwise the situation, described in the message, is considered as new and is used for updating (training) of the knowledge base.

As a result of steps described above, a formalized presentation of a situation in the form of a linear list of frames, connected with the data structures, required for the further use, procedures etc.

Conclusions. The use of the principles stated above allows creating a system for the automatic recognition and processing of the text messages, related to the safety of navigation. Such system may be rather useful as part of a ship's automation system. At present the work on such system is under way at “TRANSAS MARINE” Ltd

DYNAMIC ARCHITECTURE PROCESSORS IN VESSEL AND AIR TRAFFIC MANAGEMENT SYSTEMS.*

Victor Afanasiev, Vladislav Sapegin
Joint-Stock Company "Marine Complexes & Systems +"
37, Stachek, St.Petersburg, 198097, Russia

Valery Torgashev
St.Petersburg Institute for Informatics and Automation
39, 14th Line, St.Petersburg, 199178, Russia
e-mail: vat@mail.iias.spb.su

Abstract. Several topics associated with effective processing of radar information in vessel and air traffic management systems are considered. It is shown that usage of very high efficient dynamic architecture processors with adjustable structure depending on a problem and a purpose enables to unify systems of various purpose and, at the same time, to reduce their cost and widen their functional capacity.

Keywords. Dynamic architecture, radar processor, flexible logic, traffic management systems.

Introduction. Radar surveillance systems are important in the questions of vessel and air traffic management safety control, in air defence and frontier guard systems. As far as the radar information volume after its quantification reaches a few million values per second and each value is to be processed, it is very important to provide extremely high capacity of the radar information processing (about one billion operations per second). No existing universal processor can provide the above capacity that involves necessity to use specialized processors with the most frequent algorithm realization. The special purpose ones are expensive (exceeding $100.000), have a poor algorithm updating ability and specific purpose orientation. The described dynamic architecture processors do not have some of the above disadvantages. Their theoretical basis was worked out in the St.-Petersburg Institute of Informatics and Automation of the Russian Academy of Sciences. The processors are designed by the 'Dynamic Electronic Systems' Company and developed and

* This research is conducted within research grant No SPC-96-4079 from USA Air Force.

implemented by the 'Marine Complexes and Systems Plus' Joint-stock company located in St.-Petersburg.

Dynamic architecture processor. The basic concept of the dynamic architecture processor is an universal structure based on flexible logic, registering devices and high-performance microprocessors. The structure of software is adjustable for appropriate functioning that enables to realize corresponding processor for solving any problem. For difficult problems the dynamic architecture is variable while problem solving.

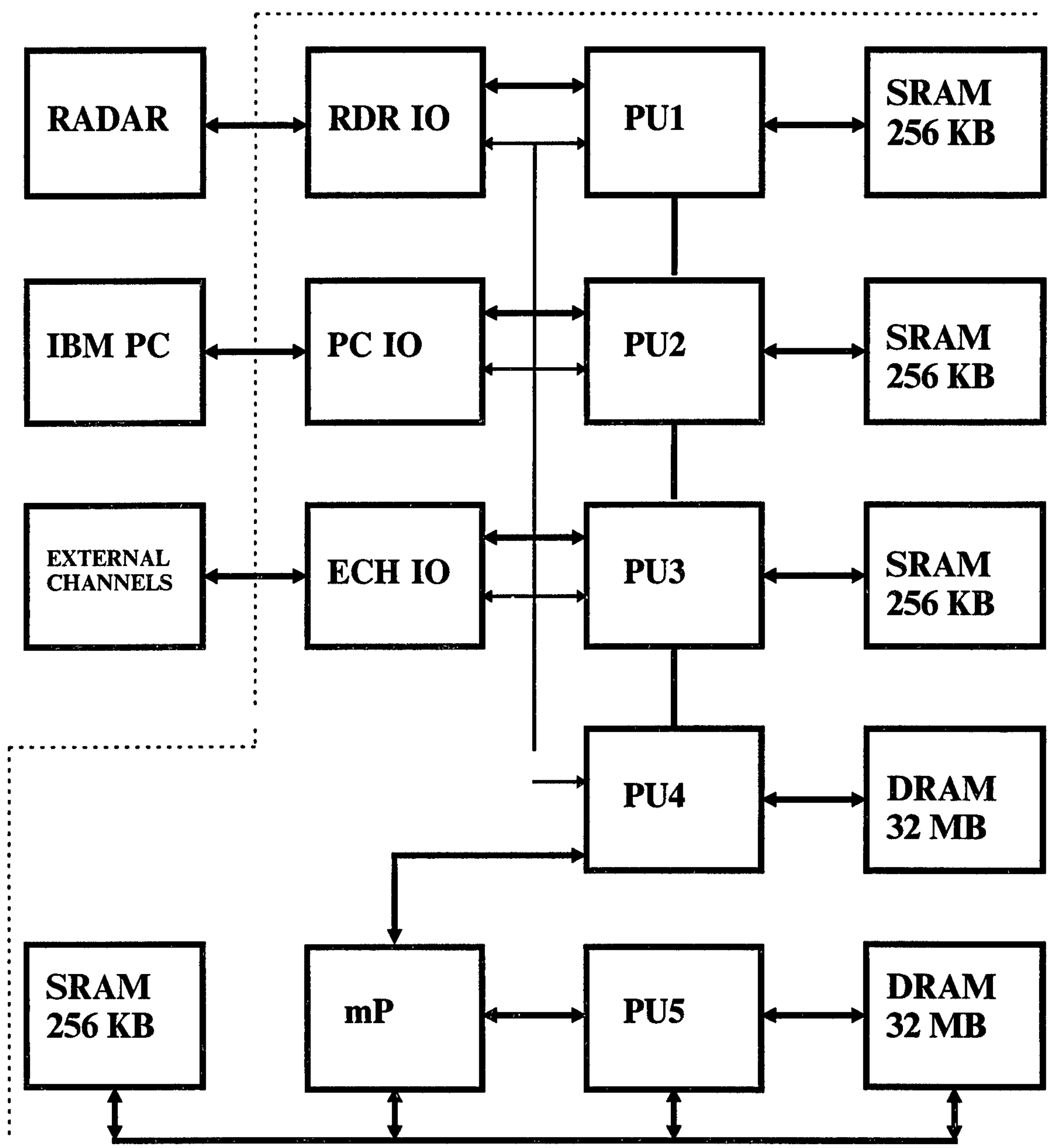


Fig. 1

Figure 1 shows the structure of the DAP-321 processor consisting of five processing units with local static and dynamic memory storages; three input-output units providing connection with a radar, a personal computer and external channels, correspondingly; high-performance microprocessor TMS320C32-60 which has a common static and dynamic memory storages with one of the processing units. All the units are connected by means of a fast bus and all the processing units have additional communication links. Each of the above units is designed as a flexible logic scheme manufactured by the 'Altera' company and adjustable for realizing different parts of the general algorithm. For instance, when using the processor in radar systems the processing units carry out the following functions:

- PU1 - interference suppression by different methods;
- PU2 - detection of targets, calculation of their coordinates and some parameters;
- PU3 - target selection and general coordination of all units;
- PU4 - image synthesis for displaying on an indicator screen;
- PU5 - in combination with the processor, secondary processing and forming target envelope.

The processor's clock frequency is 60 MHz, that enables to follow up to 60 million commands per second. The working frequency of other units is 40 MHz. Each of them carries out from 2 to 5 operations per cycle and each operation also includes a few commands of a traditional processor. Actual performance of the processor exceeds 1500 million commands per second. As far as Intel Pentium processors need two tacts for a command (average) for similar programs and often use of the main storeage, it means that the DAP-311 capacity is 10 times faster, being compared with the Pentium-200.

Traffic systems. Due to high capacity and architecture flexibility the processor is easily adaptable to any tasks of radar information processing. Using the **terms of manufacturers,** processor with such a functional capacity is called 'Radar processor'. Nowadays most radars are used as sensors of huge informational systems, for instance, a Coastal Vessel Traffic System (VTS) shown at fig.2. The radar processor is the main unit of the system. A digital array (expanse) on the processor's output contents a primary digital information of environment which is supplemented with information from other sensors:

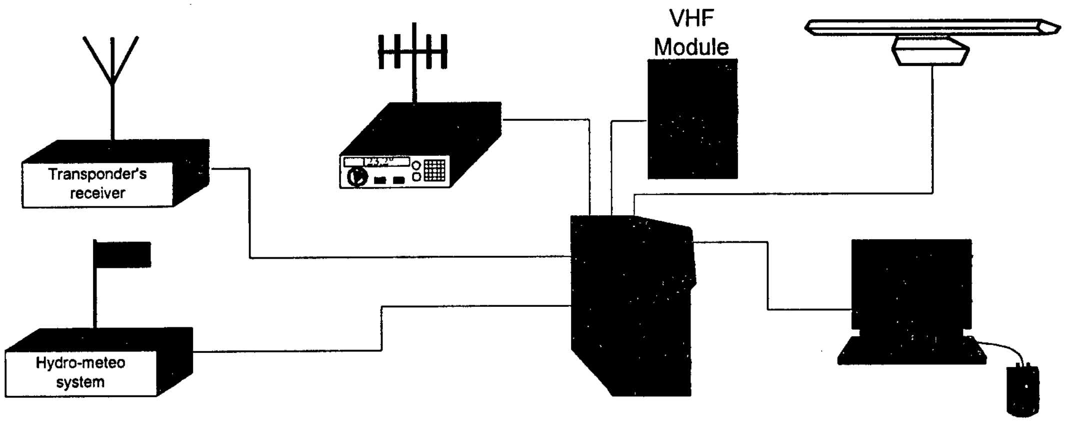


Fig. 2

- VHF direction finder,
- VHF transponder,
- automated hydrometeostation,
- VHF station.

Such a system is considered as 'open', i.e. easily supplemented with other additional sensors, including television camera. The radar processor enabled to design a new radar indicator of new high quality and low cost. In addition to above-water situation displaying, the device also carries out the radar control and a complex of other functions:

- automatic tracking;
- calculation of a target motion parameters;
- control for a target trajectory;
- control for buoyage and environment;
- control for anchorage.

Analogue-digital signal conversion improves data processing quality and antijamming capability without any change of transceiving tract. In compliance with international requirements all the displayed information has to be registered and stored for a month. Nowadays special expensive audio and video recorders are used for this purpose (about $30.000 per set). The 'MC&S+' company has designed a device with higher quality based on flexible logic. All the information is recorded and stored on a hard disc as a digital array. Special software enables to open archived records, to analyse a situation and print a copy when required. The unit has compact size being inserted into a personal computer. Audio information and secondary radar data (target envelopes) are also recorded in digital form. Such devices are usable in air defence systems, environment control systems. The extracting and tracking unit is a principle new device (figure 3) being a base of regional level VTSs.

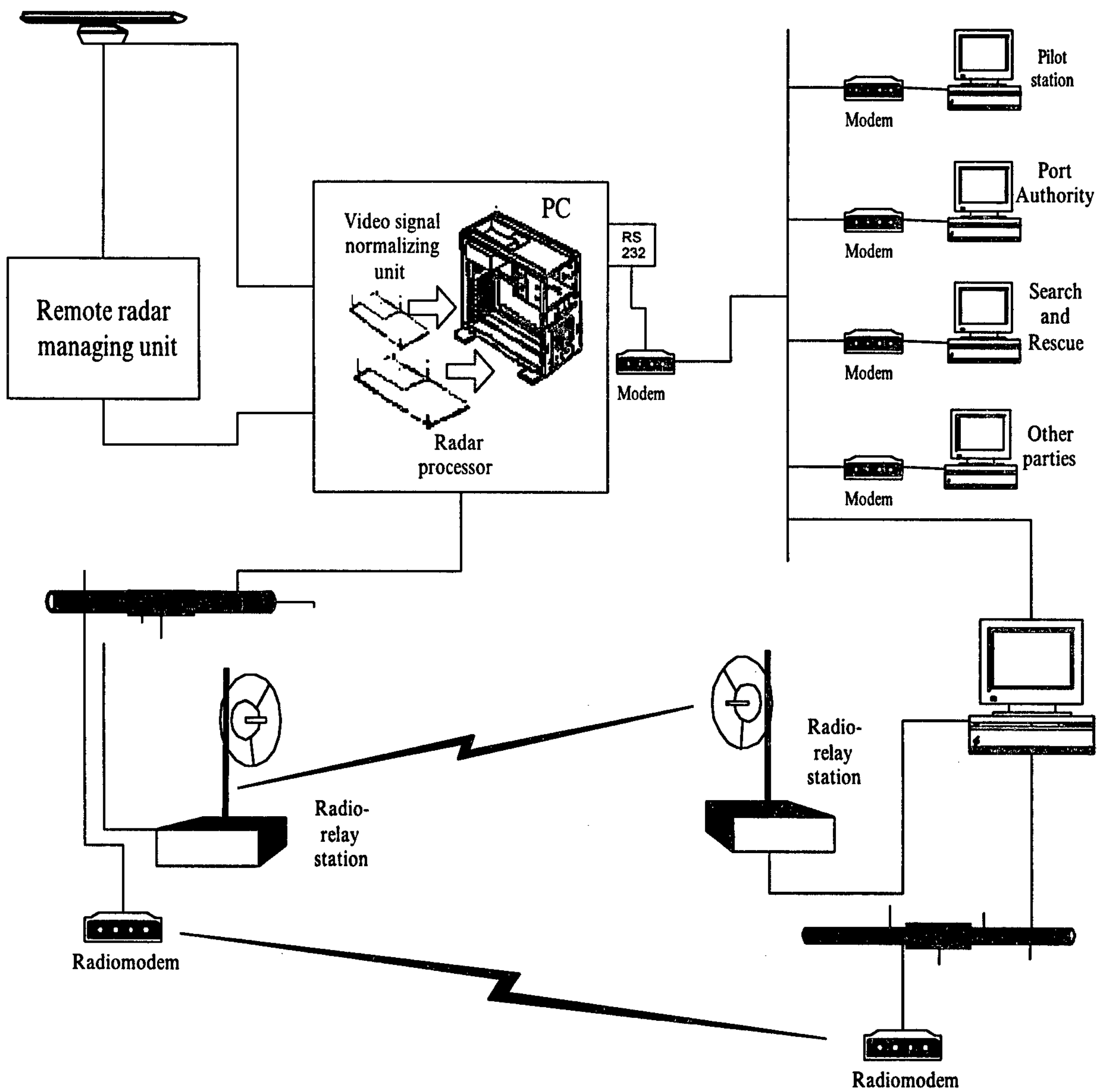


Fig. 3

As a rule a user needs an information of manoeuvring targets only. It involves another procedure. All information of coastlines, navigational features etc., is to be plotted on an electronic chart beforehand. The tracking center receives only secondary radar information (target envelopes). Remote automated radar posts equipped with target extractors are used for this purpose. It provides reducing of the information to be transmitted and using usual telephone line instead of expensive radiorelay lines.

Conclusions. Now dynamic architecture processors DAP-311 are being used in some real traffic mamagement systems such as, for example, Murmansk port VTS or distributed radar system on Norway oil platforms and they are showing very high perfomance and effectivness and some new possibilities. The authors would presume that above proposed research gives a new lead and develops in the field of new generation radar systems.

PECULIARITIES OF NAVIGATION AND CONTROL OF A RELATIVE MOTION OF THE AEROSPACE PLANE AT THE DOCKING WITH EKRANOPLANE ON A FINAL STAGE OF LANDING

A.V. Nebylov, V.V. Razumovich, N.E. Sokolov
State Academy of Aerospace Instrumentation, St.Petersburg, Russia
V.B. Diomidov, CS&R Institute "Electropribor", St.Petersburg, Russia
E-mail: alex@nebylov.usr.saai.ru

Abstract. In the paper the concept of landing of an aerospace plane (ASP) without an undercarriage on the basis of its docking on a final stage of lowering with Ekranoplane, subsonic airplane or other transport means with necessary carrying capacity, speed and stability of a motion is developed. A similar principle of landing allows to expand considerably freedom in choice of a trajectory and final landing point, and also to increase an ASP useful payload. However trouble-free docking requires very effective systems of navigation and automatic control of absolute and relative motion. A special docking system with a capability of high-precision local positioning of docking elements for parrying residual errors of relative motion control system also is necessary. Thus on ASP only the passive devices can be placed because of weight limitation and reliability requirements.

Keywords. Ekranoplane, aerospace plane, docking, navigation, video camera, identification.

1. The concept of ASP and Ekranoplane integration. The space transport systems with horizontal take-off basically have a number of functional advantages in compare with systems of vertical take-off [1-3], though yet is not quite clear and not checked experimentally the possible version of such advantages. However is clear, that ASP with a hypersonic wing for overcinking on take-off of a low subsonic speeds site requires in special take-off assist, as which it is possible to use a heavy plane or Ekranoplane. Ekranoplane in difference from a plane is capable to help not only take-off rather heavy ASP, but also it to landing, that is advantage of the Ekranoplane. The concept of use of heavy katamaran type Ekranoplane with an own take-off mass 1500t for ASP with own weight 500t assist at horizontal take-off and horizontal landing in any point of a sea surface at landing weight of approximatelly 70t is described in [3-5]. In the present publication some problems of a construction of the informational and control complex for automatic landing of ASP on Ekranoplane, exacter, docking of two these vehicles at motion with speed about 0,5M close to sea surface are considered.

2. The concept of motion control system design. The joint motion control system of vehicles during docking is a control loop many-dimensional digital automatic system.

It is clear, that for increase of quality of control and it reliability it is necessary to consider both vehicles as elements of a control loop system. The possible generalized block-diagram of such the control loop is presented in a Fig. 1.

The given system is the two-channel scheme of joint control of bicomponent dynamic object, where the first channel provides control of the ASP, and second - Ekranoplane. At preliminary research it is possible to consider ASP and Ekranoplane as oscillatory units with different time constants which differs in a factor by 10^{-1}.

The docking ASP and Ekranoplane is independed of a principle of a docking elements construction and is possible only with use of a high-precision short-range navigation system, to

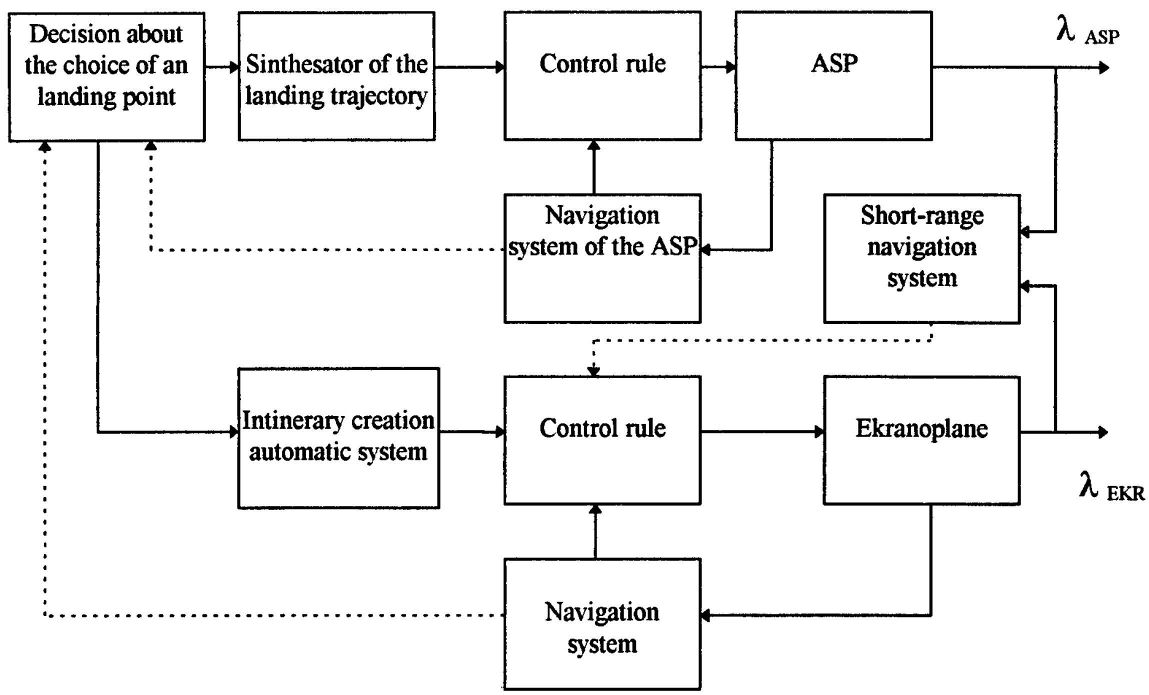


Рис.1 Block-diagram of a joint motion control system of Ekranoplane and ASP

which prefears next requirements:
- high accuracy of definition of angular and spatial coordinates ASP rather Ekranoplane (on horizontal displacement $2б \leq 10$sm, on relative altitude - $2б \leq 10$sm, on an angle - $2б \leq 30'$),
- high speed in response.

The output signals of this system are necessary for formation of pilot signals of a docking system.

3. The concept of short-range navigation system design. The similar navigation system can realize various methods of outcroping of information about relative moving parameters most of applicable from which are radar, optical and their combinations. Unconditional it is necessary to apply their integration with inertial instrumentations, i.e. to realize the integrated system of navigation.

The radar method implies either presence of special radiobeacons the radiation of wich is used for creation special radionavigation field wich permits to define a position of ASP rather Ekranoplane, or presence on ASP of corner-type reflectors or their analogs, which reflect sounding impulses from Ekranoplane. The application of the radar method is complicated with small distance between ASP and Ekranoplane at last stage of a docking and capability overreflect of signals because of the large area radioreflecing surfaces, that can conduce to error definition of coordinates. Besides the arrangement of a part of elements of a similar system onboard ASP will have negatively effect on a reliability of a docking.

That is why it seems the preferable version is the application of an optical method. The following concept of a construction is offered.

On Ekranoplane landing site are located some (for example, three) fixed video cameras directed vertically up. The video cameras work in infrared range, using a natural property of the descended vehicle - high temperature of a body (some honeycombs of degrees). It allows to receive the more contrast image, that increases speed and the quality of identification of object Probable strong disturbance infrared radiation of the Sun can be eliminated or considerably to reduce by set-up of a passband of the video cameras or special filters on predominant frequency of radiation of a body ASP. The data of the videoimage from video cameras with a step about 30 frames per second converts into the digital type with use of videocapture plates and then transfers to the computer. The computer treats the obtained data in realtime, calculats values angular and attitude ASP rather Ekranoplane. The block-diagram of the given short-range navigation system is adduced on Fig. 2.

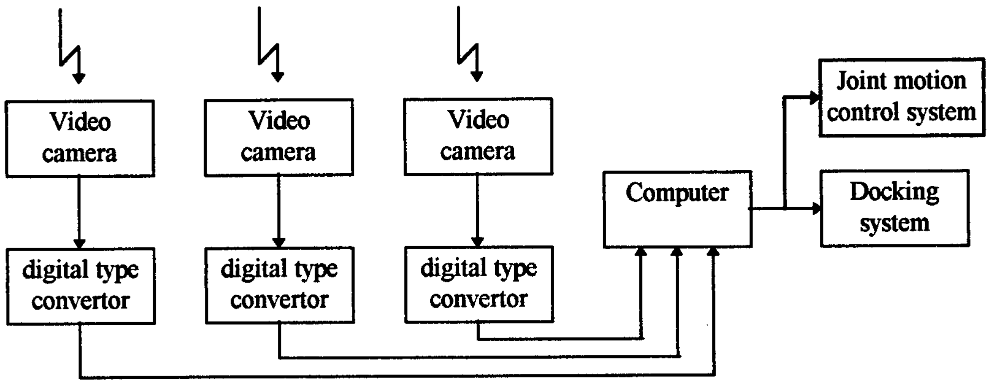


Fig. 2 The block diagram of a short-range navigation system

The given system can begin to work at altitude ASP about 100 meters above a sea level, when the initial correction of its position rather Ekranoplane is already produced and they move plane-in parallel. At a phase of landing ASP up to altitude approximately 25 meters the results of calculation are used for self-testing of activity of a system by their compare with the data obtained from ASP. Then, after switching on of ASP and Ekranoplane relative motion control loops the definition of a position ASP carries out by the optical system of navigation down to the finishing of process of a docking. Taking into account significant width of altitude range of activity of a system, it is expedient to change a focal length of video cameras depending on ASP altitude. The smooth or two-stage change of a focal length, or use of video cameras with a various focal length is possible. For facilitation of calculation and simplification of a system the second version is more preferable.

The cameras are located so that during landing they stand under the ends of the wings and under the nose of ASP. Also as characteristic surfaces ASP the specific areas possessing in high temperature can be used [6]. Range of temperatures of landing ASP body is shown on Fig.3.

Initially in a field of sight of video cameras ASP falls completely, but in accordance with lowering the cameras inspect only its part. At an active phase of the work of the system (lower 25 meters) videoimage contains the image of wings and nose on a background of the sky. Thus, working (characteristic) surfaces used for the computer analysis, are the wings ASP and it a nose, namely, their edge.

The accuracy of definition of coordinates ASP rather Ekranoplane depends on distance (altitude) up to it, resolutionability and focal length of video cameras.

As a resolutionability and focal length in a working mode are constants so accuracy of definition of coordinates is a function of distance:

$$e_s, e_{max} = F(d, f, z),$$

where: e_s, e_{max} - mean square or maximum error, d - distance up to object, f - focal length of the camera, z - resolutionability (const).

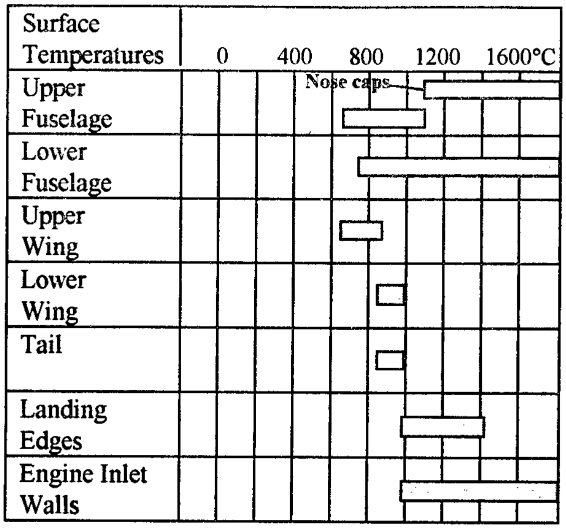


Fig. 3 Units - diagrams of the temperature range of airframe of space planes (SSTO)

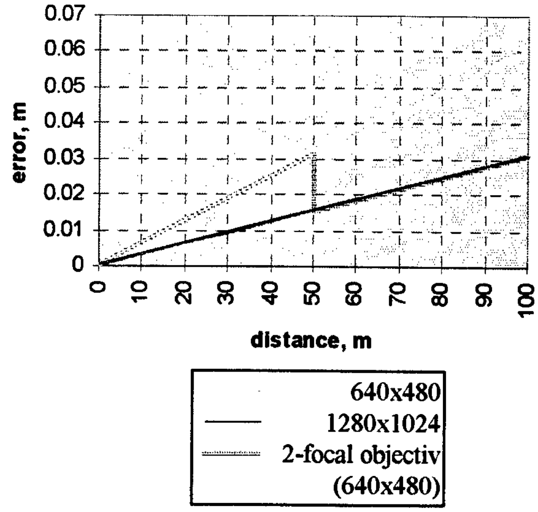


Fig. 4 Relation of a measurement error of a position ASP from distance at various resolutionability and at use of the two-focal stepwise objective

The estimated characteristics of coordinates definition quality for various resolutionability are given on Fig. 4.

As it is visible from the schedule, the use two-focal objective allows to increase accuracy of measurement on large distances. The switching of a focal length happens in a moment, when the informations about the image is unsufficient for identification because of large magnification of object. At altitude about five meters and at a resolutionability approximately 640x480 pixels can be achieved centimetric accuracy of definition of displacement ASP.

The digitization of the image is made by the plates of videocapture. Nowadays there is a great many of similar devices of the different producers, which provide digitization of the videodata with the large resolutionability in a realtime.

Digitalized videodata from video cameras trans to the computer. In the memory of the computer the original image ASP and its physical sizes are contained. The computer image analysis is made in three phases:

- Clearing of the engaging image of noise by known alhorithms [7,8];
- Finding of a original ASP image and it identification, using the original data located in memory. An results of identification is the precise outline of a wing ASP;
- Binding an outline to real physical coordinates and as arrangement of video cameras is known also calculation of attitude ASP rather Ekranoplane. As at change of bank angle or pitch ASP the image will accordingly vary pursuant to the laws of a perspective, an angular position also can be calculated.

The data of the computer calculation are used for positioning of docking elements on a final stage of ASP and Ekranoplane docking, and also as input data for a joint control system.

Conclusions. In the given paper for an estimation of control quality the ideal models of objects were used. At development of the real control system it is necessary to take into account a lot of the external and internal revolting factors: wave disturbance, turbulence of atmosphere, wing-in-ground effect, rigidity of a design, internal noise of system modules and others [9-11]. Also it is necessary to solve the problem of ASP identification stability.

References

1. Anfimov N.A. (1996) Conception of Future Reusable Space Transportation System. AIAA 7_{th} Spaceplanes and Hypersonics Tech. Conf., Norfolk.
2. Maita M., Kobayashi S. (1996) Japanesse Spaceplane and Hypersonic Technology Program Overwiev. 47th IAC, IAF 96-V.3.05, Beijing.
3. Nebylov A.V., Sokolov V.V., Tomita N., Ohkami Y. (1996) The Concept of the Ekranoplane Using as Take-off •and Landing Assist of Passengare Aerospace Plane. Transport: science, technology, management. №11. (in Russian).
4. Tomita N., Nebylov A.V., Sokolov V.V., Ohkami Y. (1996) Perfomance and Technological Feasibility Study of Rocked-Powered HTTL-SSTO with Takeoff Assist. 47th IAC, IAF 96-V.3.05, Beijing.
5. Nebylov A.V., Sokolov V.V., Tsurumaru D., Saotome T., Ohkami Y. (1996) Feasibility Study of a Rocket-Powered HTTL-SSTO with Ekranoplane as Takeoff Assist. AIAA 7_{th} Spaceplanes and Hypersonics Tech. Conf., 96-4517, Norfolk.
6. Megumi S., Akira S. (1994) The Current Status of Aerospace Plane Materials R&D in Japan. AIAA 5_{th} Spaceplanes and Hypersonics Tech. Conf., Tokyo.
7. Beits R., Mac-Donnel M. (1989) Image Recovering and Reconstruction. Mir, Moscow. (in Russian).
8. Pratt W. (1982) Imaging Digital Treats. Mir, Moscow. (in Russian).
9. Nebylov A.V. (1994) Measurement of Parameters of Flight Close to the Sea Surface. SAAI, St. Peterburg. (in Russian).
10. Diomidov V. (1996) Automatic Control of Motion of Ekranoplane. "Elektropribor" Scientific Research Institute, St. Peterburg. (in Russian).
11. Nebylov A.V. (1996) Structural Optimization of Motion Control System Close to the Rough Sea. Preprints of the 13th World Congress IFAC, Volume Q.

THE DEVELOPMENT OF PRINCIPLES RELATED TO ARRANGEMENT, PROGRAM AND TECHNICAL REALISATION OF SIMULATOR COMPLEXES FOR SHIPS' CREW TRAINING

Butner E.G., Nikitichev M.M., Novikov V.N.

Abstract. The authors have considered different approaches to arranging simulators for operators' training in relation to shipborne technical facilities control. They present a concept for arranging intellectual simulators. A set of basic tasks is formed and is used in intellectual simulators. The authors suggest ways of simulators' hardware and software realisation.
Keywords. simulation, operators' training, intellectual simulator, database, automated training system, computing complex, software.

Operators of ship's running and technical facilities' control use simulators as a means of training. Simulators have been widely used for a long period of time and have occupied an important place in the system of training complexes. Along with the complication of shipborne technology and an increase of responsibility for ships' fail-safe operation (with nuclear propulsion plants in particular), simulators require improvement both in technical equipment and in methodological aspects of crew training.

Previously, the simulators were used as a training means for receiving initial skills related to control of technical facilities, maintaining and improving operators' qualification for their employment in existing ships. Lately, simulators are spreading, in parallel with ships' design and development, faster than its initial introduction into operation. In this situation, there is no data about field tests and operation parameters, so it is necessary to pay special attention to attaining trustworthy design database, which is used as a basis for arranging simulators of technical facilities. Insufficient correctness of simulators can lead to so called negative training. As a result, operators can attain erroneous skills for situation control. In order to rise the grade of simulation trustworthiness, the designers as a rule use the following methods for attaining initial data:

- data gathering directly from reference installation or stand models of simulated objects;
- data reception by using the method of engineering analysis of theoretical calculations, which precede the design of shipborne organic technical facilities;
- data receipt from similar existing models;
- specification of parametric dependencies through experts' appraisal, who have considerable useful experience in controlling similar systems.

Along with the employment of various methods in order to provide quality of simulation, it is necessary to make a provision for specification of the simulator. It is achieved by simulator arrangement in accordance with the principles of an open system and employment of dynamic training methods (performance-based training). Performance-based training is based on the analysis of operation and assignment of modes, related to training targets, knowledge, skills and capabilities, which are necessary for these targets' achievement.

The training provides operators' comprehensive preparation, including elaboration of human engineering skills, or on the contrary, the training is limited by the tasks of intellectual preparation. Depending on training targets, simulators are being arranged with an appropriate set of equipment. In the first case, the designers pay considerable attention to the similarity of simulation equipment and organic control panels and devices, that is to say, they strive for achieving physical fidelity - similarity of simulation devices, their design outline with appropriate shipborne structures.

The second version mainly deals with operators' intellectual preparation, when the data about the "operating" installation's parameters are mapped on the display. PC peripheral devices are used as controls (keyboard, "mouse" manipulator, joystick, plotter, etc.) Such intellectual simulators (sometimes they are called "intelligible") simulators are widely used lately.

Specific features of intellectual simulators can be clearly seen in the example of simulators, related to nuclear-propulsion plant (NPP) control.

The accumulated experience of nuclear-propulsion plants' design and operation shows, that the quality of decision-making, concerning safety problems and NPP service culture to a great extent depends on experts' qualification and feasibility to "playback" different non-trivial situations of nuclear installation technical facilities' (TF) and control systems' (CS) operation in non-standard modes.

Statistical data of NPP operators' erroneous performance shows, that more than 90% of errors are made at the stage of situation analysis and decision-making. Thus, intellectual simulators, made on the base of PC can be used for skills' training, concerning analysis of problem-solving and selection of the best decision among available ones.

Such simulators can be also used not only for raising the qualification level of NPP service personnel, but also for TF and NPP CS designers, moreover, they can serve as proving grounds for checking the technical parameters of designed equipment.

Intellectual simulators (IS) match the tasks of attaining and strengthening theoretical knowledge and the practical functional skills for NPP control in different standard and emergency modes.

The universal principle of intellectual simulators' arrangement and multi-functional character of tasks being solved, make it possible to use IS both for preparation of operating personnel, which directly carry out NPP control and service functions, and for training the power complexes' command staff, who in this or that degree provide NPP safety and service culture.

Excluding training functions, intellectual simulators can attain control functions, determining the readiness level for work with NPP, accumulation of various statistic materials, dealing with results of training and research.

Intellectual simulators can fulfil a wide variety of functions. The minimum set of basic tasks, which are being solved by IS can be as follows:

- research study of the common terminology and basic principles of antinuclear safety on board the ship, requirements of NPP systems of control, shipborne systems, nuclear safety control systems, systems of NPP technical facilities;
- study of principles, concerning systems and units' arrangement and operation, physics of the processes, which flow in these systems in normal and emergency modes of operation;
- study of technical diagnostics and NPP objects technical state forecasting' methods from nuclear safety point of view;
- formation of problem-solving' analysis skills, assessment of different methods for installation removal on safety level, the selection of the best available solution;
- attainment of skills, needed for problem-solving of designed and out-of-design emergency situations, damage control.

Each of the above-mentioned directions of training can differ based on the rate of task difficulties. It enables the use of IS both for initial training and for the improvement of professional skills, and also for reproduction of NPP operation modes, which can lead to a severe after-effect on practice. Information, available to the trainee, must correspond with the requirements of trustworthiness, availability for perception, completeness, surety and shortness. Thus, when the intellectual simulator is being designed, the most important task is to create the program interface, which will be able to provide the realisation of the above-mentioned requirements.

Specific requirements in intellectual simulator design, relate to the training guidance system (TGS).
The variety of tasks assigned to intellectual simulators requires the provision of feasibility to organise both self-training and training under the guidance of a trainer (training instructor).
The efficiency of training by a trainer is determined by the arrangement of instrumental facilities, which enables the trainer to give away routine operations during the training mode assignment, input of emergency related data, control over trainee's actions, recording of training results.
In order to provide the self-training mode on IS and training process automation there is a system of automated training. Automation of instructor's influences relates to the formation of automated and time successive input of emergency data and the formation of trainees' activities assessment. This direction of research has been sufficiently studied and widely illustrated in related literature. The problem of PC-based automated training systems is less studied.
When the intellectual simulator is used for knowledge level control and certification of trainees' level, it is necessary to have a legal base for the correlation of IS criteria of training assessment and NPP safety requirements and service culture.
This legal base can be shaped on the initial stage of designing, by experts in the field of nuclear power engineering.
Later on, data on statistics will be accumulated and the database will be specified and supplemented. At this stage, it is desirable to link common theoretical NPP parameters with exact shipborne installation, the results of users' testing must be registered in limited access files.
The concept of IS arrangement must envisage the feasibility to expand its hardware-software facilities. It is necessary due to the assignment of extra requirements in the process of combat tasks' expansion and improvement of training methods.
The computing complex (CC) is a central element of the intellectual simulator. CC is arranged as a local network, or incorporates a separate PC, depending on its functional tasks.
A set of IBM-compatible PCs with a sufficient memory size and processing speed can be used as a basis unit.
Arrangement of a PC-based local computing network is reasonable for the realisation of diversified training programs. From the point of technical and economical expediency, this variant of arrangement can be considered only in relation to large training centres.
Intellectual simulator software (SW) must be arranged in compliance with the block-module principle.
In particular, we can recommend the following composition of SW:

1. Program interface, providing self-training of the trainee;
2. Modes control program, which envisages:
   - input of training type;
   - input of initial conditions of the assigned mode;
   - input of standard and emergency data;
   - assignment of emergency data automated input;
3. Analysis and trainees actions' assessment automated system;
4. Objects' imitation models, including situation models of emergency modes' development;
5. Reference database, standard database, database for analysis and trainee's assessment;
6. Graphic data block for monitor mapping.

The following modes can be realised during operation on an intellectual simulator:

- demonstration mode
- teaching mode,
- training mode,

- control mode.

Demonstration mode presupposes employment of a demonstration program and output on monitor display of different related data arrays. In particular, the minimum demonstration batch can include the following sub-programs:
- demonstration of training system functioning - acquaintance of ordinary user with the structure and operation modes of a simulator;
- demonstration of bench chain of actions, changing of controls and devices' modes in assigned mode.

The training mode consists in compliance with the following levels of data saturation:
- passive, when the trainee works in the mode of dialogue with data-reference system;
- active, when the trainee is proposed to "control" the installation (there can be a help message on the display and comments about trainee's activities).

During the training mode, the trainee receives different training tasks for NPP, its elements' control in standard or emergency modes (the training process is accompanied with an errors readout on the display and assessment of his activities).

Besides current control of operator's activities in training mode, the mode of control envisages the mode of theoretical knowledge testing, the ability to adequately react in non-standard situations, understanding of NPP dynamics in different normal and emergency modes. The control mode helps to form a partial and general assessment. The final results are printed.

The tasks of training are reasonable to form in compliance with the "from simple to complex" principle. Simultaneously, the basic set of tasks must include tasks, which provide initial stage of operators' training, data arrays, which correspond to the physics of processes, which flow in NPP and its elements, and also personnel testing tasks and their correspondence to standard requirements of job instructions.

Depending on intellectual simulator employment, the basic set of tasks can be supplemented with monitor, functional and control tasks. These tasks correspond to the higher requirements of the exact type of training and provide achievement of deepened understanding and skills of trainees.

The arrangement of functional and system software must provide expansion of SW with the appropriate enlargement of simulator's hardware by computing means, sufficient for exact simulator.

When the functions of IS expand, we suppose, that giving the features of a self-training system to the simulator is very perspective.

The intellectual simulator enables the user to work in a self-training mode, practically, it is an automated training system (ATS), presenting the dialogue model "instructor-trainee".

Feedback is an essential condition for effective IS-based training. Feedback presupposes control over the training process and presentation of theoretical data and trainee's activities results in a comfortable form. The basis of modern IS and ATS is expert database systems, which use theoretical knowledge of the exact field of activities and knowledge of qualified operators.

During knowledge recovery (in passive mode) the trainee works with a data reference system; during training, (active mode) help messages, analysis and assessment of operator's activities are compared with the standard ones in database.

The standard chain of activities, time intervals for its execution, limit values of processes' parameters are entered into the database beforehand. Then, they are compared with current values, help messages (hints) and activities' assessment is formed.

The structure of related expert systems includes the following main elements:
- database for decision-making regulations;

program facilities to provide dialogue (program interface);
- interpreter (program), which provides analysis of trainee's activities, comparison with the standard and result forming;
- models of object and control system simulation.

The functional system of the IS expert system is given in fig. 1 (self-training mode).
Figure 2 gives an example of one variant for expert system main menu realisation.

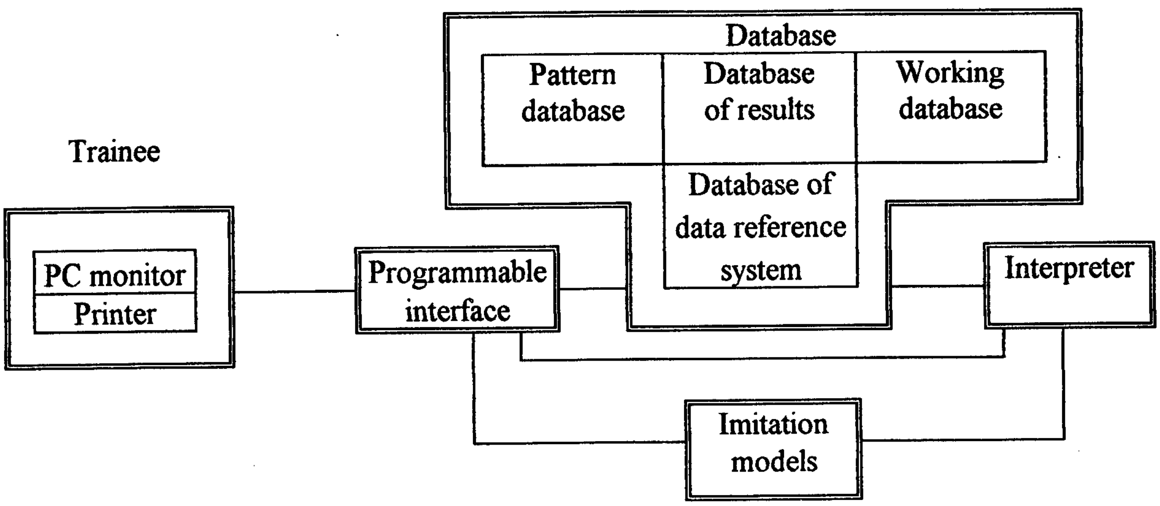


Fig. 1 Functional scheme of IS expert system

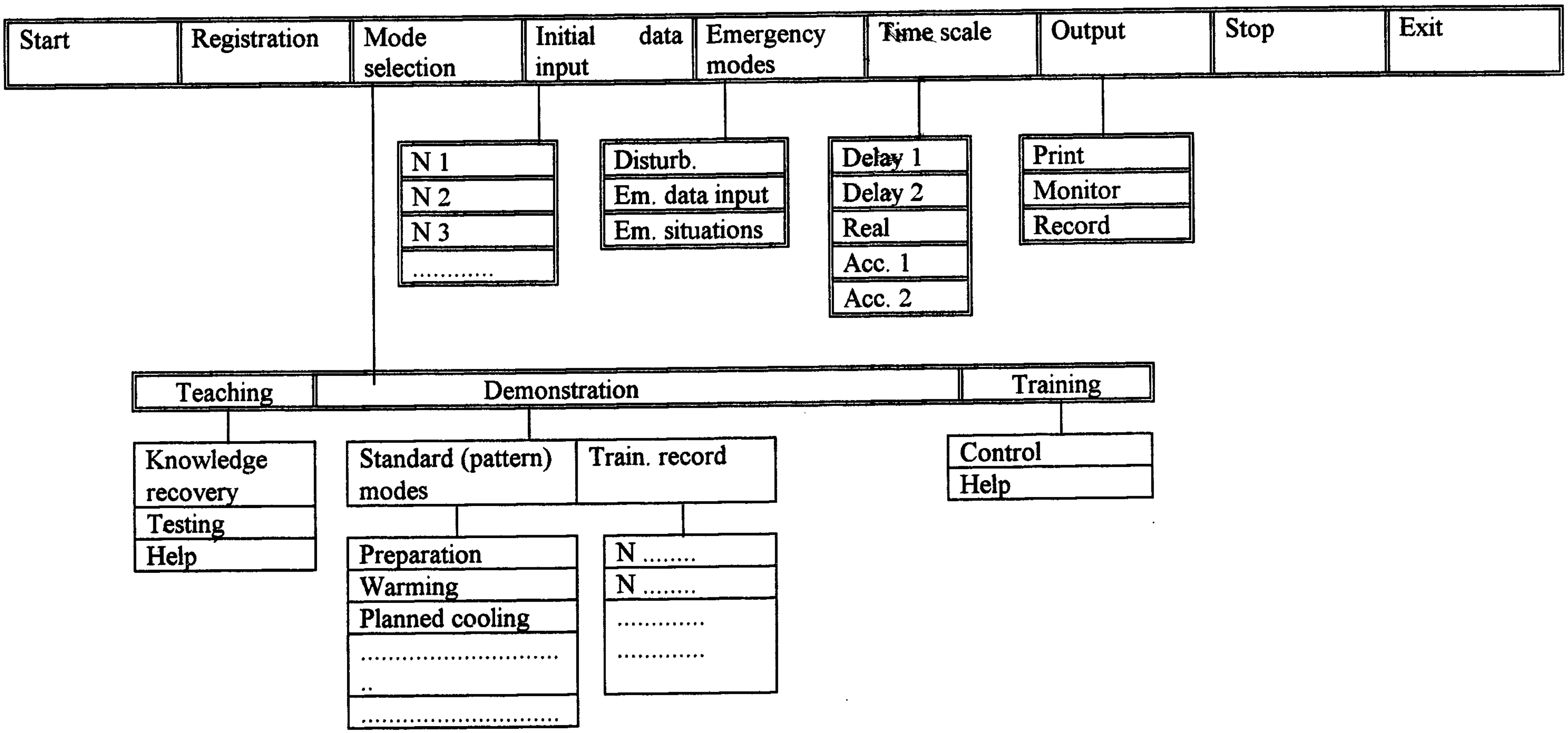


Fig. 2 The structure of IS expert system related main menu

SURVIVABILITY FIGHT SUPPORT INTEGRATED EXPERT SYSTEM FOR WARSHIP COMMAND COMPLEX

Anatoly I. Indeycev, Alexey G. Sergeyev
"AURORA" Concern
Institute of Problems of Mechanical Engineering (Russian Academy of Sciences)
St. Petersburg, e-mail: alex@math.ipme.ru

Abstract. An integrated expert system for warship captain decision making support is described. The system is based on a formalized description of the damaged warship behavior and the set of typical evaluations and recommendations. The expert system involves the sets of primary and secondary factors, the set of evaluations of warship state, the set of recommendations to the captain, the set of knowledge bases, the interface procedures, etc.

Keywords. Expert systems, navy, warship control, captain decision making support.

The Russian admiral Sergey Makarov defined a specific feature of warship — "survivability", which became an integral part of warship combat readiness' common evaluation during combat (emergency) damaging. The remaining "survivability endurance" is a potential of its operating efficiency. For that reason, survivability as any other feature, related to warship quality must be controlled and preserved. Achievement of high functional stability and information support of all warship's sub-systems is an indispensable condition for its execution.

One of the critical directions in performing of these requirements to survivability is introduction of new information technologies, stipulating the development of automatic collecting and about combat (emergency) damages data processing systems as well as the automation of working out recommendations and decisions at the fight for survivability of warship.

Introduction of new information and intellectual technologies is one of the decisive trends in order to meet the survivability requirements. These technologies envisage the development of automated systems, related to the gathering and processing of combat (emergency) damages, presentation of recommendations for survivability control decision-making.

The concept of such systems' forming is determined by modern views on the dynamics of emergency development (after-effects of combat damage). Both home and foreign experience on this problem says, that prompt warship wreck very seldom occurs after the first damage, as a rule, it comes in the case of a direct underwater explosion in the middle portion of the warship, ammunition detonation in space below the waterline (WL), which leads to the collapsing destruction and flooding of a warship's compartments below the waterline, more than 20% of its length.

The final result of damage control (DC) depends on the opportunity and efficiency of warship captain decision-making for DC strategy and tactics' selection, psychological and professional training, and to a large extent from the operating and hydrometeorological situation in the emergency area.

The analysis of any major warship wreck testifies, that there is a considerable gap between the remaining DC strength and facilities after the emergence (combat damage) and the demand for them , which appears during development of secondary damages, floods and fires. The spontaneous development of secondary damages determines the main difficulties in the organization of DC:

— the character and the extent of accident (combat damage) is never known;

— as a rule, the decision (first, in particular) on the localization of damages, propagation of fire and water on the warship is made on the base of incorrect information and with a considerable delay in time;

— the guidance of fighting against water and fire is made through a method of trial and error, very often without analysis at the place of the accident;

— normally, the situation is complicated by the occurrence of several locations of the emergency situation in different compartments of the ship.

The analysis of these difficulties gives ground to make a conclusion, that delays in receiving of timely and complete information from the scene of action then entail the issue of delayed decisions for reconnaissance of strength and facilities in the new places of emergency. It makes it impossible to prevent the development of secondary damages. Thus, reduction of time response for information about initial damage and the increase of information authenticity becomes a very urgent task.

At the same time, modern research studies dealing with the reasons of damage control' ineffectiveness in emergency situations on the ships of home and foreign navy makes it possible to make a conclusion, that the factors, stipulated by various aspects of operator' (captain) intellectual activities are rising.

Explanation of incompetence to localize this or that emergency situation is determined by the above-mentioned factors and individual psychophysiological features of each operator. The influence of these reasons can be diminished by the introduction of new information technologies, based on the employment of integral simulation methods, methods and means of artificial intellect (AI).

The authors of this study suggest an alternative method for the realization of this technology as a system of intellectual support (SIS), which relates to the class of expert systems (ES) [1, 2]. The necessity of the warship captain in SIS is determined by the following factors:

— character of control tasks, which are being solved for damage control and specific heuristic methods for its solution, based on the previous experience and specific features of a current emergency situation;

— complicated dynamic of processes going on during an emergency and in the ship's control systems along with continuous development and interference of the set of factors, complication of control processes information simulations, wide range of tasks, including interpretation of ship's condition and forecast of emergency situation development in its sub-systems, necessity to plan the outcome of the emergency situation by using the frames routine.

Nowadays, information independent SIS is realized. It functions independently and has no connection with the transducers of controlled objects. Information about surface ship (SS) and its sub-systems' condition is entered directly by the operator, according to reports from the command post (CP) and ship's action stations [3].

The main characteristic of ES-based SIS is its database, (DB). If DB is correctly structured, its merit can be defined as an ability to present its content evidently, and organize data invocation in such a way, that decision-making will be simplified.

The efficiency of such a system depends on knowledge quality, which saturates DB in the frame of this or that problem, expert's experience and skills in the exact field of application and also completeness and consistency of database (DB) about the object (ship). The feasibility of DB and action stations (AS) expansion becomes very urgent. It means further improvement of the system, along with the acquisition of new experience, specification of data related to the object and wishes of the customer. Besides it, SIS possess the following:

— wide capabilities for forecasting. This is vital for the development of a methodological basis to forecast probable survivability of the warship in different emergency situations;

— capability to specify undefined initial data according to the sequence of its entry, showing, how the decision must be changed in a new situation;

— readiness to give a detailed explanation of how a new situation leads to changes due to the influence of newly developed and previously undetermined factors;

— competence, which is determined by its new feature, specific for expert system (ES) — institutional memory, which is understood in this study as a set of knowledge, developed in coordination with grounded experience criteria, checked by TF long-term operation, research in an appropriate field of science and experience of training and development of emergency situations for ship's SC exercise and other tasks of combat training.

Such a set of knowledge is shaped into a summary of qualified judgments and expert evaluations, with a continuous renewal reference of the best strategy and tactics in the selection of surface ship's SC decision.

The basis of ES and its DB' demonstration prototype constitutes the chains of discussions about probable versions of different combat and emergency damages, when combined in this or that emergency situation can wreck the ship, or threaten to wreck the ship and cause complete or partial loss of fighting capacity.

The system incorporates the elements of "explanation routine" to make evaluation of changing the ship's condition. It is obviously shown, why the way was lost and de-energizing took place, why the reserve of buoyancy and metacentric height achieved critical limits, why the danger originated in ammunition spaces' (AS) detonation and other critical moments in the conditions of the emergency situation. Timely employment of this system's feature is one of the most important factor, which helps to prevent "human error" during SC decision-making. These errors are practically inevitable in the stress conditions of both the ship's captain and the crew, in the first moment of the harmful effects' impact (emergency situation). That's why, confidence appears and the system is considered as an "adviser", providing more calm perception of information and speeding up the process of analysis and decision-making.

The general concept of SC decision forming presupposes, that simultaneous activities for water, fire and TF SC are not feasible. It is necessary to determine the priority decisions, which define the main directions, vital for the ship's survivability. Measures against the ship's flooding and maintaining of combat stability are the keys to damage control. The captain of the ship must be sure, that crew members know what is necessary to report to the central control station (CCS) or power and damage control center (PDCC) at the first moment of an emergency, and he must understand the sequence of crew activities related to the captain's decision.

During decision-making, SIS helps the captain to sufficiently consider the set of the most important factors, related to the provision of decision optimization and also, helps to avoid major errors. Recommendations from database for each simulation of SIS information frame, simultaneously play the role of psychophysical stabilization of operator's heuristic. It gives an additional positive effect to the possible result of decision-making.

The system has a set of so called "demonstration scenarios" of emergency situations, compiled on the basis of well-known emergencies, combat and emergency damages, which took place in home and foreign navies.

The basic software of the expert system

The software was developed as an integrated expert system. In compliance with the assigned task for receiving ship condition statistical expert evaluation, the system includes the following valuable information elements:

— indexed set of initial and secondary factors;

— a set of ship condition evaluations;

— a set of recommendations to the captain;

— a procedure of correlation related to initial factors' values;

— databases in the form of sets of regulations for receiving expert evaluations, related to secondary factors and recommendations;

— the main knowledge database for the evaluation of ship's condition and corresponding recommendations;

— procedures for interface arrangement (acquisition of initial factors' values, output of evaluations and correlated initial factors);

— reset procedures of factors' values and evaluations to initial (most favorable) state, additional procedures for menu arrangement, installation of variable environment, etc.;

— a set of demonstration scenarios in the form of input data description and its corresponding fixed sets of initial factors' values;

— a set of reference information displays (on damages volume forecasting, minimum of information required for captain, examples of ship condition forecasting arrangement on the basis of specific and non-specific occasions of scuttling, related to different ship's design, etc.);

— procedures to organize interaction with other programs (for survey of layouts and cross sections related to ships' separate designs, dynamic simulation of wrecked ship behavior, schemes of scuttling , etc.).

The term "factor" here is understood as an event, which is vitally important for ship condition evaluation, (for example, the flooding of some compartment,) or event characteristics (rate of water filtering: fast — slow, type of scuttling: collapsing-multi-deck, etc.). Initial factors mean the factors, which values' input is made by the operator on the basis of incoming reports, visual observations, devices' readings, etc. Secondary factors mean the factors, which are determined by the system itself, on the basis of initial factors.

The above-defined array is loaded in the memory at the beginning of system's operation. Initial factors' values are interconnected by correlation dependencies, stored in a special procedure file. This procedure can be considered as an object-ship simulator.

On the basis of initial factors' condition, the state of the secondary factors is determined. The execution of this task is made by a corresponding sets of regulations.

There is a targeted variable in the main knowledge database — the evaluation of ship's condition in actual form (symbol variable). Simultaneously, an appropriate recommendation corresponds each evaluation, also in actual form. Totally, there are six global evaluations and corresponding recommendations:

— *the ship in the condition of wreck/rescue of the crew;*

— *threat of ship's wreck/to determine and localize the factor, leading to the wreck;*

— *the ship's combat efficiency is lost;*

— the ship suffered a damage, but is able to recover and maintain combat efficiency/damage control after damage;

— fire safety and/or readiness to exert resistance to flooding are reduced/ acceptance of measures for recovery of damaged systems' operation;

— the ship is capable of withstanding combat or emergency damages/ acceptance of measures to provide survivability before damage occurs.

Some sets of regulations, in order to determine values of secondary factors, have valuable evaluations for the captain, and concern the ship's condition. Thus, for example, the set of regulations for analysis of radioactive danger in the list of evaluations adds the additional values: "Radioactive danger" and "Threat of radioactive danger". Recommendations, which are elaborated during determination of secondary factors (e.g., " Slow down", "Go to echelons partial operation", etc.) are also added to the main set of recommendations.

This system also determines "Ship condition forecast", which depends on the state of crucial (for exact damages) systems. Thus, in the case of loosing this or that grade of unsinkability, the crucial point is the condition of bilge-pumping, water-sluicing and ballast-pumping systems. When the ship's running is in disorder, the main provision for forecast receiving is the state of such system, as central circulation system, the condition of stuffing boxes, etc., depending on main propulsion plant (MPP) type.

Generalized algorithm of the system's operation is given on Fig. 1.

Conclusion

The defined damage control information support command system represents a prototype, designed for the demonstration of capabilities related to technology of intellectual software — expert systems to solve tasks, which arise in front of the captain, when the warship received combat or emergency damages. Databases, a set of initial factors were developed in relation to a whole class of Navy combat facilities — surface ships and thus, can be considered generalized.

In the process of the system's design, the guiding materials of the Russian Navy, data from publications on foreign navies, accidents, which were described in scientific literature and also one author's personal experience of Navy service have been used.

The integrated expert system for the decision making support at the fight for survivability allows to decide a complex of problems both in everyday conditions of usage, and in combat conditions.

The following problems are decided in everyday conditions:

— checking for captains order of fight for survivability readiness of the ship control during the preparation to a battle and sail;

— checking for security arrangements and conservation of warship survivability on the motion under different conditions of readiness;

— working out of recommendations on checking and navigability conservation in complex conditions of sail (storms, at the sail in tropics and in north latitudes);

— unceasing checking of engine modes quality, that is an essential condition for fighting efficiency optimization during long sails.

The problems decided by the system during war period are:

— collection, processing and presentation information about a mature and volume of damage;

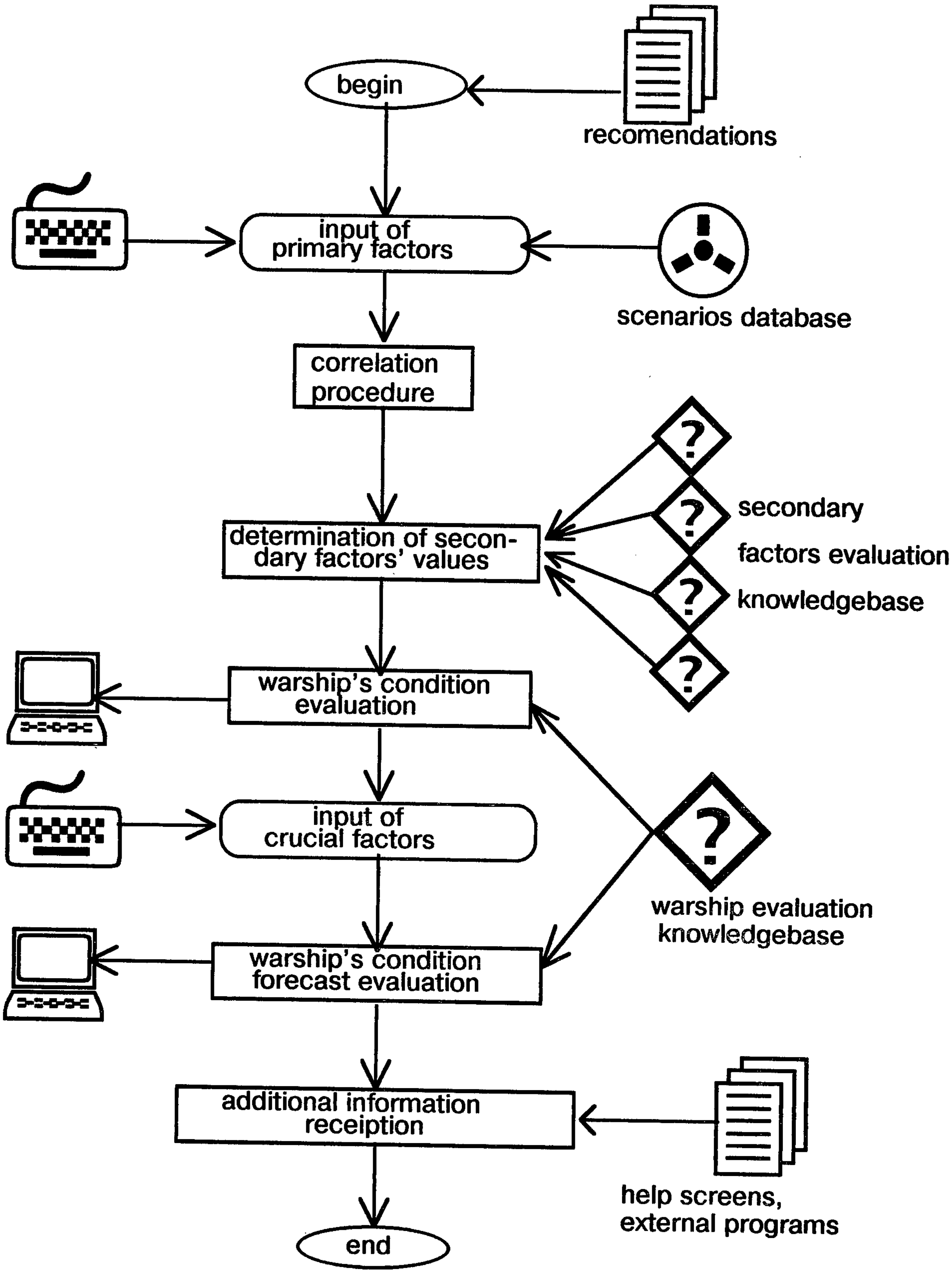


Fig. 1. Generalised algorithm of system's operation

— analysis of damage forecast models;

— forecast of changing a warship state in cases of changing a possibility stayed in the formation of power and fight facilities for survivability;

— recommendations about the fight for survivability at the development of events during the fight with flood, struggles with fires, losses of motion and provision by the electric power of combat consumers.

The results received during the SIS pilot sample check-out revealed a number of problems, connected with its adaptation to the captain's control activities as an operator. The essence of these problems can be briefly summed up in the necessity for:

— the revision of accepted patterns and content of reports from CP, damage-control parties and action stations related to damages, in order to achieve its conformity with the advised frame of description model;

— the elaboration of new requirements to the accuracy of reports presented by damage-control parties, i.e., the volume, character of the initial damage. It is necessary to consider the expansion of existing means to examine the damage area;

— the research of new concepts on employment of control panels for MPP, EES, OCS centralized systems, to achieve earlier determination of emergency situation "mature", which can lead to loss of operation efficiency of a sub-system of ship's power plant.

References

1. Waterman D. Expert systems. Moscow, Mir, 1989 (in Russian).

2. Kobzev V.V., Mironenko G.M., Shilov V.A. Navy expert systems. St. Petersburg, Dzerzhinsky VVMIU, 1993 (in Russian).

3. Indeycev A.I., Sergeyev A.G. The survivability fight support intellegent system//The methods and means of information support of warship survivability fight. St. Petersburg, IPMash RAN, 1995, p.p. 15...35 (in Russian).

INFORMATION SUPPORT SYSTEMS FOR WARSHIPS' COMMAND PERSONNEL DECISION-MAKING

Volsky V.E., Kirshner A.L., Kurochkin A.I., Naumov M.V.
"Aurora" Scientific-Production Association of
Ministry of Defence Industry of Russian Federation
15, Karbysheva Street, 194021, St.Petersburg, Russia
Telephone: +7(812)316-82-95

Abstract. The authors have considered the tested approach towards information support systems' algorithmical provision related to ships and vessels' command personnel decision-making in difficult weather conditions and damage control.

Keywords. Information support, decision-making, damage control, emergencies.

The analysis of emergencies and wrecks shows, that a considerable part of them is linked with errors made by command personnel when decisions on ship's control are made in difficult weather conditions and during damage control (passing through narrow passage waters, divergence, leakage, fires, etc.). Thus, lately, in home and foreign shipbuilding a new trend in automation of shipborne technical facilities has appeared and is intensively developing (TF), based on employment of modern computing techniques and information technologies. It is directed for the elimination of errors, made by decision-makers (DM) in such situations of ship's control. A new class of information-computing systems has appeared, which provides intellectual and information support for DM. The most widely used name for such systems is "information support systems" (ISS).

The main task of ISS is to provide DM with information, which capacity and form enable him to make timely and valid decisions for TF control, activities of the crew during damage control (DC), in different emergency situations, including advice on automated forming, taking into account the real technical state of TF and real development of an emergency.

Initially, theoretical studies in the field of DC ISS were aimed at employment of expert systems (ES) principles. In its classic form, ES includes:
- database, which is formed with consideration of the subjective experience of experts in their specific field; at the same time, data in database are stored in the form of rules, which are presented as productions, frames or networks;
- interpreter (logic output device), which serves for problem solution reception as advice to the user;
- the block of knowledge acquisition, serving for the automation of knowledge input in the system. It is made by the user-expert;
- linguistic processor (interface), providing dialogue interaction of user with the system;
- the block of explanations, designed for elucidating, why the system arrived at this or that decision;
- operating memory, which serves to store data and facts, needed for current task solving.

Real practice of command personnel ISS development realises a slightly different approach, which is aimed at the provision of requirements concerning the search rate and information mapping convenience for DM in the process of emergency development. ISS information and algorithmical base, which is realised within this

approach, can be used for class systems development up to the level of expert system (ES). Below, the authors give a brief description of the main aspects of this approach, based on generalised experience in the field of shipborne ISS design and development.
ISS are developed for the most complicated and frequently met emergencies, the number of which is fixed for each exact system. Depending on ship's (vessel) type and assignment, these emergencies may be as follows :
- fires in ship (vessel) compartments and spaces;
- water access to the compartments;
- ruptures of high pressure air manifolds;
- access of a large volume of steam in compartments and rooms;
- severe leaks and ruptures of nuclear power plant first circuit;
- de-energising of TF;
- short circuits in the main electric power equipment;
- breaks of hydraulics system manifolds;
- complex emergencies, connected with the impact of several effects of a different physical nature (fire and flooding, etc.).

The basis for information support development is:
- existing naval guiding and regulation documents (The rules of the Register, Manuals, Manual of damage control (MDC), etc.), which constitutes a legal basis during decision-making related to ship control and crew activities;
- shipborne project and maintenance documentation (key diagrams, equipment compartment layout, systems and equipment maintenance and service manuals), which serves as a knowledge source about ship (vessel) and its TF;
- mathematical simulators (of fire, flooding, etc.), designed for emergency development evaluation.

The ISS distinguishing feature is necessary to provide a multi-task operation mode in a real-time scale. First of all, it is connected with the necessity to calculate several simulators at a time; for example, when a submarine encounters a complicated emergency in a submerged condition it is necessary not only to elaborate emergency fight recommendations, but to solve the task for its surface exposure, to calculate the simulation of genetic relations for forecasted possible deterioration of the ship's technical characteristics.

Besides provision of DC modes, lately it has become evident, that there is a necessity to use these systems in standard modes of operations, such as:
- berth on the base;
- anchorage, mooring, re-mooring;
- navigation in difficult navigational conditions (passing of narrow passages, fog, ice conditions, etc.)

And finally, the presence of adjustable models in ISS algorithmical support enables the use of these systems as shipborne simulators.

Depending on the ISS introduction stage (construction, repair, modernisation), and also on the TF automation level and availability of computing means, the systems can be divided into the following two modifications:
- central co-ordinating control systems (CCCS), designed on the base of microprocessor means and PCs, with the employment of network hardware and software, which provides automatic data input from ship (vessel) functional and control systems, incorporated in CCS TF;

- ISS with manual or semi-automatic data input, installed on existing ships and vessels, the employment of which was not envisaged initially. In this case, the system is considered as a special-purpose PC, incorporated into a special-purpose structure.

The main principles of DM information support and DC ISS provision are:
1. The hierarchy principle of IS arrangement, when each lower level receives more detailed and more specialised information. We must consider at least three hierarchy levels:
- generalised information on object's current situation;
- DM behaviour scenarios in situations being considered;
- DM information support (IS), envisaged by scenarios.

Depending on object type and situation, additional (intermediate) levels of hierarchy can be used. For example, DM IS (the third main level of IS) can include two sub-levels.

Generalised information indicates type, place and time of emergency origination. It is formed by signals, coming from transducers and alarm warning devices (flame, smoke, water, radiation, etc.) or through reports of personnel. It is mapped, in order, on a plotting board, with the readout on the display panel. The readout of the current parameter values, which characterise the mode of operation and the hull position: cruise speed, capacity of the main propulsion plant, roll, trim difference, draught, etc., also is presented on the display.

Behaviour scenarios are shaped on the base of naval regulating documentation and determine the sequence and direction of DC; simultaneously, recommendations for the alteration of activities' sequence are formed on the basis of control and forecasting results. Thus, scenarios "control" the emergency situation and are automatically adjusted along with its alteration.

IS of DM activities in accordance with scenarios in each envisaged emergency situation is developed on the base of shipborne design documentation and experience of experts in the field of maintenance and DC. It contains various reference design information related to the object and its TF, current values of controlled parameters, recommendations on TF methods of employment and other related information, which is necessary for DM's understanding of recommendations, formed by the system and decision-making of activities, envisaged by scenarios.

2. IS volume and composition optimisation, which stipulates automatic "dosing" of the system information in any exact situation, i.e. information, which according to the opinion of experts is necessary for DM in this exact situation, is automatically included in IS composition, allowing the possibility to receive any interesting extra information on call. Thus, when the layout of emergency spacing is included in DM IS composition during execution of activities for personnel recovery from damaged spacing in case of fire, the plan includes information about the layout of the main large-sized equipment and the travel line and has no information, for example, about fire extinguishing equipment; at the same time, on the contrary, during the execution of activities for fire extinguishing, the plan of this deck includes information about fire extinguishers' location, but has no information about a travel line.

3. The employment of the main expressive and laconic forms and methods of information presentation to the DM, which provide higher rates and ease of

information perception: colour, light and sound encoding, a wider use of pictograms and other graphic forms of data presentation instead of text form, etc.

4. Minimisation of the number of manual operations during DM and ISS interaction, including:
- max. available replacement of data manual input by automatic (from transducers or simulators) or semi-automatic (with the help of the menu);
- optimisation of the ratio between the number of ISS controls and the number of DM manual operations during his interaction with the system for task execution.

We can show, that the described approach towards development of shipborne ISS can also be used for designing similar systems related to complex industrial objects (heat producing and nuclear power plants, chemical plants, drilling platforms, etc.).

NEUROINFORMATION AND FUZZY TECHNOLOGIES IN SHIP CONTROL SYSTEMS

Gennady S. Osipov
Institute of transport problems, RAS
199178, 12 line, #13, St. Petersburg, Russia

Abstract. The bases of construction and usage of ship's neuro-intellectual control systems and systems with fuzzy logic are considered.

Keywords. A marine vessel, control systems, neural networks, fuzzy logic.

Introduction. A vessel is a maritime dynamic object, operated in constantly varied and a priori unknown external effects. The main assignment of the vessel is satisfaction of needs in maritime transportation under safety of navigation and efficient transportation process. However, according to statistics of breakdowns and accidents, a tendency towards decrease of safety level at sea is observed.

Direct dependence of safety of navigation on quality of control is revealed, therefore an actual problem is remained in construction of systems for ship's and her technological processes control, the systems that provide a high quality of control under sufficiently rigid requirements to time of decision making. It is quit necessary to supply intellectual information support system to a navigator, assisting him to make the operative decisions in conditions of uncertainty and inexact initial data. The most important problem is an increase of information support reliability in emergency situations, when the ship's crew life and ecological safety of environment are deeply dependent on a speed and quality of control effects.

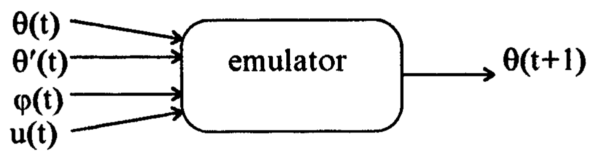
Shortening the time of generation and increasing control effects quality by means of ship's neuro-intellectual control systems and systems with fuzzy logic is a solution of problems in safe navigation maintenance.

The principles of construction. Ship's dynamics in rather general way can be submitted as follows:

$$\theta(t) = \int_0^t \int_0^t f\,(u(s), \theta(s), \theta'(s), \varphi(s))dsds,$$

where θ- vector of state, u - vector of control, φ- disturbing effect

Let us construct a neural network control system - three layer emulator, performing in a look-ahead mode.



For training of such a network, it is necessary to define a hypercube of variables change ranges ($[0,1]^4$), quantity of neurons in an intermediate layer, number of trainee images and allowable error of training.

On Fig. 1 results of comparison of regular governor - ship's rolling stabiliser and neuro-governor with 20 intermediate neurons are indicated, trained on 100 random generated images at allowable goal error 0.01 (training of a network was carried out by gradient method). Here θ - current value of rolling angle, degrees; u - torque, created by rolling stabilisers, measured in degrees of list angle; φ - effective angle of wave slope, dependent of waving intensity, ship's speed, and her course angle in relation to waves combs, degrees. The governor generating control law $u=-\theta$ was used as a ordinary one, that was provided by modelling of an initial equation of dynamics.

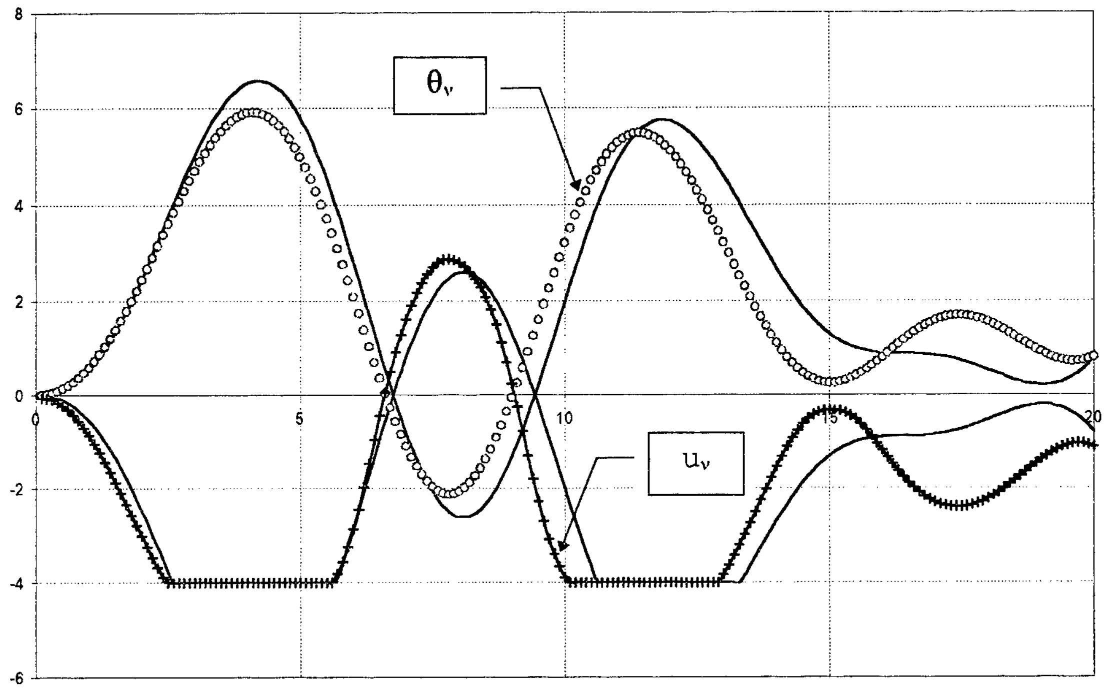


Fig. 1 Neuro-emulator and ordinary governor at regular disturbance.

Trained neuro-governor - three layer emulator provided control of the object in look-ahead mode according to the quantity of rolling angle deviation. To produce a proper control effect, an inlet signal was supplied to a network $\theta=\theta(t)$, $\theta'=\theta'(t)$, $\varphi=\varphi(t)$ and $u=0$. Further, in relation with response $\theta(t+1)$, the control effect u_v, was determined.

The emulator provides quality of control, in general case, not worse than ordinary governor, the response to zero control effect is received by means of a network, which is trained and stores a system's dynamics in it self. It permits to receive a decision in general case faster than solving an initial equation of unknown structure dynamics. The emulator - in it's idea is a model of object, that does not need, in a priori, to be trained by images $(\theta, \theta', \varphi)_t \rightarrow u_t$. The emulator is rather quickly trained and stores the object's dynamics, however it is necessary to set dependence of control effect quantity from adjusted parameter. In a considered example, there was accepted $u=-\Delta\theta$, where $\Delta\theta$ - is value of the list angle deviation from the setting at the next step and at zero control effect.

In general case, the problem of control effect formation in relation to the (deviation) quantity of governed parameter can be not trivial. Cardinal exit from such situation is use of self-organising networks. Let's construct an a self-organising control system of ship's rolling stabiliser on a base of three dimensional lattice, which every point is associated with inlet and outlet weight vectors: $w^{in}=(\theta, \theta', \varphi)$ and $w^{out}=u$.

For an example, a lattice 10^3 was taken and training was carried out on an evenly distributed random quantities (from 10^3). In this case, the training was based on a principle of competitive struggle between lattice points for the right of response to an inlet signal. As a measure of nearness of inlet signal and inlet weight vector, the Euclidean distance was taken for a search of the winner. The trained in this way a self-organising network is capable to produce it's own control strategy, and quality of such control is comparable to realisations of previously trained emulator, working at the look-ahead mode.

For considered example of rolling stabiliser control system there is a well proved practice of optimum governor construction, providing the process of governing based on rolling angle speed change, that equals the damping torque change. On a Fig. 2, system's dynamics is submitted at control of optimum governor, providing the control law $u=-2.1\theta'$ (with derivative noised up to 5%) in comparison with control, produced by self-organising network, trained with goal function $-|\theta_v'|\rightarrow\max$.

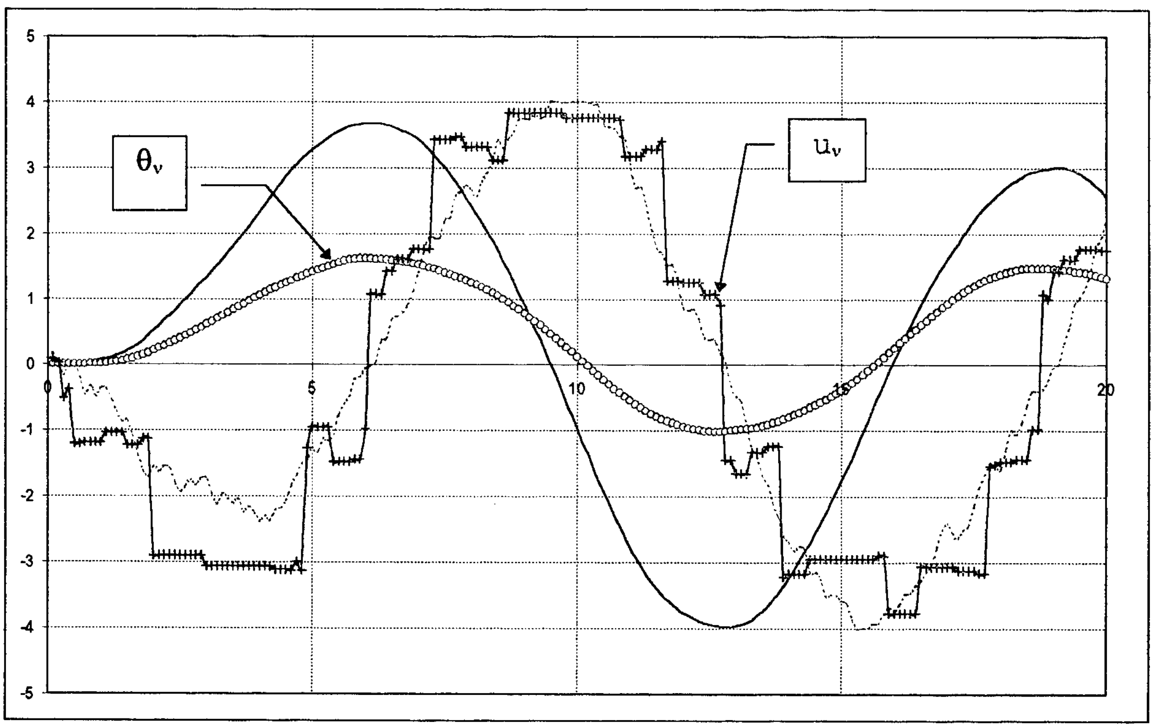


Fig. 2 Optimum governor with an error in derivative of 5% and self-organising system.

Sharp jumps in a signal, generated by self-organising system (transition occurs from one cluster and winner to another) are explained by the fact that in the given elementary case the response of only (one) victor was taken as a control effect. Taking the weighted responses of the winner and his neighbors, it is possible to smooth this function. We shall once again emphasise, that self-organising networks do not require preliminary training on examples - beforehand prepared pairs: a system current condition - required control.

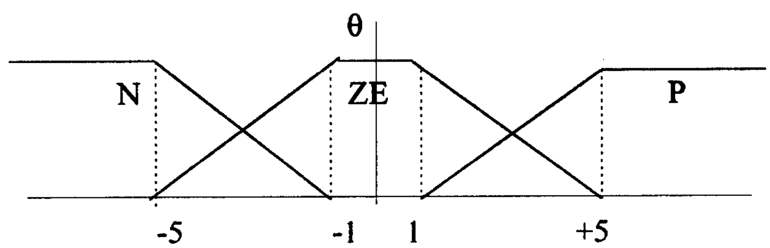
If neural networks are the structural model of nervous system, the systems with fuzzy logic simulate process of thinking. Such systems are simple in organisation, usually require smaller labour input in construction, if to compare with conventional variants, allow to provide a laws of control, based of human creature reasoning like: if the rolling is too sharp, it is necessary to decrease initial stability..

The problem was researched about synthesis of the system with fuzzy logic of stabilisation parameter θ at limitation on control u and random disturbances φ. The control system should generate regulation control effect at the excess of some of the threshold outlet value of the fuzzy logic system. The rule of control system behaviour with fuzzy logic can be given with matrix:

	θ			
		N	ZE	P
θ′	N	PL	PS	NL
	ZE	PS	ZE	NS
	P	ZE	NS	NL

The idea of the components of such associative matrix is rather obvious: so, if parameter θ is positive (P) and derivative θ′ - speed of change is positive also, it is necessary to apply larger negative (NL) control effect and etc. Associative matrixes (in general - multi dimensional) for the other state vector components are constructed by analogy.

Membership functions were set in a kind:



On the Fig. 3 a fragment of performance of the elementary stabiliser with fuzzy logic is submitted. It is accepted, that control effect can have only three values -1, 0 and +1, and disturbance -2, 0 and +2. We shall note, that control system with fuzzy logic could be easily constructed and modified, and despite a very simple principle of control effect generation, provides a rather high quality of governing. In a considered example, membership functions were chosen “manually”, and, in general, they can be optimised with the help, for example, of genetic algorithm.

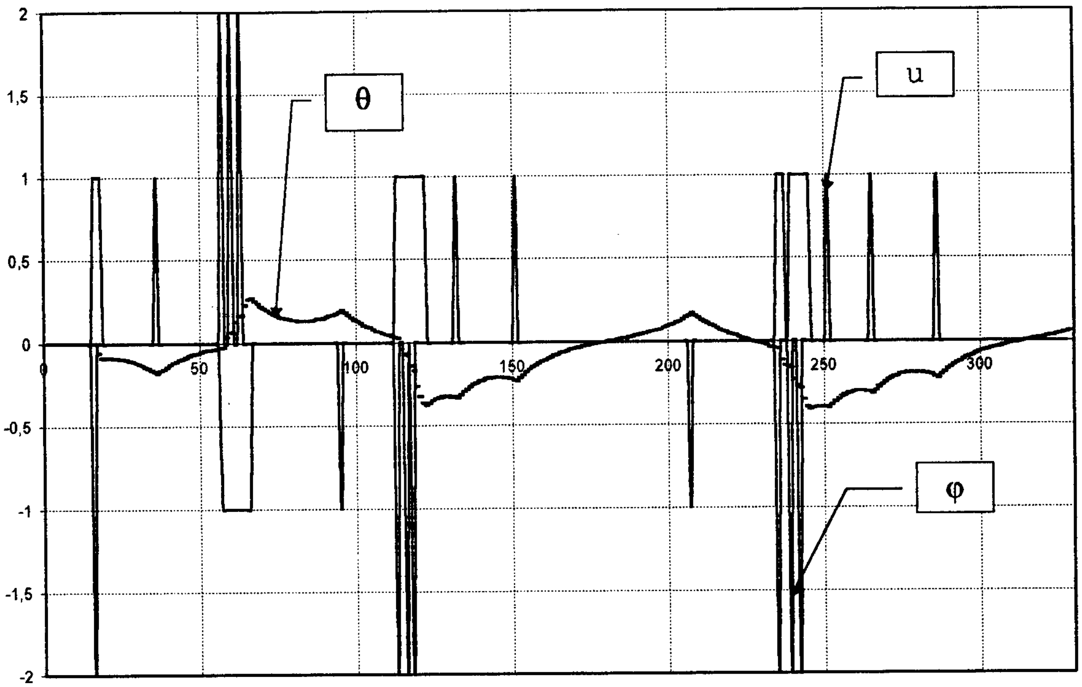


Fig. 3 Systems of stabilisation with fuzzy logic.

Conclusions. Thus, in ship's control systems operated at normal conditions, is the most advisable to use elementary and quickly trained neural network control systems, constructed by a principle of the emulator, performing in look-ahead mode. The parallel neural network emulator structure provides high speed of reception of response on trial control, the reaction on which is used for generation of appropriate control effect. Where a priori, it is impossible to form the law of control effect quantity dependence from reaction to trial signal, the emulator has to be connected to neural network governor.

Preliminarily trained multi-layer neural network control systems and systems with fuzzy logic can provide the highest quality of control. However, it takes much more time to train multi-layer multi-network systems than one emulator, and determination of structure of system with fuzzy logic (membership function etc.) in complex systems can appear as a not trivial problem. Simultaneously, practice of application of systems with fuzzy logic in passenger train control, machinery brake systems and etc., testifies the capability to provide qualitative (smooth) control.

The self-organising control systems are irreplaceable in extreme, non typical situations, when strategy of control is not prepared beforehand. A network, trained with reward function to help vessel out of area of inadmissible states, starts "immediately" to generate control effects, appropriate to goal setting.

THE SOLUTION OF A RECIPROCAL RECOGNITION PROBLEM AS A METHOD FOR CONTROL THE INFORMATION PATTERN OF MARINE OBJECTS

Beskid P.P., Rogalev Y.N. and Sajdun Mefty
St. Petersburg Electrotechnical University
Department of Marine Information Radioelectronic Systems
5, Popova st.
St. Petersburg 197376, Russia
e-mail: rogalev@mirs.etu.ru

Abstract. This paper considers a problem of control the information pattern of marine objects. For this purpose a reciprocal recognition problem is decided. To find the unique solution of a reciprocal recognition problem a numerical method is used. The block diagram illustrates the solution of this method. Mathematical expressions and graphic charts are given.

Keywords. Pattern recognition, optimization problem, resolution, unique solution.

Introduction. At the decision of an optimum organization of counter flows of vessels on the rivers with system of drawbridges, on internal channels with system of sluices, on waterways (sea channels) and in straits with narrows there is the necessity of recognition (identification) vessels. The efficiency of a recognition problem with the help of known means is much increased, if to execute control of an information object pattern. The information pattern is understood as distribution of measured object parameter values on spatial coordinates - φ_1 , φ_2 , Fig. 1. The measured parameter values $x_{i,j}$ are statistical. These values are subordinated under some distribution law of random variables. The set of actions, directed on change of an information object pattern, makes a procedure of control by an information pattern. The control of an information object pattern solves a number of problems, among which: allocation of object among similar, masking object under assigned object, elimination of differences between observable objects and etc. We shall hereinafter consider a case of object discrimination among similar.

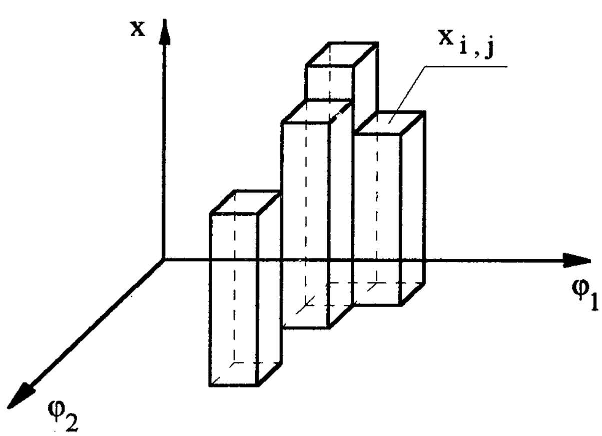


Fig. 1 Representation of an information object pattern

Environmental Impact. We allow there is some number of the same type objects A_k , $k = \overline{1, m}$. To each object there corresponds the information pattern I_{A_k}. The recognition implements by comparison of an observable information pattern I_{A_S} , object $A_S \in \{A_k\}$, with an information pattern I_{A_P} , object $A_P \in \{A_k\}$, where $s \neq p$. For two of the same type objects A_S and A_P the information patterns appropriate to them will be similar to each other, i.e. $I_{A_S} \approx I_{A_P}$, $s \neq p$. Thus the recognition probability of object A_S - $P(A_S)$ concerning object A_P , $s \neq p$ can be less permissible - P_{pmb} , i.e. $(A_{S \neq P}) < P_{pmb}$. To supply recognition probability s-th of object among m similar with required value it is necessary to change an information object pattern. For definition of quantitative change value of an information object pattern a reciprocal recognition problem is decided. The reciprocal recognition problem is understood as finding of such information pattern $I_{\overline{A}_S}$, where $\overline{A}_S$ - means "not A_S" , for which $P(\overline{A}_S) \geq P_{pmb}$ among objects of A_k . Hereinafter the index s can be omitted, i.e.

$$P(\overline{A}) \geq P_{pmb}. \quad (1)$$

The set of changed object patterns - $\{I_{\overline{A}}\}$ correspond to a condition (1). Then the solution of a reciprocal recognition problem consists from:

1. Definition of set of patterns $\{I_{\overline{A}}\}$ satisfying to a condition (1).
2. Selection from $\{I_{\overline{A}}\}$ unique solution - $I_{\overline{A}}$.

We shall consider 1-st a part of a problem. We shall assume, that the recognition is conducted on one spatial coordinate, i.e. there is the one-dimensional information pattern

$$I_A = \{x_1, x_2, \dots, x_N\},$$

where N - number of resolution elements on spatial coordinate; x_i - the measured parameter values, which are independent and subordinate to the normal distribution law - $\mathrm{Norm}(x_i; a_i, \sigma_i^2)$.

We shall enter value λ . It will name as parameter of control, equal [1]

$$\lambda = \sum_{i=1}^{N} \left(a_i^A - a_i^{\overline{A}}\right)^2 = \sum_{i=1}^{N} \Delta_i^2 ,$$

where a_i^A, $a_i^{\overline{A}}$ - expectation normalized to variance of measured parameter value of an information object pattern, which correspond i-th of the resolution element, before its change and after accordingly.

The value λ_{min}, at which permissible recognition probability P_{pmb} is provided, is from expression [2]

$$P_{pmb} = 1 - \beta(\lambda_{min}) = \int_{\chi^2_{\alpha,N,\lambda=0}}^{\infty} d\chi^2(N,\lambda_{min}), \quad (2)$$

where $\chi^2_{\alpha,N,\lambda=0}$ - there is $100 \cdot (1-\alpha)$ - the percentage point central χ^2 - distribution.

The results of calculating $P_{pmb}(\lambda_{min})$ are shown on Fig. 2.

Parameter of control λ_{min} is integral parameter, for which it is possible to define set of patterns

$$I_{\bar{A}} = \{(a_1 \pm \Delta_1)\cdot\sigma_1^2;(a_2 \pm \Delta_2)\cdot\sigma_2^2;\ldots;(a_N \pm \Delta_N)\cdot\sigma_N^2\}$$

At equality

$$\sum_{i=1}^{N} \Delta_i^2 = \lambda_{min}.$$

We shall proceed to 2-nd parts of the solution of a reciprocal recognition problem. The plurality of the solution is eliminated, if for various measured parameter values of researched object there are the limitations on their change, stipulated for example, functions of the penalty $C_i(\Delta_i)$. So, if in quality of parameter at radar target recognition, value of radar cross section of a configuration object item is used, on its change with the help of various possible methods (change of the element form, screening, cover of an element the reflecting material and etc.) impose limitations as cost, power, time expenses and etc. Then the finding of desirable values of change - Δ_i is made in view of these limitations and is reduced to the solution of an optimization problem:

$$Q = \min_{\{\Delta\}} C = \sum_{i=1}^{N} C_i(\Delta_i)$$

under condition of

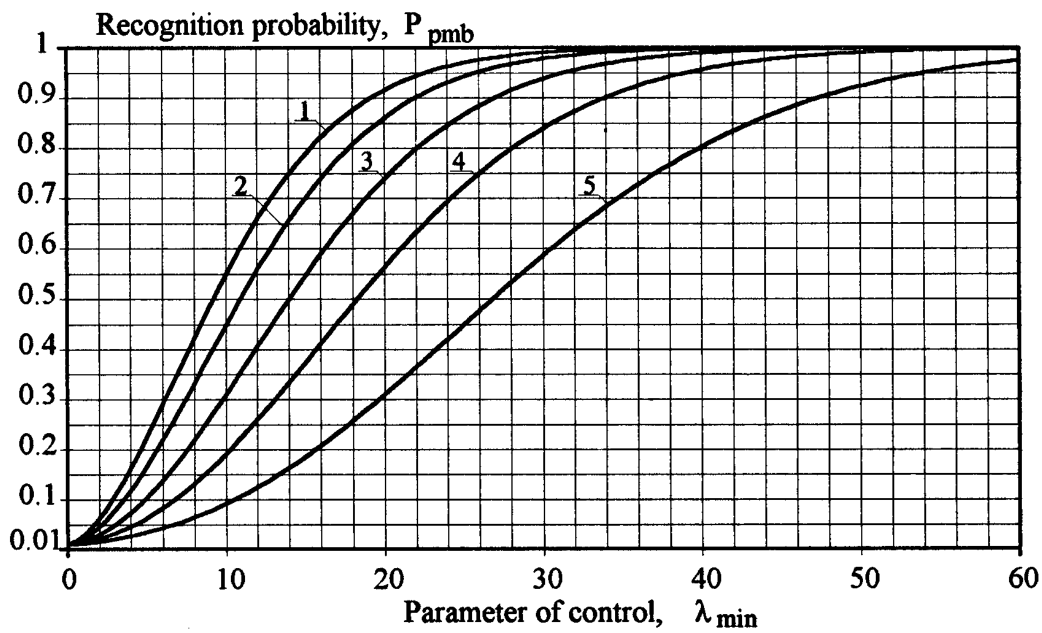
$$\sum_{i=1}^{N} \Delta_i^2 = \lambda_{min}. \quad (3)$$

The indicated problem of optimization in a general view of the analytical solution has not. The block diagram of the numerical solution of an optimization problem is shown on Fig. 3.

In the block 2 input of control parameter $\lambda = f(N, \alpha_0, P_0)$, calculated under the formula (2), of penalty function $C_i(\Delta_i)$ on all N to resolution elements, is made. The cutting off numbers n for a range of value Δ_i is introduced, which is selected depending on required accuracy of an evaluation. All C_i initial zero values, except one, for which $C_i = C_i\left(\sqrt{\lambda_{min}}\right)$ are appropriated.

In the block 3 a step $d = \sqrt{\frac{\lambda_{min}}{n}}$ of increment Δ_i is calculated.

Since the block 4 on 10 the circuit of exhaustion search of all possible change distributions Δ_i on $i = \overline{1,N}$, with check of a condition (3) is formed. At fulfilment of this condition calculation $C_k = \sum_{i=1}^{N} C_i(\Delta_i)$, where k - all possible existing methods of the information characteristic change, on found values Δ_i and storing of results is made. After this value k is increased per unit of.



False-alarm probability: $\alpha = 0.01$

Number of resolution elements: 1 - N=3 2 - N=5 3 - N=10 4 - N=20 5 - N=50

Fig. 2 Characteristic of control

The blocks 4,5,6 (7,8,9) illustrate the circuit for searching of values Δ_i on 1-st (N) resolution element. The number of the instructions of blocks 4,5,6 is equal to number of resolution elements N. After the process of calculation will reach the block 10, in memory of the computer all possible functions of the penalties C_k will be stored at appropriate variants of values Δ_i.

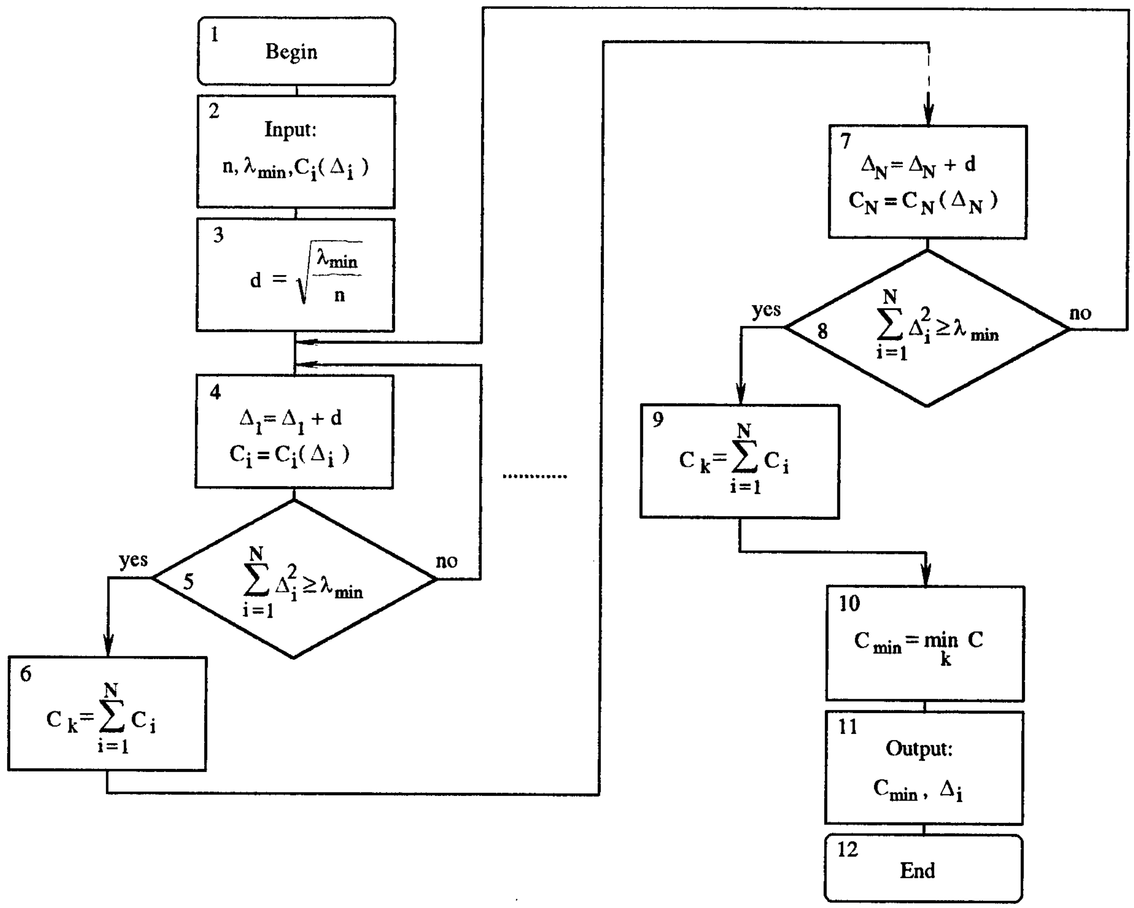


Fig. 3 Block diagram for the solution of an optimization problem

In the block 10 searching minimum value Q from k of penalty functions and by that definition of values Δ_i is made, at which the sum - $\sum_{i=1}^{N} C_i(\Delta_i)$ accept the minimum values.

In the block 11 results of the solution for optimization problem are printed out.

Conclusions. Thus, the solution of reciprocal recognition problem, with the help of a designated technique, issues the necessary and sufficient requirements to change of information object patterns, ensuring assigned probability of their identification.

References

1. Левин, Б. Р. (1989) Теоретические основы статистической радиотехники.- 3-е изд., перераб. и доп.-М.:Радио и связь, 656 с.

2. Бескид, П. П., Ю. Н. Рогалев (1996) "Управление информационными образами объектов в задачах распознавания". Сб.науч.тр.: Информационные технологии на транспорте. СПб: СПГУВК, с.155-159.

DESIGN OF DECISION MAKING INFORMATION SUPPORT SYSTEMS FOR COMPLEX TECHNICAL OBJECTS

Sergey O.Shaposhnikov, Ilja V.Stepanov, Sergey N.Turusov
St.-Petersburg Electrotechnical University, St.-Petersburg, Russia
E-mail: is@fvlfkea.etu.spb.ru

Abstract. The paper deals with the main problems of design of decision making information support systems for problem-oriented hierarchical organizational and technical systems.

Principles of decision making are considered, the system of preferences for the decision making person (DMP) for coordinated control of technical facilities is proposed. Decision making information support systems design is proposed to be fulfilled using the combined formal method based on the use of the principle of decomposition of a complex system into functional modules, i.e. simplexes. The use of semiotic models for description of the internal structure of simplexes is based on implementation of the modified Petri nets. The processes of readiness keeping and control of performance of complex technical objects are considered as subjects for decision making. Development of the information components for decision making support systems is also considered, and the main directions for the future research are stated.

Keywords. Organisational and technical systems, decision making, information support systems, Combined Formal Method, Petri net.

Introduction. The growth of problems solved by the personnel for readiness keeping of organisational and technical systems (OTS) causes sufficient growth of technical facilities used for decision making information support systems. That is why the person responsible for desision making along with the pressure caused by the responsibility for choosing the most effective,safe and economical decision is additionally pressed by evaluation of possible effects of decisions.

Decision making information support systems (DMISS) are used to provide the DMP with the necessary information on the processes taking place in the complexes of technical facilities and the environment as well as on the possible comsequences of decisions realisation.

DMISS are used to provide the person responsible for decision making with the necessary information about processes which take place in technical facilities complexes and in the environment, and about the effectiveness and possible consequences of realization of control decisions. Such systems solve the following tasks: monitoring and processing of information on the object elements states and environment, evaluation of the present states of elements, prediction of possible changes in elements states, and presentation of recommendations to the person-in-charge.

The main components of DMISS are:

- information system which fulfills monitoring and preprocessing of information;
- decision making system (DMS) or expert component (EC) which, on the basis of data about the object, models, quality criteria and priorities inferres rules for the further functioning of the object elements;
- querry control system which provides presentation of data obtained from DMS in the necessary form.

The paper deals with the main problems of DMISS development:

- functions and tasks of DMISS;
- the combined formal method for DMISS analysis;

- a method for development of DMISS structure

1. Functions and tasks of DMISS as an organizational and technical system. In general, DMS can be considered as as complex of organizational and technical structures of the object, provided to fulfill its tasks with a preset volume of resourses. The organizational structure reflects interaction of the staff who controls the complex object, its hierarchy, and instructions on the object control. The organizational structure, being to the sufficient degree a reflection of the technical structure, determines the set of restrictions laid upon the process of readiness keeping.

The technical structure is determined by those functions of the object and tasks which are to be solved by technical facilities. The technical structure includes technical facilities and their control systems, and reflects the structure of the technical facilities complex (TFC). Tasks of TFC are oriented on fulfilment of the global goal function.

TFC connections structure is based on resourses which are used to provide functions of a complex object (human, material, energy, information). Information resourses include special means and methods which bind the organizational and technical structures into a totality, i.e.systems for monitoring and processing information operators (object state indicators, systems for generation of control decisions and presentation them to operators, control and information pannels).

TFC structure reflects the global goal function (GGF), so the main principle for decomposition is to be the "technical" principle. It means, that the basis for decomposition is to be the technical structure. The main decomposition criterion is to be the requirement to construct the structure of technical facilities, each element of which will solve one of tasks or subtasks of TFC as a total. The tasks solved by the elements of TFC will be denoted as local goal functions (LGF). It is worth mentioning that decomposition shall be fulfilled under restrictions laid by the organizational structure, as it reflects the preset order of equipment exploitation and resourses use.

The basis for decomposition of TFC is the Combined Formal Method (CFM) which provides a system approach to TFC structure analysis and DMISS development [1].

2. The Combined Formal Method. The Combined Formal Method has 2 levels of presentation of information on the TFC: the abstract and the concrete. The abstract level allows to decompose complex systems into simplexes, i.e. modules which describe the internal structure, inputs and outputs of objects produced by decomposition. Simplexes are used to construct the complex of coordinated control. The complex is considered as the basic element of the system, i.e. any part of the system can be considered as a simplex of a 2-level complex (fig.1). The model provision is built as a stratified structure which allows and gives the necessary level of description. The only requirement is compatibility of formal descriptions of strats at different levels.

In the Fig.1 we denote:

W - the controlled object (CO) with its environment (if necessary);

M - CO model, i.e. the knowledge base on the structure, the present state and the functioning laws W;

F - the decision making model for CO control, i.e. control law as a totality of control procedures of finite number, each of them aimed to solve the certain local control task;

A - the adapter which selects a set of rules for decision making depending on the state W in accordance with the given algorithm;

I - the interpreter which reconstructs the model M depending on the real state W and the control instruction formed by the F block. Its main task is to interprete the environment and the

CO reactions for instructions of the control system and processes in the W by terms of the M block. This is fulfilled by means of special procedures which are realized in the interpreter, such as selection of the reason-consequence chains, determination of regularities, etc.;

D - facilities for control of the CO state;

R - facilities for interfacing the control system to controls;

$\mathbf{a}_w$ - CO state;

b - backloop signals;

$\mathbf{x}_m$ - model state;

$\mathbf{x}_A$ - adapter state;

$\mathbf{x}_I$ - the signal formed by the adapter;

$\mathbf{a}_F$ - the simplex functioning goal stated by the upper level;

$\mathbf{a}_m$ - information on goals and changes in the environment and the system as a total;

$\mathbf{g}_I$, $\mathbf{g}_F$, $\mathbf{g}_m$ - signals of strict control for blocks I, F, and M which are produced by the upper level.

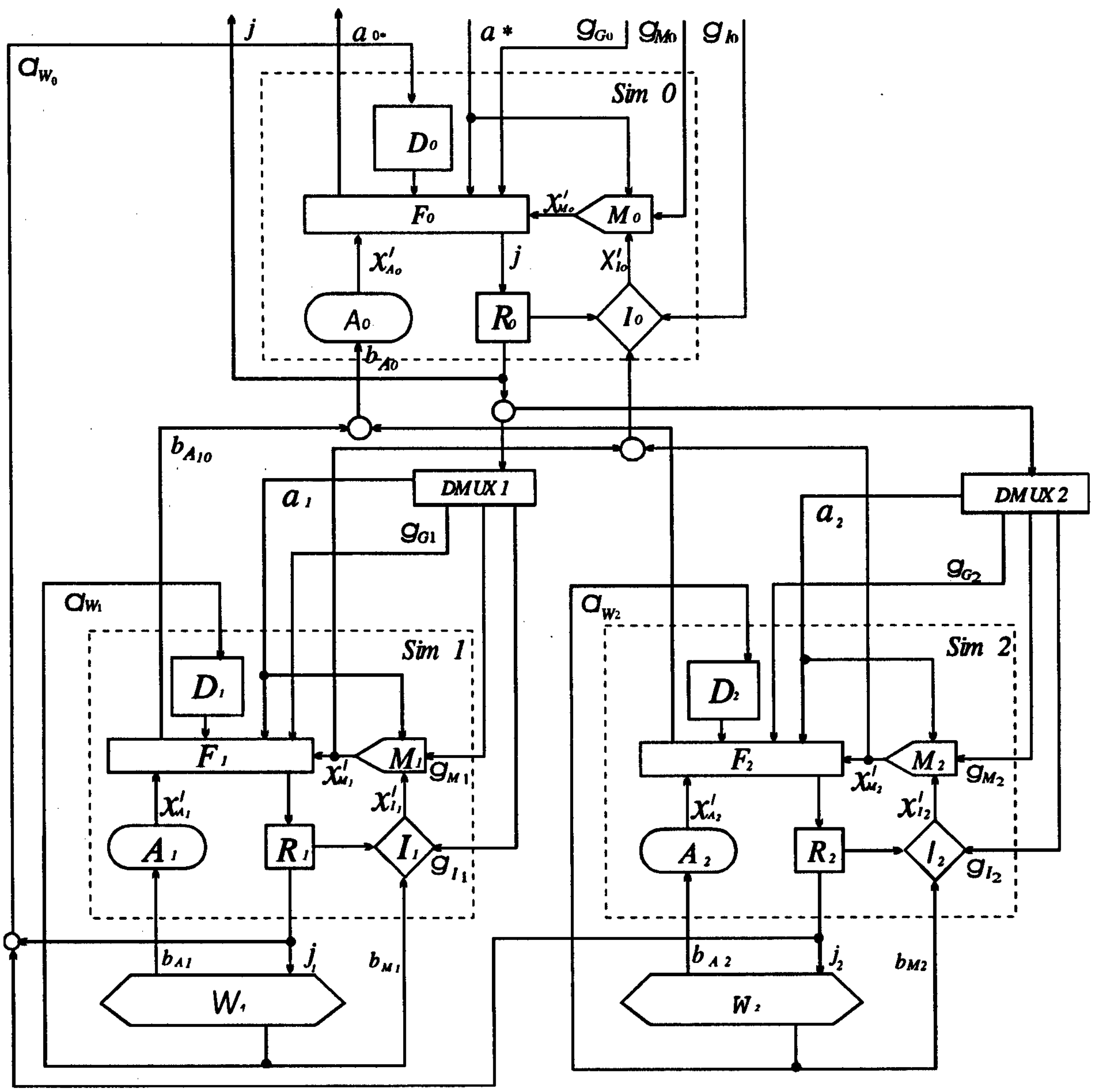


Fig. 1. The canonical complex

The concrete level of CFM provides reflection of processes which take place inside simplexes. It describes 2 interconnected models: the model M of the CO functioning and the model F of decision making for this object control.

CFM allows to consider the parts and structures of complex systems, processes which take place inside them, and processes of decision making for control of these systems. The simplexes produced by KFM are typical elements described in the unified mathematical language. It makes possible to have a system approach to the object functioning analysis which allows to model processes state changes for organizational and technical system structures as well as decision making processes for control of these structures and coordinated functioning of the lower hierarchical levels depending on requirements and restrictions set by the upper level.

It is possible, by means of CFM, to construct the generalized hierarchy of CTF (fig.2) deriving the following main levels:

- the systems of the complex level as a functionally connected aggregates of CTF elements which fulfill a part of tasks solved by the complex taking into consideration the GGF;
- the subsystems or groups of elements level, i.e. system elements which fulfill autonomous tasks in the frames of the system goal function;
- the local systems level, i.e. functional elements which fulfill their tasks according to LGFs.

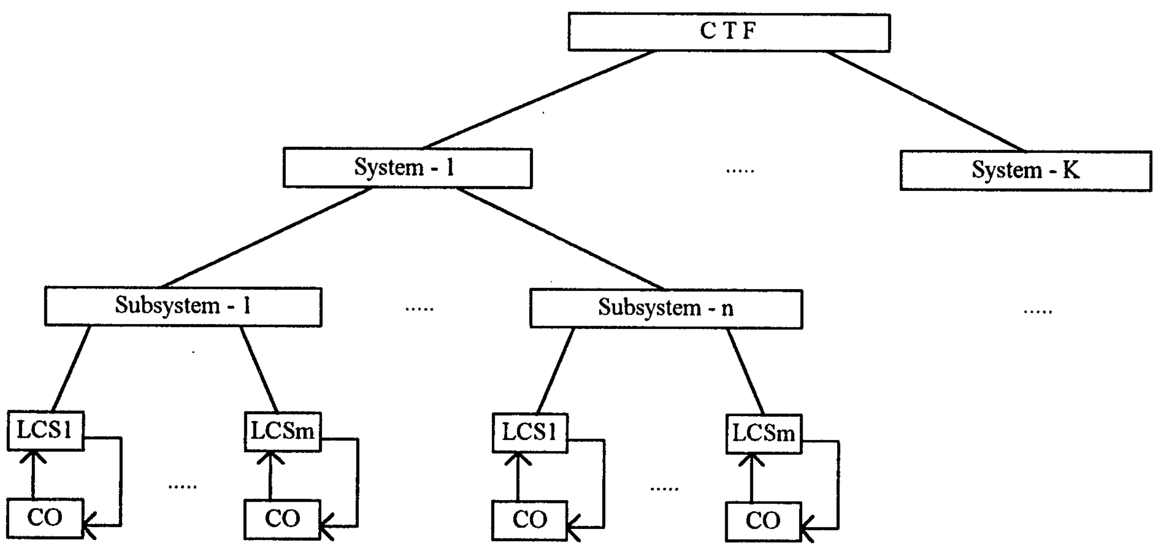


Fig.2. Hierarchy of the technical facilities complex

Organizational and technical systems (OTS) belong to situational control systems, and semiotic models play the main role for them. That is why the core of the CFM concrete level is the immitational model constructed using semiotic models. Semiotic models make possible to simulate the specific simplex of OTS, i.e. the person-in-charge who makes the decision and coordinates fulfilling of tasks on the KTF control as well as quality and timely decisions. Semiotic models are formed and modified by means of certain language facilities that are used to describe the system, environment and processes inside the OTS, and which reflect regularities taking place inside them. Models for decision making and evaluation are also based on semiotic representation of data.

In semiotic models, the dynamics of the controlled object is presented by situations description terms. The situation corresponds a certain system state or state of the object in to general. A situation in a model can be presented by a syntagmatic chain. One of possible versions of a syntagmatic chain is the Petri net (PN) which is of universal use.

The main advantages of PN are possibilities to describe complex asyncronous and parallel processes taking place in hierarchical OTS, relatively simple syntax and semantics determined by PN language and its interpretation, as well as easy-to-grasp kinds of models with wide functional possibilities. PN allows to consider the problems of software provision, instrumental support, and information processes in a system from a unified point of view. Substitution of temporal dependencies by cause dependencies allows to describe the structural features of the modelled system in a form easy to grasp.

PN formalises the concept of an asyncronous abstract system of dynamic nature which consists of events and conditions. The event is a concept which characterises the state of a system or its components. The condition is a concept that determines situations in which the event can be realised.

PN execution is controlled by the number and distribution of PN markers. The markers are located in circles, and they control transitions in the PN. PN execution is started by the transitions start. Any transition is started only in case it is allowed, i.e. if any of its input positions has the number of markers equal, at least, to the number of arrows from the position to the transition. Markers which are located at the input positions and which allow the transition are called the allowing markers. A transition is started by deleting of all allowing markers at its input positions, and by creating new markers which are placed into all output positions. The start of a transition changes the marking M of the PN to a new marking M .

To evaluate the modelling capabilities of the PN, the interpretation is used, i.e. translation of the initial description of the modelled system in a natural language into a description by means of mathematical language of formalisms. Construction of the description of the modelled system is an informal action carried out as a dialog between the "developer" and the "customer". Certain events which take place in the system as well as their conditions are described by means of PN terms. Transitions and positions in the interpreted PN are marked by some symbols which reflect the semantics of the system elements. Markers at certain positions reflect the true value of concrete sentences dealing with interpretation of the model states.

Thus, the CFM is used as a formal method for analysis of complex OTS. All the variety of immitation models of the CFM is described by means of a unified formal language. As a language for knowledge representation, syntagmatic chains are used which are based on PN. Modification of the PN provides interfacing of formal descriptions of M and F blocks inside any simplex as well as of simplexes of different levels. It allows to propose a unified methodological approach to the construction of knowledge models of structures and behavior laws for any OTS elements as well as to get the formalised descriptions of structures of complex OTS.

3. Structure of the DMISS

3.1. Tasks solved by the DMISS. The general structure of the DMISS [2] is presented at fig 3. Now, consider 3 main groups of tasks:

- information processing;
- elaboration of decisions and recommendations;
- presentation of decisions and recommendations to the person-in-charge.

The first group of tasks includes monitoring, receiving of information, its preliminary processing, and information storage. Some operations on the input data base and transformation of data into forms convenient for decision making procedures can also be included into this group.

By the object of decision making (ODM), we will denote the part of the CTF to which the certain decision making procedure is emplied.

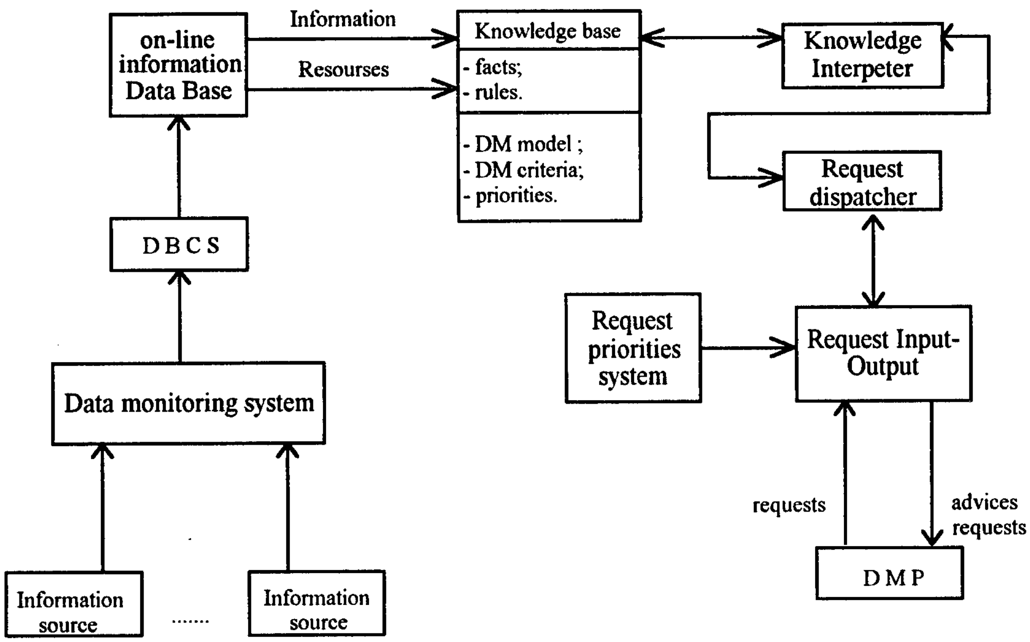


Fig.3. Generalized structure of the DMISS

Monitoring and receiving of information is done by a special system which includes measuring elements, interfacing facilities, communication lines, and a module for information collecting.

Initial processing of information includes procedures of comparing of input data with the preset values, and its compliance with the chosen data formats, evaluation of its reliability (quality evaluation), transformation of input data into formats which are used in the DMISS, and storing data in the on-line data base.

The on-line data base of the DMISS is a special data structure which keeps all information on the processes taking place in the ODM. Access to the data base is done by requests which are served by the data base control system (DBCS).

The second group of tasks includes the tasks of decisions and recommendations synthesis for the person-in-charge, prognosis and modelling of the ODM processes, the dialog with the person-in-charge and processing of requests from the staff at the local technical facilities.

The process of decision making is a complex of procedures which provides the ODM status evaluation, determination of necessity to interfere into the ODM functioning, preparation of the corresponding control commands, and transmittion of these commands to the activators or to the staff of local technical facilities.

Decision making is done according to the general strategy of the DMISS. The strategy is stated by the person-in-charge and depends on the availability of resourses needed for decisions implementation. Decision making is based on transforming of data (facts) stored in the data base, according to the rules stated for the system and stored at the knowledge base

(KB), and to restrictions inplied by the state of the ODM, available resourses, and by the person-in-charge .

Interactively, DMISS interpretes the requests, produces the responses, and provides the person-in-charge with information in the natural language or in a graphical and analytical form.

3.2. DMISS as a distributed computer system. Proceeding from tasks solved by the DMISS, we can list the following main elements of the DMISS:

- the system for monitoring, receiving and processing of information. It includes measuring elements (sensors), interfacing facilities, information transmission lines, and algorithms for control of receiving and preliminary processing of information.
- on-line data bases which include the data base structure, data base control system, interpretation procedures for requests from allied systems.
- decision making system which includes the knowledge base, the system for knowledge control, and the inference engine.
- the system for requests processing which includes the language of requests and algorithms for interpretation of the person-in-charge requests and answers generated by the system, as well as failities for presenting of the system's answers to the person-in-charge.

Thus, the DMISS can be treated as a computer system. The DMISS structure can be determined by superimposing of the organisational structure on the hierarchical structure of the CTF.

Development of the DMISS as a distributed computer system is caused by the DMISS hierarchical character as well as by the necessity to automate monitoring of information about the elements of the CTF and to bring the control and decision making system closer to the controlled ODM.

The main tasks of the ODM elements control are solved at the lowest level of hierarchy. For example, the level of local control systems (LCS) is the lowest for a shipborn complex of technical facilities. The upper level, i.e. group control systems coordinates several LCS according to the task stated for this level by the one stage upper level, i.e. the level of centralised control. Systems at this level get only that part of information which is needed to fulfill effectively the coordination of LCS systems. The level of the centralised control also gets only that part of information which allows it to make decisions concerning control of the shipborn complex of technical facilities in general.

The local control systems solve tasks of monitoring information about the controlled object (CO), evaluation of its state, and elaboration of control commands according to requirements and restrictions stated by the upper level. The upper level gets information on the main parameters of the CO and a generalized conclusion on the quality of control results.

The group control systems (GCS) are a distributed microprocessor system which is coordinated by the central computer. The GCSs monitor information from the LCSs, evaluate the state of the controlled system, produces tasks for the LCSs, and keeps information on states of LCSs. The GCSs provide the upper control level with information on the main system parameters and conclusions about the quality of results of GCSs control acts.

The systems for control of complexes (SCC) deal with tasks similar to those of GCSs. They also have some additional functions as they include elements of an expert system and possess the necessary knowledge data. SCCs analyse states of controlled systems and produce recommendations for the staff on organization of the system control and maintenance.

All SCCs are united into a shipborn territorial computer net which includes:

- the computer net server which keeps the data base and the knowledge base, the system manager, and the information interchange administrator.

- servers of local conputer nets of SCCs and servers of GCSs distributed microprocessor systems.
- the control pannel of the person-in-charge.
- net interconnection facilities.

Development of such distributed systems can be justified only under the following conditions:

- availability of a unified local control system module which has technical features enough to solve most tasks of local control.
- availability of unified hardware and software facilities at computer net servers and LCSs.
- availability of unified protocols for information interchange.

3.3. Information flows analysis of DMISS. The information flow analysis of the DMISS is oriented onto determination of characteristics for the nets included in in distributed computer system. The results of the analysis are:

- information needs of any hierarchical level of the DMISS.
- protocols of cooperation for the DMISS elements.
- configuration of nets included in the DMISS as well as the general architecture of the distributed computer net.
- evaluation of nets and their components characteristics.

Using the proposed method for analysis of the OTS by means of the CFM, it is evident that the lower the level of hierarchy the greater the role of technical criteria, and the higher the level of hierarchy the the greater the role of organizational criteria.

At the lower levels, decision making deals mostly with control of certain technical objects. It mainly consists of producing the control command which will transfer it into the desired state. The object state can be described as follows:

$$\begin{cases} x' = \mathbf{A} \cdot x + \mathbf{B} \cdot U \\ y = \mathbf{C} \cdot x + \mathbf{D} \cdot U \end{cases} \qquad (1)$$

The decision making task is producing such a control act which under minimal consumption of resources R (time, energy, fuel, etc.) will provide a transfer from the present state S_1 to the pseset state S_2

$$\exists \mathrm{U}_{\mathrm{opt}} \Rightarrow S_1 \xrightarrow{\min R} S_2 \qquad (2)$$

For upper levels of the hierarchy, decision making deals mostly with search of the strategy which will allow to transfer a set of elements into the desired state under minimal expanses. Thus, on the basis of knowledge on the state of the ODM and the desired goal it is needed to find the optimal strategy for lower levels which will meet the preset level of the goal function.

The structure of the distributed data base as well as of the knowledge base will depend upon the hierarchy level. For LCS, the volume of information necessary for its normal work can be determined as

$$\dim I = \dim y + \dim X_* + \dim U + \dim e \qquad (3)$$

where

I is the full vector of information used by the LCS which can be expressed as $I = [y, X_*, U, e]$;

y is the vector of observed, or measured state variables of the controlled object;
X_* is the vector of estimations of the controlled object state variables;
U is the produced vector of control comands for the object;
e is the vector which determines the strategy of the LCS stated by the upper element of decision making.

For GCS, the full information vector I is formed of vectors I (information on LCSs controlled by this GCS) as well as of matrixes formed during decision making. The decision making process can be described as

$$\begin{cases} f = t \cdot (K^T \cdot I) \\ S = r(f, X) \\ S_{opt} = optim(S \mid (R \& P)) \end{cases} \qquad (4)$$

where

$I \in R_{n \times 1}$ - the vector of information on the ODM state;
$K \in R_{n \times m}$ - a set of criteria for decision making, $K = [K_1, K_2, \ldots, K_m]$ (all $K_i \in R_{n \times 1}$);
$t \in R_{m \times m}$ - the matrix reflecting the task of decision making;
$f \in R_{m \times 1}$ - the final evaluation of the ODM state;
$S \in R_k$ - the set of acceptable decisions;
$X \in R_m$ - the scales for criteria evaluations;
S_{opt} - the optimal decision under restrictions on resources R;
r - the decision rule;
P - priority system.

Similarly, it is possible to determine information need for upper levels of DMISS hierarchy. It is necessary to mention that the vector of information needs I will have its own value for every level. It will be determined by decision making tasks which are stated for certain decision making elements of the DMISS.

Thus, the best type of the data base is the distributed one. Decision making elements at each level shall have its own local data base which containes all the necessary information for decision making as well as the local knowledge base containing criteria of decision making, decision quality estimations, etc.

Distributed type of data and knowledge bases allow to reduce sufficiently the information interchange between the decision making elements at different levels of DMISS. The decision making element transfers into the net of the certain level only information needed by other elements to elaborate control or coordination decisions.

Based on the information need of the decision making element at a certain level of hierarchy, one can determine the architecture of the net for each hierarchy level, protocols for information interchange, and the passing capability of the net communication channels.

References

1. On Formal Design of Decision Support Systems for Marine Technical Facilities Maintenance. //A.Bulatov, S.Shaposhnikov, I.Stepanov, S.Turusov /Thes. of First International conference on Marine Industry, Varna, 1996.

2. Концепция построения судовых систем интеллектуальной поддержки. //А.В.Булатов, И.В.Степанов,С.Н.Турусов,С.О.Шапошников /Тез. IV Международной конференции "Региональная информатика-95", Санкт-Петербург, 1995 г.

COMPREHENSIVE COMPLEX SYSTEM FINALIZING BY THE SEMI-SCALE SIMULATION METHOD AND CRITICAL REALIZATION SELECTION FOR THE STATISTICAL EXPERIMENT SPEED-UP

V. A. Nickoltsev, Dr. Sc., Electrical Engineering,
Yu. F. Podoplekin, Dr. Sc., Engineering Sciences,
E. A. Voinov, Ph.D., Engineering Sciences,
V. R. Andreyevsky, Ph.D., Engineering Sciences.

The report is devoted to the experience of complex system finalizing at the "GRANITE" Central Research Institute both as to the determination of scope of problems of the semi-scale simulation centre and its facilities and as to the statistical experiment methodology.

A possibility of application of the random factor critical realization selection has been considered for the solution of a general problem of the probabilistic quality index assessment of the system being investigated. It is shown how the critical realization selection is combined with other labour content reduction methods of statistical experiments: "main part selection" and "stratified sample".

1. THE SEMI-SCALE SIMULATION CENTRE'S ROLE IN THE COMPLEX SYSTEM DEVELOPMENT.

Being the developer of automatics and telemetry systems, weapon and movable object control systems, environment and target indication lighting facilities, monitoring and test equipment and simulators simultaneously, the "Granite" Central Research Institute (one of the oldest defense industry enterprises in St. Petersburg) has accumulated a broad experience of development of complex radio-electronic systems of military and civil purpose. Such systems can be developed quickly and qualitatively with the simultaneous application of two outwardly controversial principles:

- decomposition into subsystems and devices with a high degree autonomy of their development;
- component characteristic optimization by efficiency criteria of solution of the whole problem.

On the one hand, such dialectic approach requires a multi-aspect mathematical simulation, accumulation and utilization of the previous development experience not only in the form of repetition and development of successful solutions but also in the form complex mathematical models, which have proved its usefulness. It requires to exert operative control over the course of development in order to secure an up-to-date design of the overall system structure, the formation of partial requirements and a quick response to their corrections. The developments appear to be bound by common concept and calculation.

On the other hand, a thorough check of the collective effort results is required as to the conformity with this common concept. A total control is required of the conformity of the component development results with the initial requirements and of their lack of features and characteristics capable to aggravate an integration at the final stage, the lack of problems with the hidden failure diagnostics. The mathematical models themselves being the basis of the optimal system synthesis on the whole should also be checked for their adequacy to the partial development results.

At the initial stages, such control is secured by a simple collation of mathematical models and requirements used. Subsequently, an hierarchy of the subsystem and device developers is being established, which is closed by an integration stand and a dynamic test facility. The main role of the semi-scale simulation centre is to check and confirm initial data, models and results of previous mathematical simulation with the completeness necessary to the customer. The simulation centre, where the real apparatus belonging, for example, to rockets, interact, in the course of "electronic start-ups", with simulators and mathematical models of environment and counteraction, plays the role of an "electronic proving ground", on which tests predominate full-scale tests as to the information content. Of course, there always be a demand for rocket start-up flying tests in order to test rockets and their carriers under real start-up conditions. However, there are three aspects, in which the semi-scale simulation remains beyond competition.

First of all, it is a relative cheapness of coverage, by "electronic start-ups", of the whole range of performance conditions, which is established by the requirements to the system being developed. The statistical simulation formulation is possible and is even being implemented in practice, which meets probabilistic requirements to general system characteristics with a high confidence level.

There is no alternative to the simulation centres' ability to secure tests under extreme and catastrophic conditions, this being done without any risk.

At last, it is difficult to overestimate a possibility of multiple test repetition with the same apparatus set under well controlled conditions when reason diagnostic problems of an off-design behaviour of the system or its components arise. It is often that the reason appears to be an interface inconsistency between subsystems, which appear only under conditions of joint functioning. One of the problems becomes a sort of system "accompanying" during its full-scale tests.

Such problems seem to be justified as extra ones for the simulation centre, though, as a rule, the optimal synthesis is to be excluded from the list of main problems of semi-scale simulation owing to the necessity of quick completion and payoff of funds invested into research, design and pilot apparatus production. In the first place, negative test results may occur. And it is the same requirements of quickness and economy that make the simulation centre to be a convenient place for verification of solutions, which improve the situation. In the second place, a similar situation arises at the necessity to investigate possibilities of meeting supplemented and extra requirements, at the development of proposals on improvement of the existing systems.

If one takes into account that customers may need to confirm and elaborate the system characteristic assessment under changing conditions and to compare them with potential new developments or purchases by the "return/expenses" criteria, then we shall extend our representation about possible problems of the semi-scale simulation centres.

2. THE DYNAMIC TEST COMPLEX OF THE "GRANITE" CENTRAL RESEARCH INSTITUTE

In the diagram, the structure of the dynamic test complex (DTC) of the "Granite" Central Research Institute is shown for the control system finalizing of rockets with multichannel self-guiding heads (MCSGH) with the condition imitation of discovery, classification, selection and accompanying of targets as well as that of disturbances and arranged counteraction. The DTC provides an interaction check of the control system components at their joint functioning under the conditions of "electronic start-up" with an accurate simulation of the control object response to the control action.

In the structural diagram, the auxiliary power supply systems comprising an electric aggregate one, a compressor one and a hydraulic aggregate one as well as software development means for the DTC computer systems have been dropped.

On the whole, the DTC is a complex system of computer, radio electronic and electromechanical means as well as their programs interacting, in real time, with the apparatus being tested.

Conducing the semi-scale simulation at the DTC is a final assessment stage of the performance and the disturbance-protection of the systems being developed. The utilization of the semi-scale simulation, which increases the completeness of knowledge about the system characteristics, enables to significantly economize on the full-scale test amounts. In this case, the development and test cycle reduction is being reached.

3. CRITICAL REALIZATION DISCRETION TECHNIQUE

A critical random factor realization discretion technique excluding an area of obvious (uninteresting) outcomes in statistical simulation was offered in the 60s by Mr. N. P. Yakovlev, an employee of the Granit Research Institution, as a means to cut down effectively labour intensity in estimation of unfavourable system function probability. An efficiency of this technique was demonstrated using a n example of investigating the missile control dynamic properties at a start site. The idea of the technique was described later [1] 1975. Subsequently, this technique was invented by other research workers. A misleading reference is known which was made in the book [2] to the paper by V. V. Kouzmin, V. A. Yaroshevsky in *The Ucheniye Zapiski CAGI* (1984, Volume 15, No. 2) as the priority in this area.

The idea of the technique is to establish the equivalence between two statistical experiment schemes: i.e., a 'conventional' one where all the random factor realizations obtained as a result of 'occasional game' are investigated on the full-scale, and a 'critical' one where those 'unnecessary' are sieved out which do not establish opportunities for an outcome interesting to the investigator.

In view of the fact that the statistical testing technique is usually employed to verify a rather developed system in the course of design which behaviour may result in doubts only in case of significant deviation, say, in excess of two thirds of the allowed figure, and at least two determining factors simultaneously towards an unfavourable direction, then a reduction in the number of actual verifications while preserving a given estimation reliability is great, as high as by the factor reverse to a probability of a random option getting into the critical area. The critical realization discretion technique is worth being at the top position on the list of known techniques employed to cut down a statistical experiment labour intensity. First an foremost it is related to mathematical and derivative simulation (along with an increase in the space dimensionality of 'random' factors controlled in the course of an experiment, an efficiency of the technique is growing also).

In the original version, the critical realization selection method does not pretend for solution of problems in general formulation – the mathematical expectation evaluation of a certain system quality functional.

We shall show that it is possible to combine ideas of other methods: "main part selection" and "stratified sample" with the critical realization selection principle for the solution of the problem in general formulation.

4. ON A POSSIBILITY OF COMBINATION OF THE CRITICAL REALIZATION SELECTION METHOD WITH OTHER METHODS

The most popular are the variants of the "main part selection" method ("correlated sample", combined method of determination of probabilistic characteristics) [4,5]. The main idea of this group of methods is the utilization of a simplified (ideally – analytical system model taking into consideration the influence of main random factors upon its behaviour. This enables to reduce handling of the statistical experiment results with a more accurate model or with the system itself to the evaluation of corrections to the probabilistic characteristics of the system defined on the simplified model. The assessment of the mathematical expectation of the quality index vector obtained on the approximate model, $m_{y,м} = M[Y_м]$, is being elaborated with the help of the sample $\{Y\}$ of realizations of a more accurate model or a full-scale experiment with the formation of the difference sample between quality indexes of the system and its simplified model $\Delta Y = Y_м - Y$

Then let us suppose that $m_y = m_{y,м} + M[\Delta Y]$ and use that the dispersions of differences ΔY are significantly less than the dispersions of the quality indexes themselves (the most essential in the system behaviour is that the model, even a simplified one, should predict). As $D[\Delta Y] << D[Y]$, then $D[\hat{m}_y]$, which is approximately equal to $\frac{1}{n} D[\Delta Y]$, where n is a sample size, becomes much less than under the direct evaluation $\hat{m}_y$ over the sample $\{Y\}$. The sample size required, which is evaluated by the given relative error of the result, is so many times less as $D[\Delta Y]$ less then $D[Y]$. In this case, the way of selection of random factor combination is of secondary value. The main thing is that they should coincide for the full (system) and the simplified models and that the model adequacy should secure the correlation coefficient between Y and $Y_м$ to be greater than 0.5.

The introduction of the critical realization selection into methods of the main part selection becomes possible at the natural, as it is seen, elaboration of the requirements to the approximate model adequacy. If potential deflections ΔY are significantly less than at the ends of the tolerance field, they will not affect the results, and one may drop checking the corresponding realizations and assume for them that $\Delta Y = 0$.

In order to avoid the necessity to prove the insignificant amount of a possible $\hat{m}_y$ evaluation displacement, the notion of the critical realization region can also be made more accurate. Everything that remains after the exclusion of the symmetrical region, adjacent to the nominal characteristics, of the simplified model adequacy is expedient to be related with it (an elaboration remark – only at the general set periphery). As the main probabilistic mass is concentrated near the nominal characteristic point, the essential research labour intensity reduction effect, which is characteristic for the critical realization selection, is being secured.

The utilization of the approximate model is a significant element of another effective method – "stratified sample" too (in [6], there is a detailed bibliography together with the statements of basics). Let X be a random vector of the system parameters with the probability density function $W(x)$ and $Y = f(X)$ be a quality index, then the problem in question reduces to the evaluation of an integral $J = \int_{\Omega} f(x)W(x)dx.$

In the stratified sample method, the set of possible values of the random vector X is being divided into L non-intersecting "layers" $\Omega_1, \ldots, \Omega_L$, and a simple random sample of volume n_h is being taken from each layer Ω_h. As an estimate of the probabilistic characteristic of the system, a weighted sum of the in-layers assessment is used

$$J = \sum_{h=1}^{L} \omega_h \hat{J}_h, \qquad \text{where} \qquad \hat{J}_h = \frac{1}{n_h} \sum_{i=1}^{n} y_i,$$

$$y_i = f(x_i) \text{ under the selection of } x_i \in \Omega_h, \qquad \omega_h = \int_{\Omega_h} w(x)dx.$$

As the layers do not intersect, $D[J] = \sum_{h=1}^{L} \omega_h^2 D[\hat{J}_h] = \sum_{h=1}^{L} \frac{\omega_h^2}{n_h} D_h$.

(Here $D_h = D[f(X)]$ under the selection of $x_i \in \Omega_h$).

Under the corresponding selection of breakdown into layers and numbers $n_1, n_2, \ldots, n_L$, the dispersion $D[J]$ is significantly less than under a simple random sample and the same total number of tests and it is so much the less of the magnitude D_h, i.e. the more homogeneous is the sample in each layer.

At the stratified sample planning, the most complicated problem is that of the determination of layer boundaries. And it this problem that requires attraction of an approximate (simplified or using previous experience) model. And if the approximate model allows to obtain a good estimate for J_h even in one layer, then it is possible not to carry out research in this layer with the accurate model or with the system itself without the risk to disturb the general evaluation of the probabilistic characteristics.

5. CONCLUSION

The finalized design system for big complexes with the use of powerful stands of mathematical and semi-scale simulation, the availability of modern technologies and research personnel enable the "Granite" Central Research Institute to maintain leading positions in Russia and at the export market in the field of control systems for rocket complexes of marine and universal basing.

BIBLIOGRAPHY

1. Bogdanov V. S., Yakovlev N. P. (1975). On a certain approximate method of probability calculation at the non-linear system research. Izvestia AN SSSR, Technical Cybernetics, No. 5.

2. Malyshev V. V. Et al. (1989). Optimization of the observation and monitoring of flying objects. M., Mashinostroyenie.

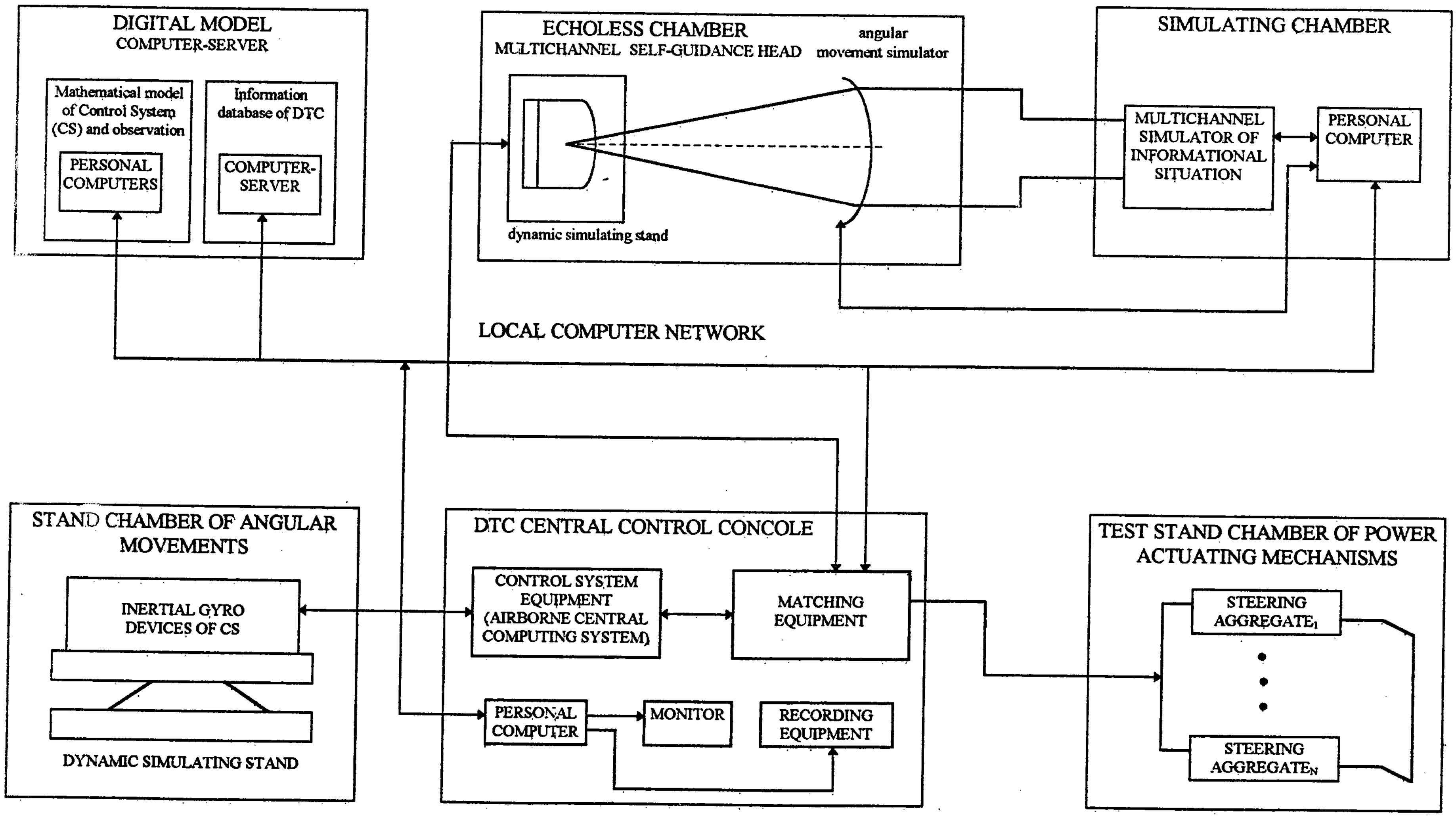
3. Yermakov S. M., Mikhailov G. A. (1982). A course on statistical simulation (a teaching aid). 2nd edition. M., Nauka.

4. Gnedenko B. V., Kovalenko I. N. (1987). Introduction to queuing theory. 2nd edition. M., Nauka.

5. Pougachov V. N. (1973). Combined determination methods of probabilistic characteristics. M., Sovietskoye Radio.

6. Vassiliev D. V., Sabinin O. Yu. (1987). An accelerated statistical simulation of control systems. L., Energoatomizdat.

DYNAMIC TEST COMPLEX (DTC)



DEVELOPMENT OF MATHEMATICAL MODELS AND ANALYSIS METHODS OF SYSTEM RELIABILITY, SURVIVABILITY AND SAFETY OF AUTOMATED TECHNICAL COMPLEXES WITH HIGH OPERATIONAL RISK AND THEIR CONTROL SYSTEMS.

Igor I. Simakov
Naval Academy named after Kuznetsov N.G., St. Petersburg

Igor P. Simakov
Scientific&Production Association "Aurora"
of the Ministry of Defence Industry of the Russian Federation,
15, Karbysheva Street, St. Petersburg, 194021, Russia
Telephone: +7(812)535-97-23(home), E-Mail: ppnovod@dux.ru

Abstract. The authors present the theory, elaboration methods of mathematical models of different types and assignment (determined and random, continuous and discrete, dynamic and static) and methods of complex analysis and quantitative assessment of system reliability levels, survivability and safety of a complicated structure and potentially hazardous multi-unit technical systems with higher technical, environmental, social and economical operational risk.

Keywords. reliability, survivability, safety and efficiency of technical systems, control systems, structural organisation, dependent failures, cascade damages, dynamic and structural stability, structural functions, ranges of degradation, dynamic reliability, self-organisation, self-algorithmization, homeostasis features, nuclear-powered ship power system, power unit of nuclear power station (NPS).

Introduction. The necessity for basic development of well-known patterns and methods of complex analysis of reliability, survivability, safety and efficiency of modern automated technical complexes (ATC), such as technical facilities complexes (TFC) and nuclear-powered vessels` power systems, power units of nuclear power stations, (NPS) comes from a number of their most important features and peculiarities, which are stated below. Along with this, the proper analysis needs to consider arrangement (structure and behaviour) of control systems (CS), which play a decisive role in the provision of high levels of safety for technical systems and complexes with higher technical, economical, environmental and social operational risk. *The main features of automated TFC, which determine methods of its analysis and synthesis are*:

1) considerable power consumption of the units, strict limitations in the accessible spectrum of physical processes co-ordinates` alterations, in compliance with fail-safe (safety) operation, which require realisation of regulation functions in a control system, emergency protection to prevent damages and failures of units under conditions of disturbances and faults in technologically connected units;

2) a large number of units, interconnected by common energy and substances flows, requiring the realisation in CS functions of co-ordinated control, in order to maintain dynamic balances of energy and substance flows between units;

3) disorders in TFC (failures of some units, complex partial damages under the impact of hazardous environmental situation) can cause sudden imbalance of energy and (or) substance flows between units, transitional process propagation in TFC, which can lead to overloading of serviceable units and their further malfunctioning. It determines the necessity to realise developed functions of anti-emergency and co-ordinated control in CS, which provides prevention of cascade (catenary) transfer of failures from one unit to another, and TFC transfer to safety mode with probably lower level of operational

quality (output, power, manoeuvrability, accuracy).

Also, let us formulate in this study the main specific peculiarities, connected with the problems of assessment and provision of TFC and control systems survivability during damage control, aimed at the preservation of possible levels of working capacity (efficiency of operation), provision of fail-safe operation and safety.

4) Indefiniteness of time, place, size (sphere of action), duration, intensity (destruction force) and physical nature of harmful effects (HE), its unique nature, priori indefiniteness, absence of probability measure, unpredictable character.

5) Combinations of different harmful effects in nature (complex emergency impact), which, as a rule, originate in consecutive time order, when the first HE, causing damages of some components (stuffing boxes, seals), units and mechanisms, "calls for "HE of different nature (e.g. fire - break of high pressure air tube - access of outboard water and so on).

6) Under the impact of HE unpredictable volume of TFC components` damages (units, mechanisms, automatics devices, cables, pipelines and so on), unpredictable combination of initial damages and different TFC components failures.

7) Appearance of non-standard (not envisaged by design) emergency situations, which are not controlled by algorithms, used in traditional control systems and safety control systems, and the absence of technological regulations for operational staff (appearance of so called "algorithmic disorder "of the control system).

8) Sudden, unforeseen and unpredictable character of the object, its "accidental" degradation, manifestation of "sleeping" links.

9) Possibility of a cascade development of damages and failures in interconnected multi-unit TFC ("domino" effect).

10) Indefiniteness of situation and unpredictable character of final after-effects of damages' cascade development in TFC (strong influence of number of hidden failures and exact points of "accumulation" in emergency protection facilities and safety systems, ambiguity and unpredictable character of object's further degradation trajectories).

11) Incompleteness, unauthenticity and quick obsoleteness of information about an emergency situation, HE parameters, and also TFC dynamic structural states.

12) Incompleteness of information caused by incomplete stability of information sources, transducers and signal lamps, communication lines and other components.

13) Strong influence on automated complexes of technical facilities topology survivability level, its territorial distribution in the ship's compartments, and also the structural organisation of CCS TF.

14) The necessity to make decisions or automatic work out of control influences, in "unplanned" or non-standard emergency situations, along with the beforehand unpredictable damage combinations or with arbitrary combinations of multiple failures of components in separate TFC or in complexes, which are interconnected by TFC technological process.

15) The problematical character of production decisions, concerning switching or TF structures reconfiguration is caused by :

- large dimensions of TFC mathematical simulators,
- impossibility to make a TFC structure regular,
- presence of "long-chained" feedback,

16) The tasks to control reconfiguration of TFC structure along with "unplanned" combinations of failures refer to overhaul, possess combinatory complexity and are NP-difficult.

17) Possibility of TFC automatic shutoff in order to prevent its existence, violation of safety conditions for the crew and elimination of risk for environment.

18) The necessity to register natural unreliability and insufficient stability of the system, its components and links to harmful effects.

19) Limited possibilities for recovery and repair.
20) The necessity to work out expensive design decisions.

The features of safety, survivability, efficiency and system reliability, which are being discussed in this study, are the development of notions of the traditional theory of automatic control, concerning stability of movement and quality of control processes for a wider grade of disturbances, causing structural disturbances in the operation of "Control object (CO) + Control system (CS)" complex as a whole.

System of criteria and values of reliability, survivability, safety and efficiency assessment. The criteria, which the systems must meet, and which have to be divided, at least into two groups, are of vital importance for the designer. The groups are :
a) *criteria, which have to be unconditionally executed,* not depending on ship assignment, at any (practical) expenses, and also not depending on software used in the designed system; an examples of such criteria are the criteria of complete safety of nuclear-propulsion plant (NPP) and cruise safety at any, at least single, sudden fault in object operation, control systems or related support systems;
b) *criteria, which implementation is necessary to strive for,* but the grade of its implementation (or its weight factors) must be determined depending on ship assignment and available resources.

It is important, that these criteria and values take into consideration the influence of technical facilities, which are the objects of control, their structural and dynamic features and characteristics, "adaptation" for control automation. For the criteria of group 2 it is evident, that it is necessary to consider the tasks being solved by the ship, tactical and technical, technical and economic, military and economic values and values of production effectiveness and so on. The solution of tasks, concerning provision of high quality levels for CS structural organisation, on one hand, leads to the necessity of conducting a thorough analysis of possible faults' consequences for TFC operation quality, and, on the other hand, for losses level, caused by emergency situations and the risk level of a safety conditions' violation. Such a comprehensive analysis can be exercised only on the base of the research of "Object + Control System" complex simulators, considering it integral to any inner and outer disturbances of a different nature:
(a) in the process of long-time working life inevitably failures appear, related to object components and its control systems, caused by its natural low reliability or production and design defects, including accumulating hidden failures in safety systems, emergency protection systems, systems of access control, etc. ;
(b) possible ruptures and short circuits in communication lines and power supply circuitry;
(c) faults in operation of information processing facilities or errors in software;
(d) disorder in operation of sub-systems of object and equipment power supply;
(e) possible partial damages of TFC and control systems and (or) main cables, caused by extraordinary environmental damages which are not predetermined by conditions of normal operation.
On the basis of methodology, designed in SPA "Aurora" for selection of technical decisions for structural organisation of automated TFC and its control systems the following hierarchy of criteria were used [1,2,8] :
A. *Determined criteria of structural stability*, which are unconditionally executed (without considering the costs of resources) :
A1. *Criteria of "complete" safety. (fail-safe operation)* at any sudden moment, caused by a natural low level of reliability, including possible accumulation of hidden faults in the systems of safety and emergency protection.

A2. *Criteria of survivability occurring after partial damages* of equipment and cables and realisation of any single emergency environmental impact from the standard multitude of mutually exclusive possible hypothesises.

Input criteria must be appraised and fulfilled not statistically (on probability or on averaged from the criteria of many systems), but determinably for each exact system and to be acknowledged not by resource tests, but through direct simulation of an object with failures determined by multiplicity simulation.

Additional criteria in this group are the criteria of failure stability (for example, criterion of single failure in nuclear power engineering), the multiplicity of which is determined with the consideration of allocated resources.

B. *Probability criteria and values of functional efficiency and TFC and CS system reliability and also probability (arbitrary and absolute) values of risks,* concerning violation of safety conditions (for example nuclear), losses resulting from a low level of reliability and total expenses for system creation and operation. It is important, that in these values, the characteristics of the complex's "TF + CS" different features (dynamic, reliability systems of technical maintenance) are linked up. The selection of rational variants from the multitude of alternative variants is made in compliance with Pareto criterion in axis type "efficiency - full expenses" or "efficiency - safety (risk)".

C. *Criteria of multi-level optimisation type "efficiency - expenses",* which provide the solution of optimal distribution of limited resources between systems of two subordinate ranks.

D. *Criteria type "time of development - labour costs",* which connects the requirements of TF CS with the state of automation devices, possibilities of design departments and producing factories.

The criteria, which determine the structural organisation are the group A and B criteria, which presuppose the conducting of an analysis and quantitative assessments of the determined and probability values of safety, survivability and efficiency of CS, and also the system values of reliability through the results of object functioning.
In order to fulfil the following analysis patterns of a different nature, level and assignment are being used [8]:
- simulation patterns of physical processes dynamics in order to receive information (knowledge, facts) about the consequences of faults;
- structural patterns in the form of a logic equations' system or Boolean differential equations, in order to receive structural functions of reliability, survivability and safety, and also a graph of system degradation, which is the initial factor for the calculation of both determined and probable values.
For probable values estimation, patterns and programs for PC have been worked out on the basis of logical and probability methods, patterns of marked processes and accelerated statistical simulation [7].

Assessment patterns of reliability, survivability, safety and efficiency. Quantitative assessment of safety levels, efficiency and system reliability is made on the basis of analysis and information processing, which is laid out in the graph $\Gamma(S,U)$ of "Control object (CO) + Control system (CS)" integrity complex state of evolution, depicting all possible trajectories $\{X(t)\}$ in the process of its degradation and recovery. Multitude of apexes of graph S (complex states) differ :

- by combinations of failed and operative components of CO and CS with regard to failures in time order,
- presence of evident (displayed on CO behaviour or detected by means of control) or "hidden" failures of components,
- by the level of quality $\Phi[X(t)]$ of TFC functioning (output, capacity, manoeuvrability, dynamic accuracy),
- amount of losses of Πij during transition from i to j status and other criteria.

In the multitude of S states are distinguished :

- sub-multitudes of efficient states Sp with different accessibility for the system's higher quality ranks of the object functioning Φ,
- sub-multitude Sof of the states of functional failure, where TFC functioning quality level is lower than the accessible level Φ_0,
- sub-multitude Soe of the CO + CS complex emergency states, characterised by irreversible loss of TFC workability (failure) or violation of safety conditions for the staff or related systems.

In the sub-multitude of efficient states Sp can be a particularly selected sub-multitude of pre-emergency states Srp/e with hidden failures of emergency protection means or reserve means for TFC safety provision, in which a long-time uncontrolled stay is potentially hazardous due to the possibility of a sudden transition in the sub-multitude Soe of emergency states, under the condition of the emergency situation in TFC or appearance of "evident" failures in CS.

The studies [1,4,7] present a method of formalised plotting of the graph of CO + CS complex degradation from the state of full efficiency to the final states of functional and emergency failures, which application does nor require direct combinatory re-selection of all possible states.

The values of a fail-safe condition (failure stability) allowance type d - and V-fail-safe condition (failure stability) are being used as calculated determined values of group A, where d- max. number of degradation transitions on the graph $\Gamma(S,U)$, which in any combination it does not lead the CO + CS complex into the sub-multitude of failure states Soe (or into the sub-multitude Sof - for failure stability assessment); V - the ratio of number $||Sp \cup Sof||$ of efficient states to the total number $||S||$ of CO + CS complex states (for assessment of functioning failure stability - ratio numbers $||Sp||$ in respect to $||S||$).

During calculation of the d value - of the fail-safe condition, transitions into Soe through sub-multitude Srp/e is reasonable to consider as one transition.

Supplementary values used to assess the safety level are the values of the whole probability.

$$\bar{P}_\pi(t) = \text{Probability}\ \{\bar{\exists}\ \tau \in (0, t)\ :\ X(\tau) \subset Soe\}$$

and also the values of conditional probability $P_y(t)$ stating, that under the condition of the appearance of any sudden fault in TFC equipment, the trajectory $X(t)$ of CO + CS complex does not come into S_{oe} sub-multitude of emergency states.

Probability values of automated TFC and its CS efficiency, which include assessment values of functional efficiency E and the values of the whole (or the whole relative) expenditures Z during the whole working life of CS TFC.

E values of functional efficiency are the quantitative features of stochastic process of time alterations of the TFC functioning quality level $\Phi(t)$ under the random character of failures and operation recovery of CO and CS workability and look like

$$E = M \{ F [\Phi (X(t))] \},$$

where $X(t)$ - random trajectory of CO+CS complex alteration of states on the graph $\Gamma(S, U)$;
$F[\ .\ ]$ - function, characterising the type of value, which is being assessed;
M - mathematical expectation on process of $F[\Phi(X(t))]$ value change realisation.

The important values for assessment of CS TFC structure and determination of resources in stock are the values of functional efficiency

$$E \equiv \bar{P}(t) = \text{Probability}\{ \exists\ \tau \in (0, t)\ :\ \Phi [X(\tau)]\ < \Phi_0 \},$$
$$E \equiv K_r(t) = \text{Probability}\ \ \forall\ \tau \in (0, t)\ :\ \Phi [X(\tau)] \approx \Phi_{nom} \},$$

Speaking about the values of the whole expenditures it is important not only to assess the losses for recovery of equipment workability, but also taking into account the volume of losses, averted by the functioning of CS TFC, i.e. failures and emergency situations, appearing in TFC due to natural low reliability of equipment. Under assigned functional and algorithmical structures of CS TFC, the values E and Z are converted in the values of CS TFC system reliability, which assess the reliability of the whole system, but not in accordance with the combination of completion probabilities of functional-self-independent operations (FSIO).

During calculation of values for integral assessment of multi-functional CS TFC reliability and efficiency, the following specific features of CS and CS + TF complex in whole are being considered:
1. multi-functional and multi-mode character of CS and TFC;
2. interconnection of physical processes in TFC, the possibility of a cascade (domino effect) character of units` damages development during failures in CS and TFC components;
3. different types of component failures, selected according to classification features (evident and hidden); failures of non-operation, excessive operation and faulty operation; safety (non-emergency, or simple) and emergency hazardous operation, etc.);
4. all possible sequences of appearance or the sequence of time order of evident and hidden failures of CS and TFC components;

5. malfunctioning failures, which occur in electronic devices and microprocessor devices (with regard to the results of actions, made by the means of its detection, to protect and recover information, and to exert influence on the progress of algorithms and program execution);
6. characteristics of failure, faults and malfunction detection means and information recovery, in particular :
- completeness and authenticity (with division on errors of I and II class) of proper functioning control,
-completeness and authenticity (with division on errors of I and II class), periodicity and efficiency of readiness control or control for devices workability, which operate in standby modes for the receipt of applications for functioning,
- the time for malfunction detection and resolving power of diagnostical control, etc. ;
7. characteristics of algorithms and programs to solve the tasks of information processing and control;
8. influence of function features and (or) emergency protection sub-systems, anti-emergency control and localisation of emergency situations on reliability and fail-safe operation (safety) of CS and automated TFC in whole;
9. dynamic characteristics of automated TFC in emergency situations, under conditions of failures in TFC and CS;
10. real abilities of operator to make a well-timed transition to reserve types of control or realisation of CS reconfiguration during failures of components and partial damages of the system and lines of communication;
11. different types of units` and executive bodies of TFC redundancy, which are being used in TFC and CS (time, algorithmical, functional, information, structural, "fragmentation" of power (output));
12. limitation of resources for recovering (list and volume of spare parts and accessories, quality of "repairmen" or repair units, accessible duration of recovery).

The models of dynamic reliability and safety. Reliability, as an independent criterion, can be insufficient for some types of control systems. Let us take one typical example. For instance, the process, being under control is displayed in phase spacing, and the task of control is the transition of the object from some initial point A to the terminal point B in compliance with some criterion of an optimum state. It is absolutely not important why this event does not take place (arrival of the object to the point B). Maybe the reason for that is the low level of reliability of the control system, or the system was not accurate enough, and on the contrary, only approximate (as a result, the phase point moved far apart from point B), or the control system was accurate, but not reliable, thus the movement up to point B did not take place. For this type of system it is reasonable to unite the reliability criterion and accuracy criterion into one criterion of efficiency, indicating the possibility of system arrival in the assigned area, surrounding point B.

For a number of automatic regulating systems (systems of TFC co-ordinates dynamic stabilisation under conditions of continuous impact of environmental disturbances), the final result of its functioning is not determined by the reliability features alone. Also, the task can not be solved due to insufficient accuracy (statistical, dynamic). An increase of accuracy, normally, leads to the reduction of system reliability (replacement of simple measuring equipment or transducers for more complicated ones, and thus less reliable; complication of information processing and control algorithms, which means the reduction of equipment's safety). In such situations, particularly when there are limitations for weight-dimensional characteristics and reservation possibilities, there is a necessity to solve "contradictions" between the necessity to improve dynamic features and provision of reliability. The outcome - to find such value, which will characterise the probability of successful task solution and which will organically consider the characteristics both for dynamic features and equipment reliability.

Let us consider the system of stabilisation of some controlled object, under the impact of environmental disturbances. The task for control is to hold the co-ordinate X (t) in the zone $-a < X(t) < +a$. A violation of the indicated zone shall be considered a malfunctioning (or object failure). The reason for that "malfunctioning" may be the imperfection of a control algorithm or equipment failure.

Suppose that an accidental change of X (t) co-ordinate is stationary and normal, so it can be completely determined by correlation to function K (t). If the boundaries $[-a, +a]$ are far enough away from the probable process value m, in other words the conditions are executed

$$m_x + \sigma_x < a, \qquad m_x - \sigma_x > -a,$$

where σ_x - root-mean-square deviation. Thus, the distance between adjacent points of X (t) co-ordinate's pass out zone boundaries will exceed the time of process correlation. It is known, that $a > 0.5\sigma_x$ moments, when the process passed out the zone boundaries can be considered to be uncorrelated. At these conditions, "the flow" of X (t) passing from the zone $[-a, +a]$ is a stationary Poisson, or the most simple, where the possibility of appearance in Δt interval, which is arbitrary placed on the interval (0, t) is equal to k deviations and is determined by an expression

$$P_k(\Delta t) = \frac{(\lambda_d \Delta t)^k}{k!} e^{\lambda_d \Delta t}, \quad (k \geq 0),$$

where λ_d - the intensity of X (t) process deviation flows over the accessible zone per time unit. Index "*din*" means, that λ_d is determined by dynamic features of an automatic control closed system. It is evident, that the quantitative value λ_d is lower, when control algorithms are more correct.

The task to determine λ_d results to well-known task of function of a random variable theory about average number of intersections of assigned level by random process.
For stationary and normal processes the intensity λ_d in this case is determined by an expression (Stratanovich R.L.)

$$\lambda_d = \frac{1}{\pi}\sqrt{-\frac{K''(0)}{K(0)}} \cdot \exp\left[-\frac{(a - m_x)^2}{2 \cdot \sigma_x}\right]$$

The truth of the hypothesis, concerning the simplest character of deviation flows confirms the statistical analysis of real technological processes (see, e.g. "Lectures on mass service theory" by Gnedenko B.V., Kovalenko I.N. KVIRTU, 1963).

In some cases, the characteristics of deviation duration ("natural recovery") may be of some interest.

The above-mentioned ratios enable us to write down the expressions for probabilities to hold the X (t) co-ordinate in accessible, for example, safety conditions, zone $[-a, +a]$ with the combined consideration of :
- system dynamic features,

- environmental disturbances characteristics,
- equipment fail-safe characteristics.

In the simplest occasion, on the interval (0, t), the possibility of a successful task solution of object stabilisation with *i* alternate variant of algorithm A_i for information processing by the system, is determined by an expression :

$$Pi(t) = \exp\{-[\lambda d(Ai) + \lambda e(Ai)] \cdot t\},$$

where λe (A_i)- intensity of equipment failures, which realises *i* variant of control algorithm.

The values of this kind should be called the values of technical efficiency.
It is evident, that the supposed efficiency values are an important means to find a "harmonic" combination of the system dynamic and reliability features, to make assessment of development expediency and improve the information processing algorithms.

Formalised methods of complicated systems` structural functions derivation during analysis of its reliability, survivability and safety. One of the important stages, which precedes the direct calculation of determined and probable values of safety, reliability and efficiency of automated technical complexes (ATC) is the stage of its structural analysis. The structural analysis is aimed at receiving the appropriate structural functions (SF), safety functions (SF), workability functions (WF) in the form of logic algebra functions (LAF).It is made under conditions, when the levels of ATC safety, workability and efficiency are independent of its components` failure sequence, or the graph of degradation (GD) is [1-3]. The tasks of receiving SF and GD plotting of multi-unit ATC with an irreducible structure are excessive and have combination complexity and are related to the NP-complexity grade [5].

The analytical approach towards receiving SF and GD plotting for a wide range of ATC is developed in [3, 4] studies. It was suggested by [6].They are characterised by the presence in their structure:
- closed circuits, which provide self-support of energy processes and independent sources of power supply,
- "bridge" type components, which provide both bi-directional transfer of signals and signal transfer (energy, substance) in one of the available directions, which is chosen by a control system in regard to the state of ATC,
- reserve units (mechanisms, devices), which are automatically put into action when the main becomes inoperative,
- units, which require functioning of emergency protection sub-systems (EPSS) to eliminate the domino effect of failure in ATC,
- units of automatic reconfiguration of ATC structure, which algorithms exert a decisive influence on ATC workability and safety level during any combination of failures.

The nature of developed methods is to elaborate a binary genetic simulator of ATC components' mutual functioning. It represents a closed system of logic equations (SLE). Each equation connects an indicator of normal functioning of the corresponding first component with an indicator of its personal workability and indicators of component aggregate normal functioning. These indicators are directly connected with the first indicator and support and stipulate its operation. Self-checking indicators of emergency hazardous components are connected by logical conditions with indicators of EPSS normal functioning. Methods of SLE successive lowering, iterations, Boolean determinants and others can be used to solve SLE. It enables the reception of necessary SF in an analytical form.

For ATC, which require GD plotting, a specific procedure is proposed, which includes:
- acquisition of SF and reduction to inverted form (ISF),
- ISF transformation into SDNF,
- presentation of each conjunction in the form of conjunctions` aggregate with all possible transpositions of failed components` indicators,
- simplification of received expression in compliance with Boolean algebra,
- presentation of received LAF in bracket shape, which analytically depicts the ATC degradation graph.

Suggested original plotting method of TF + CS functioning structural simulator in the form of closed systems of time, multi-cycle logical equations enable the consideration of the cascade (domino effect) character of emergency situations` development, when components and its links become inoperative. There was developed a number of analytic solutions of logical equation systems (a method of successive lowering of rank, method of iterations and others) for formalised derivation of logical functions. These functions characterise the conditions of TF + CS complex as a safe and fail-safe operation and, also the preservation of accessible levels of operation efficiency at any number of combinations of fault components and complex links.

One of the works being fulfilled was the creation of a multi-purpose program complex (MPC) designed for qualitative assessment of nuclear safety levels (NS), system reliability and automated nuclear-propulsion plant survivability. The NPP has different assignments (shipborne, stationary [7]. The developed MPC is designed to solve the following tasks (solution of functions), which are aimed at the provision of modern requirements to safety levels, reliability and survivability of NPP:

1. Qualitative assessment of NS project levels, NPP reliability (including durability) and survivability during the design period (modifications, reconstruction). It is also made during expert evaluations and comparison of alternative versions of NPP separate systems and its control system and the acknowledgement of advanced projects to the requirements of the technical assignment, specifications and existing domestic and international standards (norms).
2. Quantitative assessment of NS existing levels, NPP and its systems reliability on the base of received systematised data concerning failures and equipment running time, volume and completeness of recovery and preventive studies, including the assessment of NPP and ACS TP equipment resource remainder.
3. Prompt assessment of current and forecasted risks' level in case of violation of NS conditions. This assessment is made for decision-making, concerning a ship or a NPS power complex structure reconfiguration with mapping of current and proposed changes in structures and risks' level assessment results. For this purpose, means of technological diagrams (in the form of generalised and detailed mimic flowsheets and histograms with logarithmic scales) on the terminals of the operating staff or station shift chief's terminal (chief executive) and NPS executive staff are being used.
4. Training (retraining) of NPP and control systems designers, operational, service and command personnel, having the purpose to master updated methods of NS levels, reliability, survivability, VAB methods, composition of formalised analysis patterns (dynamic, structural, probable) assessment and direct calculation.

When development of MPC was carried out, the designers took into account current domestic (branch and state), international (IAEA) standards, standardised, guiding and methodical documents concerning NPS safety, reliability and survivability requirements, range of values and its calculation methods, also updated results of safety, survivability and reliability theory, related to multi-unit technical complexes with a high risk operation.

Designed MPC differs from well-known blocks, which implement so-called PSA-technology and has the following features:

1. Automatic plotting of structural patterns (e.g. in the form of logical equations system) to analyse NPP and control systems` safety, reliability and survivability, related as a whole complex due to its actual structure diagrams. The computer input of these diagrams is made by using specific graph editing program.
2. Automatic output with the usage of developed mathematical instrument concerning safety structural functions (reliability, survivability) in the form of logic algebra functions (minimum sections of danger or failures), complex structural states degradation graphs, trees of failures or events with its mapping on displays. The form of display can be in analytic, graphic and other suitable forms of perception.
3. Provision of user interface with a computer in a dialogue mode with developed graphical form. In particular, during correction of danger trees, (failures or events), which had been received formally with the input of additional, hard to formalise initial (actuating) events, which appears due to the operator's possible errors, extraordinary impact of an emergency environment and resulting failures.
4. Calculation of risk levels concerning NS conditions` violation as a very rare event, and also determined fail-safe operation and failure safety values of type d- and v- with indication of their above-critical values and "narrow" places in the current structural state. It is necessary for decision-making, related to the allocation of extra anti-emergency resources from the available types of redundancy.
5. Accomplishment of NPP efficiency and systematic reliability values` qualitative assessments and assessments of fail-safe operation and NPP readiness integral values, which are calculated in the form of determined functionals from random processes of unit power levels alterations.

The block contains: NPP typical components' database (DB) (more than 1000 components) with characteristics concerning the component reliability (fail-safe operation, durability, fitness for repair or recovery, retention), and also with characteristics, concerning the component stability and communication tracks of damaging environmental impacts, such as fire, flooding (using reference, test and design data). The editing program with an analysed system database, consisting of typical components and stored in DB of typical elements, elements, which are input from the keyboard; editing programs for the creation of failure and event trees; calculation programs for calculating probabilities of end states of failure and event trees; program for input and visualisation of component failures on a mimic flowsheet; program for communication with the database; program for assessment of current risks and their presentation on the display in the form of histograms. It is presupposed to make an input of statistical data concerning running time and failures of NPP structure elements, which are received from maintenance and repair services. It is necessary to assess values of practical reliability, levels of risks and resource remainder.

MPC is developed with the use of PC type IBM PC/AT, monitor VGA and mouse in MS DOS 3.3 and higher, Clipper 5.0 and its minimum volume is 2 MB on hard disk and up to 0.5 MB of RAM. It is possible to transmit the processed data through telephone channels by using a modem. Optionally, the data can be transmitted to other Centres. The link with common ships` or common stations` communication line is envisaged. It is presupposed, that MPC will use multi-task and multi-user network UNIX-like real-time operational system type QNX (quickUNIX), graphical system of X-Windows standard. WATCHOM SQL and other basic software. It is also presupposed to conduct development of switching means of engineering patterns (patterns of physical and technological processes in NPP, durability patterns, etc.) in order to receive knowledge about end failures in non-standard emergency situations, taking place during NPP operation.

Analysis and synthesis methods of control systems' structures on the basis of survivability criteria at different levels of information about environmental emergency impacts. In order to solve the tasks of analysis and synthesis (selection) concerning functional, organisational, topological and other structures of multi-unit and distributed TFC control systems it is vital to meet the requirements of survivability. These requirements are based on the necessity to provide survivability of automated TFC and to the ship as a whole.

CS survivability implies [1-3] a quality integral part of the system structural organisation. It is displayed during partial damages of equipment and lines of communication and means the ability of the system to maintain TF units' workability, which had not suffered the direct damaging impact of environmental emergency situations. It also means the capability to provide the complexes' fail-safe operation and the possibility to control its operation modes, maintaining extreme levels of quality and efficiency (output, capacity, manoeuvrability, dynamic accuracy), which match the efficient components' characteristics.

The analysis of CS survivability considers the behaviour of TF + CS integrity as an integral information and energy closed system. The values of potential and real survivability are being calculated, taking into account the complete and real stability of equipment and CS communication lines.

The values or functionals are being used for the survivability assessment. These values on one hand characterise the quality level or TFC efficiency, on the other hand, the level of fail-safe operation at different environmental emergency impacts.

The studies [1-3] give a classification of assessment methods, concerning survivability of technical systems with a complicated structure. This classification covers the questions : survivability analysis approach (determined, probability and the approach, which is being developed by the authors, which is based on the ideas and methods of the mathematical theory of decision-making in the conditions of indefiniteness, and which is the most adequate to the nature of the survivability problem; the types of operation simulators (dynamic and structural) and consideration of the reliability characteristics during the survivability assessment.

The studies give an original method of TF + CS complex operation structural simulator. It presupposes its presentation in the form of a closed system of logical equations, which enables the consideration of the cascade (domino effect) character of emergency situations (ES) development, when the components and links become damaged. The authors developed a number of methods concerning the analytical solution of a logical equation system (the method of level successive lowering, method of iterations and so on). These methods are used to make formalised derivation of logical functions, which characterise the conditions of safe or fail-safe functioning of TF + CS complex and preserve the accessible levels of operation efficiency at any combination of failures.

The tasks of CS structures selection (synthesis) comply with the traditional scheme : the selection of structure, which is most favourable by some criterion and best of all matches the essence of design task. This selection is made among the multitude of structure variants $S= \{ S_1, ...S_i, ... S_n \}$ on the multitude of incompatible states $H= \{ H_1,$

....Hk, ...Hm } of an emergency environment (EE) and at the presence of Function values multitude Y= { Yik }. Its elements characterise the levels of workability (efficiency, fail-safe operation, safety) at the i-variant of structure and k-realisation of EE. The authors studied several typical situations with different levels of AB indefiniteness [10]:

a) a priori is known from the distribution of conditional probabilities { Pk } of the event appearance { Hk } (and);

b) environment actively counteracts;

c) priori probabilities` preferences are known from physical premises $\pi_1<\pi_2<...\pi_m$;

d) there is complete confusion about the probable realisation of ES from { Hk }.

In situations (a) and (b)"revealing" of indefiniteness during system structure selection is made with the help of Buys criteria.

$$S i_0 = \arg\max_{s_i \in S}\{\sum_{(k)} \pi_k Y_{ik}\}$$

and criteria

$$S i_0 = \arg\max_{s_i \in S} \min_{H_k \in H} \{Y_{ik}\},$$

which are similar to the rules of antagonistic plays` theory. For the less favourable situation, from an information availability point of view (d), the authors proved the reason to use Hojes-Leman type criteria:

$$S i_0 = \arg \max_{s_i \in S^* \subset S} \{\sum_{(k)} \hat{\pi}_k Y_{ik}\},$$

where S* - sub-multitude of structures, for which

$$Y_{ik} \geq \alpha, \quad \alpha = const, \quad \min_{i,k} Y_{ik} < \alpha \leq \max_{i,k} Y_{ik}, \quad \hat{\pi}_k = \frac{1}{m}.$$

For situations (c) in Hojes-Leman criterion, it is reasonable to follow the recommendations of Fishborn

$$\hat{\pi}_k = \frac{2(m-k+1)}{m(m+1)}.$$

To make the assessment of survivability under well-known EE for the whole system and with the of "point-like"EE it is necessary to use values of the survivability endurance type d- and V-survivability.

Conclusion. By now, the experts of the Scientific and Production Association "Aurora", in co-operation with the specialists of the Institute of Control Problems, Russian Academy of Science, Naval Academy and other institutions of the Ministry of Defence, worked out the theory, methods of mathematical simulators` plotting and information technology concerning complex analysis and quantitative assessment. The assessment concerns : reliability, survivability and the safety level of the structure complicated and potentially hazardous multi-unit systems with a high risk of technical, ecological, social and economical operation. The above-mentioned studies were widely tested in the course of solving a considerable number of practical tasks for assessing reliability, survivability, safety (including nuclear and radiation), efficiency of nuclear-powered ships and vessels. These features are considered to be an aggregate of a large number of interacting systems, which are created by various organisations and departments and provide a complex assessment of the whole project on the level of Chief (General) designer and are able to find "narrow" places, also possible "faultiness" of the structural system, which are invisible at the level of separate sub-systems. They provide co-ordination of works in designing such large objects, with development of mutually matched requirements and/or determining (principle) technical decisions with rational distribution (re-distribution) of limited recourses among sub-systems. It starts from the initial stages of designing - i.e. stages of concept development, vanguard project, sketching the technical project. The structure, as an instrument of a Chief Designer can be fully used in designing different objects of nuclear power engineering, in particular floating nuclear power stations, electric power stations, nuclear ice-breaking and cargo ships, during projects` expertise, in relation to separate systems, which are completed for NPS programs of re-equipment.

Reference.

1. Simakov I.P. Criteria of safety, survivability and efficiency of control, consecutive synthesis for control system formation. - in Collection : IX AllUnion Conference on control problems. Moscow : Nauka, 1983. (in Russian).
2. Astrov V.V., Simakov I.P. The methodological basis of synthesis concerning control systems` structure on the basis of safety, reliability and survivability criteria. In Collection : X All-Union Conference on control problems. Moscow, 1986. (in Russian).
3. Simakov I.P., Astrov V.V. Methods of control systems structure selection, on the basis of survivability criteria at different level of information availability, concerning environmental emergency impacts. In collection : XI All-Union Conference on control problems. Moscow : Nauka, 1989. page 232 - 233. (in Russian).
4. Simakov I.P. Analytical methods of structural analysis, concerning automated technical complexes for the assessment its reliability, survivability. In Collection : Conference on reliability, survivability and safety of automated technical complexes. - Suzdal, 1989. (in Russian).
5. Gery M., Johnson D. Computer devices and tasks difficult to solve. Moscow : Mir, 1982. (in Russian).
6. Astrov V.V. Simakov I.P. Cherkesov G.N. Utilisation of probability logic methods and operations research concerning the survivability of distributed power systems. - In Collection : Reliability of power engineering systems. Issue 20, "Survivability of power engineering systems". - Irkutsk, USSR Academy of Science, 1980. (in Russian).
7. Simakov I.P. Program complex for probability analysis of safety, related to automated shipborne NPP and NPS during its design, operation and direct control. In Jubilee scientific and technical Collection of "Aurora" SPA, St. Petersburg, 1995. (in Russian).
8. Voitetsky V.V., Symakov I.P. Development of methodology, theory and principles of naval ships CCS TF and control systems ACS TP, technical complexes with operation of a high level risk. See the same, p. 46-48. (in Russian).
9. Bochman D., Posthoff C. Binare dinamische systeme. Academie - Verlag, Berlin, 1981.
10. Трухаев Р.И. Методы исследования процессов принятия решений в условиях неопределенности -Л.: ВМА, 1972.
11. Ryabinin I.A. The concept of the logical - probability theory of safety. - Gears and control systems, N10, 1993, pp. 6-9. (in Russian).

NOOSPHERE AND INFORMATION

Academician V. V. Demianov

Noosphere is literally a sphere of Intellect. But a sphere is primarily a space. Does it mean that this is a ***space of Intellect***? "How can it be?",- this will be a "rightful" question from a person who, since the early childhood has been taught to consider "Spirit", Conscience, Intellect and their products - ***thoughts*** - to be some ideal ("certainly non-material") aspect of the human existence physiology. This is one of the worst mankind's delusions which have roots stretching back into antiquity. Because just like in those times, now each person who can count candidly believes that the count of noospheric values equals to the opus of a savage using the finger for counting: ***"1 finger+1 finger=2 fingers"***. And this is the way the people, up till today and with the condescension of the greatest winners of Noble Prize in Economy, continue the works of the savages, being sure that 1 dollar + 1 dollar = 2 dollars (in the sphere of monetary relations), that 1 bit + 1 bit = 2 bits (in the sphere of information exchange), etc. And it does not occur to anyone that in a complete analogy to the electrodynamic reality of the "physical" part of the multidimensional World (where it has long been known that $1_{\text{active Volt}} + 1_{\text{reactive Volt}}$ is always determined by the quadrature of $1_{\parallel}\hat{A}+1_{\perp}\hat{A}=1{,}41\hat{A}$) respectively, in its "spiritual" spheres of information exchange (circulation), commodities/money turnover and calculations of other spiritual values (N), more often than not there are similar orthorelations which produce $N_{\parallel}+N_{\perp}<2\bullet|N|$. Up till now, unfortunately, noosphere, just like the concepts of "sphere of knowledge", "sphere of activities" and many other "spheres" of informational manifestation of the experience and thought, have been understood by people figuratively, i.e. as certain spatially unreal images of intellect. Euclidean sphere, discourses a "materialist", is quite a different matter: this is an " objective reality given to us in our perception", being allegedly "absolutely obvious" and not requiring any proof of its reality and its predetermination.

Whom is it given by? It is in the absence of an answer to this question where lies the entire confusing essence the way the reality is understood my "materialism" and its "shadow" in the people's existence: "common sense". This "viewpoint" of materialism is in no way more sensible than Berclean view of the World (as a sum of non-material images of "sense"). Moreover, Berclean "fleshless" abstraction of intellect, with deep scientific approach to the problem of the World's spatiality, proves to be as close to the reality as the "soulless" materialistic concreteness of forms observations in the Euclidean space. This report is a short Prologue in the part of some problems which I raised in the first part of my book "Noosphere Evalectics" (Novorossijsk, 1995; you may look through this book here at the Conference), to its second part which I hope will be issued next year.

I consider noosphere as the most real (of all the known realities) information/code space which includes both, space and time. Noosphere is formed by the vector sum of three "sixspheres": space/pulse, time/frequency and probability/information. Each of these components if formed by a "direct" and "reversed" threespheres inserted in each other normatively locked on each other, as all the ("direct") originals are naturally locked on their ("reversed") images. The mankind started studying the first sphere (Euclidean sphere) a long time ago, at first understanding it as a Ptolemaic "flat reality") - before the 1st century B.C., then as a sphere to a full measure perceived by Euclid, after the 1st century B.C., and, finally, since recently as Minkovsky tetrasphere not represented by anyone (since the 20th century AD); I guess this tendency of the World's spatiality gnoseology will remain for some time. As far as the third (information) sphere is concerned, it has not yet been perceived by the real space whatsoever.

In fact, the noosphere's ontology as a complex space of ***expanse/time flow/information*** ($r \perp t \perp \Psi$) counts as many billion of years, as the time of Universe's existence (about 15 billion years). I consider it as a "straight" relict thrice folded sphere f the multi-dimensional World locked on the reversed sphere with the respective orthoaxes ($\hat{E} \perp \grave{A} \perp Z$), which contain, in the encoded object and wave forms, the entire ontological code of "the laws of nature" in the amplitude-kinetic, energetic and higher symmetry "simple" Constructs, capable of being synergetically arranged in the "complex" Synthesis of things, beings and subjects of Intellect. The second "reversed" orthoaxis (A) genetically connected with its "direct" (t) generating cause, is a consequence of the virtual rotation of the latter which is its abstract-image "project" synthesized in the Intellect per se. This "Berclean idealism" will seem nonsense to any "materialist". However, in the aforementioned first part of my book (and in more detail in the second part) I am proving that a still greater idealism is several thousand year old delusion that there allegedly exists some infinite "empty" receptacle of the World in the form of Euclidean space.

There is no such space and there has never been! Furthermore, there is not space filled with the Field, as there is no Field itself in any of its kinds "invented" as of today. With such a prospect of collapse of the entire "fundamentals of life" (and the space/time today is, in everyone's opinion, this very fundamental principle of any kind of existence), I believe even "materialism which cannot give up its principles" will have to think hard on what should be used for plugging the whole in the "lost reality given to us in the senses. It will be here that it will become clear that only the Intellect will be able to help us to understand which reality is detected by our senses. It will at last become clear that not the "Physiological existence determines the mind", but it is determined by the ***Spiritual existence***, which means that the mind is determined by the Spirit of the Nature whose active (ontological) essence consists in that it has brought to life all the levels of the World Freedom: both the substantial (solid) ***spatiality***, and the substantial/phase ***time flow***, and the substantial/information *contents* (in the laws of nature) of the World. The mind space ontology is in the basis of the manifestation of noosphere - sphere of Intellect (with material orthoaxes $\hat{E} \perp \grave{A} \perp Z$).

Evolutionally growing vector multiplication of objects of two base noosphere axes $[\hat{E} \times \grave{A}] \approx Z$ gradually develops the potentiality of mind of the Intellect subjects and noosphere's self-consciousness providing the phenomenon of the nature's self-gnoseology. So any act of perceiving something *new* =Z is only possible on the basis of the final experience =$\hat{E}$ and theoretic (image) knowledge =$\grave{A}$, mastered by the mankind as of the given moment of time, and is an elementary act of generation (extension $=\Delta$) of the "volume"(V_i) of the sphere of Intellect($\Delta V = \Delta \hat{E} \bullet \Delta \grave{A} \bullet \Delta Z$, where $\Delta Z \approx [\Delta \hat{E} \times \Delta \grave{A}]$). The triad ("Trinity") of the noosphere vectors has been spiritualising the life of all the things in existence for more than 15 billion years, gradually bringing into the harmonic space of nature the Intellect of mankind which originated from its heart.

One of the most mysterious niches of noosphere is manifested by the spiritually creative ***sphere of values*** of the Intellect's activity: this niche is a sixsphere of two interlocked threespheres: "direct" three-dimensional space of the "object-commodity mass" of the intelligent ontology of developing "complex" values from the "simple" raw material units of the nature (atoms, molecules, organisms, etc.) and its closing "reversed-image" information three-dimensional space of value-monetary (essentially informational) associate of this commodity mass which is a sphere in the orthobasis of the active ($N_{\parallel}$), reactive ($N_{\perp}$) and newly created ($N_{\approx}$) material/spiritual and "physical" values. In this sphere's space (in the noosphere) only the complex logic of Synthesis is true: $1_{\parallel}+1_{\perp}=1{,}41$; $1_{\parallel}+1_{\parallel}=2_{\parallel}$; $1_{\perp}+1_{\perp}=2_{\perp}$, etc. "Money" is the information associate of the produced "physical" values/commodities in the informationally whole space of the vector structural prices of these commodities. The noospheric mentality of the modern civilisation should be extended by the view of the presentation of ***Freedom*** not only as the space of expanse and time flow of the ***"Flesh"***, but primarily as the sphere of Contents (information realisation) of the ***"Spirit"*** of the World.

SOME ASPECTS OF THE DEVELOPMENT OF MARITIME AREAS IN THE SOUTH OF RUSSIA BY SINGLING OUT THE TRANSPORT AND ECONOMIC STATUS OF NOVOROSSIJSK REGION

V.V.Popov, V.V.Kashitsin, S.E.Ivanova, G.E.Panamarev, V.V.Demianov

1. Introduction

By signing after 1991 all the international documents governing the modern development of the Global Maritime Distress and Safety System (GMDSS) in seafaring, Russia assumed certain obligations to the World Community. As of today the radio communication is the only means which ensures a properly arranged assistance to a vessel in distress. Until recently this assistance was based on the special national emergency radio communications facilities; today we are faced with the problem of uniting world-wide and of working out a system approach to the development and use of the emergency communications facilities to enhance efficiency of the rescue and salvage operations at sea, as the losses caused by the accidents at sea have started to significantly slow down the progress of the merchant marine. The global coordination of work to prevent the distress situations and to ensure the rescue at sea, will allow creating an essentially new system (referred to above as GMDSS), which will fully duplicate (and many times) and secure the traditional use of manual distress radio communications.

Since 1991 this work has been coordinated within the framework of the world-wide program of GMDSS development; at the moment (up to year 1999) we are in the third stage of its implementation. Operating today, however, are more than half (and this comes to several hundred thousand) of the principal and backup shipboard and shore based radio facilities which do not meet the present-day IMO requirements to the GMDSS equipment. This is entirely true for Russia which has inherited from the former USSR the lag in this sphere of ensuring transport activities at sea. To fulfill the aforementioned obligations Russia shall bring its national coastal navigation areas (areas A1 about 30-50 miles in width and areas À2 about 150-200 miles in width from the own coasts) in line with the IMO international documents during the Third stage of implementing GMDSS all over the world (before 1999).

Due to the aforementioned historic circumstances, Russia received in the south "absolutely bare" (from the viewpoint of IMO's GMDSS requirements) coasts from Taganrog ("north" sea border between Russia and the Ukraine) to Sochi ("south" Russian sea border with Georgia); and Russia will have to build everything anew along the entire coastline specified above, about a thousand kilometres long, rather than "bring something in line with something". And Russia will have to do it by stages between 1996 and 1999 by establishing the Novorossijsk Centre of GMDSS Communications Control as early as in the first half of 1997. This article provides a brief description of the technical and economic problems involved in this construction.

2. On the Economic Feasibility of Implementing the GMDSS Network in the South of Russia

The southern coast of Russia from Taganrog to Sochi stretches to about 1000 kilometres and is fully devoid of any infrastructure which could perform functions resembling those of the GMDSS. If the tactics of referencing the newly created GMDSS network to the infrastructure existing in the area is to be ruled out, the expenses involved in establishing this network are estimated at 1 trillion roubles. In connection with the economic crisis, it becomes essential to estimate both, the economic feasibility of creating such system by the country as a whole, and the region's technical capability of carrying out the minimum required schedule of works in the aforementioned limited time interval.

Unfortunately, the possibility of widely using the tactics of referencing the newly created GMDSS network is considerably limited by the specific features of the latter (there has been no such system in this country before), by the inadequacy of the region in providing the public communications (such as TV centres and re-transmitter, broadcasting stations, trunk and cellular telephone networks, radio relay lines, etc.). Even the radio centre which remained after the former Novorossijsk Shipping Company and which ensures radio communications in areas À1, À2 and À3, is now in such poor technical condition that cannot be looked upon even as an initial infrastructure for the further works. The requirements to the national GMDSS networks are such, that the Maritime Administration of the Port of Novorossijsk (MAPN), together with the research and training institutions in the area (South Marine Geology, Novorossijsk State Maritime Academy, etc.) have to start developing the first stage of this network from the surveys. This is why the budget for these works includes the entire range of works starting from "zero":

– expenses on the research and survey;

– expenses on the scientific studies in connection with the Technical Specification for the future GMDSS network Project in the south of Russia;

– expenses on the science expert assessment of the development;

– expenses on the draft pre-project studies of the network's first stage fragments;

– expenses on the detail designs of the Novorossijsk area GMDSS network segment.

The total amount of all the Novorossijsk segment expenses listed above (from Anapa to Gelendjik) is estimated at about 40 billion roubles (in the prices which existed at the end of 1996).

Nonetheless, it can be seen from the aforementioned research works which served as the basis for developing a schedule (see drawing 1) of the first stage of the GMDSS network's Novorossijsk Segment (with other adjacent border segments), some options for using the existing infrastructure were found for implementing the cost reduction tactics. And although the existing infrastructure does not meet the optimum conditions in many parameters, it will be shown below that a way has been found to significantly reduce the cost of the first stage (almost by an order of magnitude relative to the cost of project "from zero" mentioned above), all the system's principal parameters remaining within the permissible (by IMO requirements) values.

At the same time, even with the cost reduction tactics intended to be used, the creation of GMDSS system amounts to the expenses so great that they do not allow the aforementioned segment to be now implemented by the resources of a single ministry or region. In this connection there arises a problem of searching for the economic models of the system implementation within the specified short period of time. It is obvious that they should not be confined to involving as investors the state maritime administrations of ports in the respective coastal transport junctions only. A number of organisations interested in the implementation of this project could be mentioned here: Federation Constituents, state agencies (especially the Ministry of Transportation of Russia, subjects of foreign economic relations (town and regional level, as well as major Russia cities, especially Moscow), representatives of small business and public: in a word everyone who is interested in the implementation of this project. It is necessary to remember the considerable role which could be played by the foreign investors, both actual and prospective. Such widening of the scope of participants in the project makes utterly important the problem of responsibility in the effective control and use of the invested funds, which is directly connected with the problem of their diversification in the process of the operation of the project which will follow.

It is clear that out of the prospective investors listed above, it is the direct users of the GMDSS systems, and the persons or entities who suffer quite tangible losses due to its absence, who should be involved in the project first of all. These are primarily "Novoship" joint stock company, "Novorossijsk Merchant Port" joint stock company. Novorossijskrybprom, all the transport enterprises of the Novorossijsk transport junction, as well as all the national and foreign companies operating regular services or/and export/import operations in the system's operating zone. There may be various forms, including the establishing of a state joint stock company (with the state warranty) with e.g., MAPN holding the control package.

The municipal authorities may be represented in the project by their property, transport-infrastructure network which they control. Besides they can participate through the municipal loan, issue of securities which could be distributed among both, physical persons and legal entities. The project could be supported by the district agencies and authorities who could participate in it with their relevant property, supply lines, and, what is most important, that share in the collected mass of taxes which they have at their disposal (e.g., the amount of taxes sent to the district centre from Novorossijsk comes to about 50 per cent of the taxes collected here). There may be, therefore, various forms of the financial support for the project: distribution of securities in the offset of tax payment, and the investment credit, and tax advantages, etc.

Correspondingly, a still more active support could be provided by the formation of free economic territories in the areas of the project sites. It is with this kind of approach that the interests of the Centre, Federation Constituents and Municipal Councils can be optimised and used in many ways. Specifically, as of today the most tangible source of funds for implementing any regional (let alone Federal) projects - the tax money (in Novorossijsk, e.g., 35 per cent of the collected tax money which goes to the Federal budget, could be in part directed to the financing of regional projects of GMDSS network type), obtain an effective return in the project exactly within the framework of a free economic territory. Within this same framework the investment credit, securities turnover, loans, insurance investments of the commercial banks, funds, formation of corporations and associations, etc. are most effectively used.

So much emphasis is placed on attracting the national investors to the project because, as shown by the international practice, in the overall amount of investments the share of foreign investors makes up about 15 per cent (even within the framework of a special economic territory). For the same reasons special attention should be paid to attracting the population's money, as this is one of the largest prospective investment bulk. Here a flexible system of control over the population's invested money, rate of interest, warranties and obligations is of special importance.

These are, in our opinion, some efficient market options of attracting funds for implementing recouped projects where a not entirely insignificant role is played by the problem of ensuring the profitability of a functioning GMDSS system.

3. Characteristics of the Science and Technical Potential of Regions in the South of Russia in the Context of Possibility to Create a GMDSS Network

The course to the development of the open economy and the involvement of the domestic manufacturers in the process of extending the economic relations makes a vital issue of expediting the engagement of the Novorossijsk potential in the market structure of the world economy, thus creating important pre-requisities for a fast self-financed development of a powerful production and economical complex in the south of Russia which would allow its influence in the Black Sea and Mediterranean area to be considerably enhanced. Assigning to Novorossijsk the status of a territory with a special legal and economic regime for conducting commercial activities will allow creating a favourable economic climate for the development of the Azov Sea areas adjacent to it, whilst the attraction of non-governmental investments in the industries and fields of activities having a strategic importance for many countries would dramatically accelerate the processes of reconstructing a new configuration of economically efficient of all the transport arteries in the south of Russia linking them to form single transport junctions around the advantageous sea ports in this area.

The port of Novorossijsk is historically the largest of such transport junctions. Therefore, as mentioned above, the Novorossijsk sea area is, for the time being, the only area whose resources can be used for starting the development of an appropriate federal program of constructing the GMDSS network in the south of Russia. The scope of these works is so great that it is hardly possible to provide the necessary financing of this large project without creating marketable mechanisms of self-development at the regional level. Under the present conditions it is possible to obtain rapid self-financing mechanisms' by reviving this program only through "Free Economic Area" institute in the Novorossijsk region. Here is where the interests of the federal centre and remote Russian sea region coincide.

Basically, all the forthcoming construction program of GMDSS network along the south coast of Russia may be regarded, ***first,*** as part of the country's conversion programs, as up to 1991 all this coastal area, especially along the Black Sea coast, was "militarised" with strict frontier regulations. All the best heights for placing the coast radio communication stations are now occupied by the military radio technical structures whose number is (in the modern conception of Russia development), no doubt, redundant and to state budget's disadvantage.

The coordination of the Conversion Program for Naval constructions with the building of national GMDSS network along the south coasts (under simultaneous reconfiguration and optimisation of the transport system which Russia gained "torn" in 1991) could essentially lighten this capacious complex of state and regional construction. The fact that at present the power of the state budget is not sufficient, makes this matter all the more topical, as the resources of the regions (Taganrog, Azov, Yeisk, Temryuksk, Anapa, Novorossijsk, Tuapse and Sochi) of Rostov district and Krasnodarski krai for all these programs have never been adequate, and naturally, are not sufficient today. One needs a new fundamental idea which would be useful not only for the Centre, but also for all the aforementioned regions of the south of Russia.

It could be possible to solve this entire set of state and regional problems in such a short period of time (without withdrawing the resources withdrawal from one company and sending them other companies) by including the whole coast line with the territories of the aforementioned seacoast cities of the south of Russia into a separate economic area with marketable economy temporarily relieved from taxes. The area budgets self-financing factors of the aforementioned cities, brought up by such steps (in the "coast garland" of the south of Russia area) will allow ***to implement,*** in the shortest possible time, the program of GMDSS network construction along Azov-Black sea coast without any load to the State Budget, ***to solve*** part of conversion programs with re-configuration of the military coastal objects of the Black sea coast in connection with the geolocation of the Navy battle array which have changed since 1991 with simultaneous re-structuring of the made discord in the transport complex functioning in this part of the country, caused by the loss of Crimea.

At the moment, practically all the mentioned coastal cities have "Port Authorities" of the appropriate titles; they are charged with implementing he plans of GMDSS network construction in the south of Russia. However, as mentioned before, only Novorossijsk Port Authorities together with the scientific resources of the region (Juzhmorgeologia, Novorossijsk State Marine Academy) proved to be the only capable of functioning structure which had begun, basing on its local facilities (collected from the maintenance of the old Novorossijsk port constructions, from anchorage of Tsemessky bay, etc.), the implementation of the GMDSS network construction in its responsibility area (from Anapa to Dzhubgy with the length about 200 km). That is the reason why the first turn's implementation of the Program under discussion is made by Novorossijsk Port Authorities efforts. For the structural diagram of the informational GMDSS flow organisation for Anapa-Sochi sub-system see figure 1.

4. Assessment of Costs of Creating a GMDSS Network Segment in MAPN's Zone of Responsibility

4.1. Costs. As the GMDSS network is only being created at the moment, this study will be based on the current costs of developing only one segment (in MAPN's zone of responsibility, see the diagram in Fig. 1) of the future network, whilst the overall costs of creating the entire GMDSS network in the south of Russia will be estimated by extrapolating these costs to the whole geographic area. The assessment is directly connected with the analysis of the useful properties of these facilities inherent in any value in use, and is made primarily at the stage of the buyer's "intention" and manufacturer's "offer", but is finally ascertained in the process of operation and is based on the unity of the value in use and the established duration of the equipment operation whereby the invested funds are recouped in the form of certain property or values. The time factor has been included in the expression of the value in use due to the fact that the equipment's being able to meet certain public needs implies a certain duration of its use during a normative operating period when this equipment performs a strictly specified production function, whereby its particular value in use obtains its final expression. For a manufacturer of a certain product who sells this product (tools, equipment), this product does not have any value in use. For this manufacturer the output equipment only has an exchange value, where the measure of value is the correlation of the full production cost and wholesale price at the sales market.

The assessment of the technical product's value in use is made by the buyer. It is the buyer who assesses both, the equipment's operation efficiency and its consumer qualities which enable its purpose to be fulfilled in the spheres of its use. Apart from the value in use, the price at which this new equipment is bought is also essential for a buyer. In the price structure a full cost's share is high, but the prices should be determined not only by the cost of production, but also by the useful effect which determined the demand for this equipment.

The fact that the prices for new equipment take into account the useful effect allows its potential economic efficiency to be most fully determined, and the condition to be fulfilled whereby the resultant part of the new equipment should exceed the cost part. When the new equipment is created and when a buyer is selecting between its various samples at the market, it is important that, as early as at the stage of its pre-design study, the assessment of its up-to-dateness be made, all the types of costs involved in both its design, building up and future operation conditions, be comprehensively determined. It is only the deliberate isolation of the equipment's consumer qualities which enhances its ability to fulfill its purpose, provided the safety, operation reliability, ecological and operation efficiency and other requirements are met, that ensures the economic interest in the purchase and introduction of new equipment.

The contents of the useful effect could be presented as the implementation of the comprehensive manifestation of the new equipment's value in used. Expenses (i.e. the cost) involved in its production should imply and facilitate the conditions of realising the fullest possible useful effect and should, therefore, be functionally related to the expected level of the efficiency factor. This is what stipulates the necessity of the tentative (preliminary) assessment of the efficiency level in the use of new equipment , as its costs of creating are this equipment are compared to its useful effect in the process of its operation.

The Global Maritime Distress and Safety System (GMDSS) in the south of Russia should be built up in accordance with a two-level schedule:

The first, "external network" level: the base stations "Ploskaya", "Lysaya", "Verkhnebakansksys", "Doob", - will insure the information unity of all the base stations combined into a single regional GMDSS network in the south of Russia;

the second, (internal" level: Communication Control Radio Centre (CCRC) and Rescue Coordinating Centre (RCC) deployed in Novorossijsk.

The expert assessment shows that the overall cost of constructing the control radio centre and GMDSS base stations only in Novorossijsk region amounts to about 15 billion roubles. Besides, the costs of the survey works, pre-design works and the detail design on creating the GMDSS in the south of Russia have come to about 1–1.5 billion roubles. These costs have included: development of the Technical Specification of the design of the GMDSS system in the south of Russia; expert and pre-design study of the GMDSS system, elements. The aforementioned assessments of the costs on creating Novorossijsk segment of the system are the base ones in this research and are used for extrapolation assessments of the full cost of the entire GMDSS network in the south of Russia.

4.2. On the recoupment of costs. The recoupment of the aforementioned costs can be only determined by the magnitude of the useful effect which has a social character of the national and international levels, as the use of GMDSS equipment for alerting on the distress at sea will provide certain guarantees of safety of a human life at sea. This primary effect will consist in the reduced number of human casualties at sea due to the timely alerting on the distress at sea and reduced time of the search and rescue operations. The international trust to the operation of the GMDSS network deployed in the south of Russia will increase the trade turnover on the country's seaways and the carriage of passenger in the development of tourism and recreation business.

Another aspect of the effect of introducing the GMDSS network for the efficient prevention of and assistance during the distress at sea could be considered as an ecological effect which manifests itself in forestalling, preventing or reducing the environmental pollution which is commonly associated with the accidents of the oil carrying vessels. So, in the cases of oil spills, unless they can be prevented by fast alerting on the emergency at sea by way of transmitting accurate enough coordinates of the distress position, the development of the sea pollution processes can usually be prevented or considerably reduced by stopping an oil spill, treating the oil spills from the aircraft, etc.

No less important is the third aspect of the recoupment of expenses on the construction of national GMDSS networks. It is necessary to add to the reduced social and ecological consequences, the effect of decreasing time for performing the search and rescue operations, which produces direct cost reduction due, primarily, to the lower expenses involved during thus reduced period of operation, as well as the direct reduction of material losses due to the timely assistance which usually enables a considerably part of material values to be saved on the distress ship which received such timely assistance. The efficient reduction of costs involved in the search and rescue operation is achieved in the following cases of their efficient organisation at the CCRS and RCC of the regional network:

– with the maximum shortening of time required for the operation, achievable exclusively with the highly reliable operation of all the network's technical facilities and efficient organisation of all its services. The advantage will be gained due to the reduced time required for alerting and minimised time required for the search of the distress object, as well as due to the fast and timely alerting and accurate enough position fix.;

– with the smaller number of special rescue vessels and the vessels within the distress area, whose operation is forcibly (or due to the poor organisation) disturbed by them being assigned with extrinsic functions during the rescue operations. Here the economy of material values can be obtained due to the reduced running costs of the rescue vessels and reduced costs involved in deviating the vessels which alter their course and proceed to the rescue area.

5. Extrapolated Assessment of the Cost of the Entire GMDSS Network in the South of Russia

In Fig 1 above of the GMDSS subsystem we can distinguish between two unequal topological structures of the network: Novorossijsk segment (it is marked off with a dashed frame), which is a logically closed subsystem capable of functioning on its one in sea areas À1 and À2; and the rest of the structure which reflects the geographic configuration of the Southern coastal area of the country, but which, at this stage of the project, is not yet closed to form a single network and not capable of performing the minimum number of functions required for the GMDSS networks. The idea of extrapolated assessment of the cost of the entire GMDSS network in the south of Russia is this reduced to the possibility of extending the already known costs in Novorossijsk subsystem configuration (Novorossijsk segment) to the other geographical regions taking into account the identical and differing topological features of these subsystems. We will consider here two more segments approximately equal in length to the Novorossijsk segment (about 300 km): Tuapse-Sochi segment and Azov segment.

When substantiating the costs of creating an intended GMDSS station network on the Azov part of the coast, it is necessary to note the principal pre-requisities which we used as the basis of the extrapolation method of the cost assessment of their implementation. The analysis of the practice of creating GMDSS networks, as it exists today, prompts that it is necessary to take into account two groups of criteria implying the ***similarity*** of conditions and their ***difference.*** The first group includes the single equipment type, similarity of its functional purpose, single requirements to the reliability, similar operation and servicing conditions, similarity of climatic conditions, etc.; all this allows using for extrapolation the normal law of distribution of qualitatively similar costs of approximately the same average scale. The second group may include considerable geographic differences (the height of the terrain), which imply the construction of tall tower like buildings in the low land Azov area and, therefore, higher cost of providing the necessary height of the antenna installations in the radio centres of the low land areas as compared to Novorossijsk mountainous area; these higher costs (item 4 in tables 2-5) form a factor which may basically increase twice or even thrice the cost of creating equivalent (in their range) antenna units . However, in the cases of referencing to a more developed infrastructure of the Azov coastal area of Russia the aforementioned expenses involved in the construction of tower like antenna foundations can be significantly reduced. This is why in assessing the costs of building the first stage of GMDSS network in Azov segment, the costs may be considered as close to the estimated costs for Novorossijsk segment (e.g., with higher coast co-efficient of about 1,3–1,4).

As far as Tuapse-Sochi segment of GMDSS network is concerned (about 300 km long), with a considerable similarity of the geographic, climatic and technical and operational conditions, the costs (within the specified framework of the first stage) can be assumed to be approximately equal to those estimated for Novorossijsk region. Therefore, the national GMDSS network which has to be deployed in the south of Russia before 2005, at this stage of study has been conventionally divided into three segments approximately equal to each other from the viewpoint of expenses: Novorossijsk, Tuapse-Sochi and Azov segments (in Fig. 1 they are marked as S_1, S_2 and S_3 respectively). The actual costs of creating one of these segments Novorossijsk, S_1), the construction of which is being underway now, have been fundamentally determined. With regard to the above, the total expenses involved in the construction of the first stage of the entire GMDSS network in the south of Russia (from Taganrog to Sochi), whose structure was minimised as shown above, will come to about 100 billion roubles (in the prices which existed at the end of 1996).

These are "net" expenses on the construction of the national GMDSS network material basis which do not take into account the costs of personnel training, organisation and operation expenses in the systems of manned and remote control base stations deployed along the thousand lm Azov and Black Sea coastal area). These costs along with the purchase and mastering of the appropriate software products for the corporate computer networks and costs of adequate personnel training are comparable to the costs involved in the construction of networks themselves. By introducing correction co-efficient to the aforementioned additional coasts (ê=1,5–1,8), we will obtain a final extrapolated assessment of cost of the first firs stage in the GMDSS network in the south of Russia (from Taganrog to Sochi) which comes to about 150–180 billion roubles (at the exchange rate which existed at the end of 1996).

DEVELOPMENT OF THE FEDERAL GMDSS NETWORK IN THE SOUTH OF RUSSIA IN SEA AREAS À1 AND À2

V.V.Popov, A.P.Litskevich, L.N.Yusupov, V.G.Globa, V.G.Senchenko, V.V.Demianov

1. General Information

The geographic position of Novorossijsk, town and port, determines its importance for the country in the new political conditions as the "south entrance to Russia". The course to the development of the open economy and the involvement of the domestic manufacturers in the process of extending the economic relations makes a vital issue of expediting the engagement of the Novorossijsk potential in the market structure of the world economy, thus creating important pre-requisities for a fast self-financed development of a powerful production and economical complex in the south of Russia which would allow its influence in the Black Sea and Mediterranean area to be considerably enhanced. Assigning to Novorossijsk the status of a territory with a special legal and economic regime for conducting commercial activities will allow creating a favourable economic climate for the development of the Azov Sea areas adjacent to it, whilst the attraction of non-governmental investments in the industries and fields of activities having a strategic importance for many countries would dramatically accelerate the processes of reconstructing a new configuration of economically efficient seaways in the south of Russia.

Until now the significance of Novorossijsk for Russia has been limited primarily to its function of a major (though not principal) port and to the availability of resources for the competitive export of the best brands of cement. Other industries played an auxiliary role and had a secondary importance. In the Soviet times the principal emphasis was laid on the development of ports in the Ukraine. As a result of such policy not only the cost of export via Odessa was considerably higher than of that via Novorossijsk (on account of the extra kilometres), but all the facilities have in the long run been lost for Russia, as they have turned to have been constructed on the territory of a foreign country. For more than 20 years the southern ports of Russia did not receive necessary financing for the development of the supply line infrastructure which were of strategic importance for the country.

By signing after 1991 all the international documents governing the modern development of the Global Maritime Distress and Safety System (GMDSS) in seafaring, Russia assumed certain obligations to the World Community. Novorossijsk sea area has so far been the only one whose resources could be used for starting the development of an appropriate federal program for the construction of GMDSS network in the south of Russia. As a whole the scope of these works is rather large and covers the sea areas from Taganrog to Sochi; therefore, even with the favourable financing conditions, the works will most probably be carried out segment by segment.

In view of extensive research made on the Novorossijsk Port and a vanguard bulk of scientific work already carried out here in the fields of radio wave propagation, electromagnetic compatibility, problems of GMDSS network reliability, account of the way these features are affected by the mountain terrain and climatic conditions prevailing in the area, it has now become possible to form technical specifications for this network project not only in Novorossijsk area where the construction works are already underway, but also to the north of this place in the Azov sea area, and to the south: in Tuapse and Sochi sea areas.

The further construction of GMDSS network, according to the expert assessment, as well as the already available practical experience of the program implementation, have to be developed on the already elaborated theoretical and practical basis for the part of the network known as Novorossijsk Segment. Such approach allows performing structural and segmental identification of new areas with the already developed network, and forecasting the maximum number of the determining parameters of the entire created network. Ò This identification should refer primarily to the technical/economic and reliability/quality parameters and features of the entire GMDSS network in the south of Russia. The tentative assessment of the technical/economic factors was already considered in the first report.

Along with the technical and economic aspects of the GMDSS network development, another important, and often a decisive aspect of the development by analogy, is a distinct reproduction, in the new conditions, of the requirements to the principal qualitative parameter of any information communication network: ***the reliability of its functioning***. The problem of ensuring the reliability of a newly created network is aggravated by the number of the specific features of climate in the area. These specific features stipulate higher requirements, both to the functional readiness of the entire GMDSS network (timely rescue of the people in distress at sea), and to the reliability of its elements and subsystems with regard to the main parameter: frequency of unexpected failures. In the areas with adverse weather and climatic conditions the insurance against the suddenness factor is of primary importance.

A unique feature of the Novorossijsk climate is strong north-west winds called "bora". These winds gain the speed of 50 m/s and inflicts considerable damage to the city's and port's economy. The average duration of the annual period with a temperature of above 0^0Ñ is about 231 days; it should be noted that in this century the following numbers of such days were observed: the maximum - for 275 days, the minimum for 186 days a year.

The aforementioned conditions prevail, as an average, over the entire Black Sea Coast enabling the around-the-year navigation without the sea surface getting frozen, from the Kerch Strait to Sochi. The seaways on the Azov Sea which is in the winter time covered with ice for the period from one month (in the south) to four months, require the development of a light ice breaker fleet to keep them open for navigation round the year.

The features of such climatic factor as the winds should be dwelt upon in more detail, as it is the winds which in the bulk of cases cause the emergencies and accidents. It is evident from the materials of many year of observation that in Novorossijsk area the prevailing direction of the wind is north-east. An average wind speed over a month varies from about 8,0 m/s in January/February to 2,7 m/s - 6.1 m/s during the rest of the months. The largest observed wind speed (more than 35 m/s) is during the "bora", i.e. with the north-east wind. The winds of south-east direction gain the speed of up to 50 m/s. As an average there are 30-35 days a year without wind, in 27.63 per cent of all the cases south-west wind is observed, in 27.23 per cent of cases - north-east wind, in 11.57 per cent of cases - south wind and in 10.26 per cent of cases - north-west wind, in 27.23 per cent of cases - north - east wind, 11.57 per cent – south wind and 10.26 per cent - north-west. The share of north directions makes up 45.27 per cent, and the share of south directions – 43.94 per cent. The stormy winds (i.e., winds of more than 15 m/s) with the directions from sea (south-eat, south, south-west) make up 0.82 per cent, whilst the winds of 20 m/s and more – 0.29 per cent.

Prevailing in the spring time are the winds of south-east direction, their recurrence is up to 38.31 per cent of all the observed cases, the winds of north-east direction are also highly recurrent - 23.92 per cent of cases, the recurrence of the south direction winds is 12.53 per cent. In the spring time the recorded winds with a speed of 20 m/s were those of north-east. south-east and south directions. Prevailing in the summer time are also the winds of south-east direction, their recurrence is 30.03 per cent, with their recurrence decreasing to 19,90 percent and the recurrence of north-east winds growing to 30.36 per cent. In the autumn the winds with a speed of more than 15 m/s are observed from the north-east, south-east, south, south-west and west directions, whilst the winds with a speed of 20m/s and more are observed from the north-east and south directions.

To ensure a reliable GMDSS network functioning it is important to take into account the temperature and humidity of the ambient air which have a considerable effect on both, the operation of the external radio electronic system units, and on the conditions of radio waves propagation.

According to the observation data from Novorossijsk hydrometeorological station, over the last 80 years an average annual air temperature has been +12.8^{0}Ñ. The highest air temperature of +39.4^{0}Ñ was recorded in June 1957, and the lowest temperature of –24.0^{0}Ñ in January 1907. So the absolute range of the temperature oscillation is 63,4^{0}Ñ. An average monthly air temperature in all the months of the year is above zero, but in some years it may obtain negative values. In Novorossijsk area the average number of days with negative average daily temperatures is up to 20 a winter. It should be noted that the changes of temperature over the day and night are not so distinct as in the areas remote from the sea. The first autumn frosts occur in Novorossijsk in the middle of November (as an average) the late frosts stop at the end of March.

An average annual value of the absolute air humidity is 8.6 mb. The changes of the absolute air humidity over a year are associated with the changes of air temperature. The maximum value of the average monthly absolute air humidity was recorded in August (16.9 mb), whilst the minimum value was recorded in February (2.5 mb). In the transition periods the average monthly absolute air humidity varies from 11.6 mb to 3.5 mb. The average annual relative humidity is 72 per cent. Its changes over the year are reciprocal to the air temperature. In the winter and spring time (from December to May) the average monthly relative humidity reaches 81-82 in per cent, in the summer time it drops down to 75-59 per cent.

The fogs in the Novorossijsk Harbour are caused by the frequent north-east winds, proximity of sea and availability of cement works (which generate the aerosol environment which facilitates the fog formation). The fogs are caused by the advance of cold air to the warm sea surface. According to the Novorossijsk Hydrometeorological station there are about 36 days with fog in a year. The recurrence of fogs increases in spring (in April-May). At this time they are stable. In these months the number of days with fog grows to 5.

A specific feature of the area is possible thunder storms in any winter month. In the area of Tzemess Harbour, as an average 24 days with thunder storm are noted in a year. The maximum number of thunder storms falls on June-August. The largest number of days with thunder storm (43 days) was recorded in 1960. In some cases thunder storms are accompanied with hail. Hail falls in some places in comparatively small areas. Most often the hail falls in winter.

The ice crust is usually forming on the ground in the transition and winter months and is characteristic for the mountain areas. This phenomenon is normally connected with the change of weather, with the temperature transiting over 0^{0}Ñ in the lower direction. In most cases the ice crust is formed when the sky is cloudy, the relative humidity if over 90per cent and a temperature is from –0.1^{0}Ñ to –5^{0}Ñ. The ice crust is mostly noted when the wind speed is under 10 m/s, but it can often form with the wind speed of above 10 m/s. The average duration of the stay of ice crust on the ground is 30 hours (note that in December 1959 the duration of ice crust remaining on the ground reached 160 hours). In view of the climatic features of the area listed above and a high probability of emergencies and disasters connected with them, we should consciously formulate the fundamental requirement to the enhanced reliability of the designed GMDSS network.

The problem of insuring the reliability of networks as a whole is fundamental and covers all the stages of its development and construction. The GMDSS concept defining its functional purpose unambiguously formulates the following features which should be inherent in the national networks and some of its segments:

– insure steady reception (with 10^{-4} -10^{-6} probability of error) of all the distress and safety messages from areas À1 and À2;

– insure fast and reliable alerting (by re-transmitting the received message) of all the vessels in the vicinity within areas À1 and À2;

– provide effective help in the rescue of people in distress using the facilities of the coastal Rescue Coordinating Centre in the MAPN's GMDSS responsibility area.

From the viewpoint of these problems, the analysis of the developed system for insuring the reliability cannot but be of comprehensive character, for the reliability itself as the principal feature/quality of the system normally has an integral character. The regional maritime communication system as a rule includes a large number of complicated and diverse elements combined in a single information system, and before formulating the reliability's quality criteria, it is necessary to clarify the specifics and particulars of the created object, and corresponding characteristics of its elements.

The principal feature of the system which is being developed is it being structurised. Three subsystems can be singled out in its structure. The ***first*** subsystem is formed by the operation personnel which has a considerable effect (up to 10-20 per cent) on the quality of the entire communication system's functioning. The optimum could be found in selecting the ratio of automated (unmanned) coastal stations of the system to the number of manned stations. The ***second*** subsystem is a technical complex which includes the radio equipment, antenna-feeder lines and power supply systems. Failures in each of the aforementioned systems may cause the failure of the entire system as a whole. The ***third*** subsystem is the tropospheric segment of the radio communications where the radio waves propagate between the transmitting and receiving antenna.

The condition of tropospheric medium has a significant effect on the reliability of communications, causing the breaks in the message transmission due to the fading of the signal level on the noise and clutter background, sometimes interrupting the operation of the system as a whole. Reasoning from the communication system structure specified above, the analysis of its functioning reliability should take into account the dynamics of systems interaction within a single "Human being - equipment - propagation medium" system. In order to solve a specific problem: to insure the reliable maritime regional communication system within MAPN's responsibility area, it is necessary (with regard to the aforementioned specific features) to identify the priorities of directions and concentrate the efforts on the solution of problems precisely in the determined directions. The analysis of experience accumulated over many years of operating the communication systems shows, that it is necessary to pay attention to the following: in the 10 years (in the 80-ies) the percentage of idle time caused by fading made up 29 per cent in the overall share of idle time at the stations (71 per cent. It should be noted that the idle time of equipment available at the stations, by groups, could be divided as follows:

– idle time due to the failures of power supply - 35per cent ;

– idle time of communication equipment – 26 per cent ;

– idle time caused by the personnel's inexpert work – 15 per cent .

The statistics quoted above allow identifying the particular directions which the efforts should be concentrated on in insuring the system's reliability. Specifically, in view of the entire communication system's reliability depending directly on the power supply units, it is necessary to concentrate efforts on enhancing the latter's reliability. In this case it would be efficient to increase the number of feeder power lines from the independent substations of the state power supply network, as well as to install autonomous uninterrupted power supply sources based on the diesel generators with multiple redundancy.

The second in importance factor of insuring the reliability of the communication system is preventing the radio wave fadings in the "vessel-shore" segments of areas À1 and À2. An efficient measure would be a diversity reception on the "shore" antenna units and discarding the use of closed channels in all the segments of radio wave propagation. The use of diversity reception (space or frequency) provided the diversity value is selected correctly, ensures the static independence (orthogonality) of idle of inactivity of two elements connected in parallel. There is quite a considerable dependence between the reliability of the entire communication network and the reliability of the radio equipment installations (up to 26 per cent); it means that the enhanced reliability should be looked for on the ways of using up-to-date communication equipment based on the integral units with a high degree of integration and with the modern functioning principles (optimum methods of reception, digital processing of data, structural redundancy, broad system adaptability, etc.).

When developing regional GMDSS systems it is also necessary to use a considerable reserve (15 per cent) for enhancing the reliability which lies in the efficient organisation of the network servicing. The data set forth above shows that 15 per cent of idle time is caused by the inexpert actions of the service personnel: violation of the Rules of Technical Operation of Communication facilities, inability to foresee and identify the causes of failures, errors and omissions in the system operation. To ensure the reliability in this direction it is necessary to provide the equipment with the fast automatic identification and recording of failures. In this connection it would be worthwhile, in the development of the communication system, to design into it such a system of the shore based unmanned centre diagnostics which would include, as its integral part, the equipment for the automatic identification of failures and recording of idle time.

Introduction of these measures will enable the average time for the equipment repairs to be reduced, whilst the timely information on the causes and place of failures will allow taking efficient measures to prevent them; this would reduce the overall number if failures and increase the equipment service time, i.e. would enhance the system's reliability. Another reserve for enhancing the reliability is the development and implementation of maintenance schedules, as well as scientifically justified calculations of the backup equipment and instruments redundancy, and certainly the time delivery of spare parts to the places where the backup equipment is exhausted.

The measures to enhance the reliability of operation listed above requires due attention to be paid to the initial and advanced training of the specialists capable of expertly operating the complicated communication systems, forecasting their technical condition, perform the failure diagnostics and prevent them, identify and eliminate in time.

Formulated above are the purport directions in insuring the reliability of the developed system. We will now provide a quantitative criterial factor of reliability in the operation of the communication system which is being designed. It normally includes the following reliability criteria: P(t) – probability of trouble free operation over t time (communication session); $\hat{E}_{\tilde{a}}$ – readiness coefficient; $\hat{E}$ – idle time coefficient ($\hat{E}=1-\hat{E}_{\tilde{a}}$); Ò – average time of trouble free operation.

On the basis of analysis of the requirements to the communication systems' reliability and their elements, and the identified priorities in the actuation of factors of the unreliability in their operation, the requirements to the system which is being designed can be formulated. In view of the specific features of the developed system (which insures the safety of a human being at sea), it is required to both, have a high level readiness coefficient, and insure a high probability of fault free operation during each communication session. In this case both of these factors should be independent in all the modes of the GMDSS network operation. We can thus finally present the criterion of the reliability of the entire system's operation in the following logic and operation form:

$$V_t \leq t_{max} \{(K_{ã} \geq \alpha) \Lambda (P(t) \geq \varepsilon)\} = 1,$$

whose meaning is reduced to the following: the obtained readiness coefficient of the designed system $\hat{E}_{ã}$ should not be less than α set in the technical specification with the concurrently and independently implemented probability of fault free operation Ð(t) of the entire system during longest communication sessions t_{max}, determined in the technical specification and should never be lower than the set level ε in accordance with the technical specification.

CHAPTER II

TELECOMMUNICATIONS AND COMPUTER NETWORKS

THIS CHAPTER INCLUDES PAPERS

PRESENTED AT THE CONFERENCE SESSION:

TELECOMMUNICATIONS AND COMPUTER NETWORKS

Organized by: ***Acad. Nikolay A. Kuznetsov,***

Dr. Richard M. Soley,

Prof. Alexander A. Hagin.

QUALITY ESTIMATION OF PACKET-SWITCHED VOICE TRANSMISSION IN INTEGRATED SERVICES DIGITAL NETWORK

Dr. Ahmad Ali Al-Noman
Yarmouk University
Communication Department, Hijjawi College For Applied Engineering
Jordan - Irbid
tel. 271100 telex 51533

Abstract. Packet-switched voice transmission belongs to a category of traffic that has strict requirements on time delays during the transmission of messages in real time. If the admissible time-delay limit is exceeded, then packets can be lost and the voice transmission quality deteriorates. This paper describes a method for estimating the quality of packet-switched voice transmission using the speech intelligibility criterion, this is based on calculating the energy spectrum of voice signal, taking into account the impact of packet rejection due to various reasons. Among them time delay, quantization noise, and digitization noise. In general, the described method allows estimating the quality of transmission in packet-switched communication networks.

Keywords: Packet switching, Voice transmission, Quality characteristics, Quality estimation, Signal-to-Noise ratio, Intelligibility.

Abbreviations

CCITT	International Telephone and Telegraph Consultative Committee.
I	Phrase Intelligibility.
ISDN	Intergrated Services Digital Network.
LPF	Low Pass Filter.
SDNR	Signal -To -Digitization Noise Power Ratio.
SNR	Signal -To- Noise Power Ratio.
SQNR	Signal-To-Quantization Noise Power Ratio.
S	Syllable Intelligibility.
TDMA	Time Division Multiple Access.
W	Word Intelligibility.

Introduction. Packet-switched voice transmission imposes restrictions on time delay during the transmission of packets between two terminal units. This is because the two-way conversation must not be interrupted with long intervals (pauses). In addition, time fluctuations in packet-switched transmission causes phase vibrations in packet-reception time which increases delay over the network, so it is necessary to enlarge elastic memory capacity.

In network operation, situations will inevitably occur where some of the packets are delayed so long, that they would reach the terminal unit at a time, when the time assigned for decoding of the packets has expired, and are consequently rejected. This results in breaks over the active-voice portions reconstructed by the terminal unit and deteriorated intelligibility.

These situations are obviously due to network overloading because delays in packets transmission increase rapidly, even with a slight excess in actual load intensity compared to the estimated one. Therefore, packet-switched voice transmission is much more sensitive to time delays than packet-switched data transmission [1].

For this reason, in designing packet switching networks, one must take into account the changes in speech intelligibility characteristics resulting from network operating conditions, particularly, the changes in load intensity.

The discussion here concerns with the method for estimating the quality of packet-switched voice transmission in digital networks.

Models For Calculating The Quality Of Packet-Switched Voice Transmission. To solve the above problem, it is assumed that the initial delay of packets in buffer memory units of the destination point equipment is selected, for a normaly operated network, to avoid distortions in the course of voice reconstruction.

All the troubles occur when, by the time of decoding a regular packet content, the packet is missing due to excess delay in the course of its transmission through the network, subsequently, X packets may be rejected in turn, where X = 1, 2, 3 ...... .This results in the loss of mX = m, 2m, ... samples and the need to regenerate the voice signal over the missing time intervals:

$$\frac{mX+1}{f} = \frac{m+1}{f}, \frac{2m+1}{f}, \ldots, \ldots \qquad (1)$$

The regeneration is accomplished using the terminal unit interpolating filter. In Eq. (1) m is the number of samples in a packet (the packet length), and f is the speech signal sampling frequency.

We shall use the approximation of energy spectrum [2], assuming that the voice signal spectrum is restricted by an ideal low pass filter (LPF) with cutoff frequency f_c:

$$G(\omega) = R(1+M)\sigma^2\left(\frac{1}{\omega^2 - 2B\omega + R^2 + B^2} + \frac{1}{\omega^2 + 2B\omega + R^2 + B^2}\right) \qquad (2)$$

Where:

$R = \frac{4\Delta f}{1+g^2}$;

$B = gR$;

$M = 1 - \frac{1}{\pi}\text{arctg}\left(\frac{\Omega - B}{R}\right) - \frac{1}{\pi}\text{arctg}\left(\frac{\Omega + B}{R}\right)$;

$\Omega = 2\pi f_c$;

g is a factor;

Δf is the effective spectrum width [3];

σ^2 is the signal variance.

For this signal model, it is not difficult to obtain a relationship between signal power and the power of the noise caused by digitization and reconstruction processes (SDNR). Taking into account the results of packet rejection the SDNR could be given by:

$$\mathrm{SDNR} = 10\lg\left\{\sum_{X} p(X)\frac{1+M}{2\pi}\left\{\sum_{l=1}^{Q}\left|\mathrm{arctg}\left(\frac{\Omega' - C}{R}\right) + \right.\right.\right.$$

$$+\mathrm{arctg}\left(\frac{\Omega' + D}{R}\right) - \mathrm{arctg}\left(\frac{2\pi l\cdot f/(mX+1) - \Omega - C}{R}\right) -$$

$$\left.\left.\left. -\mathrm{arctg}\left(\frac{2\pi l\cdot f/(mX+1) - \Omega + D}{R}\right)\right]\right\}\right\}^{-1} \qquad (3)$$

where:

p(X) is the probability distribution of random variable X ;

$\Omega' = 2pF'_c$;

F'_c is the cutoff frequency of an ideal LPF used for reconstructing;

$C = 2\pi f + B$;

$D = 2\pi F - B$;

$Q = \frac{(\Omega' + \Omega)(mX+1)}{2\pi f}$

In estimating the noise caused by voice amplitude quantization, the time-average approximation of uniform probability distribution density is used, which is given by:

$$W(\alpha) = \left(\frac{K}{\sigma}\right)\left\{\exp\left[-(\pi-1)\left(\frac{K|\alpha|}{\sigma}\right)\right] + \exp\left[-30\left(\frac{K|\alpha|}{\sigma}\right)\right]\right\} \qquad (4)$$

Where:

K = 0,64, is the normalization factor. The signal -to -quantization noise power ratio (SQNR), using the A-law companding characteristics [4] for approximation given in Eq.(4), is determined in the form of

$$\mathrm{SQNR} = 10\lg\left\{\frac{\Delta^2(1+\ln A)^2}{6\sigma^2}\left\{\frac{1}{A}\left|\frac{1}{\pi-1}\right|1 - \exp\left(-\frac{(\pi-1)K}{\sigma A}\right) + \right.\right.$$

$$+\frac{1}{30}\left[1 - \exp\left(-\frac{30K}{\sigma A}\right)\right]\Bigg| - \exp\left[-\frac{(\pi-1)K}{\sigma}\right]$$

$$\left[\frac{1}{\pi-1} + \frac{2\sigma}{(\pi-1)^2K} + \frac{2\sigma^2}{(\pi-1)^3K^2}\right] - \exp\left(-\frac{30K}{\sigma}\right)$$

$$\left[\frac{1}{30} + \frac{2\sigma}{30^2K} + \frac{2\sigma^2}{30^3K^2}\right] + \exp\left[-\frac{(\pi-1)K}{\sigma A}\right]$$

$$\left[\frac{1}{(\pi-1)A^2} + \frac{2\sigma}{(\pi-1)^2KA} + \frac{2\sigma^2}{(\pi-1)^3K^2}\right]$$

$$\left.\left. + \exp\left(-\frac{30K}{\sigma A}\right)\left[\frac{1}{30A^2} + \frac{2\sigma}{30^2KA} + \frac{2\sigma^2}{30^3K^2}\right]\right\}\right\}^{-1} \qquad (5)$$

Where:

Δ is the average quantization step size.

A is the parameter used to define the amount of compression when using the A-Law companding characteristics, and is defined by CCITT.

Examples for Estimating The Quality of Packet-Switched Voice Transmission. The results from the analysis performed by the Administration of Communication of Norway [5, 6] are used. Extrapolation of these results are presented in Fig.1. in this figure SNR denotes the equivalent signal-to-noise power ratio. The dependence of syllable intelligibility (S) on signal-to-noise power ratio (SNR) is shown in Fig.2 [5]. Here, SNR = ΣB_v - ΣB_n, where ΣB_v is the total voice level and ΣB_n is the total noise level, both of them expressed in dB. In this graph we assume that $\Sigma B_n = 60$ dB. Figure 3 gives the relationship between syllable (S) and word (W), syllable (S) and phrase (I) intelligibilities.

Example 1: Let us estimate the quality of a packet-switched voice transmission in a cellular digital communication system with a narrow-band TDMA (TDMA being time-division multiple access). This estimation has been performed for the following parameters: m = 64 samples, f = 8 KHz. The packets are transmitted, along with additional data, in one of the n = 10 segments arranged in succession, each lasting for T = 0,8 ms [7].

The first approximation of packet rejection probability can be determined using the following formula [8]:

$$\theta = 1 - \left(1 - \frac{1}{n}\right)^{qN-1} \tag{6}$$

Where: q is the probability of mobile stations operated in transmission mode.

N is the number of mobile stations in a cell.

As an example, for n = 10, q = 0,024, N = 42, we obtain $\theta = 8 \cdot 10^{-4}$.

If the attempts to realize algorithms for access to the transmitting medium in adjacent cycles are presumed to be independent, then the geometric probability distribution p(X) can be used for packet non-rejection (for the first time after X trials) [7]

$$p(X) = \theta^{x}(1 - \theta), \quad X = 0, 1, 2 ..., \quad 0 < \theta < 1 \tag{7}$$

The results obtained from calculating the transmission quality characteristics using Eqs. (2) - (5), (7) and the graphs 1 depicted in Fig.1 for the conditions: f = 8,000 Hz, $\Delta = 1/128$, A = 100, P = 0 dB ($\sigma = 0,33333$), here P denotes the voice signal power, $F_c = 3,400$ Hz, $F'_c = 3,700$ Hz, $\Delta f = 1,800$ Hz, g = 2,5, and m = 64 samples, are shown in the table below:

Table: The results of calculating the packet-switched voice transmission quality characteristics

	Voice Transmission Quality Characteristics			
θ	SDNR, dB	SQNR, dB	SNR, dB	S, %
0.01	5.14	39.27	5	48
0.001	15.18	39.27	15	70
0.0001	25.18	39.27	25	82
0.00001	35.18	39.27	35	93

. Data in the table demonstrates the relationship between syllable intelligibility (as the basic quality characteristic in voice transmission) and the probability of message-packet rejection. Superior quality can be obtained with fairly liberal requirements on packet rejection probability.

Example 2: We shall demonstrate the impact of packet length on quality characteristics of voice transmission, all other factors being equal. We assume that the processes of voice transmission in the network are independent. This assumption will allow the use of formula (7). The results of calculations obtained from Eqs. (2) - (5), (7) and the graphs depicted in Fig.1 for the conditions: f = 8,000 Hz, Δ = 1/128, A = 100, P = 0 dB (σ = 0,33333), F_c = 3,400 Hz, F'_c = 3,700 Hz, Δf = 1,800 Hz, g = 2,5, are shown in Fig.4.

The horizontal dashed line shows the recommended standard used in Rec. G.712, adopted by CCITT (for the primary digital transmission system with digitization and reconstruction noise negligible). The curves, similar to those in Fig.4, can be used to determine the packet length. Using the graphs shown in Fig.2 and Fig.3, we can find the dependence of speech intelligibility characteristics on message-packet length (m).

Example 3: We shall estimate the effect of voice signal power on voice transmission quality characteristics. Let us assume that the processes of voice transmission in the network are independent. This will allow the use of formula (7). The results of calculations obtained from Eqs. (2) - (5) and the graphs (Fig. 1) for the following conclitions: f = 8,000 Hz, Δ = 1/128, A = 100, F_c = 3,400 Hz, F'_c = 3,700 Hz, Δf = 1,800 Hz, and g = 2,5, are presented in Fig.5. The solid and dashed lines correspond to the cases of P = 0 dB (SQNR = 39.27 dB) and P = -50 dB (SQNR = 17,48 dB), respectively. As shown in Fig.5, the impact of voice signal power on quality characteristics increases with the decrease in the probability of packet rejection.

Example 4: Let us estimate the quality of voice transmission for the system [9], operated in a packet-switched synchronous mode, using the virtual voice- and- data channel, with priority to voice packets servicing, The packet length is 80 samples (10 msec). In [8] it is shown that, with the voice signal factor ρ = 0.56 and the number of originating voice channels N = 15, the probability of voice-packet rejection will be 0.0001. From the graphs in Fig.4 we obtain SNR = 24 dB, and from Fig.2 S = 80%.

Conclusions. In evaluating digitization techniques for speech, we need a measure of success. The signal-to-quantization noise ratio, a measure that has been discussed earlier, is related to the meansquare quantization error. However, in speech transmission, subjective measures often take precedence over these mathematical measures. For example, word or sentence intelligibilities and subjective qualities of the speech become important.

The methods presented in this paper permit estimating the quality of packet-switched voice transmission in terms of syllable intelligibility, depending on the influence of polytypic time-probablity characteristics of digital communication networks. These methods can be used for perfonnance analysis in both existing networks and those under development. Potentials of the developed mathematical models can be applied to other problems outside the scope of the present paper.

References

1. Bellamy, G., Digital Telephony, Moscow, Radio i Svyaz, 198G, 544 p.
2. Velichkin, A. I., Theory of Digital Transmission of Continuous Messages, Moscow, Sov. Radio, 1970, 296 p.
3. Gorelov G. V., Non-Regular Digitization of Signals, Moscow, Radio i Svyaz, 1982, 256 p.
4. Leon W. Couch, Digital and Analog Conununication Systems, New York, Macmillan Publishing company, 1993, 827 p.
5. Vemian, G. V., Voice transmission Thnrugh Coirtmunication Networks, Moscow, Radio i Svyaz, 198S, 272 p.
6. "CCITT Contribution Com. XII-N99" CCITT, Geneva, Study Period 1976-1980.
7. Lindell, F., Swerup, J., and Uddenfeldt, J., "Digital Cellular Radio for the Future", Ericsson Review, No Band 3., 1987, pp. 48-56.
8. Gorelov, G. V., "The burst Aloha Algorithm in Statistic Multiplexing of Audio Information Signals", Trans. of Intern. Conf. On Local Networks, Riga, 1990 pp. 38-41.
9. Choi, J. K., Seo, J. U., and Un Ch. K., "Performance Analysis of a Packet-Switched Synchronous Voice, Data Transmission System", IEEE Trans. On Comm., Vol. 38, No 9, September 1990.

LAND MOBILE CELLULAR NETWORK DESIGN

E. P. Vishnevsky, A. S. Anikonov, C. U. Ushakov
St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Science (SPIIRAS)
V. G. Maslov
S. I. Vavilov State Optical Institute
e-mail: lai@mail.iias.spb.su

Abstract. The alternative optimal cellular network design adaptive concept has been formulated which not subjected to the disadvantages of conventional concepts. The essence of the concept constitutes the consecutive execution of the spatial and frequency planning in complex urban environment using terrain digital model, and its conceptual flow is rigorously determined.

Keywords. Terrain digital model, area of interest, grade of service, cellular network design, mobile telephone switching once, Rician channel, channel assignment, frequency assignment, grade of service.

Integrated networking is the unifying theme of telecommunications systems evolution in the final quarter of the twentieth century. The purpose of an integrated network is to collect information from diverse sources and deliver it to designed destinations. In networks with wide area coverage, advances in technology are coordinated under the heading integrated services digital network(ISDN). ISDN provides a structured set of protocols for moving and processing information in public networks. Standards are now in place for voice communications and for the exchange of data at low and moderate speeds. Thanks to recent work on broadband ISDN [1], a network architecture for communication of information from high-rate sources is emerging and keeping pace with advances in the technologies of transmission, switching, and signal processing.

Wireless communication networks use ultrahigh frequencies to link terminals with the fixed infrastructure of base stations and switches. The propagation of radio signals between a wireless terminals and a base station is strongly influenced by

- the terminal-to-base distance;
- terrain, architectural features, and other environmental details;
- interference from other transmissions in the same system; and
- noise of many varieties.

All of these characteristics are unpredictable at the beginning of a communication session and are subject to substantial and rapid changes during the progress of the communication. Thus, in contrast to other communication networks, wireless networks are subject to continual and rapid changes in configuration. Call set-up and handoff (the process of switching the communication from a channel served by one cell to another channel served by an adjacent cell) are currently performed by a mobile telephone switching once (MTSO) that controls all the calls in a large geographical area [5].

The cellular telecommunications service offers high teletraffic capacity, continuous connection over a large area of interest (AOI), and a high Grade of Service (GOS). The key factors that constrain the optimal cellular network design (CND) in complex urban environment are:

1. communication channels linking base station and terminals are hardly predictable and highly variable with time;

2. a wireless network is highly dynamic, its configuration changes many times each second as wireless terminals change locations;

3. spectrum reuse is essential to high throughput because the radio bandwidth for wireless networking is strictly limited;

4. the propagation of radio signals between a wireless terminal and a base station is strongly influenced by: a) the terminal-to-base distance; b) terrain, architectural features and other environmental details; c) interference from other transmissions in the same system; d) noise of many varieties.

5. a call could be handed off to the neighboring cell to provide uninterrupted service to the mobile subscriber crossing a cell boundary and moving into adjacent cell while the call is in progress.

All of these characteristics are hardly predictable at the beginning of a communication session and are subject to substantial and rapid changes during the progress of the communication. Together they create a challenging transmission environment.

At present the CND problem is solved using different statistical models of electromagnetic wave (EMW) propagation, Rician channel model and simple geometrical configurations of cellular grid (triangular, rectangular, hexagonal). The main disadvantage of conventional CND concept is an imperfect accounting of the above factors which leads to overestimating of the required power, frequency and technical resources.

A new optimal CND adaptive concept is proposed to overcome these disadvantages and incorporates the following main features:

1. It takes into account the complex topological structure of urban environment.

2. It takes into account the fine structure of EM field in complex urban environment on the basis of exact consideration of main physical mechanisms of EMW propagation (diffraction, multiple reflection and diffuse scattering).

3. It relies on realistic model of dynamic multipath channel predictable in every point of urban environment for arbitrary base station (BS) site.

4. It provides means for some intelligent network reactions to the environment changes (traffic dynamics) within the frequency domain.

5. It guarantees the prespecified levels of grade of service (GOS) standards (power, information, handoff, interference), as well as it provides maximum of spectrum efficiency and minimum of required number of base stations by optimal manner.

Basic Data.

1. exact Terrain Digital Model (TDM) of the considered urban environment.

2. spatial variation of noise distribution over the Area of Interest (AOI);

3. spatial-time variations in traffic; the traffic over the area of interest may have an arbitrary distribution; it is assumed to be quantized into L levels so that the whole (AOI) may further be subdivided into a number of subareas with approximately constant traffic.

Generalized Spectrum Efficiency. In all new cellular systems to be implemented, a fundamental role in network design is played by efficient use of the spectrum and, consequently, many researches have been developed so far in the field of digital modulation systems in order to obtain "good" systems from a spectral point of view. It must be immediately emphased that the actual spectral efficiency (SE) of a cellular system should be defined on the basis of its behavior in the presence of adjacent and cochannel interferences in the cellular mobile radio topology.

The main reasons for extending the definition of spectrum efficiency to complex urban environment are:

1. propagation conditions are usually very irregular and therefore interference levels do not depend on distance ratio only;

2. cell shapes are very irregular and strongly depend on base station(BS) site;

3. spatial variations in traffic lead to a channel demand varying from cell to cell;

4. environmental constraints often impose limits on the usability of certain frequencies.

It should be clear that if any of these phenomena were present in a mobile radio network, the conventional CND concepts could not be applied any more. Since most real-world problems suffer from all of them, the need for a more general approach to CND and its quantitative characteristic, generalized Spectrum Efficiency, is obvious.

So, consider a given area of interest (AOI) divided into a certain number of cells sharing the same frequency resource. Let

N be a number of cells in AOI;

S_i (km^2) be an area of an i-th cell;

$A_{c(i)}$ (erlang) be a carried traffic in an i-th cell;

$n_{c(i)}$ be a number of channels assigned to an i-th cell;

W (MHz) be a bandwidth per channel.

The spectrum efficiency E_s of a cellular system is usually represented by the product of spectrum efficiencies against time, frequency and space. Conserving the physical meaning of its components, the conventional expression of spectrum efficiency may be extended to a cellular system in complex urban environment in the following manner:

$$E_s = \frac{1}{N^2}\sum_{i=1}^{N}\frac{A_{c(i)}}{S_i\, n_{c(i)} W} \quad \left[\frac{erlang}{MHz \cdot km^2}\right] \tag{1}$$

$n_{c(i)}W$ being the frequency resource used in an i-th cell. In the simplest case of an uniform cellular grid with uniform traffic the cells of AOI are grouped in a certain number of equal clusters sharing the same frequency resource. Thus, in this case, the spectrum efficiency of the cellular system is given by E_s of a cluster derived from (1) with $N = C$:

$$E_s = \frac{A_c}{C\,W\,S\,n_c} \tag{2}$$

C being a number of cells in a cluster.

It is the conventional formula for E_S. The extended spectrum efficiency E_S given by (1), and the GOS standards are used in our further consideration as optimal criteria of CND.

Main Features of the Concept. The essence of any CND concept constitutes the spatial and frequency planning. Figure 1 gives an overall outline of optimal CND general problem adopted in our concept.

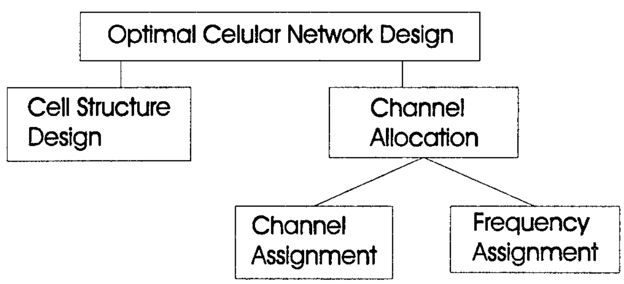


Fig. 1 The CND general problem

Spatial planning is considered as a part of cell engineering process dealing with BS allocation within the coverage area. This problem (Cell Structure Design) is solved only once (at the beginning of CND process) and provides the invariable (static) part of CND.

Frequency planning deals with Channel Allocation problem that could be split into two different levels of planning:

1. Channel Assignment (CA) problem – finding a number of required channels for each cell

2. Frequency Assignment (FA) problem – finding a compatible assignment of carrier frequencies to radio cells. The frequency planning process is performed in adaptive manner and provides variable (dynamic) part of CND.

Algorithms of Optimal Cell Structure Design. To solve the BS allocation problem, the heuristic algorithms are proposed which use the basic data described above. We now turn to the formal specification of the optimum BS allocation algorithms in the sense of maximizing the spatial fractions of the extended spectrum efficiency, under power and information GOS constraints. The algorithms are based on the following assumptions:

1. exact TDM of high resolution for a given AOI is known;
2. noise distribution over AOI is known;
3. initial traffic data for AOI are given;
4. BSs and mobile terminals are omnidirectional;
5. BSs are allocated at some height h_B (model parameter) above the roof of a building;
6. simulation process "Cell" is used to find a cell for a given BS site.
7. simulation process "Traf" is used to calculate the carried traffic for a given cell.

Discussion. The problem of BS allocation is solved by the elaborated heuristic algorithms in optimal manner to maximize the spatial fraction A / S of the extended spectrum efficiency for each cell and provide the power-information GOS standards. But there are principle difference between these algorithms to be accounted in its further application.

First, the problem of AOI coverage is solved in a superfluous manner, from the beginning, in the "Cell Extraction" algorithm while AOI coverage is performed consecutively in the "Top-to-Bottom" algorithm.

Second, the AOI coverage process in "Top-to-Bottom" algorithm is controlled by the model parameter ε_{ov} (threshold of cell-overlapping percentage) related to the selection of intersecting cells in accordance with its values of A / S. In contrast, the percentage of cell-overlapping is the intrinsic characteristic of the completed cell-extraction process which cannot be arbitrary predetermined.

The last circumstance should be accounted in the further applications of these base station allocation (BSA) algorithms when the possibility to vary cell overlapping may be important in Frequency Planning of CND, e.g. in the interference problem solving.

Simulation Results. For any BSA algorithm cell determination for a given BS site is required. It is an elementary algorithmic step of fundamental interest subjected to the influence of many factors of different kind (see Figure 2). So, first of all, consider the cell formation in detail.

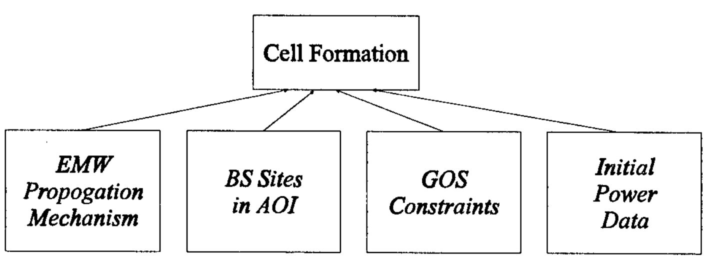


Fig. 2 Basic factors of cell formation

1. For a given BS site and for any physical mechanisms of EMW propagation accounted, the proper procedure of cell formation is subjected to prespecified levels of power-information GOS standards, i.e. power threshold of received signal $P_{r,th}$, signal-to-noise ratio snr_{th}, and bit-error rate P_e. For each GOS factor we obtain the appropriate domain in AOI. Then, the required cell is formed by simple intersection of the above domains in which the GOS constraints are satisfied simultaneously.

In order to give a synthetic view of the obtainable cellular structure design (CSD) performance, let us fix system parameters as is usual in mobile radio systems (see Table 1).

Table 1 The average values of model parameters in BSA simulation process

HBS	ε_{ov}	$P_{r,th}$	snr_{th}	$P_{e,th}$	P_n	$P_{0,BS}$	$P_{0,MT}$
15 m	15%	$3\cdot10^{-5}$ W/m^2	10 dB	10^{-3}	$3\cdot10^{-6}$ W/m^2	40 W	10 W

The main task of present research is optimal allocation of base stations in a given AOI in accordance with the accepted Adaptive Concept of CND. For this, the BSA algorithms have been elaborated, and their efficiency should be studied. So, let us define the numerical characteristics to be used for their comparison. They are:

a) N_{BS} - the total number of BSs required for complete coverage of AOI;

b) $\overline{\left(\frac{A}{S}\right)}$ - the average meaning of spatial fraction of the extended spectrum efficiency over the set of cells obtained by BSA algorithm, i.e.

$$\overline{\left(\frac{A}{S}\right)} = \frac{1}{N_{BS}} \sum_{i=1}^{N_{BS}} \frac{A_i}{S_i} \qquad (3)$$

c) T (sec) - the required operation time.

The value $\overline{\left(\frac{A}{S}\right)}$ is the principle characteristic of a BSA algorithm as follows from the Adaptive Concept. The value N_{bs} characterizes the required technical resource (cost-factor) for complete AOI coverage by a BSA algorithm, and may be confidence as an implicit optimal criteria of BSA algorithm. So, further a BSA algorithm will be characterized by the «fair» of number (T, NB_S, $\overline{\left(\frac{A}{S}\right)}$).

Turn now to basic factors determining the BSA algorithms. The two first values H_{BS} and ε_{ov} are the model parameters of the developed BSA algorithms. H_{BS} is the height of a BS above the building's roof, and ε_{ov} is the threshold of cell overlapping percentage (only for «top-to-bottom» algorithm). These model parameters of great importance permit to control the BSA process in order to attain the optimal solution for CND.

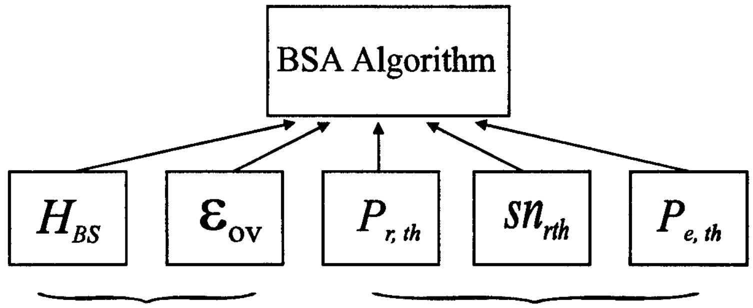


Fig. 3 Basic factors of BSA algorithms

At the other hand, the last three factors at Figure 3 are related to GOS constraints prespecified at the beginning of BSA process. These quantities characterize the required quality of wireless communication and reveal themselves implicitly in cell formation process. Precisely, they influence on the shape and area of cells arising in BSA process, and, finally, on characteristics $(T, N_{BS}, \overline{\left(\frac{A}{S}\right)})$.

Conclusions.

1. The alternative Optimal CND Adaptive Concept has been formulated which not subjected to the disadvantages of conventional CND concepts. The essence of the concept constitutes the consecutive execution of the spatial and frequency planning in complex urban environment using TDM, and its conceptual flow is rigorously determined.

The Spatial Planning is based on the principal assumption that BS site completely determines the cell shape in complex urban environment. It is the main difference with the conventional BS allocation approach. At this point, the initial distribution of traffic carried over covered area is assumed to be known from any exact traffic model or any traffic data for a given AOI.

The BS allocation is performed in optimal manner to maximize the spatial part of the extended spectrum efficiency, to minimize the number of required base stations, and to provide the prescribed levels of power-information GOS standards. The spatial planning is performed only once at the beginning of CND and results in fixed BS allocation and corresponding cell pattern of a given AOI using TDM.

The Frequency Planning is reduced to Channel Assignment and Frequency Assignment problems which provide variable (dynamic) parts of CND adaptive to spatial-time variations of traffic data. The optimal Channel Assignment is performed to maximize the extended spectrum efficiency and to meet some GOS standards. The Frequency Assignment is the final step of the conceptual flow and is performed in optimal manner to meet the prescribed levels of GOS standards for interference. The principle feature of Channel Allocation in our concept is the absence of clusters usually appearing in conventional concepts due to regular cell geometry used.

2. The exact formula for extended spectrum efficiency has been derived and was taken as an optimal criteria in the CND Adaptive Concept. The formula makes it possible to estimate

quantitatively the efficiency of reusing the full frequency resource in a dynamic urban environment.

3. "Top-to-Bottom" (TB) and «Cell Extraction» (CE) algorithms of BS realizing the first part (spatial planning) of Adaptive Concept have been elaborated. These heuristic algorithms result in a fixed BS network giving a set of BS sites and the related cells covering AOI and satisfying the above constraints. The problem of BS allocation is solved by the above heuristic algorithms optimally to maximize the spatial fraction A / S of the extended spectrum efficiency for each cell and meet the power-information GOS standards.

4. In the result computer aided prediction system (CAPS) prototype realizing two proposed algorithms of BS allocation (TB and CE) has been created.

5. The task-effectiveness and time-efficiency of TB and CE algorithms of BSA have been studied. For this purpose, at first, the fundamental process of cell formation has been thoroughly considered. The simulation results showed the decisive role of BS site and GOS constraints on the shape and area of a cell. The detailed analysis of simulation results makes it possible to compare the efficiency of TB and CE algorithms for a wide set of basic factors.

References

1. Armbuster H., Arndt G. (1987) Broadband communication and its realization with broadband ISDN, IEEE Commun. Mag., vol. 25, pp. 8-19.

2. Mac-Donald V. H. (1979) The cellular concept, Bell Syst. Tech. J., vol. 58, no. 1, part 3, pp. 15-42.

3. deBrito G. (1988) Low bit rate speech for Ran European mobile communication system, in Proc. 38th IEEE Vehic. Technol. Conf., Philadelphia, PA, pp. 147-152.

4. IEEE Project 802, Local Area Network Standards, IEEE, New York, Aug. 1984.

5. Vishnevsky E., Anikonov A., Maslov V., Ushakov C. (1996) Optimal Cell Structure Design with Using Terrain Digital Model: The Adaptive Concept. Quarterly Technical Report, 91 P.

6. Vishnevsky E., Anikonov A., Maslov V., Ushakov C. (1995) High Resolution Digital Mapping Pilot Project. Final Technical Report. St. Petersburg, 113 P.

7. Vishnevsky E., Anikonov A., Maslov V., Ushakov C. (1995) Modeling and Calculation of Electromagnetic Fields in 3D Cities Using Terrain Digital Data. Final Technical Report. St. Petersburg, 65 P.

8. Vishnevsky E., Anikonov A., Maslov V., Ushakov C. (1995) Updating Shannon's Laws for Nonlinear Stochastic Processes and Under Dynamic Conditions. Final Technical Report. St. Petersburg, 295 P.

CHANNEL MODEL AND INFORMATION MAPS IN PERSONAL COMMUNICATION SYSTEMS

E. P. Vishnevsky, A. S. Anikonov, C. U. Ushakov
St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Science (SPIIRAS)
V. G. Maslov
S. I. Vavilov State Optical Institute
e-mail: lai@mail.iias.spb.su

Abstract. The relevant channel model for communication medium using its terrain digital model is introduced. On this basis we shall find the appropriate channel capacity, separating the principal ray-theoretical waves which define the deterministic part of the channel, and the non-theoretical waves which determine the fading part of the channel. In calculations we used the central part of Dallas square but in addition to electromagnetic power (for certain source position) we calculated some information values: channel capacity C and irreducible error probability P_e in every surface point.

Keywords. Terrain digital model, channel model, Rayleigh-fading, Rice-fading, multipath, Gaussian noise, intersymbol interference.

The general problem of statistical communication theory consists in a design of optimal communication system for a given source of message subjected to a chosen optimal criteria under certain restrictions. In many cases it is possible to define an ensemble of signals and to find a message – signal transformation which solve this problem. The finding of optimal transformation is considerably facilitated by its splitting into two operations - encoding and modulation. This permits at first to choose an optimal ensemble of signals taking into account the channel's properties and the restrictions without consideration the source of message. Then one finds an optimal encoding method considering the channel's characteristics as well as statistical properties of a message source. It should be noted that the great diversity of practical restraints in communication channels and optimal statistical criteria make impossible the general solution of the considered problem.

So, efficiency increase for each communication system may be performed in a numbers of ways

1) precising transform of message into signals

2) precising a signal transform

3) precising mathematical model of channel

The first way is related to finding "good" encoding schemes to control errors in a transmitter.

The second way is related to finding "good" transmission taking into account the channel's properties and the signal's restrictions.

The third way is related to objectively invariable part of the communication system. Particularly, the communication channel plays the most important role in design of real communication systems. Channel's capacity, as follows from the fundamental Shannon laws, determines maximal transmission rate of information which could be attained with any desired error frequency. Moreover, the channel statistical properties are one of important factors in calculating the error probability in a given communication system.

Thus, mathematical modeling of communication channel is the most difficult and important part in optimal design of communication systems. Moreover, we shall focus our attention on dynamic multipath communication media as channels with electromagnetic waves (EMW) as physical carrier of information transmitted through the channel. We meet this case in many actual communication systems. The next important feature of our approach consists in using terrain digital model (TDM) of the communication media based on high resolution Russian space images.

So, the priority purpose of this paper is to introduce the relevant channel model for communication medium using its TDM. On this basis we shall find the appropriate channel capacity, separating the principal ray-theoretical waves which define the deterministic part of the channel, and the non-theoretical waves which determine the fading part of the channel. Such separation will enable to precise the calculation of channel capacity and other important information characteristics. That is, for a particular communication system defined by its Protocol, precised error probabilities will be given for each point of the communication media using its TDM.

While discussing characteristics of a multipath channel, we have been tacitly assuming a number of paths of approximately identical strength for which the central limit theorem may be rigorously applied. However, rather often, a medium may be characterized, at least part of the time, as a medium having a few major strong stable paths with a number of additional weak paths. The stable paths often appear as a result of a stable EMW reflection or diffraction giving rise to the "principal ray-theoretical waves". Only using TDM of a communication medium, we may strictly identify and calculate these waves in each point of three-dimensional medium.

In such a case, it appears that the composite signal at the input of the Receiver in a fixed point will be composed of the sum of a steady (deterministic) signal $S_d(t)$, a Rayleigh-fading signal $S_f(t)$, having the usual Gaussian quadrature components, and additive bandpass Gaussian noise $n(t)$.

For some applications, the additive noise is also not stationary or not Gaussian. However, all our later analyses are confined to models involving stationary additive Gaussian noise. One reason is that this accurately describes additive noise in a broad scope of radio applications where the dominant noise is being originated in receiver "front ends." Secondly, even when external noise is dominant, it still often bears a Gaussian character within bandwidths of interest, with a relatively slow change of noise power level. The latter is similar in effect to a long-term signal fading in both those above described relatively slow changes in signal-noise ratio which must be compensated through providing an adequate system margin.

So, the composite received signal will have the form

$$S(t) = S_d(t) + S_f(t) + n(t) \tag{1}$$

where

$$S_d(t) = \sum_{k=1}^{N_d} \alpha_k(t) S_0(t - t_0 - \tau_k); \tag{2}$$

$$S_f(t) = \sum_{k=N_d+1} \alpha_k(t) S_0(t - t_0 - \tau_k) \tag{3}$$

Here N_d is the number of principal ray-theoretical waves and α_k are the path transmission factors for the *k*th path. The mean power of deterministic signal P_d may be strictly calculated using TDM on assume

$$P_f = \gamma P_d; \qquad 0 \le \gamma < 1 \tag{4}$$

for the mean power P_f of fading signal. The values N_d, α_k, ($1 \le k \le N_d$) and γ are assumed to be known in each point of the medium and may be considered as model parameters.

Under above assumption, the deterministic part of the received signal is the simplest linear transformation of an emitted signal $S_0(t)$, the statistics for which may be easily obtained.

The envelope statistics for Rayleigh- fading part of the received signal is described by

$$p(v) = \frac{2v}{P_f} \exp\left(-\frac{v^2}{P_f}\right) \tag{5}$$

v being "instantaneous root mean square (rms) value of Fading signal. In a particular case N_d the sum of two first terms in (1) gives "Rice-fading" signal, the envelope statistics of which is described by

$$p(v) = \exp\left(-\frac{P_d}{P_f}\right) \cdot \frac{2v}{P_f} \exp\left(-\frac{v^2}{P_f}\right) I_0\left(\frac{2v\sqrt{P_d}}{P_f}\right) \tag{6}$$

where v is an instantaneous rms value of a composite signal (without noise).

The formula (1) for the composite received signal in each point of a communication medium with a given TDM is a basic one for our further considerations of information characteristics of a communication system with dynamic multipath channels.

As mentioned above, the most important measure in the comparison of communication systems is the information transmission rate which can be attained. In this statement we imply that this information transmission rate be either free of errors or remains within some special limit of error frequency. A system capable of transmitting without errors at the rate C (channel capacity) is called an ideal system. Such a system cannot be achieved with any finite encoding process, but can be approximated as closely as desired as follows from fundamental Shannon laws.

We now derive the exact formula for the capacity of a dynamic multipath channel whose mathematical model is described above. The variations of EMW propagation characteristics of the dynamic multipath communication medium are assumed to be slow relatively to a signal duration.

According to the basic formula (1), we can treat the multipath channel as the product channel operated by independent sending a signal through each *k*th propagation path with an appropriate transmission factor α_k.

Then the capacity *C* of the product channel is the sum of the individual capacities C_k related to each *k*th propagation path [6].

The expression for C_k is given by [1].

$$C_k B \ln\left(1+\alpha_k^2 \frac{P_0}{P_n}\right) \tag{7}$$

where P_0 is the mean power of the emitted signal, P_n is the mean power of noise. Applying this formula we must distinguish the deterministic part of composite signal with α_k known exactly and its fading part for which α_k are random values with Rayleigh distribution [1].

Thus, we get

$$C = B\left[\ln\prod_{k=1}^{N_d}\left(1+\alpha_k^2\frac{P_0}{P_n}\right) - \exp\left(\beta^2\right)E_i\left(-\beta^2\right)\right] \tag{8}$$

where

$$\beta^2 = \left[\gamma\frac{P_0}{P_n}\sum_{k=1}^{N_d}\alpha_k^2\right]^{-1} \tag{9}$$

$$E_i\left(-\beta^2\right) = \int_{-\infty}^{-\beta^2}\frac{e^t}{t}dt \tag{10}$$

and γ is defined by (4). Formula (8) is required for the capacity of a dynamic multipath channel with additive Gaussian noise. In the particular case, when $N_d = 1$, $\alpha_1 = 1$ and $\alpha \to 0$ from (8) follows

$$C = B\ln\left(1+\frac{P_0}{P_n}\right) \tag{11}$$

It is the classical Shannon's formula for the capacity of a loosless channel.

Consider now in detail the parametric dependence of the channel capacity *C*. We see that

$$C = C\left(B, \frac{P_0}{P_n}, \gamma, N_d, \alpha_k\left(1 \le k \le N_d\right)\right) \tag{12}$$

For a given dynamic multipath communication medium with appropriate TDM the first and the second parameters are constant while the other parameters vary from point to point of the medium in a known way. So, for a dynamic multipath medium with its TDM the appropriate channel capacity may be calculated from (8) with a high accuracy in each point of the considered communication medium.

Discussion. The following sections consider a performance estimation for multi-user frequency hopping (FH) spread spectrum mobile radio system with binary frequency-shift keying (BFSK) modulation and, more specifically, an estimation of the effect of time delay spread on the above performance. The following consideration is closely related to the paper [2], however, approaches introduced by us and final formulas are essentially different.

We consider a multi-user FH-BFSK mobile radio system with S simultaneous users, and we will analyze the signal received by the s^{th} test user. The S simultaneous users share a common band with a width of about 10-20 MHz which is further divided into Q channels. Each channel consists of two FSK tones. Under a transmission and interference environment similar to the FH-MFSK system, this system can accommodate from 170 to 270 simultaneous users [2].

Because of the time delay spread, one user may be submitted to intersymbol interference (ISI) from other users. For example, let us consider three simultaneous users. A probability exists that they occupy a same channel q in succession at $(i-1)^{th}$, i^{th} and $(i+1)^{th}$ time intervals. At the i^{th} interval, user 2 uses the channel. At the beginning of this interval, he will be submitted to ISI caused by the delayed portion of the previous user signal (user 1). At the end of this interval, before he hops into a new channel, he may be submitted to the first arriving components of the next channel occupier (user 3).

We assume that for each user the system has a perfect address assignment, which is a specific time frequency hopping pattern with a length of L hops, and L>>1, and also guarantees that the chance of any two users using the same channel during the same hopping interval is zero. At each hop, according to this unique address, each user will occupy a new channel and he can possibly use all Q channels available. It will be also assumed that system synchronization is perfectly maintained. The purpose of the above assumptions is only to limit the multi-user interference to the ISI caused by the time delay spread. However, these assumptions do not represent fundamental restrictions, and the analysis can be extended to cover other cases.

Using above described model, the received signal could be written as follows:

$$R(t) = r_0 S(t) + \sum_{m=1}^{M} r_m S(t - \tau_m) + N(t) \tag{13}$$

where $S(t)$ is the transmitted signal and r_0 is a changeable coefficient which can take either 1 or 0 values. Summation is over all scattered components including both deterministic and Rayleigh-fading ones considered above in this paper.

In order to illustrate the developed approaches we have undertaken the attempt to realize some of them numerically for a real city square using TDM data. In our calculations we used the central part of Dallas square but in addition to EMW power (for certain source position) we calculated some information values: channel capacity C and irreducible error probability P_e in every surface point (at a given height above surface). We used the same ray approximation having considered all multiple diffractions on building roofs and in addition not more than one (for simplicity) reflection on building side walls. As a result of calculations every receiver point was characterized by a number m of rays (direct, reflected or bent by diffraction) falling to it. For every ray respective amplitude attenuation r_m and path length l_m were also calculated. In this realization only deterministic signal part was taken into account: we assumed $\gamma = 0$ and $\beta = 1$.

For channel capacity C equations (8, 9) were used (with $\gamma = 0$). We substituted all α_i by normalized amplitude attenuation values r_m which are assumed to be equal to 1 in point at a receiving height just bellow the source. The signal to noise ratio P_0 / P_n also was assumed to be related to the same point. The noise power level was supposed to be constant at every point of the space being studied. Frequency bandwidth B in our calculations may be selected arbitrarily and in the present calculations is assumed to be 50 kHz in all cases.

All calculations were performed for above described FH-BFSK system in full-occupancy case. For generality the formulas for absence of LOS signal were used. In our calculations it was possible to select one of three formulas for $\bar{t}$ calculations: $\bar{t} = \tau_0$, $\bar{t} = \tau(\max r_k)$, $\bar{t} = \sum_m \tau_m r_m / \sum_m r_m$. The selected rule was used for all receiver points. The signal to noise ratio was treated in the same way as for calculation of channel capacity. We used certain asymptotic formulas [11, 12].

All the following results are represented as pseudocolor maps. All intermediate values are displayed by mixed colors in correspondent ratios. Here we present results for two different source locations.

(a) Power maps. The power values are displayed in dB as ratios to the value in a point directly bellow the source: the highest value is 0 dB and the lowest one is -75 dB.

(b) Channel capacity maps. The channel capacity distributions calculated as described above. The values of the used parameters are as follows: frequency bandwidth B=50 kHz, signal to noise ratio $P_0 / P_n = 70$ dB. The channel capacity values are displayed in logarithmic scale: the highest value is 3000 kBit/s, the lowest one is 0.001 kBit/s. Areas where $R < C$ (R is a data transmission rate) are displayed for three different values of R. The comparison of these maps with power distribution shows that usually channel capacity is high where signal power is also high. However, a more attentive consideration shows that the value of C is determined not only by a summarized signal power but also by a signal "origin". In points where the signal originates from several reflected rays C is much higher than for points where only one ray falls on even if the power of this one ray is higher than a summarized power of several rays in the first case.

(c) Irreducible error probability maps. Signal to noise ratio was assumed the same as in calculations of C. $\bar{t}$ was calculated as $\tau(\max r_k)$. Values of P_e were calculated and displayed only in points where the Shannon conditions $R < C$ are valid. P_e is displayed in negative logarithmic scale: the lowest value is 2×10^{-7} and the highest value is 0.5. The consideration of presented results shows that P_e values for comparatively small R (30 kBit/s) reflect mainly the power distribution. But for higher values of R (approaching to 1 Mbit/s) the behavior of calculated P_e becomes in a certain sense analogous to that of C: for signals of multipath origin P_e becomes higher than for same power signals of one-path origin. (We do not discuss here whether it is possible or not to use a real FH-BFSK system in such frequency regime.

Note that the idealized qudrature detector does not contain any frequency selective elements and, as a consequence, frequency bandwidth is absent in equations for P_e). Such P_e behavior is originated from ISI effects peculiar for FH-BFSK-type systems. Because of these effects in points, where C is high due to multipath nature of signal P_e, values are also rather high and communication may becomes unreliable.

Results. Our channel's model makes it possible to render more precise calculation of different information characteristics in each point of communication medium and to make the optimal design of communication systems with nonlinear stochastic processes under dynamic conditions more realistic.

1. The capacity of a dynamic multipath channel is considered in details. This capacity is the most important information characteristic of a communication system which is embedded in Shannon Laws, and its analytical study is extremely complicated. On the basis of our channel's model the exact expression for channel capacity is derived under reasonable assumption of slow channel dynamic relative to the signal duration. This formula enables to calculate with high accuracy the values of channel capacity in each point of the communication medium using its TDM, and, thereby, to create the "map of channel capacity" for the considered city domain.

2. Analytical technique for the evaluation of the average error probability of incoherent FH BFSK communication over a dynamic multipath channel is elaborated on the basis of our mathematical channel's model, and presented an example of its application. The appropriate formulas may be used to calculate the values of error probability in each point of the communication medium using its TDM.

3. The prototype of computer-aided prediction system (CAPS) software has been elaborated to compute the maps of an information channel capacity and a communication reliability (irreducible error probabilities) in a central part of Dallas. The calculations were performed on the basis of TDM previously elaborated by us and our above analytical results. Main results are represented in a form of information maps.

Thus, our mathematical channel's model, the adequate analytical formulas for channel capacity and the error probability, and the modified computer aided prediction system (CAPS) prototype enable to refine and visualize the fundamental Shannon Laws for nonlinear stochastic processes under dynamic conditions, i.e., to show the domain of reliable communication and the appropriate spatial distribution of error probabilities over a dynamic multipath communication medium for a given position of information source in its TDM. Such "information maps" provide us with a powerful tool for optimal design of mobile communication systems.

References

1. Fink, L. M. (in Russian) (1970) Theory of Discrete Information's transmission, Moskva.

2. Niyonizeye, G., Lecours M., Huynh H. T.(1985) Mutual Interferences in a Binary FH-FSK Spread Spectrum System for Mobile Radio, IEEE Trans Veh. Technol., vol. VT-34, no. 128-34.

3. Shannon, C. E. (1948), A Mathematical Theory of Communication, Bell System Tech. J., 27, 379-423 (Part I), 623-656 (Part II). Reprinted in book form with postscript by W. Weaver, Univ. of Illinois Press, Urbana, 1949.

4. Shannon, C. E. (1949) Communication in the Presence of Noise, Proc. IRE. 37, 10-21 .

5. Shannon, C. E. (1956), The Zero Error Capacity of a Noisy Channel, IRE Trans. Inform. Theory, IT-2, 8-19.

6. Shannon, C. E. (1957), "Certain Results in Coding Theory for Noisy Channels," Inform. and Controll, 1, 6-25

7. Shannon, C. E. (1959), Coding Theorems for a Discrete Source with a Fidelity Criterion, IRE Nat. Conv. Records, Part 4, 142-163. Also in Information and Decision Processes, R. E. Machol, Ed., McGraw-Hill, New York (1960)

8. Shannon, C. E. (1959), Probability of Error for Optimal Codes in a Gaussian Channel, Bell System Tech. J., 38, 611-656

9. Shannon, C. E. (1961), Two-way Communication Channels, in Proc. Fourth Berkely Symp. on Prob. and Stat., 1, 611-644, University of California Press, Berkeley, Calif.

10. Shannon C. E., Gallager R. G., Berlekamp E. R. (1967) Lower Bounds to Error Probability for Coding on Discrete Memoryless Channels, Inform. and Control, 10, 65103 (part I), 522-552 (Part-II).

11. Schwartz, M., Bennett W. R., and Stein S. (1966) Frequency-hopped Multilevel FSK for Mobile Radio, Bell Syst. Tech. J., vol. 59, 1257-1275.

12. Wang, K. D., Lecoure M. L.(1994) The Effect of Delay Spread on FH-FSK Spread Spectrum Mobile Radio System Over Frequency-Selective Fading Channels, IEEE Trans. Com., vol. 42, no. 2,3,4, 312-324

QUALITY OF SERVICE MANAGEMENT ARCHITECTURE FOR DISTRIBUTED APPLICATIONS

Maxim Alexandrov, Alexander Hagin, Nikolai Hagin, Yuri Karpov, Vladislav Voinov

St.Petersburg State Technical University
Distributed Computing and Networks Dept.
Politekhnicheskaya ul., 29
St.Petersburg, Russia 195251
E-mail: ma@dcn.nord.nw.ru, alexandr@hagin.spb.ru, nikolay@hagin.spb.ru,
karpov@hpc.spb.su, vlad@dcn.nord.nw.ru

Abstract. In the paper distributed applications (DA) Quality of Service (QoS) management problems are described and an architecture for automated QoS monitoring and operational control is presented. Monitoring and operational control schemes using model-based reasoning are discussed. QoS model building method includes a hierarchical set of user request processing models and resource models. A task-to-resource mapping model allows to construct so-called clusters of resources and tasks, which are represented by performance models interconnected through parameters. We demonstrate the proposed architecture and methodology applied to QoS management of a Web system. QoS models of a Web system and user HTTP request processing, and their application to response time prediction and operational control are explained. Finally, main results and possible future directions of the research are discussed.

Keywords: Quality of Service management, management architecture, monitoring and operational control, distributed applications management, Web system management.

Introduction. A distributed computing environment comprised of dozens of computers, local and wide area networks, peripherals, operating systems, middleware and applications and where resources are commonly shared, often carries a heavy bundle of high complexity, so generally distributed systems are not easy to manage. A considerable research contribution to the DA management was made over the past several years within the field of QoS management, in particular for multimedia applications.

In [1] an OSI compatible QoS architecture that reconciles the user-oriented QoS requirements with the network-oriented QoS environment for distributed multimedia applications is defined. In [2] QoS adaptive transport system is implemented so that it incorporates a QoS oriented API and a range of mechanisms to assist multimedia applications in exploiting QoS and adapting to fluctuations in QoS.

QoS specification and layer-to-layer QoS mapping are discussed in [3]. A DA logical structure and QoS negotiation approach and a corresponding negotiation and resource reservation protocol NRP are presented in [4].

Mapping DA logical structure to computer network permits to find the optimal placement plan of DA topology into distributed computer system (DCS) taking into account the resources required by the DA and DCS resources available in an offered time interval of the DA session. Some polynomial complexity algorithms for solving the problem are proposed in [5, 6]. In [6] an effective algorithm for DA to DCS re-mapping is proposed.

In [7] a loosely hierarchical architecture based on a distributed approach is proposed to support end-to-end QoS management functions in a multi-domain for multimedia applications. The domain QoS manager and QoS agent concepts, and a hierarchical negotiation protocol are introduced to build the proposed architecture.

Management strategies on the system as a coherent entity are discussed in [8]. An approach to distributed systems management based on system policies is described and a methodology of their expression by rules is presented. Implementation and interpretation of policies in Marvel software development environment are shown. An architecture for monitoring agents is presented and smart agents methodology is discussed.

Despite of the importance of QoS DA management and need for respective policies, that now are widely discussed, the problem hasn't been solved in general yet. No generic QoS DA management architecture including administrative, management, control and measurement systems interconnected with each other has been defined to support in a flexible way different QoS monitoring and management goals at various phases of the DA lifecycle.

In this paper we propose an architecture for Quality of Service management for DAs. An approach to building monitoring and operational control schemes using model-based reasoning and QoS models is discussed. The methodology is then applied to Web system management on the base of QoS models, which are developed in accordance to the methodology. The experiments conducted on a testbed proved acceptable accuracy of the models that can be used for Web QoS prediction and provision.

The paper is structured as follows. First, we describe a QoS management architecture. The model-based approach to building monitoring and operational control schemes is proposed. Then we present a method for model building and performance evaluation of DA. The method includes a hierarchical set of user request processing models and resource models. A task-to-resource mapping model allows to construct so-called clusters of resources and tasks, which are represented by parametrically interconnected performance models. Further sections demonstrate the proposed architecture and methodology applied to QoS management of a Web system. Here QoS models of a Web system and user HTTP request processing, and their application to response time prediction and operational control are explained. Finally, we give some concluding remarks on our results and future research directions.

Model-based approach. Application users think in terms of high-level transactions rather than in terms of computers, networks, hubs, switches, etc., and their capacities and workloads. The user perceives so-called Quality of Service (QoS), which the system provides to him or her. This QoS is expressed in requirements (QoS goals) to the transactions, such as response time, and their costs, and to the results of these transactions, such as content of the response (e.g. results of a search in a database), its size, image colorfulness, etc.

In a given example (see Fig. 1) there are several machines that have WWW and Mail client software installed. Through a network they are connected to the Mail server, the Web Server and the Database Server (note that the latter two reside on the same machine). User QoS requirements focus on the response time he or she perceives performing requests to remote servers. The user goal can be expressed as some constraints on the response time (e.g. ResponseTime < 200 ms). This time is comprised of:

1. Time client-side software spends parsing the user request and preparing the request that will be sent to the server.
2. Network delay of the request delivery.
3. Request processing and response generation on a server side.
4. Network delay for delivering the response.
5. Response processing on a client side (e.g. displaying the result).

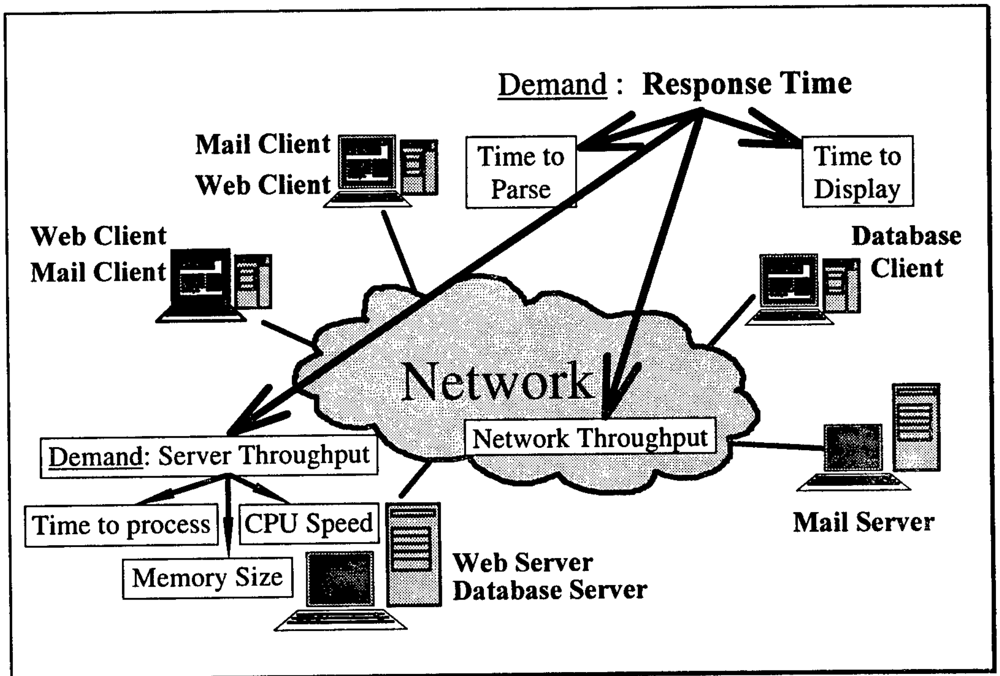


Fig. 1 QoS requirements and resource sharing

Therefore, response time constraints introduce some constraints on those processing delays listed above, and, subsequently, on performance parameters of corresponding subsystems that perform processing. For instance, user QoS requirement may impose requirements on the network (throughput, error rate), the server (CPU speed, size of available memory, etc.) If something goes wrong the administrator should be able to trace down the cause of the problem (perform diagnosis), i.e., see which constraints on application and system elements configuration and behavior are imposed by user QoS requirements, whether they are satisfied and find a good solution to eliminate QoS degradation.

In general the problem is quite complex due to a high complexity of DAs and systems, because they often are heterogeneous and may operate in different configurations, so that all the particulars cannot be specified and evaluated a priori. Moreover, resources are commonly shared between different applications and users, so that it's not sufficient to consider only one application, or just one particular user of that application.

All DA management tasks, such as DA installation, QoS monitoring, prediction, diagnosis, etc., have a lot of common. To perform any of them one should know the dependencies between QoS parameters, that comprise QoS goal, and system parameters at various levels of the system architecture, plus properties of the parameters (measurable or not, controllable or not). These dependencies and properties are interpreted specifically for a particular management task and management policies.

These parameter dependencies and properties are put into models, that describe DA and system component structure and behavior. The completeness and level of detail of those models depend on goals of management and existing measurement and control tools. Depending on a management task appropriate models should be combined to provide calculations and reasoning needed for the task. Requirements to models and model builder are discussed in further subsections.

Management Architecture. Fig. 2 outlines an architecture for DA management. At first, to be able to track perceived QoS levels *monitoring* should be provided. Given a particular goal from the Administrative System the Management System should build a *monitoring scheme*, so that the goal expressed through high-level application parameters could be evaluated by means of the Measurement System. Thus, the Management System has to find out, which parameters to measure and how to derive high-level application parameters from those measured. Such a mapping is performed by the Model Builder, which uses various system and DA models for constructing a model (e.g. system of equations), which binds QoS parameters with measurable ones.

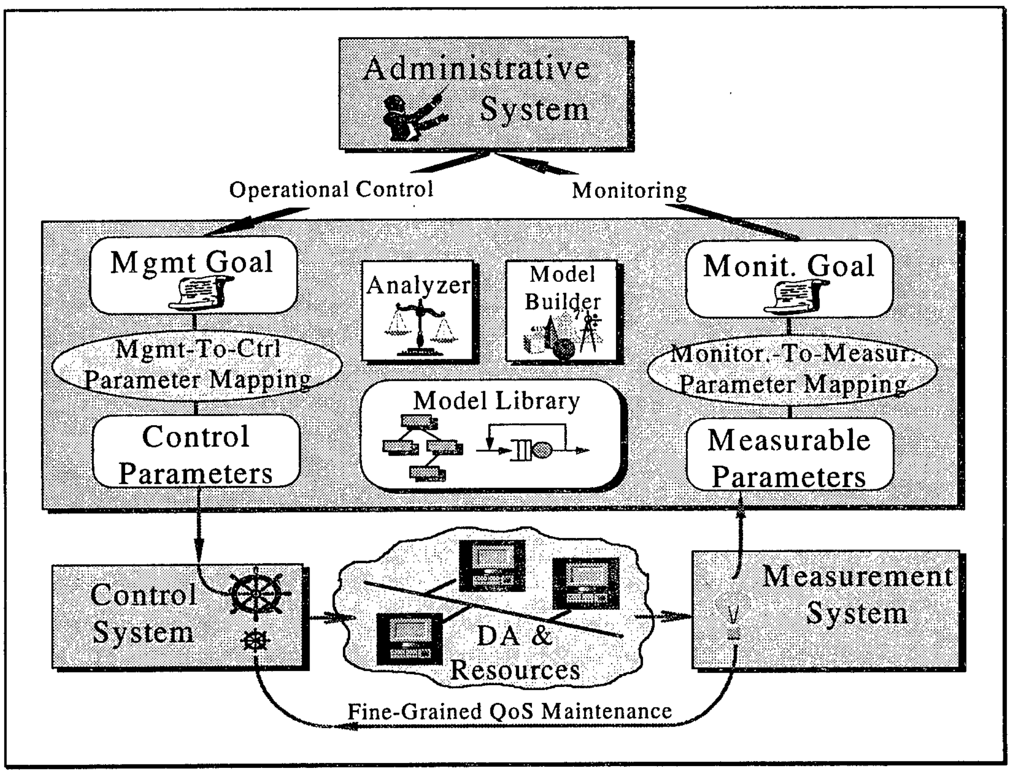


Fig. 2 DA Management Architecture

Once it's done, the Administrative System is able to check whether specified QoS goals are satisfied or not. If a goal is violated, the Administrative System can issue a command (with help of administrator, or automatically) to bring QoS back to the desired level. Such a command may also be specified as a goal defined on high-level parameters, for instance, "make the average response time less than a threshold."

Having received a command the Management System builds an *operational control scheme*, which would result in a set of application and system (controllable) parameters, that could be changed in order to satisfy the goal. Thus, the Management System should map QoS parameters specified in management goal to those, which can be set to satisfy the goal. With help of the Measurement System and management policies the Analyzer performs diagnosis to track down, which performance parameters caused QoS degradation, and using the Model Builder maps them to controllable parameters and determines their new values to set.

When an appropriate control as a set of parameters and their new values is found, the Control System is activated to conduct fine-grained resource adjustments tuning resource primitives and application parameters. Thus, the Control System directly regulates application execution and operates in real-time implementing dynamic QoS maintenance mechanisms.

If the Management System cannot achieve the goal changing controllable parameters, it issues an indication to the user. If the user has offered the 're-negotiation' policy in the service contract, then he or she gets degradation indication and the re-negotiation request. In this case QoS re-negotiation is executed and new agreed level of QoS requirements is established. Then application session is continued with a new level of QoS.

The Administrative System supervises all these activities forming and pursuing management objectives. It connects monitoring and control branches in the management loop formulating management goal when the monitoring goal becomes false.

Described "big" management loop is run in a time scale of seconds to dozens of seconds. To be able to perform QoS management dynamically, a "small" management loop is executed in a time scale of milliseconds, when the Monitoring System performs monitoring of QoS levels, and if they become unsatisfactory, it issues a command of fine-grained QoS maintenance to the Control System. If the Control System cannot adapt the system, then "big" management loop is executed.

Monitoring scheme. QoS monitoring is used during an application execution to track the ongoings of QoS levels with the help of the Measurement System and to compare them with the requested QoS levels. A goal of monitoring is formulated as a sentence (function) through QoS parameters and their thresholds. The Management System checks a value of the goal while the application is running. When the value of the goal becomes false, the Administrative System is immediately notified about goal violation.

From the view of monitoring, all the parameters can be classified in three groups: measured , computed (derived) and assigned (e.g. default) ones. The monitoring goal has to be expressed through QoS parameters, each of which is either measured or computed. If a parameter is immeasurable, then it has to be computed with help of a model through other parameters, that can be measured or are assigned in advance.

Thus, one of the problems of monitoring setup phase is to build *hierarchical* models parametrically connected with each other, that allow to compute immeasurable QoS parameters used in the goal sentence. An algorithm for building the monitoring scheme was built and its effectiveness and applications are under investigation.

Operational control scheme. One of the reasons, why the operational control is executed, is monitoring goal violation. For example, one of the monitoring parameters violates its threshold, and the monitoring goal becomes false. Then the problem is to find controllable parameters, varying of which allows to change the monitoring parameters, so that the goal will become again true.

In the algorithm for monitoring goal evaluation the controllable parameters are assigned ones, i.e. input ones. Therefore, the high-level parameters directly used in the monitoring goal sentence are not directly controllable. The Analyzer takes the model generated and embedded in the monitoring scheme, and looks for controllable parameters in this model. If such parameters are found, the Analyzer uses the model to determine needed values of these parameters to achieve the management goal. Note, that the model is used in the inverse mode to inference the contribution of the arguments to the function value.

If changing the controllable parameters does not allow to achieve the management goal, then the Analyzer asks the Model Builder to increase the model to compute parameter used in the goal sentence through other controllable parameters, i.e. to find another parameter decomposition. An approach can be as follows: controllable parameters must be as high as possible at the decomposition level and provide the most effective control mechanism. According to this scheme the model enhancing is started from high-level parameters.

An algorithm that was developed for operational control scheme building also uses one of the model-based approaches for algorithm construction to make a management decision. Authors plan further investigation of the efficiency and development of the algorithm.

A model building method. Proposed method for model building and performance evaluation of DA is based on [9].

A model of *user request processing* (shortly, task model, or T-model) uses the oriented graph notation for describing a stage (task) sequence of user request processing and parameters of the task interconnections with each other. The model is constructed as a hierarchical system by sequential decomposition of the task processing. The lowest level of the decomposition depends on the system level which permits to obtain input data for models (e.g. with help of the Measurement System), and on the objective of the modeling, too. In particular, T-models are used for computing an average system response time, task active and waiting times.

A *resource* model (shortly R-model) is presented as a hierarchical set of models, too. It describes a set of resources, such as computer-client, computer-server, transport system, CPU, disk etc. and their parameters, which are used for task execution. The R-model decomposition is done in accordance with the task model decomposition, so that every task has a resource, where the task is executed. Note, that generally the number of levels in the T-model is not necessarily equal to the number of levels in the R-model.

A decomposition of task and resource models for a Client-Server system is presented in Fig. 3. At the level 1 task Th represents user thinking process, when he or she prepares a current request to the system. This task is executed by resource User, i.e. the user himself executes this task.

Task Serv is a complex one and represents all the actions that a service system performs during user request processing. Resource Sys of the service system is used by the task. At level 2, task Serv is decomposed onto three tasks: Cl_Rq for the request preparation by the client endsystem, Cl_Dsp for displaying a system response to the user by the client endsystem, and Tr_S for the transmission, client request processing and the system response. This task decomposition permits to decompose resource Sys onto two resources: Client of the client system and Net_S of the rest of the service system (consisting of the network and the server system).

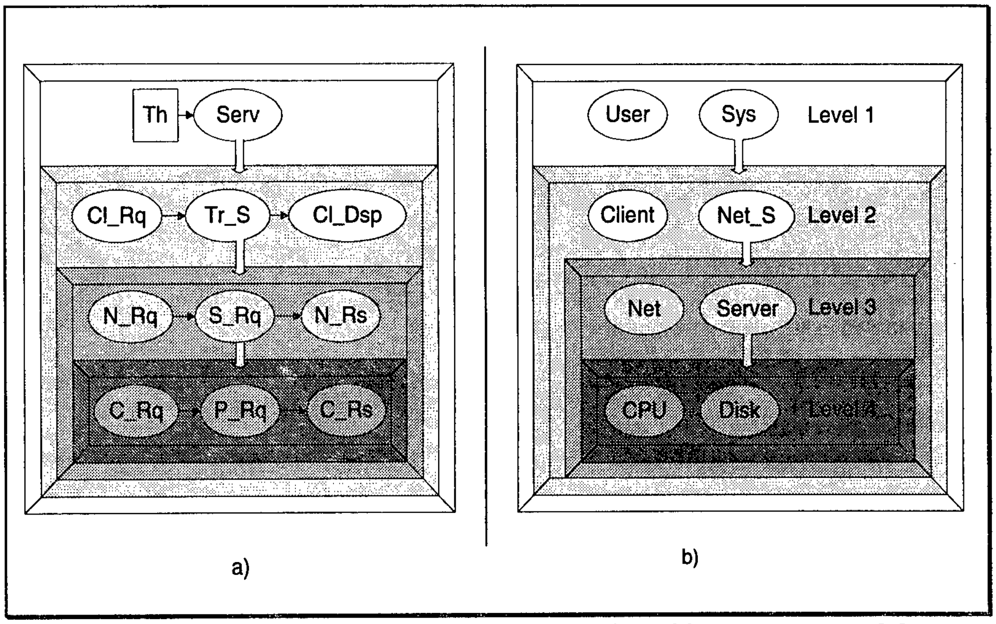


Fig. 3 Decomposition processes for a) task and b) resource models

Let us give an explanation to the notation used in Fig. 3:

- Cl_Rq - the task of request preparation by the client system,
- Cl_Dsp - the task of displaying a system response to the user by the client system
- Tr_S - the transmission task
- N_Rq - the task of client request transmission by the network
- N_Rs - the task of server response transmission by the network
- S_Rq - the task of client request processing by the server
- C_Rq - the task of client request reception by the server
- C_Rs - the task of response sending by the sever
- Client - the client system
- Net_S - the service system consisting of the network and the server system
- Net - the network resource
- Server - the server computer resource
- CPU - the CPU resource of the server computer
- Disk - the disk resource of the server computer

The T-model and the R-model can be constructed using different decomposition level for different tasks and resources minding that an ancestor and its heirs cannot be represented in the same model.

A *task-to-resource mapping* model (shortly, TR-model) represents a distribution of tasks among resources. The TR-model permits to form so-called *clusters* of resources and processes that are represented as separate, interconnected performance queuing network models. The tasks mapped onto the same resource are gathered into one group. Then all the resources, to which tasks of this group are mapped, are included into the group. If added resources have one or more tasks, that are mapped to them, and are not in the group, then these tasks must be added to the group, too. Then this process can be repeated for the new tasks included in the group. The result of this process is a cluster of interdependent tasks and resources.

Fig. 4 illustrates the TR-model. The T-model and the R-model consist of the leaves of the decomposition tree for tasks and resources depicted in Fig. 3 (a) and (b) correspondingly. The dashed arcs show mapping of every task to corresponding resources. A task can use one or more resources; one resource can be used by one or more tasks. For example, task S_Rq uses resources CPU and Disk, and resource Net is used by two tasks N_Rq and N_Rs.

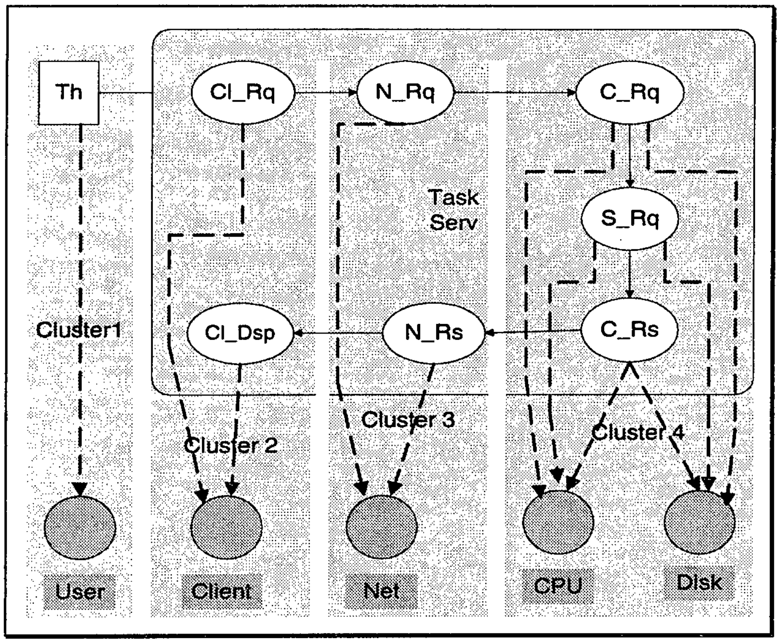


Fig. 4 Task to resource mapping model

For every cluster corresponding performance queuing network model is built. For example, for cluster 4, a heterogeneous closed queuing network with three classes of customers corresponding to the three tasks executed on the resources can be constructed. The customer population in every class is determined by the processing concurrency level for corresponding task. Resources are represented by queue nodes in the model. One queue node imitates the waiting time of every task process, when there are no requests to the task. Such kind of the network queuing model is presented in [9].

The proposed method is close to the Method of Layers [10]. However, it has some differences such as:

- construction and computation of the hierarchical task models differ from the hierarchical software models proposed in [10];
- using TR-model permits to construct, by a natural way, a set of separate, but interdependent performance models for DA systems. A set of different mapping models can be constructed by considering tasks and resources at various decomposition levels;
- set of performance models can be derived from the corresponding mapping model by clustering tasks and resources of different decomposition levels. Not all the tasks and resources must be presented in the performance models as in the Method of Layers.

Using exact computational and approximate techniques [10-14] general queuing network models can be computed. The set of computed parameters is usually enough to calculate task and resource parameters for DA and system performance evaluation. As shown above, for real systems often the service time parameters and transition probabilities can be computed through the parameters of tasks and resources, plus average number of task visits to each resource.

Experimentation. The proposed architecture and methodology were applied to QoS management of a Web system. Corresponding QoS models of user HTTP request processing were built and applied for prediction of Web system response time for a user and operational control by varying Web server performance parameters, such as bandwidth level and maximum concurrency level. The experiments conducted on a testbed proved acceptable accuracy of the models. Developed QoS models can be used for Web QoS prediction and provision.

Fig. 5 depicts the plots T of the average system response time, τ_{S_Rq} of the average residence time of the client request in the Web server and CurB of CurrentBandwidth as functions of BandwidthLevel. The curves were got by the means of corresponding performance models. Measured values of the system response time T are shown in the figure with symbol *. In the experimentation MaxConcurrency was equal to 4, and the client number N was 20.

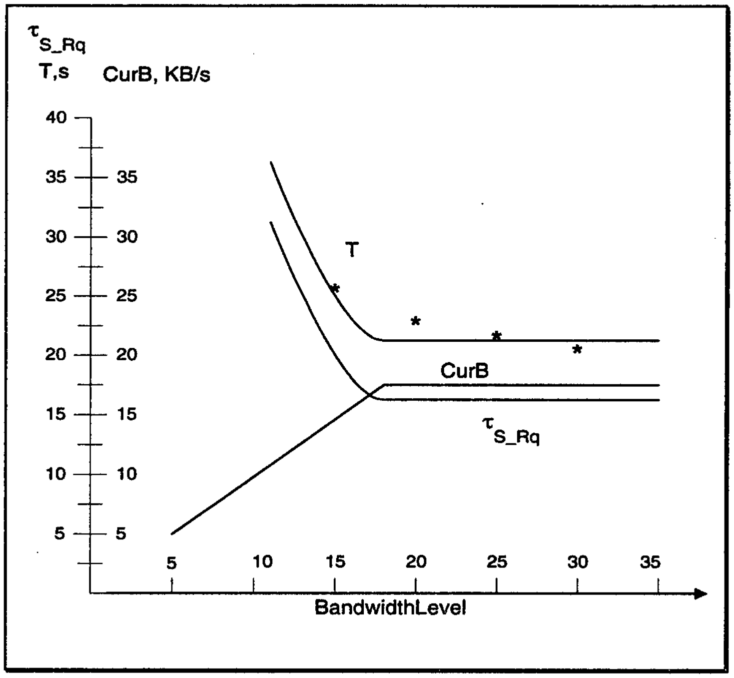


Fig. 5 Web system performance characteristics

The plot T shows that there is an interval of BandwidthLevel from 5 KB/s to 18 KB/s where changing BandwidthLevel causes changes in system response time. On this interval the server limits the CurrentBandwidth, so that it is always equal to BandwidthLevel, thus utilizing no more than BandwidthLevel of network bandwidth and letting other networked applications use the network. This is achieved by delaying user requests by the Web server, which causes system response time increase.

Conclusions. In the paper we have demonstrated a methodology for Quality of Service management for distributed applications. Major elements of the architecture were described, and a brief description of the QoS models for DAs, in particular, task processing model, resource models and task-to-resource mapping models, their building methodology and application were presented. Finally, we described how the architecture and QoS models were applied for QoS management of the Web system.

In the future we would like to widen the coverage of the QoS DA management tasks based on our approach, for instance, the further investigation and development of algorithms for building monitoring and operational control schemes, QoS prediction and adaptation, and apply our theoretical results to Web management. Another area of the research is application of different models for behavior specification and elaboration of the requirements to the performance Model Builder, including questions about model sufficiency and completeness.

References

[1] Campbell, A., Aurrecoechea, C., Hauw, L. (1995). Architectural Perspectives on QoA Management in Distributed Multimedia Systems. Proc. Second Workshop on Protocols for Multimedia Systems "Mozart on Multimedia Highways", Salzburg, Austria, pp. 329 - 353.

[2] Campbell, A., Coulson., G. (1996). A QoS Adaptive Transport System: Design, Implementation and Experience. Proc. ACM Multimedia 96, Boston, pp. 117-127.

[3] Dermler, G., Fiederer, W., Barth, I., Rothermel, K. (1995). A Framework for Negotiable Quality of Service in Distributed Multimedia Applications. Tech. Report, Fakultatsbericht, No. 10, Universitat Stuttgart, p. 19.

[4] Dermler, G., Fiederer, W., Barth, I., Rothermel, K. (1995). A Negotiation and Resource Reservation Protocol (NRP) for Distributed Multimedia Applications. Tech. Report, Fakultatsbericht, No. 11, Universitat Stuttgart, p. 27.

[5] Iqbal, M.A., Hagin, A. (1996). Partitioning and Mapping Techniques for Distributed Multimedia Applications. Tech. Report, Fakultatsbericht, No.14, Universitat Stuttgart, p. 23.

[6] Hagin, A., Dermler, G., Rothermel, K. (1996). Mapping of Distributed Multimedia Applications Based on a Sequential Method. Tech. Report, Fakultatsbericht, No. 16, Universitat Stuttgart, p. 51.

[7] Hafid, A. (1995). *Hierarchical Negotiation for Distributed Multimedia Applications in a Multi-Domain Environment.* Second Workshop on Protocols for Multimedia Systems "Mozart on Multimedia Highways", Salzburg, Austria, pp. 397-409.

[8] Koch, T., Krämer, B. (1996). Rules and agents for automated management of distributed systems. Special issue of the Distributed Systems Engineering Journal on Distributed Systems Management, The Institution of Electrical Engineers and Institute of Physics Publishing.

[9] Hagin, A., Klimovskiy, O.V., Cherkesov, G.N., et al. (1991). Algorithms and Software Design and Implementation to Support Required Reliability of Telecommunication Management System Based on Computer System SM 1425. Tech. Report, No. 23290, Institute of Machine and Technology Reliability, Russian Academy of Science, Leningrad.

[10] Rolia, J.A., Sevcik, K.C. (1995). The Method of Layers. IEEE Transactions on Software Engineering, Vol. 21, No. 8, pp. 689 - 700.

[11] Bryant, R.M., et al. (1984). The MVA Priority Approximation. ACM Trans. On Comp. Systems, Vol. 2, No. 4, pp. 335 - 359.

[12] Kapelnikov, A., Muntz, R.R., Ercegovac, M.D. (1989). A Modelling Methodology for the Analysis of Concurrent Systems and Computations. J. Parallel and Distributed Computing, No. 6, pp. 568 - 597.

[13] Basharin, G.P., et al. (1989). Queuing Analysis for Computer Networks. Theory and Computational Methods. Nauka, Moscow, p. 336.

[14] Nelson, R.D. (1993). The Mathematics of Product Form Queuing Networks. ACM Computing Surveys, Vol. 25, No. 3, pp. 339 - 369.

Component Identification in the Heterogeneous Processes Arising in High-Speed Networks †

Denis A. Khotimsky (dkhotimsky@lucent.com)

Bell Labs – Lucent Technologies
101 Crawfords Corner Road,
Holmdel, NJ 07733-3030, U.S.A.

Abstract. Performance measurements in high-speed networks may often be interpreted as samples from a mixture of state-dependent distributions. The set of states, which may correspond to components of user profile, levels of congestion, status of the network elements, etc., is fixed but hidden from the observer. To generalize such phenomenon, we introduce an abstract model of *heterogeneous process* which is a doubly-stochastic process with hidden state $\{\Omega_n\}_{n=1}^{\infty}$ and the observable state-dependent random output $\{T_n\}_{n=1}^{\infty}$. To exercise efficient control, a network element must be able to estimate the expected value of the output process and to refine the heterogeneous process model as the new observations become available. This paper make use of a tracking procedure based on exponential smoothing to estimate the expected output and propose a new algorithm for the identification of the heterogeneous process component distributions. The algorithm uses the smoothed output values as a basis for the conditional sampling, or censoring, of the observations.

Keywords: Traffic measurement, Exponential smoothing, Hidden Markov chains, Heterogeneous processes.

1 Introduction

Performance measurements in high-speed networks may often be interpreted as samples from a mixture of state-dependent distributions. The states of a system may correspond, for example, to active components of a user traffic profile, levels of congestion, variations of network topology, status of the network elements, etc. These states belong to a fixed set, but remain hidden from the observer. A system sojourns in state i for some random interval, generating the output T drawn from the distribution $F_i(x)$, and makes a transition to a new state j afterwards. A notion of *heterogeneous process* is introduced to generalize such type of environment. A heterogeneous process is a doubly-stochastic process with hidden state $\{\Omega_n\}_{n=1}^{\infty}$ and the observable state-dependent random output $\{T_n\}_{n=1}^{\infty}$. The problems that arise in the context of high-speed network management involving heterogeneous processes are

(i) Given the model of a heterogeneous process, including the parameters of state process and the output distribution functions, find the probability of the observed output sequence

(ii) Given the model parameters and the output sequence, reconstruct the most likely sequence of hidden states

†This research has been completed by the author while he was with the Computer Science Department, University of California, Santa Barbara, using solely resources and equipment of UCSB, and was supported by the NSF Grant NCR-9103998.

(iii) Construct an estimator that would track the expected output of a heterogeneous process

(iv) Given the output observations, find (refine) the heterogeneous process model

The formulation of these problems suggests that the classic method of the hidden Markov chains be examined as a possible approach to their solution. Indeed, the framework of Hidden Markov models provide an efficient way to compute the probability of observed output given the process model and to refine them as a new set of data becomes available. This result can be further used by a Viterbi-type decoder to find the most likely hidden state sequence. However, these methods are computationally intensive and require iterative off-line processing of large sets of data [5]. Therefore, their applicability for the control and management of the high-speed networks is limited, and the fast on-line alternatives should be sought.

An exponential smoothing estimator, that combines the observations in a linear combination with the weights that fall in geometric progression with the age of the observation, provides a foundation for such an alternative.

Ahmadi and Kermani [1] considered the problem of real-time traffic load estimation in a packet-switching networks. They proposed an adaptive by-modal estimator, a combination of the optimal linear filter and the exponential smoothing, pointing out the problem of appropriately choosing the smoothing constant. Khedkar and Keshav [4] described an exponential averaging predictor operating under a fuzzy logic controller to determine the value of the smoothing parameter that gracefully responds to the changes in the system behavior. Warfield, Konheim et al [8] suggested a doubly-recursive variation of the exponential smoothing estimator for the cell arrival rate in an ATM network with multiple underlying states. The parametric view on the traffic model estimation using EM algorithm was examined by Yegenoglu and Jabbari [9].

This paper presents a new algorithm for non-parametric identification of the component distributions of the heterogeneous process. The algorithm employs the exponential smoothing of the output as a basis for the conditional sampling, or censoring, of the observations.

The rest of the paper is organized as follows. Section 2 formalizes the abstraction of heterogeneous process and outlines the properties of an estimator based on exponential smoothing. Section 3 presents a component identification algorithm. Simulation results are discusses in Section 4, followed by concluding remarks.

2 Abstraction of a Heterogeneous Process

Let $\{\Omega_n\}_{n=0}^{\infty}$ be a stochastic process taking values in the finite set $E \subset \mathbf{N}$.

Definition 1 (Heterogeneous Process) *A stochastic process* $\{\Omega_n, X_n\}_{n=1}^{\infty}$ *is a heterogeneous process if*

(h1) *the distribution of* X_n *depends on the state* Ω_n *and does not depend on the epoch* n, *i.e.* $\{\Omega, X\}$ *is time-homogeneous:*

$$\mathbf{Pr}\{X_n \leq x \mid \Omega_n = k\} = F_k(x); \tag{1}$$

(h2) *given the sequence of states* $\bar{\Omega} = \{\Omega_i\}_{i=1}^{n}$, *the corresponding values of* $\bar{X} = \{X_i\}_{i=1}^{n}$ *are conditionally independent:*

$$F_{\bar{X}|\bar{\Omega}}(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} F_{\Omega_i}(x_i). \tag{2}$$

The process $\{\Omega_n, X_n\}_{n=1}^{\infty}$ is *M-phase heterogeneous* if the range of Ω_n has cardinality M.

In a broad sense, $\{X_n\}_{n=1}^{\infty}$ may be viewed as a sequence of observations drawn from a distribution which is chosen according to the state Ω_n. The index n is *an observation epoch.*[1] An important special case of the heterogeneous process is obtained if the underlying state process $\{\Omega_n\}_{n=0}^{\infty}$ is a finite-state discrete-time Markov chain.

Definition 2 (Markovian Heterogeneous Process) *A heterogeneous process $\{X_n\}_{n=1}^{\infty}$ is* Markovian *if the underlying state process $\{\Omega_n\}_{n=0}^{\infty}$ is a Markov chain.*

It can be immediately observed that for any state $j \in E$ of a Markovian heterogeneous process, the length L_j of a sojourn in state j has a geometric distribution:

$$\mathbf{Pr}\{L_j = k\} = (1 - q_j)q_j^{k-1}, \tag{3}$$

where $q_j = \mathbf{Pr}\{\Omega_{n+1} = j \mid \Omega_n = j\}$.

Definition 3 (Semi-Markovian Heterogeneous Process) *A heterogeneous process $\{X_n\}_{n=1}^{\infty}$ is* semi-Markovian *if the underlying state process $\{\Omega_n\}_{n=0}^{\infty}$ is a finite-state discrete-time semi-Markov process.*

In the case of a semi-Markovian heterogeneous process the sojourns have a general distribution that is dependent on the state.

3 Heterogeneous Component Identification

Let $\{\Omega_n, T_n\}_{n=1}^{\infty}$ be an M-phase heterogeneous process with the hidden (unobservable) states $\{\Omega_n\}$ and the output observations $\{T_n\}$ drawn independently from the state-dependent distributions $F_k(x)$, $k = 1, \ldots, M$. Let the component distribution function $F_k(x)$ have finite moments

$$\mu_k = \int x F_k(dx) \tag{4}$$

$$\sigma_k^2 = \int (x - \mu_k)^2 F_k(dx). \tag{5}$$

Given the sequence of observations $\mathcal{T}_n = \{T_i\}_{i=1}^{n}$, we want to det ermine the set of expected observation level estimates $\{\hat{\mu}_k,\ k = 1, \ldots M\}$ corresponding to the M phases of the heterogeneous process. The approach to this problem lies in the proper choice of the set of sampling criteria $D_k(\mathcal{T}_n)$ that can be evaluated based on the observed history up to the n-th observation epoch, and used instead of the condition $\Omega_i = k$ in the empirical distribution functions computation:

$$\hat{F}_{k,n}(x) = \frac{\sum_{i=1}^{n} \chi_{\{T_{\Omega_i} \le x,\ D_k(\mathcal{T}_i)\}}}{\sum_{i=1}^{n} \chi_{\{D_k(\mathcal{T}_i)\}}}. \tag{6}$$

The proposed algorithm employs the exponential smoothing of the output observations and uses the assumption that for the large class of heterogeneous processes (which

[1]To avoid confusion caused by the fact that in many practical cases, the output observations of a heterogeneous process have dimensionality of physical time, an attempt is made to circumvent time-series terminology where possible.

includes the practically important case with relatively large sojourns) there exists a value of the smoothing parameter ϑ for which the number of modes of the estimate's probability density function corresponds to the number of essentially distinct phases of the heterogeneous process.

The two elements of the method are the construction of the *peak density intervals*, and the conditional sampling or *censoring* of the observations.

3.1 Peak Density Intervals

3.1.1 Definition

Let $g(x)$ be a multimodal mixture of the unimodal probability density functions $g_k(x)$ on the real line and let ξ be one of the local maxima of $g(x)$. We call the open interval $\mathcal{I}$ *a peak density interval* if $\xi \in \mathcal{I}$, and for any $\mathcal{I}'$, such that $\xi \in \mathcal{I}' \subset \mathcal{I}$

$$\frac{1}{\mid \mathcal{I} \mid} \int_{\mathcal{I}} g(x)dx < \frac{1}{\mid \mathcal{I}' \mid} \int_{\mathcal{I}'} g(x)dx \tag{7}$$

It is easy to see that the maximum number of peak density intervals is equal to the number of modes and that any pair of intervals has an empty intersection.

If $\{\xi_1, \xi_2, \ldots, \xi_M\}$ is the set of local maxima of the probability density function $g(x)$ and
$\{\eta_1, \eta_2, \ldots, \eta_{M-1}\}$ is the set of local minima, such that $\eta_i \in (\xi_i, \xi_{i+1}),\ \ i = 1, \ldots, M-1$, then

$$\{\mathcal{I}_k \mid \mathcal{I}_1 = (-\infty, \eta_1);\ \mathcal{I}_M = (\eta_{M-1}, \infty);\ \mathcal{I}_k = (\eta_k, \eta_{k+1}),\ 1 < k < M\} \tag{8}$$

is the set of largest possible peak density intervals. In practice, however, it may be advantageous to choose smaller PDI.

The set of peak density intervals to be used for censoring may be found off-line by using the finite observation history, or obtained from the known characteristics of a heterogeneous process. For example, the user profile collected over previously completed calls can provide necessary information for the processing of the newly established connection.

3.1.2 Exponential Smoothing in Construction of PDI

The observation of the heterogeneous process up to epoch N $\{T_i\}_{i=1}^{N}$ leads to computation of the exponential smoothing estimator

$$U_n(\vartheta) = \vartheta \sum_{i=0}^{n-1} T_{n-i}(1-\vartheta)^i = \vartheta \sum_{i=1}^{n} T_i(1-\vartheta)^{n-i}. \tag{9}$$

and construction of the aggregate empirical distribution function

$$G_{N,\vartheta}(x) = \frac{1}{N} \sum_{i=1}^{N} \chi_{\{U_i(\vartheta) \leq x\}}, \tag{10}$$

which is a finite sample approximation of the true distribution

$$G_\vartheta(x) = \mathbf{Pr}\{U(\vartheta) \leq x\} = \sum_{k=1}^{M} \pi_k G_{k,\vartheta}(x) \tag{11}$$

If the sojourns are large enough then for sufficiently small ϑ, the distribution $G_{N,\vartheta}(x)$ is a mixture of M unimodal distributions. Our objective is to construct the set $\{\mathcal{I}_k,\ k = 1, ..., M\}$ of disjoint intervals around each of the modes, so that most of the mass of the conditional densities falls into the respective interval and the conditional probability $\mathbf{Pr}\{\Omega_n = i \mid U_n \in \mathcal{I}_k\}$ is maximized when $i = k$.

3.2 Censoring

Let $\{\Omega_n, T_n\}_{n=1}^{\infty}$ be a heterogeneous process. In this section we show how the peak density intervals of the exponential smoothing estimate can be used to specify the sampling criteria that identify the moments of the heterogeneous component distributions (6).

Assume that an appropriate value of the smoothing parameter ϑ has been chosen and the peak density intervals $\{\mathcal{I}_k,\ k = 1, ..., M\}$, based on the estimate $U_n(\vartheta)$, have been constructed. Although the PDI set satisfies the conditions set in section 3.1.2, the straightforward use of the indicator $\chi_{\{U_n(\vartheta) \in \mathcal{I}_k\}}$ as a decision function for sampling of T_n to obtain the empirical conditional distribution of output observations in state k $\hat{F}_k(x) = \Pr\{T \leq x | \Omega = k\}$ is not satisfactory. The reasons are twofold. First of all, the estimate $U_n(\vartheta)$ depends on the entire observation history and, therefore, the sampled value is not independent of the decision function. Secondly, the result depends critically on the choice of the PDI set. Errors of two types can occur while a decision is made whether the current hidden state is k, and therefore whether or not the current observation should belong to the k-th output sample. A type I error arises when at the n-th observation epoch the sampling rule incorrectly rejects the hypothesis $\Omega_n = k$; a type II error is associated with the state being estimated as k whereas the actual state of the system is different.

To reduce the effect of type I and type II errors, the sampling rule $D_k(T_n) = \chi_{\{U_n(\vartheta) \in \mathcal{I}_k\}}$ undergoes the following modification.

At each observation epoch n the heterogeneous process output T_n is observed and the new value of the exponential smoothing estimate U_n is computed. The k-th peak density interval $\mathcal{I}_k$ *is hit* if $U_n \in \mathcal{I}_k$; otherwise, a *miss* has occurred. Thereafter, the number of recent interval hits and misses is analyzed. If within the R most recent observation epochs U_n has been hitting some $\mathcal{I}_k$ often enough, the hypothesis $\Omega_n = k$ is accepted and the observed value T_n is included into the k-th output distribution sample. If on the other hand the number of misses is large, the decision is made to reject the hypothesis $\Omega_n = k$.

Definition 4 *The sampling procedure that selectively assigns the output observations of the heterogeneous process $\{\Omega_n, T_n\}_{n=1}^{\infty}$ to one of the M output distributions, or discards them, based on the criteria derived from the history of peak density intervals hits and misses is called* conditional sampling, *or* censoring.

3.2.1 Basic Sampling Algorithm

Two related sampling criteria formalize the rules suggested in the previous subsection. Let $\|U_n^m\|_k$ denote the number of times the estimate belongs to the peak density interval $\mathcal{I}_k$ between the observation epochs n and m, inclusively.

Consecutive events censoring The *consecutive event* criterion operates with two threshold parameters for each k, $k = 1, .., M$: the re-entry threshold r_k and temporary leave threshold l_k. If the previously estimated state is other than k, and r_k consecutive hits are observed, the current state is presumed to be k. If the previous estimated state has been k and the number of consecutive misses does not reach l_k, the state remains the same. When the number of misses is at least l_k, a transition is presumed to have occurred.

(i) [Entry rule] **if** $D_k(\mathcal{T}_{n-1}) = 0$ **and** $\left\| U_{n-r_k}^{n-1} \right\|_k = r_k$
then $D_k(\mathcal{T}_n) \leftarrow 1$;

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-1}) = 1$ **and** $\left\| U_{n-l_k}^{n-1} \right\|_k = \left\| U_{n-l_k-2}^{n-1} \right\|_k = 0$
then $D_k(\mathcal{T}_n) \leftarrow 0$.

To ensure consistency of the basic sampling criterion, its parameters might be subjected to the following requirement:

$$r_k > l_m, \quad \forall k, m;\ k \neq m.$$

Moving window censoring The second alternative is the *moving window* criterion with parameters w_k and a_k. The rule employed is: if the estimator value U_n belongs to the interval $\mathcal{I}_k$ for at least a_k out of w_k most recent arrival epochs, the state is presumed to be k; otherwise, the state is undecided.

(i) [Entry rule] **if** $D_k(\mathcal{T}_{n-1}) = 0$ **and** $\left\| U_{n-w_k}^{n-1} \right\|_k \geq a_k$ **then** $D_k(\mathcal{T}_n) \leftarrow 1$

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-1}) = 1$ **and** $\left\| U_{n-w_k}^{n-1} \right\|_k < a_k$ **then** $D_k(\mathcal{T}_n) \leftarrow 0$

The consistency condition is

$$a_k > \frac{w_k}{2}; \quad \forall k.$$

3.2.2 Properties of the Censoring Algorithm

Both versions of the algorithm described above use two parameters and are the special cases of the more general model with separate two-parameter windows for the entry and exit rules. Setting the required number of hits and misses to the size of the corresponding window defines the consecutive event censoring, whereas using the same parameters for entry and exit leads to the moving window approach.

To compare the performance qualities of both versions, set $r_k = a_k$ and $l_k = w_k - a_k$. The entry part of the moving window criterion is weaker, i.e. every observation sequence admitted by the consecutive event criterion is admitted by the moving window as well, whereas the converse is not true. An example is provided by the groups of $r_k - 1$ hits interleaved by a single miss. On the other hand, the moving window has a more stringent exit part, as for example, a group of $l_k - 1$ misses followed by a single hit is tolerated by the consecutive events criterion but triggers $D_k(\mathcal{T}_n) \leftarrow 0$.

Since by the model assumption, the output observation T_n is conditionally independent of the previous history $\mathcal{T}_{n-1}$, it is also independent of the decision rule, thus independence of the procedure is achieved. If the observation history is long enough, reducing the size of the sample does not play significant role and a type I error may be ignored to a certain extent. A type II error, on the other hand, which is due to the inertia of the estimator and the associated lag before the exit decision is made, is highly undesirable as it jams the sample distribution function and affects computation of the moments.

3.2.3 Modifications of the Basic Algorithm

Preemption The consistency requirements can be relaxed if the sampling rules are modified by adding the preemption clause: if at some arrival epoch n, the previous estimated state $D_j(\mathcal{T}_{n-1}) = 1$, $j \neq k$, and the entry condition for the k-th state is satisfied, then the transition from j to k is presumed to have occurred without meeting the exit condition for state j, the ties being broken by the peak density interval to which the current estimate belongs.

Retraction The major cause for a type II error is the inertia of the estimator, the smaller parameter ϑ, the higher. After the hidden process $\{\Omega_n\}$ makes a transition from state k, it has been sojourning in, to some other state, the value of the estimate U_n may stay in the peak density interval $\mathcal{I}_k$ for a few observation epochs longer. Besides it takes l_r or w_r additional epochs to detect the change of state. The modification of the basic algorithm, which allows adjustment of the output distribution sample retroactively at the epoch when the state transition is detected, is called *retraction.*

The exit rule for consecutive event censoring is modified:

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-1}) = 1$ **and** $\left\| U_{n-l_k}^{n-1} \right\|_k = 0$
then $D_k(\mathcal{T}_i) \leftarrow 0. \quad \forall i,\ n - l_k - \alpha \leq i \leq n.$

α is a parameter that defines the depth of the retraction.

In case of moving window censoring, the exit rule allows for the retroactive exclusion from the k-th distribution sample of the β most recent observations when fewer than a_k hits of the $\mathcal{I}_k$ interval are detected within the w_k most recent epochs.

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-1}) = 1$ **and** $\left\| U_{n-w_k}^{n-1} \right\|_k < a_k$
then $D_k(\mathcal{T}_i) \leftarrow 0 \quad \forall i,\ n - \beta \leq i \leq n.$

By retroactive exclusion of the observations that *precede* the epoch when the decision on the change of the system state is made, type II errors can be reduced substantially. However, the drawback lies in the fact that the sample independence condition no longer holds. The output observations that largely deviate from the expected level for the given phase are likely to cause frequent misses of the corresponding peak density interval and therefore to be excluded from the sample.

Two-sided criteria Instead of changing the decision retroactively, it is possible to delay its acceptance until the later observation epoch. The history of peak density intervals hits and misses is analyzed *before* as well as *after* the observation epoch in question. It is natural to apply the same conditions to both *ante factum* and *post factum* observations, requiring that both were met to accept the hypothesis $D_k(\mathcal{T}_n) \leftarrow 1$ and *at least one* was met in order to make the opposite transition. The rules for the consecutive event approach become

(i) [Entry rule] **if** $D_k(\mathcal{T}_{n-r_k-1}) = 0$ **and**
$\left(\left\| U_{n-2r_k}^{n-r_k-1} \right\|_k = r_k \text{ and } \left\| U_{n-r_k+1}^{n} \right\|_k = r_k \right)$
then $D_k(\mathcal{T}_{n-r_k}) \leftarrow 1;$

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-l_k-1}) = 1$ **and**
$\left(\left\| U_{n-2l_k}^{n-l_k-1} \right\|_k = 0 \text{ or } \left\| U_{n-l_k+1}^{n} \right\|_k = 0 \right)$
then $D_k(\mathcal{T}_{n-l_k}) \leftarrow 0.$

Whereas the moving window censoring is now defined by

(i) [Entry rule] **if** $D_k(\mathcal{T}_{n-w_k-1}) = 0$ **and**
$\left(\left\|U_{n-2w_k}^{n-w_k-1}\right\|_k \geq a_k \text{ and } \left\|U_{n-w_k+1}^{n}\right\|_k \geq a_k\right)$
then $D_k(\mathcal{T}_{n-w_k}) \leftarrow 1$

(ii) [Exit rule] **if** $D_k(\mathcal{T}_{n-w_k-1}) = 1$ **and**
$\left(\left\|U_{n-2w_k}^{n-w_k-1}\right\|_k < a_k \text{ and } \left\|U_{n-w_k+1}^{n}\right\|_k < a_k\right)$
then $D_k(\mathcal{T}_{n-w_k}) \leftarrow 0$

4 Simulation

The strengths (and weaknesses) of the proposed component identification algorithm are illustrated by the simulated example. The realization is based on Markovian heterogeneous process with four phases, where the output distributions belong to the two-parameter Weibull family with shape factor $\alpha = 2.0$ and variable scale factor β.

The value of the smoothing parameter ϑ was choosen from a discrete set with granularity 0.05. The choice of the basic censoring parameters (re-entry and temporary leave thresholds for the consecutive event approach as well as the window size and threshold for the moving window approach) in each case is stipulated by the value of the smoothing parameter and the relative size of the corresponding peak density interval.

The retraction depth parameters are the result of a compromise between the reduction of type II errors, on one hand, and the risk of suppression of the output distribution tails, on the other. In our simulation, the number of the observations excluded retroactively was set equal to the number of the most recent *consecutive misses* for either version of the algorithm. In the two-sided modification of censoring, the *ante factum* and *post factum* conditions were made symmetric.

The simulation results are given in Appendix A. Quantitatively, the following parameters are used for evaluation and comparison:
- N_k: the total length of sojourn in k–th state,
- $\hat{N}_k$: the number of observations identified as belonging to state k,
- K_I and K_{II}: type I and type II error coefficients defined, respectively, as the ratio of the number of correctly identified observations to the length of total sojourn in the corresponding state, and the ratio of the number of correctly identified observations to the total number of observations referred to that state,
- $\varepsilon_{\hat{\mu}_k}$: the relative error of the estimated component mean,
- KS, AD: the Kolmogorov-Smirnov and Anderson-Darling distances between the estimated component distribution function and its reference which are computed as

$$KS = \max_j \left[\hat{F}(j) - F(j)\right], \tag{12}$$

$$AD = \sum_j \left[\hat{F}(j) - F(j)\right]^2 \frac{\hat{f}(j)}{\hat{F}(j)(1-\hat{F}(j))}. \tag{13}$$

Here $\hat{f}(j)$ is an estimated probability mass function and $\hat{F}(j)$ - a distribution function.

The two blocks of data listed for each modification of the censoring algorithm correspond to consecutive event and moving window censoring.

Discussion

Simulation experiments with the proposed heterogeneous component identification algorithms support the following recommendations.

The basic censoring algorithm makes it possible to recover the general shape of the component output distribution, yet is prone to type II errors. Specifically, the observations that immediately follow the state transitions tend to be identified incorrectly, which result in heavier tails of the estimated distribution function and, consequently, in large errors in estimation of the mean output of the individual components. This problem can be overcome either by using the retro-active adjustment of the output distribution sample, or by making a delayed decision conditioned by both the foregoing and succeeding events. Although these two modifications clearly outperform the basic version of censoring, they achieve that at the expense of independence of the sample and therefore entail a risk of excessive tail suppression.

The consecutive event criterion is easier to implement, but it performs slightly inferior to the moving window criterion, especially if the component distribution functions possess large variance and heavy tails.

An infinite increase in the length of the observation history is not sufficient for the convergence of the estimates. The state sojourns must grow to infinity as well. Besides, if the sojourns possess finite moments, there is always a possibility that two phases with different expected output observation levels will be indistinguishable.

5 Conclusions

The paper has presented a new algorithm for non-parametric identification of the component distributions of heterogeneous processes. This model can be very useful in describing a wide variety of of processes arising in the control and management of high-speed networks, where the time-constraints make infeasible the traditional estimation techniques that involve extensive off-line computations.

The proposed algorithm make employs an exponential smoothing estimator as a fast on-line alternative to those computationally intensive techniques, and makes use of it as a basis for the conditional sampling, or censoring, of the observations. The advantages and disadvantages of the algorithm has been discussed and illustrated with simulated examples. To improve performance of the algorithm, a set of modifications has been proposed. The simulation shows that the proposed algorithm may present a viable technique for identification and estimation of the individual components of heterogeneous processes.

References

[1] H. Ahmadi, P. Kermani. Real time network load estimation in packet switched networks. *Proc. of IFIP,* pp. 367–380. Elsevier North-Holland, 1991.

[2] R. J. Elliot, L. Aggoun, J. B. Moore. Hidden Markov models. Estimation and control. Springer-Verlag, New York, NY, 1995.

[3] W. Feller. An introduction to the probability theory and its applications, II. 2nd ed. Wiley, New York, NY, 1971.

[4] P. Khedkar, S. Keshav. Fuzzy prediction of time series. *1992 IEEE International Conference on fuzzy systems.* San Diego, CA, 1992, pp. 281–288.

[5] D. A. Khotimsky, A. G. Konheim. Estimating the rates of heterogeneous ATM traffic using hidden Markov model. Extended abstract. *Proc. IEEE ATM'95.* Washington, DC, Oct.–Nov. 1995. pp. M3A-39–50.

[6] D. A. Khotimsky, A. G. Konheim. Estimating the rates of heterogeneous ATM traffic. Submitted to: *IEEE/ACM Transactions on Networking.*

[7] L. B. Klebanov, G. M. Maniya, I. A. Melamed. A problem of Zolotarev and analogs of ifinitely divisible and stable distributions in a scheme for summing a random number of random variables. *Probability Theory and Applications.* vol. 29 (1984), pp. 791–794.

[8] R. Warfield, S. Chan, A. Konheim, A. Guillaume. Real-Time Traffic Estimation in ATM Networks. *Proceedings of ITC-14.* Elsevier, Amsterdam, Netherlands, 1994. pp. 907–916.

[9] F. Yegenoglu, B. Jabbari. Maximum likelihood estimation of ATM traffic model parameters. *IEEE Globecom'94, Communication Theory Mini-Conference record.* San Francisco, CA, 1994, pp. 34–38.

Appendix. Simulation Results

Sample moments and parameters of the algorithm

k	N_k	μ_k	σ_k^2	$\mathcal{I}_k$	r_k	l_k	w_k	a_k
1:	135869	0.758	0.137	(0.600, 1.000)	8	4	12	8
2:	94115	1.292	0.423	(1.120, 1.660)	10	4	14	10
3:	115561	2.174	1.229	(1.800, 2.560)	14	2	16	14
4:	142226	3.235	2.776	(2.740, 3.800)	18	1	19	18

Censoring — basic version

k	$\hat{N}_k$	K_I	K_{II}	$\hat{\mu}_k$	$\varepsilon_{\hat{\mu}_k}$,%	$\hat{\sigma}_k^2$	KS	AD
1:	55285	0.336	0.826	1.099	44.956	1.374	0.111	0.149
2:	88034	0.488	0.521	1.405	8.686	1.333	0.074	0.036
3:	84685	0.468	0.639	2.179	0.230	2.000	0.069	0.015
4:	65369	0.348	0.758	2.940	9.118	3.062	0.093	0.025
1:	56310	0.342	0.826	1.098	44.889	1.369	0.111	0.149
2:	90835	0.503	0.522	1.408	8.937	1.336	0.073	0.036
3:	88145	0.485	0.635	2.178	0.205	2.005	0.070	0.015
4:	67152	0.358	0.759	2.941	9.094	3.063	0.093	0.025

Censoring — retraction

k	$\hat{N}_k$	K_I	K_{II}	$\hat{\mu}_k$	$\varepsilon_{\hat{\mu}_k}$,%	$\hat{\sigma}_k^2$	KS	AD
1:	44315	0.314	0.964	0.789	4.126	0.222	0.013	0.001
2:	64915	0.474	0.687	1.291	0.097	0.824	0.061	0.011
3:	69770	0.457	0.757	2.160	0.624	1.665	0.043	0.005
4:	58216	0.339	0.829	3.099	4.195	2.954	0.046	0.006
1:	38137	0.275	0.981	0.775	2.273	0.204	0.007	0.001
2:	49106	0.440	0.844	1.316	1.829	0.763	0.028	0.004
3:	49918	0.405	0.938	2.182	0.388	1.394	0.009	0.000
4:	37646	0.254	0.960	3.259	0.735	2.827	0.006	0.000

Censoring — two-sided conditions

k	$\hat{N}_k$	K_I	K_{II}	$\hat{\mu}_k$	$\varepsilon_{\hat{\mu}_k}$,%	$\hat{\sigma}_k^2$	KS	AD
1:	44932	0.315	0.954	0.799	5.473	0.242	0.019	0.003
2:	58507	0.466	0.750	1.340	3.696	0.859	0.032	0.007
3:	63493	0.440	0.800	2.159	0.706	1.625	0.040	0.004
4:	51374	0.303	0.839	3.099	4.196	2.918	0.045	0.005
1:	36936	0.267	0.982	0.777	2.547	0.217	0.007	0.001
2:	47649	0.434	0.857	1.320	2.157	0.774	0.026	0.003
3:	48816	0.398	0.941	2.185	0.502	1.405	0.009	0.000
4:	36691	0.248	0.962	3.259	0.756	2.837	0.006	0.000

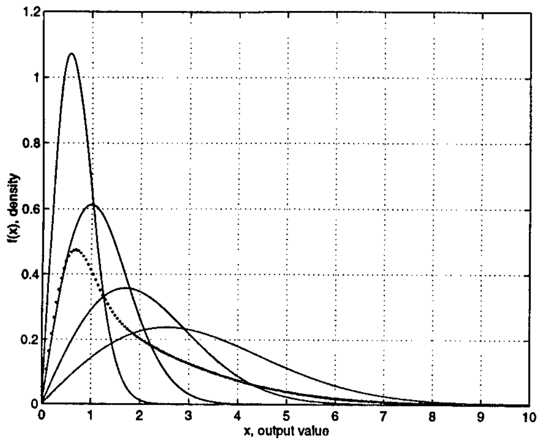


Fig. 1: True output densities of individual phases.

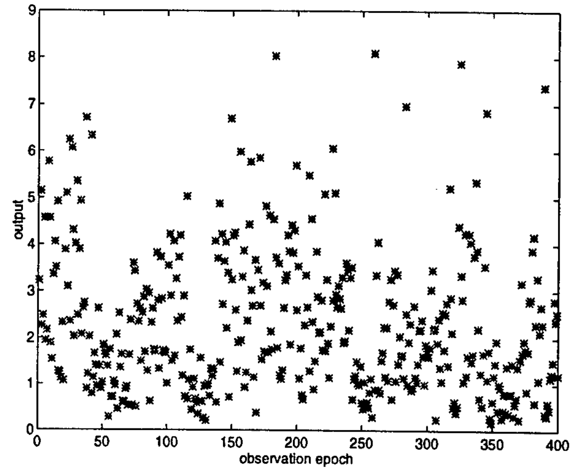


Fig. 2: Realization of the output process.

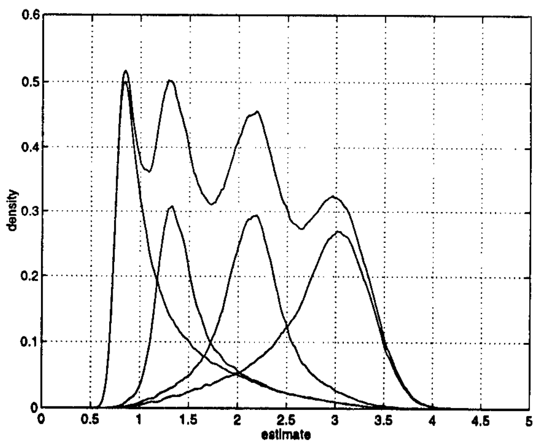


Fig. 3: Conditional densities of the estimate $\vartheta = 0.05$

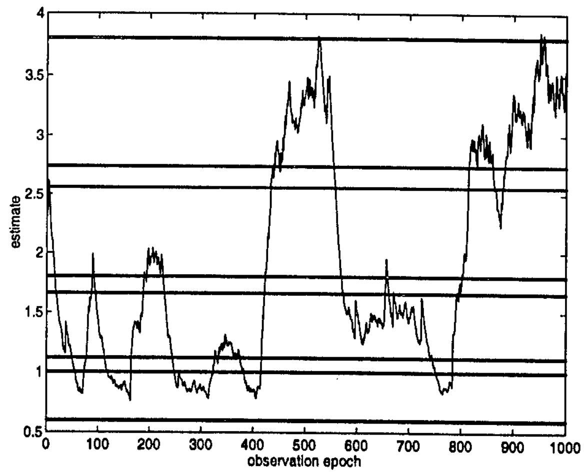


Fig. 4: Realization of the estimate and the PDI's.

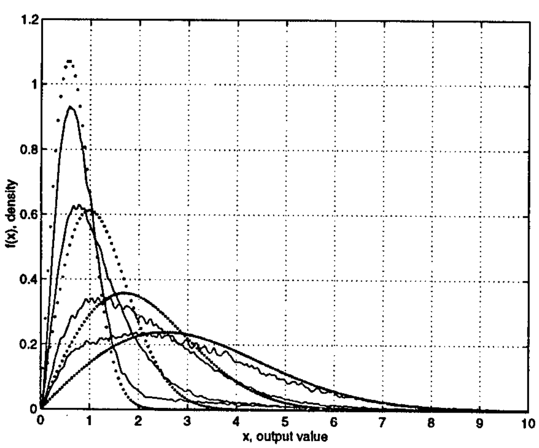


Fig. 5: Consecutive event censoring — basic version.

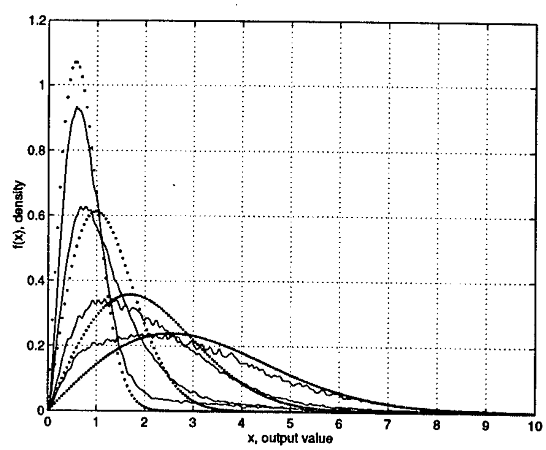


Fig. 6: Moving window censoring — basic version.

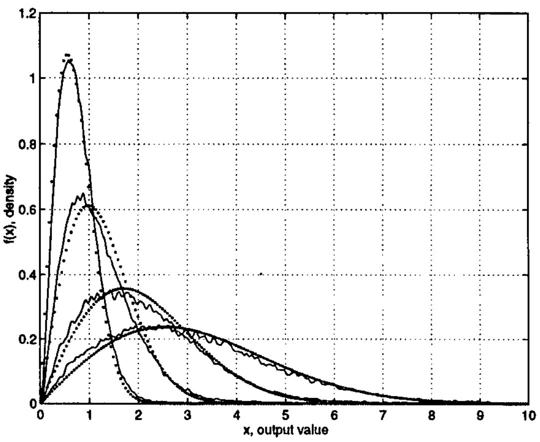


Fig. 7: Consecutive event censoring with retraction.

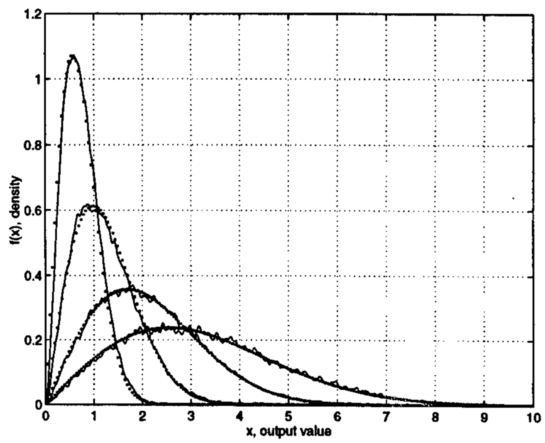


Fig. 8: Moving window censoring with retraction.

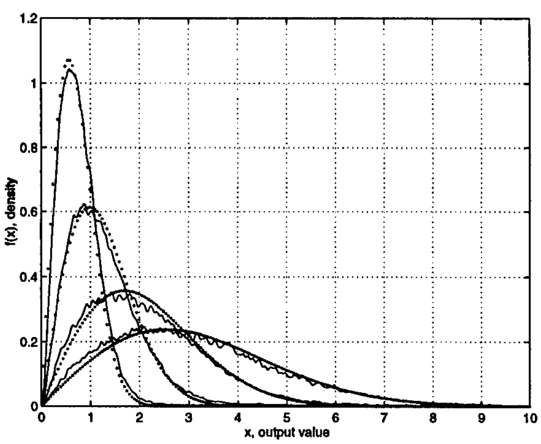


Fig. 9: Moving window with two-sided condition.

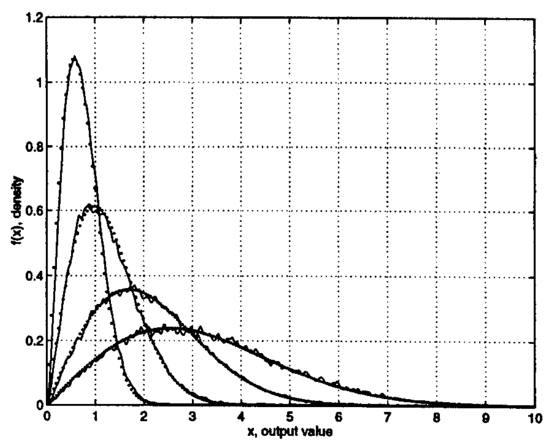


Fig. 10: Consecutive event with two-sided condition.

TEMPORAL CONSTRAINT LOGIC PROGRAMMING FOR ANALYSIS OF BEHAVIOUR OF SPACESHIPBORN EQUIPMENT BY TELEMETRY

Sergey Bogomolov and Alexander Yankovsky
Mozhajsky Military Space Engineering Academy
13, Zhdanovskaja Street , St.Petersburg, 197082 Russia
e-mail(via mediator): gor@iias.spb.su

Abstract. The aim of this paper is to present an approach to *temporal reasoning* in *telemetry* control system of *spaceshipborn equipment* based on conception of *constraint logic programming*. The temporal constraints of diagrams of functionig processes of spaseshipborn systems lie in the class of closed linear inequalities involving no more than two temporal variables. This restriction allows to propose effective procedure of computation of answer to temporal constraints query. The *temporal constraint logic program* is defined by set of the rules with temporal constraints given in the form of temporal constraints graphs. The temporal reasoning scheme combines the temporal constraints consistency procedure with cubic complexity and nondeterministic choice procedure. New flexible computation strategies are proposed on the base of accuracy analysis of temporal constraints of Horn formulae. The accuracy of temporal constraints depends on the length of intervals loads of edges. The size of search space is cut down much more large if there are more accurate temporal constraints in an initial query and rules from temporal constraint logic program.

Keywords. Temporal Reasoning, Constraint Logic Programming, Telemetry Processing, Spaceshipborn Equipment.

Introduction. An analysis of spaceshipborn equipment behaviour is based on the telemetry processing. Usually an expert-operator manipulates the assertions about some states of observed telemetry parameters. These assertions may be true in the some time intervals and be false in another time intervals. Therefore in terms of telemetry parameters states the functioning processes of the spaceshipborn systems may be presented as the temporal diagrams of the dynamic assertions. The effectiv reasoning system above assertions with time binding and temporal constraints allows to give new flexible tools for telemetry processing and analysis of spaceshipborn systems.

The conception of constraint logic programming (**CLP**) developed by J.Jaffar and J.-L.Lassez allows to connect the paradigms of constraints solving and logic programming [1][2]. The basis for this conception is following idea. The assertions of logic program are complemented by constraints. The last-mentioned ones are presented as equalities and inequalities over real, integer or boolean numbers. In this case, each computation step includes constraints consistency test and nondeterministic choice. For linear inequalities over real numbers, the consistency test is based on well-known in the operations research simplex-method with exponential complexity. This method have been used in the known **CLP(R)**-scheme [2].

The temporal constraints of diagrams of functionig processes of spaseshipborn systems lie in the class of closed linear inequalities involving no more than two temporal variables. This restriction allows to propose the effective computation procedure for the logic program with temporal constraints. The relative arrangement of temporal intervals corresponding to predicate atoms is seen as loaded oriented graph named **TC**-graph. The edges of **TC**-graph can be related to the elementary temporal constraints in the terms of closed linear inequalities over temporal intervals boundaries. The logic program with temporal constraints constitutes a set of pairs *"Horn formulae + TC - graph"*. Given paper is an extension of results of early investigations of authors [3] and[4].

Graphs of Temporal Constraints over Predicate Atoms. In this section we consider a model of temporal constraints and a consistency method of temporal constraints that will be used in temporal constraint logic programming. The predicate atoms are defined in usual fashion. The some temporal interval is associated with each predicate atom. The restriction on the relative arrangements of temporal intervals corresponding a finite set of predicate atoms can be presented in the form of closed linear inequalities system or in the form of temporal constraints graph. The temporal variables corresponding the temporal intervals boundaries are introduced explicitly for the case of inequalities systems. We use only binary or unary temporal relation between the intervals boundaries moments and therefore each inequality of the system involves no more than two temporal variables. The temporal variables must be inserted into predicate as terms to link the temporal intervals with corresponding predicate atoms. The second form is free from explicit temporal variables and may be expressed in terms of loaded oriented graph with two-type edges called next as *graph of temporal constraints (TC-graph)*. The **TC**-graph nodes are pairs **<begin,A>** and **<end,A>**, where **A** is an atom from a finite set of atoms. We use the double directed edges in the form

A
<begin,A>• ==========>• <end,A>
[U',L']

and the single directed edges in the form

<*bound1*, **A** >• ————————→ • <*bound2*, **A'**> .
[U,L]

The symbol parameters *bound1* and *bound2* are from a set {**begin,end**}. The symbol parameters *begin* and *end* serve for the begin and end boundaries indication correspondingly. The parameters **U'** and **L'** are non-negative real numbers ($0 \leq U' \leq L'$). The parameters **U** and **L** are real numbers ($U \leq L$). The double directed edge means that interval duration $T = t' - t$ corresponding to atom **A** must satisfy the condition $U \leq t' - t \leq L$, where **t** and **t'** are the begin and end boundaries time moments of time interval **T**. The single directed edge means that the time moments **t** and **t'** for the interval boundary *bound1* of an atom **A** and the interval boundary *bound2* of an atom **A'** must satisfy also the last-mentioned condition. The **TC**-graph is constructed by junction of one-named nodes.It is apparent that the **TC**-graph corresponds to some system of closed linear inequalities involving no more that two temporal variables in each inequality. Further we use only the graph form of temporal constraints over set of predicate atoms $A_1,..., A_m$ denoted by $G[A_1,..., A_m]$ or simply by **G**.

The **TC**-graph is named *temporal consistent* one when for any edge follow equality holds

$$[U,L] = \bigcap_k \left[\sum_i U_{ik}, \sum_j L_{ik} \right],$$

where **[U,L]** is non-empty interval load of edge, index **k** belongs to the index set of all kinds of paths between nodes of edge, **[** U_{ik},L_{ik} **]** is a interval load of some edge on the path indexed by **k**, $\bigcap$ is intersection operation of intervals. Notice that for given **TC**-graph it may be nonexistent the temporal consistent **TC**-graph. To say that the corresponding system of closed linear inequalities is inconsistent. Other wise it may be exist several temporal consistent **TC**-graphs to be partial ordered by insertion relation above corresponding interval loads of the edges. In the last case there exists the only greatest (in the sense of insertion relation) temporal consistent **TC**-graph in after referred to as *normal* one. To transform a given **TC**-graph to normal **TC**-graph the method of transitive closure may be used [3]. Also the adaptation of Floyd-Warshall's all-pairs-shortest-paths algorithm with cubic complexity is suited to this aim [6]. Notice that the related graphs named temporal constraints networks was proposed before by R.Dechter, I.Meiry and J.Pearl for temporal reasoning[6]. The innovation of our approach calls for the incorporation of temporal constraints graphs and temporal consistency procedure into constraints logic program and inference search scheme correspondingly [3] and [4].

Temporal Constraint Logic Programs. The *temporal constraint logic program* is defined by set of the rules with temporal constraints (named the *TC-rules*) in the form

$$A_0 \leftarrow A_1,\dots,A_m \parallel G[A_0,A_1,\dots,A_m],$$

where $A_0 \leftarrow A_1,\dots,A_m$ is a standard rule of classic logic programming (logical component), $G[A_0,A_1,\dots,A_m]$ is a temporal constraints system of the rule atom represented as **TC**-graph (or as set of **TC**-atoms). For **m=0** we have a special case of the **TC**-rules called the *TC-facts*. The temporal constraints system of **TC**-fact is used for time binding of temporal interval boundaries of predicate atom. In this case we assume that the **TC**-graph **G[A**$_0$**,Time(t), Time(t')]** includes the special point time atoms *Time(t)* and *Time(t')* for time binding specification. **t** and **t'** are some real numbers. The second type of temporal constraints system of **TC**-fact is used if we are in the dark about the time binding of predicate atom temporal interval. In last case we assume that the **TC**-graph **G[A**$_0$**]** is the only double directed edge labelled $[0,+\infty]$. Also we define the query with temporal constraints named *TC-query* and presented as pair

$$\leftarrow C_1,\dots,C_k \parallel G[C_1,\dots,C_k,B_1,\dots,B_n],$$

where $\leftarrow C_1,\dots,C_k$ is a typical query of classic logic programming (logical component of **TC**-query), $C_1,\dots,C_k$ are the predicate atoms named *active* ones and included in logical component, $B_1,\dots,B_n$ are the predicate atoms named *passive* ones and excluded in logical component, $G[C_1,\dots,C_k,B_1,\dots,B_n]$ is a temporal constraints system involving active and passive predicate atoms. If the **TC**-query is free from passive atoms, then one is said to be *initial*. If the **TC**-query is free from active atoms (logical component is empty query $\leftarrow$ **[ ]**), so one is named *final*. With the availability of passive and active atoms the **TC**-query is said to be *intermediate*. We suppose that the temporal constraints systems for any **TC**-rules, **TC**-facts and **TC**-query are the normal consistent ones.

Temporal Reasoning Scheme and Strategies of Inference Search. Let given the temporal constraint logic program **TP** involving finite set of **TC**-rules. Also let given initial **TC**-query (or simply query) Q_0 to program **TP**. The *successful computation* of answer to initial query Q_0 is a finite sequence of queries $Q_0,Q_1,...,Q_N$, so that query Q_N is final query and query Q_{i+1} is derived from query Q_i by using the computation step (inference rule) in the following form

$$\frac{\begin{array}{l}\leftarrow C_1,C_2,...,C_k \| G'[C_1,C_2,...,C_k,B_1,...,B_n] \\ \quad A_0 \leftarrow A_1,...,A_m \| G''[A_1,...,A_m]\end{array}}{\leftarrow \Theta\,(A_1,...,A_m,C_2,...,C_k) \| G[\Theta\,(A_1,...,A_m,C_2,...,C_k,B_1,...,B_n,C_1)]},$$

where Θ is unifier of predicate atoms A_0 and C_1, that is, $\Theta\, A_0 = \Theta\, C_1$, the temporal constraints system **G** of query Q_{i+1} is an union of **TC**-graphs **G'** and **G''** presented also in normal consistent form. We may use **TC**-rule if there exists unifier Θ for predicate atoms A_0 and C_1 and exists normal consistent temporal constraints system for union **G'** and **G''** of temporal constraints systems of current query and **TC**-rule. The unification procedure is used for construction of unifier [7]. For second condition test we use the modification of Floyd-Worshall's procedure [3] (see up).

Let $\leftarrow C_1,...,C_k \| G_0 [C_1,...,C_k]$ denotes the initial **TC**-query Q_0 to program **TP**. Let $\leftarrow [\ ] \| G_f [B_1,...,B_n]$ denotes the final **TC**-query Q_N of successful computation $Q_0,Q_1,...,Q_N$. Let Θ be a composition of unifiers $\Theta_1,...,\Theta_n$ obtained with successful computation. Then the answer to **TC**-query Q_0 is a pair $\{\Theta\, C_1,...,\Theta\, C_k\} \| G [\Theta\, C_1,...,\Theta\, C_k]$, where $\{\Theta\, C_1,..., \Theta\, C_k\}$ is the set of predicate atoms of query with desired constant terms, $G[\Theta\, C_1,..., \Theta\, C_k]$ is the final **TC**-graph $G_f [B_1,...,B_n]$ restricted by set of atoms $\Theta\, C_1,....,\Theta\, C_k$. Notice that there is $\{\Theta\, C_1,..., \Theta\, C_k\} \subseteq \{B_1,...,B_n\}$. And also for any edge of **TC**-graph G_0 with label $[\Delta_0, \Delta'_0]$ there exists the edge of **TC**-graph G_f with the same pair of nodes and label $[\Delta_f, \Delta'_f]$ that there is insertion relation $[\Delta_0, \Delta'_0] \subseteq [\Delta_f, \Delta'_f]$.

Above we have supposed that the only temporal interval is associated each predicate atom. To distinguish different appearances of temporal intervals that correspond fixed predicate atoms we use indexes. Predicate atom **A** with index **i** is denoted by $A^{(i)}$. The empty index may be omitted, so that by default **A** is $A^{()}$. Let $C^{(i)}$ be a selected predicate atoms of logical component of current **TC** - query Q_0 and let $A_0^{(i0)} \leftarrow A_1^{(i1)},...,A_m^{(im)}$ be a logical component of used **TC**-rule. With due regard for indexing the computation scheme is refined thus:

- the predicate atoms $A_1^{(i1)},...,A_m^{(im)}$ and $C_1^{(i)}$ of resulting **TC**-query **Q'** are replaced by the predicate atoms $A_1^{(i1.i)},...,A_m^{(im.i)}$ and $C_1^{(i.i0)}$ correspondingly,

- there are not any replacements of another predicate atoms.

The search of successful computation is performed by back-tracking method. But the use of temporal constraints consistency procedure with cubic complexity allows to cut off non-promise ways of search. The size of search space is cut down much more large if there are more accurate temporal constraints in initial **TC**-query and **TC**-rules from temporal constraint logic program. The accuracy of temporal constraints depends on length of intervals loads of **TC**-graph. The best (absolute) accuracy of temporal constraints system takes place when all interval loads are degenerated into point loads.

Then the system of closed linear inequlities corresponding to temporal constraints system transformed into system of equlities.The logic programs with temporal constraints permit to use new flexible computation strategies based on analysis of accuracy of temporal constraints systems of **TC**-rules and **TC**-query. The computation strategy defines an order of round of search tree and includes a selection step of active atom from current query and a suitable rule choice from admissible rule list. The simplest strategy is to select such atom $\mathbf{C_1}$ from query and such rule with head atom $\mathbf{A_0}$ that intersection of intervals corresponding to atoms $\mathbf{C_1}$ and $\mathbf{A_0}$ is minimal.For selection of suitable rule a more compound strategy may use the ordering of admissible rules set based on insertion relation of interval loads.

Analysis of behaviour of spaceshipborn equipment by telemetry date. An analysis of spaceshipborn equipment behaviour is based on the telemetry processing. An expert-operator manipulates the assertions about some states of observed telemetry parameters. These assertions may be true in the some time intervals and be false in another time intervals. Therefore in terms of telemetry parameters states the functioning processes of the spaceshipborn systems may be presented as the temporal diagrams of the dynamic assertions. The effectiv reasoning system above assertions with time binding and temporal constraints allows to give new flexible tools for telemetry processing and analysis of spaceshipborn systems.

The proposed temporal reasoning scheme is used in the prototype of a telemetry control system of spaceship-born equipment (SBE). Intelligent control system of SBE involves four levels of telemetry processing: the top level making the current control goals (by script or operator request), the second levels selecting the relevant knowledge and planning inference (by control knowledge) , the third level performing inference search, the low level making timed facts (by pre-processing collected original telemetry date). The processing of temporal knowledge takes place on the second and third levels. The knowledge base of expert analysis system is represented as set of TC-rules. The data base of expert system is described by the set of TC-facts with temporal constraints in time binding form.

The examples of a TC-rule and a TC-fact are given below

$$\mathbf{STATE(D1,ON) \leftarrow P1(CLOSED) \;||\; G1[STATE(D1,ON), P1(CLOSED]} \quad ,$$

$$\mathbf{P1(CLOSED) \leftarrow \;||\; G2\,[P1(CLOSED\,), 5,24]} \quad ,$$

where **G1** and **G2** are the TC- graphs presented in the Fig. 1 and Fig. 2. The example of the initial TC-query **Q0** is following expression

$$\mathbf{\leftarrow STATE(D1,ON), MODE(D2,x) \;||\; G0,}$$

where **G0** is the temporal constraints system of query represented as TC-graph in Fig. 3. The TC-query **Q0** describes a set of questions to the temporal constraint logic program :

a) Is the device **D1** turned on?
b) Which mode is there the device **D2** in ?
c) How long time is the device **D1** turned on ?
d) How long time is the device **D2** in unknown mode **x**?
e) What's the temporal arrangement of assertions about states of devices **D1** and **D2** with each other?

The questions (a) and (b) are referred to a logical component of TC-query. The questions (c), (d) and (e) are relevant to temporal constraints system of TC-query.

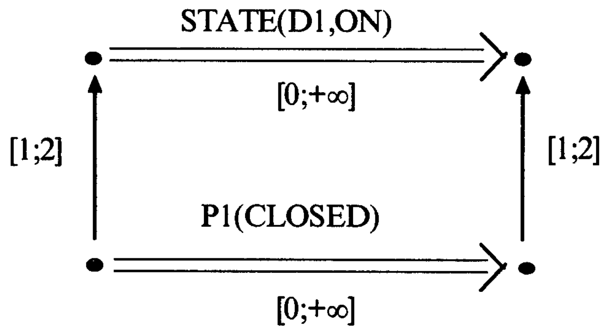


Fig. 1. TC-graph **G1**

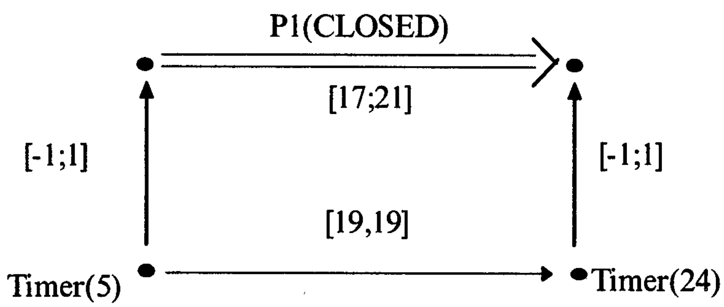


Fig. 2. TC-graph **G2**

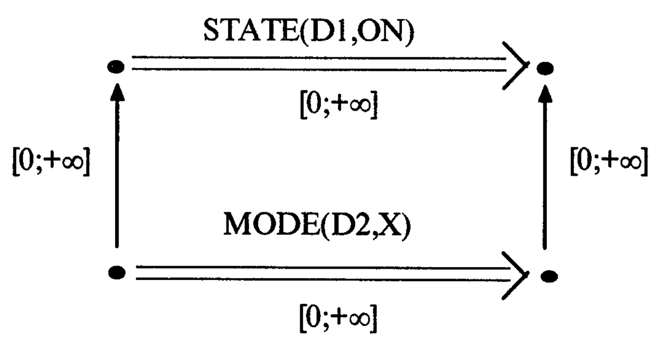


Fig. 3. TC-graph **G3**

Conclusion. We have presented an approach to temporal reasoning in telemetry control systems based on the conception of constraint logic programming. The time aspect was introduced into logic programs by constraints systems presented in the form of temporal constraints graphs over predicate atoms that correspond to system of closed linear inequalities involving no more than two temporal variables. The temporal reasoning scheme combines the temporal constraints consistency procedure based on modification of Floyd-Warshall's all-pairs-shortest-paths algorithm with cubic complexity and nondeterministic choice procedure based on **SLD**-resolution. The consistency procedure allows to cut off nonpromise ways of search. The related idea was proposed independently of authors by C.Brzoska in [5]. In that paper the formulae with temporal operators of modal logic are translated into formulae of first-order logic with inequalities between temporal arguments. But the types of temporal relations in our approach are the more general then ones in last-mentioned paper. Besides we use the graph's model of temporal constraints. It allows to construct new flexible computation strategies based on analysis of accuracy of temporal constraints systems of Horn formulae. The accuracy of temporal constraints depends on length of intervals loads of **TC**-graph. The size of search space is cut down much more large if there are more accurate temporal constraints in an initial query and rules from temporal constraint logic program. The best (absolute) accuracy of temporal constraints system takes place when all interval loads are degenerated into point loads. Then the system of closed linear inequlities corresponding to temporal constraints system transformed into system of equlities. The computation strategy defines an order of round of search tree and includes a selection step of active atom from current query and a suitable rule choice from admissible rule list.

References

1. Jaffar,J. and J.-L.Lassez (1987) Constrained logic programming, in: Proceedings 14 th ACM Symposium on Principles of Programming Languages, Munich, Germany, 111-119.
2. Jaffar,J. and S.Michaylov (1987) Methodology and implementation of CLP system, in: Proceedings of 4th International Conference on Logic Programming, Melbourn, Australia.
3. Bogomolov S.Ye. and A.G.Yankovsky (1993) Representation and processing of dynamic knowledge in expert systems , Technique Cybernetics, Bulletin of Russian Academy of Science. **5**, 17-23. (In Russian).
4. Bogomolov S.Ye.and A.G.Yankovsky (1995) Temporal constraint logic programming in modelling of distributed real-time systems, in: Proceedings of the First International Workshop on High Speed Networks of Open Distributed Platforms, St.-Petersburg, Russia, 195-198.
5. Brzoska C. (1991) Temporal logic programming and its relation to constraint logic programming. In Proc.of the 1991 Logic Programming Symposium, San Diego.
6. Dechter R., I.Meiry and J.Pearl (1991) Temporal constraint networks , Artificial Intelligence, **49(1)**, 61-95.
7. Lloyd J.W. (1984) Foundations of Logic Programming, Springer Verlag, Berlin.

COMBINED METHODS OF INCREASE OF INTERFERENCE DEFENCE IN SPACE CHANNELS OF INFORMATION TRANSMISSION

Alexander Kuzichkin, Ilya Gorelov, Victor Gribin, Igor Khvorov
Pushkin Higher Military College of radioelectronics of Air Defence
Department of Defence, Pushkin-9, St.-Petersburg, Russia

Abstract. In difficult signal-interference situation the solution of a problem of an interference defence in space channels of the information transmission is actual. The choice of a way of the solution depends on many factors and arises on the first development cycles. In this report such methods, as application of spread spectrum signals (SSS), adaptive spatial-time processing, rejection of concentrated interference are considered. It is shown, that no method separate is capable to meet the requirements of an interference defence in view of various restrictions on their application. The method of a combined signals processing (CSP) is offered and it integrates the aforesaid. The realization of elements and algorithms CSP on various element base is considered.

Keywords. Space channel of information transmission, adaptive algorithm, spread spectrum signals, combined signal processing, combined system of signal processing, rejection concentrated interference, spatial-time filter, direct correlation matrix inversion.

Introduction. Now as basic elements of space channels of the information transmission (SCIT) adaptive antenna systems (AAS) with the phased arrays with spatial-distribution sensors are widely used. They are used in a satellite, and in a ground received station. The problem of increase of interference defence especially is actual in application of above-stated systems in a difficult constantly varying signal-interference situation, whose characteristics are multi-regularity, ambiguity, and weak predictability of processes and situations, and also incompleteness and limited reliability of the *a priori* information about a structure, characteristics and methods of the interference application. The problem of choice of a way to provide of an interference defence arises on the first development cycles of elements SCIT, and the solution is provided with consideration of limitations in specific systems. The methods to provide the interference defence are subdivided into organizational, technical and organizational-technical, and technical methods, which as a rule, are basic in the solution of these problems and it are applied widely [1].

Basic methods of an interference suppression in the SCIT with the AAS are considered below.

1. Method of application of spread spectrum signals

This method is applied widely. At the realizable pass bands of radio electronic systems which participants in an information transmission, and transfer rates of the information from 100 up to 2000 bits per second the prize in an interference defence can achieve 50... 70 dB with use of spread spectrum signals (SSS).

SSS have received wide application as provision of:

--- High interference defence;

--- Compatibility of information transmission with measurement of parameters of motion;

--- Electromagnetic compatibility SSS with narrow-band systems of radio communication and broadcasting;

--- Best usage of frequency spectrum on limited territory as compared with narrow-band systems;

It characterize by high power reserve.

It is known, that an interference defence properties SSS are improved with increase of their base. In many practical cases it is required to apply SSS with size of base of the order 10^5 and more than [1]. However, with increase of signal base the time of their searching is also essentially increased which can be commensurable or even exceed necessary time for information transmission.

The application of digital processors for searching SSS is restrained by overall dimensions, high consumption of energy and cost, that especially is actual for a space element of SCIT. Recently there have been developed engineering solutions permitting on an acoustooptical element base to execute algebraic operations such as vector-matrix multiplication, including digital signals [2]. The potential response of such processors (up to 10^{15} oper/sec) allows to realize high-speed adaptive algorithms of signal processing and shows that application of acoustooptical and acoustoelectronic element base as processors for a satellite AAS has perspective, and it solves a problem of searching SSS.

Maximum accessible factor of interference suppression at use of a method with application SSS has the order 50... 70 dB [1], it is clear, this method separately taken will not solve the problem of defence SCIT from radio electronic suppression.

2. Methods adaptive spatial-temporary processing of signals.

These methods provide high effectiveness of suppression of interference which attitude different from arrangement of a useful signal and have the following important advantages:

--- They are easily combined with other methods of suppression;

--- Do not result in thickening electromagnetic situation;

--- Allow even without the *a priori* information to discover and to suppress deliberate interference;

The following can be considered as disadvantages:

--- Impossibility of suppression of number of interference greater, than number of degree of freedoms of the array ($N_{suppress} \leq N_{degree\ of\ freedom} - 1$);

--- Impossibility of suppression of interference coming from directions, which are close to a direction of arrival of a useful signal;

--- Long adaptation time of antenna array.

The above-stated singularities do not allow to use this method as a universal means of an increase of interference defence, though it allows to achieve suppression of interference up to 30... 35 dB.

3. Method of rejection of concentrated interference

This method provides application of specialized devices for suppression of certain interference and it has a series of virtues:

--- High degree of suppression of concentrated interference (30... 60 dB at narrow-band, 70... 80 dB at pulse interference);

--- Simplicity of a realization and a shared use with other means;

But also a lot of shortages is inherent in this method:

--- The suppression of wide-band interference and interference of large duration is impossible;

--- In a some cases the appearance of collateral effects, for example such as, origin of pulse interference is possible at suppression narrow-band etc.

4. Methods of combined signal processing.

The analysis of capabilities and tendencies of development of systems and means of radio electronic warfare [3] allows to draw a conclusion, that in situation of a mass use of means of radio electronic suppression no aforesaid method allows independently to ensure a required level of interference suppression by virtue of limitations. It, in turn, indicates necessity of a shared use of known methods of interference suppression with the purpose of a decrease of influence of the limiting factors on suppression quality and reaching of a maximum interference defence. Thus application of a method of combined signal processing (CSP) is most perspective.

The system which include an adaptive array, SSS processing unit, unit of rejection of narrow-band interference and unit of adaptive spatial-time processing of signals is considered. A high degree of suppression of narrow-band interference (in relation to frequency band of a useful signal $\Delta f_{narrow\text{-}band} << \Delta f_{useful\ signal}$) is provides by adaptive rejection filters. With wide-band interference ($\Delta f_{wide\text{-}band} = \Delta f_{useful\ signal}$) influence the effectiveness of adaptive rejection filters is low and other methods are required, for example, with spatial distinguishing. The adaptive spatial filters allow effective struggle with spatially delivered (with a useful signal) interference [4,5]. However, the adaptive spatial filters have a limited number of degree of freedoms, therefore it is expedient to apply these filters to suppression only of wide-band interference, but to suppress narrow-band interference adaptive rejection filters are necessary. In paper [6] the suppression both narrow-band, and wide-band interference in a combined system consisting of adaptive rejection filter and adaptive spatial filter is described. It is shown, that their joint application allows to arrange functions of interference suppression depending on their characteristics so, that the wide-band interference are suppressed in adaptive spatial-time filter, and narrow-band are suppressed in adaptive rejection filter. From a practical point of view the capability of economies of spatial channels of processing is important, as the number of them is determined only by quantity spatially-discernible wide-band interference.

Basic problems of application of a considered combined system of signal processing (CSSP) are such, that:

--- The joint application of composite units of CSSP results in necessity of consideration of influence of subsystems against each other, and, consequently, as a whole on effectiveness of all multilevel protective system from radio electronic suppression;

--- The high-power deliberate interference act to elements AAS, as a rule, and for its steady operation in situation of radio electronic suppression the use of the whole complex of known methods and means of interference defence is necessary, that considerably complicates a realization of CSSP.

--- At the presence of large number of interference, the time of adaptation of CSSP can exceed time necessary for processing, that results to a loss of useful signal .

The choice of qualitative algorithm of adaptation of CSSP in the given signal-interference situation for the solution of problems of practical application is necessary. Only a CSSP allows to realize a shared use of algorithms of the direct correlation matrix inversion (DMI) and gradient algorithms, since a part of a components can realize adaptation on algorithms of DMI, and other part can realize adaptation on gradient algorithms. These algorithms are most flexible, as a use of the common processor allows to redistribute computational capabilities for the benefit of components, where changes of signal-interference situation are most fast and essential, that allows to decrease number of operations. The application of acoustooptical and acoustoelectronic processors decreases a time of adaptation of CSSP.

Conclusions. Thus, considered in this report method of application of CSSP with combined algorithm of its adaptation combining merits of DMI-algorithms and a gradient algorithms allows to increase the interference defence of space channels of information transmission.

References

1. Кузичкин, А.В., С. Е. Кондаков, В. А. Григорьев (1994) Проблемы обеспечения помехоустойчивости систем передачи информации. Проблемные вопросы сбора, обработки и передачи информации в сложных радиотехнических системах : Сб. науч. тр., Выпуск 1. Санкт-Петербург. Часть 1, 113-120.
2. Acoustooptics: Researches and development (1990) Schoolseminar. Leningrad.
3. Крылов, В. В., К. Ю. Никашов (1988) Перспективы развития техники и технологии систем РЭБ. Зарубежная радиоэлектроника., 6, 3-12.
4. Монзинго, Р. А., Г. У. Миллер (1986) Адаптивные антенные решётки. Радио и связь, Москва.
5. Журавлёв, А. К., Е. И. Глушанков, А. П. Родимов и др. (1986) Адаптивные радиотехнические системы с антенными решётками. Изд-во ЛГУ, Ленинград.
6. Григорьев В.А. (1995) Многоканальная обработка сигналов на фоне комплекса помех. Известия ВУЗов. Приборостроение., 9, 32-35.

MODEL-FOLLOWING DECOUPLING CONTROL FOR 2 AXIS BEAM STEERING MIRRORS FOR SATELLITE-BASED LASER COMMUNICATION SYSTEMS

Victor A. Skormin and Timothy E. Busch
Watson School of Engineering, Electrical Engineering
Binghamton University, Binghamton, NY 13902, USA
Space Communication Branch, US Air Force Rome Laboratory
Rome NY 13441-5700, USA
e-mail: vskormin@binghamton.edu

Abstract. Accurate beam positioning is crucial for laser-based free-space communication systems, capable of very high bit-rate data transfer. Performance of beam steering systems is adversely affected by satellite jitter, insufficient bandwidth of a beam steering system and cross-coupling between its dynamic channels. Several model-following control techniques, alleviating these problems, are considered herein.

Key words. Satellite laser communication, decoupling control, model reference approach

Introduction. Space based optical communications offer several advantages over traditional radio frequency systems. They include: smaller beam divergence, smaller antennas, higher data rates, low probability of intercept, reduced electromagnetic interference, and low probability of jamming. Additionally, due to the much higher frequencies present in the optical spectrum, this technology has the potentials for light weight, small volume, and low power terminals. Since antenna size, beam divergence, and information bandwidth are all wavelength dependent, it is desirable to operate at as high a frequency as possible. These advantages make laser communications a consideration for such intersatellite communication programs as the Teledesic and Iridium. For the satellite designer, this technology allows for the minimization of the volume, weight, and power required by the terminal thus allowing for more freedom in other payloads or for launch costs reduction.

Interconnecting satellites by way of cross links can provide for a high speed, flexible communication network that is independent of ground station location. In this network, information would be transmitted up to a satellite from either ground-based or airborne terminals, routed by satellite nodes, and brought down anywhere in the world. Various architectures have been proposed containing configurations of satellites in various orbits. A low Earth orbit (LEO) constellation might call for tens or hundreds of satellites located in polar orbits at altitudes on the order of 800 Km. The longest link range associated with this geometry is about 15000 Km. These terminals will be required to acquire and track satellites in adjacent rings in a highly dynamic environment.

Geosynchronous Earth orbit (GEO) constellations might consist of a few satellites in an equatorial plane at altitudes of 40,000 Km. Their longest link range will be approximately 80,000 Km. Since these satellites are in fixed positions, relative to one another, their pointing acquisition and tracking systems don't face the same dynamic environment as those in the LEO orbit. However, because of the long intervening distance small disturbances caused by solar panel movements, spin stabilization or any small vibration can affect the success of these links.

Accurate beam positioning, summarized as the pointing, acquisition, and tracking (PAT) task, is crucial for laser-based free-space communication systems. However, this technology is dependent of our ability to reduce effects of satellite jitter, and assure high performance of mirror steering systems. There are several types of errors associated with PAT in laser communication systems. Pointing errors are associated with imprecise knowledge of the attitude and position of the transmitting terminal and the initial position of the receiving satellite. The errors associated with tracking are boresight misalignments, satellite jitter, and point-ahead errors. They are due to the relative motion of the two platforms and the propagation delay caused by the large intervening distance. These errors are often combined in an RMS fashion to provide an estimate of the systems PAT error [1], [2], [3].

Commonly a high bandwidth feedback control is used to stabilize the beam steering elements of a PAT system. The critical element to the use of this type of control is the beam steering element. Much work has been devoted to the development of both high bandwidth steering mirrors and other non-mechanical beam steering devices. Currently, these devices are available only as developmental models and consequently are very expensive. Applications of advanced control techniques, such as model reference (MR) control, applied to commercially available mirrors provides a means of achieving the required performance at a reasonable cost. While this approach improved the performance of a piezo-electric actuated mirror, it was hampered by strong cross channel coupling [4]. In fact it became necessary to implement an explicit decoupler. A method for achieving both decoupling and model following for a 2 input 2 output coupled system is presented herein.

Application of Model Reference Control. A typical beam steerer includes a high performance optical mirror, drives, position sensors, and feedback control circuitry [5]. The mirror has two control channels, elevation and azimuth. The beam position, defined by a quadrant detector, is controlled by the application of the appropriate reference signals. While it is desired that dynamic channels be controlled independently, the cross-coupling effects, which can be viewed as the "exchange of bending modes" could be always observed. Analyses of closed-loop dynamics of particular channels indicate nonlinearity and quite limited bandwidth of the appropriate frequency responses. In addition, satellite jitter constitutes a permanent source of beam positioning errors.

Application of MR control schemes has the potential for alleviation of the listed above problems [6]. Conceptually, it implies that a reference model representing the desired closed-loop system performance be established in the form of an analog circuit or a computer code. In a MR scheme, an additional control signal defined by a specially designed controller, "forces" the beam steerer to follow the response of the reference model. Understandably, a reference model not only has a perfect dynamic characteristics, but cannot be affected by satellite jitter and hardware deterioration. Therefore, in addition to improved dynamics, this approach results in jitter rejection and increased system robustness. Limitations of the MR approach are related to the limited magnitude of control effort which can be applied to a beam steerer, and the complexity of its implementation.
Conventional MR schemes, capable of decoupling, extended system bandwidth, reduction of jitter effects, and high system robustness, require state-variable implementation and are overly complex (numerically intensive) for many applications. In many instances, when some of the above listed effects are not required, particular simplifications of a "general" MR control can be suggested.

Described below, is a model-following, rather than MR, control scheme utilizes so-called System Error Equation Method. This method results in performance modification and jitter rejection, but not in robust performance. Such a system can be recommended when system dynamics is fairly well defined and assumed to be time-invariant. The system utilizes readily available input-output signals.

The second system, described herein, utilizes an adaptive predictor and is intended for digital implementation. It exhibits all features of a MR scheme, but does not require a state observer.

System Error Equation Method of Model-Following. Consider closed-loop dynamics of a mechanical beam steerer described in the form,

$$Y(s)=G(s)U_P(s) \quad (1)$$

the reference model is defined as follows,

$$Y_M(s)=M(s)U(s) \quad (2)$$

the output error is defined as,

$$E(s)=Y_M(s)-Y(s) \quad (3)$$

and the controller implementing the MR concept is,

$$U_P(s)=U(s)+W(s)E(s) \quad (4)$$

where

$G(s)=[g_{ij}(s), i,j=1,2]$ and $M(s)=[m_{ij}(s), i,j=1,2]$ are appropriate transfer matrices; it is expected that G(s) is non-diagonal, and M(s) is a diagonal matrix,

$Y=[y_1\ y_2]'$, and $Y_M=[y_{M1}\ y_{M2}]'$ are elevation and azimuth position components of the existing beam steerer and the reference model,

$U_P=[u_{P1}\ u_{P2}]'$, and $U=[u_1\ u_2]'$ are control signals applied to the existing beam steerer and to the reference model (also known as the commanded input), and

$W(s)=[w_{ij}(s), i,j=1,2]$ is the MR controller.

Rewrite equation (1) as follows,

$$Y(s)=G(s)U(s)+G(s)W(s)E(s)=G(s)U(s)+G(s)W(s)[Y_M(s)-Y_P(s)]=$$
$$G(s)U(s)+G(s)W(s)Y_M(s)-G(s)W(s)Y_P(s),$$

then

$$Y(s)+G(s)W(s)Y_P(s)=G(s)U(s)+G(s)W(s)Y_M(s), \text{ or}$$
$$[I+G(s)W(s)]Y(s)=[G(s)+G(s)W(s)M(s)]U(s)$$

Finally,

$$Y(s)=[I+G(s)W(s)]^{-1}G(s)[I+W(s)M(s)]U(s) \quad (5)$$

Redefine the output error as

$$E(s)=\{M(s) - [I+G(s)W(s)]^{-1}G(s)[I+W(s)M(s)]\}U(s) \quad (6)$$

Define an arbitrary transfer matrix

$$Q(s)=[q_{ij}(s), i,j=1,2] \quad (7)$$

representing the desired dynamics of the "commanded input - error" channels. The implication of the word "desired" are,

- transfer matrix Q(s) must be stable,
- settling time not to exceed 1/2 of the settling time of the reference model, and
- sufficiently small steady-state values of the error response, which can be assured by ||Q(0)||N 1

Note, that there is no need to achieve the error decoupling.

Then the following equation should be established,

$$M(s) - [I+G(s)W(s)]^{-1}G(s)[I+W(s)M(s)]=Q(s) \quad (8)$$

which results in the definition of the controller W(s),

$$W(s)=G(s)^{-1}M(s)Q(s)^{-1} - G(s)^{-1} - Q(s)^{-1} \quad (9)$$

It should be noted that the major advantage of this approach is in a very high flexibility in the particular choice of error equation matrix Q(s). Indeed, upon exploration of equation (9) allows for the choice of such Q(s) which would result in the sufficiently simple definition of the controller W(s). The disadvantage of this approach is quite typical of any linear model-following design: it implies that matrix G(s) is accurately known and time-invariant. However, according to our preliminary results, the developed model-following scheme can be successfully extended to a MR scheme utilizing input-output signals.

The following numerical example is based on a simplified mathematical description of a beam steerer. Particular elements of the matrix G(s) are,

$$g_{11}(s)=2.6186E8/(s^2+444s+2.1822E8)$$
$$g_{12}(s)=2.0731E8/(s^2+444s+2.1822E8)$$
$$g_{21}(s)=1.E8[1.005s^2+428.39s+1.3735E8]/[(s^2+471s+1.5824E8)(s^2+400s+1.24E8)]$$
$$g_{22}(s)=-1.E8[1.067s^2+463.16s+1.4985E8]/[(s^2+471s+1.5824E8)(s^2+400s+1.24E8)] \quad (10)$$

The reference model transfer matrix is defined as,

$$m_{11}(s)=m_{22}(s)=3000/(s+3000)$$
$$m_{12}(s)=m_{21}(s)=0$$

The following is the definition of the "error transfer matrix" Q(s),

$$q_{11}(s)=q_{22}(s)=.001*s/(s+6000)$$
$$q_{12}(s)=q_{21}(s)=0$$

leading to the definition of controller W(s),

$$w_{11}(s)=-1.E11w_0(s)[2.6s^4+1.925E4s^3+1.87E8s^2+1.23E12+1.49E15]$$
$$w_{12}(s)=37.1w_0(s)[(s^2+426s+1.37E8)(s^2+444s+2.18E8)(s^2+400s+1.14E8)]$$
$$w_{21}(s)=76.5w_0(s)[(s^2+471s+1.58E8)(s^2+400s+1.14E8)(s^2+400s+1.24E8)]$$
$$w_{22}(s)=-1.E11w_0(s)[3.24s^4+3.29E4s^3+2.72E8s^2+1.86E12+5.09E15]$$

where

$$w_0(s)= [s(s+3000)(s^4+831s^3+2.56E8s^2+1.059E11s+1.62E16]^{-1}$$

Computer simulation results indicated that the obtained controller, even after reasonable simplification, was capable of "enforcing" the required system dynamics on the beam steerer and reduced the effect of satellite jitter on the beam positioning accuracy.

MR Control with an Adaptive Predictor. Consider a beam steerer described by a transfer matrix G(s) as in (1). Assume that elements of the same row of matrix G(s) are brought to a common denominator. Convert matrix G(s) to its discrete-time equivalent, G(z), assuming that the steerer is driven through a zero-order-hold. Then the following difference equations could be used for the description of the beam steerer,

$$y_1(i)=-\Sigma a_{1j}y_1(i-j) + \Sigma b_{1j}u_{P1}(i-j) + \Sigma c_{1j}u_{P2}(i-j)$$
$$y_2(i)=-\Sigma a_{2j}y_1(i-j) + \Sigma b_{2j}u_{P1}(i-j) + \Sigma c_{2j}u_{P2}(i-j), \quad (11)$$

where i is the discrete-time index. One can see that summation in (11) is performed with respect to time index j=1,2,3,... , and coefficients a_{1j}, b_{1j}, c_{1j}, a_{2j}, b_{2j}, and c_{2j}, are the parameters of the z-domain transfer functions $g_{11}(z)$, $g_{12}(z)$, $g_{21}(z)$, and $g_{22}(z)$, and as such are dependent on the time discretization step.

Note that since y_k, k=1,2 are azimuth and elevation positions of the laser beam accurately defined by the quadrant detector, and u_{Pk}, k=1,2, are known control signals, parameters a_{1j}, b_{1j}, c_{1j}, a_{2j}, b_{2j}, and c_{2j}, can be continuously updated by the application of a finite-memory recursive least squares procedure [6]. Although the success of the parameter updating can be assured by good initial values of these parameters and low measurement noise, it is known that when required, the procedure can be enhanced by orthogonalization of variables $y_1(i\text{-}j)$, $y_2(i\text{-}j)$, $u_{P1}(i\text{-}j)$, and $u_{P2}(i\text{-}j)$ and consequent modification of the obtained estimates. Implementation of the parameter updating procedure is instrumental in situations when parameters of the controlled process are poorly known and/or time-dependent due to hardware deterioration and environmental effects.

Availability of accurately estimated parameters of the equations (11) facilitates the use of these equations for the prediction of azimuth and elevation positions. Then continuous parameter updating provides the assurance of the prediction accuracy. Indeed, if (11) holds than the following "one step ahead" prediction can be achieved as

$$y_1(i+1)=\Sigma a_{1j}y_1(i\text{-}j+1) + \Sigma b_{1j}u_{P1}(i\text{-}j+1) + \Sigma c_{1j}u_{P2}(i\text{-}j+1)$$
$$y_2(i+1)=-\Sigma a_{2j}y_1(i\text{-}j+1) + \Sigma b_{2j}u_{P1}(i\text{-}j+1) + \Sigma c_{2j}u_{P2}(i\text{-}j+1) \qquad (12)$$

Introduce a reference model in the form of two fixed-parameter difference equations, reflecting the desired closed-loop dynamics of the beam steerer,

$$y_{M1}(i+1)=-\Sigma m_{1j}y_{M1}(i\text{-}j) + \Sigma n_{1j}u_1(i\text{-}j)$$
$$y_{M2}(i+1)=-\Sigma m_{2j}y_{M1}(i\text{-}j) + \Sigma n_{2j}u_1(i\text{-}j) \qquad (13)$$

One can see that equations (13) represent a decoupled system. Computation of the predicted azimuth and elevation, $y_1(i+1)$ and $y_2(i+1)$, and their reference values $y_{M1}(i+1)$, $y_{M2}(i+1)$ facilitates computation of the expected position errors,

$$e_1(i+1)=y_{M1}(i+1)-y_1(i+1) \qquad (14)$$
$$e_2(i+1)=y_{M2}(i+1)-y_2(i+1)$$

Define increments of control signals, $u_{p1}(i)$ and $u_{p2}(i)$ to assure that position errors have zero values,

$$b_{11} \quad u_{P1}(i) + c_{11} \quad u_{P2}(i) = -e_1(i+1)$$
$$b_{21} \quad u_{P1}(i) + c_{21} \quad u_{P2}(i) = -e_2(i+1) \qquad (15)$$

Then control signals $u_{P1}(i+1)$ and $u_{P2}(i+1)$ can be defined as follows,

$$u_{P1}(i+1)=u_{P1}(i+1)+ \quad u_{P1}(i)$$
$$u_{P2}(i+1)=u_{P2}(i+1)+ \quad u_{P2}(i) \qquad (16)$$

The proposed MR control procedure performing parameter updating, solution of equations (15), and definition of the control signals as per (16) at each step of the discrete-time procedure, i=1,2,3,..., has been implemented in software and tested by simulation. It is important that the proposed decoupling MR technique can be extended from two to any practical number of input-output channels.

The following are difference equations representing the beam steerer, originally defined by transfer functions (10), at sampling frequency 32.5 KHz,

$$y_1(i)=1.7857y_1(i\text{-}1)\text{-}.9865y_1(i\text{-}2)+.1207[u_{P1}(i\text{-}1)+.9954u_{P1}(i\text{-}2)]+.0956[u_{P2}(i\text{-}1)+.9954u_{P2}(i\text{-}2)]$$

$$y_2(i)=3.7122y_2(i-1)-5.41815y_2(i-2)+3.6628y_2(i-3)-.9736y_2(i-4)+ .0466[u_{P1}(i-1)-.8649u_{P1}(i-2)-.8649u_{P1}(i-3)+.9825u_{P1}(i-4)] - .0495[u_{P2}(i-1)-.8612u_{P2}(i-2)-.8616u_{P2}(i-3)+.9823u_{P2}(i-4)]] \quad (17)$$

The reference model is represented by expressions

$$y_{M1}(i)=.912y_{M1}(i-1) + .08799u_1(i-1)$$
$$y_{M2}(i)=.912y_{M2}(i-1) + .08799u_2(i-1) \quad (18)$$

Computer simulation results indicate that recursive least squares procedure allows for fast and reliable convergence of the parameter updating process, both in the estimation and tracking modes. The control procedure, especially upon convergence of the parameter updating process, allows for successful model following. Utilization of the continuous parameter updating provides high degree of robustness of the resultant control system.

References

1. Barry, J.D., Mecherle, G.S. (1985) Beam pointing error as a significant design parameter for satellite-borne, free space optical communication systems. Optical Engineering, Vol. 24(6), 1049-1054
2. Held,K.L., Barry, J.D. (1988) Precision pointing and tracking between satellite-borne optical systems. Optical Engineering, Vol. 27(4), 325-333
3. Chen, C.C., Gardner, C.S. (1989) Impact of random pointing and tracking errors on the Design of Coherent and Incoherent Optical Inersatellite Communication Links. IEEE Transactions on Communications, Vol. 37(3), 252-260
4. Busch, T.E., Skormin, V.A. (1997) Experimental implementation of model reference control for fine tracking mirrors. Free-Space Laser Communication Technologies IX, G. Stephen Mecherle Ed., Proc. SPIE Vol. 2990
5. Dumas, R., Laurent, B. (1990) System test bed for demonstration of the optical space communications feasibility. SPIE Vol. 1218, Free-Space Laser Communication Technologies II, 398-411.
6. Astrom, K.J., and Wittenmark, B., (1990) Adaptive Control, Addison-Wesley Publishing Company

MODELLING OF PULSE-WIDTH-MODULATION SYSTEMS BASED ON THE CONVERSION OF THE MODULATION TYPE

V. V. Anokhin
State Academy of Aerospace Instrumentation
67, Bolshaya Morskaya
St.-Petersburg, Russia

Abstract. Modelling of radioelectronic objects, operating with the pulse-width modulation is examined. It is shown that they can be represented as the amplitude-pulse systems, if switching frequency is more greater than cut-off frequency of the output linear analog part. This approach significantly increases response speed of models and effectiveness of the objects dynamic properties investigation.

Keywords. PWM — pulse-width modulation; APM — amplitude-pulse modulation; IR — impulse response; RAP —reduced analog part; LAP — linear analog part; ARMA — autoregression with moving average.

Introduction. PWM is extensively used in modern radioelectronics for in this case energetic parameters, mass and weight efficiency improve significantly. There are modulators, pulse amplifiers, automatic switching regulators, switched-mode power supplies in radioelectronics, where PWM is used.

Non-linear character of the conversion concerned with PWM essentially complicates both physical interpretation of the electrical processes interaction, which takes place in the researching devices, and their mathematical description. The models of non-linear pulse systems are often characterized by too high complexity and low response speed, or they have such assumptions, which limit their applications. Therefore designing of such objects has partly empirical character, which results in of the dynamic and stability parameters getting worse.

Description. In this paper the method of synthesis of the discreet mathematical models of PWM-objects, having the response speed and accuracy, compared with the analogous characteristics for linear systems is derived. Such approach is correct for the objects under condition

$$f_T = 1 / T >> f_{max} \tag{1}$$

where f_T — switching frequency, T — switching period of pulse element, f_{max} — cut-off frequency of the device output LAP.

If the condition (1) is correct, synthesis idea consists of changing PWM by APM. Source discreet PWM-model is shown in Fig. 1

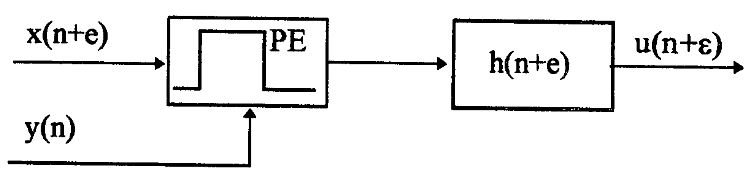


Fig. 1.

This model is a non-linear pulse automatic regulator, where pulse element (PE) generates square pulses with the frequency f_T. The amplitude x(n) and the duration $\theta(n)T$ depend on the input voltage $x(n+\varepsilon)$ and the control signal y(n) respectively.

The output signal is the convolution of the IR of the LAP and the input signal [1]:

$$u(n+\varepsilon)=\begin{cases} T\left[\sum_{m=0}^{n-1} x(m)\int_{0}^{\theta(m)} h(n-m+\varepsilon-\tau)d\tau + x(n)\int_{0}^{\varepsilon} h(\varepsilon-\tau)d\tau\right], \\ \qquad 0\le\varepsilon\le\theta_N, \\ T\sum_{m=0}^{n} x(m)\int_{0}^{\theta(m)} h(n-m+\varepsilon-\tau)d\tau, \quad \theta(n)\le\varepsilon\le 1. \end{cases} \tag{2}$$

Equations (2) may be represented as

$$u(n+\varepsilon)=\begin{cases} \sum_{m=0}^{n-1} x(m)h_p[n-m+\varepsilon,y(m)] + x(n)h_p[\varepsilon,y(n)], \\ \qquad 0\le\varepsilon\le\theta(n), \\ \sum_{m=0}^{n} x(m)h_p[n-m+\varepsilon,y(m)], \quad \theta(n)\le\varepsilon\le 1, \end{cases} \tag{3}$$

where

$$h_p[n-m+\varepsilon,y(m)] = T\int_{0}^{\theta(m)} h(n-m+\varepsilon-\tau)d\tau \tag{4}$$

— non-linear IR of the RAP of the object [2]. Variable upper limit in (4) leads to the low speed response of model (3). Therefore, it is a good idea to avoid the dependence $\theta(n)$.

The IR (4) may be represented in terms of poles s_k and the residues c_k' of the transfer function of the LAP:

$$h(n-m+\varepsilon-\tau)=\sum_{k=1}^{2} c_k' e^{s_k T(n-m+\varepsilon-\tau)}. \tag{5}$$

Substitution (5) into (4) results:

$$h_p(n-m+\varepsilon,y(m)] = T\sum_{k=1}^{p} c_k' e^{s_k T(n-m+\varepsilon)}(1-e^{-s_k T\theta_N}e^{-s_k Ty(m)}]/(s_k T) \tag{6}$$

If (1) is correct, then $|s_k Ty(m)| << 1$ and therefore

$$e^{-s_k Ty(m)} \approx 1 - s_k Ty(m).$$

Then (6) is represented as

$$\begin{aligned} h_p[n-m+\varepsilon,y(m)] &= \sum_{k=1}^{p} c_k e^{s_k T(n-m+\varepsilon)}(1-e^{-s_k T\theta_N}) + y(m)\sum_{k=1}^{p} Tc_k' \times \\ &\times e^{s_k T(n-m+\varepsilon)}e^{-s_k T\theta_N} = h_r(n-m+\varepsilon)+y(m)h_m(n-m+\varepsilon), \end{aligned} \tag{7}$$

where

$$h_r(n-m+\varepsilon)=\sum_{k=1}^{p} c_k' e^{s_k T(n-m+\varepsilon)}(1-e^{-s_k T\theta_N})$$

— the IR of the RAP with pulse element, generating square pulses at the beginning of each switching interval with the duration $\theta_N T$,

$$h_m(n-m+\varepsilon)=\sum_{k=1}^{p} Tc_k' e^{s_k T(n-m+\varepsilon)} e^{-s_k T\theta_N}$$

— the IR of the RAP with pulse element, generating δ-pulses with constant delay $\theta_N T$ in each switching interval.

According to (7) PWM-model shown in Fig. 1 may by presented as APM-model (see Fig. 2).

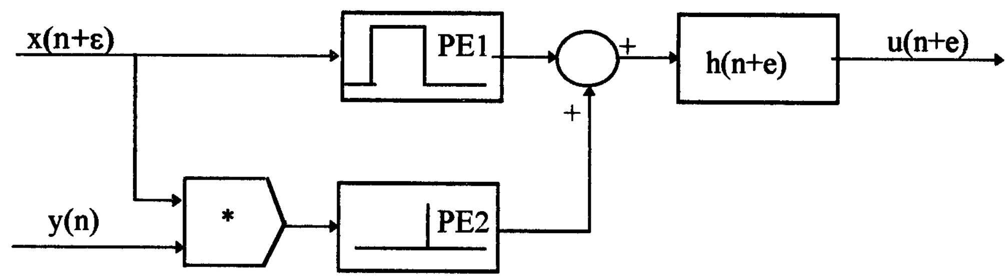


Fig. 2.

The IR of the RAP (7) characterizes APM-system, input of which is connected to the parallel pulse elements PE1 and PE2. Combination (3) and (7) results

$$u(n+\varepsilon)=\begin{cases}\sum_{m=0}^{n-1} x(m)h_r(n-m+\varepsilon)+x(n)h_r(\varepsilon)+ \\ +\sum_{m=0}^{n-1} x(m)y(m)h_m(n-m+\varepsilon)+x(n)y(n)h_m(\varepsilon), & 0\le\varepsilon\le\theta(n), \\ \\ \sum_{m=0}^{n} x(m)h_r(n-m+\varepsilon)+\sum_{m=0}^{n} x(m)y(m)h_m(n-m+\varepsilon), & \theta(n)\le\varepsilon\le 1.\end{cases} \tag{8}$$

It is the final discreet APM-model of PWM-object, which may by represented in recurrent form:

$$u(n)=\sum_{i=1}^{q} b_i x(n-i)-\sum_{j=1}^{p} a_j u(n-j)+\sum_{m=1}^{r} d_m x(n-m)u(n-m)$$

with constant coefficients $\{b_i\}$, $\{a_j\}$ and $\{d_m\}$ [3].

Conclusions. The conversion of the modulation type leads to removing of the PWM-non-linearity, but in this case it results in the non-linearity as a product of an input and a control signals. In spite of this the final model (8) has two important advantages: a) for small signal mode it may be simply linearized; 2) it may be represented as ARMA-model with additive non-linear part and constant coefficients. The high speed response and accuracy of the model (8) make it very useful for investigation of transient responses and identification of devices parameters.

References

1. Y. Z. Tsypkin, Y. S. Popkov (1973) The theory of non-linear pulse systems. Nauka, Moscow (In Russian).
2. Y. Z. Tsypkin (1963) The theory of linear pulse systems. Phismatgiz, Moscow (In Russian).
3. V. V. Anokhin (1993) Modelling of transient processes in switching-mode power supplies. International symposium on electromagnetic compatibility. St-Petersburg, Part 2, pp. 409—413 (In Russian).

A REAL TIME COMPUTING SYSTEM FOR THE MEASURING SIGNAL CONTROL

Vladimir Kryukov and Karen Shachgeldyan
Vladivostok State University of Economics
Department of Computer Engineering
41, Gogolya Street, Vladivostok, Russia
e-mail: kryukov@vvsu.ru

Abstract. Questions of software developing for a dual-processor computing system intended for digital signal processing are considered. It is shown that both fast digital processing algorithms and organization of concurrent operation of devices are necessary for creation of effective measuring real-time systems on the base of multiprocessor computing platform. It is confirmed that the use of high level languages in real-time systems allows to simplify creation of the applied software and raise its operational reliability. Testes of the developed software of spectral estimation and signal filtering on the base of the dual-processor computer system with Intel 80x86 and Texas Instruments' TMS320C30 are described.

Keywords. Real-time system, multiprocessor, data acquisition, fast Fourier transform (FFT), signal filtering, spectrum estimation, digital signal processing (DSP).

Real-time system. The real-time computer systems are used in various fields of science and engineering. The formal definition of real-time systems can be given as follows: the real-time system is a system of excitation-response (cause-and-effect relation) elements where excitation-response relations are time limited [1]. The system must process events which do not necessarily occur in stated order. Often a real-time computer system is understood as a hardware/software complex which allows to generate determined signals or estimate parameters of signals as fast as input data alters. Real-time system is a system in which data are processed as it is acquired in-stead of being accumulated and processed at a later time.

The real-time systems are closely associated with processes and phenomena of the real world. Any action or inactivity of such a system always cause some consequences. Therefore high efficiency and reliability of software of real-time systems are required. The programs must handle errors which may occur as the system operates to reduce data loss. The organization of system's reactions to parallel inquiries appears to be extremely complex, i.e. special efforts should be taken for synchronization of processes in real-time systems. The real-time measuring systems must satisfy the requirements of real physical experiment and work under strictly time-critical conditions. Such systems have to operate during a long time, therefore they must be capable to self-recover if a failure or malfunction of the equipment occurs. At last, allocation of resources and their manageability are especially important problems for the real-time software.

One of the most serious problems of real-time systems is time efficiency. Both fast algorithms of data processing and organization of actual synchronization of different parts of a system are necessary. However, there is no need in fast processing algorithms if a program is awaiting data. Programs which operate in a single processor system may be not effective or not working at all in a multiprocessor system. The real-time systems demand a lot of computing resources. With such systems, the problem of reallocation and reinitialization of resources is the most significant, since the programs have to process parallel inquiries, which sometimes contradict each other.

Hard dependency of software on the type of a processor is the serious imperfection of real time systems. The DSP hardware independence can be realized by means of operating system and (or) joint efforts of user and programmer. If a program allowed the user to allocate resources on processors (to choose the particular processor on which the software block will run) and to profile the block (how long processor works or awaits), the user could adjust allocation of blocks on processors.

Testing of real-time software is not a simple problem. Analysis of results in time-critical conditions and in conditions of concurrency is significantly more difficult than in serial systems. In a case of error detection it is rather difficult to find out unit of malfunction. Besides in some areas the reliability requirements can be so great that the testing of quality is not a guarantee.

Use of high-level languages for creation of real-time systems allows to simplify developing of software and to improve its reliability. Now low-level tools like interrupts are often used for maintenance of synchronization. Special software has been developing recently and such software allows to organize threads. This approach implies that any part of a program is designated as a thread and then the operating system runs it on a processor which is idle at the current moment. Both developer's contribution and support of development software and operating system are necessary for maintenance of the actual synchronization.

Highlights of developing the real-time applications. Real-time and multitask mode programs are the most significant part of data acquisition and information control systems. Personal computer's operating system is the main obstacle for their realization. The real-time is needed because it is required to recognize processes and react to events immediately. There is no sense in analyzing characteristics of process after it has changed. Besides the system must have be able to run several tasks simultaneously. For example, an emergency has occurred, and some adjustment of the system and getting printed document are necessary. Additionally, a human operator has to enter data from keyboard and to control additional variables of this or other processes.

After the data acquisition is completed it appears necessary to present data in graphics, to perform some mathematical operations (transforms, calculations), to print data or write them to disk. Writing thousands of similar records would be waste of memory and time. Statistical and spectral analysis is used for reduction of data volume and for saving the memory and disk space. Multitasking in real-time allows operators and experts make decisions during system operation. Also it allows to analyze data retrospectively to look backward at the event which has been elapsed very fast, to perform statistical or spectral analysis without the data acquisition being halted.

Processors for digital signal processing. The microprocessors with the architecture and the instructions set specially intended for digital signal processing applications were released at the beginning of 80's. Since DSP processors have been designed by many companies: Texas Instruments, Analog Devices, Motorola, AT&T and others.

The DSP microprocessors support parallelism and pipeline organization for increase of operating speed. Pipeline is a high-performance processor structure in which the completion of an instruction is broken into its elements so that several elements can be processed simultaneously from different instructions. The majority of DSP chips has Harvard architecture (dual program and data access). Separation of the data bus and the address bus allows to process data simultaneously with loading next instruction. DSP chips have a number of specialized processing devices, each being able to work in parallels with the others. 32-bit floating point DSP processors are the powerful tool of the 90's. They are applied to the various of areas, from recognition of speech to supercomputer processing, and first of all to real-time systems design.

The DSP- processors are not conventional RISC or CISC processors. Actually, they are similar to the original microprocessors designed for solving math-intensive problems. Programming of DSP-chips includes learning the new ways of processing loop design.

Last years 32-bit DSP-processors have united advantages of 32-bit addressing in RISC-processors and number-crunching ability of near-supercomputer vector processors [2]. Larger memory size allows DSP-processors to use high-level programming languages like C. C makes it possible to manage resources of a computing system, to organize data input/output and to write effective programs for digital signal processing.

New DSP-chip architectures lend themselves to multistage processing as well as to parallel operations. For example, Texas Instruments' TMS320C40 has communication ports, which can link a single processor to as many as six other processors for data exchange and work coordination. DSP-chips become the key tool in the multimedia applications, such as combining voice and video information.

Data processing system. Hardware/software complex was developed at the department of Computer Engineering of the Vladivostok State University of Economics. It is intended for data acquisition, control, data processing, analysis and data presentation. It contains the data acquisition board ADC12/200 and digital signal processing board on the basis of Texas Instruments' TMS320C30 processor. The software of this complex allows:

- to input data using up to 16 channels,
- to establish various modes of input (to assign sampling frequency, gain, number of channels an so on);
- to present the acquired data in an oscilloscope mode and color frequency representation (using zero crossing);
- to perform statistical preprocessing (to remove trend, to do smoothing);
- to store data to disk in various formats (pack, unpack, wave format);
- to process data using methods of spectral analysis (averaging periodograms method, auto-regressive, moving average, Prony and MUSIC are used for power spectrum estimation, cross-spectrum and coherence function);
- to estimate statistical parameters (mean value, variance, deviation and others) and to build histogram of distribution function;
- to design digital filters and to perform signal filtering.

The software contains modules working on the DSP-processor. Calculation of the spectral characteristics and signal filtering may be carried out in a multiprocessor mode or only on the main processor. We will pay great attention to the developing of the digital signal processing software for dual-processors multitasking mode.

The operating system DEASY is intended for coding, debugging and running DSP programs for TMS320C3x processors modules. DEASY is a single task simple operating system of real-time [3]. DEASY libraries' functions were used for organization of synchronously performing processes.

For organization of a multitask mode it is necessary to pass messages, which display a current condition of program operation. For example, when the calculations is completed by DSP processor (DSPP) it is necessary to send message to another (CPU) to tell it may read the results and to initiate the continuation of the DSPP calculations. Often it uses the format of programming known as event-driven programming. An event is any operation that requires some response from the application program. An event produces a message whenever the event occurs.

System functions libraries allow to generate interrupts between processors and to arrange interrupts processing. The pass of messages and data was organized on the basis of DEASY libraries' functions with the use of interrupt handling. The library for CPU is intended for management of the DSP-board carried out by CPU running user's program written in C language. The library is delivered for compilers Borland C/C ++, Microsoft C. It was updated by the authors for Zortech DOS, 32-bit DOSX and for MS Windows 3.1.

The realization of algorithm with use of the DSP processor. The program of calculation of the spectral estimations is intended for calculation of an auto-spectrum, cross-spectrum and coherence function. The calculations can be performed with the processor TMS320C30 or without it. In the first case DSPP performs the most part of calculations. CPU inputs data and outputs results.

Classical periodogram technique of estimation of the spectral characteristics [4] is used (other spectral methods are performed by only CPU). In this method a signal is divided into overlapping segments, and each segment is windowed prior of computing the periodogram. Periodograms are calculated by FFT, and spectrum estimation is the average of periodograms. The results of intermediate accumulation are displayed during calculation. The user can halt process of calculation, look through the results, write the results to file, continue or stop account.

Let us consider the program. The algorithm of spectrum estimation may be broken into three basic steps:

1. Selection of the i^{th} data segment (from board or reading from file).
2. Calculation of spectral characteristics of the i^{th} segment (removal of mean value, windowing, FFT).
3. Accumulation and representation of the results.

It is obvious that item 2 should be performed by the DSPP at the same time when item 1 and item 3 were carried out by CPU to provide maximum efficiency.

The mechanism of pass and processing the messages is used for realization of parallel work of two processors. Both processors are constantly in a waiting mode, pass or processing the messages. Spectrum estimation of the i^{th} segment may be performed simultaneously with the display of the results of accumulation up to the $(i\text{-}1)^{th}$ segment and reading of the $(i+1)^{th}$ data segment. DSPP performs all basic computing functions: calculation of window function and energy of window, windowing of data, FFT and calculation of i^{th} segment spectrum. DSPP may average periodograms and calculate logarithm but it is memory size is limited (in our case) thus CPU performs these program functions.

Often it is necessary to pass a signal through a finite impulse response filter with a fixed number of taps. The filtering operation must be computed piecemeal. The technique known as the overlap-add method [5] is used for signal filtering. The algorithm can be broken into following items:

1. Selection of the i^{th} data segment (from board or reading from file).
2. Calculation of FFT and performance of point-by-point product with transfer function.
3. Calculation of inverse FFT.
4. Addition of overlapping parts of output blocks.
5. Write and display of filtering results of the i^{th} segment.

Items 1 and 5 are carried out by the CPU while items 2-4 are performed by the DSPP concurrently to provide maximum efficiency. Signal filtering on the i^{th} segment can be performed simultaneously with write and display of the filtering results for the $(i\text{-}1)^{th}$ segment and reading of the $(i+1)^{th}$ data segment.

Testing the developed software. Efficiency of the developed digital processing software depends on throughput of chosen processors, on the task to be solved, and on how the program shares system resources. We have tested processor pairs Intel 80386 (40 MHz) + TMS320C30, Intel 80486 (50 MHz)+ TMS320C30, Intel 80486 (100 MHz) + TMS320C30. The program has significant time advantage with pair Intel 80386 + TMS320C30 in comparison with Intel 80386 alone. However, it may work slower with pair Intel486 (50) + TMS320C30 than with Intel 80486 (50) alone. It is caused by the fact the CPU carries out its own part of the job faster than TMS320C30 does, therefore the latter makes the CPU stand idle waiting for the results of processing. Obviously, the more effective CPU is used, the more effective DSPP is necessary to be chosen. The program of signal filtering with display of processing results fits pair Intel 80486 (100 MHz) + TMS320C30 most of all.

The program of calculation of the spectral characteristics was tested on Intel 80386 (40 MHz) and Intel 80486 (50 MHz) processors. The time (in seconds) of the program's operation with TMS320C30 as well as without it, is shown in the Table 1. The signal was read from file by 512 samples portions. The signal length was 100 Kwords. The processing occurred either on the CPU, or on the CPU and TMS320C30 in parallels. The CPU displayed intermediate results in graphic form while TMS320C30 was processing. As it follows from the Table 1, spectrum estimation is calculated significantly faster on TMS320C30 and Intel 80386 than on single processor Intel 80386. With Intel 80486, the operating time with and without TMS320C30 are nearly the same. This circumstance is explained by the fact that Intel 80486 processor is almost as fast as TMS320C30. Therefore speed of periodogram calculation has to match speed of data acquisition and of result output precisely, otherwise the processor Intel 80486 will stand idle.

Table 1 Operating Time of Power Spectrum Estimation Program

i80386	i386 +TMS	i486 (50)	i486 (50)+TMS
25	18	14	13

The tests were performed for various data length, overlapping factors, amount of spectral samples. They have shown that the speed ratio with and without TMS320C30 are the same.

The filtering program for Windows 3.1 was tested on Intel 80486 (50MHz) and Intel 80486 (100 MHz) processors. The time (in seconds) of the program's operation with and without TMS320C30 is shown in the Table 2. The signal was read from file by 1 Kwords portions, processed and written to disk. The signal length was 100 Kwords. The processing occurred either on CPU only or on CPU and TMS320C30 in parallels.

Table 2 Operating Time of Filtering Program

	i486(50)	i486(100)	i486(50)+TMS	i486(100)+TMS
without display	12.4.	4.8	6.4	4.7
with display	32.5	10	25	6.2

Let us consider the signal filtering program without display results. As it follows from the Table 2, with Intel 80486 (50), the signal filtering program using DSPP requires twice less time than on single processor Intel 80486(50). With Intel 80486(100), the speeds of calculations with and without TMS320C30 are nearly the same. The last tend takes place because the Intel 80486(100) carries out its own part of the job (reading and writing of data) faster than TMS320C30 does, therefore the latter makes the CPU stand idle waiting for the filtering results.

The program on pair processors Intel 80486(50)+TMS320C30 is performed for 6.4 seconds but CPU stands idle for 3.4 seconds. This time can be spent on display of output data. But data display is a very long procedure for the Intel 80486(50) without graphical accelerator. The processor Intel 80486(50) requires 19 seconds for data display in this test. Signal filtering time is the small part of common execution time. The similar test on Intel 80486(100) gives about 3.9 seconds for display and 1.6 second for reading of data and writing of results. Thus the filtering program on the pair processors Intel 80486 (100) and TMS320C30 can be organized with data display in parallels.

Conclusion. Software must be organized so that both processors work in parallels without idle. The real-time software will develop as well as all areas of programming. Complexity of processes and phenomena controlled by real-time systems increase. It raises the level of software requirements. The development of real-time systems is connected with parallel calculations using of the object-oriented programming. This direction is of great interest for improving of algorithms efficiency and system characteristics.

References

1. Bologna, S., W. Ehrenberger, J.R. Taylor (1986). In W.J. Quirk (ed), Verification and Validation of Real-Time Software. Springer-Verlad. Berlin, Heidelberg, New York, Tokyo.
2. Weiss R. (1991) 32-bit floating-point DSP processors. EDN, November 7, pp.127-146.
3. DEASY User's Guide (1994) InSys Corp. Moscow.
4. Marple S. L. Jr. (1987) Digital Spectral Analysis with applications. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
5. Rabiner L. R., B. Gold (1975) Theory and Application on Digital Signal Processing. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

TRAFFIC MODELS AND MANAGEMENT IN HIGH-SPEED NETWORK

Vladimir Zaborovski, Sergey Yegorov
State Technical University of Saint-Petersburg
Polytekhnicheskaja Street, 29, Saint-Petersburg, 195251, Russia
e-mail: vlad@stu.neva.ru

Yuri Podgurski, Yuri Shemanin
Institute of Robotics and Technical Cybernetics
Tikhoretsky Avenue, 21, Saint-Petersburg, 194064, Russia
e-mail: yuri@neva.ru

Abstract. The ability to access the right information at the right time and in the right place is a key feature of the new information technology. To make an efficient networking possible the intelligent management functions are needed. The development of information systems based on high-speed computer networks leads us to the discussion about what does "efficient management" mean. Basing on example of remote real-time control via ATM network we present a modeling technique for actual traffic in high-speed networks and propose a formal approach to analyze and design congestion control mechanisms.

Key words: traffic model, network management, optimization procedure, automatic control.

Introduction. The development of information systems based on high-speed computer networks has a special place among the trends of modern technical systems development. From system analysis point of view such systems can be represented as complicated distributed hierarchical objects whose parameters and structures are time dependent. This feature considerably complicates tasks of control and providing quality of service (QoS) needed. In contrast to phone and TV systems, computer systems use sophisticated protocols for access to resources. The feature of those protocols is their interaction with distributed data bases keeping information describing state and logical connectivity of network structure. That feature of computer communications allows using of facilities and methods of intelligent control for adaptation of applications to used network infrastructure. Such shift towards intelligent control creates basis for integration of existing communication systems into united Intelligent Information Infrastructure (III) of computer telecommunications. This paper presents the results of current network management research holding within RUSnet project [1]. The concept of the project is based on the III-model and Telenetics paradigm [2,3]. The main areas of the project's research are dedicated to the principles of management in computer network as well as to traffic congestion control algorithms based on optimization and adaptation procedures.

Network management design

The most difficult aspect of high-speed network development and hotly debated topic is management design. The main goal in designing is to maintain the QoS while attempting to make maximal use of network resources. To define the requirements to network control a network management system may be decomposed into three parts: congestion control techniques, traffic management functions and information access/search algorithms. The selection depends upon the duration and severity of management. Congestion control deals only with problem of reducing load during overload. Traffic management functions have to be intelligent procedures in their nature since they try to provide needed QoS. For this purpose the network continuously monitors its traffic, provides real-time data mining and feedback to the source end systems. The amount of information available in networks is growing very fast. In early information retrieval systems, users need to know which set of information servers is most likely to have the desired information, and how to interact properly with them. In order to decrease user load the intelligent solutions with data mining algorithms may be used as a part of network management system.

New management requirements of future network infrastructure can be made basing upon the analysis of application tasks that use the new abilities new networking technologies have. As an example of such task we will discuss the task of remote control of the space arm via ATM network. The 15-meter arm had been built in RTC Institute for Russian space shuttle "Buran" is the unique technical device and able to position objects with a mass up to 15 tons in space environment. In case of including that unique complex into global cyberspace its abilities could be used not only by Russian researchers but as well by scientists of others countries for realizing their national space programs.

Let us discuss the technical aspects of such kind of control technique and basing on them outline the requirements to network management design.

There are following principles of building the system:

1. Transport network has to provide defined QoS for real time multimedia traffic;
2. Intelligent network agents have to be used for resource control;
3. User interface to system resources has to be provided by standard Internet facilities.

As the access to system resources has to be provided from several points, James ATM network has been chosen as a transport network fit to principle 1. Point of presents (POP) at Helsinki has been chosen to establish connectivity with James. ATM E3 channel with PVC 2 Mbps was established from POP to Petersburg. Decision has been made to use TCP/IP for providing standard interface. That is why switching equipment used to carry on the experiment had to support IP over ATM specification. In that case, it was possible to use standard interface facilities such as WWW browsers, HTML-3.2 and video transmission system VIC. It worth to note that VIC is available in sources as well as binaries for all platforms including UNIX, Win95, WinNT.

Handling of large volumes of data and providing real time control demands to use intelligent agents and expert systems. An access to such an agent can be based on using WWW server cgi-scripts. Experiments of building such system and therefore approving the principles chosen has been carried out in Nov. 1996. Remote control station resided at Torino (Italy). It was a workstation running WWW browser Netscape-3.0 and videoconferencing system VIC-2.8. Solaris-2.4 (for SUN SS20) and Linux 2.0.26 (for PC) were used as operating systems. Application server resided at Saint Petersburg (Russia). It was a PC in following configuration: Pentium 133, 32 MB RAM, two framegrabbers Matrox Meteor RGB, OS Linux 2.0.26. Apache-1.0 has been used as WWW server. Two copies of VIC were running on the same PC for transferring video from two cameras simultaneously. First camera was placed at space arm manipulator and the second was giving the overview of arm's environment. The presence of visual feedback allowed operator to control the arm movement.

The input of information by operator was based on standard facilities of HTML-3.2. Cgi-scripts (C-written) handled forms and image maps. The same cgi-scripts communicated with intelligent agent -- program of estimating extrapolating and simulating of robot's dynamics that put control to special control computer via serial RS-232 interface. It ought to note that performance of PC was more than enough to run WWW-server, intelligent agent and two videoconferencing programs (format 1/4 PAL, 8-15 fr/sec). System debugging and testing was carried out in RUSnet network, full-scale experiment has been organized with participation of specialists and specially dedicated resources of Telecom Italy, Telecom Finland, James ATM, and RUSnet networks and provided by Lenenergo Ltd. fiber infrastructure.

As we have shown above, new network application continue to migrate to media-rich and interactive forms of content. In order to get such application to work a combination of congestion control techniques, traffic management mechanisms, and information access algorithms is used. For this purpose it is possible to make it effectively using extended model "client-intelligent agent - server". The basic unit of information for such a system is a message of defined inner structure. Having received such a message, intelligent agent working independently executes actions requested by client and returns response message to client. For comparative analysis of these schemes a formal approach is needed.

Traffic management: a formal approach. To specify complex quality of service requirements a network management model should be decomposed into several parts and represents it as a hierarchical system with different levels of control function and a formal definition of each management level. For many traffic management control functions, such a hierarchical presentation is natural and even unavoidable. For example, a complete traffic management strategy should include several congestion control and avoidance schemes that work at different protocol levels, and that can handle congestion of varying duration [4]. From the automatic control point of view, it means that there is a set of automatic control systems that have different control rates and respond to the input force changes in appropriate time intervals (short or long). Despite the differences in management goals and time scale of levels, we propose to give the formal definition of every management level and to formalize the task of control synthesis based on:

- the model of input variables,
- the model of object under control,
- the formal definition of efficiency criterion,
- the methods of the optimal regulator deriving,

This leads us to the discussion: what does efficient management mean? What are the numerical metrics of management efficiency in general and of each control mechanism in particular? Below, without loss of generality, we discuss our considerations of choosing models and efficiency criteria oriented to high-speed network with cell-based traffic (ATM-technology).

A primary role of traffic management in ATM networks is to protect the network and the end-system from congestion in order to achieve network performance objectives and use network resources efficiently. The ATM Forum has defined a set of main traffic management functions that may be used in appropriate combinations depending on the service category [5].

Note that the input variables for these functions can be divided into individual source traffic (for example, for Usage Parameter Control) and total network workload (for example, for ABR Flow Control). In order to take into account data stream characteristics while synthesizing the control law, we have to develop models of individual traffic source (IND-model) as well as multiplexed traffic source (MIX-model). The robust model has to be built basing upon actual traffic parameters. In accordance with ATM Forum specifications [5], each actual connection is specified by traffic contract which includes the negotiated values of traffic (Traffic) and Quality of Service (QoS) parameters, as shown in the Table 1.

Table 1.

Constant Bit Rate (CBR)		Variable Bit Rate (VBR)		Available Bit Rate (ABR)		Unspecified Bit Rate (UBR)	
Traffic	QoS	Traffic	QoS	Traffic	QoS	Traffic	QoS
PCR	CLR	PCR	CLR	PCR	CLR	PCR	-
	CDV	SCR	CDV	MCR	-		
	CTD	MBS	CTD		-		

PCR- Peak Cell Rate
SCR- Sustainable Cell Rate
MCR- Minimum Cell Rate
MBS- Maximum Burst Size
CLR- Cell Loss Ratio
CDV- Cell Delay Variation
CTD- Cell Transfer Delay

That is why the ability to handle those parameters to create different types of control schemes is very important. But the traffic parameters above describe the worst case restrictions of traffic contract and are not sufficient to define the traffic source behavior as a function of time and to synthesize an optimal control system. The additional information is needed to describe the stochastic nature and time dependency of a cell flow for the connection. To create an automatic control first of all the adequate model of the traffic source is needed.

Traffic source model. The individual source traffic can be defined as a discrete parameter stochastic process $\{X_n\}$ representing the spacing of cell flow. The process parameters have to satisfy the restrictions derived from the traffic parameters (Table 1): Mean $\{X_n\}=1/SCR$, $X_n \geq 1/PCR$, maximum length of continuous sequence of minimal values $(X_n=1/PCR)$ must be not greater then MBS. The analytical expression of a such cell flow is absent.

As a first step of developing unified formal IND-model we propose a new Petri net based model of individual traffic source, which takes into account the real traffic parameters (Table 1) and may be used for both simulating and some analytical investigations. Petri net is a well-known formal tool for modeling the behavior of discrete systems and have an analytical and graph interpretation [6,7]. The model is defined as Petri net (Fig.1) with time parameter T_0.

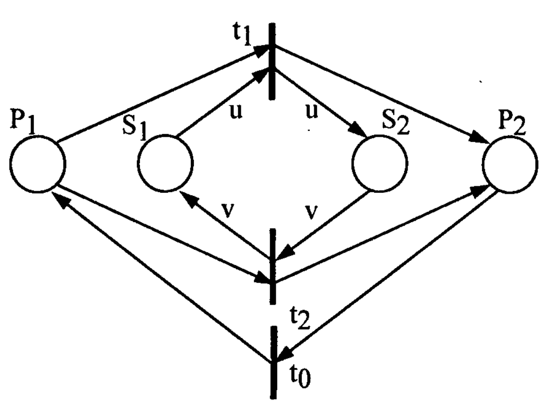


Fig. 1

The arcs **(s_1,t_1)** and **(t_1,s_2)** of the model both have weight **u**, and the arcs **(s_2,t_2)** and **(t_2,s_1)** both have weight **v**. Hence, whenever **t_1** fires, **u** tokens move from **s_1** to **s_2**, and whenever **t_2** fires, **v** tokens move from **s_2** to **s_1**. The initial marking **M_0** of Petri net is

$M_0=(M(p_1)=1, M(s_1)=z\geq u+v-1, M(s_2)=0, M(p_2)=0)$.

The values **$M(s_1)$, u, v** and **T_0** are the functions of service category and traffic parameters.

The transition **t_0** may be considered as an internal clock of the source and has a finite firing time **T_0**. The interval between two consecutive occurrences of **t_0** then is taken to be a time slot of internal clock. Transitions **t_1** and **t_2** fire instantly after enabling. As **t_1** and **t_2** may fire only alternately with **t_0** if **$M(p_1)+M(p_2)=1$** the firing of **t_1** may be viewed as emulating the sending of a cell whereas the firing of **t_2** emulates the time slots during which no sending takes place.

It may be shown [8] that model regulates, for reproducing processes **Pr**, the firing count ratio between **t_1** and **t_2** as **$h_1(Pr)/h_2(Pr)=v/u$**, where **$h_i(Pr)$** is the component of reproduction vector **h(Pr)** associated with the transition t_i. The Petri net model is obviously live if **$z=M(s_1)\geq u+v-1$**. The actual number of tokens **z** which constitutes a live marking of model determines how rigorously the firing count ratio **$h_1(Pr)/h_2(Pr)=v/u$** is enforced. This may be illustrated by mentally executing the Petri net with stepwise increasing the initial number of tokens in place **s_1**. We may observe that we introduce an increasing degree of freedom as to the choice of either firing **t_1** or **t_2** under a particular token distribution. A quantitative measure for this phenomenon is a synchronic distance formally defined in [8]. The synchronic distance reflects the degree to which the two transitions may deviate from the mean firing count ratio as specified by the respective components of reproduction vector. It may be shown [8] that for Petri net of Fig.1 and a live marking with **$M(s_1)=z-q$**, **$M(s_2)=q$** and **$M(p_1)+M(p_2)=1$**, the synchronic distance between **t_1 and t_2** is given as

$$Syn(t_1,t_2)=\lfloor (z-q)/d \rfloor + \lfloor q/d \rfloor,$$

where $\lfloor x \rfloor$ denotes the greatest integer $\leq$**x** and **d** is the greatest common divisor of **$h_1(Pr)$** and **$h_2(Pr)$**. The synchronic distance fits very well to be a measures of maximum burst size. The different classes of traffic source may be efficiently modeled by using the proposed model.

For example, consider the VBR class traffic source model. The relations between the Petri net characteristics and traffic source parameters are as follows:

$$M(s_1)=MBS*(PCR-SCR)/w,$$

where **w** is the greatest common divisor of **SCR** and **(PCR-SCR)**,

$$u=M(s_1)/MBS,$$

$$v=M(s_1)*SCR/[MBS*(PCR-SCR)],$$

$$T_0=1/PCR.$$

Then the firing sequence of transition **t_1** will exactly correspond to the VBR traffic with

$PCR=1/T_0$, **$SCR=1/T_0*v/(u+v)$,** **$MBS=M(S_1)/u=Syn(t_1,t_2)/u$.**

For instance, let the traffic parameters of VBR source are as follow: **PCR, SCR=PCR/3, MBS=4**. From the equations above we have:

$M(s_1)=8$, $v=1$, $u=2$, $T_0=1/PCR$.

The possible firing sequence which starting with initial marking Mo and time diagram of cell sending are shown in the Fig. 2,

M(s1)	8	$t_1>$	6	$t_2>$	7	$t_1>$	5	$t_2>$	6	$t_2>$	7	$t_2>$	8	$t_1>$	6	$t_1>$	4	$t_1>$	2
M(s2)	0		2		1		3		2		1		0		2		4		6

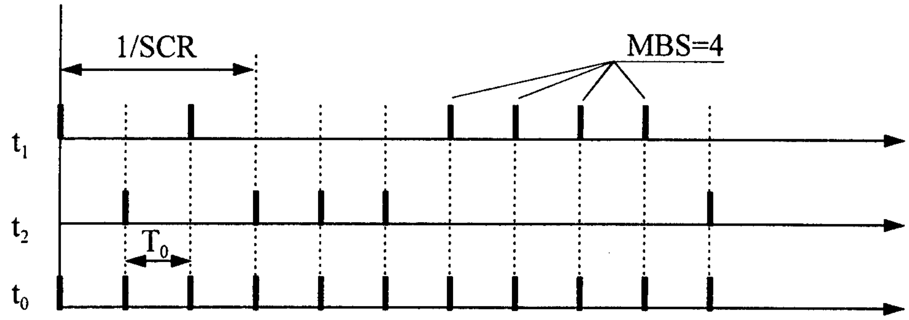


Fig. 2

where $\begin{matrix}1\\0\end{matrix} \overset{t_i>}{} \begin{matrix}0\\1\end{matrix}$ denotes the change of the marking **M(s$_1$)=1**, **M(s$_2$)=0** into **M(s$_1$)=0** and **M(s$_2$)=1** due to the firing of the transition $\mathbf{t_i}$.

The same Petri net may be used to model the CBR and UBR class traffic sources. To model the ABR class traffic source the above Petri net model is suitable with the extension as follows: the firing time $\mathbf{T_0}$ of transition $\mathbf{t_0}$ is not constant but is enforced by ABR Congestion control scheme feedback. The range of $\mathbf{T_0}$ is from 1/PCR to 1/MCR.

To create a complete source model is an important task. The above traffic parameters describe the worst case restrictions of traffic contract. The additional information is needed to describe the stochastic nature of cell sending. It is necessary to study different application traffics to determine their consistent time and distribution characteristics as well as to evaluate the adequacy IND-model as a Marcovian realization.

An example of optimal control synthesis. Despite all its outstanding capabilities, ATM technology has some internal features that can cause traffic related problems. According to [4], even in lightly loaded cell-switching networks it is not unusual to observe cell blocking due to network congestion. The loss of a single cell in an ATM transmission leads to higher-level packet retransmission and an exponential increase in traffic through a switch. A congestion collapse is of particular concern when ATM switch relay large amounts of bursty data. In order to efficiently manage the network resources an additional research is needed.

Measurements of LAN and VBR video traffic show that data flow in packet and cell based networks differs fundamentally from conventional telephone traffic, and may be even fractal in nature [9]. In this case some conventional QoS criteria associated with actual bursty traffic may differ drastically from that predicted by conventional models. It should be taken into account while creating the MIX-model of cell flow. Below we show that the actual nature of packet traffic has serious implications for network engineering and management approach. Consider the cell discarding algorithm.

Input variable model. We will use a function of time **F(t)** as a model of input packet traffic. Let **q=1/PCR** is the elementary time unit. The function can be observed every **q** time unit during interval **T**. If the function is well-known the control decision is possible in off-line mode. Let **F(t)** is a stochastic process and we know only its autocorrelation function. Bellow this function will be used for optimal control synthesis.

According to [9] at every time scale autocorrelation function of traffic in cell-based network decays much more slow then Poisson-like law predict. Let us consider two variants of input function **F(t).** The first variant indicates the presence of long-range dependence and given by autocorrelation function

$$K_f(t) = (g+t)^{-b}$$

where $0<g<1$ parameter of regularization and $0<b<1$.

A long-range dependence may cause heavy losses and as a result long decays during bursts.

The second model defines the Poisson-like process with an exponential curve of autocorrelation function

$$K_p(t) = ne^{-bt}$$

In the frequency domain traffic, models can be characterized by a spectral density $A_f(w)$.

The model of object under control. The common requirement for mathematical model is its adequatennes to the real system behavior. To accommodate bursty traffic, several approaches are evolving: dynamic buffering mechanism with per-VC queuing, intelligently organized packet/cell discard algorithms, and ABR algorithm which based on Resource Management (RM) cells. This cells indicate the presence or absence of congestion and force end systems to change the amount of data being sent into network.

Consider the cell discarding algorithm in the buffer of the cell switch. The buffer provides additional resource for a limited time interval **T**. That resource is equivalent to temporal extending channel bandwidth to smooth the burstness of traffic. Cells from virtual circuits (VC) are scheduled in statistical multiplexing order. If any one VC sends a large burst of data then the service rate at the buffers of all other VC drops until the burst has been served. We will call the buffer for VC as a virtual buffer (VB). To protect the VB from congestion, a cell discarding scheme may be used as shown in Fig. 3.

The aims of the control are to maintain the number of cells in the VB queue at a desired setpoint, and protect buffer from overflow. Since the data flow has burst components, a border of the buffer's filling will oscillate around the setpoint value.

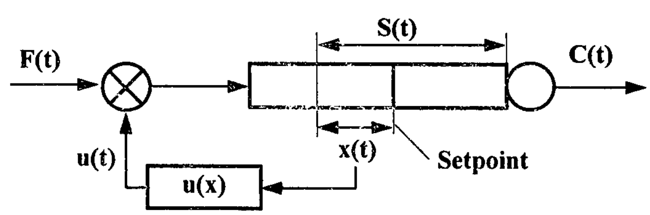


Fig. 3

The choosing of the setpoint (at the SetUp phase of connection) reflects a tradeoff between mean cell delay, cell loss and bandwidth loss. Let: **S(t)** - current buffer's filling, **<S(t)>** - mean value of **S(t)**, **<F(t)>** - mean value of **F(t)**.

If a control action is done once every **q** time units we can make the fluid approximation of data flow, ignore cells boundaries and define a speed of the buffer filling **dS/dt** by

$$dS/dt = F(t) + u(t) - C,$$

where **C** is the mean service rate and **u(t)** - control rate for discarding cells. In this equation **S(t)** can be positive as well as negative. Negative values characterize a degree of not using of available VC bandwidth. Denote the value of a variable **x(t)** by

$$x(t) = S(t) - \langle S(t) \rangle$$

and a variable **f(t)** by

$$f(t) = F(t) - \langle F(t) \rangle$$

In case of mean value **<F(t)>** equal to service rate **C**, we can write the buffer equation in derivations

$$dx/dt = f(t) + u(t)$$

Formal definition of efficiency criterion. The efficiency criteria have to reflect the aim of control and take into account the values of QoS parameters needed by the particular traffic contract (see Table 1). In addition while choosing the criterion the designers have to minimize the complexity of the control value computation. In our case we have following functional

$$J = \frac{1}{T}\int_0^T Q(x(t))dx \quad .$$

If $Q(x)=x^2$ then **J** denote as

$$J=\langle x^2 \rangle \quad .$$

The requirements to minimization of CLR value can be expressed by means of integral limitation $\langle u^2 \rangle \leq N_u$ and efficiency criterion expressed as weighted sum

$$J = m^2 \langle x^2 \rangle + \langle u^2 \rangle,$$

where m^2 - weight coefficient. Its value depends on value of N_u.

The first item in **J** enforce the minimizing of Cell Delay Variation (CDV) the second item enforce the minimizing of Cell Loss Ratio (CLR). The choosing the value of N_n reflects the relative significance of CLR and CDV regulating.

The calculating of optimal regulator characteristics. The depth of criterion minimum that can be reached depends on information of input data streams. In order to compute the minimum of **J** it is possible to use spectral density $A_f(w)$, that is Fourier transform of autocorrelation function $K_f(t)$.

In case of $K_f(t) = (g+t)^{-b}$ the $A_f(w)$ is a generalized function, and value of $\langle u^2 \rangle$ could be calculated

$$\langle u^2 \rangle = \left(\frac{m}{2}\right)\int_0^\infty (1 + tm)(g + t)^{-b} \exp(-tm)dt$$

as well as **J**

$$J = m\int_0^\infty \exp(-tm)(g + t)^{-b}dt$$

This value of **J** is achieved by using a linear feedback law

$$u = m^2 \int_{t_0}^{t} x(\tau)d\tau - 2f(t_0)$$

Note that if $A_f(w)$ satisfies a Poisson-like distribution with parameter **n=1**, then $\mathbf{J = m/(b+m)}$. Figures 4 and 5 show, in case of exponential distribution, the intensity of optimal control is lower than in case of actual burst traffic. Note that the decreasing regularization parameter **g** of the autocorrelation function causes the rising of control intensity.

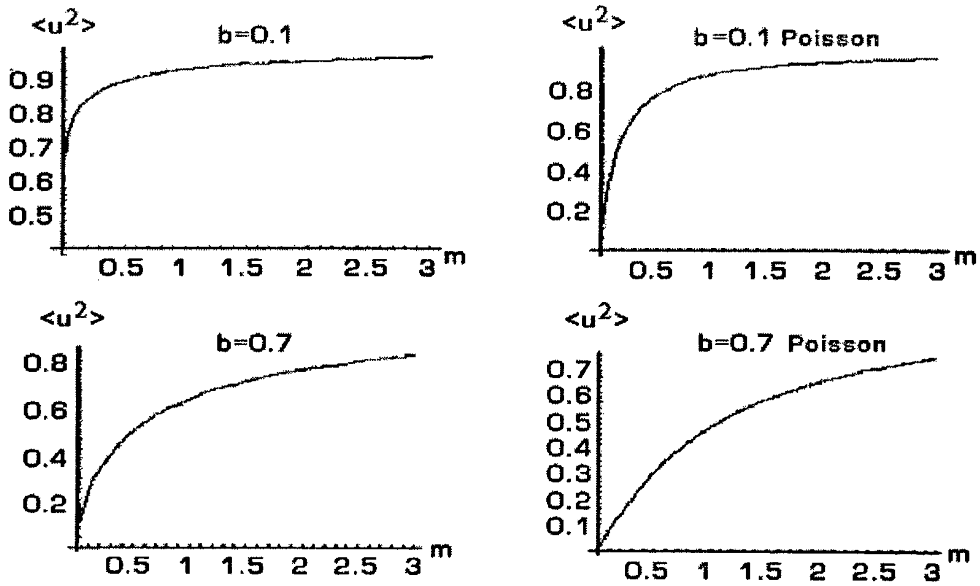


Fig. 4

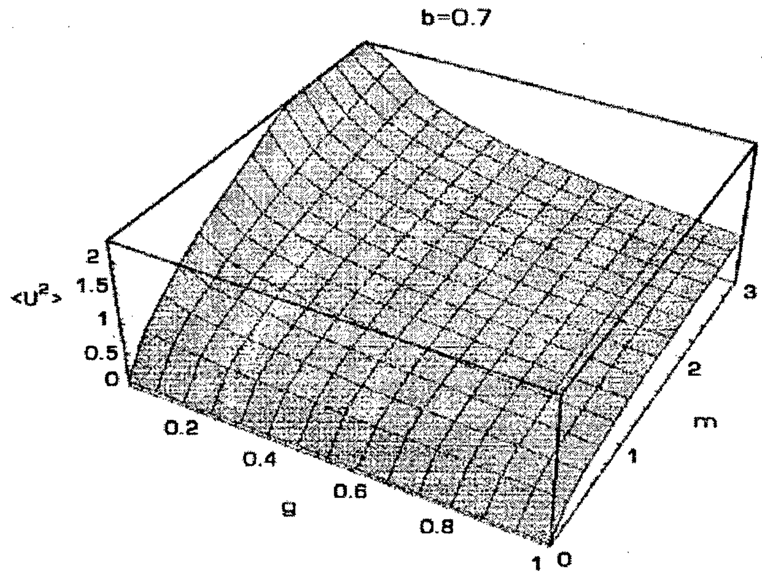


Fig. 5

Figures 4 and 5 show that not many frequency constituents of spectral density of input traffic are equally important for the system being controlled. It is enough to satisfy the conformity between the theoretical and real spectral densities in limited bandwidth only, where the module of frequency characteristic of loopbacked system can be neglected. Therefore it is possible to use Poisson-like distribution to define traffic flow characteristics for cell discarding algorithms. Note, the received value of control action **u(t)** can be used in discarding cells algorithm as well as to change the source intensity in ABR mode. This circumstance has to be taken under consideration to synthesize the controls needed.

Conclusions. Automatic control approach is proposed to formalize traffic management system design. The main steps of automatic control methodology include the hierarchical representation of management system and the formal definitions of input variables, object and goal of control of each management level. Requirements to the input variable models are discussed. We propose to distinguish two classes of input variables models. These models should be based on the actual individual traffic descriptors (IND-models) and on the real network workload parameters (Mix-models). A Petri net model of individual traffic source based on ATM forum specification is presented. An example of an optimal control scheme for cell discarding algorithm is presented. It is noted that the current set of traffic parameters recommended by ATM-forum is not enough to synthesize optimal control system. It is necessary to explore each applications' traffic to determine its time and distribution characteristics.

References

1. Lopota, V. A., V. S. Zaborovski (1995) Bringing Internet to North-West of Russia - RUSnet N/Wproject. Proceedings of INET'95, Honolulu, USA, pp. 855-857,. (http://info.isoc.org/HMP/PAPAR/115/abst.html)

2. Zaborovski, V. S. (1995) The main aspects of construction of the Northern-Western segment of Russian University and Scientific Network. Posters of JENC6. Tel Aviv. .

3. Zaborovski, V. S., A. Vasilev (1996) Cooperative Network Engineering in N/W of Russia. Proceedings of MULTICUBE Collaborative Workshop, Torino - Madrid.

4. Jain R. (1994) Myths About Congestion Management in High-Speed Networks.. (http://www.cis.ohio-state.edu/~jain).

5. ATM Forum Traffic Management Specification. Version 4.0. (1996) (http://www.atmforum.com)

6. Podgurski, Y., V. Nikonov (1984) Petri Nets. Theory and Applications. Zarubegnaja Radioelectronika ,4, pp. 28-59, (in Russian).

7. Podgurski Y., V. Nikonov (1986) Petri net Applications. Zarubegnaja Radioelectronika , 11, pp. 17-37, (in Russian).

8. Kluge, W., K. Lautenbach (1982) The Orderly Resolution of Memory Access Conflicts Among Competing Channel Processes. IEEE Trans., C-31, 3, pp. 194-207.

9. Leland W., E. (1993) On the Self-Similar Nature of Ethernet Traffic. Proc. SIGCOMM'93, v. 23, N 4.

MATHEMATICAL MODEL OF AN ISDN DIGITAL SWITCHING SYSTEM

Ignatiev V., Kolbahev M.
St.Petersburg University of Telecommunications
Department of Information Control Systems
61, Mojka emb. tel. 213-3409

Abstract. Brief description of mathematical model of an ISDN Digital Switching System (IDSS) is given. It allows to carry out estimation of the main probabilistic and time parameters of system performance, stated in the ITU-T Recommendation Q.543: blocking probabilities and average delays for calls within the system. The considered model is based on an information platform of development of models and methods of estimation of probabilistic and time characteristics of the IDSSs.

Keywords. Digital Switching System, parameters of Grade of Service, models and methods of estimation, control subsystem, controlled subsystem, protocols of interaction, information model, reference loads, traffic, blocking probabilities, probabilities of inadequately handled calls, delays of system response.

Introduction. At present a number of domestic switching systems of the third generation - digital telephone exchanges with modular architecture and distributed microprocessor control - are developed and practically mastered in manufacture in Russia. On the basis of these exchanges the development of digital switching systems (DSS) belonging to the next generation - hybrid digital switching systems for narrow-band ISDNs (ISDNs-N) which use of a circuit switching mode for speech messages and a packet switching mode for data - is carried out. At the same time the research works on creation of experimental DSSs of the fifth generation for broadband ISDNs (ISDN-B) which switch all kinds of messages in a packet switching mode are carried out.

The ITU-T recommendation Q.543 [1] specifies : values of reference loads for various types of lines, digital switching systems and their work modes; a list and values of performance parameters of DSS for various types of calls and signals (signaling messages) and for various kinds of system environment. Main of these parameters are the parameters of Grade of Service (GOS), which determine allowable values of probabilities of inadequately handled calls, average delays of system response on input signals (messages) and delays, which should not be exceeded with probability 0,95.

Reference loads and the GOS parameters given by the recommendation can be used by the manufacturers at designing IDSS and Administration at an estimation of the design of specific system or for comparison of the various designs with the purpose of a choice of the design for potential use. However the methodologies, described in the annexes A and B of the recommendation Q.543 do not give opportunities the manufacturers and Administration to calculate the specified GOS parameters for the concrete IDSS design.

Thus, there is practical need for development of mathematical model of an IDSS and numerical method based on it which allow to estimate GOS parameters, taking into account of the main structural and function features of equipment and software of the system, and also probabilistic and time characteristics of its environment, including the users.

In a general case the development and practical usage of mathematical methods for an estimation Probabilistic and Time Characteristics of information system provide for:

making up an Information Platform of Development;

development of the mathematical or simulation models of an IDSS based on Information Platform of the Development;

development of numerical methods and software for an estimation Probabilistic and Time Characteristics of the system on the base of its mathematical or simulation models;

preparation of source data for the specific system, necessary computing and the analysis of results.

In the present paper the brief description of mathematical model of IDSS developed by the authors satisfying this practical need is given. This model is based on an information platform of an IDSS, and results [2,3], received by the authors earlier.

1. Information platform of development

At making up the Information Platform of the Development of an IDSS oriented to the problem of an estimation of its Probabilistic and Time Characteristics and containing full information that is necessary for such an estimation the system is subdivided on a control subsystem and a controlled subsystem. The first one carries out functions of supervision, control and coordination of functioning of objects of the IDSS which provide the establishing of connections between users, the second one - contains objects, which are used by a control subsystem for providing required connections.

Information models of the control and controlled subsystems and also an information model of protocols of interaction of an IDSS with other elements of a network and with users are used as elements of the Information Platform of the Development. In the basis of developed models the principles of a constructing of appropriate models for TMN (recommendation ITU X.720, X.721 and X.722) are put. Herewith features of a construction and function of digital switching systems and purpose of the given research were taken into consideration. The information models of the control and controlled subsystems give the description of a list of physical and logic resources, their parameters and methods of access to them. The examples of resources of the control subsystem are microprocessors and call handling programs. Main parameter of microprocessors is their speed from the point of view of calculation of the Probabilistic and Time Characteristics of the system. Key parameters of the programs are an average number of the operators carried out by them at implementation of the call handling function and their priorities in relation to other programs of the same microprocessor. Access to resources of the control subsystem is carried out with use of waiting discipline.

The information model of a controlled subsystem describes resources which are seized by the control subsystem during call handling. The examples of such resources are timeslots of PCM trunks for ISDN-N, service circuits and also memory areas to store the information on current statuss of calls in progress. Major parameters of resources of a controlled subsystem are their amount and allowable number of seizure attempts and a feature - access to them according to loss discipline as a rule.

At constructing the information model of the protocols of interaction it is supposed that the users any moment generate the demands on connections and also other input signals which come at an IDSS in the form of the input signaling messages and require performing these or those actions of call handling. The IDSS forms the answer signals as a reply to the input signals and transmit them in the form of the output signaling messages to users. These answer signals are demands of the system on the new information from users that it is necessary to continue further call handling.

Thus, the process of interaction between users and the system is described as process of interchange by the signaling messages and represented in the appropriate information model as a sequence of actions carried out by them within time intervals between moments of appearance of new signals. With the purpose of the present research the process of interaction can be subdivided into two groups of stages.

The beginning of stages of the first group is moments of receiving of the input signaling messages and their termination is moments of output of the appropriate answer signal messages. All actions related to call handling during these stages are carried out by a control subsystem. After identification of a kind of input signaling message the control subsystem accomplishes a choice of a kind of the answer signaling message that should be given out in reply. The kind of the answer signaling message depends on a resource status (free or busy) of a controlled subsystem. After the choice of the answer signaling messages the control subsystem occupies and (or) exempts resources of a controlled subsystem, changing thus its status. At the new status of the controlled subsystem the chosen answer signaling message is transmitted to an environment of the IDSS. The stages of this group are stages of transitive statuses of a application process. We shall name their stages as transition stages.

The beginning of stages of the second group is moments of output the signaling messages and their termination is moments of receiving of the input signal messages. On these stages the status of the controlled subsystem remains without any change, so we shall name these stages as stable stages of the application process.

Each transition stage is implemented with the help of a number of programs which can be carried out by different microprocessors and interchanged internal messages, so the stage is subdivided into a number of phases. Just as it was made at definition of stages a set of phases is subdivided into two groups: processing phases and transmission phases.

The processing phases are initiated by the internal input messages. Call handling on each processing phase is performed by one of the programs related to input, output or processing of the information. Transmission phases are initiated by the internal output messages and related to transmission of internal messages between call handling programs located in the same microprocessor or in different microprocessors which communicate each other via a local transmission network.

The information model of the protocols of interaction is represented at two strata. At the first strata IDSS is considered as "black box" the process of functioning of which is reduced to an interchange of the input and output messages with environment via signaling channels and, accordingly, is given in the terms of the protocols of a network level. A sequence of changing of stages at first strata of the description conveniently is set through the tree directed loaded graph $G_{c.h.p.}$ with two kinds of nodes:

$b_j(j=\overline{1,N})$ which correspond to transition stages and

$a_i(i=\overline{1,M})$ which correspond to stable stages.

The load on the b type node sets for the appropriate transitive status a stochastic restriction $F_j(u)$ on response time of IDSS on the input signaling message initiating this status. In a general case it is supposed that

$$F_j(u)=1-\sum_{m=1}^{m_j}\eta_{jm}e^{-k_{jm}\gamma_{jm}u}\sum_{n=0}^{k_{jm}-1}\frac{(k_{jm}\gamma_{jm}u)^n}{n!}.$$

It allows by the appropriate choice of parameters $m_j,\ \eta_j,\ k_j,$ and γ_j to approximate by function $F_j(u)$ any distribution function.

The load on the a type node sets for the appropriate stable status of application process a probability and an average time of generating by a user that or other the input signaling messages.

At the second strata each transition stage is represented as the sequence of change of cooperating phases of receiving, processing and transfer of messages which is described in the form of a graph $G^{(j)}_{t.s.}$ designed similarly to the graph $G_{c.h.p.}$. Nodes $b_{j,k}(k=\overline{1,N_j})$ of this graph correspond to the processing phases and nodes $a_{j,l}(l=\overline{1,M_j})$ correspond to phases of transmission of the messages.

The load on nodes $b_{j,k}$ gives types of resources of a control subsystem which are used during call handling on the appropriate processing phase and types of resources of a controlled subsystem which are seized and (or) released by the control subsystem on this phase.

The load on nodes $a_{j,l}$ gives average time of transmission of the appropriate internal message.

The Information Platform of Development as a whole including three submitted above information model the relations between which is established through a system of graphs $G_{c.h.p.}$ and $G^{(j)}_{t.s.}$ $(j=\overline{1,N})$ describes IDSS as large real time information system providing an interchange by the signaling and information messages between the users according to the accepted protocols. Herewith the presence of resources necessary for functioning IDSS at various stages of process of interaction of the users with the system and with each other is taken into consideration. The Information Platform of Development systematizes and represents in a formalized form the information necessary for estimation of the Probabilistic and Time Characteristics of an IDSS.

2. IDSS mathematical model

In a basis of developed mathematical model of an IDSS the use of idea of probabilistic graphs applied at calculation of blocking in multistages switching networks lays. According to this approach the problem of calculation the probabilistic and time GOS parameters of an IDSS is reduced to a problem of finding out probabilistic and time characteristics of paths in the graphs $G_{c.h.p.}$ and $G^{(j)}_{t.s.}$ $(j = \overline{1,N})$ from root nodes to terminal nodes.

The mathematical model of an IDSS has hierarchical structure. The mathematical model of each level represents analytical functional dependencies of characteristics of the given level upon characteristics, described by the model of the following lower level, i.e. the last model is embedded into the first one.

Taking into account a restriction on a size of the published paper, elementary results, necessary for understanding of the main idea of constructing the IDSS mathematical model, are only given below.

The mathematical model of an IDSS at the first level gives the following main functional dependencies:

$$P^{sys}_q = \sum_{k \in \omega^{sys}_q} p^a_k / \alpha_q \, , \; p^a_k = \prod_{\substack{(i,d)\in\mu^a_k \\ (j,r)\in\mu^a_k}} p^{ab}_{id}\, p^{ba}_{jm} \, , \; q = \overline{1,Q};$$

where P^{sys}_q - the probability of inadequate q type call handling;

p^a_k - the probability of achieving of the stable stage k by a call during handling (the node a_k of the graph $G_{c.h.p.}$);

α_q - a fraction of calls of the q type in a total flow of incoming calls;

ω^{sys}_q - a set of numbers of terminal nodes of the graph $G_{c.h.p.}$, which correspond to stages of inadequate termination of the q type call handling; p^{ab}_{id},

p^{ba}_{jm} - probabilities of transitions along branches with the order numbers d and m , which go out accordingly from nodes a_i and b_j of graph $G_{c.h.p.}$;

μ^a_k - a path in the graph $G_{c.h.p.}$ from the root node a_0 to the node a_k.

The transition probabilities p^{ab}_{id} depend upon the probabilistic and time characteristics of IDSS environment, whereas the transition probabilities p^{ba}_{jm} depend upon probabilities of call blocking because of absence of free resources of the necessary types at the j transition stage of a call handling process $(j = \overline{1,N})$ and also upon probabilities of excessive delay of a system response at this stage on incoming signals (signaling messages).

The mathematical model of an IDSS at the second level gives functional dependencies for these probabilities. So, probability P_{jr}^{block} of blocking a call at the j transition stage for the lack of requested n_{jr} free units of the r type resource ($r = \overline{1,R}$), access to which is performed in a mode with losses of the requests and the seizure time of which has any probabilistic distribution , can be found under the first Erlang formula (B-formula by Erlang): $P_{jr}^{block} = E_{V_r - n_j}(Y_r)$ where Y_r - an offered traffic intensity for the r type resources.

For probabilities P_{jm}^{delay} of excessive delay of the system response at the j transition stage causing transition of the system to the m stable stage (the transition from node b_j of the graph $G_{c.h.p.}$ along the branch m) using the assumption that the allowable delay of response is stochastic value with a distribution described by a mix of Erlang distributions, the following expression is valid:

$$P_{jm}^{delay} = 1 - \sum_{k=1}^{k_j} \pi_{jk} \sum_{i=0}^{i_{jk}-1} (-1)^i \frac{s_0^i}{n!} \tilde{t}_{jm}^{(i)}(s_0),$$

where $s_0 = i_{jk} \gamma_{jk}$;

$$\sum_{k=1}^{k_j} \pi_{jk} / \gamma_{jk} = T_j^{max};$$

T_j^{max} is a maximum allowable delay of system response at the j transition stage according to the recommendation Q.543;

$\tilde{t}_{jm}^{(i)}(s_0)$ - a value in the point s_0 of i-th derivative of Laplace-Stiltjes transformation (LST) for a distribution function of delay of response of the system at the j transition stage, that was finished by transition to the m stable stage.

The mathematical model of an IDSS at the third level uses the graphs $G_{t.s.}^{(j)}$ ($j = \overline{1,N}$) and gives functional dependencies for the traffic intensity Y_r, LST $\tilde{t}_{jm}(s)$ and their derivatives in the point s_0, and also for average delays $\overline{T}_{jm}$ of response of the system. s_0, for the transformation $\tilde{t}_{jm}(s)$ we have:

$$\tilde{t}_{jm}(s) = \prod_{b_{jq}, a_{jl} \in \mu_{ji(m)}^{b}} \tilde{t}_{jq}^{b}(s) \tilde{t}_{jl}^{a}(s),$$

where $\mu_{ji(m)}^{b}$ - the path in the graph $G_{t.s.}^{(j)}$ from the root node b_{jl} to the terminal node $b_{ji(m)}$ corresponding to the branch m that goes out from node b_j of the graph $G_{c.h.p.}$;

$\tilde{t}^b_{jq}(s)$, $\tilde{t}^a_{jl}(s)$ - LSTs for distribution functions of delay of a call (signaling message) accordingly on the processing phase q and the transmission phase l of the j transition stage.

The mathematical model of an IDSS at the f ourth level represents a set of functional dependencies of transformations $\tilde{t}^b_{jq}(s)$, $\tilde{t}^a_{jl}(s)$ and appropriate average delays $\overline{T}^b_{jq}$, $\overline{T}^a_{jl}$ upon hardware and software parameters of a control subsystem of an IDSS.

So, at deriving of functional dependencies for transformation $\tilde{t}^b_{jq}(s)$ it was supposed that the processing phase q of the j transition stage is carried out by the control unit (processor) of the type n_{jq} with the use of the program of the type m_{jq} .The program has the priority $p_{m_{jq}}$ and is related with the resource of the type $i_{m_{jq}}$ access to which is performed in a mode with waiting for a free unit of the resource of this type. The analytical expression for transformation $\tilde{t}^b_{jq}(s)$ includes as a component the LST $\tilde{w}_{ni}(s)$ of a waiting time distribution function of the i type resource unit. For the LST $\tilde{w}_{ni}(s)$ at a constant seizure time t_i of the i type of resource we have:

$$\tilde{w}_{ni}(s) = \frac{s(v_i - Y_i)(Y_i)^{v_i}}{Y_i^{v_i} e^{-s}} \prod_{k=1}^{v_i-1} (1 - \frac{s}{s_k}),$$

where s_k are roots of the equation $Y_i^{v_i} e^{-s} - (Y_i - s)^{v_i} = 0$, $Y_i = \lambda_{np} t_i$; λ_{np} is a rate of coming requests to the program with the priority p carried out by the n type processor.
The set of mathematical models as a whole represents a large dimension system of nonlinear equations for the decision of which an iterative method can be used.

Conclusion. The developed mathematical model of an ISDN Digital Switching System allows to carry out calculation of the main GOS parameters, to make clear dependency of the probabilistic and time characteristics of a call handling process upon parameters describing their logic and dynamics when the circuit switched mode and packet switched mode are used and to estimate mutual influence of these characteristics on each other. The mathematical model of an IDSS suggested by the authors is applicable for an estimation of system performance for a wide scope of DSSs for ISDNs-N and ISDNs-B the architecture and function process of which can be described by information models.

References

1. ITU-T Recommendation Q543.- ITU, 1994.
2. V.Ignatiev, M. Kolbanev. Calculation method for the probabilistic and time characteristics of ISDN-N switching system.- Sredstva svyazi (Communication means),1990, v.4, pp.75-79.
3. V. Ignatiev. Characteristics calculating method for an ISDN Basic Access System.- Proceedings of International Conference on Informational Networks and Systems (ICINAS-94).- St.-Petersburg, 24-28 October 1994, pp. 199-205.

STRUCTURE MANAGEMENT IN LOCAL INFORMATION NETWORKS

Chugreyev O., Verzoun N.
St.Petersburg State University of Telecommunications
Department of Information Control Systems
61, Moyka embankment, tel. 213-3409

Keywords. Integral local network, management informational system, probable-temporary and cost characteristics, uncompressed continuous information, multicriterion choice of set optimum variants, set of Pareto.

One of the main functions of administrative management information networks is management of configuration of their structure. In this work object of research is a subsystem of reference model of opened system interaction of local information network with integration of services, which contain two bottom levels: physical and channel. The last one consists from sublevels of logic channel management and access by transfer environment. Configuration structure management consists in choice of protocols set and its parameters; topological structure, depending on external effects (flows of information, network geography, number of users) and requirements to services quality during network functioning. The structure reconfiguration on the basis of choice problem of decision should be made at change of specified effects and requirements.

The structure of multichannel integral local network is represented in a general kind by a vector

$$W=[W_1,\dots,W_j,\dots,W_J];\ j=\overline{1,J},$$

where subvektor $W_j=[Y_j,X_j,E_j,A_j]$, $j=\overline{1,J}$, displays the structure of j subnetwork ILN, Y_j is the type of topological structure, X_j - type of protocols, E_j - type of element base and A_j - type of administrative management system of j subnetwork.

At construction of system mathematical model from microdescription of physical ILN it is expediently to pass to its macromodel, submitted by system model with W structure, possessing by U input and V output.

For ILN the input of the model is submitted by a vector

$$U=[U_1,\dots,U_j,\dots,U_J];\ j=\overline{1,J},$$

where U_j, $j=\overline{1,J}$, is connected with set external parameters for j network, for example, number of N_j stations, sequence $f_{aj}(n_{aj})$ of distribution of n_{aj} intervals between incoming messages (applications), their λ_j intensity, k_j length, messages of j type, coordinates of accommodation of j subnetwork users on the territory of object and so on.

The output of the model is submitted by a vector

$$V=[V_1,\dots,V_j,\dots,V_J];\ j=\overline{1,J},$$

where V_j, $j=\overline{1,J}$, displays probable, probable-temporary and cost characteristics of ILN j subnetwork. We shall notice that the components W_j, U_j, V_j for some subnetworks can coincide for the considered network. So in view of that in multichannel ILN a unified topological structure is used $Y_j=Y_i=Y$, $j=\overline{1,J}$. We shall hereinafter consider also that all subnetworks are realized on identical element base $E_j=E$, $j=\overline{1,J}$, and use a uniform management system $A_j=A$, $j=\overline{1,J}$.

We are addressed to by vector making further. A communication system of ILN, containing two bottom level of program structure in accordance with the primary model of opened system interaction is considered in given work. Then we have:

$$X_j=[x_{j1c}, x_{j1p}, x_{j2l}, x_{j2a}];\ j=\overline{1,J},$$

where x_{j1c}, $j=\overline{1,J}$, displays the type of transmission environment, x_{j1p} - the type of physical level, determining V_c speed of transfer in environment, x_{j2l}, $j=\overline{1,J}$ - type of protocol of second level logical channel and x_{j2a} - type of the protocol of sublevel access.

We shall further consider that uniform means of transfer and the protocol of physical level are used in monochannel ILN and three types of information packages, such as data D, compressed continuous information C, uncompressed continuous information N, are transferred, that is J=3, and $j=\overline{D,C,N}$. And then

$$X_j=[x_{1c}, x_{1p}, x_{j2l}, x_{2a}];\ j=\overline{D,C,N},$$

We shall designate through W' set presented to choice of variants of w structure W, $w\in W'$ produced by variants of presented types Y_{v_y}, $v_Y=\overline{1,N_y}$; $x_{1c}v_{1c}$, $v_{1c}=\overline{1,N_{1c}}$; $x_{1p}v_{1p}$, $v_{1p}=\overline{1,N_{1p}}$; $x_{j2l}v_{2l}$, $v_{2l}=\overline{1,N_{2l}}$ and $x_{2a}v_{2a}$, $v_{2a}=\overline{1,N_{2a}}$.

$V=f(U,W')$ for entered system module.

It is obviously that if there is a set of U external parameters then for presented set W' can be found V - probable, probable-temporary and cost characteristics.

Now we shall consider a problem of multirange of optimum choice of w variants set from presented W' set, set by N_y types of topology, N_{1c} types of transfer environment, N_{1P} types of physical level, N_{j2l}, $j=\overline{D,C,N}$ types of logical channel protocols and N_{2a} types of protocols of access at fixed element base and management system.

As components V_j, $j=\overline{D,C,N}$ we shall choose: probable-temporary characteristics [1] : $\overline{t_{qd}}$, $\overline{t_{qc}}$, $\overline{t_{qn}}$ - average times of delays of D, C and N packages delivery, $\overline{P_{qd}}$, $\overline{P_{qc}}$, $\overline{P_{qn}}$ - probabilities of duly delivery packages of the messages for the types of packages of information, indicated above. We shall choose P_r - a profit on ILN operation as a cost characteristic

$$P_r = I - E_a t - C,$$

where I is an income of ILN operation, $E_a=0{,}15C$ - annual operational charges on ILN maintenance, t - quantity of operation years, C - capital costs of ILN introduction,

$$I = I_p + 12\, t\, U_p + I_a t,$$

where I_p is an installation payment, U_p is a user's monthly payment, I_a - revenue of payment for information transfer

$$I_a = P\, F\, (N_d k_d \lambda_d + N_c k_c \lambda_c + N_n k_n \lambda_n),$$

where P is byte tariff for information transfer, F is a factor, taking into account the number of \working days in a year, recalculation bit in kilobytes and other, k_d, k_c, k_n - are the lengths of packages informational parts for three subnetworks, N_d, N_c, N_n are the number of stations in every subnetwork, λ_d, λ_c, λ_n are intensity of messages flows from one transfer subnetwork station D, C and N accordingly,

$$C = D_m C_k + N_d C_d + N_c C_c + N_n C_n,$$

where C_d, C_c, C_n is the cost of every subnetwork one station, D_m is the length of the connection channel (in meters), C_k is the cost of one meter of cable.

Then the problem of multicriterion choice of set optimum variants is reduced to search

$$C(W') = \underset{w \in W'}{\arg\max}\, P_r(w), \qquad C(W') = \underset{w \in W'}{\arg\min}\, \overline{t_{qd}}\,(w),$$

$$C(W') = \underset{w \in W'}{\arg\min}\, \overline{t_{qc}}(w), \qquad C(W') = \underset{w \in W'}{\arg\min}\, \overline{t_{qn}}(w),$$

at restrictions $\underset{w \in W'}{\overline{P_{qd}}(w) \geqslant \overline{P_{qd з}}}$, $\quad \underset{w \in W'}{\overline{P_{qc}}(w) \geqslant \overline{P_{qc з}}}$, $\quad \underset{w \in W'}{\overline{P_{qn}}(w) \geqslant \overline{P_{qn з}}}$,

where $\overline{P_{qd з}}$, $\overline{P_{qc з}}$, $\overline{P_{qn з}}$ - are allowable significance of probabilities of duly delivery of packages D, C and N accordingly.

Two-stage procedure of optimum choice is offered. At the first stage we divide multicriterial choice of optimum set, using a principle of decomposition, on 3 two-criterial concerning P_r and $\overline{t_{qi}}$ procedures. We find set of optimum variants for each subnetwork, using Pareto proportions. Set of Pareto contains those elements of w* of set W', for which top section is empty, concerning element w*.

At the second stage the association of the decisions of optimum choices, made at first stage , is performed. Thus by using crossing operation to found sets of Pareto for each subnetwork we shall separately receive set of optimum variants for all network which will consist of ILN construction variants, fallen in sets of Pareto for separate subnetworks simultaneously.

References

1. Chygreyev O., Krotova N. Analysis of Packet Transmission Continuous Information with Compression of Local Networks with Integration of Services // Proceedings of International Informatisation Forum III International Conference on Information Networks and Systems. ICINAS - 94. St. Petersburg, October 24-28, 1994, pp. 165-168.

RESEARCH AND ACADEMIC COMPUTERNET OF THE INTERNATIONAL UNIVERSITY OF KYRGYZSTAN AND ITS APPLICATIONS

Valery P. Zhivoglyadov, Svetlana A. Yampolskaya and Vitaliy A. Kuznetsov
International University of Kyrgyzstan
255, Chui Prospect
Bishkek 720047, Kyrgyzstan
e-mail: zhi@iukr.bishkek.su

Abstract. The goal of the report is to discuss some new basic approaches to the development of information systems and solving research problems of models identification, prognosis and decision making using the computernet of International University of Kyrgyzstan as example. The computer net created in the International University of Kyrgyzstan IUK is directed fowards supporting research and academic activity in the field of Management Information System, Systems Analysis and Design, Control Theory, etc. In particular researches and developments in the field of administrative informatization have been carried out using this system. The conception of administrative information has been worked out and given to the Administration of the President of Kyrgyzstan. With the help of CASE methodologies, design by prototyping, there has been developed an Information System of Financial Control of the city administration. Water-level of Issyk-Kul lake prognosis models was proposed.

Keywords. Information system, system analysis, design by prototyping, administrative informatization, prognosis models, Internet, network, local area network.

Introduction and General Problem Statement. The problems of systems analysis and design, development of effective Management Information Systems and Decision Support Systems have attracted the attention of many researchers and attained impotance in modern fields of information science and control [1-4]. The traditional approach to algorithms and information systems development is based on the premise that exact information processing needs and can be identified during the early stage of a development project. Our investigations focus on some problems dealing with essential uncertainties in the initial requirements.

The problem statement is as follows:

- to create IUK research network,
- to accumulate procedures, algorithms, CASE tools for providing short-term development and implementation of systems,
- to develop and study applied information systems for management and prognosis.

We are developing a hybrid approach that combines methodologies of structural design, design by prototyping, and Computer-Assisted Software/System Engineering. The authors of the paper are participating in some applied research projects during 1996-1997.

The paper is organized as follows. In the next two sections, the description of IUK computer network system and types of activities are presented. Then, some applications of this approach to investigations of different problems and some results are given.

Description of IUK Network System. The computer net created in the International University of Kyrgyzstan IUK is directed towards supporting research and academic activity in the fields of Management Information Systems, System Analysis and Design, Control Theory, etc. IUK Computernet includes three local area networks (LAN) using multimedia systems, CD-ROM, VideoPhone, scanners, net printers: LAN of 25 computers PENTIUM-100, PENTIUM-133, LAN of 25 computers COMPAQ, small net of POWER MACINTOSH (6 computers), computer-assisted presentation system (CAPS).

The IUK-network host providing stable connection according TCP/IP (AsiaInfo - leased line) and UUCP (Kyrgyzstan Freenet - switched line). The server of the University has WWW-server and FTP-server with address: iukr.bishkek.su.

The structure of the IUK-network is shown in Fig 1.

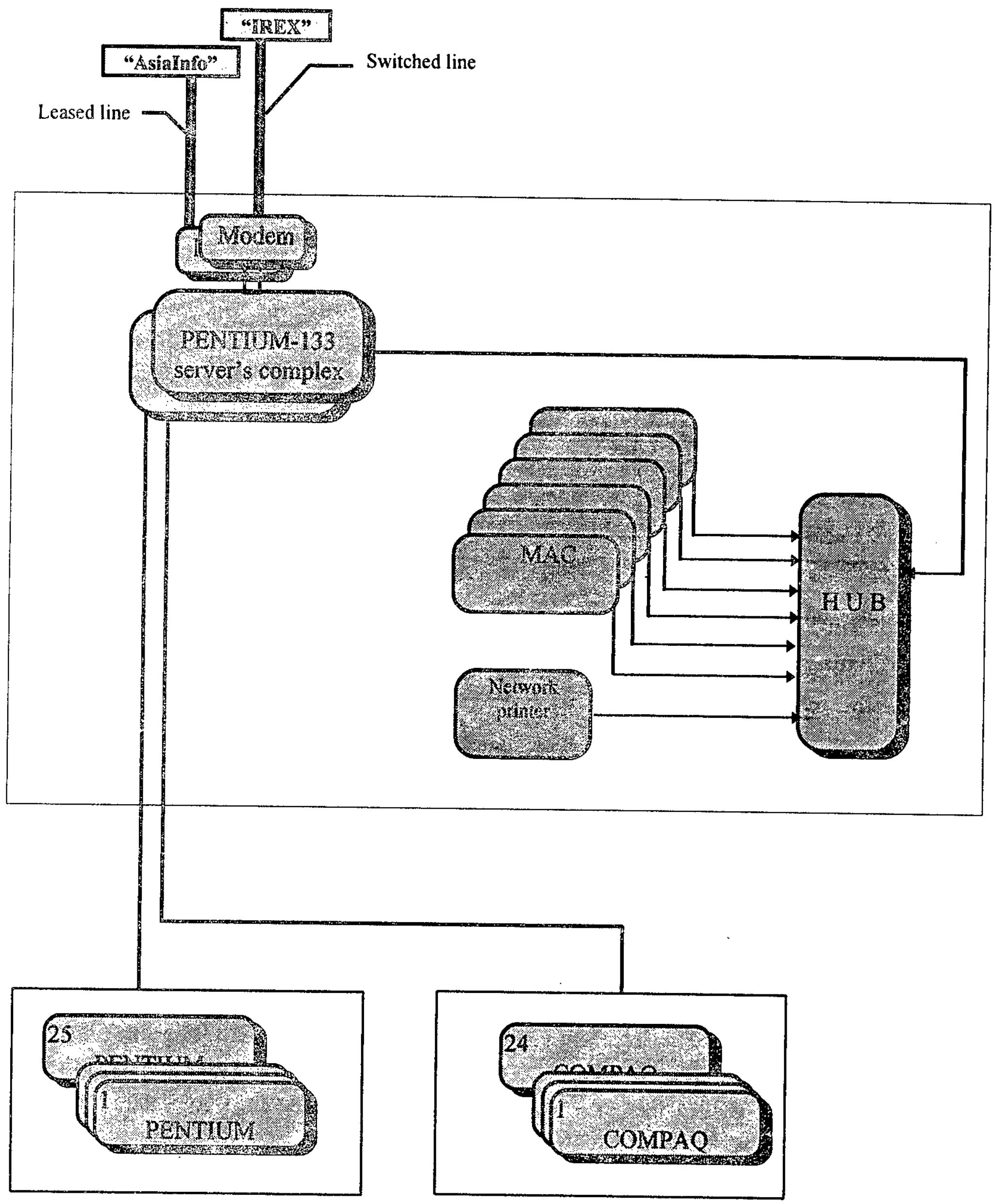


Fig. 1. The structure of the IUK-network

Three computer LANs are connected to the server's complex. The first net is the laboratory of the office systems. The second and third nets are the LAN of two computer laboratories . There are about 60 computers connected to the server's complex.The IUK-network provides investigations and development of information systems and control systems (basic stages and types of activities):

- planning of investigation and preparation of project proposals,
- developing and using distributed data bases,
- experimental information processing,
- modeling and simulation of processes and systems,
- multicriterial optimization of parameters for control systems,
- development of information systems for governmental management and expert systems,
- supporting group work in the IUK-network,
- using telecommunication channels for information exchange for searching nesessary scientific information in the WWW and FTP and electronic journals, sending information to teleconferences,
- preparation of research reports and presentations using computer projectors, etc.

The main directions of academic activity are:

- Business computer information systems,
- International computer nets, Internet,
- Control theory and control systems
- Information science,
- Computer graphics,
- Management information systems,
- Data bases
- Office systems, etc.

The goal of IUK-network is to provide scientists and students with a comprehensive tool for information processing and systems investigations. We use the design by prototyping technique (DPT) and Computer-Assisted Systems Engineering (CASE) as powerfull analysis, design, and development tools and the Systems Development Life Cycle (SDLC) as a general methodology for managing complex applied software project. These CASE tools are designed to help scientists and students meet using contemporary approaches for systems study and building effective Information Systems.
The electronic transparencies enable the dynamic presentation of graphics, text, and animation.

The user's access to the IUK-network. At present, the International University of Kyrgyzstan network allows the use of all resources of Internet in educational and scientific research processes. Currently, there are about 80 registered users in the IUK-network, they are the researchers, scholars, administrative workers, and students of IUK.
Access to the net is possible only by permission of the System Administrator. Each user must be registered in the system. After registration, each user gets his own User Identificator (UID), password for access to the system and some disk space on the server's hard drive, on which he keeps his information. Each user has his own rights in the system. Each user can do the following:
- use E-Mail;

- subscribe on USENET's conference (only for educational and scientific purposes);
- use on-line mode for connection to remote servers;
- use FTP only in LAN, for transport of their files from their peronal computer to the server.

A friendly interface has been created for nonprofessional users, which enables them to easily work with Internet services. Professional users have the possibility to connect to remote servers, to send and receive files, and different kinds of information.

The IUK-network server supports access to the Internet by switched line from any place in the city. Any users of the IUK-network can work with Internet from home computers but only from 19:00 till 7:00. The traffic size, in megabytes, during this time is limited.

The administrative group provides the stable working of all of the system's parts, especially the server's complex. They make sure that all the servers and clients PCs are connected and talking to each other at all times. The security problem is also very important. They use advanced techniques, such as firewall and shadow's password to protect the system from different kind of attack.

Some Applications. Research and development has been carried out using this system. Some examples are:

- The concept of administrative informatization for Kyrgyzstan has been worked out and given to the State Administration;
- Analysis of the information processing requirements of local government management ;
- Experimental data processing, modeling and simulation of processes and systems;
- Information systems development , using telecommunication channels for information exchange for searching nesessary scientific information in the WWW and FTP.

We consider these examples in further detail.

Concept of Administrative Informatization. The government of Kyrgyzstan is now required to ensure efficient, comprehensive and flexible public administration. It is necessary to have full and exact information about the economic and social spheres. In order to meet these requirements, "The Concept of Promoting Administrative Reforms through Informatization" has been established. The main goal is to promote informatization in public administration by applying results of new information and telecommunications technologies to various fields of public administration. Within the Concept, administrative informatization is regarded as a compex of phenomena which have important consequences for the functioning and organization of public administration.

The Concept of administrative information has been given to the Administration of the President of Kyrgyzstan.

Research project "Data processing requirements analysis of the Karakol city administration". The principal purpose of this research project was to analyze the needs and requirements of the Karakol City Administration in data processing using new information technologies, concentrating on improving practice, efficiency and manageability in the areas of finance, budget and taxation .

It completed a functional - structural analysis of the city administration and an analysis of possibilities of information systems development with estimation of necessary resources, costs and terms for development and implementation of systems and their practical usefulness.

We developed prototype functional structures for several selected information systems, proposals for software and database selection and development, investigation of the design method based on the application of "fast prototype" and modern software, which allow considerable reduction of systems implementation and development terms.
Information given in the research project report can be used for motivated selection of projects and plans of computer information systems implementation for the Karakol City administration.

Structural Design of Information Systems. Structural Design is one of the most important stages of the System Development Life Cycle (SDLC), that is formation of the functional structure of the system.
Initial reasons for execution of this stage are the following:
- determination of the end-users and departments where the system is to function ;
- clarification and analysis of requirements for data processing users and system requirements;
- determination and analysis of the information flow ;
- determination of the list of functions and development of information systems providing these functions .

In the previous chapter the necessary information for design has been discussed. Considering limitation of resources & MFM project terms we propose using of "design by prototyping" (design with the use of "fast prototype"). In this case the structural design of "fast prototype" is offered first in accordance with CASE methodology. Then, after the "fast prototyping" test performance, definition and modification of the end-users requirements, the design of structures of databases, subsystems & information systems on the whole is being executed .

Information System of Financial Control (ISFC). ISFC-I (see Fig. 2) was developed and implemented for the Karakol city financial department by the IUK School of new information technologies research group during May-June, 1996. ISFC-I system is responsible for execution of the following functions:
1. Enter and storage of the information about the city departments and organizations, keeping of organizations/departments information guide with the system of codes both in Russian and English.
2. Keeping of economic classification directories:
    - actual in Kyrgyzstan, approved by the KR Ministry of Finance;
    - proposed in MFM project in accordence with requirements accepted by international organizations.
3. Allocation of funds orders data enter, storage of information and delivery of output operative documents for any inquired period for registration of confirmed assignations and costs.
4. Enter, storage and presentation of budget examination list for the current year and introduction of changes in budget.
5. Formation and delivery of budget implementation control reports:
    - for any month;
    - from the beginning of the year;
    - for any requested dates.
6. Protection of information from unautorized access by outsiders.

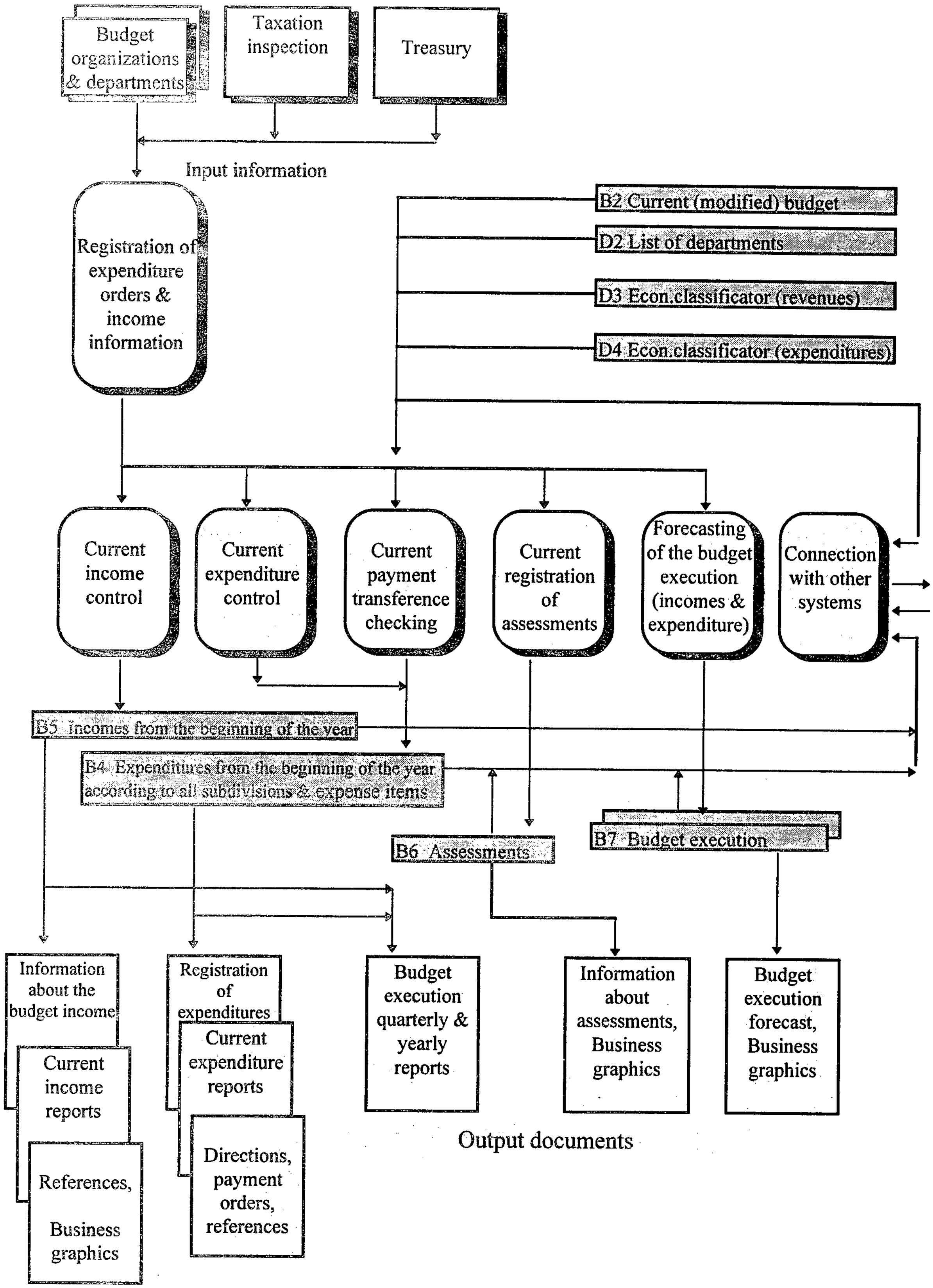


Fig.2. Functional structure scheme of the Information System of Finance Control

Access to tables and possibility of an information system data base adjustment is determined by stipulated level of users access (particular level for a particular user).
The system hardware: COMPAQ PROLINEA 5100E, PENTIUM 100mhz,16Mb RAM,HDD 620Mb.
The system software: applied software of the system is developed using ACCESS 2.0, under the WINDOWS 95, (WINDOWS 3.11) control. The project proposes to implement development of a widened version of the system with high capacity. In contrast to ISFC-I advanced information system of financial control ISFC-A (Fig.2.) includes the following supplementary functions:

- Current budget incoming control.
- E-mail communication exchange among local network users.
- Decision making support using business graphics during preparation and signing of orders for assignation.
- Expansion of information directories.
- Reports of various forms.
- Reports of finance control with quarterly distribution of budget assignations.
- Reports of dual-level scheme finance control (city and regions).
- Aggregation of information by codes of the city departments and economic classifications.
- Exchange of information with other information systems.

Prognosis of water-level of Issyk-Kul lake. Prognosis of water-level of undaugings Issyk-Kul lake was based on the analysis of the water balance dynamic between input (in form of rain or snow) and output (in the form of evaporation) of water resources.
Prognosis algorithm has the following form:

y[s] =F(nn, q[s-1],..., q[s-k], x[s], x[s-1],..., x[s-n], **c**),

where s is discrete time, s = 1,2, ..., ; y[s] is prognosis of water-level; nn is trend;
q[s], q[s-1],..., q[s-k] are the real water-levels for different years;
x[s], x[s-1],..., x[s-n] are the precipation values; **c** is an adjustment parameter vector;
F is a known function.

Long-term hydrologic characteristics were used in statistical hydrological computations. The accuracy of the prognosis models were checked by the following method: prognosis models parameters were calculated using the same learning sets of data - 30 points (hydrologic characteristics at the period from 1933 year to 1963 year) and 40 points (hydrologic characteristics at the period from 1933 year to 1973 year), but examinations of models were made using the examinating sets of data (hydrologic characteristics for the period from 1974 to 1996). The results of computer simulation and prognosis are shown in Fig. 3.

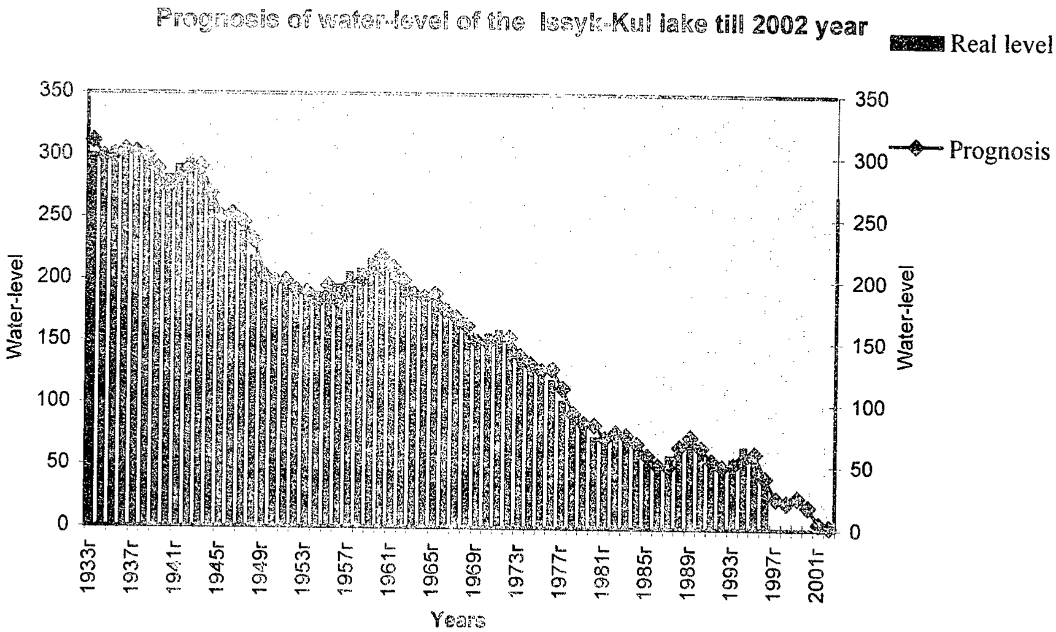


Fig. 3. The results of computer simulation and prognosis

Conclusion. Researches and developments in the field of administrative informatization, Information System design by prototyping, model identification and prognosis models development have been carried out using this system. A friendly interface has been created for nonprofessional users, which enables them to easily work with Internet services.

References

1. Long, L. (1994) Introduction to Computer and Information Systems. Prentice-Hall, Inc.
2. Whitten, J.L., L.D. Bentley, T.I.M. Ho (1986) Systems Analysis and Design Methods. Times Mirror / Mosby College Publishing.
3. Madron T.W. (1992) Network Security in the 90's. Issue and Solutions for Managers. John Wiley & Sons, Inc, New York.
4. Zhivoglyadov, V.P., G.P. Rao, N.M. Filatov (1993) Application of b-operator Models to Active Adapive Control (AAC) of Continuous-time (CT) Plants. Control-Theory and Advanced Technology, Vol. 9, N 1, pp. 127-137

THE TRAFFIC SELF-GUIDANCE FOR MULTIMEDIA COMMUNICATIONS

Valeriy P. Vidomenko
SE "Rubin" RI
Kantemirovskaia st.5, 197342 St.Petersburg, Russia.
E-mail: rubin@sovan.com.

Abstract. It is shown in this paper, that modern technologies, including fiber optics and photonics, cannot eliminate the disparity between today's standard network topologies and multimedia communications needs of the future. The regularization of network topologies is necessary for this problem solving. The main advantage of regular topologies is the traffic self-guidance. It makes the large-scale network supervisible and comprehensible. The current assortment of network topologies is too miserable to solve this problem. The main part of this paper has focused on two vast classes of new original topologies which provide both the traffic self-guidance and the gradual augmentation of networks within a practically unlimited range.

Keywords. Multimedia applications, network topology, telecommunication, traffic self-guidance, self-organization.

1. INTRODUCTION

Current standard network topologies and communication protocols are not well suited to the multimedia communications needs of the future, despite of non-stop in base technology progress. Standards for construction of worldwide B-ISDNs are permanently being worked out and up-dated in various international standards bodies. Transmission media technology has made a great leap in speeds and reliability of networks, moving from copper wire telephone lines to fiber-based B-ISDNs. Emerging new transport technologies such as Synchronous Digital Hierarchy (SDH) and Asynchronous Transfer Mode (ATM) provide efficient and flexible means to carry diverse information. Photonic technology promises Teraflop computing and Terabit transfer. The synergism of computers and communications has been arisen up to symbiosis in a number of outstanding implementations of the C&C (Computing & Communication) conception.

Communication network topologies are likely the only area to find itself beyond the technological progress. They stay put in ancient times when trees, rings, and busses had children's sizes and there was need of neither large scale networks. At present, the stagnation in topologies area became an obstacle for efficient implementations of the technological achievements of other networking areas, since any real network has to lay on Procrustean bed of a chosen topology.

Such a state of affairs seems especially curious in the light of the system-forming role, network topologies have acted. One can imagine the fundamental significance of network topologies, remembering the burst development of the ultracomputer industry. A single topology of hypercube became a catalyst of the burst which resulted in multiprocessor production of such firms as Thinking Machines, NCUBE, Intel and some others.

The traffic self-guidance, intrinsic in the regular topologies, introduces elements of self-organization which is necessary for large-scale networks operation to do them supervisible and comprehensible independently of their size. Besides, the traffic self-guidance allows to get rid of routing overhead since it makes unnecessary routing tables which demand to concentrate a lot of processor and network resources while increasing scales of networks entailed more and more demands for switch processing power. The growth of networks is much stimulated with multimedia applications development. By present, specific expenditures on switching subsystem exceed ones on cable and transfer subsystem, put together. They come close to civil engineering expenditures of a typical exchange system.

This paper briefly reviews some network problems, induced by multimedia nature of information objects, and their solution ways in a frame of topological approach. Also, main achievements and gaps in regular topology area is analyzed. In the main part of the paper, some new classes of original topologies useful for large scale networks with self-guided traffic are described in comparison with existing ones.

2. NETWORK PROBLEMS OF MULTIMEDIA APPLICATIONS

There is a number of network problems, induced by modern multimedia applications, which may become a challenge in computer networking of the nearest future. It is especially related to multimedia transport systems, including transport and access networks and a protocol stack for OSI layers 2, 3, and 4. The problems originate, firstly, from the multimedia nature of information objects, and secondly, from inability of widely using primitive topologies to support large scale, expandable, logically complex multimedia networks with demanded efficiency within their life cycle.

A multimedia-induced problem results from differences in new application requirements for service, flow control, buffer management and connection management. There is a lot of multicast applications like teleconferencing and telemedicine which integrate several information types such as video, sound, data, still images, and moving pictures under computer control. Different information type characteristics and requirements vary too greatly to support all possible multimedia applications equally well over a large variety of service requirements and all possible network topologies with a single super transport protocol stack. A way to simplify protocols passes through network topologies unification and through network operation on principle of self-organization. The first step on this way is likely to be the traffic self-guidance of regular topologies.

Another multimedia-induced problem arises, resulting from the very large size of a multimedia data object, which may occupy too many blocks in secondary storage. Furthermore, unlike conventional data models, multimedia objects can have a complex logical structure and can be shared among other objects. These factors entail the increasing in number and size of repositories and in complexity of flow control. That, in its turn, increases the network traffic complexity and intensity, requiring the proper increasing of network through-put and routing efficiency, unattainable for primitive topologies.

Definitely, one can relay upon lightwave technology which promises a Terabit capacity. But it should be meant, that photonics firstly demands to overcome a number of principle restrictions, including the power restrictions and the electronic bottleneck. Architectural solutions, usually considered as the best ones, are: a star topology to the power problem, and a banyan-like topology to the bottleneck problem [1]. As a rule, these considerations proceed from a more than scanty background of contestant topologies, usually consisting of ring and bus only. The topologies above put together nearly exhaust the current assortment, hit in the system architect's field of vision. In fact, one has not sufficient reasons to discuss a topological superiority, facing the practically nonalternative choice.

The currently available WANs have mostly a bus topology, and most of current high speed LANs have a ring topology. Moreover, as it were for escalation of the differences mentioned above and thereby to harden the networks integration problem, the bus topology WANs have connection-oriented interface, based on either physical or virtual circuits, meanwhile the ring topology LANs have broadcast-type interface which tend to provide simple connection-less services at the lower protocol layers [2].

As it was noted by P.Fletcher [3], communication development today is inhibited by two kinds of rigidity: a rigidity in network topology, and mental rigidity. Despite of these rigidities, the understanding of necessity to regularize network topologies is expanding as well as their assortment, at any rate in regard to the scientific research realm. The rectilinear grid is used as a base for MANs development [4,5]. The space grid of more complexity, based on bus topology and called Spanning Bus Hypercube, is used for solving multicast problem, urgent for multimedia applications such as teleconferencing, replicated databases, and distributed operating system services [6]. Multicast communications is characterized in three ways: temporal localization, complete accessibility, which is a logical equivalent to complete graph, and finally the benefit of a tree topology due to its routing simplicity. Spanning Bus Hypercube is used as a meta-network to construct multicast trees in it. It is far not a simple problem as tree and complete graph topologies occupy the polar positions regarding to network complexity.

During research in the area of high traffic capacity networks, banyan-based (shuffle-exchange) topologies, including tree as a special case, attract more and more attention [2,7,8]. Sometimes, banyans are proposed for raising performances of communication systems by distribution of process recourses in the ultracomputing style. A number of parallel disc systems has been investigated, ranging from the available Connection Machine`s parallel 40 discs system (Data Vault) till optical disks using multiple beams with the potential for massive data rates. In the frame of alternative approach, namely topological one, a Heterogeneous Multimedia Distributed architecture was proposed [9]. It consists of a number of servers, a global communication network, and a central controller for distributed objects. The multimedia server has a distributed architecture too. But, unlike the global network (a B-ISDN of optical MAN technology on dual bus topology), it is based on a shuffle-exchange topology which is far from primitive. Moreover, each server incorporates two the shuffle-exchange networks. The solution to distribute the server architecture entails the side effect of more irregularity in the topology of global MAN-LAN-server system, which increases waste internetworking expenditures.

The most positive effect, produced by works like [4-9] developing topological approach is in attraction of system architect's attention to new opportunities which the regular topologies open.

3. PROGRESS AND GAPS IN THE REGULAR TOPOLOGY AREA

The first progress in the theory of network topologies is connected with the works by N.G.de Bruin ([10] 1946), C.Clos ([11] 1953), V.E.Benes ([12] 1964) and Ju.P.Ofman ([13] 1965). Their names became names of networks, whose widespread popularity started much later, in the seventies-eighties range only, when the flow of published works in regular topology area grew avalanche-like. The special service belongs to C.Clos. He was the first who has set up and solved the problem of nonblocking topology construction. Reconfigured Benes network is used in current switching fabrics, e.g. in System 12 ITT (Alcatel).

The seventies and eighties are characterized with a plethora of works in theory and applications of layered topologies, banyan-based particularly. Layered (multistage) networks include the majority of ones, widely used in theory and practice, e.g. trees, banyans, Clos-, Benes-, Ofman-networks, and a number of others, including hypercube. As defined by G.R. Goke and G.J. Lipovski [14, 1973], the banyan is a switching network which provide a single path for each pair of opposite terminal ports (network poles). Many independent studies had resulted in banyans of different kind, until C.-L. Wu and T.-Y. Feng [15, 1980] proved those networks to be topologically equivalent. At present, three main types of topologically equivalent banyans are known to differ by the shuffle pattern. They are actually banyan [14], baseline [15] and omega [16].

In last years, an interest has arisen in banyan topology for B-ISDN with ATM. In large scale networks, the banyans are attractive owing to simplicity of the traffic self-guidance and moderate expenditures on the switching equipment. Unfortunately, high blocking and vulnerability, connected with one-path property, confined the application area of banyans, as a rule, in theoretical research. However, the symmetric collage of the two banyans on one common layer (the Benes network), implemented in switching fabric of System 12 (ITT-Alcatel), proved again that nothing is more practical than a good theory.

Though up to now, the Benes network is out of competition on the network topologies market-place (even the ultracomputer GF11 IBM was developed on its base), it has two intrinsic shortcomings: configuration rigidity and blocking of free poles by connections, set before. The both shortcomings is a heritage of the Benes network after its parents - the pair of banyans. Between a single switching element and maximal configurations of the System 12 switching fabric, there are only two complete configurations. The capacity and number of alternative internodal paths for a pair of adjacent configurations differs 8-multiply and complexity (the general number of nodes) differs about 10-multiply. If there is a demand to have configurations with intermediate values of the parameters, then one can to meet it at the cost of the extra switch equipment which will be never in operations. The Clos network has the same configuration rigidity as the Benes network has.

As the analysis shows, the topologies, widely used in the modern network technology and research, do not match to the multimedia communications of the future, since they do not meet the complex of demanded characteristics including unlimited and gradual augmentation of capacity and throughput, reliability, versatility, flexibility and high routing efficiency. Neither ring nor bus, nor star, nor banyan-based topology can support coming multimedia networks. In this connection, the development of new topologies is of paramount urgency.

4. NEW REGULAR TOPOLOGIES

Successive employment of regular topologies in ultracomputers and switching fabrics stimulated numerous attempts to use them in other application areas, including LAN, MAN and WAN construction, neuroinformatics, physical experiments and control of complex objects in energetics, astronautics, aircraft, fleet and the like.

However, the attempts to widen the application area of regular topologies, capable to provide the traffic self-guidance, do not achieve a grate success. Main reason is a poor choice of using topologies and the absence of a general theory to synthesize new ones. In particular, there is a need of a mathematical apparatus, oriented towards research and design of network topologies. This apparatus should be based on strictly defined notions and enable to classify well-known topologies and synthesize new ones.

A number of works within recent years [17-24] enabled to fill up this gap in certain extent, and to enlarge the set of topologies, which provide the traffic self-guidance. The specially developed calculus allows to define formally a layer-homogeneous network by three parameters: the number of the shortest alternative paths, network capacity and its configuration. Two of this parameters may be chosen independently. It gives the opportunity of a flexible response to differences in the task settings of a network topology development or analysis. For example, equated to one the number of alternative paths in the parametric definition of a layered network gives the parametric definition of a banyan.

The new original topologies excel banyan-based networks and, especially, hypercube in the versatility and flexibility of network configurations to a grate extent. They give an opportunity evenly vary network parameters without breaking of configuration completeness and without reconstruction of the starting network. In the same time, the new topologies do not yield to the known ones over the traffic self-guidance simplicity. Fig.1. presents the 3-layred 24×24 quasi-banyan, scaled with one extra layer to increase the alternative paths number (APN). The technique developed in [18,19,20] allows to drive an orthogonal search of a network configuration on the "capacity versus APN" plane, on which the Clos's, Benes's, Ofman's and other banyan-based networks are only some dots few and far between.

Among the other topologies on the "capacity versus APN" plane, the topologies of nonblocking networks has a special meaning. Definition of the generalized Clos network as a collage of two banyan on one common layer significantly enlarges the Clos's class of nonblocking networks. Another class of nonblocking networks, derived analytically, is presented by nonblocking banyans (Fig.2). It should be noted, that the banyan and blocking notions implied to be inseparable up to the present. The self-guidance procedure is common with self-routing on usual banyans.

Up to now, we have considered two-dimensional cut of the set of network configurations. However, there is a third dimension in the space of configurations. To see it, one can consider a two-layered section of multi-layered network as an isolated network (Fig.3). Banyan-based sections are out of interest due to disintegration on primitive components in a form of bipartitely-complete graphs (a special case of bipartite graph). A nonbanyan network section is connected and may be presented as a torus. (Hence, the multi-layered network of the nonbanyan type is a polytorus.)

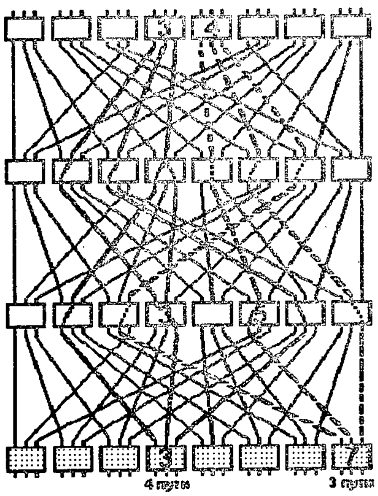


Fig. 1. Quasi-banyan with extra layer

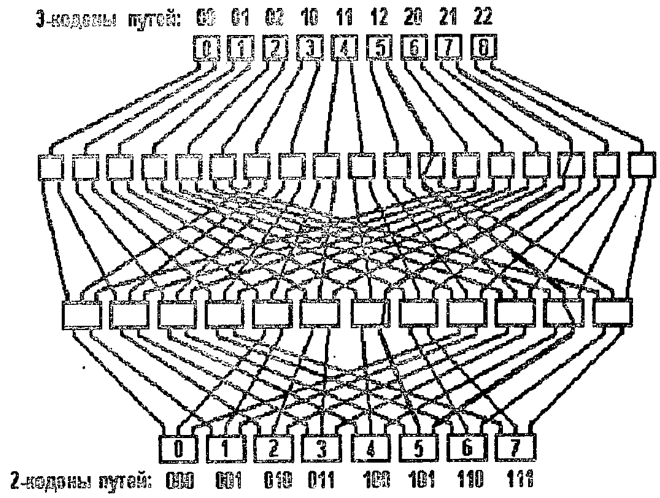


Fig.2. Nonblocking banyan

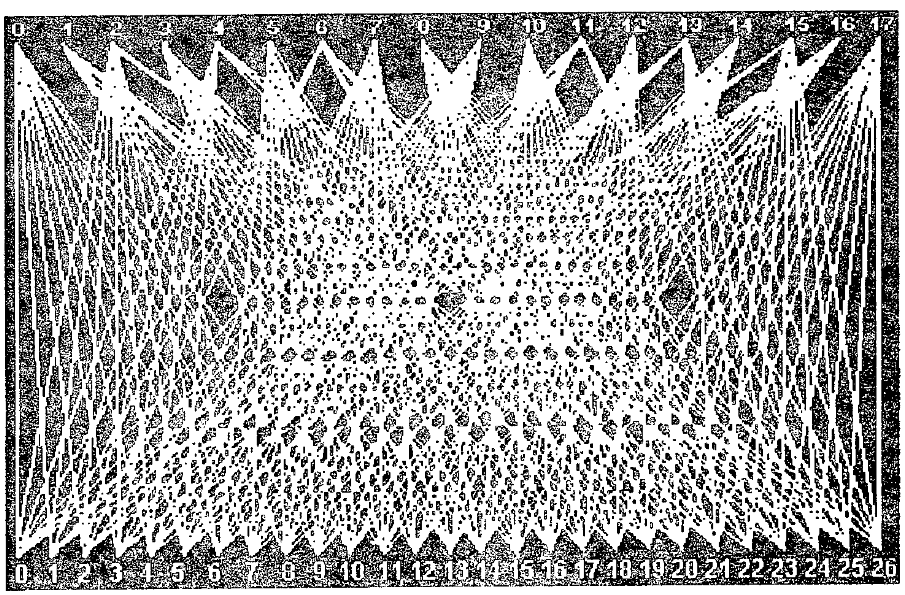


Fig.3. 2-layered network in the unified shuffle form

The theory of two-layered networks is expounded in works [21.22]. Such a network is a product of a simple torus on a bipartitely-complete graph and may by presented in a form of hypergraph as its special case a hypertorus (Fig.4).

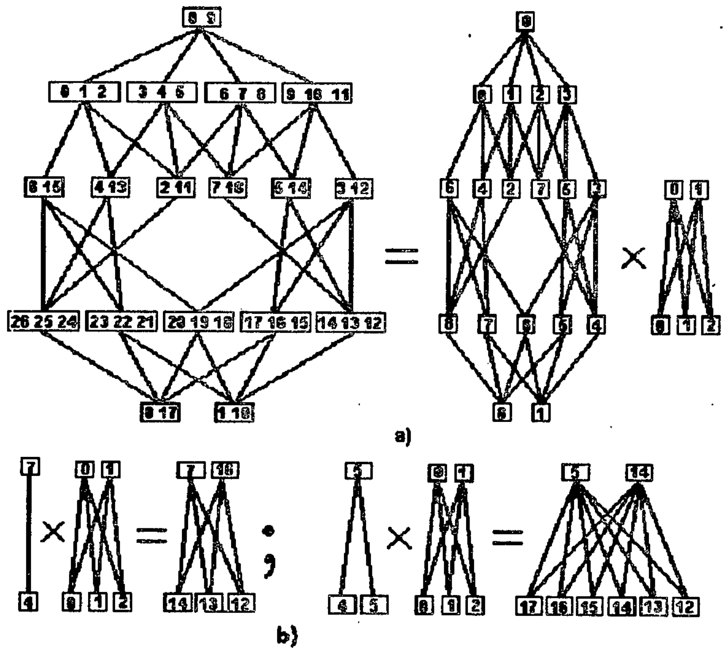


Fig.4. Hypertori algebra (a) and two algebraic examples (b)

The special algebra [22] with numerical operations on parameters was developed to analyze and synthesize two-layered networks. The algebra-based research was undertaken to derive new topologies by means of topologically equivalent transformations of the initial network. These transformations brought the fourth dimension in the space of topologies.

The set of topologically equivalent but visually very different networks offers the more extensive practical choice for system architects. The large two-layered network like one in Fig.3 is too complex for visual analysis. Topologically equivalent transformations may bring it to the more robust form of tori (Figs.5, 6). Its structure is so obvious that may stimulate the system architect's imagination and make his work more effective. The most simple configuration is like a matrix (Fig. 7).

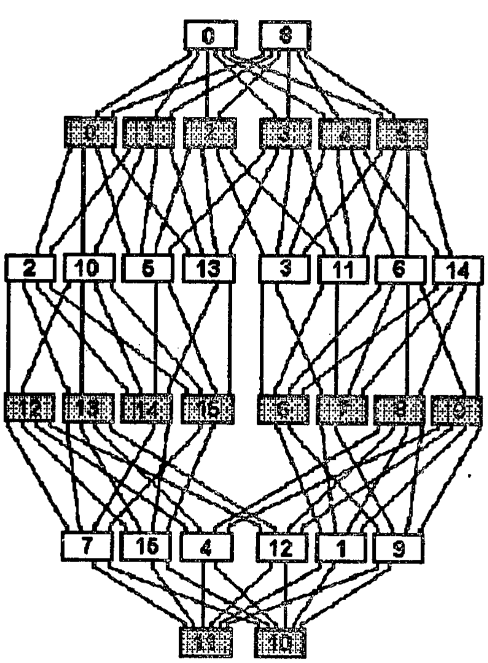


Fig.5. A torus network of //(16, 6; 16, 6)

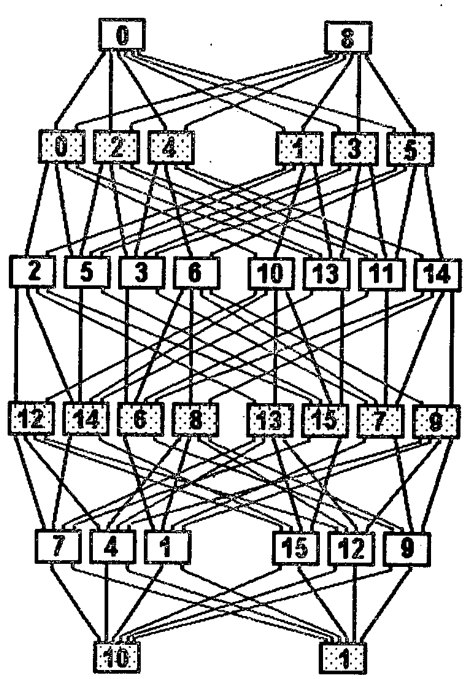


Fig.6. The torus in a "catamaran" form

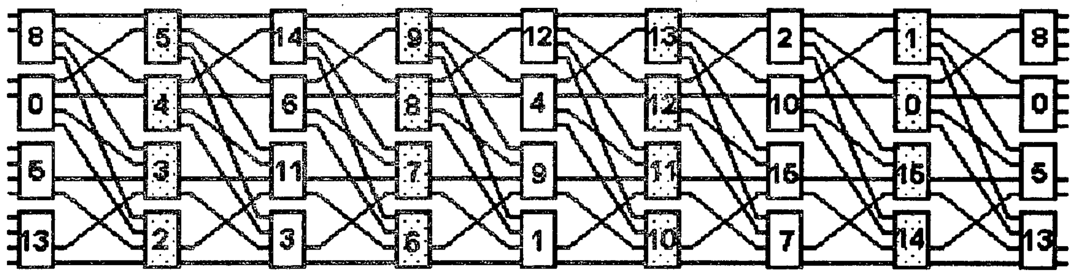


Fig.7. The torus in a "bracelet" form

So the considered 4-dimensional space of network topologies includes the networks, configurations of which one can get, varying: capacity and the alternative paths on multi-layered networks, parameters of tori and, finally, torus topology by means of topologically equivalent transformations. Among the remaining topologies, we consider only one, which may be useful for the development of networks on a plane. The topology in question is a plane tetraangulation with layer-homogeneous nodes. It is a kind of the generalized grid including common rectilinear grid as a special case.

All above considered networks had a layer-homogeneous configurations. It means identity of all the nodes, which belong to the same layer, regarding the numbers of inlet and outlet ports separately, that is regarding the node configuration. If the node configuration is indifferent, unlike the whole quantity of node ports, then such a network is called layer-homogeneous (regardless of node configurations) as the plane tetraangulation in question is. If all nodes have the same number of ports on each layer, the network is called simply homogeneous.

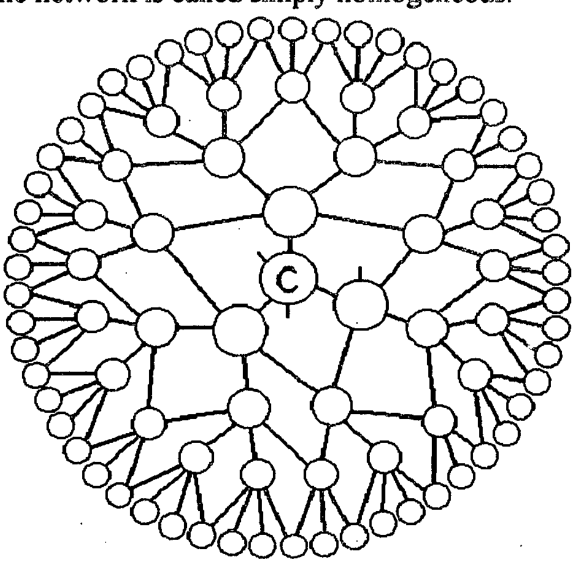


Fig.8. Plane tetraangulation on 5-ports nodes

Unlike the common rectilinear grid, which can have 4-port nods only, the plane tetraangulation may be constructed on nodes with any number of ports equal to or more than four (see Fig.8). It is much like a tree. It easily scaled like. a tree from the center (root) to the periphery. Its traffic self-guidance is like the one of a tree. It may be used as a kind of stencil or metanetwork for the development of a network with the traffic self-guidance. A hypothetical tetraangulation may envelop a designed or even really existing irregular network. In this case, the embedded network gets opportunity to be interpreted as the regular one so far as it coincides with a fragment of the regular network. It means, that this irregular subnetwork allows to realize the traffic self-guidance. The plane tetraangulation may be convenient for the simulation of large-scale LANs, MANs and WANs topologies. It easily simulate trees and rings with all kind of practically inevitable chords and cross-pieces.

CONCLUSION

The capacity, bandwidth, functions and services provided by existing high-speed networks, including FDDI, DQDB, B-ISDN, do not sufficiently match the requirements of the future multimedia communications. Even united forces of traditional technologies, new standards, fiber optics and photonics cannot eliminate this unmatching. For this problem solving, the regularization of network topologies is a matter of necessity. The main advantage of regular topologies is the traffic self-guidance that makes the large scale network supervisible, comprehensible and self-organizing. The current assortment of network topologies is miserable. It is nearly exhausted by bus, ring, tree (including star), rectilinear grid and banyan-based topologies. This paper has focused on two vast classes of new original topologies which, due to their regularity, provide the traffic self-guidance and gradual augmentation within practically unlimited range of parameters. To search an optimal configuration, the class of networks with layer-homogeneous configurations delivers the 4-dimensional space of capacity and number of alternative paths in multi layered networks or in tori parameters with a number of topologically equivalent transformations. The class of plane layer-homogeneous tetraangulations allows to construct plane grids of tetragonal meshes, using nodes with any, exceeded 3, number of ports. The offered topologies can be used not only for the construction of real networks. They can act the role of hypothetical metanetworks enveloping real irregular networks. In this case, the embedded irregular network acquires the traffic self-guidance ability.

REFERENCES

1. Henry P.S.(1989). High-capacity lightwave local area networks.- IEEE Communication Magazine, v.27,no.10, pp.20-26.
2. Hehmann D.B., M.G. Salmony, H.J. Stuttgen, (1990). Transport services for multimedia applications on broadband networks.- Computer Communications, Special issue: Multimedia communications, v.13, no 4, pp.197-203.
3. Fletcher P.(1991) A self-configuring network.- Connect. Sci. (UK), v. 3, no.1, p.35-60.
4. Maxemchuk N. (1987). Routing in the Manhattan street network.- IEEE Trans. Commun., v.C-35, no 5, pp. 503-512.
5. Borgonovo F., E. Cadorin, (1990) Packet-switching network architectures for very-high-speed service.- Proceed. Intern. Seminar on Digital Communications, Zurich, 5-8 March 1990. New York, NY, USA: IEEE 1990, p.336-351.
6. McKinly Ph.K., W.S. Liu (1990).Multicast tree construction in bus-based networks.- Communications of the ACM, v. 33, no.1, pp.29-42.
7. Lee T.T. (1990). A modular architecture for very large packet switches.- IEEE Trans. Commun., v. 38, No 7, pp. 1097-1106.
8. Lea. C.-T. (1991). Design and performance evaluation of unbuffered self-routing networks for wide-band switching.- IEEE Trans. Commun., v.39, no 7, pp.1075-1087.

9. Berra P.B., C.Y.R. Chen, A. Ghafoor, C.C. Lin, T.D.C. Little, D. Shin, (1990). Architecture for distributed multimedia database systems.- Computer Communications, Special issue: Multimedia communications, v.13, no 4, pp.217-231.

10. de Bruin N.G.(1946). A combinatorial problem.- Koninklijke Nederlands Akademie van Wetenschappen, Proc. v. 49 (part 2), pp. 758-764.

11. Clos C. (1953). A study of non-blocking switching networks.- Bell Sys.Tech. J., v. 32, pp. 406-424.

12. Benes V.E.(1964). Optimal rearrangeable multistage connecting networks.- Bell Sys. Tech. J., 1964, vol.43, no 7, pp.1641-1656.

13. Ofman Ju.P.(1965). Universal automaton.- Proc. Mosc. Math. Soc., v.14, pp.186-189.

14. Goke G.R. and G.J. Lipovski, (1973). Banyan networks for partitioning multiprocessor systems.- Proc. 1st Ann. Symp. Comput. Arch., pp. 21-28.

15. Wu C.-L. and T.-Y. Feng, (1980). On a class of multistage interconnection networks.- IEEE Trans. Comput., 1980, v.C-29, no 8, pp.694-702.

16. Lawrie D.H.(1975). Access and alignment of data in an array processor.- IEEE Trans. Comput., v.C-24, no.12, pp. 1145-1155.

17. Vidomenko V.P., (1993). Clos's networks 40 years later.- II-nd Conference "Informational Networks and Systems, CINS-93". St.Petersburg, 18-20 November 1993. Proceed., pp.42-44. (Russ.)

18. Vidomenko V.P., (1994). A layered networks topology.- International Informatization Academy Forum III. International Conf. on Informational Networks and Systems, ICINAS-94. St. Petersburg, 24-28 October 1994. Proc., pp. 359-264. (Russ.)

19. Vidomenko V.P., (1995). Layered network topologies calculus.- The 4-th St.Petersburg international conf. "Regional Informatics-95". St.Petersburg, 1995. Abstr., P.1, p.99.

20. Vidomenko V.P., (1991). Unified shuffle networks.- XVI All-Union school-seminar on computer networks. Reports, p.1. Moskva - Vinniza, 1991, pp.17-22. (Russ.)

21. Vidomenko V.P., (1989). Homogeneous bipartite networks.- XIV All-Union school-seminar on computer networks. Reports, p.1. Moskva - Minsk, 1989, pp.29-34. (Russ.)

22. Vidomenko V.P., (1992). Algebra of bipartite-homogeneous networks with unified shuffle.- XVII All-Union school-seminar on computer networks. Reports, p.1.Moskva - Alma-Ata, 1992, p.37-40. (Russ.)

23. Vidomenko V.P., (1987). Analytical model of a large homogeneous network.- Avtomatica i vychislit. technica, 1987, no.4, pp.36-41. (Russ.)

24. Vidomenko V.P., (1989). Combinatorics of plane delta-homogeneous tetraangulations.- Cybernetica, 1989, no.4, pp.64-69. (Russ.)

CHAPTER III:

INTELLIGENT MANUFACTURING PROCESSES AND SYSTEMS

THIS CHAPTER INCLUDES PAPERS

PRESENTED AT THE CONFERENCE SESSION:

INTELLIGENT MANUFACTURING PROCESSES AND SYSTEMS

Organized by: ***Prof. Alexander V. Smirnov,***

Prof. Peter Groumpos.

PHILON - A NEW HOLISTIC DESIGN METHOD OF AUTOMATION CONCEPTS

Alfred Erismann
DELEPROJECT AG
CH-3114 Wichtrach
Switzerland

Abstract. The PHILON-Method concerns the design of automation concepts for process- and production automation, specially for chemical and mechanical processing applications. The method was designed by the author based on the conclusions of more than twenty-five years of practical experience in a great variety of automation projects all over the world. PHILON tries to show new methods to design better concepts due to enterprise strategy and goals. Focal points of PHILON are a principal strategic approach to automation problem solutions , based on a set of systematic perspectives of automation problems, applied on an abstract automation model. The final evaluation of a designed concept version to be realised is based on the calculation of the best automation net yield rate.
The method is named PHILON in memorial of Philon Byzantius, an early Greece philosopher and engineer who was the author of the first known compendium about *Automates*.

Keywords. Automation concepts, automation model, automation perspectives, automation net yield, strategic approach, holistic design method,

Introduction. During the past two decades, the degree of automation in production plants of the western industries had an enormous increase, not only for the reason of manpower shortage but also for security and quality reasons and the operability of the mass production plants. A result of this trend was a shortcoming of adequate and consistent theory of automation and methods as well as the connection with the principal economical and entrepreneurial questions. So automation concepts were usually designed empirically bottom-up, based on individual personal experiences and electronic innovations and not by a good and fundamental designing practise.

In fact, automation as an essential element of today's industrial production can only be studied and threated in a holistic view. Enterpreneurship and management are called up to integrate automation in their entrepreneurial paradigm.

Understanding the problematic nature of automation concept decisions. In the first place, automation concept design is an entrepreneurial challenge to the enterprise- and factory management itself. PHILON postulates, that *automation* is a tool to reach enterprise targets and therefore has to be an explicit subject of enterprise strategy. The approach to automation concepts has to be done top-down, starting by interpretation of enterprise needs and ending by a concrete concept. It is also a matter of course that automation projects have to be subject of investment calculation methods and return-on-investment. Automation has to prove an economical benefit.

The lack of a comprehensive theory of automation. Some high sophisticated special fields as control theory, fuzzy logic or cybernetics are cultivated but no general theory of automation as a performativ act in a global point of view is existing.. Because of the impact of automation concepts on economical and entrepreneurial effects, this lack can be completed with an abstract automation model as a general conception of problem to be solved.

Development of expedient design procedures. As a result of the above mentioned facts, we are missing today straight and appropriate problem solving procedures. PHILON focuses on the three following aspects:

**Working of human intellect and mind.* The practised linear and algorithmic procedures of design- and engineering are not appropriate to the real working of human intellect and mind. The methods must be adapted to the working of human intellect and mind and not vice versa.

**Consensual harmonisation of objectives.* Normally a wide circle of interested parties with more or less divergent interests inside the enterprise are involved in automation projects. These partial interests have to be articulated, discussed and prioritizated in a consensual way allowing the involved persons to find and to accept a satisfying agreement, documented by the attracts of automation

**Anticipation of future trend aspects.* Automation concepts are long living in its results and effects according to the individual life cycles of the components of automation system as plant, hardware etc. These life cycles differ from 12 to 15 years of functional lifetime for hardware up to 25 or 30 years for the plant itself. So foreseeable trends have to be respected in concept design best possible to prevent or reduce a future impossibility to migrate the concept in the needed direction. PHILON includes a method to trace out the adaptability of automation concepts in a projective way by considering the relevant future trends of product markets, customer needs, hard- and software innovations, education etc.

Principal strategic approach to automation problem solutions. The starting point for automation considerations is principally the enterprise strategy. Normally today there is no explicit reference on automation matter. The respective implicit facts have to be figured out as a derivate of interpretations of different other topics. This derivate of interpretations has to be confirmed by the authorised management. Subjects of strategy based interpretation are:

**Amortisation targets.* Delay of amortisation (in years) gives important indications of demanded automation net yield and focal points of rationalisation. Important variable of investment calculation.

**Capacity politics.* Constant , seasonal or occasional use of production capacity important for design of MMI (Man Machine Interface) and amortisation targets.

**Autarky politics.* Autarky concerning maintenance and service, storage, resources, energy etc.

**Employment philosophy.* Manner of employment, specially defined in operating- by-exception, operating-by-routine, operating-by-milestones etc.

**Purchase politics.* Purchasing capital goods as equipment's and automation systems, e.g. phased extension in the dimensions of: system integration, operating comfort, performance and output.

**Marketing strategy.* The wide range of possibilities of marketing strategy between mass production and filling the gap with special or individualised products requires different strategies of automatization with fix or recipe controlled programs.

**Personnel politics.* In the sense of personnel qualification and staff as well as type of personality psychograms, important not only for the design of MMI but also for an anthropocentered adequate division of labour between man and machine, which is the cardinal question of automation.

**Safety politics.* Safety and security priorities to observe. Relevant for MMI and system architecture.

The derivate of interpretations must form a consistent, intact and plausible integral whole.

Phenomenology of automation problems. The problems of automation presents themselves as multidimensional, complex and interdisciplinary. To fulfil the claims of a good automation concept, it's indispensable to include all perspectives and aspects, to consider all connections and the environmental context at beginning of a projetc. This manner of thinking, considering and imagination is called *holistic reflection*. A *holistic* reflection demands for polyvalent concept engineers with the ability to abstract. *Reflection* means the inexistence of linear algorithmic thinking paths. Reflection is an intellectual procedure with highly contemplative character, which needs a certain maturation period.

PHILON gives the following pattern of four defined perspectives as *phenomenological spectrum:*

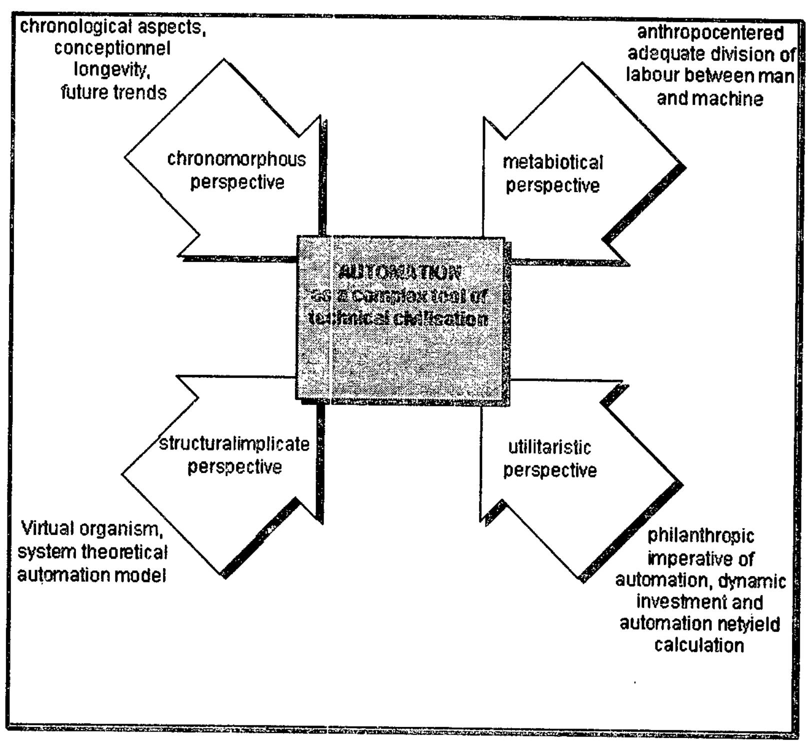


Fig. 1 Phenomenological Spectrum

* *Chronomorphous Perspective.* The chronomorphous perspectives show the inherent chronological aspects of automation concepts. The aspect of *conceptionnel longevity* of automation concepts and it's cost and benefit determing effects which are variables of the investment calculations. Changes on the concept level are often impeded by expensive costs. The aspect of *asynchronism of life cycles* is evident specially because of the gap between life cycle of the plant and life cycle of the automation system. A further aspect is the *anticipation of future trends* as mentioned above. Evident trends are: global trends (social, world economy, global resources), consumer trends (consumption trends, fashion and taste), marketing trends (availability, marketing mix, customer needs), technological trends (materials, procedures, quality), production strategies (complementary or substitutes on the same plant), automation technological trends (innovations on hard- and software, methods), labour market trends (job situation, education, job descriptions).

* *Metabiotical perspective. Metabiosis* means the coexistence of two different types of species, from which only one is a beneficiary of it. Self evident, the machine as a human creation is invented to complement the man with it's capabilities. Therefore in the case of automation the beneficiary should be the human. That's why the metabiotical perspective considers the cardinal question of *anthropocentered adequate division of labour* between the metabiotic partners man and machine. A powerful metabiosis between man and machine amplifies the characteristic strengths and neutralises the characteristic weaknesses of both. It's an important act of balance and optimisation. The visual product of this act is the design of the MMI. This is one of the most important problem in automation which has to be solved with accuracy. Following some types of metabiotical effects: simulating, leading, serving, amplifying, substitution, supplement, safeing.

Utilitaristic perspective. Utilitarism is a philosophical doctrine which postulates usefulness as the main decision criteria. We have to accept this philosophical position for automation purposes.

Utilitaristic postulate of automation. In the first place, automation has to serve the enterprise goals. That means, that automation must achieve economical benefits, maximised for the whole period of life of the automated productionplant. The influence of automation concepts can be calculated in the well known dynamic investment calculation by the capital value method.

Philanthropic imperative of automation. We suggest the following categorical imperative of automation to prevent an eventual unscrupulously profit of automation in the sense of misuse and ignoration of the dignity of man:
All operations, functions and jobs which are definitively inhumane must be automated. All which are definitively humane are prohibited to automate and all which are neither definitively humane nor definitively inhumane are allowed to automate.

Even automation is a moral question after all.

Dynamic investment calculation. Characteristics of this method are the balance of the costs as well as the proceeds of each phase of the lifecycle, including the phase of investment and realisation, the phase of production and the final phase of desinvestment. The resulting amount has to be capitalised with compound interest. The result is the investment net yield. The costs and proceeds are to break down in a sufficient range of items that allows to consult the punctual influence of conceptional components in the widest sense.

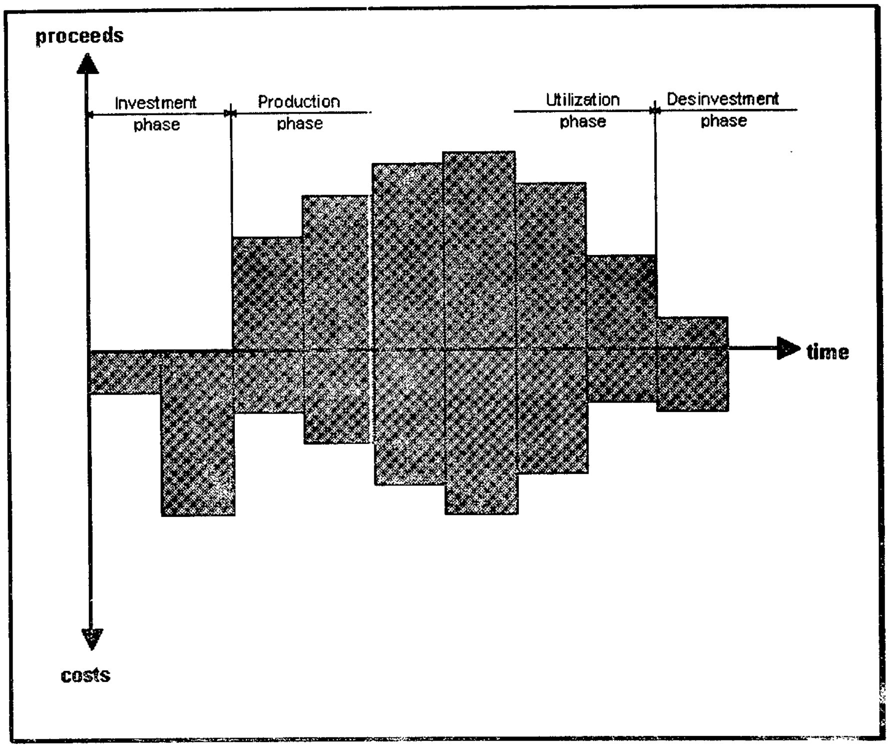


Fig. 2 Dynamic Investment Calculation

* ***Structuralimplicate perspective.*** Recognised structures and patterns of automation problems are projected on the multidimensional system theoretical automation model of PHILON. The relations and degree of congruency between patterns of different layers and the interactions and relationship between the system elements are typical and allow concrete conceptional conclusions. The whole enterprise with it's production plants and organisation is shown as a *virtual organism* in the original sense of the word. The structuralimplicate perspective includes the following main aspects: structural aspect (structures, architectures, patterns), functional aspect (functions, principles, effects, procedures, methods), operational aspect (operating, manipulation, handling, steering), dispositional aspect (logistics, disposition, availability, communication, traffics). In the axial dimensions, the projection leads right through the different levels of automated systems: physical process level, plant level, actor-sensor level, automation system level, automatotomical level, operator ergonomic level, organisational level and manufacturing tactical level. In the philon automation model, the holistic reflection is institutionalised.

Automation model. The PHILON automation model is a babuschka-like four-dimensional model with actually three spheres, corresponding partially with different levels of the automated system, boxed each into another.

The automation model has a tetraedrical form, determined by four aequilateral triangles, representing the four dimensions. The intellectual penetration of the model can be done in an endogenous (from the inside to the outside) or exogenous (from the outside to the inside) way.

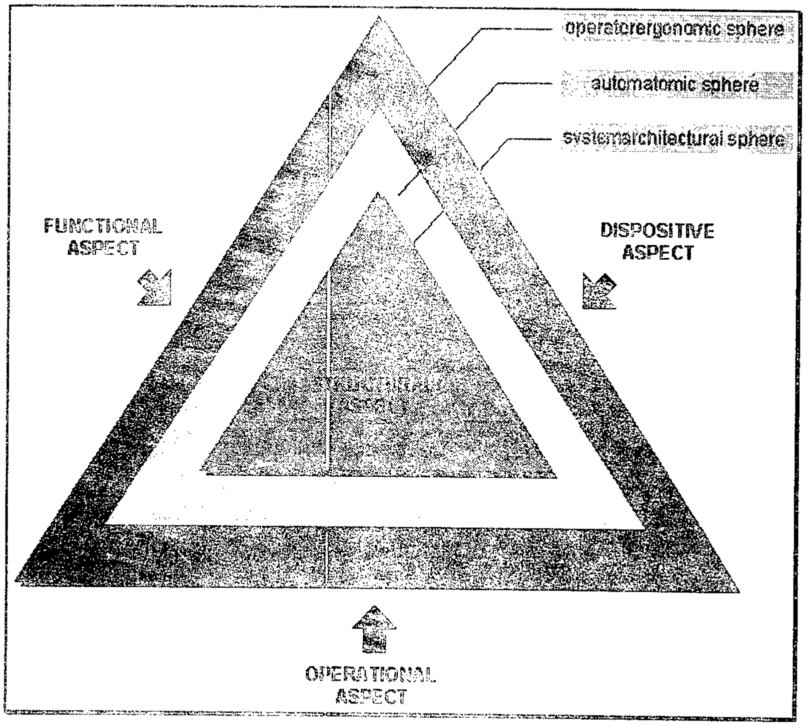


Fig. 3 Automation Model

The following table presents the substance of the aspects and spheres of the automation model in a concentrated form:

Table 4 Configuration of automation model

Spheres / Aspects	operatorergonomic	automatotomic	systemarchitecture
Structural	*Ergonotectonics* Layout operatorpanels Topology of pictures Competence levels	*Structography* Structure, hierarchy, redundancy, standard modules	*Systemarchitecture* Architecture & layout hardware, networks, typical
Functional	*Imaginography* Iconography, colours, ani= mation, architectonics of pictures, alarming	*Functionality* Control functions, batch control, PID, sequences, emergency procedures	*Functional design* Functional principles effects, design of com= ponents
Dispositive	*Histography* Datahandling and protocols, process documentation, event= and alarmprotocol, statistics	*Organography* Datacomputing, intra= modular communication, information types	*Systems interactivity* Interactivity of system components, interfaces, datahandling
Operational	*Manipulation* Picture callup, inputprocedures, job- & receipehandling, alarm= handling, override, emergency operations	*Status transition* Automate status, status transitions, conditions of transition	*Systemhandling* Boot, restart, emerg= ency functions, power-& systemfailure

Automativity. This term is a neologism of the author and means the *extent of automation efforts*. The term of automativity replaces the term of *degree of automation*, which is the coefficient of the numbers of operations, that could be automated and the number of the really automated ones. This term of degree of automation is an only retrospective and as a value rated index inappropriate.

Automativity is represented in a spiderweblike dodecagonal diagram with twelve scaled rays of automation criteria. The dodecagonal diagram gives a qualitative as well as a quantitative impression of automativity.

The *profile of automativity* shows the qualitative aspect of the extent of automation efforts. The *area* of the profile of automativity gives a certain measure of the intensity of these efforts.

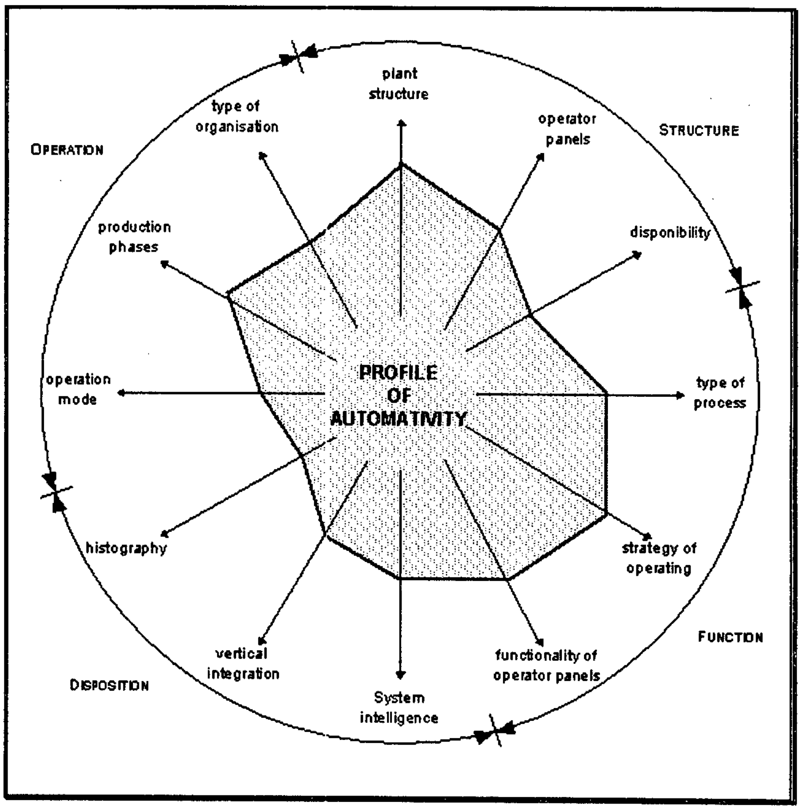


Fig. 5 Profile of Automativity

Adequate profile of automativity. The question for the adequate profile of automativity is an individual and must be answered projectrelated. The primary tricky question of H.P.Huber Novartis Switzerland : *„ Which is the adequate degree of automation?“* can never be answered by a simple value rated index. The answer on this question is decided by the designed concept version with the maximal *automation net yield.*

Automation net yield. The automation net yield calculation is based on the above mentionned dynamic investment calculations for the designed concept versions. This calculation needs as a reference value a *cero option of automation concept.* This may be a realisable concept version with an absolute minimum automativity.

i0 = investment net yield value of zero option version
in = investment net yield value of concept version „n"
an = automation net yield value of concept version „n"
ax = maximal automation net yield as concept version „x"

$$\mathbf{an = in - i0}$$

The best concept version is the version „x" with the maximal automation net yield of the whole progression, for example:

$$\mathbf{ax > am > ap > az > af > ac \ldots}$$

The version „x" represents the concept with the *adequate automativity* however this individual concept may be configured.

Summary. The PHILON-Method offers a new and holistic approach to automation concepts starting by the goals and needs of the enterprise and gives also a tool to harmonise the relation between human and technological resources in production ,based on an automation net yield calculation.

ORGANIZATION MANAGEMENT DECISION SUPPORT TOOLS FOR MANUFACTURING SYSTEMS RE-ENGINEERING

Aleksy Barski[1], Edward Michalewski[1], Irina Rakhmanova[2], Alexander Smirnov[2]
[1]Systems Research Institute of the Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland
e-mail: michalew@ibspan.waw.pl
[2] St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences
39, 14-th Line, 199178, St.-Petersburg, Russia,
e-mail: smir@iias.spb.su

Abstract. This paper investigates how Decision Support Systems (DSS) technology can be applied to manufacturing systems organization management at different stages of conceptual design, strategic planning and organization management to meet the requirements of the flexible business environment. The approach proposed implements the technologies of group problem solving with mechanisms of conforming decision making and simulation model designing with generation of computerized information subsystems for organization management. The approach is illustrated by an environment **PQE** developed in the St. Petersburg Institute for Informatics and Automation Russian Academy of Sciences on FoxPro 2.5 under Windows 3.11, and **DIANA-9** environment elaborated at the Systems Research Institute of the Polish Academy of Sciences as program environment for IBM PC 386/486 with CASE-tool properties.

Keywords. group decision support system, organization management, enterprise re-engineering, diagnostic analysis.

Introduction. Enterprises today are characterized by ever-increasing levels of information complexity and turbulence. The modern business environment is undergoing a change at a continuously increasing speed. In order to survive in the competitive, global market of today, enterprises must be flexible, and capable of making effective decisions quickly.

Today's enterprise represents a complex organization, that brings together people, resources and facilities in order to develop, manufacture and market products profitably. In order to maintain its profitability, the enterprise also must flexibly react to changing market conditions. Manufacturing companies need methodologies and tools that cover all levels of the organization, from management down to operational levels to support the strategy formulation and implementation process [1-3]. There is no single universal tool that would solve this problem in its entirety. The organization management approach proposed includes the following general stages:

- evaluation of external enterprise requirements,
- mapping external requirements on internal requirements (manufacturing system goals),
- configuring flexible organization management system.

Therefore the principle objective of the interactive environment to be considered is to support enterprises in their continuous process of adapting themselves to a changing business environment. For this purpose DSS have become widely used. They represent computer-based applications for the decision making process.

The main features of advanced information technology to be proposed include group decision support facilities, intelligent support for domain model design, quantitative and qualitative information to process, multifaceted diagnostic analysis, C language program generation facilities.

"Re-engineering is the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service and speed" [4]. The re-engineering problem can be formulated in the following way: to specify the requirements to a system and to compose a system from a number of template solutions to meet these requirements. To realize the above decision making scheme in the considered domain the following methods are proposed:

- hierarchies analysis method with conformation of decisions for enterprise goals specification, evaluation and choice [5, 6],
- domain model designing mechanism,
- multilevel polyhierarchical information network process,
- multifaceted diagnostic analysis [7],
- multiple criteria choice methods.

The Management System Model. The GDSS tools proposed are based on the management system model, represented as multilevel units hierarchy with information network (Fig. 1).

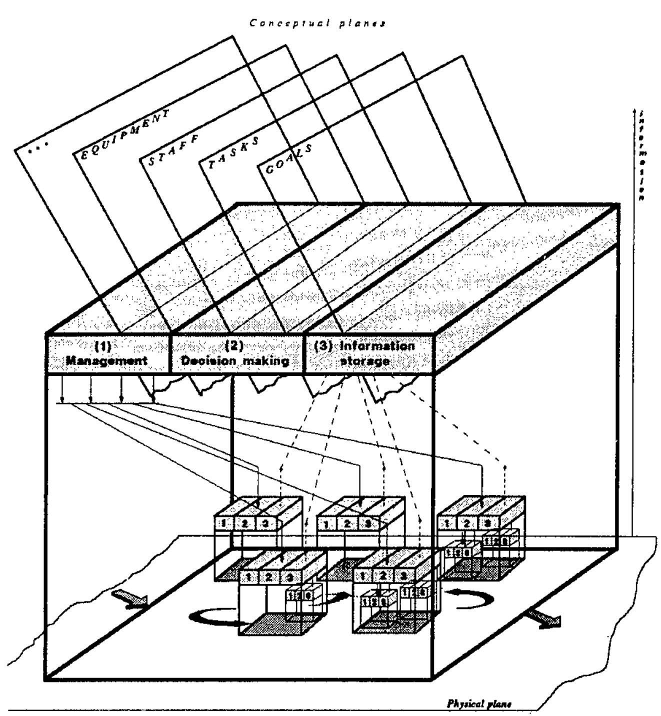


Fig. 1 The organization management system model

Every unit (block) the hierarchy composed describes a partly autonomous controllable system with it's own goals, tasks, staff, equipment etc. There are three main unit functions (block top):

- to store information about it's lower level units behavior (states),
- to support the decision making process, based on the external goals and internal states evaluation,
- to produce the control impacts to the lower level units according the best solution obtained. The control impact result is the goal oriented lower level units interaction.

The bottom of each block is a part of the physical plane, where all physical entities are. The block volume is the information space.

The hierarchy conceptual plane projection represents the conceptual (plane) model of the organization management system. This plane model also has multilevel hierarchy structure with information network. To analyze different aspects of the organization management system we need to consider the composition of the interrelated plane models.

The Intelligent GDSS Tools General Structure. There are four main components of the intelligent GDSS (Fig. 2):

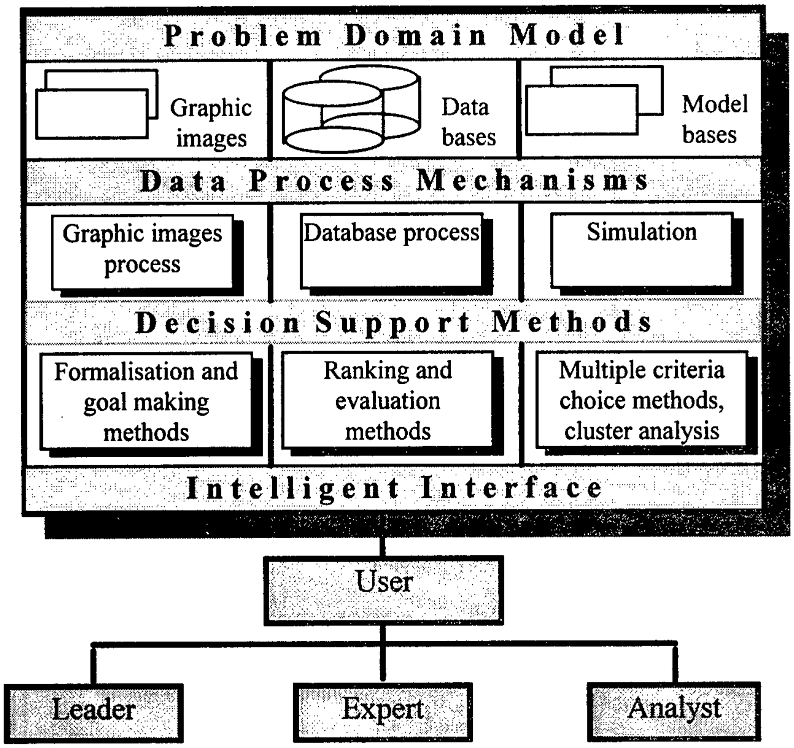


Fig. 2 The intelligent GDSS general structure

- Problem domain model, represented by basic information types: image (knowledge), data and different kinds of models.
- Suitable information processing mechanisms such as image process, database management systems for data and simulation systems for computer models.
- Decision support methods such as intuitive (empirical) methods for difficult problem normalization and goal making, mathematical methods for multiple criteria choice, factor and cluster analysis etc. and combining methods such as hierarchies analysis method for conformed expert decision making.
- Intelligent interface includes the graphic representation tools, help support, demo examples for comprehensive education, Windows similar screens layout, report support.

Three kinds of GDSS users are considered:

- Team leader, who have to obtain the best decision on the concrete problem area (management object).
- Experts as the domain experienced employees of the management object considered.
- Analyst as a specialist in the area of computer models and mathematical methods.

For the enterprise management organization re-engineering purpose the following steps are proposed (Fig. 3). Two following tools are proposed to realize the above decision making scheme in the considered domain:

1. Program environment DIANA-9 makes it possible to carry out the multifaceted diagnostic analysis of the management system considered, to implement - on the diagnostic results basis - the improving changes and to design the new organizational structure with the possibility of testing the changes effectiveness introduced on the model. The program environment allows also to design the computerized information subsystems for selected activities, generated in C language. For enterprise re-engineering purpose the DIANA-9 provides the capacity of verifying various strategies and, after having chosen the most appropriate one, adapting the organizational structure to it.

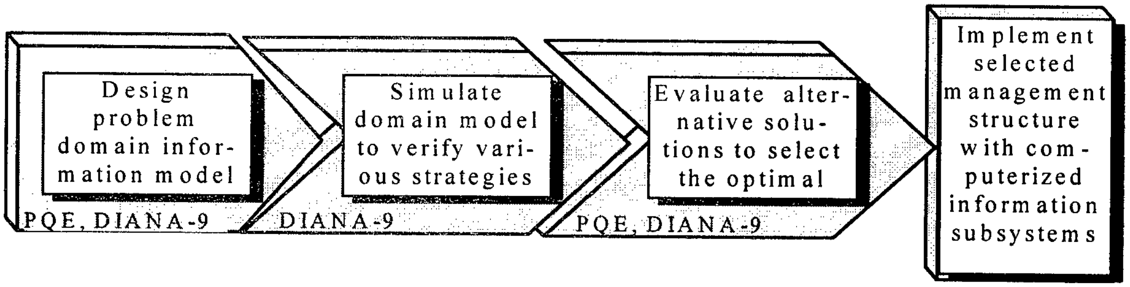


Fig. 3 Designing steps for the optimal organization management structure implementation

2. PQE has been developed to sample and order internal goals, tasks, particular solutions attributes in dependence upon the external requirements and to order and choice obtained alternative organization management solutions. For this purpose experts estimate alternative sets of goals or solutions from different viewpoints reflecting the design aspects. The initial GDSS is interfaced by appropriate techniques to an intelligent front-end for group expert decision support for functional and quality attributes sampling, weighing and evaluation with subsequent acceptance of conformed decisions.

Both environments are based on the organization management system model above. The domain information model designing is considered as a process of conceptual planes projection to obtain the sets of interconnected hierarchies: goals, tasks, resources, organizational units, elementary operations (Fig. 1).

To evaluate goal tree, to determine the tasks and attributes weights PQE environment is applied. For updating or creating new organization management system model with simulation and diagnostic analysis the DIANA-9 environment is suitable. The PQE and DIANA-9 environments interaction scheme is represented in Fig. 4.

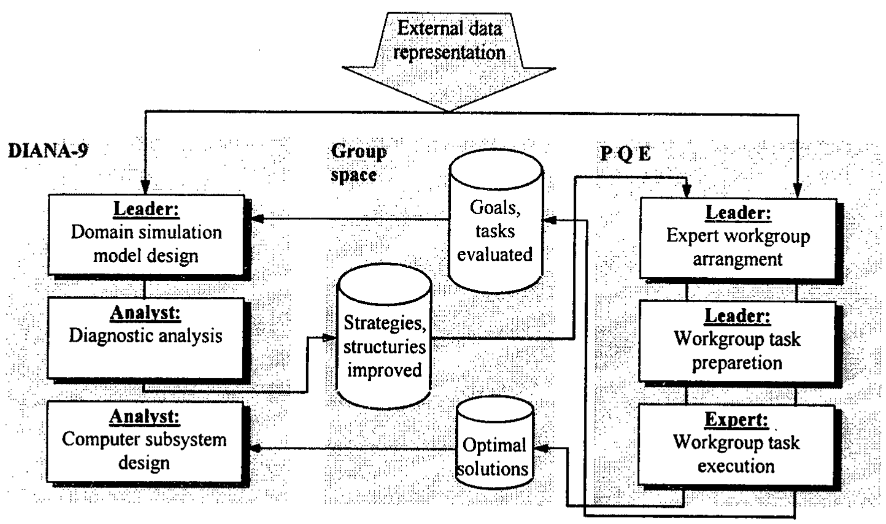


Fig. 4 PQE and DIANA-9 interaction scheme

Group Decision Support Tools For Hierarchies Evaluation. The initial GDSS is interfaced by appropriate techniques to an intelligent front-end for group expert decision support for functional and quality attributes sampling, weighing and evaluation with subsequent acceptance of conformed decisions. Fig. 5 represents PQE external environment [6].

To obtain conformed expert decisions two levels of abstraction are distinguished: local level (subsystem "Expert") and group level (subsystem "Leader") aggregating local expert opinions.

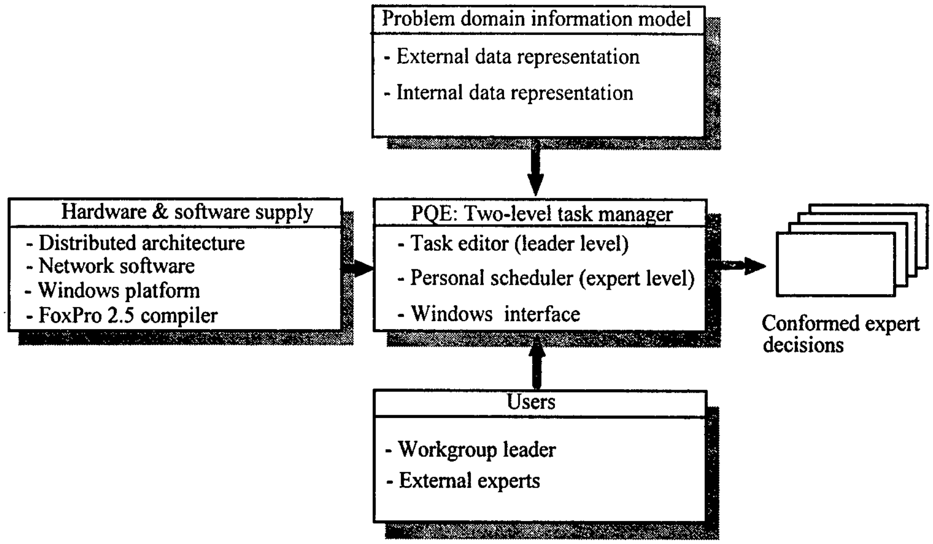


Fig. 5 PQE external environment

PQE external problem domain data consists of alternative sets of goals or solutions, represented as:

Project $P = \{F, V\}$, where F - attributes glossary, V - attributes values.

Attribute glossary $F = \{C, N, Type, Um\}$, where C - attribute codes, N - attribute names, *Type* - attribute types, *Type = {Qualitative, Quantitative, Classifying}, Um* - units of measurement.

Internal problem domain data is represented by templates as: $\Phi_1 \colon F \rightarrow T_g$, $\Phi_2 \colon V \rightarrow T_v$, where T_g - template of attributes glossary, T_v - template of attributes values.

The distributed environment of the group expert evaluation of organizational decisions quality is based on the two-level coordination mechanism. The work coordination mechanism at the group level is associated with the activity of a workgroup leader and provides for creation and maintenance of analysis and decision evaluation tasks submitted to the group of experts. The mechanism is realized as a task editor and a system of supervision of the task execution by the expert workgroups. The task is considered to be formed if the following sets are defined:

$T = \{S, P, E\}$, where S - scenarios, P - projects, E - experts (workgroup);
$S = \{M, D\}$, where M- evaluation methods (models), D- stages date;
$M = \{St, Sc\}$, where St - stages, Sc - scales.

PQE implementation is based on the algorithm for creating the goals specification and alternatives evaluation. The stage by stage algorithm operation is guided by the set of programming and information models, used by workgroup leader during the task scenario synthesis.

The enhanced hierarchy analysis method is used to obtain conformed expert decisions. This method is based on iterating procedure of local decisions aggregation to come into an agreement with that of a workgroup, being processed. In this method conformation attributes are used as well as graphic representation schemes of various type using verbal and numeric ranking scales. A feedback is applied to different stages of an algorithm: group decision aggregation, initial data representation and expert agenda.

To obtain conformed decision an expert changes iteratively attribute weights or group estimation is involved as an additional attribute. A wholesome function is used to select a decision. In the case conformed decision is not obtained after a number of iterations, conformed procedure leads to partially conformed decision. Multiple criteria choice methods are applied for solutions ranking.

The work coordination mechanism on the local level is associated with the work of experts group and is implemented as a task manager. The task manager performance is based on the analysis procedures of interconnected instructions, functions, states and messages sets. The shared work coordination calendar provides for the activity of different workgroups in the task space being tied to terms of particular stages. Meanwhile, the individual workgroup members decisions are being kept non-transparent until they are submitted as particular opinions for a formation of a group decision. The approach presented is applied also for ordering and choosing obtained alternative organization solutions.

The Diagnostic Analysis Tools. The DIagnostic ANAlysis and design management system tools (DIANA-9) enable:

- to carry out the multifaceted diagnostic analysis of the management system model considered,
- to implement the improving changes on the diagnostic results basis,
- to design the new organizational structure with the possibility of testing the changes effectiveness introduced on the model.
- to design the computerized information subsystems for selected activities, generated in **C** language.
- for enterprise re-engineering purpose the DIANA-9 provides the capacity of verifying various strategies and, after having chosen the most appropriate one - of adapting the organizational structure to it [7].

The fundamental DIANA-9 functions are represented in Fig. 6.

The block of computer-assisted diagnostic analysis. This block identifies 64 symptoms of incorrect situations on the particular network levels, like, i.e.

- information blind alleys,
- doubling of operations,
- "bottlenecks",
- lack of temporal synchronization,
- uneven charge on the units,
- malfunctioning/dysfunction,
- sources of errors and delays,
- hidden conflict situations,
- inconsistent hierarchy of positions
- inadequate propensities,
- inadequate assignment of staff,
- inadequate division of cells,
- lack of connection with goals,
- inadequate resources, etc.

The symptoms identified are transmitted to the diagnostic matrix, which determines the syndrome of causes bringing about these malfunctions. This procedure facilitates the identification of incorrect functioning source locations and allows to correct corresponding malfunctions. The obtained designs of organizational improvements are sequentially tested on the model. The best variants tested is then implemented in reality. The DIANA-9 environment enables to design and simulate new parts of the system considered.

The block of computer-assisted design of organizational structures. This one makes use of the so called *germs* - the positions which are the essential ones for the designed organizational units, determined by the team of designers (the particularly experienced employees of the management object considered). The software, when carrying out the "cluster-analysis" type of al-

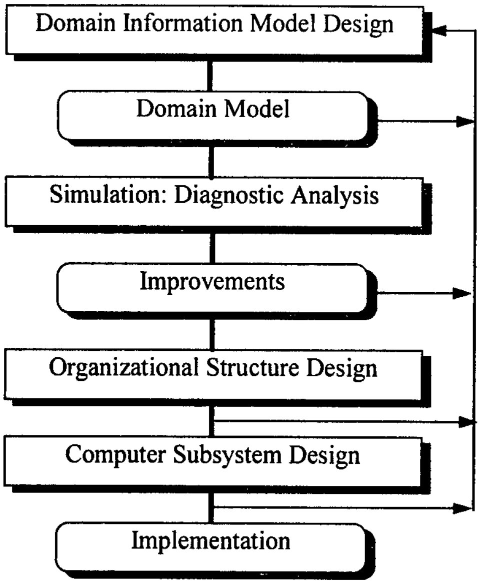


Fig. 6 DIANA-9 functional scheme

gorithm, concentrates around these germs other positions which are most strongly connected with them. The *quality measure* of the units designed is provided by the so called *linkage strength*, which reflects the compactness of the operations performed within the units, while the so called *dispersion measure*, characterizing connections among units, corresponds to the quality of the whole design.

During the design process we tend to maximize the strength of linkages and to minimize the dispersion measure. In a more vivid language we aim at preservation of the "closed quarters" principle: most of the matters are settled by employees (clerks) in their offices-organizational units, and only when matters are actually settled, the results are transmitted to another office-unit. Implementation of this process for the consecutive levels of the structure, assisted by the DIANA-9, gives the possibility of obtaining a complete organizational design for the analyzed management system.

The block of computer-assisted design of computerized information subsystems. This one enables to automatically distinguish from the whole of the network of informational connections the part which is best fit for computerization. It is for this part, when the formal description has already been introduced (the environment contains a number of facilitating devices, making it easier to carry out this stage of the process) that we can simulate the functioning of the future computerized information subsystem. When we obtain satisfactory results we do generate the subsystem in the source code (C language). Depending upon the degree of advancement of computer literacy and equipment in a given management system we can either include the generated software in the existing subsystems or use it, after compilation, as an autonomous subsystem.

The general scenario of the organization management structure improvement. The general scenario is carried out on the management system model according to the following sequence:

1. After the diagnosis the changes are introduced into the model, and this step is repeated until all shortcomings are eliminated.
2. The multi-variant design of the new organizational structure is performed - the best variant is introduced into the model and the diagnostic cycle (above) is repeated again.
3. Design of the computerized information subsystems.
4. The changes which would be brought about by the subsystems meant to be implemented are introduced into the model and the diagnosis is carried out yet again.

Upon the completion of this phase the DIANA-9 can be used as the organizational consultant in the design of the developmental and restructuring undertakings, enabling simultaneously carrying out of the on-going monitoring of the way the management system functions.

Conclusion. We have considered an approach, which covers the main stages of decision making process and well do for re-engineering problem solving. The environment designed comprises means for group decision support for goal and task specifications, multi-aspect diagnostic analysis and simulation of the organization management system model, the alternative-solutions-oriented design of the enterprise re-engineering and project quality evaluation. The GDSS tools proposed do not only fully correspond to the capacities offered by the most modern instruments (e.g. CASE-tools), but gives also entirely new, unique capacities, namely

- adaptation of the organizational structure to the selected restructuring design,
- consideration of the thus important factor of the management systems as the *human factor* (including, for instance, identification of personnel conflicts or of the lack of work satisfaction, achievement of the rational assignment of tasks and of a more effective manpower economy);
- securing of the answer to the apparently most critical question in the process of implementation of the computerized information systems: *what will happen next?* - how will the management system react through its changes in terms of goals, structure and functions, what will be the new rational assignment of tasks and people? this answer can be gotten not *post factum*, but yet at the stage of design, and so the possibility arises of generating alternative solutions to the problem and finding the best of them;
- the working of the DIANA-9 constitutes the common *integrating platform* for the specialists from various domains, not only computer specialists, but also managers, lawyers, economists, financial experts, sociologists, engineers and specialists in technology.

Acknowledgments. The paper is due to the research which is carrying out as a part of the joint project "Methods, models and environment for manufacturing system configuration design and organization management" in correspondence with cooperation agreements between Polish and Russian Academy of Sciences in 1996 - 1998.

References

1. Lepikson, H.A. (1995) Core Competence for Flexibility in Product Design and Manufacturing: One Approach for Long Term Competitiveness. Proc. of the II International Conf. on Concurrent Engineering, Washington, pp. 541 - 548.
2. Sanchez, J.M., W. A. Priest (1994) Heuristic Based Approach for Producibility Analysis in Concurrent Engineering. In A. Paul, M. Sobolewski (ed), Concurrent Engineering: Research and Applications (CE'94), Pittsburgh, Pennsylvania, pp. 401 - 405.
3. Fernihough, A.M., G.W. Owen, A.R. Mileham, S.J. Culley (1995) The Development of a Technique for the Evaluation of Business Strategies. Proc. of The Conf. Concurrent Engineering (CE'95), McLean, Virginia, pp. 521 - 528.
4. Hammer, M., J. Champy (1990) Re-engineering the Corporation. A Manifesto for Business Revolution. Harper Business, 215 p.
5. Saaty, T.L.., K.P. Kearns (1985) Analytical Planning: The Organization of Systems. Graduate Scool of Business, University of Pitsburg, Pennsylvania, 222 p.
6. Smirnov, A.V., L.B. Sheremetov, I.O. Rakhmanova, P.A. Turbin (1996) GDSS for Re-eggineering of Production Systems. Proc. of the Sixth International Conference (IPMU'96), Granada, Espania, pp. 313 - 318.
7. Michalewski E. (1992) Multilevel polyhierarchical model for organizational decision support implemented on IBM PC type package DIANA-9. Proc. of the International Conf. Support Systems for Decision and Negotiation Processes, Warszawa, pp. 253-256.

SIMULATION-BASED PLANNING AND CONTROL OF PRODUCTION FRACTALS

Dr. Wilfried Sihn, Dipl.-Ing. M. Lickefett
Corporate Management Department
Fraunhofer Institute for Manufacturing Engineering and Automation (IPA)
Nobelstr. 12
D-70569 Stuttgart, Germany
e-mail: whs@ipa.fhg.de

Abstract. In order to increase their competitiveness, companies optimize their manufacturing by adapting structures to the requirements of technology and market. On the other hand, they try to fulfil specific requirements by extreme customer orientation and manufacturing on demand. This leads to new organizational structures and manufacturing systems. Appropriate organization forms, such as the fractal company, are distinguished by high flexibility, rapid adaptability and exploitation of human potentials. However, they also require appropriate PPC-systems (Production Planning and Control system) or rather new methods of PPC. In addition to a high planning reliability these systems have to ensure high flexibility, distributed processes of planning and decision-making, general orientation towards common goals and an information flow conform to the requirements. The most important elements of such a PPC-system are long-range analysis of the production structure, medium- and long-range planning for the entire production area, short- and medium-range coordination among production fractals and short-range shop-control inside the fractals.

At the Fraunhofer Institute of Manufacturing Engineering and Automation simulation-based systems are being developed to support the planning and control of such organization forms within the production. Thereupon the user is put into the position of planning job orders reliably and optimizing production, concerning cost and schedule reliability, by making use of the advantages of decentralized structures.

Keywords: Fractal Company, Decentralized Production Planning and Control

Introduction. In order to remain competitive, enterprises are forced to optimize their production by constantly adapting to the highly dynamic environment. In today's economy, companies are facing increasing competition and high cost pressure in an ongoing globalization of the market. The product life time, the quantities, and batch sizes have decreased. At the same time the service life of products and the number of variants has risen continually. Complex corporate joint ventures and a radical change in customers' behavior represent further evidence of an increasing complexity. A shift is recognized from a sales-oriented market to a customer-oriented market. More and more, customers expect services and products to be tailored to their specific needs. High price/quality and high service/delivery performance is expected from the supplier [1]. In Fig.1 the essential characteristics of the market conditions are shown.

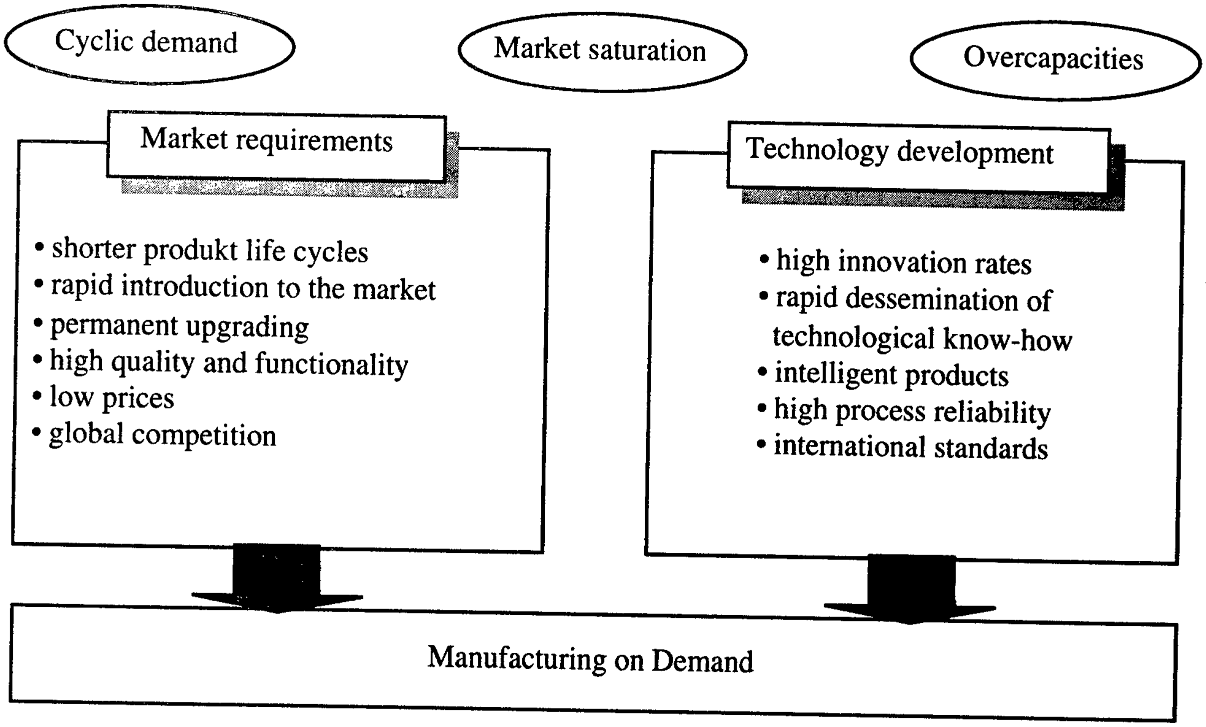


Fig. 1: Customer-oriented manufacturing under turbulent market conditions

Corresponding to the turbulent market conditions, flexibility has become an important factor in the development of new production concepts such as, for example, the Fractal Company [2]. The concept is based on small but powerful and dynamic manufacturing units, the fractals, which operate in a very customer-oriented way within a production network. Manufacturing on customer demand requires flexible production facilities which can quickly adjust to the market changes.

The management of the production of these fractal manufacturing networks is the key to a companies' competitiveness. Besides an adapted product structure there are two important aspects to be met when realizing this type of production:

1. Production structure based on the system technology in order to obtain a comprehensive and autonomous performance optimization.
2. Integrated, decentralized production planning and control [3] in a flexible and dynamic manufacturing network.

Fractal Production. Manufacturing enterprises are traditionally structured on the basis of labour division. Their structure still follows the philosophy of optimization through detailed planning and application of the classical methods of factory focusing on the goal of maximum utilization of the existing resources as regards technical capability and time. For the first time, the philosophies of the fractal [2] or agile enterprise [4,5] are breaking with this tradition as they favour self-organisation and self-optimization: The employees themselves are responsible for the layout of the performance centres.

Figure 2 illustrates the principles of the fractal factory as defined by Warnecke [2]. Each business unit acts as an autonomous factory which is integrated within a communication network. In the view of organisation the business units are similar. They optimize themselves and are integrated in a hierarchical system of objectives and managed in an organisation network.

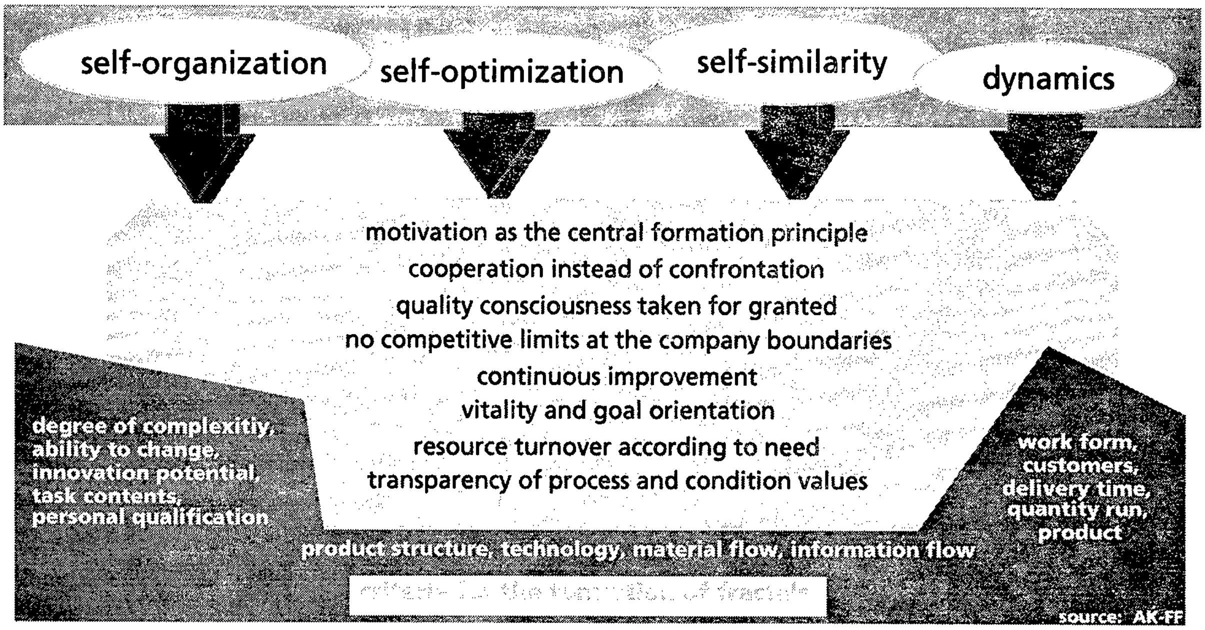


Fig. 2: Principles of the Fractal Factory [2]

A field of production such as the mechanical production, is a complete system consisting of several subsystems, for example the manufacturing cells. The subsystems again can be divided into further subsystems. They are linked with each other according to certain laws like the workpiece or tool cycle. The methods of Total Quality Management (TQM) are applicable to the customer-supplier relationship between the individual elements.
Today, some enterprises are already entrusting the suppliers of the machines and facilities also with their operation including the provision of technology, tools and equipment as well as the personnel. Thus, the machine supplier himself becomes an operator. In this value adding process, he takes the responsibility for the utilization and optimization as a provider of services with the machines produced by himself.

Fractal Production Planning and Control. Market demands for short delivery times make planning and control of multi-stage, customer order-oriented production difficult and complex. So far, neither rough scheduling PPC-systems nor disjoint local control stations, based on conventional forms of production organization, have been able to manage the enormous coordination effort for planning multi-stage linked production. Thus, the objective was to develop an integrated system from globally coordinated planning and short-term production control with distributed cooperative local control stations, on the basis of the Fractal Company concept. The expected benefits of such a system are reduced inventories and lead times, and an increase of the delivery reliability. In the past few years, the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Stuttgart has developed a system architecture which meets these demands. A prototype is currently being used successfully in the industry.

System Architecture. The system architecture for the production planning and control of fractals consists of three levels (Fig. 3). The long-term evaluation of the production structure takes place on the first and upmost level where the intensity of relations among fractals is being analyzed. The degree of intensity is measurable by the number and the nature of planning conflicts.

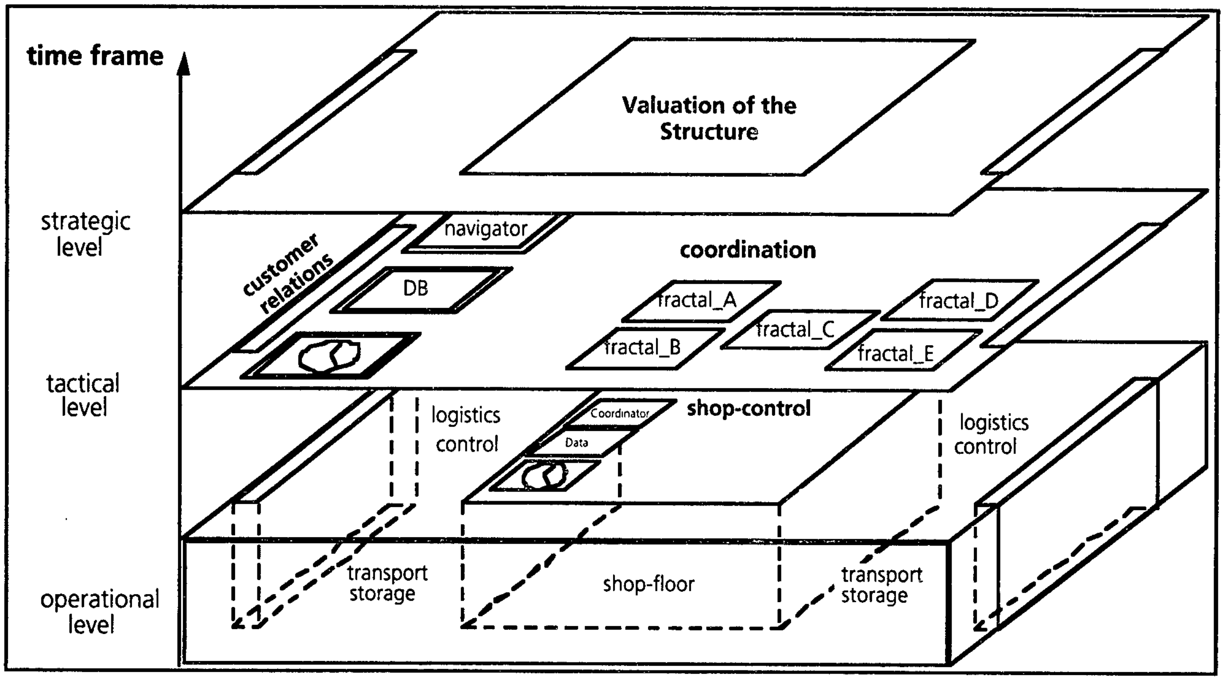


Fig. 3: Planning levels

The second level contains an outline model of the entire production. The function of this level is to visualize a holistic view of the production, in order to determine the overall guidelines for the decentralized units.

The function of the third level with its detailed, autonomous models of fractals is to carry out short-term planning and control. The guidelines are transformed into feasible plant programs, replannings are issued in case of sudden breakdowns, and conflicts among fractals are resolved.

Modeling. The generation of models happens in two steps. The first step is to develop a detailed model of each individual production unit, called fractal. Subsequently, the outline model of the entire production area derives from these models.

The fractal models are used to represent the production processes within the individual fractals. The fractal models consist of essential elements for the planning and control of the production, such as production, transportation and inventory systems, as well as strategies to control the material flow within these systems. They also contain additional functions, such as statistics and various means to evaluate simulation runs. A predefined, configurable fractal model facilitates the user's task to generate a fractal model.

After the individual fractals have been modeled, the outline model of the entire production area is set up. Here, too, a predefined model of material flow and information flow interfaces exists, of strategies to control the material flow and additional functions, such as statistics and the means to evaluate simulation runs.

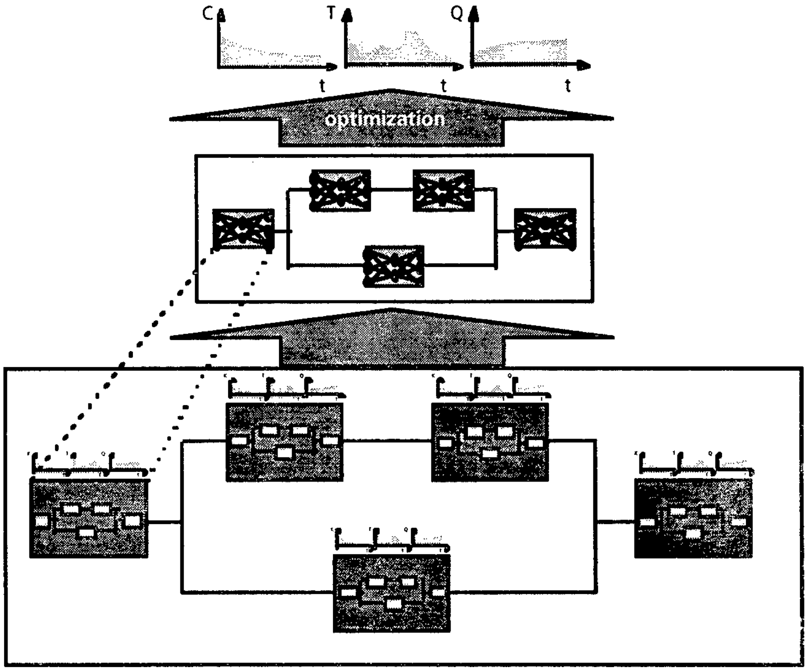


Fig. 4: Hierarchy of modeling

Translating and aggregating the detailed fractal models in the outline model happens automatically by means of soft-computing algorithms (Fig. 4).
A third element in generating models is to build temporary coordination models. In the case of self-organized coordination, a model may consist of two fractal models. This modeling type is prompted by conflicting planning operations of fractals with direct customer-supplier-relations. The components of the coordination model correspond with those of the fractal model.

***Planning procedure*.** In contrast to the modeling procedure, the planning cycle starts with rough planning in the production model (Fig. 5). The main function of rough planning is to schedule customer orders. In a backward scheduling operation based on delivery dates the latest possible planning date is determined. The resulting order sequence is used as a basis for a subsequent forward simulation to schedule the fractal orders, while taking into account the current stock situation within the fractals. A fractal order is equivalent to a 'time window,' defining the time during which a customer order is processed in a fractal. The fractal order is used as a guideline for detailed planning in the fractal model. In general, this provides sufficient freedom for planning and scheduling, so that the fractal planner is able to achieve internal optimizations. The rough planning process is followed by detailled planning in the individual fractal models. If it is not possible to find a planning variant satisfying the overall guidelines (i.e. time allowances for fractal orders) during detailled planning, this might affect the schedules and planning of subsequent fractals. In this case, the coordination model comes into effect, which has two strategies available to resolve the conflict. The first one is called Program-controlled coordination: it combines all planning variants of conflicting fractals (generating a disposition window that considers both fractals) to find out the optimal combination still complying with the overall guidelines (Fig. 6).

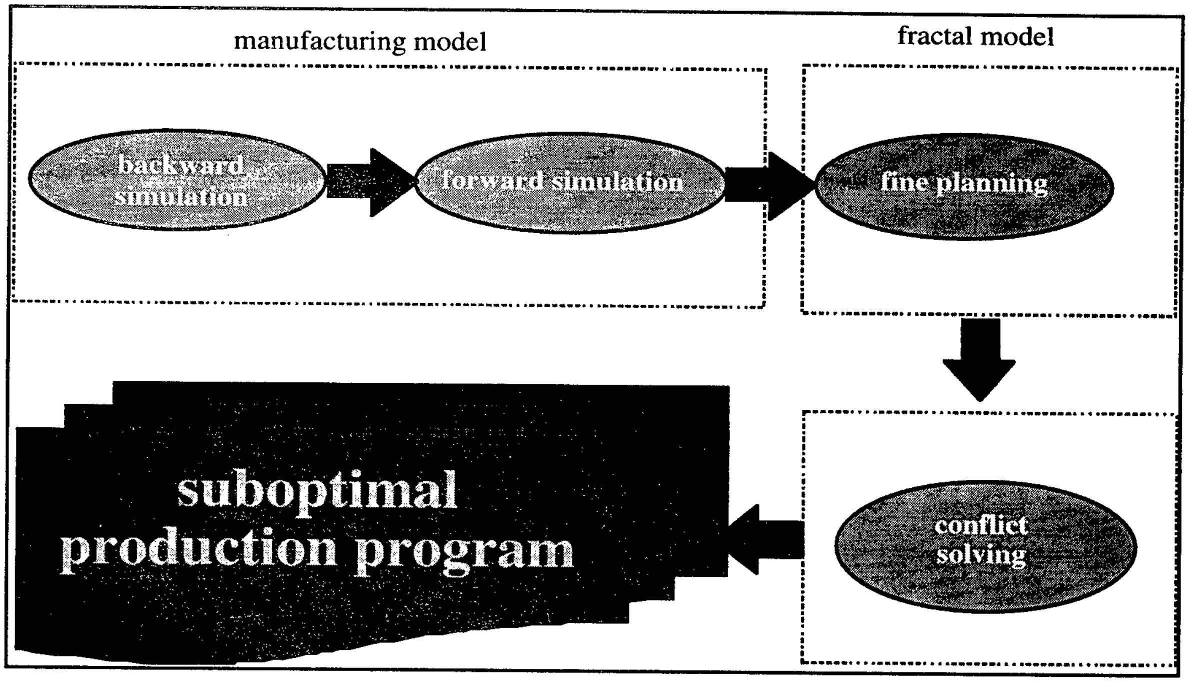


Fig. 5: Planning Procedure I

If there is still no solution to be be found without neglecting the overall guidelines, then the second strategy for problem-solving becomes effective. It is based on a method called Self-organized coordination. The fractals now make use of a model containing all elements of the conflicting fractals that are relevant for the period of conflict. The target of this joint planning act is to comply with the combined disposition window of both fractals.

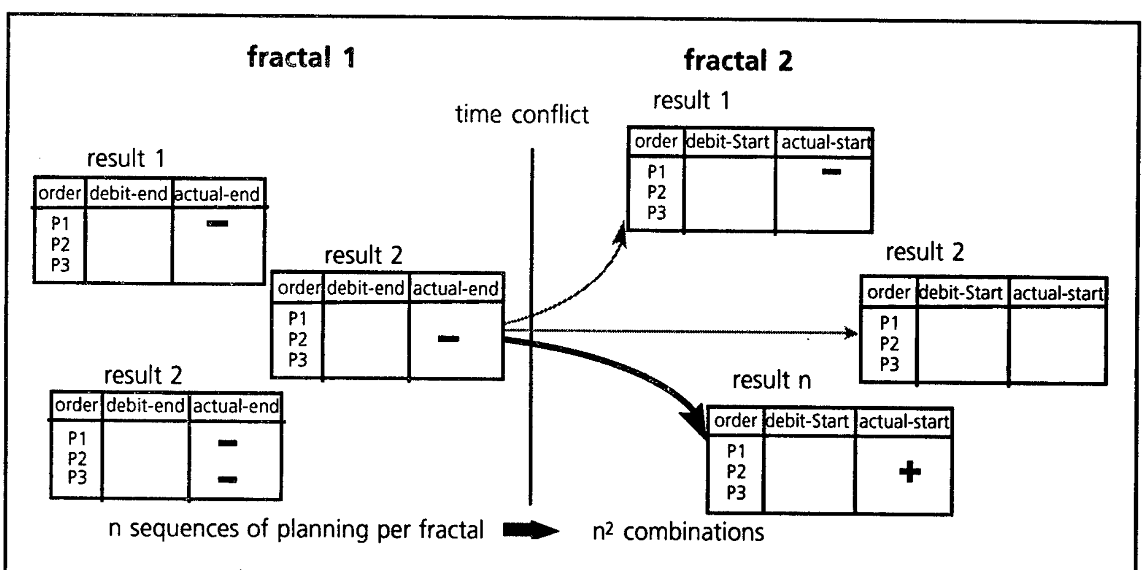


Fig. 6: Programme-controlled coordination

Benefits and Applications. The advantages of the described system for the planning and control of production fractals are, for the most part, a result of its architecture.

Much of the modeling, model maintenance and planning tasks is shifted to the production area. This increases flexibility and up-to-dateness of the planning system when changes in production occur. The know-how existing in the production area can be exploited for planning operations. With the distribution of planning activities over several production units, a simultaneous run of planning processes is achieved that results in an accelerated planning procedure. Since the fractal staff themselves determine the planning results, the planning is more accepted and more reliably fulfilled. The scope of action is enhanced by transferring planning responsibility into the production area and, in effect, stimulating a continuous improvement of the production process.

The automatic generation of a model for the entire production area ensures that a global optimum is pursued. Enveloping complex structures into fractals leads to more transparency and facilitates the analysis and evaluation of production.

Clearly defined customer-supplier-relations among the fractals have the effect that weak points are discovered on time. The tool to settle conflict situations (cooperation model) supports the user in eliminating the deficiencies and also promotes cooperation and the common understanding that solutions can only be found in a joint effort.

A practical example illustrates the benefits of the described system:

In a large European enterprise the networking of decentralized units through information technology changed the whole planning system from an accumulation of disjoint production units to a customer-oriented planning cosmos. Increased planning reliability and the early recognition of conflicts and bottlenecks together with the simultaneous problem-solving support of the planning system created the necessary foundation of trust for a stronger networking of production. Transparency and orientation of all planners towards customer demands had an impact on the management ratios of the considered enterprise.

A more consistent networking of local planning units and the global coordination of planning resulted in a 30 % reduction of work in progress. Since production in the described enterprise is aligned to customer orders, the reduction of work in progress-inventories is directly connected to a reduction of lead times (by 30 %) and delivery times. As the work in progress amounts to the value of $ 100 to 150 million, the reduction of inventory already results in savings of several million dollars each year.

It is important to make sure that the utilization of the plants and the total throughput is not affected negatively by the above described measures. The consideration of technological restrictions within the global planning coordination even achieved and improved composition of batches and reduced the percentage of set-up times. At the same time, the early recognition of conflicts and the common target alignment increased the schedule reliability for customer orders. In the past, very few customer orders were delivered on time to the finished products storage. Today, a delivery reliability of over 90 % has been achieved. This, in return, affected the distribution, so safety buffers between the customer due date and the delivery to the finished products storage hardly exist any more. Inventories could be further decreased, resulting in a very positive effect on the capital tie-up costs at the end of the value creation process.

Due to the improvement of these performance measures, the enterprise became more flexible and agile on the market and economically more efficient in production. The described optimization of logistics parameter induced IPA to advance the architecture that has been described above. At present, the focus of work is engaged in a heterogeneous and international production networks. Companies are particularly interested in the rapid adaptability of these networks.

Conclusion. This paper describes a strategy for the future production based on the idea of fractal and agile enterprises. The actual aim is to realize a type of enterprise operating very closely at the market and only on customer demand at the shortest delivery times possible. To achieve this, the adaptability of the capacities to specific demands has to be increased, and the production planning and control has to be modified correspondingly. Thus, the objective was to develop an integrated system from globally coordinated planning and short-term production control with distributed cooperative local control stations, on the basis of the Fractal Company concept. The expected benefits of such a system are reduced inventories and lead times, and an increase of delivery reliability. In the past few years, the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Stuttgart has developed a system architecture which meets these demands. A prototype is currently being used successfully in industry.

References

1. Matthyssens, P. , C. Van den Bulte (1994). Getting Closer and Nicer: Partnerships in the Supply Chain Long Range Planning. Vol. 27, No. 1, pp. 72-83,

2. Warnecke, H.J. (1993). The Fractal Company, A Revolution in Corporate Culture. Springer-Verlag. Berlin, Heidelberg, New York.

3. Zäpfel. G. (1989). Dezentrale PPS-Systeme – Konzepte und theoretische Fundierung. Neuere Konzepte der Produktionsplanung und –Steuerung. Zäpfel (Eds.) pp. 29-59.

4. Noaker, P.M. (1994). The Search for Agile Manufacturing. Manufacturing Engineering, 11, pp. 40-43.

5. Kidd, P.T. (1994). Agile Manufacturing Forging new frontiers. Addison Wesley, Wokingham.

MULTI-TARGET CONTROL SYSTEMS FOR COMPLEX MANUFACTURING PROCESSES

Abducappar A.Ashimov, Svetlana P.Sokolova
Institute of Informatics and Control Problems,
Ministry of Sciences - Academy of Sciences
Kazakhstan Republic, Almaty 480072, KZ
e-mail: LNS@ipic.academ.alma-ata.su

Abstract. Undesirable environmental effects of industrial development can be significantly reduced by advanced process control strategies. Although criteria of particular process control problems are often mutually contradictive, a reasonable compromise solution reflecting manufacturing and environmental considerations can be found. This problem presents a challenging task due to the high complexity of typical manufacturing processes. Development of the theoretical fundamentals for the solution of relevant mathematical modeling, control, information, software development and implementation problems is proposed. The study utilizes the principle of binarity formulated for variable-structure systems, Razumihin's principle and vector Lyapunov's functions applied to multi-target automatic control systems.

Keywords. Multivariable, nonstationary processes, binarity principle, Razumihin's principle, vector Lyapunov's functions.

Introduction. Typical manufacturing processes in metallurgy, chemical and oil-processing industry, notorious for undesirable environmental effects, are multivariable, spatially-distributed, nonstationary, dynamically-uncertain processes with delays and high level of noise in the information and control channels. These processes are subjected to a wide variety of operational conditions, including emergency-type situations and transient regimes when pollutance can drastically exceed nominal levels.

Mentioned specific properties of manufacturing processes as objects of control determine the nature of the decision tasks which are being made on correspoding levels of the control hierarchy. For example, the automatic control task includes definition and application of control efforts maintaining the process operation within a permissible operational region, established according to the required quality and quantity of the end product. Property defined, control efforts not only result in the elimination of the coordinate errors, detected by primary sensing elements, but also guarantee the acceptability of resultant transient processes. The processes operation task includes selection and coordination of the permissible operational regions assigned to particular control systems. Solution of this problem presents an opportunity for process optimisation with respect to various "manufacturing" and "environmental" criteria.

Designers of control systems, responsible for the formulation and solution of control and operation problems, are usually dealing with a number of uncertainties associated with:

- the absence of the information about all possible disturbances affecting the controlled process (coordinate and parametric),

- the choice of the configuration of mathematical models of the process,

- the choice of the technical means, implementing the required control functions,

- the definition of the parameter correction laws of the mathematical models, and

- specific difficulties of the various stages of control system design.

One of the most prospective ways to control nonstationary dynamic processes under uncertainty is the method utilizing the theory of binarity [1] in automatic control systems (ASC), formulated for variable-structure systems.

This theory presents a basis for the suggested heuristic method of structure synthesis of control systems which implies consistent selection of different sorts of uncertainties, and the choice of a corresponding class of feedback operators. The control law establishes a correspondence between the elements of the class of operational uncertainties and the class of feedback operators thus facilitating an automatic selection of particular feedback control procedures. This approach the structure synthesis of automatic control systems provides an opportunity for further development of the theory of binarity in automatic control systems by widening the class of feedback operators for each sort of uncertainties.

An intention to establish a design methodology for automatic control systems entended for nonstationary manufacturing processes with incomplete information, suitable for particular operational regimes including normal and emergency operational conditions leads to the concept of multi-purpose control system design.

Multi-purpose control implies:

- decomposition of the initial control task into subtasks so that particular operational regimes could be consistent with their own well-coordinated subtasks,

- relative independence of a particular operational regime defined according to its independent subpurpose and individual requirements (desirable dynamics and quality criteria),

- unique representation of a particular operational regime by a mathematical model reflecting its peculiarities and resulting in a specific formulation of a control problem.

Development of control systems of this class requires the formulation of the most sophisticated control methods, capable of purposeful change of the subsystem structure on the basis of interaction of subsystems. The selected class of control systems is known as "hybrid" systems with logical-dynamic principle of operation. This term implies the concept of hierarchy of system complexity levels, the highest of which defines the logical control task of the system structure.

Main results. The goal of this project is the development of the methodology for the design of multi-target automatic control systems including the information technology and software design.

Specific scientific research tasks are:

- formulation of the heuristic procedure for structural synthesis of multi-target automatic control systems,

- definition of the existence conditions for task decomposition of the parametric synthesis problem for corresponding one- and multi-dimension control subsystems,

- definition of the aggregated comparison system for multi-target automatic control systems and investigation of its dynamic properties,

- development of the information technology of the software intended for multi-target automatic control systems.

A heuristic procedure for structural synthesis of the automatic multi-target control system is developed. The system defines particular subpurposes of control subsystems depending on the operational regime using the binarity principle and requiriments for reliable feedback formation. The logic control of the system's configuration is based on the information on the desirable dynamics of system's motion within the corresponding regions of operation, and on the dynamic state of the controlled process. This problem is solved using the binarity principle. The control law, prescribing the feedback operator selection in the corresponding control subsystems, is defined via piecewise continuous functions, and implemented by control algoritms with the pulse-frecuency modulation.

Such mathematical peculiarities of the described multi-target control systems as variable dimension of the state space, significant nonstationarity, and discontinuity of the motion trajectory is addressed. Investigation of the dynamic features of one- and multi-dimensional control subsystems is carried out by direct Lyapunov's method application with the scalar and vector Lyapunov's functions. The Razumihin's principle [2] is extended for irregular automatic control systems of multi-target class.

Definition of the solution existence conditions facilitates the development of the computational algorithms of the parametric synthesis which can be utilized for automatic control systems design. The parametric synthesis algorithms is applied via a specially designed comparison system (reference model-type). The design procedure of the aggregated comparison system and investigation of dynamic properties of the resultant multi-target automatic control systems is perfomed.

References

1. Emelyanov, S.V. (1985) Binary control systems. IRIMS, Moscow.
2. Razumihin, B.S. (1988) Irregular systems stability. Nauka, Moscow.

AN ENVIRONMENT FOR DEVELOPMENT OF RULE-BASED PRODUCTION PROCESS CONTROL SYSTEMS

V.M.Shpakov and M.V.Shpakov
St.-Petersburg Institute for Informatics and Automation of the Russian Academie of Sciences
39, 14th line, St.-Petersburg, 199178, Russia
e-mail: vlad@isa.iias.spb.su

Abstract. The report is devoted to description of a set of interacting tools which composed some environment (named ENVICON) which is to facilitate development of production process control systems. The control system (CS) development with ENVICON consists in working out executable rule-based specifications of CS and simulation model of the object controlled and in testing the system with the production process imitation. When developing ENVICON we first of all kept in mind discrete - continuous production processes which take place in a chemical industry and the like.

Keywords. Control Systems, Simulations, Discrete-Event Systems, Production Processes, Rule-based programming.

Introduction. There is a great demand for developing production process control systems and consequently for means which can facilitate control system designing. There are a lot of different CAD means for control system analysis and design. It is especially so for conventional or classical control systems and maybe a bit less for modern ones. But it must not be stated that the means market is sated. As far as we can judge there are some niches which are void still.

The fact is that the good means must meet some requirements which are not easy to be satisfied. First of all it must be a set of interacting and coordinated tools i.e. an environment. The environment is used not only for CS creation but also for its exploitation and maintenance. In other words the control system can not exist without the environment and therefore the environment cost is a part of the CS first cost. Now let us point to main requirements to the environment. It must:

- provide for executable specifications of rather wide class of control algorithms;
- provide for executable specifications of production process simulation models and support the process imitation to give an opportunity of stepwise testing the control system under design;
- be easy to use by the process engineer without special training;
- not be too expensive;
- provide for necessary high performance and precision of the process control.

As far as we know G-2 of Gensum is one of the best of existing CACSD. It is a real time expert system environment which has very convenient visual and object-oriented tools for CS creation. It provides development of a rule base to specify algorithms and models, the rule structure being rather arbitrary. When designing CS of very complicated and expensive object (such as a rolling mill, for example) G-2 usage is a completely good choice. But it is rather expensive and one must spend time to study it. On the other hand there are many production processes which are not so complicated and there is a great variety of requirements to the quality of controlling them. A competitive decision may consist in developing a set of

CACSD environments oriented both to kind of production processes and to kind of control. So each of these environments may be based on a limited work space for the process model and the control system states representation and on a limited set of strict rule structures and functional dependencies for specification of the process and CS. We consider here a structure of such a simple CACSD environment and discuss its using for CS development. The environment oriented on discrete-continious production processes was implemented and is now under research.

Specification. Developing CACSD environment we must ensure specification not only the CS itself but the production process simulation model (the object model) also. Fortunately the specification of the controller usually not much differs with the one of the object controlled. So hereinafter we consider specification problems as a whole, doing remarks if some means of specification has to do only with CS or with the process model. A structure of any model depends on a real world entity simulated and on the goal of simulation. The only goal of the process modelling is to imitate it for testing the control system. A great variety of approaches was proposed for dynamic system specifications [1].

Choosing proper specification formalism we based ourselves on developing CS for industry installations. Such an object has different devices and mechanisms, which can be switched on or off (open or close), and which functioning associates with changing some continuous parameters (pressure, temperature, level and the like). So current discrete-state technology is suited for this case and combined discrete event and differential equation system specification formalisms (DEv&DESS) may be used [1, 2]. This formalism is fit for specification both of control systems [3, 4, 5] and of object models [6, 7, 8]. As to the process model specification it is desirable to reduce this formalism in order to simplify its implementation. It is possible since the process simulation is necessary only for testing CS under design and there is one and only question the model must to answer, namely "what is the process state after time augmentation". Let us consider specific forms of the formalism we used in our CACSD environment prototype ENVICON.

The work-space of both the CS under design and the object model (OM) consists of a state vector SV with Boolean coordinates to represent discrete-event part of the system and a parameter vector PV with real coordinates to represent continuous one. The control system functioning (production process imitating as well) consists in evolving (updating, transforming) its state in response to current state changes and to external impacts. So a global system state comprises not only internal state variables and internal parameters but external discrete and continuous impacts also. For OM external impacts are CS discrete commands and its continuous settings. For CS they are OM output states and parameters getting by the feedback. Besides there may be operator's impacts. The whole system (CS and OM) evolving is a stepwise process, on each step of which after time augmentation the current states and parameters must be inferred from the previous ones. The shorter the time interval is the better the process precision. To speed up the inference procedure we divide state and parameter vectors in some subvectors depending on the way of their coordinates updating. So the state vector SV includes the following Boolean subvectors: external independent impacts X1d, external feedback impacts X2d, internal device (or "hard") states S1 and internal parameter (or "soft") states S2. The parameter state (a coordinate of S2) is a Boolean variable which value depends on the parameter being in some predetermined range. As to the parameter vector PV let's at first consider the object model parameters. Their current values in many cases are not meaningful for CS testing. An event in the system occurs only if some parameter enters some predetermined subrange. So OM parameter changes may be simulated more or less approximately. It is rather effective and admissible to simulate most of OM parameters changing in accordance with the following linear equation:

$$p_i = p_{i_0} + k_i \cdot \Delta t,$$

where p_{i_0} and p_i are current and updated values of parameter p_i, k_i is an assigned real coefficient (p_i change rate), and Δt is time augmentation. Obviously k_i must be dependent on the system state. The parameters which changes may be represented with a rate we name kinematic parameters and include them in subvector P1. All the kinematic parameters change rates make a set K. This way is also suited for CS time relays representation. The parameters which must be represented with differential dependencies (servo drives in OM and PID controllers in CS for example) we named dynamic ones. As a result in the parameter vector PV we discern the external feedback subvector Xc, the kinematic subvector P1 and the dynamic one P2. Now let's consider means of the control process (as well OM imitation) specification. The first type rule determines the current "device" states depending on the previous global states. It is a kind of "if ... then" rules.

$$s1_{i_1}, s1_{i_2}, \ldots \leftarrow s_{j_1}, s_{j_2}, \ldots, \tag{1}$$

where $s1 \in S_1$, $s \in (S_1 \bigcup S_2 \bigcup X1_d \bigcup X2_d)$. Any variable in an action left hand side (lhs) of the rule is true if a conjunction in its condition right hand side (rhs) is true. The second type rule determines the current "parameter" state depending on previous values of some parameters. The conditional part of the rule has a form of an inequation and make it possible to specify system constraints.

$$s_{2_i} \leftarrow (p_{j_1} \geq a + p_{j_2}) \text{ and } (p_{j_1} \leq b + p_{j_3}), \tag{2}$$

where $s_2 \in S_2$, $a, b \in R$, $p \in (P1 \bigcup P2 \bigcup X_c)$, The "parameter" state s_{2_i} is true if the rule rhs is true. This type rule enables to make an event in the system when some parameter enter the range predefined relatively to the other parameter. There is a default rule concerning these two rules: any state variable (s_1 or s_2) is false if there is no rule (1) or (2) which makes it true.

Let's move on to means of the parameters evolving specification. The third type rules determine the kinematic parameters change rates depending on the global states.

$$(k_i = c + p_j) \leftarrow s_{j_1}, s_{j_2}, \ldots, \tag{3}$$

where $k \in K, c \in R, p \in (X_c \bigcup P1 \bigcup P2), s \in (X1_d \bigcup X2_d \bigcup S1 \bigcup S2)$.

One can see from (3) that the change rate value of one parameter may be dependent on the value of another parameter (not only kinematic one). If rhs of the rule is false then the rule determines a zero rate and the parameter does not change on this step. If there are several rules (3) concerning the same parameter then its rate is the sum of the rates determined with every rule.The rate k of any kinematic parameter p1 is equal to zero if there is no rule (3) which makes it different from zero.

As to the dynamic parameters they are specified as functions of other parameters or of the parameters and of time.

$$p2_i = f_n(p_j, t), \tag{4}$$

where $p2_i \in P2, p_j \in (P1 \cup P2 \cup X_c)$, t - time. The set of the functions f_n may be oriented to the specific application. So for developing PID controller it is necessary to have proportional, integral and differential dependensies. For servo drives imitation may be necessary non-linear functions (relays, saturations and other), aperiodic and oscillatory dependencies. To realize various logic control laws the rule of a "contact" type must be included in the rule base.

$$(p2_i = p_j) \leftarrow s_k, \tag{5}$$

where $s_k \in SV$ - a condition of the contact switching. This rule is necessary also to imitate switching devices on and off. Considering all aforesaid the formalism used by ENVICON may be represented in form of the following tuple:

ENVICON = $(X1_d, X2_d, X_c, S_1, S_2, P_1, P_2, K, Y_1, Y_2, \delta_1, \delta_2, \delta_3, F, \lambda_1, \lambda_2)$,
$X1_d$ - external discrete independent impacts;
$X2_d$ - feedback discrete impacts;
X_c - feedback continuous impacts;
S - "hard" or device state variables;
S_2 - "soft" or parameter state variables;
$P1$ - kinematic parameters values;
$P2$ - dynamic parameters values;
K -kinematic parameter change rates;
Y_1 - discrete outputs;
Y_2 - continuous outputs;
δ is a function, state dependent state transition specification, $\delta: (X1_d \cup X2_d \cup S_1 \cup S_2) \rightarrow S_1$;
δ_2 is a function, parameter dependent state transition specification, $\delta_2: (X_c \cup P_1 \cup P_2) \rightarrow S_2$;
δ_3 is a function, parameter change rate transition specification, $\delta_3: (S_1 \bigcup S_2) \rightarrow K$;
F is a set of functions determining dynamic parameters time variations;
λ is a discrete output function, it is a mapping from $(S_1 \bigcup S_2)$ to Y_1;
λ_2 is a continuous output function, it is a mapping from $(P_1 \cup P_2)$ to Y_2.

This formalism may be presented as a structure shown in Fig. 1.

Specification Example. To demonstrate using the ENVICON specification an example of a well known system is presented. It is a model of a controllable barrel filling process used in a brewery and chemical industry [1].

Here is a description of the process in English. If there is a permission to start working and the barrel is ready then it is necessary to begin opening the tap. When the tap opening is more zero the level begins to go up. The level raising rate is proportional to the tap opening. When the level achieves the pre-assigned value it is necessary to begin closing the tap and do it until the tap is close.

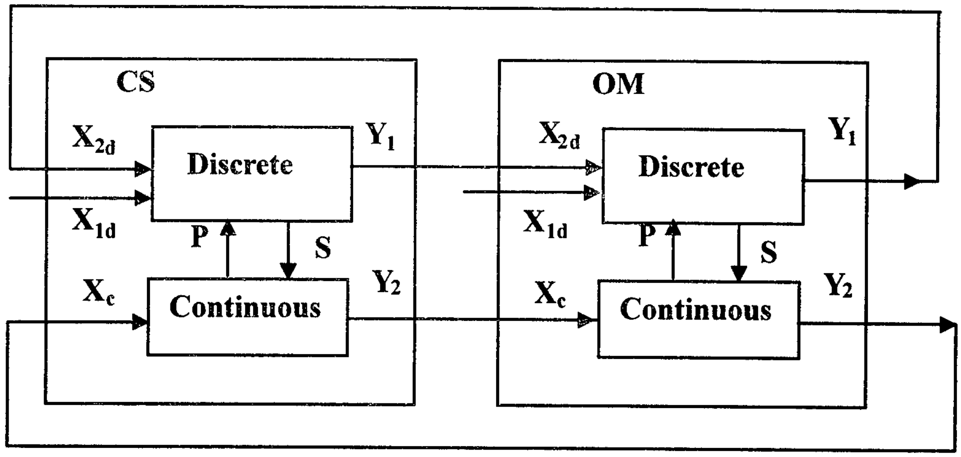


Fig. 1 Specification formalism structure

Now let's do this system specification by means of the formalism considered. We hope that names of the variables make their application purposes clear enough.

Here are the subvectors:

X_d = {*Barrel_is_ready, Start*} X_c = {*Level_assigned*}
S = {*Open_Tap, Close_Tap*} S_2 = {*Tap_is_open, Tap_is_close, Enough*}
P = {*Level, Tap_opening*}

Here are the rules of three types determined:

Open_Tap ← *Start, Barrel_is_ready*, not *Tap_is_open*
Close_Tap ← *Enough*, not *Tap_is_close*

Tap_is_open ← *Tap_opening* > 0.99
Tap_is_close ← *Tap_opening* < 0.01
Enough ← *Level* > *Level_assigned* + (-0.6)
Tap_opening, k = 0.1 ← *Open_Tap*, not *Tap_is_open*
Tap_opening, k = -2.0 ← *Close_Tap*, not *Tap_is_close*
Level, k = *Tap_opening* ← *Start*, not *Tap_is_close*

The initial variables values may be arbitrary. The following combination makes sense: *Tap_is_close* is true, the rest Boolean variables are false, *Level* = 0, *Tap_opening* = 0, *Level_assigned* > 0.

Control and production process imitation algorithms specification with the state transition rules has some peculiarity. The fact is that the set of rules has no time as global independent variable which always is present in the specification of the procedure type. A succession of actions can not be assigned by the designer directly, but only through events which are to take place. So there is necessity to develop programming technique based on some set of rules. It may be stated now that some discrete-event algorithms may be represented with help of the rules rather easily and naturally whereas for the others (control structures, for instance) it is in prospect to find expressive patterns of rules. Let us consider an example of a "while" type cycle realization with rules described. Let it is necessary to run a time relay when some condition is false, then after assigned time expiring check the condition and if it is false run the relay again. And do so while the condition became true. The process may be interpreted as a sand clock running. Here is rule based program of the algorithm.

X_d = *{ Condition }* S = *{ Decrease, Increase }*
S_2 = *{ Position_1, Position_2 }* *P1 = { Sand }*

Increase ← Position_1, not Condition
Increase ← Increase not Position_2
Decrease ← Position_2, Not Condition
Decrease ← Decrease, not Position_1

Sand rate = 1 ← Increase
Sand rate = -1 ← Decrease
Position_1 ← Sand <= 0
Position_2 ← Sand >= 1

ENVICON Implementation. ENVICON is coded in Turbo Pascal 7.0 under DOS. Each variables is represented by a record type which has a string field for its name and another field for its value, Boolean for the discrete variables and real for the continuous ones. The record for the continuous variable besides that has a real field for a value change rate. Any set is represented by an array of corresponding records.

Each rule is represented by a record. The record for the first type rule has two pointers to state variable lists, one for lhs and another for rhs of the rule. The record for the second type rule contains fields for used variables and constants. For variables there are fields only for the appropriate array indices. The record for the third type rule carries fields for the parameter index, the parameter change rate value and the pointer to the state variables list for rhs of the rule. The whole rule base consists of three lists of aforesaid rule records. When editing the rule base or the model state it is much more convenient for the user to deal with variable names. But when updating the state it is less time consuming to deal with the indices of the variables.

The procedure of the model updating is rather simple. It consists in time augmentation and substituting at the rule rhs the values of the variables computed on the previous step and computing new values of the variables of the rule left hand side.

The worked out control system and the object model may be saved in files, each a subvector and a rule list in a separate one. All the files related to OM and to CS will be saved in separate directories.

ENVICON Description. ENVICON has means for creation and editing the above-mentioned subvectors and rules, means for controllable production process simulation handling and for forming and displaying a trace of discrete-event process simulated. A reduced architecture of ENVICON is presented in Fig. 2. ENVICON's main menu has the following items: **CS, OM, IniState, Edit, Time, Pause, Control** and **Results**.

CS and **OM** are for choosing one to deal with.

IniState is for choosing an initial state. The fact is that when simulating the process may be suspended and the state occurred may be saved to use afterwards as an initial one.

Edit is for composing the vectors and the rules and for assigning values to variables.

Time is for choosing time mode. ENVICON provides two modes of time augmentation. In real time mode the timer reading and the user defined coefficient are used for Δt calculation.

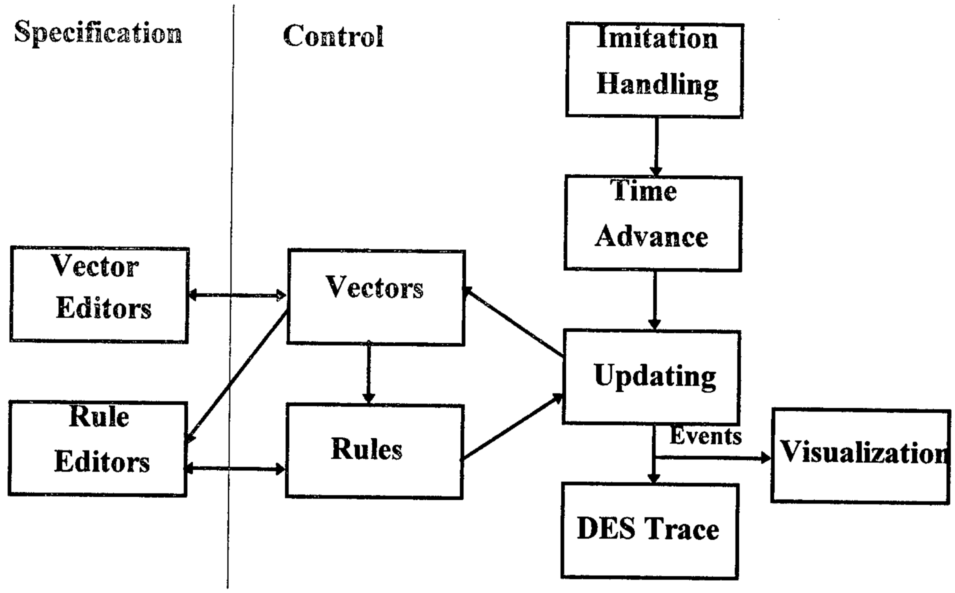


Fig. 2 ENVICON architecture

In this mode the simulation time is directly proportional to the real object time and accuracy depends on the internal performance of the computer and the model dimensions. In model time mode Δt is assigned by the user. In this case accuracy is inversely proportional Δt and the simulation time depends on the computer performance and the model dimensions.

Pause is for the specification of stop rules. They are the situations which if occur must pause the simulation process.

Control is for the control and the production process imitation. When this item is chosen the values of any variables both of CS and OM may be displayed.

Results is for displaying a list of events (trace of the process) which occurred separately in CS and in OM during the simulation.

Conclusions. A control system for an industry installations of compressed gases pumping and issuing was worked out and implemented with ENVICON. It was also used for testing a control system of a refuel plant and proved its efficiency. It turns out that ENVICON is rather helpful for joint simulation of the plant and the control system while designing one or another or both together. There are some directions of improving ENVICON. One of them consists in providing it with ability to compose a complex control system (object model) out of subsystems (submodels) specified and tested beforehand. Now we also work to supply it with built-in procedures of the rule base verification and validation.

References

1. F.Pichler, H.Schwartzel (eds.) (1992). CAST Methods in Modelling. Springer Verlag, Berlin, p.376.
2. F.E.Cellier (1986). Combined Continuous/Discrete System Simulation - Application, Techniques, and Tools. Proceedings 1986 Winter Simulation Conference, Washington D.C., pp.24-33.

3. Kimon P. Valavanis, George N. Saridis (1990), A Review of Intelligent Control Based Methodologies for Modeling and Analysis Systems. Proceedings of the 5th IEEE International Symposium on Intelligent Control, Philadelphia, Penn.
4. Д.А.Поспелов (1986). Ситуационное управление: теория и практика. Наука, Москва, 288 с.
5. Peter J.G.Ramadge (1989). The Control of Discrete Event Systems. Proceedings of the IEEE, Vol. 77, No. 1, pp. 81 - 98.
6. Kim Tag Gon, B.P.Ziegler (1990). The DEVS-Scheme Modelling and Simulation Environment. In Knowledge-based simulation: methodology and application, eds. Paul A.Fishwick, Richard B.Modjeski, Springer Verlag, pp. 20 - 35.
7. B.P.Ziegler (1989). DEVS Representation of Dynamical Systems: Event Based Intelligent Control. Proceedings of the IEEE, Vol. 77, No. 1, pp. 72 - 80.
8. J.Rothenberg (1990). Knowledge-Based Simulation at the RAND Corporation. In Knowledge-based simulation: methodology and application, eds. Paul A.Fishwick, Richard B.Modjeski, Springer Verlag, pp.133 - 161.

INFORMATION TECHNOLOGY OF AUTOMATED SCIENTIFIC RESEARCHES IN SHIPBUILDING

Alexander I. SUVOROV

The First Central Research Institute of RF Defense Ministry
Department of Scientific Research Automation
30, ul. Chapaeva, St. Petersburg, Russia

Abstract. In this report some problem items of automated scientific researches of complex objects properties are considered. Directions of their decision are formulated.

Keywords. Information technology, scientific researches, automation, shipbuilding, complex object.

Introduction. The organization of work in the field of scientific researches on problem of the military shipbuilding (MSB) provides fulfillment just as of fundamental and search work, so applied researches. It in turn predetermines presence of two basic purposes:

- Reception of new knowledge;
- Development of methods, models and techniques of an estimation and analysis of properties of naval ships and vessels on the basis of realization of the appropriate experimental and theoretical researches.

The effective achievement of the purposes formulated above is possible on the basis of modern information technologies (IT) researches, which frequently named as new information technologies (NIT). The structure of concept formation NIT is considered by the author in [1], in [2] the basic distinctive features, allowing to speak about fundamental changes in IT researches, occurring in the last years, are reasonable. These changes in methods and means of scientific researches render decisive influence to efficiency of their realization, radical improvement of their quality, completeness and timeliness.

We shall formulate and reveal a number of the basic thesis, forming a basis of the NIT concept.

The thesis №1 " About plurality of mathematical models ". Plurality of models we shall consider from two complementary positions. First - plurality of models, describing the same process, is caused by research of its various sides, application of series of criteria for synthesis of models, in the greatest degree satisfying the purposes in hand of models use.

The second position is defined by the fact that for the decision of one research purpose, for example, forecasting models of various classes (the differential equations, statistical models, MGUA models and so on), with the set of inherent properties can be used:

- Adequacy;
- Efficiency;
- Complexity;
- Completeness;
- Adaptation;
- Stability;
- robustness NOISE stability;

- Sensitivity & other.

In this connection that the achievement of all properties by one model is impossible, at different stages of modeling, coordinating with the local purposes in hand, a few models are developed. Each of such models satisfies in the greatest degree one or group of properties, they are not the competitors and are used with greater efficiency for the decision of particular problems. True enough it is necessary to note the special place of models adequacy property, which as a rule, has to be present as unifying property.

Thus, for achievement of the put purposes of models use there will correspond a few MM and each of them will at the same time have some set of desirable properties. Concurrence of the MM purposes and properties to the researcher requirements gives required model.

It is the author's opinion that the conclusion about plurality of the researches purposes and models properties, which results in plurality of the mathematical description of one object (process), is fundamental for ideology of development and methodology of models qualimetry use.

The realization of the given thesis together with natural desire of the researcher to receive model with beforehand given, desirable properties assumes an opportunity of an estimation of MM properties at various stages of its development and use, i.e. creation of some formalized apparatus of MM qualimetry.

Definition. Qualimetry of models - set of methods and criteria of estimation of MM essential properties, being subject to the analysis and comparison for selection of models with desirable properties.

MM qualimetry is based on methods of statistics and numerical modeling, theory of sensitivity, automatic control and information and so on.

It is expedient, for greater efficiency of criteria use, estimation of MM properties to make at various stages of information technology (IT) modeling. By such stages it is possible to refer following ones [3,4]:

- Processing experimental data (ED);

- MM formalization;

- MM algorihtmization;

- Computing experiment (CE);

- Analysis of models.

For reception of MM with desirable properties various criteria of structure formation and parameters definition are used, for example, set of MGUA algorithms [5], and in case of CE planning - various criteria of the plans optimality, that also gives an opportunity for MM synthesis with diverse properties.

The thesis №2 " About integration of researches ". The problem of integration in a complex of shipbuilding researches can be treated, as a minimum, in two aspects.

First assumes development of mathematical models, uniting some fields of knowledge or aimed at reception of a cumulative estimation of the ship as a whole, that allows to speak about allocation of these models in a certain class system models of the analysis of the ships properties [6]. The second aspect is directed on the decision of a problem of integration of results of various kinds of researches of one subject area. In particular it is possible to speak about a problem of analytical models parameters correction, received in theoretical researches (TR), on data of experimental researches (ER).

The unity of experimental and theoretical activity is realized, on the one hand, in integration of specialized kinds of activity, and with other, that the experimental activity always includes theoretical one, and theoretical activity should take into account results of experiments. The good theory proves to be true by results of experiment, on the other hand, success experiment can become stimulus and basis for updating the theory.

For considered problem areas of research of shipbuilding complex objects properties in rather general statement we shall accept, that the theoretical basis for a various type of researches are the theories:

a) For the ER-mathematical theory of experiment, including methods of planning of experiment and statistical processing of tests results;

b) For the TI-theory of modeling, including methods of synthesis and the analysis of models, their transformation, qualimetry of models, computing experiment. The integrating theory for ER and TR are, apparently, theory of identification, allowing to estimate parameters of theoretical models on data of experiments and to correct their known meanings, and also methods qualimetry of models, intended for an estimation of properties of models, and first of all of their adequacy with ED use. In development of ideas, stated in [8], following method of correction of parameters of theoretical models on ED is offered.

Let nonlinear on parameters the model of object looks like

$$y = F(x; a), \quad X \in Dx, y \in Dy, a \in Da, \qquad (1)$$

Where x - vector of entrance coordinates;

y - Target coordinate of object;

a - Vector of parameters, found in result of analytical designing of model. We shall consider it as initial approach, which should be specified , using data nature of experiments. It is supposed, that there is some a*-optimum meaning of a vector and, at which model we shall consider exact. Thus one of problems consists in construction and realization of the optimum plan of experiment, other - in an estimation of errors of model on Dx and Dy infinite sets.

It is offered following the double-stage approach to the decision of a problem.

At the first stage linear on parameters statistical model of a kind is built

$$Y = \sum_{i=1}^{M} a_i z_i(x) = \sum_{i=1}^{M} a_i z_i, \qquad (2)$$

Where a - required factors;

$Zi=Zi\ (x)$ - some, usually sedate, function of coordinates of a vector x.

As the result of full scale of experiment realization there are the sets of experimental data - final sets $wZ_i = \{z_{ij}\}$,

$wy = \{y_j\}, j = 1, N,$ on which systems of the equations of a kind are formed

$$Yi = \sum_{i=}^{M} a_i z_{ij} + \ _i, \quad = 1, N \qquad (3)$$

Or in the matrix form

$Y = Fa + \xi,$ (4)

Where F - the matrix of experiment plan or is simple a matrix of the plan. The estimation of a vector ***a*** defined is in result of the following extreme problem decision

$$a = Arg\ min\ (Y\text{-}Fa)\ T\ (Y\text{-}Fa) = arg\ min\ \xi\ T\ \xi \quad (5)$$

The accuracy of parameters estimation depends on a matrix or choice of points $z_j = \{z_{ij}\}$. In the theory of planning of experiment [9] problems of optimum (on various criteria) choice of these points are solved. Not in detail of various aspects of a problem, we shall note only one important rule of the theory of experiment planning.

For linear regress model of given structure it is possible to construct the optimum plan before realization of experiment. In case of nonlinear models consecutive strategy, requiring of realization additional experiments in the points, not belonging to the optimum plan, are usually applied. Thus, it is possible to hope, that the offered approach allows, basically, to lower the volume of experimental researches at the first stage.

We shall assume, that the model constructed on experimental data is adequate, in statistical sense, researched object on final sets. For a considered class of polynomial models an opportunity to investigate their properties rather full is available and on infinite sets. In case of a positive conclusion about precision of the characteristics of input - output models we shall consider them as reference.

The second stage of the problem decision consists in optimum approximation of model (2) to reference. We shall enter ratio with this purpose

$$J[Y(X;\alpha) - Y(X)] = J[f(X;\alpha) - \sum \alpha_i \varphi\ (x)] \quad (6)$$

And we shall formulate mini-maxi problem of parameters correction

$$\alpha = arg\ min\ max\ J[Y(x,\alpha) - Y(x)]$$

For its decision it is possible to use one of known methods, in particular of casual search. During search the point is searched simultaneously with a vector

$$X = arg\ Jm = arg\ min\ J(\alpha, x), \quad (7)$$

In which the error of function approximation (2) to reference will be minimum, in sense of an accepted measure of affinity. If the condition is carried out

$$Jm <= J\grave{o}on, \quad (8)$$

that a problem of vector corrections consider solved.

The thesis №3 " About IT intelligent support ". One of the basic distinctive features of new information technology is use of software, based on technologies and methods of artificial intelligence, to which recently refer [10]:

- Expert systems (systems, based on knowledge);

- Neuron networks and "illegible" logics;

- Natural - language systems.

ES in NIT, however as in other fields of knowledge, are intended first of all for automation of the non-formalistic problems decision and, thus, much expands the circle of automation means use.

The effective realization of information technology of scientific researches and first of all of modeling requires intelligent support as expert system. It is caused by multialternativeness of possible realizations of the IT decisions and trajectories for particular problems of modeling, various preparing of the users and their orientation in problems of modeling, necessity of accumulation and analysis of the accepted decisions and other reasons.

We shall consider the basic directions of ES use in experimental and theoretical researches in shipbuilding.

For experimental researches questions of statistical data processing, the methods of experiment planning, synthesis of experimental - statistical models, informal analysis of allowable set of the decisions and other require the intelligent support.

In theoretical researches it is possible to allocate 3 basic classes of operations and procedures:

- Analytical researches (formation of conceptual model, formalization of model and its transformation);

- Computing experiment;

- Qualimetry of models.

Separately it is possible to note a problem of optimization of models. For ES TR and ER the following peculiarities are characteristic:

- In basic, they carry out intelligent support of processes of synthesis, analysis, interpretation of models and acceptance of the decisions;

- As a rule we are dealing with a choice of an expedient method;

- Support of technology and correct use of methods at several user levels;

- Realization of a special subsystem " of the store of experience " uses (both unsuccessful, and successful) various methods in different situations;

- On the basis of the analysis of this experience determination of methods ratings and "trajectories" accumulation of acceptance of the correct design and research decisions;

- Application of combined methods of representation of knowledge (systems of products, mechanisms of algebra frames and apparatus of semantic networks).

Generalizing aforesaid, it is possible to note, that we are dealing with intelligent support and purely formation of strategy of the decision of a put problem and first of all of modeling problems. In the integrated plan it is possible subdivide this strategy on problems of a choice of the purpose of a soluble problem, class of model and method of the decision; methods of processing and analysis of data; methods of modeling; strategy of planning nature and computing experiments; strategy of models transformation.

The thesis №4 " About alienation of models ". In the fundamental, academic plan the problem concerns to questions of "another's" knowledge use. Here we shall examine problem aspects of alienation of models from their developer. We shall note, that the basic aspect is the aspect of trust to model, received from other researchers.

For the scientist could concerns to "another's" models with trust, he should have some additional information, which, as a rule, is absent. To such information it is possible to refer :

- Estimation of adequacy to MM and area of its use;

- Methods, algorithms and techniques, fixed in a basis of its development;

- Methods, algorithms and software, used by work with MM on the COMPUTER;

- System of designations and identification of parameters, settlement circuits and restrictions, and also other characteristics, accompanying creation and use of model. The base for mistrust arises, when the listed and other information is inaccessible to the user. One of possible directions of alienation of models from their developer is the approach, offered in [11,12].

It is based on concepts of MM "text" and "context" and methods of their algebra and logic.

The decision of a problem of alienation of models from the developer consists in complete definition contextual variable, as singe, frequently non-considering of a context MM conducts to incorrect its interpretation or to mistrust.

Expediently to enter three levels of alienation:

a) Parameters of models;

b) Mathematical models;

c) Techniques on base MM.

Each of levels has independent meaning, and the realization should go consistently from alienation of parameters to alienation of techniques.

The program realization of the formal apparatus of alienation can be directed on creation uniform parametrical basis-universum for MM, used in NSB. Some approaches to such realization on base SCDB MM MAMOD are resulted in [8,13].

The stated thesis together with a number of the other approaches are fixed in a basis of realization of automated system of scientific researches (ASSR) in NSB as to a component of system of automation of research designing (SARP) "Chertjozh-4", developed and founded practical application in the 1 CSRI DM RF (NSB).

In SARP "Chertjozh-4", ASSR is intended for support of the basic purposes, that is provided by expedient interaction of a lot of subsystems. On the basis of automation of processes of statistical processing of experimental data, planning nature and computing experiments, statistical analysis and modeling, qualimetry of models, realization of computing experiments with mathematical models, transformation of models and their alienation from the developer, advanced means of the graphic analysis and display of the information, and also use of expert system for a choice of expedient strategy is possible to ensure essential increase of efficiency in interests of the military shipbuilding.

References

1. Суворов А.И. О понятиях "технология", "информационная технология" и "новая информационная технология" //Программные продукты и системы - 1996, №2

2. Суворов А.И. Концепция новой информационной технологии научных исследований. Тез. докл. V межд. конф. " Региональная информатика ", Санкт-Петербург. 1996.

3. Юсупов Р.М., Иванишев В.В., Костельцев В.И., Суворов А.И. Принципы квалиметрии моделей. Тез. докл. IV Межд. конф. "Региональная информатика", Санкт-Петербург, 1995, с.90-92.

4. Суворов А.И. Проблемные вопросы автоматизации научных исследований в кораблестроении //программные продукты и системы. - 1996, №4, с.2-8.

5. Ивахненко А.Г., Степашенко В.С. Помехоустойчивость моделирования. - Киев. - Наукова думка, 1985. - 200с.

6. Захаров И.Г. Особенности моделирования задач предметной области в интересах разработки системных моделей исследовательского проектирования (учебное пособие). - С-Пб. I ЦНИИ МО, 1993.

7. Г. Гёрц. Соотношения между экспериментом, моделью и теорией в процессе естественнонаучного познания. //Сб. статей "Эксперимент, модель, теория" - из-во "Наука", Москва-Берлин, 1982.

8. Суворов А.И., Валькман Ю.Р., Соломаха О.Н. Информационная технология единого комплекса исследований в военном кораблестроении. //Программные продукты и системы. - 1994, №4.

9. Налимов В.В. Теория эксперимента. - М., Наука, 1971.-207с.

10. Попов Э.В., Фоминых И.Б., Кисель Е.Б. Статистические и динамические экспертные системы. - М., 1995. - с.125.

11. Валькман Ю.А. Принципы построения алгебры и логики текстов и контекстов математических моделей //тр. III конф. по искусственному интеллекту (КИИ - 92). Тверь, 1992. Т.1.

12. Валькман Ю.Р., Рыхальский А.Ю. Исчисление моделей в исследовательском проектировании. Сб. трудов Межд. конф. KDS - 95, т.2, с. 324-334.

13. Валькман Ю.Р., Суворов А.И. Принципы построения интегрированной системы баз данных моделей. //Программные продукты и системы. - 1993, №3

INFORMATION TECHNOLOGIES IN INITIAL STAGES OF SHIPBUILDING OBJECTS DESIGN

Roald Allik, Mark Radoushinsky
Krylov Shipbuilding Research Institute
196158, St.Petersburg, Russian Federation
e-mail: albert@krylov.spb.su

Abstract. The paper contains a brief review of information technologies used in the initial stages of shipbuilding objects design. Reviewing the technologies the authors have to give a general description of design procedures in the cases specific for shipbuilding not fully known to representatives of other professions.

Keywords. Shipbuilding objects, life cycle, initial stages of life cycle relating to design, information technologies in design, design procedures.

Introduction. Under the shipbuilding objects here one can understand the main production of shipbuilding industry: ships of different types, floating constructions, submersible vehicles and other similar technique which usually are considered as complicated technological systems.

The information technologies are used as for the process of manufacturing such systems so for their operation as far as their utilization saying by other words in all stages of life cycle (LC) the system.

The key object for consideration will be information (electronic) model of shipbuilding object which usually is called a product model.

The paper deales with initial stages of LC relating to design.

In modern market conditions a hightened interest to the initial stages (IS) is connected with necessity to respond effectively and quickly to market needs as well by maximum satisfaction of customer wishes. To make a contract for delivery it is necessary to work out quickly and with quality a contract design (CD) which is one of the main stages of object LC and determines to a large extent the efficiency of its operation. Working out the contract design in some countries as well as in Russia is made with CAD systems used different methods and means of information technologies.

The works begin by making a "zero" variant of information object model. These works are very poorly formalized and their results are largely determined by competence of persons taking decisions.

In the process of designing of a shipbuilding object an essence of carrying out design procedures in IS consists in checking compatibility of interincluding customer requirements and basing their technical realization. In the result of iterations mutually acceptable decisions will be received. After this one can consider that an initial variant of information subject model is made. It must be developed and specified in future but as a rule will not be changedgreately. So the main decisions on design shipbuilding object are taken in initial stages. It is precisely these decisions that determine a future "life" of the object and that is why the price of them is very high.

It should be mentioned that the majority of national and foreign systems of automatic design of shipbuilding objects are oriented on most labour consuming works connecting with production of working drawings. The tasks of initial stages of design are automated much weaker because of difficulties to formalize these works and deficiency of information to make a decision.

Characteristics of the Object and Design Procedure. The shipbuilding object the LC initial stages of which are under discussions corresponds to all formal signs of complicated technical system. Applying to these systems usually one considers a logically connected pair "design object-design procedure" that needs consideration of two models - object and design procedure. The sequence of carrying out design procedures in determined by the peculiarities of the object and presence of the information. But in all cases specification of the object model by stages must be considered as a result. Iterative character of the procedure in IS is illustrated by well known "design spiral" (slide No 1).

Design Procedures and Information Technologies.

Simulation Modelling. When taking design decisions in IS in order to reduce a risk level and uncertanities means of simulation modelling of functioning of the design objects in the systems of higher level such as operation of a ship in a definite production-transport system are developed. As an example a task solution on performance evaluation of investment design (ID) in shipbuilding is considered.

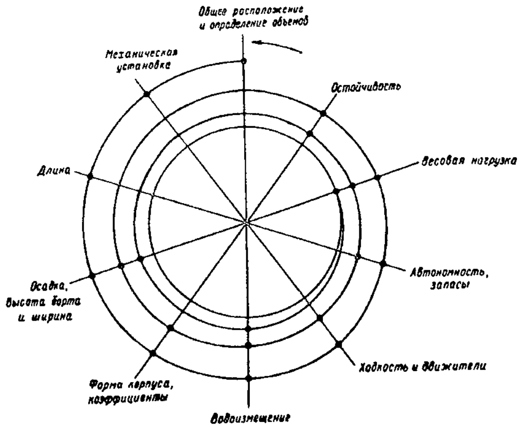
For potential investor the performance evaluation of ID has a paramount significance. A criterion system reflecting correlation between expenses and received results (effects) during LC of ID is used. Specific character of shipbuilding is connected with peculiarities of production and operation of objects: small serial production, large capital expenses, long cycle of building, long operational period and complicated methods of calculations of different kinds of effects. These peculiarities prevent using known universal methods and means of performance evaluation of ID supposing existance of all necessary initial information.

On the strength of specific character of shipbuilding objects the initial information shall be preliminary calculated. The task is solved by development of simulation models ensuring on realization receipt of data on incomes and expenses concerning to a particulary taken ship, transport or transport technology system. The slide No 2 shows schematicly a transport-technology system connected with Shtokman Gas Deposit (Russian Arctic Region). For this system a simulation model having a universal character in the part of its use in different information situation is developed. Having distribution laws of initial parameters and using the model data of expenses and incomes in form of laws of their distributions have been received. After processing the received information is presented to an investor. The model is used and in the absence of distribution laws of initial parameters that corresponds to known situation "game vith nature" (uncertainty conditions). In this case they get a matrix of ID performance evaluation variants. To choose a pereferable variant of investment processing of the matrix is done using known methods of operation theory.

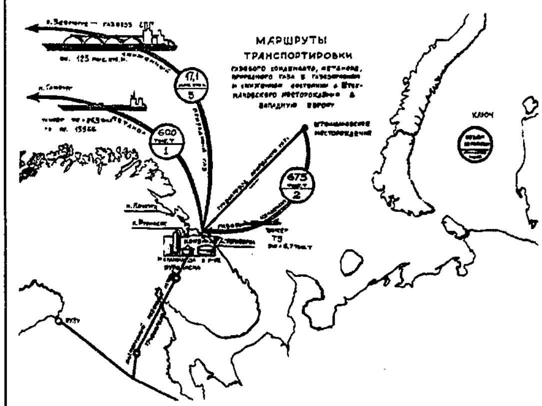
Using simulation modelling is not restricted by the shown example.

Surface Design of External Lines of an Object. External lines being an information base of an object design as a whole shall correspond to many requirements and are a subject of thorough processing. In the majority of the cases to describe the shipbuilding object surface by analitical method is impossible. A lines drawing (slide No 4) is a graphic presentation of complicated three dimensional (3D) surface of the object (slide No 3).

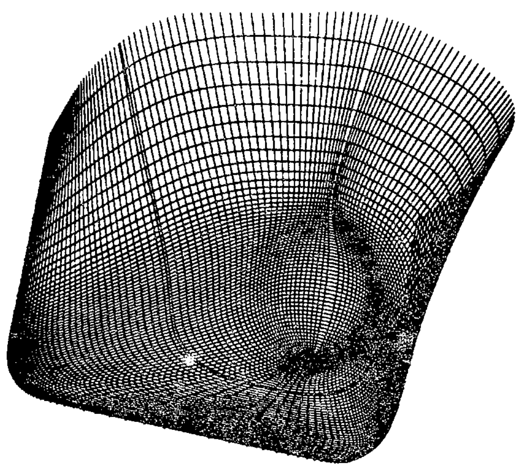
Design and generation of external lines surface is differed. CAD system ensures building of surface in an interactive regime and properly is a mechanical variant of manual work. Usually one has a frame model drawn by hand which in computer is "improved" and can be used as a base for stretching surface. Finally one gets coordinated and smoothed surface. It is necessary to take into consideration that during the process of surface design automatically or manually a whole number of shipbuilding parameters are controlled and "held". So the work does not reduced only to geometrical procedures.



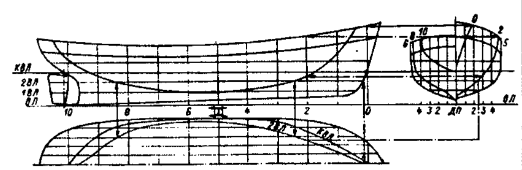
Slide № 1



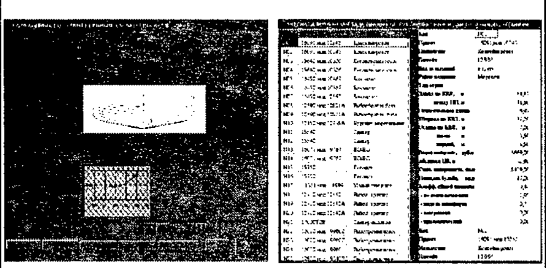
Slide № 2



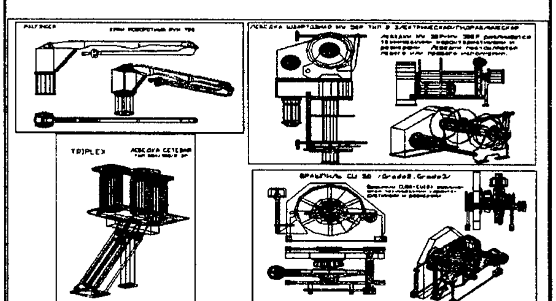
Slide № 3



Slide № 4



Slide № 5



Slide № 6

Generation as opposed to design consists of surface forming according to given main design characteristics of the object. There are several subsystems of generation. In one of them a mathematical description is done on the base of nonalgebraic functions. There is a special procedure ensuring transformation of well known to a designer main hull characteristics to a set of parameters suitable to generation. As in the case of surface design it is not described by analytic methods. Design of external lines can be done with use of prototype which is selected with special information searching system (slide No 5). After selection of the prototype an optimization task on transformation prototype external lines to required ones is solved. The optimization criterion is quantity of object movement resistance within given restrictions.

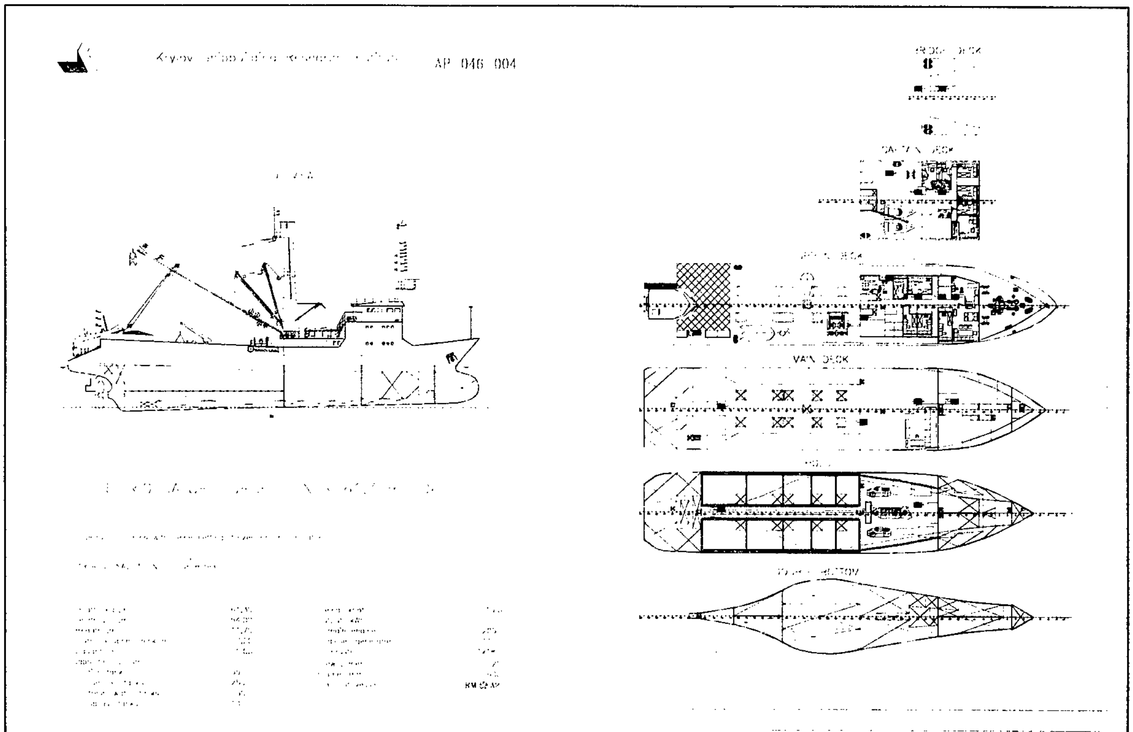
Three-dimensional Modelling. Three-dimensional modelling supposes building of three-dimensional object model. There are three types of such models: wire frame model, surface model and solid model.

During shipbuilding object design usually all three mentioned types are used. Equipment solid models in particular deck equipment are used in IS as during design as well as during making advertisment booklets (slide No 7), video and so on.

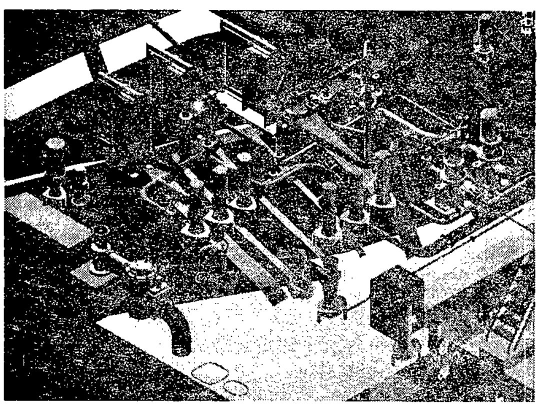
A fragment of solid model of machinery compartment is shown on slide 8. It is clear that during its working out solid models of separate mechanisms saved in special data base (DB) are used. More detailly such DB will be described later. Before implementation of solid modelling expensive models were manufactured. With their use checking of acceptability of taken decisions partly for presence of passages and maintenance zones, noncrossing of pipe lines traces and cables and so on for some rooms with a big quantity of equipment and communications have been made necessary dimensions have been taken from the model to drawings. Use of electronic solid models allows to reduce this expensive process requiring much time.

Scheme of General Arrangement. Surface of external lines is a boundary of an object and external medium. A scheme of general arrangement gives division of internal volume by crossing planes. As a result in IS main watertight compartments and the most large premises shown in drawings are marked out. The drawings are considered to be necessary documents presented in IS (slide No 7). Division of internal volume is done with the help of geometrical modelling system. Marked out premises are divided into two groups. Rooms for which norms and recommendations on volumes, squares and other characteristics exist correspond to the first group. The second group is represented by rooms marking out of which shall precede definition of their characteristics, choosing of main equipment placed there and studying of arrangement. The solid models are done for such rooms. Such models are analogy to shown on slide No 8 but with less degree of detailing. An information searching system used as an auxiliary procedure for forming rooms is developed. Data bases of this system contain norms, rules, recommendations corresponding to forming premises hypertext processing of information taken into the data base is realized. Fragments of software including in ensuring Internet are used.

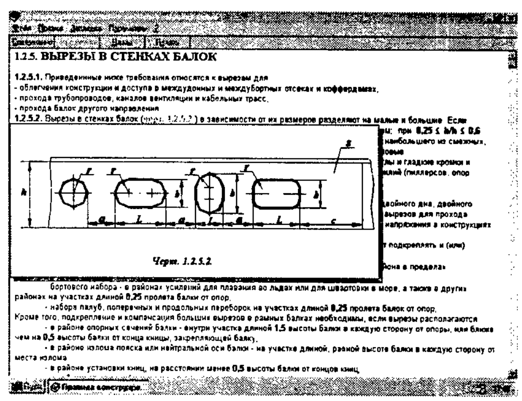
Hull. Till now external lines surface as been considered as some theoretical casing. After hull form processing "materialization" of this casing begins. It means choosing according definite rules material thickness, framing. Then structural drawings of different hull fragments (overlaps, frames and so on) are developed. Necessary strength calculations are made. Under form processing completed in IS one can understand numerical checking of its compliance to many requirements. In complicated cases a desision to carry out necessary model tests is accepted. Drawings according to which the object model will be made are turned out. Numerical form processing allows before model manufacturing and carrying out expensive tests to exclude in advance not perspective



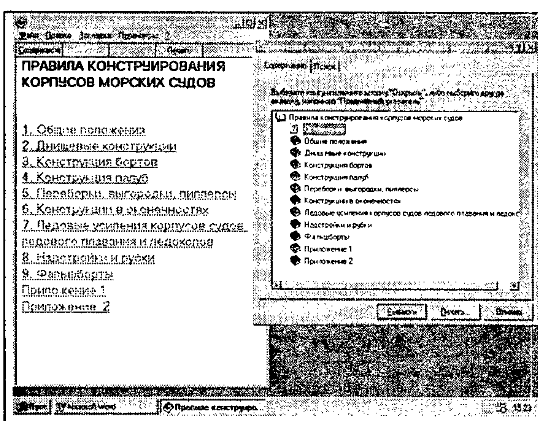
Slide № 7



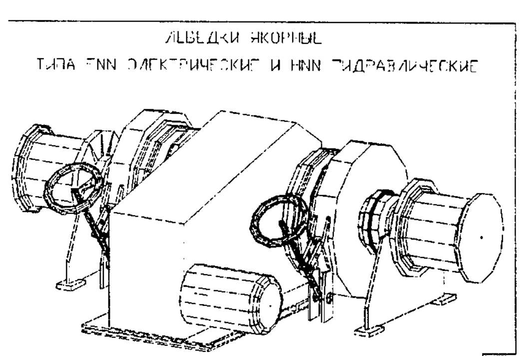
Slide № 8



Slide № 9



Slide № 10



Slide № 11

variants of lines. For this purposes a special complex of software and information searching system on model tests results is used (slide No 5).

"Materialization" of casing is done with the help of special electronic reference book accumulating design and operation practice for slip hulls (slide No 9 and 10). Hypertext processing of information included in the reference book is realized.

Selection of Equipment. Selection of main equipment of different purposes shall be done in initial stages to get data on mass loading, price, needs in energy and mediums and also for ensuring possibility of preliminary scheme development for arrangement in rooms corresponding to the second group. Very different equipment is selected. Among it there are energetics and electroenergetics, ship systems, deck equipment (anchour-mooring line, loading, rescue and so on equipment), technological equipment (fishery, production technological complexes), electronic equipment (navigation and communication, fish searching and so on). Before selection different calculations beginning from definition of needs in energy and media are preceding. Received information and requirements of Classification Societies, recommendations and rules surve as a base for selecting. Detailed data on equipment, ship furniture and so on is included in data bases of special information searching systems (ISS) and information analytical systems (IAS).

In graphic parts of data bases slides containing three projections and axonometry on every mechanism and equipment are represented. Besides slides data bases contain solid electronic models intended for use during prepairing graphic documents (slide No 11). Model export to drawing systems, subsystems of prepairing advertisement booklets video and so on is ensured. Graphic information systems are included in ISS and IAS.

Pipes and cables layout. Fragments of pipe lines traces for the system of ship power plant (PP) are shown on slide No 8. Pipe lines of ship systems and PP systems, cables of electroenergetics systems are traced. To fulfil such works original automatic subsystems are developed. In their limits optimization tasks are also solved.

In this paper these works are not discussed because they are carried out in later design stages.

Calculating Medium. Calculating medium means a totality of softwares of general purpose ensuring effective use calculating resources during solving applied tasks. The work on design is done in the limits of local area network (LAN) including about 15 working stations on PC base, server and peripherals of collective use (matric and laser printers, plotters, digitizers). The network is old variant of EtherNet with transmission velocity 10 Mb per sec.

According to our organisation traditions we are oriented on base programs means of Microsoft. Originally we used 16 bit Windows 3.1 system. Taking into consideration relatively modest possibilities of working stations on PC base main programs taking a large disk volume are saved in server memory in order to make disk memory of working stations free for users.

To ensure data change between users operation system (OS) Novell NetWare demonstrating high safety and enough quick operation is used. Further development is connected with transition to 32 bit system MS Windows NT. Role of network OS in limits of NT belongs to MS Windows NT Server which together with MS BackOffice ensure integrated calculating platform. For work with geometrical information possibilities of subsystems Apirs, Deimos and AutoCad are used. Input of own system (Packet of Geometrical Buildings) PGB is planned.

In the limits of the considered system FoxPro is used as data base management system. Further this function will be transferred to MS SQL server. To ensure the succession orientation to use PC (in initial design stages) with transition to computers with super scalar processors P6 and P7 is preserved. To process geometric information at working stations it is necessary to have 3D accelerator and operative memory not less than 64 Mb.

Besides this transition to 100 MEGC Fast Ethernet with ensurance of exit to Internet is planned.

Control of Design. When a collective operates on design in the limits of LAN principal decision on effective organisation of data use means that information of the building object shall be common resource. Access to it shall be ensured for corresponding categories of staff.

In modern automatic systems data base management technology is used for data control as like as for design control. In our practice design control functions are realized by a monitor. To some extent it dublicates rules of documents turnover and signing adopted in design. The last ensures a statement of electronic document access to use at other working places though the document was worked out at one of the automated working places.

At the beginning of the paper it was mentioned about modelling of LC of the shipbuilding object. Design is properly one of the components of the object LC. The design in a general sense is considered as a totality of files which should be controlled as a whole. Modelling of design LC is realized among others by fastening to every design file or to some part of a file a status of condition or level of completion and also access control keys and further use.

Conclusion. In conclusion we will mark that considered in the paper means and technologies of automatization of initial stages of shipbuilding objects design are constantly improved and developed. On their base native integrated CAD system for shipbuilding is under realisation and partly is already used.

INTELLIGENT CONFIGURATION MANAGEMENT OF MANUFACTURING SYSTEMS

Alexander V.Smirnov
St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences,
39, 14th line, St.Petersburg, 199178, Russia
e-mail: smir@mail.iias.spb.su

Abstract. Reconfigurable Manufacturing Systems are the most flexible and productive systems because they are based on standard/template sub-systems and elements. Configuring requires an analysis of existing enterprise facilities along with its future goals. The results of such an analysis will be required for determining an appropriate reference configuration to build an enterprise. During the configuring, structure and elements of systems must be modified and adapted to new goals and new sub-systems or elements must be added. The organization of these processes, called Configuration Management, constitutes the main focus of this paper. The general objectives of the paper are to determine a methodology, a set of models, and a decision support system architecture based on AI techniques for configuration management.

Keywords. Configuration management, manufacturing system, AI, reengineering.

Introduction. Increasing market demand in conjunction with the modern information technologies requires the development of new manufacturing concepts. In the last decade, a big number of manufacturing concepts have been developed, including: World Class Manufacturing, Agile Manufacturing, Concurrent Engineering, Customer Driver Manufacturing, Business Process Re-Engineering, Learning Organization, etc. Modern manufacturing systems must encompass flexibility, re-configurability, plug-ability, easy maintenance, user friendly, low investment cost.

Development of information technologies (constraint and object-oriented programming, object-oriented DB, multi-agent systems, high speed networks, open distributed platforms, multimedia technology, etc.) introduces a broad range of new approaches to Decision Support System (DSS): from formal methods of operational research to AI tools (task solving virtual team technique instead of personal task solving, constraint satisfaction models instead of optimization models, etc.).

It is common knowledge that the most efficient methods are those applied at preliminary engineering stages, which enables to form the concept of the system "product - process - resource" satisfying the constraints on manufacturing resources, such as investment costs, production area, capacity, and load time. The solution of the problem of rapid, flexible, and accurate manufacturing system configuration will be derived through AI-based tools. Configuring is one of the first and most successful applications of AI-based systems. The use of the latest computer science technologies, in particular Concurrent Engineering, Constraint Satisfaction, and Object-Orientation, facilitates the implementation of these technologies in complex fields like value engineering, design, and manufacturing. It provides for the integration of cost/benefit analysis and manufacturing, merging the process of reengineering with manufacturing implementation and complete cost accountability.

Configuration Management Methodology for Business Processes Reengineering. Product life cycles and pay-off periods have been steadily reduced in recent years, and product complexity has been growing. Thus, it is harder and harder to meet the rapidly changing requirements using any conventional product development process. Nowadays price and quality are no longer the only relevant criteria which guarantee the economical success of a product. The right timing for a product introduction becomes more important, and this usually means that the introduction should take place as early as possible [1].

New market opportunities require constant increase of product quality and decrease of cost in the rapidly changing environment. This trend has become prevailing during 90's. Consequently, the traditional concept of stabilized business and production processes is overshadowed by concerns about flexibility and competitiveness [2, 3]. In order to cope with these new paradigms, companies need to deeply transform their product development structure and the structure of business processes, and this also must be accompanied by on-line transformations, called reengineering.

Cooperation is one of the central requirements for engineering today. Engineering paradigms like participatory design, concurrent engineering, and total quality management all focus on teamwork [4]. Participatory design supports cooperation between users and system designers. Concurrent engineering especially focuses on cooperation between design and production. Total quality management requires cooperation between all departments of an enterprise.

The area of business reengineering with application of Computer Supported Co-operative Work (CSCW) techniques such as groupware systems or the area of distributed production in temporary consortia will also be presented [5, 6]. Multi-agent system technology can be considered as the basis for component integration in cooperative reengineering process [7].

Products and components are often traditionally designed without considering constraints imposed by a manufacturing system. Most of the design-for-manufacturing literature does not consider constraints related to manufacturing systems, but rather deals with manufacturing processes. Based on concurrent engineering concept a product should be designed for a manufacturing systems, as much as the manufacturing systems should be designed for the product. The ability of a manufacturing system to be reconfigured as the production mix changes is closely related to the type of manufacturing resource selected and the arrangement of equipment on a manufacturing floor. In order to design a manufacturing system that can be reconfigured to meet the changing production demand, one has to understand the relationship between the structure of products and the structure of a manufacturing system [8, 9].

Reengineering helps to keep legacy applications in business by transforming their current architecture to a new, more maintainable one. The connections of the configuration engineering and re-engineering tools is shown in Fig. 1. The reengineering process usually requires introduction of additional knowledge at several levels of abstraction [10].

In engineering, configuration management (CM) is one of the main concerns for improving system productivity and reliability. These methodology is applied to various domains. CM is the discipline of identifying the configuration of a system at discrete points in time for the purpose of systematically controlled changes to the configuration and maintaining the integrity of the configuration throughout the system life cycle. The main elements contents of CM process will include the following [11]:

- Configuration Identification - concerned with the definition of different baselines and associated components of the system, any change made to the components of baselines.

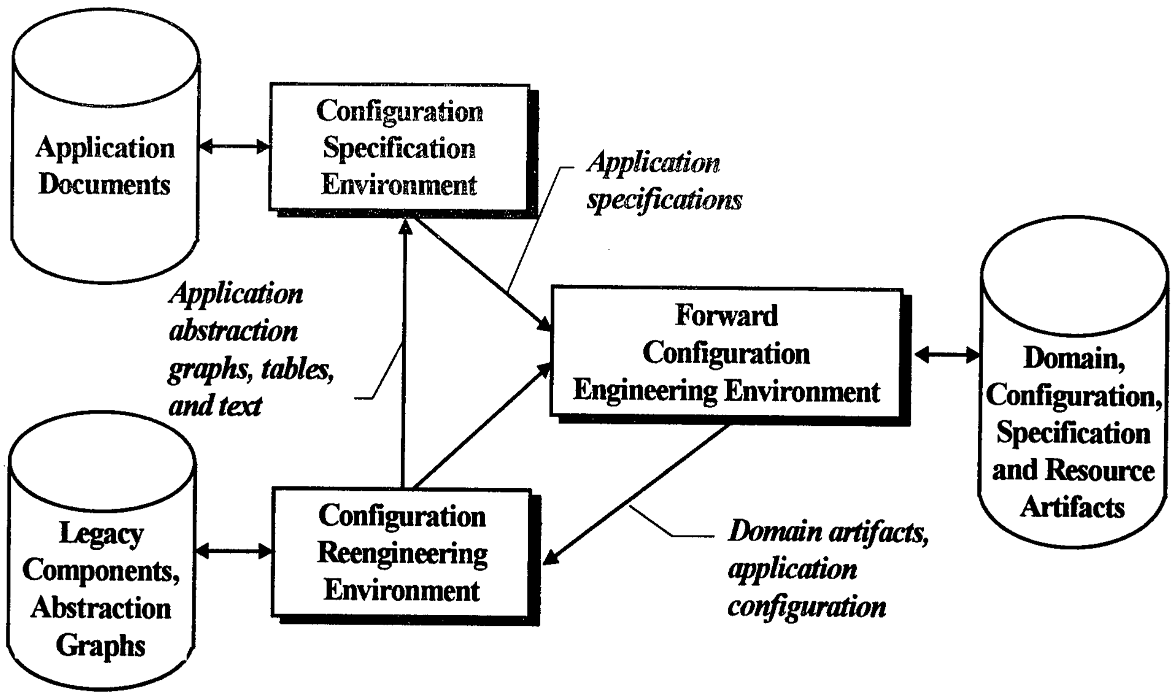


Fig. 1 Integrated Tools for Configuration Reengineering

- Configuration Control - essentially the control procedures for making changes to components and baselines.
- Configuration Status Accounting - provides an administrative history of the evolution of the system.
- Configuration Auditing - determines if the baselines meet their requirements.

These four elements are applied to three major areas of manufacturing system development - product, process, and resource. CM provides the mechanism for integrating these three areas into marketable systems.

Technical systems configuring is an ubiquitous industrial engineering task, and has been a domain for expert system application ever since the longtime paragon R1/XCON [12]. Most configuration expert systems agree in using frames or objects for the representation of factual knowledge about components [13, 14]. For the knowledge representation about the factual knowledge application production rules, semantic networks and constraint-based models are used [14 - 17].

Modern problem solving strategies in area of manufacturing are based on a combination of such methods as: constraint-based problem-solving, constraint-based heuristic search, human decision-making and etc. [18]. It is widely accepted that design activities can be regarded as involving search to satisfy constraints. This model of design has been applied to a wide variety of design domains including conceptual design of manufacturing systems [19 - 21]. One of the configuration principle as structure-oriented model of system is existence of dependence constraints between parts (components and elements) of structure being built.

A constraint network is a collection of constraints interconnected by virtue of sharing variables. The problem of variables selection and testing can be solved by a combination of computers and humans. If done by computer then it is called constraints satisfaction. If the variable values selection and testing are carried out by a combination of human and computer then it is called constraint monitoring. Such a constraint monitoring approach used in a design advise system has the potential for computation efficiency when compared with constraint satisfaction/automatic design approaches.

The statement of a constrain problem may correspond to the following kinds of needs:

- Searching for solution: finding a solution, finding all solutions, finding an optimal solution for given criterion. This approach is called the constraint satisfaction (see Fig. 2).

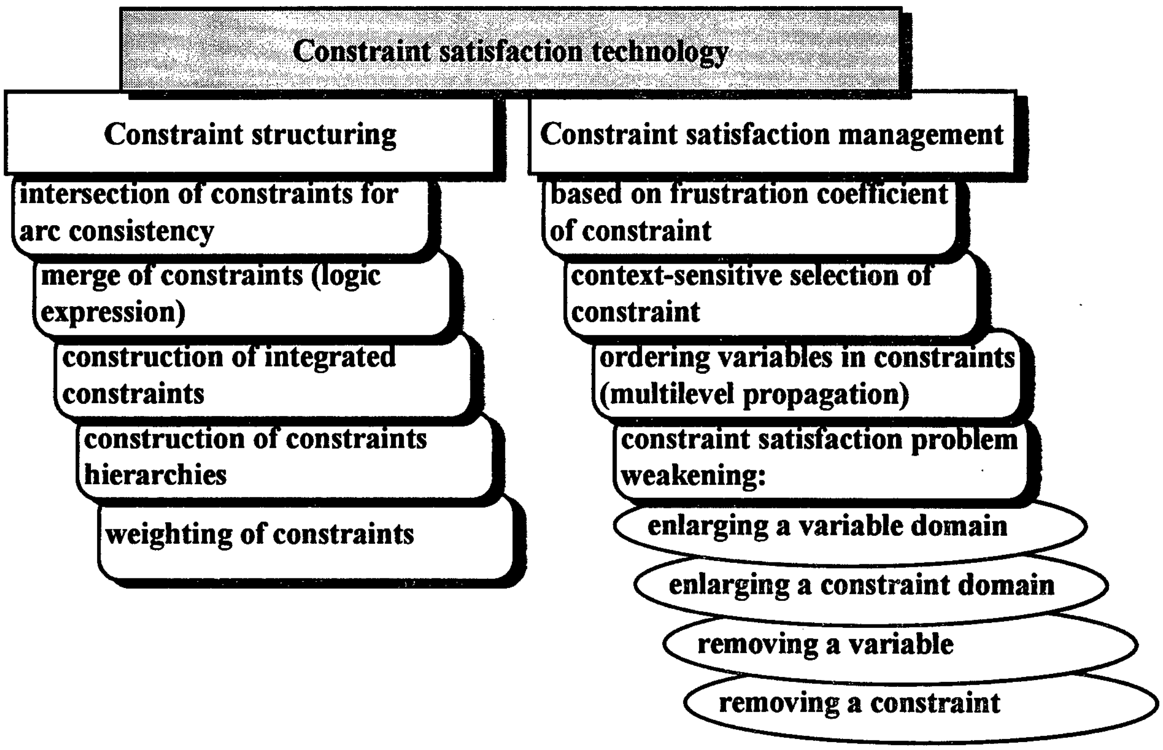


Fig. 2 Effectiveness increase methods of constraint satisfaction technology

- Modifying solutions: taking off from a solution, this method consists of finding a new solution after having modified the value of one or more constrained variables. This approach called the constraint propagation (see Fig. 3).

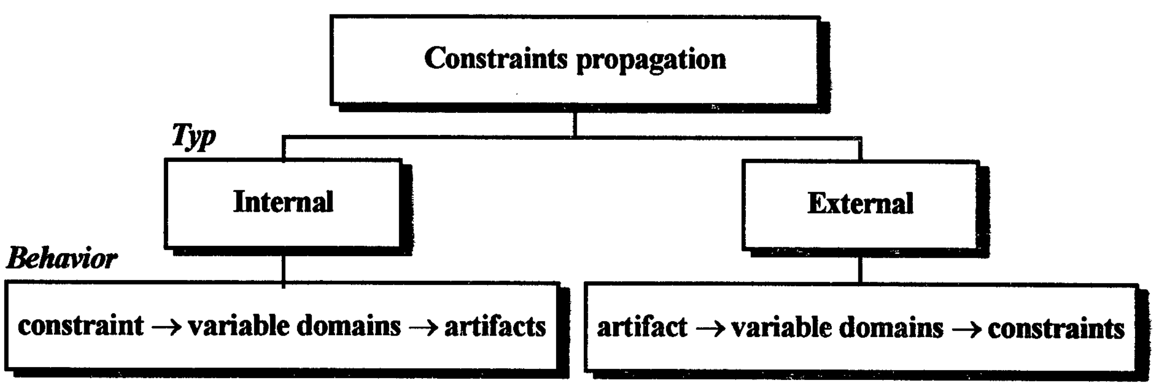


Fig. 3 Artifact-based constraint propagation schemes

The GDSS implements both approaches by means of an engine combining two kinds of features: propagation (that is the repercussion of modifying values or reducing domains) and backtracking (that is automatic management of mechanism for returning to a previous state). This engine enables the search for solutions by enumeration using such methods as generate ant test. However, relying on constraint propagation during the search for solutions makes it practical to detect impossibilities much earlier, and that fact greatly increases the efficiency of the research. This name could be qualified as "propagate and then generate" (or "test-generate").

AI-based Configuration Management Models. The main philosophy of the technology is based on the generic structure of goal-oriented systems. Any of manufacturing systems is composed of one or more configurations (baselines or versions) of the parent system. All components of a given system release or baseline can be organized and managed as a single entity that preserves the relationship between components. For the purpose of describing different versions of components the above entity will be called "virtual" configuration.

Goal-oriented manufacturing systems configuration management approach basically necessitates the realization of following major configuring life-cycle phases: identification of the problem/goal, formulation and specification of the current configuration, re-design and validation of the cost affordable virtual configuration, data collection and analysis, pertinent to the particular problem at hand; determination of parameter values, model-oriented solver selection and constraint-based generation of new appropriate configuration solutions, interpretation and analysis of the solutions, evaluation and modification of the configuration.

The generalized content of configuration management problem includes: the determination of the entity requirements, in other words, the goal, the selection and parametrisation of solutions for entity types from a given catalogue, the creation of new entity types using inheritance mechanisms.

The general scientific objective is to contribute to the building of a consistent body of new technology concept and a set of models via integrating the experience of different experts in (i) evaluation of external enterprises requirements, (ii) mapping external requirements on internal requirements (manufacturing system requirements), (iii) configuring goal-oriented manufacturing systems for decision support system of manufacturing systems configuration management. To realize the above decision making scheme in the considered domain the following methods are proposed [22]:

- contribution-margin ratio analysis and cost-volume-profit analysis for external enterprise requirements specification,
- hierarchies analysis method with conformation of decisions for internal enterprise requirements specification, alternatives evaluation and selection,
- constraint-based structural synthesis and analysis for obtaining goal-oriented manufacturing systems configurations.

Appropriate models will be developed to sample and order external requirements depending upon system goals, and to order and to choose obtained alternative configuration solutions. For this purpose external experts estimate alternative sets of requirements or solutions from different viewpoints reflecting various aspects of re-design. The GDSS is interfaced by appropriate techniques with an intelligent front-end for group expert decision support for functional and quality indices sampling, weighing and evaluation with subsequent acceptance of conforming decisions. This procedure is based on the enhanced hierarchy analysis method, used to obtain conforming expert decisions.

Some of the most important configuration re-design problems are: (i) the determination of the system structure, including its entities and their relations, (ii) decision making on how to realize the entities, (iii) checking decision of entities to guarantee their integration within the framework of the total project. They all are based on constraint satisfaction/monitoring procedure.

Problem domain is specified in terms of template component types to compose manufacturing system under configuring. Each component type is defined by attributes, arithmetic and functional constraints on these attributes and relations with other types. Object-oriented programming paradigm is used for knowledge representation. Component catalogue databases are dedicated to each component type (see Fig. 4). They represent template solutions for possible configurations. The object is represented by a constraint network. It is important that each entity has two sets of constraints. First is the set of internal constraints, second - the set of external constraints.

Co-operative Decision Support System. The general practical objective of this paper is to define tools for configuration management of manufacturing systems in a multi-agent environment that would allow distributed application domain knowledge description on dynamic constraint networks and co-operative decision making by means of local, delegated and shared task concepts.

A rapid manufacturing system configuration management GDSS is based on dynamic constraint networks model for knowledge representation, object-oriented programming technique, and co-operative decision making. The described environment provides users with facilities for requirements specification indexes sampling and weighing, and multi-level configuration synthesis. Different configuring problems are treated as separate tasks with embedded constraint satisfaction and consistency support facilities. A local DSS is an Intelligent Editor for re-design, where the editor constantly verifies that design rules are scrupulously observed by the user and, on the basis of its expertise, editor suggests reasonable solution for difficult re-design problems.

The implementation of the basic principle for the configuration manufacturing - its collaborative nature - is based on procedure distribution between different users (or different aspects) concurrently in the common knowledge space. In this case it is more natural to represent the configuration management knowledge as a set of interactive autonomous agents (see Fig. 5).

The research prototype of software environment for manufacturing system configuration management developed is based on template-based methodology. It originates by integrating the concepts of knowledge representation by hierarchical constraint model, object-oriented programming and cooperative decision making support (see Fig. 6). The prototype can be used for the solution of the following problems: control of adaptability to streamlined configuration; modeling of configuration assembly process; cooperative selection of the most efficient variant of configuration solution.

Conclusion. Manufacturing System Configuration Management is a real innovation technology in the domain of Agile Manufacturing. Using AI-based configuration management techniques enables fewer people to make fast and better-quality decisions on manufacturing configurations based on standard template solutions under rigid constraints with reduced variance. Imple

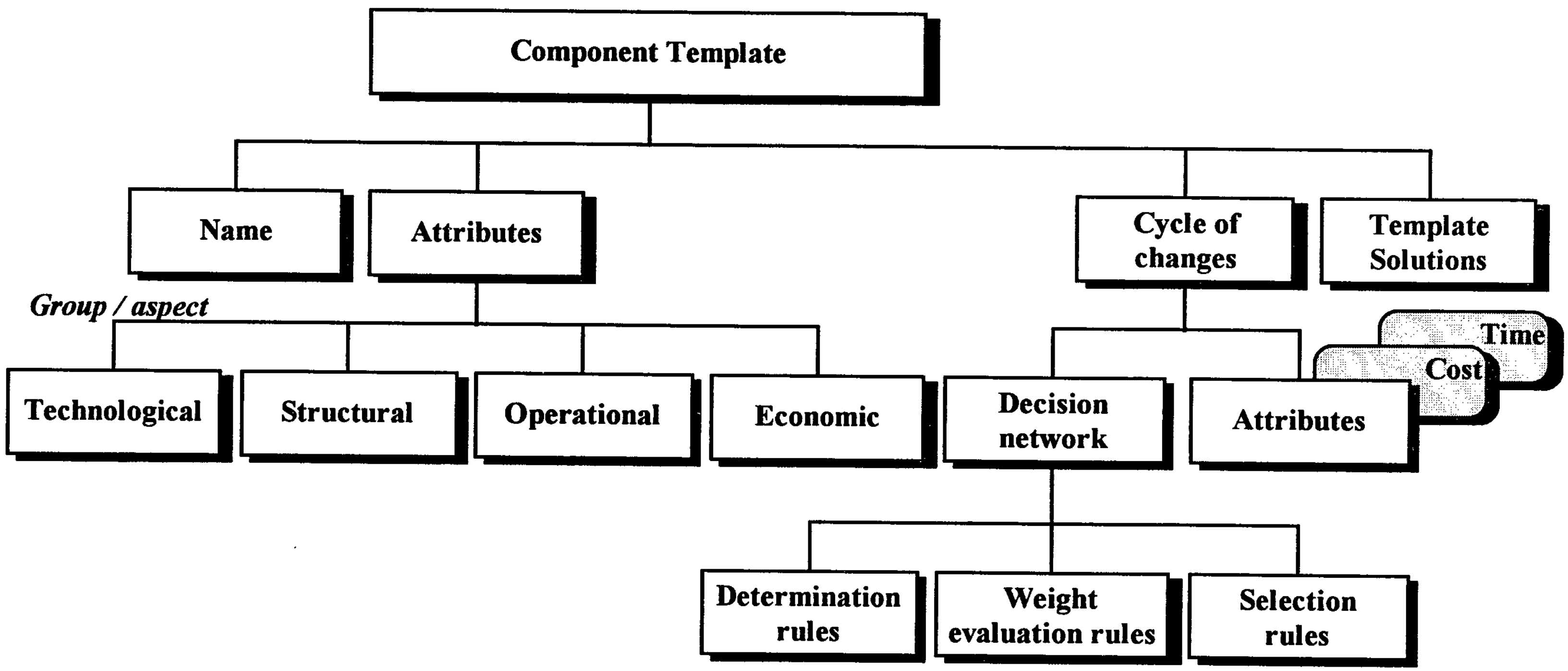


Fig. 4 Structure of "Component Template" Entity

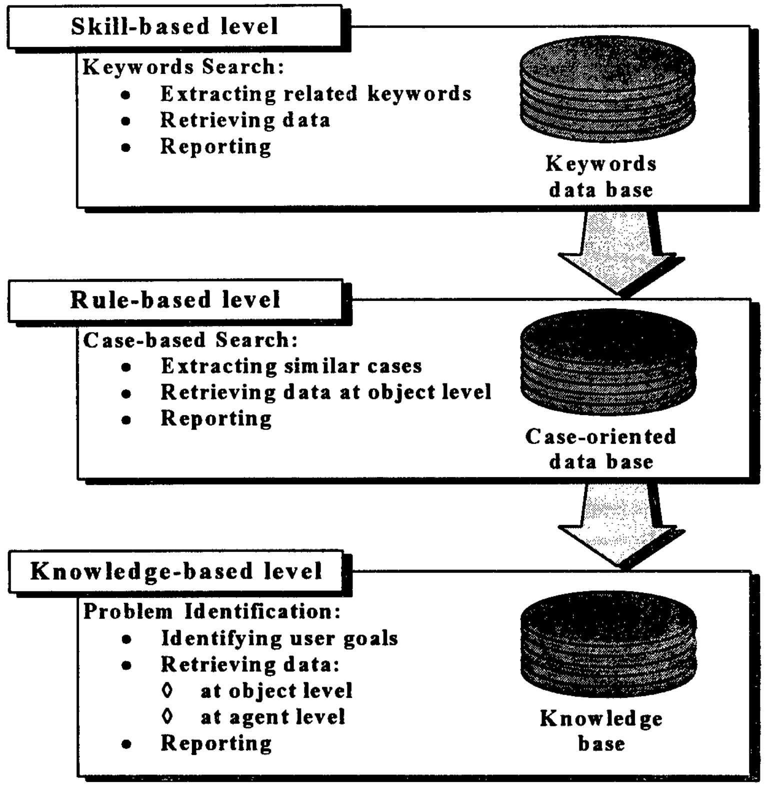


Fig. 5 Multi-level Structure of Knowledge Base and Reasoning Engine for Engineering

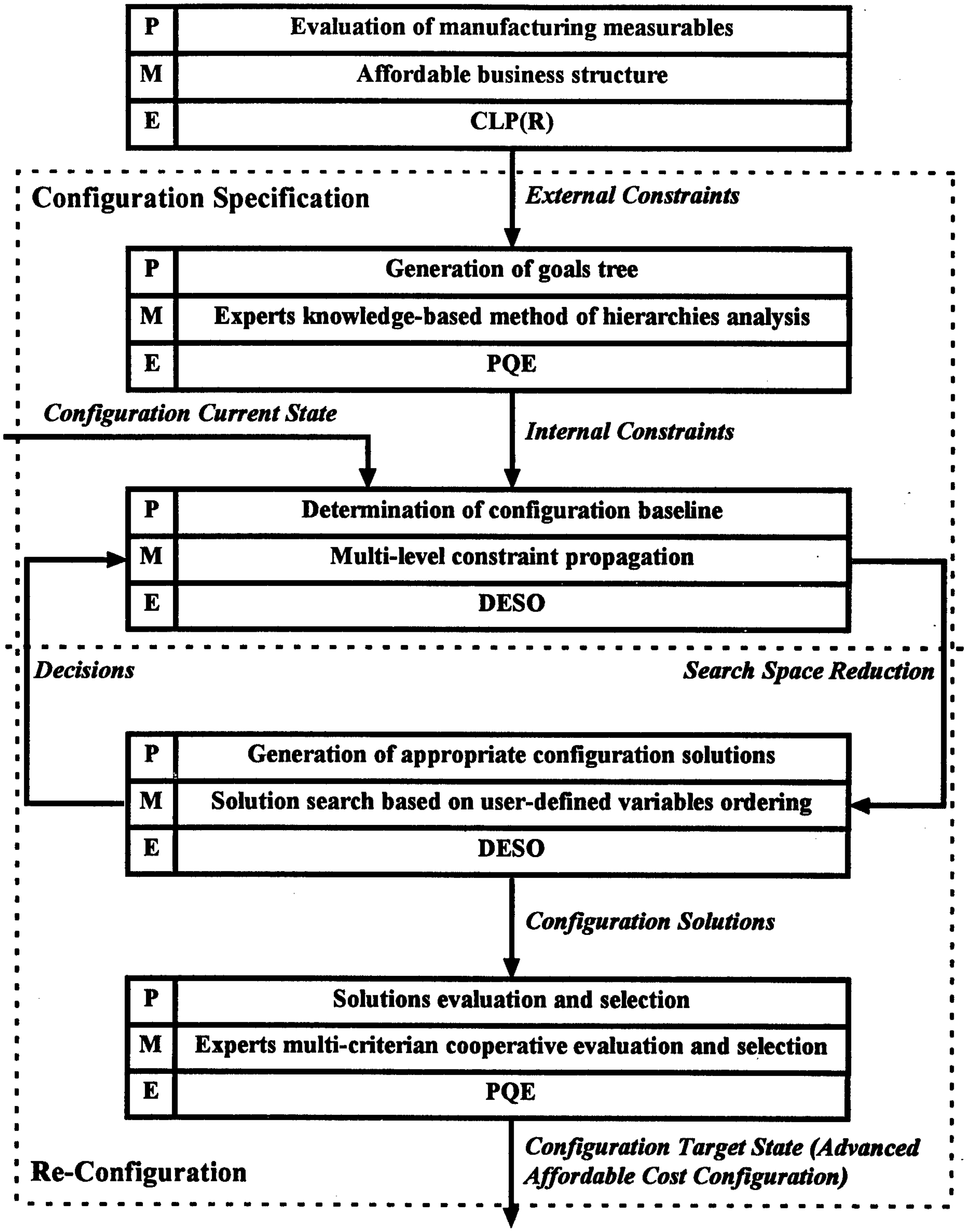
mentation of this technology will result in increased quality, reduced cost, reduced errors, decreased personnel required, better configuration solutions.

Acknowledgments. This research has been partially supported by the State Committee of the Russian Federation for Science and Technologies (grants 236.132 and 05.04.1233) in 1996-1997.

References

1. Eversheim, W. et.al. (1995) Simultaneous Engineering, Erfahrungen aus der Industrie fuer die Industrie, Springer-Verlag.
2. Hammer, M., J.Champy (1993) Re-engineering the Corporation: A Manifesto for Business Revolution. HarperCollins, New-York, 215.
3. Lepikson, H.A. (1995) Core Competence for Flexibility in Product Design and Manufacturing: One Approach for Long Term Competitiveness. Proceedings of the II International Conference on Concurrent Engineering: Research and Applications. DC Area, Concurrent Technology Corporation, Washington, 541 - 548.
4. Jacobs, S., O.Herrmanns (1995) Cooperative Design: Requirements on Network Technology and Document Architecture. In J.Barbere and J.Kiers (eds.), Proceedings of the 6th Joint European Networking Conference, Tel Aviv, Israel, 1995. England, TERENA, 223-1 - 223-8.
5. Hirsch, B. (1995) Information System Concept for the Management of Distributed Production. In Computers in Industry, Elsevier Science B.V., 26, 229 - 241.
6. Eversheim, W., T.Heuser (1994) Process-oriented Order Processing - a New Method for Business Process Re-engineering. In J.Knudsen et al. (eds.), Sharing CIM Solutions, IOS Press, 132 - 141.

7. Intelligent Agents - Theories, Architectures, and Languages, M.Wooldridge and R.Nicholas Jennings (eds.) (1995). In Lecture Notes in Artificial Intelligence, Springer-Verlag, 890, 403 p.
8. Kusiak, A. (1993) Concurrent Engineering: Models and Solution Approach. In I.Mezgar and P.Bertok (eds.), Knowledge Based Systems, Elsevier Science Publishers B.V., 27 - 49.
9. Fernihough, A., G.Owen, A.Mileham, S.Culley (1995) The Development of a Technique for the Evaluation of Business Strategies. Proceedings of the International Conference on Concurrent Engineering: A Global Perspective, McLean, Virginia, 521 - 528.
10. Gall, H., R.Kloesch, R.Mittermeir (1996) Application Patterns in Re-Engineering: Identifying and Using Reusable Concepts. Proceedings of the International Conference Information Processing and Management of Uncertainty in Knowledge Based Systems, Granada, Spain, 3, 1099 - 1106.
11. Bersoff, E. (1984) Elements of Software Configuration Management, IEEE Transactions on Software Engineering, 10, 1, 79 - 87.
12. McDermott, J. (1982) R1: A Rule-based Configurer of Computer Systems. In Artificial Intelligence, 19, 39 - 88.
13. Borgida, A. et al. (1985) Knowledge Representation as the Basis for Requirements Specification, In Computer, 18, 4, 82 - 91.
14. Yoshikawa, H. (1987) General Design Theory and Artificial Intelligence. In T.Bernold (ed.), Artificial Intelligence in Manufacturing, Amsterdam, Elsivier Science Publishers B.V., 35 - 61.
15. Searls, D., L.Norton (1990) Logic-based Configuration with a Semantic Network. In Logic Programming, 8, 53 - 73.
16. Leskin, A.A., A.V.Smirnov (1991) A Technological Knowledge Model in FMS Design System. In J.L.Alty and L.I.Mikulich (eds.), Industrial Applications of Artificial Intelligence, Amsterdam, North-Holland, 378 - 381.
17. Clarke, B.R. (1989) Prokern-XPS: a Mixed Architecture Expert System for Industrial Automation System Configuration. In G.Rzevski (ed.), Proceedings of the 4th International Conference on Applications of AI in Engineering, Cambridge, UK, 1989. Berlin: Springer-Verlag, 49 - 76.
18. Hirsch, B.E. et. al. (1994) Decentralized and Collaborative Production Management in Distributed Manufacturing Environments. In J.Knudsen et. al. (eds.), Sharing CIM Solutions, IOS Press, 43 - 52.
19. Smirnov, A.V. (1994) Conceptual Design for Manufacture in Concurrent Engineering. Proceedings of the Conference "Concurrent Engineering: Research and Applications", Pittsburgh, Pennsylvania, 461 - 466.
20. Sheremetov, L.B., A.V.Smirnov, P.A.Turbin (1995) Constraint-Based Expert System for the Design of Structured Objects. Proceedings of the International AMSE Conference on System Analysis, Control & Design, Methodologies & Examples SYS'95, Brno, Czech Republic, 2, 64 - 71.
21. Smirnov, A., I.Rakhmanova, L.Sheremetov, P.Turbin (1996) GDSS for Re-engineering of Production Systems. Proceedings of the International Conference "Information Processing and Management of Uncertainty in Knowledge Based Systems", Granada, Spain, 313 - 318.
22. Smirnov, A. (1996) Environment for Manufacturing Systems Configuration Management. In M.Jamshidi, F.Pin, F.Pierrot (eds.), Robotic and Manufacturing Systems, Proceedings of the World Automation Congress (WAC'96), Montpellier, France, 3, 729 - 734.
23. Jaffar, J., S.Michaylov, P.J.Stuckey, R.H.C.Yap (1992) The CLP(R) Language and System. In ACM Transactions on Programming Languages and Systems, 14, 3, 339 - 395.



P	- Problem
M	- Method
E	- Environment
CLP(R)	- Constraint Logic Programming Language [23]
DESO	- Design Environment of Structured Objects (Borland C++ 4.5)
PQE	- Project Quality Evaluation (FoxPro 2.5 for Windows)

Fig. 6 Basic scenario of configuration reengineering technology

METHOD OF TECHNIC-ECONOMICAL ESTIMATION OF THE WORK: CREATION AND EXPLOITATION OF AUTOMATED INFORMATION CONTROL SYSTEM.

N.G.Lipatova, V.V.Makrusev
Russian Customs Academy, Scientific Reseach Center
Komsomolsky 4, Lybertsi-city, Moscow reg., 140009 Russia
tel: (7)-559-0033 fax: 5590336

Abstract. *Method is destined for choice and determination of importance of indexes of technical-economical estimation of AUTOMATED INFORMATION CONTROL SYSTEM (AICS) on stages of plan and introduction, exploitation and development (modernize) of system.*

Keywords. Informatisation, control, information control system, creation, exploitation, approach, index, expense.

The object of informatisation of Russian Custom's Academy (RCA) is solved to purpose of creation and development of united automated information system (AICS). AICS RCA is a calculating system providing automatisation of basic educational, research and administrative Academy's functions on basis of united informative, technological and technical decision. It is classed as complicated technical system in its complication, polyfunctional, polycoherent elements, hierarchy of communications of subsystem and indexes of functional and so in other characteristic. Elaboration, introduction and exploitation of such system is connected with large material and working expense, which is predetermined importance and topical of technical-economic estimation of the work over creation AIS, especially in conditions of monetary means limit and constant inflation.

At present time methodical and standard-technical documents on determination of value of expense on creation, development and exploitation of complicated technical system is absent, that it considerably embarrassed realization of research on elaboration and estimation of technical proposition, on substantiation and choice versions of unfold and stages of creation AICS RCA.

Creation of methodical apparatus, permitting to realize scientific-substantiate technical-economical analysis of stage of creation of information-calculating system is propose of this method.

Method is destined for choice and determination of importance of indexes of technical-economical estimation AICS RCA on stages of plan and introduction, exploitation and development (modernize) of system.

Complex indexes of total value of expense on creation AICS RCA and annual expense of exploitation of system are proposed as in the capacity of indexes of technical-economical estimation of work in method.

Complex index of total value of expense on creation AICS RCA includes particular indexes on next articles of expense:

expenses on R&D and produced as result scientific-technical product;

capital expenses on technical means of system; expenses on project software of system; expenses on montage and tuning up of equipment; expenses on installation of project software; expenses on documentation and specialist's training.

Complex index of total value of exploitation of system include direct and indirect expenses of exploitation.

Complex method, including method of base expenses, expert method of estimation, method of calculation on structure of assignment and on middle perches price of technical means and software products, and also method of calculation with use of standard documents of standard work is used for determination of cost indexes of technical-economical estimation in method. Use of this method, and also completeness and reliability of introduction of initial facts permits to get sufficiently exact results. Methodical error doesn't work out more 15%.

System technical parameters, including facts on technical and program structure of examined system and technical-economical parameters, including time, financial and resource limit, acting in organizations of elaborates and customers of s stem, and also determining development's level of means of computing machinery and communication and corresponding standard documents are initial facts for realization of technical-economical estimation.

Method represent finished science technical product, keeping theoretic-methodical and applied parts.

In theoretic-methodical part it is leaded : enlistment of indexes of technical-economical estimation, list of initial facts, ways and methods of their determination, algorithm of realization of technical-economical estimation.

In applied part it is leaded : technology of automatically solution of task and examples, illustrating hard-work of method.

Technology of automatically solution of task of technical-economical estimation is oriented to unprofessional user and is realized as applied program "TEE" in programming language TURBO Pascal 7.0, functioning in environment MS-DOS and Windows-95.

Interaction of user with task realize conversational mode. With that information's exchange is happened as on user's initiative, and so on task's initiative. It is foreseen on user's initiative input initial dates and delivery output information (meanings of index of technical-economical estimation) as table, round and columnar graphs. It is given on task's initiative message about errors and about course of processing of information.

Control of reliability of initial dates is provided by user of task.

Use of method permits to decide whole complex of task of project theoretical and practical character, in private, task of structure-functional synthesis and system integration AICS RCA, optimization of index efficacy of operation and use of system, choice of appropriate strategy step by step realization AICS RCA, and also choice and introduction perspective informational and learned technology.

Method can be used in Research and Development Establishment, in universities, in organization, ordered and elaborating means of automation in realization cost estimation of expense for creation and exploitation complicated computer system.

CHAPTER IV:

NONLINEAR SYSTEMS OF AUTOMATIC CONTROL

THIS CHAPTER INCLUDES PAPERS

PRESENTED AT THE CONFERENCE SESSION:

NONLINEAR SYSTEMS OF AUTOMATIC CONTROL

Organized by: ***Prof. Ronald A. Nelepin.***

INVESTIGATON OF DYNAMICS OF A CLASS OF NON-LINEAR NON-AUTONOMOUS CONTROL SYSTEMS WITH THE POINT MAPPING METHOD

Svetlana Alexeeva, St.Petersburg Technical Forest Academy,
Department of Mathematics, Institutskiy per.5,
St.Petersburg, 194018, Russia

Abstract. The closed executive system for operation of manipulators as an example of non-linear non-autonomous control system with periodic external action has been investigated by the point mapping method.

It was shown that for these systems the steadiness of periodic programmed motions depends on location in the unit circle on the complex plane of the roots of characteristic polinomial for the point translational transformation T_τ' defined on the basis of linearized dynamic equations.

The forced periodic motions of a specific model of three-links manipulator have been investigated.

Keywords. Non-linearity, point mapping, fixed point, steadiness, control system.

Introduction. Non-linear non-autonomous systems of differential equations describe a lot of mathematical models of automatic control systems. As an example the closed executive system for manipulator control can be considered. The rigid requirements to operation characteristics of these systems cause great importance of the control system analysis and synthesis problems.

Specifically, there is a problem of working out the given periodic motions which can be organized as forced periodic motions.

Principal part. Let's consider the system of non-linear unsteady equations which describes a closed executive system for operation of a three-link manipulator:

$$\frac{dx_i}{dt} = X_i(x_1, ..., x_m, q_{lp}(t)), \quad i=1,2,...,m, \quad m=6+\sum_{l=1}^{3} n_l, \qquad (1)$$

where n_l is the order of differential equation for the l-th transmission (l=1,2,3), x_i is the i-th component of the system state vector $x=(x_1,...,x_m)^T$, X_i are given nonlinear functions, $q_{lp}(t)$ are given periodic (with the period τ) functions which determine the programmed trajectories in generalized coordinates q_l.

Let's assume that the non-linear functions $X_i(x_1,..., x_m, q_{lp}(t))$ satisfy the existence and uniqueness conditions (for increasing t) for initial data $x_1 = x_1^0, ..., x_m = x_m^0$ in $t=t_0$ and programmed trajectories $q_{lp}(t)$ in generalized coordinates q_l.

One of the methods for exploration of these systems is the strict analytical method of pointwise mapping which enables to bring the problem of investigation of phase space division structure, and hence the problem of dynamic behavior of the non-linear system (1), to the investigation of some point mappings [1,2]. That is, we intend to solve the problem of given programmed motions steadiness on the basis of investigation of fixed point steadiness for a point translational transformation T_τ. For the fixed point of the transformation T_τ to be assymptotic steady in a little it is suffice according to [3] that all the roots of the characteristic polinomial linearized in the vicinity of the fixed point

$$\det|(\partial f_i / \partial x_k)_{M^*} - \delta_{ik} z|^{m}_{k,i=1} = 0 \qquad (2)$$

lie strictly in the unit circle $|z| \leq 1$ on the complex plane z; $f_i = f_i(t+\tau, x_1, \ldots, x_m, t)$ are continuously differentiable functions which define in the phase space with dimension m+1 the pointwise translational transformation T_τ; δ_{ik} is Kronecker delta.

In other words, the investigation of manipulator periodic programmed motions steadiness can be reduced to the study of location of characteristic polinomial roots about the unit circle on the complex plane for pointwise translational transformation T_τ' connected with the linearized dynamic equations for the system.

The method above has been used to investigate the forced periodic motions of a three-link manipulator on the fixed mount.

The manipulator has a progressive kinematic pair, a rotative kinematic pair and a progressive kinematic pair (each of them has one degree of freedom) which are characterized by generalized coordinates q_1, q_2, and q_3, respectively.

The manipulator links with all their units (rotors of operating transmissions and so on) are assumed to be absolute rigid bodies. The mathematical model of dynamics is formulated subject to the assumption that the manipulator links are symmetrical about the coordinate axis of the coordinate system centered in the center of inertia of the links, using the base kinematic relations as the formulae of coordinates, velocities and accelerations transformation from the rectangular coordinate system related to the link to the inertial coordinate system related to the mounting.

The dynamic equation for the manipulator in the form of Lagrange II kind equations are

$$(J_1 n_1^2 + a_1)\ddot{q}_1 + a_5 \ddot{q}_2 + a_6 (\dot{q}_2)^2 = n_1 (M_{1D} + M_{1C})$$

$$(J_2 n_2^2 + a_2)\ddot{q}_2 + a_5 \ddot{q}_1 = n_2 (M_{2D} + M_{2C}) + a_7 (p_2 + 2(p_3 + p_r)) \qquad (3)$$

$$(J_3 n_3^2 + a_3)\ddot{q}_3 = n_3 (M_{3D} + M_{3C})$$

$$a_5 = (l_2/2)(m_2 + 2m_3) \cos q_2 \,;\; a_6 = -(l_2/2)(2m_3 + m_2) \sin q_2 \,;$$

$$a_7 = -(l_2/2) \cos q_2 \,.$$

Here a_i (i=1,2,3) are the functions of the masses m_i, of the lengths l_i and of the moments of inertia J_i ; q_i is generalized coordinate; M_{iD} is the moment of electric motor of the i-th operating transmission; M_{iC} is the moment of resistance force working on the i-th motor shaft; p_2, p_3, p_r are the weights of the 2th and 3rd links and of the load in the grip, respectively; n_i is the reducer transmission ratio of the i-th transmission; J_i is the moment of inertia of the i-th rotor (i=1,2,3).

It is assumed that the dynamic properties of the i-th electric transmission governing the rotation or translation of the i-th link can be written in the form

$$M_{iD} = q_i^* \, d_i - h_i \dot{q}_i, \ (i=1,2,3) \qquad (4)$$

where d_i, h_i are the construction parameters of the i-th electric motor; q_i^* is the control potential at the entry of the i-th motor.

The law of control potentials is supposed to be as follow:

$$q_i^* = K_n^{(i)} (q_i - q_{ip}) \qquad (5)$$

where $K_n^{(i)}$ is the given amplification; q_{ip} is the programmed trajectory in the generalized coordinates q_i as a periodic (with a period τ) function of t.

As the differential equation for the generalized coordinate q_3 is linear and q_3 does not enter in the first two equations of (3) we suppose that the parameters of the local closed subsystem controlling the third link are chosen so that the programmed trajectory are assymptotic steady while control algorithm is (5).

To investigate the dynamic properties of the local closed subsystems for the first two links the differential equations have been reduced to the Cauchy normal form and linearized in the vicinity of programmed trajectories for the generalized coordinates and velocities. The resulting differential equations for deviations in the matrix form are as follow

$$du/dt = A(t)u \qquad (6)$$

A(t) is the matrix [4 x 4] ; $u = (u_1, u_2, u_3, u_4)^T$ is the vector of deviation of the generalized coordinates and velocities from their programmed values. In the new variables u_i (i=1,2,3,4) the zero solution of (6) corresponds to the periodic solution $q_{ip}, \dot{q}_{ip}$.

In order to find the linearized point mapping T_τ' the numerical approach have been used; namely, the program for numerical search of the point mapping T' and the roots of its characteristic polinomial have been created.

On the base of this program the numerical simulation have been carried out and the parameters values that provide the assymptotic steadiness of the given periodic programmed motions have been found.

Conclusion. This method enables to analize the alternations of control actions from the point of working out the given programmed periodic trajectories $q_{ip}(t)$ in the generalized coordinates $q_i(t)$ and can be used for solution of control laws synthesis problems.

References

1. Butenin N.B., Yu.I Neimark ,N.A. Fufal (1976). Introduction in the theory of non-linear oscillations. Nauka, Moscow .
2. Neimark Yu.I.(1972). The point mapping method in the theory of non-lineary oscillations. Nauka, Moscow.
3. Nelepin R.A. (Ed.) (1975). The methods of investigations of non-linear automatic control systems. Nauka, Moscow.

ON THE CIRCLE CRITERION FOR ABSOLUTE STABILITY.

S.V.Badmaeva, N.E.Barabanov
St.Peterburg State Electrotechical University
197376 St.Peterburg, Popov street, 5, Russia

Abstract. This paper is devoted to the problem of improving the circle criterion for absolute stability of control systems containing several time-varying nonlinearities. A new frequency domain criterion for absolute stability improve the circle criterion. It was derived by means of state space extension method. The main new idea concerns the method of construction of integral quadratic constraints on the basis of local quadratic constraints.

Keywords. Absolute stability, frequency domain criterion

Introduction. The problem of improving of the circle criterion has been studied for a long time. Its goal is to extend the domain of parameters, corresponding to absolute stable systems. The criteria must be effective and easy for testing. The new frequency domain criteria satisfy these conditions. Problem statement. We consider the automatic control system

$$\frac{dx}{dt} = Ax + B\xi,\ \xi = \begin{pmatrix} \xi_1 \\ \dots \\ \xi_m \end{pmatrix},\ \sigma = Cx = \begin{pmatrix} \sigma_1 \\ \dots \\ \sigma_m \end{pmatrix} = \begin{pmatrix} c_1^* x \\ \dots \\ c_m^* x \end{pmatrix},$$

$$\xi = \varphi_j(\sigma_j, t)\ j = 1,..,m \tag{1}$$

where A, B, C are constant matrices of dimensions $n \times n$, $n \times m$, $m \times n$ respectively; $x(\cdot)$ is n-vector function, $\xi_j(\cdot)$ and $\sigma_j(\cdot)$ are scalar functions, φ_j are measurable functions, $j = 1,\dots,m$. All quantities in (1) are real. Let $\{\mu_j\}_{j=1}^m$ be a positive numbers. We denote by M the set of functions $\varphi = \{\varphi_j\}_{j=1}^m$ such that

$$\varphi\ (z,t)(z - \mu_j^{-1}\varphi_j(z,t)) \geq 0 \qquad \text{for all } t > 0,\ z \in R. \tag{2}$$

The task is to define the conditions for parameters $A, B, C, \{\mu_j\}_{j=1}^m$ sufficient for the stability of system (1) with any nonlinearity φ of the class M. Without loss of generality we assume that the matrix A is Hurwitz. The transfer matrix function of linear block of system (1) we denote by $W(s) = C^*(A - sI)^{-1}B$. The well-known circle criterion [1] claims that the system (1) is absolute stable if there exists positive-defined diagonal matrix T such that for all $\omega \geq 0$ the following inequality holds

$$\mathrm{Re}\{T(W(i\omega) + M)\} > 0, \tag{3}$$

where $M = diag\{\mu_j\}_{j=1}^{m}$. This criterion may be obtained using the frequency domain theorem [2] for the quadratic form $F(x, \xi) = \xi(\sigma - \mu^{-1}\xi)$.

It occurs that besides the standard quadratic constraint for the class M it is possible to point out a number of new integral quadratic constraints which lead to absolute stability criteria that are stronger than the circle criterion. The essential role in it plays some sort of the state space extension method.

The method for construction of the integral quadratic constraints. Given real number we consider the quadratic form G of 6 arguments:

$$G(u_0, u_1, u_2, w_0, w_1, w_3) = (u_0 - w_2)^2 + (u_2 + w_0)^2 + \beta\left[((u_0 - w_1)^2 + (u_1 + w_2)^2\right]$$

Now we denote by M_T the set of pairs of measurable functions (f_1, f_2) having support within the interval $[0, T]$ and such that $f_1(t), f_2(t) \equiv 0$.

Then for some real number C and for all pairs $(f_1, f_2) \in M_T$ and all positive τ the following inequality holds:

$$\int_0^T G(f_1(t), f_1(t-\tau), f_1(t-2\tau), f_2(t), f_2(t-\tau), f_2(t-2\tau))dt \geq C\int_0^T (f_1(t)^2 + f_2(t)^2)dt.$$

We add new equation to the initial system (1):

$$\frac{dz_{j,k}}{dt} = \alpha_{j,k}^{-1} P_k z_{j,k} + \alpha_{j,k}^{-1} q_k \xi_k,$$

$$\frac{dw_{j,k}}{dt} = \alpha_{j,k}^{-1} P_k w_{j,k} + \alpha_{j,k}^{-1} q_k \tilde{\sigma}_k,$$

$$j = 1, \ldots, p_k, \quad k = 1, \ldots, m,$$

where $n_k \times n_k$ matrices P_k are Hurwitz; q_k, r_k are n_k-vectors; $\alpha_1, \ldots, \alpha_p$ are positive numbers; $\tilde{\sigma}_k = \sigma_k - \mu_k^{-1}\xi_k$.

Let we have a quadratic forms F_k, $k = 1,\dots,m$ depending on 2p+2 arguments:

$$F_k(\gamma_0,\dots,\gamma_p,\delta_0,\dots,\delta_{p_k}) = \delta_0 \sum_{j=1}^{p_k} C_{j,k,1}\gamma_j - \gamma_0 \sum_{j=1}^{p_k} C_{j,k,2}\delta_j,$$

where $C_{j,k,1}, C_{j,k,2}$ are given coefficients. Then [3] for all solutions of the system (1), (4) with initial conditions $z_{j,k}(0) = 0,$ $w_{j,k}(0) = 0, j = 1,\dots,p_k$

the following identity holds

$$\int_0^T F_k(\xi_k(t), r_k^* z_{p_k,k}(t), \tilde{\sigma}_k(t), r_k^* w_{1,k}(t),\dots,$$

$$r_k^* w_{p,k}(t))dt = \int_0^\infty r_k^* e^{P_k\tau} q_k$$

$$\int_0^T F_k(\xi_k(t), \xi_k(t-\alpha_{1,k}\tau),\dots,\xi_k(t-\alpha_{p,k}\tau),$$

$$\tilde{\sigma}_k(t), \tilde{\sigma}_k(t-\alpha_{1,k}\tau), \tilde{\sigma}_k(t-\alpha_{p,k}\tau))dtd\tau.$$

Here we assume to simplify the notation that $\xi_k(t) = 0, \tilde{\sigma}_k(t) = 0$ if $t < 0.$

Now let the numbers C_k be such that we assume that for all $(f_1, f_2) \in M_\tau$, $\tau > 0$ the following inequality holds

$$\int_0^T F_k(f_1(t), f_1(t-\alpha_1\tau),\dots,f_1(t-\alpha_p\tau), f_2(t), f_2(t-\alpha_1\tau),\dots,f_2(t-\alpha_p\tau))dt$$

$$\le C_k \int_0^T (f_1(t)^2 + f_2(t)^2)dt.$$

If $\int_0^\infty |r_k^* e^{P_k\tau} q_k| d\tau \le 1$ then the following inequality holds

$$\int_0^T \{(\xi_k(t)^2 + \tilde{\sigma}_k(t)^2) + C_k^{-1} F_k(\xi_k(t), r_k^* z_{1,k}(t),\dots,$$

$$r_k^* z_{p_k,k}(t), \tilde{\sigma}_k(t), r_k^* w_{1,k}(t),\dots, r_k^* w_{p_k,k}(t))\} dt \ge 0 \qquad (5)$$

There is a new integral quadratic constraints which lead to new frequency domain criteria for absolute stability.

New frequency criteria for absolute stability. We denote $\Lambda = diag\{\lambda_k\}, Y(s) = diag\{r_k^*(sI - P_k)^{-1}q_k\}, A = diag\{a_k\}, T = diag\{\tau_k\}, M = diag\{\mu^{-1}\}$.
For any positive a we denote by z the positive root of the equation

$$z^3(a^2 - 1)z - a = 0$$

This number may be presented using well-known formula

$$z = \left(\left(\frac{a^2}{4} + \frac{(a^2-1)^3}{27}\right)^{1/2} + \frac{a}{2}\right)^{1/3} + \left(\left(\frac{a^2}{4} + \frac{(a^2-1)^3}{27}\right)^{1/2} - \frac{a}{2}\right)^{1/3}$$

Then for any positive a we define the values $S(a)$:

$$S(a) = \frac{3}{2} \text{ if } 0 < a \le 1;$$

$$S(a) = \frac{a^2 + az + z^2}{a + z} \qquad \text{if } 1 < a \le 1 + \frac{1}{\sqrt{2}};$$

$$S(a) = \frac{a+1}{\sqrt{2}} \qquad \text{if } 1 + \frac{1}{\sqrt{2}} < a \le 1 + \sqrt{2};$$

$$S(a) = a \qquad \text{if } 1 + \sqrt{2} < a.$$

Theorem. Let there exist positive number a, diagonal matrices $T > 0, \Lambda \ge 0, A > 0$, diagonal matrix function $Y(i\omega) = diag\{\tilde{y}_k(i\omega)\}_{k=1}^{m}$ where $\tilde{y}_k(i\omega)$ is Fourier transformation of the function $y_k \in L_1(-\infty, \infty)$, such that $\|y_k\| \le 1$ and for all $\omega \ge 0$ the following inequality holds

$$\operatorname{Re} T(W(i\omega) + M) - (\Lambda A - (W(i\omega) + M)^* \Lambda A^{-1}(W(i\omega) + M) + \\ + \frac{2}{S(a)} \operatorname{Im}((a \operatorname{Im}(\Lambda Y(i\omega) + \Lambda Y(2i\omega))(W(i\omega) + M)) > 0. \quad (6)$$

Then the system (1) is absolute stable in the class M

Example. Consider the following system with transfer matrix function

$$W(s) = \begin{pmatrix} \frac{1}{s+1} & \frac{1}{s+1} \\ \frac{0.5}{s+1} & \frac{1}{s^2+\gamma s+1} \end{pmatrix}$$

and nonlinearities satisfying sector conditions

$$0 \le \varphi_1(z,t)z \le z^2, 0 \le \varphi_2(z,t)z \le \mu z^2, \text{ for all } \mathbf{z,t}.$$

If $\gamma = 0.5$ then the circle criterion is fulfilled only for $\mu = 1.165$. A new criterion for absolute stability is fulfilled if $\mu = 1.594$.

Conclusion. The method of constructing of integral quadratic constraints having local quadratic constraints was used to develop new frequency domain criteria for absolute stability of systems with several time-varing nonlinearities. New criterion improves the circle criterion. The number of parameters to be chosen in new criterion depends linearly on the number of nonlinearities.

References.

1. Narendra K.S., Taylor J.H. (1973) Frequency domain criteria for absolute stability. N.Y., Academic Press.,.
2. Yakubovich V.A. (1973) A frequency domain theorem in theory of control. //Sibirski matematicheski zhurnal, N 2, pp.482-520.
3. Barabanov N.E. (1989) A new criterion for absolute stability and instability of control systems with time-varying nonlinearities // Differential equations, T.25, N 4, pp.555-563.

DESIGNING OF OSCILLATORY MOVEMENT OF A EXECUTIVE DEVICE REGULATOR WITH THE HELP OF ROTATED LJAPUNOW's FUNCTION

Valery Nikolaevih Bagdouyev
Electrotechnical State University
Saint-Petersburg
Department of Automatic and Computer
Prof Popov street, 5
Saint-Petersburg

In the article the problem of designing oscillatory movement of an executive device regulator (EDR) with given frequency and amplitude is considered. The oscillatory movement enables us to decide the problem of stabilization of nonlinear controlled objects (NCO) with the partial information and constantly acting deviations. At the designing of oscillatory movement (EDR) the idea of V.V.Nemytsky of rotated Ljapunow's function is used [1]. Use of this idea enables us to decide the problem of autooscillation of movements of an executive device regulator.

Problems are:

1. To design the rotated Ljapunow's function.
2. To define conditions of occurrence of oscillatory movement of a executive device regulator.
3. To find necessary and sufficient conditions of structural stability in relation to constantly acting perfurbation autooscillation of movements EDR with given amplitude and frequency.

We shall consider oscillatory movement as change of angular coordinate and we shall require unlimited recurrence of this angular coordinate.

Let some phase trajectory $g = f(p,t)$ is given. We shall say, that the phase trajectory defines oscillatory movement, if there is an axis, around which fluctuation occurs, i.e. thereis such direct X^-, in relation to which the following constructions are executed. We shall give an example for the three-dimentional space. Let C be an individual circle with a center on X^-, disposed in orthogonal to X^- plane.

Let $g = f(p,t)$ be some point of a trajectory. We shall design it on a plane of a circle C. Its projection is g^*. Let put a radius-vector of point g^* and note on an individual circle a point of crossing of this radius-vector with a circle C. Let it be point r.

We shall consider the display $T = r$. Changing t, we can write:

$$T(f(p,t)) = r(t).$$

Let now $t \to +\infty$. We shall assume, that $f(p,t)$ is an oscillatory mode, if a degree of display T on a circle C is aimed to $+\infty$ at $t \to \infty$, i.e. $r(t)$ will bypass a circle innumerable set of times at $t \to +\infty$. Let's represent an oscillatory movement analytically. Let a trajectory $g = f(p,t)$ and axis X^- are given. Let's make a construction marked higher for a point p and let corner between a radius-vector of point p^* and axis Z be φ; this corner is referred to as a phase of fluctuations relative to X^-. We shall design now the point g^* on the axis Y and we shall designate a projection through Y^*. If now to designate distance of point g^* from the beginning of coordinates through r^* and corner between a radius-vector of point g^* by a direction φ through α we'll receive $y^*(t) = r^*(t)\sin(\alpha(t)+\varphi)$.

At last, to present the formula in the usual way we shall write

$$y^*(t) = r^*(t)\sin\left(\frac{\alpha\ (t)}{t}\ t + \varphi_\theta\right)$$

when $0 < t < +\infty$ is amplitude of oscillation, and $\omega\ (t) = \dfrac{\alpha\ (t)}{t}$ is variable frequency. Thus it is expediently to consider $\sup\omega\ (t)$ as top right frequency of fluctuation ($0 \le t < +\infty$; $\inf\omega(t)$ as bottom right frequency of fluctuation and at negative current of time, it is possible to speak about left-hand top frequency and bottom left-hand frequency. All these sizes depend on choice of an axis X^-. Therefore the nonlinear fluctuation should be considered with respect to particular axis. Concerning one axis the mode can be oscillatory, and concerning the other --- non-oscillatory. Instead of consideration of absolute amplitude $r(t)$ it is more convenient to consider $\lim\sup r(t) = r_0$. If r_0 is the finite quantity and the number is not equal to zero, we have not faded autooscillation. We shall pass to analytical criteria of existence of oscillatory movements of a executive device regulator.

Let all points of a condition of balance NCO and the closed trajectories of an executive device regulator lie on the characteristic set I, M (individual sphere) [2]. The characteristic set I represents an intersection of successive spheres consistently endosed in each other. Let there be two closed integral surfaces ΔL^+ and ΔL^- taking place through a beginning of coordinates of saddle point [3]. Any decision, beginning on ΔL^+ is aimed to the beginning of coordinates at $t \to +\infty$. Any decision, lying on ΔL^- is aimed to the beginning of coordinates at $t \to -\infty$. Each decision, lying outside this integrated surface is aimed to infinity: as at $t \to +\infty$ and at $t \to -\infty$. In this case ΔL^+ represents steady separatrice of saddle, ΔL^- unstable separatrice of saddle. The set I consists of steady stationary points, saddles and unstable separatrice of these saddles. Any point of set I is either periodic or lies on unstable separatrice of some point being a saddle. Let these integrated surfaces fill in sphere surface I. This connection between integrated surfaces ΔL^+ and ΔL^- in given in an equation

$$V = \varphi = \frac{2}{\pi} arctg \frac{\Delta L^+}{\Delta L^-}, \tag{1}$$

ΔL^+ and ΔL^- are continuously differentiable in individual sphere and presented in a kind of the square-law forms, equal to deviations from the integrated variety M. ΔL^- is a positively determined function, ΔL^+ is an alternating function.

The surfaces ΔL^+ and ΔL^- are crossed transversally in sphere, that follows from the formula (1). ΔL^- represents a equation of spheres with a center located in saddle special point, surfaces ΔL^+ and ΔL^- having a common point only in saddle special point. And at the same time this surface ΔL^+ changes sign continuously at continuous change of φ. Moreover φ

varies from 0 to 2π, ΔL^{+} describes character of approximation of trajectories lying inside the characteristic set I to integrated variety M. Function $\varphi = V$ is alternating.

Where $C = \frac{\Delta L^{+}}{\Delta L^{-}}$ is limits of change, $-\infty < C < +\infty$; $-1 < \varphi < 1$. We receive significances of Ljapunow's function in 4 quadrants: 1 - quadrant $\varphi = 0 - 1$, $C = 0 - (+\infty)$; 2 quadrant - $\varphi = 1 - 0$, $C = +\infty - 0$; 3 quadrant $\varphi = 0 - (-1)$; $C = 0 - (-\infty)$; 4 quadrant $\varphi = (-1) - 0$, $C = (-\infty) - 0$. Function $\varphi = V$ is refered to as rotated Ljapunow's function. Thus:

1. The family $\varphi = V$ - consists of continuous surfaces and through each point of a vicinity of saddle point one and only one surface of family is passed.

2. The complete derivative function V by virtue of a system is continuous and does not come to a zero in a vicinity of saddle special point.

We shall represent equations indiguation of movement of an executive device regulator as a nonlinear similar parahyperbolic differential equation [3], i.e. possessing not more than four types of points: a point of rest in the beginning of coordinates, positively parabolic point, negatively parabolic point, negative hyperbolic points.

Let the similar differential equation of a kind describe indignation movement EDR relativeto saddle special point

$$\frac{dv_i}{dt} = f_i(v_i, \ldots, v_k) \quad (i = 1, \ldots, k) \tag{2}$$

where $f_i(v_i, \ldots, v_k)$ - similar function.

We shall find comple derivative of Ljapunow's function from the system

$$\frac{dV}{dt} = \frac{2}{\pi}\frac{d}{dt} arctg \frac{\Delta L^{+}}{\Delta L^{-}} = \frac{2}{\pi} \frac{\sum_{i=1}^{k} \left(\frac{\partial(\Delta L^{+})}{\partial V_i} \Delta L^{-} - \frac{\partial(\Delta L^{-})}{\partial V_i} \Delta L^{+} \right) f_i}{(\Delta L^{+})^2 + (\Delta L^{-})^2} > \rho > 0. \tag{3}$$

Expression (3) means, that either the trajectory in a final time comes to saddle point, or it comes from saddle vicinity, or remaining inside the ball crosses a surface ΔL^{-} infinite number of times rotating around the saddle point. It will be an oscillatory mode on M.

Also (3) satisfies a condition of the theorem of Krassowsky [4] i.e. condition of transversality (structural stability) [5].

We shall formulate the theorem of Krassowsky.

Theorem. For the purpose that in some limited area I containing not more than one saddle special point, there is no trajectory (different from saddle of the special point) enclosed in this area it is necessary and reasonable, to have alternating Ljapunow's function, with complete derivative (3) to be positive everywhere.

If the conditions of this theorem are executed, in I there can be only hyperbolic and parabolic trajectories.

The last statement can be proved from the following identity

$$V(t) = V_0 + \int_{t_0}^{t} \frac{dV}{dt} dt$$

and, hence, if along the decision

$$\frac{dV}{dt} > p > 0,$$

then $V(t) = V_0 + \rho(t - t_0)$ and $V(t)$ - beyond all bounds grows along the decision, i.e. innumerable set of times the angular parameter φ accepts all significance from 0 to 2π. The surface of level $V = I$ of rotated Ljapunow's function for the characteristic set I is equivalent to a sphere, therefore it crosses sphere on spheres and the availability of rotated Ljapunow's function provides continuous display of these spheres in themselves at shift along a trajectory and, hence, gives availability of the periodic decision (autooscillation EDR). Thus the surfaces $V = I$ have common saddle point and have no other common points, and one of family surfaces passes through each point saddle vicinity.

We shall present significance of derivative Ljapunow'c function (3) in I in a kind $\dot{V} = \Gamma(V_1,...,V_k)V^2$, where Γ satisfies the following condition

$$|\Gamma| > \alpha^2 > 0.$$

Then on a surface of level $V = I_0$, $I_0 \neq 0$ we have got

$$\left.\left(\frac{dV}{dt}\right)\right|_{I_0} \geq \alpha^2 I_0^{\,2}$$

and, therefore, for enough small periods the trajectory, leaving from points of this surface of level, will at various t be on different surfaces of level $V = I$ and the trajectory croses transversally a surface of a level and it can not return. Thus it provides the structural stability of a condition of balance. As I is a continuous and monotonous function t, the transition from $I < I_0$ to $I > I_0$ at the continuous change t can be executed only with the help of transition through the surface $I = I_0$ and it means, that $V = I_0$ is a local section. The surfaces appropriate to $\varphi = 2\pi$ coincide with a surface $\varphi = 0$ and it corresponds to surface $\Delta L^+ = 0$.

We shall consider significance of $\frac{dV}{dt}$ on a surface $\Delta L^+ = V = 0$. On this surface $\frac{dV}{dt}$ has

$$\frac{dV}{dt} = \frac{\sum_{i=1}^{k} \frac{\partial \Delta L^+}{\partial V_i} f_i}{\Delta L^-} = B(V_1,...,V_k),$$

If $B \geq \alpha^2 > 0$, the local section according to the same reasons will be the surface $V = 0$. We shall consider any point inside I, lying on the surface $V = 0$ at $t = 0$, as on a condition

$$\left.\frac{dV}{dt}\right|_{V=0} > \alpha^2 > 0$$

the phase point at $t > 0$ can not remain on the surface $V = 0$ and going on a surface to appropriate large significances I, through a final time leaves on the integrated variety M, i.e. trajectory of movement has hyperbolic character.

Thus for any I, not equal to 0 $\frac{dV}{dt} = \frac{dI}{dt}$. To confirm all the above-stated, we shall formulate the theorem of Nemytskiy.

Theorem. If through all boundary points of the spherical closed area, containing no spacial points, all ω-limiting trajectories enter inwards the sphere, and if relatively I there is rotated Ljapunow's function $V = I$, for which $\Gamma > \alpha^2 > 0$, $B > \alpha^2 > 0$ for the points I, then I contains an oscillatory mode.

For the proof of the theorem, first of all we shall notice, that in conditions of theorem

$$\frac{dV}{dt} > \rho$$

in all closed area I. If there are some intervals of change I, in specificaly $-1 \leq I \leq 1$ statement is obvious, as $\frac{dV}{dt} \geq \alpha^2 I^2$. On the surface $V = I_0$ $(\Delta L^+ = 0)$; on a condition of the theorem

$$\left.\frac{dV}{dt}\right|_{V=0} > \alpha^2 > 0.$$

We shall show: that in enough small vicinity $\Delta L^+ = V = 0$; $\frac{dV}{dt}$ will be more than $\frac{\alpha^2}{2}$. This statement follows from the fact that in up to the small vicinity $V = 0$ and I has no points of a surface $\Delta L^- = 0$. Surface $\Delta L^+ = 0$ is an equation of unstable integrated variety. $\Delta L^- = 0$ is an equation of steady variety. These surfaces can have a common point in saddle point, and therefore $\frac{dV}{dt}$ is continuous in the vicinity $V = 0$.

So, there exists such number ρ, that $\frac{dV}{dt} > \rho^2$ everywhere in I. This inequality shows, that along the trajectories $I(t)$ there is an unequivocal strict monotonous function t on this section of change t, while the trajectory will leave from the surface $V = 0$ at $t = 0$, and will not return again on this surface. Hence t is an unequivocal function I. Using this property, there will be frequency of autooscilation and its dependence on parameters of an equation of an executive device regulator.

So, along the trajectory $\frac{dV}{dt} = \frac{dI}{dt}$, w, considering along the trajectory t as function I, we can write the identity

$$dt(I) = \frac{dI}{\frac{dV(I)}{dt}}.$$

We shall show, that in conditions of the theorem the trajectory left from the surface $V = 0$ in the moment $t = 0$, through a final time will reach a surface $\Delta L^{-} = 0$. We shall find this period from the equality

$$T = \int_0^1 dt = \int_0^1 \frac{dI}{\frac{dV}{dt}}.$$

We have

$$T = \int_0^{\delta} \frac{dI}{\frac{dV}{dt}} + \int_{\delta}^{1} \frac{dI}{\frac{dV}{dt}},$$

where δ - any number $0 < \delta < 1$, as at any I

$$\frac{dV}{dt} > \rho^2 > 0, \quad \int_0^{\delta} \frac{dI}{\frac{dV}{dt}} \le \frac{\delta}{\rho^2}.$$

We shall evaluate second integral

$$\int_{\delta}^{1} \frac{dI}{\frac{dV}{dt}} = \int_{\delta}^{1} \frac{dI}{\Gamma I^2} < \frac{1}{\alpha^2} \int_{\delta}^{1} \frac{dI}{I^2} = \frac{1}{\alpha^2}\left[-\frac{1}{I}\right]_{\frac{1}{\delta}} = \frac{1}{\alpha^2}\left[-1 + \frac{1}{\delta}\right].$$

So

$$T \le \frac{\sigma}{\rho^2} + \frac{1}{\alpha^2}\left[-1 + \frac{1}{\delta}\right].$$

The same valuation takes place for intervals of change (-1;0). We shall recollect, that for the parameter of rotated Ljapunow's function the angular coordinate φ, is used and it becomes clear, that the phase point will return on an initial surface, if φ passes from 0 to 2π, i.e. C changes from 0 up to $+\infty$, from $+\infty$ to 0, from 0 to $-\infty$, from $-\infty$ to 0.

As the indicated valuation is suitable for each of these intervals of change t, through a time

$$T_0 \le 4\left[\frac{\delta}{\rho^2} + \frac{1}{\alpha^2}\left(-1 + \frac{1}{\delta}\right)\right].$$

If we put $\delta = 0{,}5$, $\rho = \alpha$ (from that follows $V = 1$), then

$$T_0 \le \frac{6}{\rho^2} \tag{4}$$

the phase point will return on an initial surface. The frequency of autooscillation can be defined under the following formula:

$$\omega = \frac{2\pi}{T_0}.$$

Accordingly the frequency with the account (4) is equal $\omega \approx \rho^2$. Propose the frequence of autooscillation, it is possible by changing a parity between ρ, α, δ to find parameters EDR. The amplitude of autooscillation is equal to a diameter of characteristic set and equals 2.

Thus, any trajectory beginning in I on the surface $V = 0$ will be oscillatory, or asymptotely is aimed to a oscillatory mode. Hence the osxillatory mode exists inside I.

References

1. Nemytskiy V.V. A method of rotated Ljapunow's functions for searching of oscillatory modes. IS GIVEN TO USSR, 1954, v.17, N 1, p.33-36.
2. Pliss V.A. Integrated sets of periodic systems of differential equations. M., 1977, p.304.
3. Reyzin L.Z. Local equivalence of differential equations. Riga: Zinatne 1971, p.235.
4. Krassowsry N.N. About the reference (manipulation) of the theorem of A.M.Ljapunow and N.G.Chetaew about instability for stationary systems of differential equations. The applied mathematics and mechanics. Izv.AN USSA, v.18(5), 1954, p.513-531.
5. Andronow A.A., Pontrjagin L.S. Rough systems. IS GIVEN TO USSR, 1937, v.14(5), p.246-249.

INFORMATION TECHNOLOGIES FOR MODELLING PHYSICAL PROCESSES AND PHENOMENA USING DIFFERENTIAL EQUATIONS WITH HIGHER-ORDER DERIVATIVES

R.Bun, H.Gudz, A.Semykina
Institute of Information Infrastructure
P.O.box 5446, Lviv, 290031, Ukraine

Abstract. In many science and technology areas the mathematical models of the investigated objects can be represented as systems of non-linear differential equations containing not only first-order derivatives of the unknowns, but also higher-order derivatives. Numerical methods for solving a system of non-linear differential equations with higher-order derivatives that does not require the equations be redused to a system of first-order equations are considered. The proposed methods are implicit many-step methods which makes possible to solve systems of equations that are not reduced to normal form, as well as stiff systems of equations. The highest of the derivatives can be different for each of the variables in equations solved by these methods. One can also solve algebraic-differential equations with higest-order derivatives. For these methods there is no characteristic initialization procedure, which is inherent in the majority of multistep methods and which consist in the fact of so-called "starting points", which can be obtained by another, most frequantly one-step, method, are necessary for the method to work.

Keywords. Information technology, mathematical moddeling, high-order differential equation, physical process, numerical method.

Introduction. In many areas of science and technology the mathematical models of physical processes and phenomena can be represented as systems of non-linear diffential equations containing not only first-order derivatives of the unknowns, but also higher-order derivatives [1]. To solve such equation, one can use the well-developed numerical methods of the Cauchy problem, but the system of higher-order equations must then be reduced to first-order equations, which increases the dimensions of the system. A more serious drawback of this approach is that it cannot be used in cases when the given system of equations is not able to be solved explicitly with respect to the derivatives of the highest order. Existing numerical methods for solving higher-order differential equations are, in principle, applicable to equations of special form. The methods proposed in the present paper, which is based on polynominal interpolation with multiple nodes, is free from the above drawbacks.

Basic notation. Let h_ν be the length of the νth integration step, n the number of equations in the system to be solved, t and t_ν the independent variable and its discrete value that determines the end of the νth integration step, $x_j(t)$ and $x_j^{(r)}$ the jth unknown function and its derivative of order r $(j = 1, 2, \ldots, n)$, $x_{j\nu}^{(r)}$ the exact discrete value of the derivative of order r of the jth unknown at $t = t_\nu$ with $x_{j\nu}^{(0)} = x_j(t_\nu)$, $\tilde{x}_{j\nu}$ and $\tilde{x}_{j\nu}^{(r)}$ the approximate discrete values of the jth unknown and its derivative of order r computed at the step $\nu, w_{j\nu}$ the degree of the algebraic interpolation polynomial that approximates the jth unknown at the step ν, ϑ_ν the dimensionless argument corresponding to the step $\nu, \mu_{j\nu}$ the number of backward steps for the approximation of the jth unknown at the step ν, ν the current integration step number, and $\rho_{j\nu i}$ the multiplicity of the ith integration node for the jth unknown at the step ν.

Formulation of the problem. We are given the mathematical model

$$f_k(\{x_j^{(r)}\}, j = 1,2,\dots,n, r = 0,1,\dots,s_{jk}; t) = 0, \quad k = 1,2,\dots,n \tag{1}$$

(where f_k are continuous functions, and s_{jk} is the order of the leading derivative of the jth unknown in the kth equation) of non-linear differential equations with a unique solution of the form $x_j(t) \in C_{sj}[t_0, t_N]$, with

$$s_j = \max_{k\in(1,2,\dots,n)} s_{jk}$$

Independent initial conditions are known in the form of the initial values of the unknowns $x_{j0} = x_j(t_0)$ and their lower-order derivatives $x_{j0}^{(r)} = x_j^{(r)}(t_0)$ for all $j \in \{1,2,\dots,n\}$, and $r \in \{1,2,\dots,s_j - 1\}$. It is assumed that the dependent initial conditions, i.e. the values $x_j^{(s_j)}$, of the highest-order derivatives, can be computed from Eq. (1) written for $t = t_0$ to within an admissible error using existing numerical methods.

We have to find the solution of (1) in the form of a set of approximate discrete values $\tilde{x}_{j\nu}$ for each of the variables in the set of discrete values of the argument $t_\nu \in [t_0, t_N]$.

The basic principles of the methods. The problem can expressed in algebraic form by replacing the exact discrete values $t = t_\nu$, in the system of equations

$$f_k(\{x_{j\nu}^{(r)}\}, \ j = 1,2,.,.,n, \ r = 0,1,\dots,s_{jk}; t_\nu) = 0, \quad k = 1,2,\dots,n, \tag{2}$$

obtained using the relations (1) for $x_{j\nu}^{(r)}$ by the approximate values $\tilde{x}_{j\nu}^{(r)}$ expressed in the form

$$\tilde{x}_{j\nu}^{(r)} = \sum_{i=1}^{\mu_{j\nu}} \sum_{l=0}^{\rho_{j\nu i}-1} \beta_{j\nu ri}^{(l)} \tilde{x}_{j,\nu-i}^{(l)} + \gamma_{j\nu r}\tilde{x}_{j\nu}. \tag{3}$$

Here $\beta_{j\nu ri}^{(l)}$ and $\gamma_{j\nu r}$ are the coefficients whose values must be determined at each step from the known values h_ν and $h_{\nu-1} = t_\nu - t_{\nu-i}$ for any $i = 1,2,\dots,m_{j\nu}$. For $r = 0$, we have $\beta_{j\nu 0i}^{(l)} \equiv 0$ and $\gamma_{j\nu 0} \equiv 1$. $\tilde{x}_{j,\nu-1}^{(l)}$ are the approximate values of the unknowns and the corresponding derivatives found at the $\mu_{j\nu}(\mu_{j\nu\leq\nu})$ preceding steps (the initial conditions are used for $\nu - i = 0$).

As a result of the above substitution, we obtain a system of non-linear algebraic equations of the form

$$g_{k\nu}(\{\tilde{x}_{j\nu}\}, j = 1,2,\dots,n = 0, \quad k = 1,2,\dots,n). \tag{4}$$

Solving this system by well-known numerical methods, one can determine the values of the variables corresponding to the discrete value $t = t_\nu$.

Before passing to the $(\nu+1)$st integration step for the system (1), it is necessary to compute, using formula (3) and on the basis of the values $\tilde{x}_{j\nu}$ obtained, those values of the derivatives $\tilde{x}_{j\nu}^{(r)}$, corresponding to the current step that will be used at the subsequent steps.

Computation of the coefficients in the expressions for the derivatives. The algorithm for computing the coefficients contained in (3) is based on the approximation of the relation $y_{j\nu}(\vartheta_\nu) = x_j(t)\big|_{t=t_\nu - h_\nu \vartheta_\nu}$, obtained as a result of introducing the dimensionless argument $\vartheta_\nu = (t_\nu - t)/h_\nu$ by the algebraic interpolation polynomial

$$y_{j\nu}(\vartheta_\nu) \approx \varphi_{j\nu}(\vartheta_\nu) = \sum_{r=0}^{w_{j\nu}} d_{j\nu r}\vartheta_\nu^r,$$

constructed on the mesh $\vartheta_{\nu i} = (t_\nu - t_{\nu-i})/h_\nu, i = 1, 2, \ldots, \mu_{j\nu}$, with $\mu_{j\nu} + 1$ interpolation nodes of multiplicity $\rho_{j\nu i}$ (the Hermite interpolation polynominal). We take $\rho_{j\nu i} \in \{1, 2, \ldots, (s_j + 1)\}$, for $i = 1, 2, \ldots, \mu_{j\nu}$, and assume that the integration node $\vartheta_{\nu 0}$ corresponding to $i = 0$ is simple, i.e. $\rho_{j\nu 0} \equiv 1$. Taking the latter into account,

$$w_{j\nu} = \sum_{i=0}^{\mu_{j\nu}} \rho_{j\nu i} - 1 = \sum_{i=1}^{\mu_{j\nu}} \rho_{j\nu i}.$$

The coefficients $d_{j\nu r}$, of the polynomial $\varphi_{j\nu}$ can be found from the condition that the discrete values

$$\varphi_{j\nu}(\vartheta_{\nu i}) = \sum_{r=0}^{w_{j\nu}} d_{j\nu r}\vartheta_{\nu i}^r$$

and

$$\varphi_{j\nu}^{(l)}(\vartheta_{\nu i}) = \sum_{r=l}^{w_{j\nu}} \frac{r! d_{j\nu r}}{(r-l)!}\vartheta_{\nu i}^{r-l}, \quad l = 1, 2, \ldots, (\rho_{j\nu i} - 1), \tag{5}$$

of this polynomial and its derivatives at every interpolation node $\vartheta_{\nu i}, i = 1, 2, \ldots, \mu_{j\nu}$ are equal to the values $\tilde{x}_{j,\nu-i}^{(l)}$, computed at the preceding steps. For $i = 0$, if $\vartheta_{\nu i} = 0$, it is then required that $d_{j\nu 0} = \tilde{x}_{j\nu}$. Taking this condition into account and also the fact that

$$\varphi_{j\nu}^{(l)}(\vartheta_{\nu i}) = (-1)^l h_\nu^l \tilde{x}_{j,\nu-i}^{(l)} \tag{6}$$

at the interpolation nodes, we can express the above requirements concerning the coefficients $d_{j\nu r}$ by means of the system of $w_{j\nu}$ of linear algebraic equations

$$\sum_{r=1}^{w_{j\nu}} \vartheta_{\nu i}^r d_{j\nu r} = \tilde{x}_{j,\nu-i} - \tilde{x}_{j\nu}, \quad i = 1, 2, \ldots, \mu_{j\nu},$$

$$\sum_{r=l}^{w_{j\nu}} \frac{r!\vartheta_{\nu i}^{r-l}}{(r-l)!} d_{j\nu r} = (-1)^l h_\nu^l \tilde{x}_{j,\nu-i}^{(l)}, \quad i = 1, 2, \ldots, \mu_{j\nu}\big|_{i\neq k, \rho_{j\nu k}=l}$$
$$l = 1, 2, \ldots, (\rho_{j\nu i} - 1),$$

with unknowns $d_{j\nu r}$. The latter system written in the matrix form becomes

$$Q_{j\nu} D_{j\nu} = X_{j\nu}^* - X_{j\nu}, \tag{7}$$

where $Q_{j,nu}$ is a square matrix of coefficients, $Q_{j\nu} = \|q_{j\nu vr}\|$, for $\nu, r = 1, 2, \ldots, w_{j\nu}$; and $D_{j\nu}$, $X^*_{j\nu}$ and $X_{j\nu}$ are column matrices of the unknown approximation coefficient and free terms, respectively.

The elements of the matrices $Q_{j\nu}$, $X^*_{j\nu}$ and $X_{j\nu}$ are given by the relations which are represented in [2].

We will express the solution of Eq.(7) in the form

$$D_{j\nu} = P_{j\nu}\left(X^*_{j\nu} - X_{j\nu}\right),$$

where $P_{j\nu} = \|p_{j\nu rv}\| = Q^{-1}_{j\nu}$. On the basis of this matrix formula, taking relations for $Q_{j\nu}$ into account, we arrive at the expressions

$$d_{j\nu r} = \sum_{i=1}^{\mu_{j\nu}} \sum_{l=0}^{\rho_{j\nu i}-1} (-1)^l h^l_\nu p_{j\nu rv} \tilde{x}^{(l)}_{j,\nu-i} - \sum_{i=1}^{\mu_{j\nu}} p_{j\nu ri} \tilde{x}_{j\nu}, \tag{8}$$

where the number of the column of $P_{j\nu}$ that contains the element $P_{j\nu ru}$ of the rth row of $P_{j\nu}$ is given by

$$v = \begin{cases} i, & l = 0, \\ \mu_{j\nu} + \sum_{k=0}^{i-1}(\rho_{j\nu k} - 1) + l, & l = 1, 2, \ldots, (\rho_{j\nu i} - 1). \end{cases}$$

Since, in accordance with equalities (5) and (6) written for the interpolation node $\vartheta_{\nu 0} = 0$, the approximation coefficients and the discrete values of the derivatives at this node are connected by the relation $d_{j\nu r} = (-1)^r h^r_\nu \tilde{x}^{(r)}_{j\nu}/r!$, on the basis of (8) we obtain the expression

$$\tilde{x}^{(r)}_{j\nu} = \sum_{i=1}^{\mu_{j\nu}} \sum_{l=0}^{\rho_{j\nu i}-1} \frac{r! p_{j\nu rv}}{(-1)^{r-l} h^{r-l}_\nu} \tilde{x}^{(l)}_{j,\nu-i} - \sum_{i=1}^{\mu_{j\nu}} \frac{r! p_{j\nu ri}}{(-1)^r h^r_\nu} \tilde{x}_{j\nu}. \tag{9}$$

By comparing (3) and (9), we arrive at the formulae

$$\beta^{(l)}_{j\nu ri} = \frac{r! p_{j\nu rv}}{(-1)^{r-l} h^{r-l}_\nu}, \qquad \gamma_{j\nu r} = -\sum_{i=1}^{\mu_{j\nu}} \frac{r! p_{j\nu ri}}{(-1)^r h^r_\nu}.$$

If the integration step is of constant magnitude, then formula, used to compute $\beta^{(l)}_{j\nu ri}$ and $\gamma_{j\nu r}$, become much simpler.

Computation of the initial approximations. The initial approximations of the unknows are necessary in order to solve the system of non-linear algebraic equations (4) at each integration step. These values can be computed on the basis of the approximate discrete values of the derivatives $\tilde{x}^{(r)}_{j\nu-i}$ for $r = 1, 2, \ldots, w_{j\nu}$ corresponding to the previouse step, which can be found from (3). In the general case the initial approximation $\tilde{x}^{[0]}_{j\nu}$ can be computed from the formula (3)

$$\tilde{x}^{[0]}_{j\nu} = \sum_{r=0}^{w_{j,\nu-1}} \frac{\tilde{x}^{(r)}_{j,\nu-1}}{r!} h^r_\nu, \tag{10}$$

obtained by representing x_j in the neighbourhood of $t = t_{\nu-1}$ by a truncated Taylor series.

A formula with a similar purpose van be obtained on the basis of the interpolation polynomial $\varphi_{j,\nu-1}$ that corresponds to the previous step after substituting $\vartheta_{\nu-1} = -h_\nu$ into this polynomial, which can be expressed by the relation

$$\tilde{x}_{j\nu}^{[0]} = \sum_{r=0}^{w_{j,\nu-1}} d_{j,\nu-1,r} \left(-\frac{h_\nu}{h_{\nu-1}}\right)^r . \tag{11}$$

One can show that formulae (10) and (11) are essentially the same, but the latter is applicable only for $\nu \geq 2$, since $h_{\nu-1}$ is undetermined at the first step.

The comparison of the predicted values of the unknowns computed by means of (10) and (11) with their refined values found by solving the system of equations (4) by numerical methods can be used to control the magnitude of the discretization step. In the case when, for each of the variables, the modulus of the relative difference between the predicted and refined values is much smaller than the given relative error of the corresponding variable, the magnitude of the step can be increased, e.g. two-fold. Correspondingly, one has to discard the results of the νth step and reduce h_ν if the modulus of the above difference exceeds the given relative error for at least one the variables.

The local error of the method, and determination of stability domains. The error due to the fact that the unknown functions are approximated by polynomials of finite degree $w_{j\nu}$ will be called local and is defined by $\Delta x_{j\nu\pi} = x_{j\nu} - \bar{x}_{j\nu}, \quad j = 1, 2, \ldots, n.$ Here $\bar{x}_{j\nu}$ is the solution of the system of equations (4) obtained by using the values of the unknowns that correspond to the exact solution of the original system (1) at the previous steps as the initial data determining the approximating functions.

As it was shown in [3] the local error of computing the unknowns can be presented in the form $\Delta x_{j\nu\pi} = O(h_\nu)^{k_\nu}, \quad j = 1, 2, \ldots, n,$ where

$$k_\nu = \min_{i \in \{1,2,\ldots,n\}, r \in \{1,2,\ldots,s_j\}} (w_{i\nu} - r + 1) = \min_{i \in \{1,2,\ldots,n\}} (w_{i\nu} - s_i) + 1.$$

In the most unfavourable case (when the one-step version of the proposed method with $w_{i\nu} = s_i + 1$ is used) the exponent k_ν is equal to two. At the cost of increasing the degree of the interpolation polynomials φ_{jr}, which can be achieved by increasing number of backward steps, the exponent k_ν can, in principle, be increased by as much as is required.

The discrete problem expressed by (4) is called stable if numbers $h_0 > 0$ and $\delta > 0$ exist such that for any $h < h_0$ and $\varepsilon \in R_n$, such that $\|\varepsilon\| < \delta$, the discrete problem

$$g_{k\nu}(\{x_{j\nu}^*\}, j = 1, 2, \ldots, n) = \varepsilon_k, k = 1, 2, \ldots, n,$$

has only one solution $x_{j\nu}^*$, $j = 1, 2, \ldots, n$, $\nu = 1, 2, \ldots, N$, which differs from the solution $\tilde{x}_{j\nu}$, $j = 1, 2, \ldots, n$, $\nu = 1, 2, \ldots, N$, of (4) by $|x_{j\nu}^* - \tilde{x}_{j\nu}| \leq C\|\varepsilon\|$ by all $j \in \{1, 2, \ldots, n\}$. Here $\|\varepsilon\|$ is the norm of the column matrix ε characterizing a perturbation of the discrete problem and C is a constant. Besides, it is customary to say that there is a qualitative correspondence between the discrete and continuous problems if their phase spaces have the same structure.

The stability of a numarical method is usually investigated in the case when the method is used to solve a system of homogeneous linear differential equations. When there are higher-order derivatives the system can be written in the form $\sum_{r=0}^{s} A^{(r)}X^{(r)} = 0$, where s is the highest order of the derivatives of any unknowns, $A^{(r)}$ are the coefficient matrices, $A^{(r)} = \|a_{kj}^{(r)}\|$, $k, j = 1, 2, \ldots, n$, $r = 0, 1, \ldots, s$, and $X^{(r)}$ is the column matrix formed by the derivatives of oreder r of the unknowns x_j, $j = 1, 2, \ldots, n$.

Connection with other methods. Depending on which of the functions and their derivatives found at the previous steps are used to compute the values of the variables that correspond to the current integratetion step, i.e. depending on the form of the interpolation mesh, various variants of the methods proposed in the present paper are possible, for example, the one-step variant with constant multiplicity of the interpolation nodes, etc.

The one-step variant is characterized by the number of backward steps being equal to one for all the variables. The digree of the interpolation polynomial, $w_{j\nu}$, to which we assign the largest possible value $w_{j\nu} = s_j + 1$ for $\mu_{j\nu} = 1$, is also constant for all steps. For the one-step variant, it can be shown that even if the magnitude of the discretization step changes, the inverse matrix $P_{j\nu}$ to $Q_{j\nu}$ is nor required when computing the coefficients $\beta_{j\nu ri}^{(l)}$ and $\gamma_{j\nu r}$ in (3), which significantly simplifies the corresponding algorithms. In the extreme case when the highest oreder of the derivatives in the mathematical model (1) to be solved is equal to one, and the equations can be written in normal form, the one-step variant of the proposed methods is indentical to the trapezium method, since, using a suitable expression for the derivative, one can obtained the relation

$$\tilde{x}_{j\nu} = \tilde{x}_{j,\nu-1} + (\tilde{x}_{j\nu}^{(1)} + \tilde{x}_{j\nu-1}^{(1)})h/2.$$

A characteristic feature of the variant with constant multiplicity of the interpolation nodes is that the number of backward steps and the multiplicity of each of the interpolation nodes, except for the node corresponding to $t = t_\nu$ are the same for all steps ans all variables, which can be expressed by the conditions $\mu_{j\nu} = \mu$ and $\rho_{j\nu i}$ for all $i = 1, 2 \ldots, \mu$. It can be shown [2] that in the extreme case when $\rho = 1$, i.e. in the case when all the interpolation nodes are simple, if the methods described in the present paper is used to solve mathematical models with first-order derivatives, it is exactly the same as that proposed in [5] which is based on the Lagrange and Newtom interpolation polynomials, and is known as the "BDF method" (the backward differentiation formula method). Consequently, the methods in the present paper is a generalization of the BDF method to the case of differential equations with derivatives of any order, which justifies us in calling it the GBDF methods (the generalized (the backward differentiation formula methods).

Conclusions. The GBDF methods presented here enable the volume of computations when solving higher-order mathematical models to be reduced considerably, since one does not need the equations to be reduced to first-order ones. There are implicit many-step methods which make it possible to solve systems of equations that are not rediced to normal form, as well as stiff systems of equations. The highest order of the derivatives can be different for each of the variables in equations solved by the GBDF methods. One can also solve algebraic-differential equations with highest-order derivatives. For the GBDF methods there is no characteristic initialization procedure, which is inherent in the majority of multistep methods and which consists in the fact that so-called "srating points", which can be obtained by another, most frequently one-step method, are necessary for the method to work. Since the GBDF methods encompasses many distinct variants including the one-step variant, it is always possible to use one of the variants at any of the steps without using other methods.

References

1. Bun, R. (1994) Information technologies for the simulation of complex nonlinear electrical circuits. Pattern Recognation and Image Analysis, 3, 298–305.
2. Bun R., Y. Vasiliev (1992) A numerical method for solving differential equations with derivatives of any order. Computational Mathematics and Mathematical Physics, 3, 317–330.
3. Bun R. (1993) An analysis of GBDF method local error. Measuring and Processing of Information, 9, 32–37.
4. Bun R., A.Semykina (1996) Numerical Methods for solving high-order differential equations using generalized backward differentiation formula. Volynski Mathematical Journal, 3, 26–30.
5. Brighton R., F. Gustavson, G. Hatchel (1972) A new efficient algorithm for solving algebraic systems of diferential equations based on the use of imlicit formulae for numerical differentiation with backward differences. Proc. IEEE, 1, 136–148.

DIFFERENCE EQUATION LINEARIZATION OF THE DISTURBED MOTION OF PULSE-WIDTH SYSTEM WITH A NON-LINEAR CONTINUOUS PART.

Andrew Korshunov
St.Petersburg University
Anatole Korshunov
Popov Naval College of Radio - Electronics
15, Razvodnay street, Petrodvorets,
198903, Russia

Abstract. Linearization of perturbation difference equations of pulse-width systems are obtained. Pulse-duration modulator's output signal contains in non-linear autonomous differential equation as additive. Found equations are necessary to stability analysis of constant pulse-width steady-state.

Keywords. Pulse-duration modulator, non-linear continuos plant, constant pulse-width, steady- state stability, difference equation linearization.

Introduction. Pulse-width system (PWS) containing linear continuous part (CP) is till now usually investigating. In reality CP of PWC is more or less non-linear. Therefore the most of obtained results may be used to PWS only in the case of weak CP non-linearity. Separate exiting DC Motor is for example such CP. But DC Motor in the case of serial existing is example of essentially non-linear CP. PWS analysis is very troubled because of absence of non-linear differential equation's analytical solution. Consequently PWS difference equations cannot be obtained and analytical stability analysis methods using successfully for linear CP are helpless.

Below matrix of linearization perturbation difference equations will be obtained.

Notation and Definitions. Let us take a view of closed loop PWS indicated on figure.

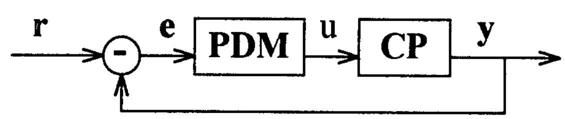


Fig. Pulse-width system.

The pulse-duration modulator (PDM) is described as follows

$$u(t) = U\left[1(t - nT) - 1(t - nT - \tau_n\right], \quad nT < t < (n+1)T \ , \tag{1}$$

$$\tau_n = Tsat\frac{\alpha e_n}{T} \text{ , if } \ e_n = e(nT) = r - y(nT) \geq 0, \quad \tau_n = 0 \text{, if } e_n \leq 0 \ , \tag{2}$$

where U, α, T are it's parameters.

The CP under consideration is assumed to be mathematically described by vector differential equation

$$\dot{x} = F(x) + hu, \tag{3}$$

where $x^T = [x_1, x_2, \ldots, x_m]$, $F^T(x) = [f_1(x), f_2(x), \ldots, f_m(x)]$, $h^T = [h_1, h_2, \ldots, h_m] = const.$

The vector-function's $F(x)$ elements are assumed to be continuous and continuously differentiated. The CP's output equation is written as follows

$$y = c^T x \ , \tag{4}$$

where $c^T = [c_1, c_2, \ldots, c_m] = const.$

At the differential equation (3) points of continuous break $t = nT + \tau_n$ the state vector x is determined according continuity, as initial mean in interval $[nT + \tau_n, (n+1)T]$ to be equal to final mean in interval $[nT, nT + \tau_n]$, that is $x(nT + \tau_n + 0) = x(nT + \tau_n - 0) = x(nT + \tau_n)$.

When $r = const$ it's assumed the PWS has desirable steady-state such that

$$\tau_n = \tau_0 = const, \, 0 < \tau_0 < T. \tag{5}$$

The co-ordinate steady alteration low

$$\bar{x}(\eta), \, 0 < \eta < T, \, \bar{x}(\eta) = \bar{x}(T) = x_\infty \tag{6}$$

may be obtained by mathematical modelling. The point x_∞ of PWS's state variables $\{x_n = x(nT)\}$ is invariant point of the system's difference equations

$$x_{n+1} = F(x_n), \quad F^T(x_n) = \left[\tilde{f}_1(x_n), \tilde{f}_2(x_n), \ldots, \tilde{f}_n(x_n) \right], \tag{7}$$

i.e.

$$x_\infty = F(x_\infty).$$

Steady-state stability is depended on linearization perturbation difference equations matrix eigenvalues.

$$z_{n+1} = W z_n \,, \tag{8}$$

where $z_n = x_n - x_\infty$, $W = \dfrac{dF}{dx_n}(x_\infty)$.

Since it is impossible to obtain vector-function's $F(x_n)$ mathematical expression, matrix W will be found by means of another manner below.

Difference equation linearization. Let us look at perturbed motion vector

$$z(\eta) = [z_1(\eta), z_2(\eta), \ldots, z_n(\eta)]^T = x(\eta) - \bar{x}(\eta), \quad 0 \le \eta \le T \tag{9}$$

and perturbed motion differential equation

$$\dot{z}(\eta) = \Phi(\eta, z(\eta)) + h\Delta u(\eta, z(0)), \tag{10}$$

where

$$\Phi(\eta, z(\eta)) = [\varphi_1(\eta, z(\eta)), \varphi_2(\eta, z(\eta)), \ldots, \varphi_n(\eta, z(\eta))]^T = F\left(\bar{x}(\eta) + z(\eta)\right) - F\left(\bar{x}(\eta)\right), \tag{11}$$

$$\Delta u(\eta, z(0)) = U[1(\eta - \tau_0) - 1(\eta - \tau_0 - \Delta\tau)], \tag{12}$$

$$\Delta\tau = -\alpha c^T z(0), \tag{13}$$

($z^T(0) z(0)^{1/2}$ is small enough)

It is apparent that $\varphi_i(\eta, z), i = \overline{1,m}$ and $\dfrac{d\varphi_i}{dz_j}, i,j = \overline{1,m}$ are continuous to z_j, $j = \overline{1,m}$ and to η in consequence continuity the solution of equation (3) to η and also $\Delta u(\eta, z(0))$ is square-impulse positive, if $\Delta\tau > 0$ and negative if $\Delta\tau < 0$ that is piecewise constant function. The matrix (8) can be evidently obtained as

$$W = \frac{dz(T)}{dz(0)}. \tag{14}$$

As the equation's (10) right side has continuity break in points $\eta = \tau_0$ and $\eta = \tau_0 + \Delta\tau$ the matrix W obtaining from variation differential equation [3] is impossible. Nevertheless the equation's (10) right side piecewise continuity and continuous differentiation to $z_j,\ j = \overline{1,m}$ permit to use solution continuous dependence on initial conditions and continuous differentiation to initial conditions in each of three intervals $[0,\tau_m),(\tau_m,\tau_M)$ and $(\tau_m,T]$ $(\tau_m = \min\{\tau_0,\tau_0+\Delta\tau\},\ \tau_M = \max\{\tau_0,\tau_0+\Delta\tau\})$. In the case of $\Delta z(0) \approx 0$ it is possible to write in the first interval

$$z_i(\eta,z(0)) - z_i(\eta,0) = \left.\frac{dz_i(\eta,z(0))}{z(0)}\right|_{z(0)=0} z(0) + 0_i(\eta,z(0)),\ i = \overline{1,m}\ , \tag{15}$$

where $z_i(\eta,0) = 0$ according (9), $\lim\limits_{z(0)\to 0} 0_i(\eta,z(0)) \Big/ z(0)^T z(0)^{1/2} = 0$

The expression (15) can be written in vector form

$$z(\eta,z(0)) = \left.\frac{dz(\eta,z(0))}{z(0)}\right|_{z(0)=0} z(0) + 0(\eta,z(0)), \tag{16}$$

where

$$H = \frac{dz(\eta,z(0))}{dz(0)} = \left\|\frac{\partial z_i}{\partial z_j(0)}\right\|_1^m \tag{17}$$

is the $m \times m$ matrix, $\lim\limits_{z(0)\to 0} \dfrac{0^T(\eta,z(0))0(\eta,z(0))^T}{z^T(0)z(0)^{1/2}} = 0.$

Since matrix's H elements satisfy variation differential equation [3]

$$\frac{d}{d\eta}\left.\frac{\partial z_i(\eta,z(0))}{\partial z(0)}\right|_{z(0)=0} = \left.\frac{\partial\varphi_i(\eta,z)}{\partial z}\right|_{z=0} \left.\frac{dz(\eta,z(0))}{dz_j(0)}\right|_{z(0)=0},\quad i,j = \overline{1,m}$$

(18) under initial conditions

$$\left.\frac{\partial z_i(0,z(0))}{\partial z_j(0)}\right|_{z(0)=0} = \begin{cases} 0, i \neq j, \\ 1, i = j. \end{cases}$$

(19) and the matrix H satisfies matrix differential equation

$$\frac{d}{d\eta}\left.\frac{dz(\eta,z(0))}{dz(0)}\right|_{z(0)=0} = \left.\frac{d\Phi(\eta,z)}{dz}\right|_{z=0} \left.\frac{dz(\eta,z(0))}{dz(0)}\right|_{z(0)=0} \tag{20}$$

under initial conditions

$$\left.\frac{dz(0,z(0))}{dz(0)}\right|_{z(0)=0} = E, \tag{21}$$

E is the identity matrix.
The matrix's columns satisfy the vector differential equation

$$\frac{d}{d\eta}\frac{dz(\eta,z(0))}{dz_j(0)}\bigg|_{z(0)=0} = \frac{d\Phi(\eta,z)}{dz}\bigg|_{z=0}\frac{dz(\eta,z(0))}{dz_j(0)}\bigg|_{z(0)=0} \tag{22}$$

under initial conditions

$$\frac{dz(\eta,z(0))}{dz_j(0)}\bigg|_{z(0)=0} = E_{*j}, \tag{23}$$

E_{*j} is j-th column of matrix E.

If $\psi(\eta)$ is fundamental matrix of equation (22) and

$$\psi(0)=E \tag{24}$$

from equations (22) and (20) it can be found that

$$\frac{dz(\eta,z(0))}{dz_j(0)}\bigg|_{z(0)=0} = \psi(\eta)E_{*j}, \qquad 0\le\eta\le\tau_m, \tag{25}$$

$$\frac{dz(\eta,z(0))}{dz(0)}\bigg|_{z(0)=0} = \psi(\eta), \qquad 0\le\eta\le\tau_m. \tag{26}$$

With taking into account the last equality the expression (16) can be written as

$$z(\eta,z(0)) = \psi(\eta)z(0)+0(\eta,z(0)), \quad 0\le\eta\le\tau_m. \tag{27}$$

As vector $z(\eta,z(0))$ is continuous to η it's initial mean in interval $[\tau_m,\tau_M]$ is supposed to be equal the final mean in interval $[0,\tau_m]$

$$z(\tau_m,z(0)) = \psi(\tau_m)z(0)+0(\tau_m,z(0)). \tag{28}$$

The vector $z(\eta,z(0))$ in interval $[\tau_m,\tau_M]$ satisfies equation

$$\frac{dz(\eta)}{d\eta} = \Phi(\eta,z(0))z(\eta)+hU sign\Delta\tau. \tag{29}$$

As according to (13) and τ_m and τ_M determination

$$\tau_M-\tau_m = |\Delta\tau| = |\alpha c^T z(0)| \tag{30}$$

the equation (20) solution may be presented as follows

$$z(\tau_M) = z(\tau_m)+\left[\Phi(\tau_m,z(\tau_m))+hU sign\Delta\tau\right]|\Delta\tau|+0'(\tau_m,z(0)) =$$
$$= z(\tau_m)+hU\Delta\tau+0(\tau_M,z(0)) = z(\tau_M+0) ,$$

where $0(\tau_M,z(0)) = 0'(\tau_M,z(0))+\Phi(\tau_m,z(\tau_m))z(\tau_m)|\Delta\tau|$.

In the third interval $[\tau_m,T]$ the perturbed motion vector $z(\eta)$ satisfies the same uniform differential equation as in the first interval $[0,\tau_m]$. Consequently

$$z(T) = \psi(T)\psi^{-1}(\tau_M) = \psi(T)\psi^{-1}(\tau_M)\left(\psi(\tau_m)z(0)+0(\tau_m,z(0))-\alpha Uhc^T z(0)+0(\tau_M,z(0))\right) =$$
$$= \psi(T)\left(\psi^{-1}(\tau_M)\psi(\tau_m)-\alpha Uhc^T\right)z(0)+0'(T,z(0)).$$

Whereas the fundamental matrix's $\psi(\eta)$ continuity and the equality (30) the last equality may be written as follows

$$z(T) = \psi(T)\left[E-\alpha U\psi^{-1}(\tau_0)hc^T\right]z(0)+0(T,z(0)). \tag{31}$$

From the expression (31) it follows that the linearization perturbation equation's matrix W has expression

$$W = \psi(T)\left[E - \alpha U \psi^{-1}(\tau_0)hc\right]. \tag{32}$$

It is to be noted that the last expression is similar matrixes of linear PWS in the case of linear CP [4]. If CP is linear it is to be written

$$F(x) = Ax, \quad \Phi(z) = Az, \quad \frac{d\Phi}{dz} = A$$

and the expression (32) gives well known formula

$$W = H(T)\left[E - \alpha U H(-\tau_0)hc^T\right], \quad H(t) = \exp At.$$

This is the indirect expression's (32) verity argument.
Without correct proof the matrix W may be obtained more simply as it is shown below.
Continuous part perturbation differential equation (10) linearization gives

$$\dot{z} = A(\eta)z + h\Delta u(\eta, z(0)). \tag{33}$$

Pulse - duration modulator perturbation linearization leads to commutation of square impulse (12) into instantaneous impulse

$$\Delta u(\eta, z(0)) = U\Delta\tau\,\delta(\eta - \tau_0), \tag{34}$$

where $\delta(\eta)$ - Dirac delta function, $\Delta\tau = \alpha c^T z(0)$.
Differential equation (33) solution [3] gives with taking into account (34)

$$z(T) = \psi(T)\left[z(0) + \int_0^T \psi^{-1}(\eta)\alpha c^T z(0)\delta(\eta - \tau_0)\right] = \psi(T)\left[E - \alpha U \psi^{-1}(\tau_0)hc^T\right]z(0). \tag{35}$$

Evidently (32) follows from (35).

Example. DC serial exiting motor velocity PWC in the case of non-saturated motor, proportional velocity controller and constant moment load, may be written by equation (1-4), where

$x^T = [x_1, x_2]$, $h^T = [1, 0]$, $c^T = [1, 0]$, x_1 is motor's current, x_2 is motor's velocity, $\varphi_1(x_1, x_2) = -(a + bx_2)x_1$, $\varphi_2(x_1, x_2) = cx_1^2 - d$, a, b, c, d are constant parameters.

If τ_0 and $\bar{x} = \left[\bar{x}_1(\eta), \bar{x}_2(\eta)\right]^T$ are known the expression (11) gives

$$\Phi(z) = \begin{bmatrix} \varphi_1\left(\left(\bar{x}_1 + z_1\right), \left(\bar{x}_2 + z_2\right)\right) - \varphi_1\left(\bar{x}_1, \bar{x}_2\right) \\ \varphi_2\left(\left(\bar{x}_1 + z_1\right), \left(\bar{x}_2 + z_2\right)\right) - \varphi_2\left(\bar{x}_1, \bar{x}_2\right) \end{bmatrix}$$

Vector $\Phi(z)$ differentiation and $z = 0$ insert bring matrix

$$A(\eta) = \left.\frac{d\Phi(\eta, z)}{dz}\right|_{z=0} = \begin{bmatrix} -\left(a + b\bar{x}_2(\eta)\right) & -b\bar{x}_1(\eta) \\ 2c\bar{x}_1(\eta) & 0 \end{bmatrix}.$$

Matrix $\psi(T)$ and $\psi(\tau_0)$ necessary for W determination may be found with numerical integration the following matrix differential equation

$$\frac{d\psi(\eta)}{d\eta} = A(\eta)\psi(\eta), \quad \psi(0) = E.$$

Conclusions. Since matrices obtained by means of mathematically correct method and without correct proof absolutely coincide more simple method matrix W obtaining is proved. Therefore matrix W may be found by means of separate PWM and non-linear CP linearization.

References.

1. Bromberg P.V. (1967). Matrix methods in pulse and relay control theory. Science (ed.), Moscow.
2. Kuntsevitch V.M., U.N. Chehovoi (1970). Non-linear pulse-width and pulse-frequency modulated control systems. Technique (ed.), Kiev.
3. Pontrjagin L.S. (1965). Ordinary differential equations. Science (ed.), Moscow.
4. Korshunov A.I. (1977). Pulse- Width second kind system's steady-state motion stability. Automatics and Telemechanics, G., 6, 31-37.

ANALYSIS AND SYNTHESIS OF THE ROBUST SYSTEM WITH THE ROOT LOCUS METHOD

Tamar Kotrikadze,
Georgian Technical University,
Department of Automatic Systems,
Kostava 77, Tbilisi, Georgia

Abstract. I offer the Root Locus Method for solving the problems of Robust Analysis and Synthesis. The Grapho-analitical Method is used to build the trajectories of the equations' roots that have the real coefficients. That is why it becomes much more convenient to solve this problem at the computer. I got very important results that make it easier the process of finding Roots' Trajectories when the several parameters of the Algebraic Equation are changing. It becomes possible to sort the Kharitonov's Polinoms with the Root Locus by order of Dominating Roots.

Keywords. Robust System, Analysis, Synthesis, Root Locus, Kharitonov's Strong Theorem.

Environmental Impact. For finding the Root Trajectories' Picture of Algebraic Equations real coefficients' Root Locus Properties are used, when a few parameters or coefficients of the equation are changing [1].

It becomes much more difficult to use them if it is necessary to find the Root Trajectories when two or more coefficients , parameters are changing. It is hardly possible to solve this problem at the computer. That is why I decided to process the new method which allows us to find the Root Trajectories with any exactness for any number of the equation's coefficients that are changing in the set intervals. I have formulated and proved four theorems that make easier finding the Root Trajectories.

Theorem 1. The equation

$$S^n + c \cdot S^k + 1 = 0 \qquad (1)$$

where $c \in R$; $n, k \in N$; $n > k$ and S is variable, may have the root with the multiple not more than 2.

$$S = \pm \sqrt[n]{\frac{k}{n-k}} \qquad (2)$$

If n is an even number and

$$S = \sqrt[n]{\frac{k}{n-k}} \qquad (3)$$

If n is an odd number.

The equation will have such root if

$$c = -\frac{n}{k}\sqrt[n]{\left(\frac{k}{n-k}\right)^{n-k}} \tag{4}$$

Theorem 2

$$S^{2n} + c \cdot S^{2(n-m)+1} + 1 = 0 \tag{5}$$

The Trajectories of the Eq. (5) when $c \in R$ is symmetrical to the Roots Complex Plane and to the Imaginary Axis. In Eq. (5) $n \in N$ and $m \in \left\{1,2,\ldots,\left[\frac{2n+1}{2}\right]\right\}$ where $\left[\frac{2n+1}{2}\right]$ is $\frac{2n+1}{2}$ number's integer part.

Theorem 4.

$$S^n + c \cdot S^k + 1 = 0 \tag{7}$$

where $n,k \in N$ and $n > k$, are the initial roots. When $c = 0$, they are located on the unit circle, when $c \to \infty$ from n initial roots $(n-k) \to \infty$.

Beside these theorems for finding the Roots' Trajectories I use their equations and the relations between equations' coefficients and roots.

Let say we have reduced standard equation:

$$t^n + a_1 t^{n-1} + a_2 t^{n-2} + \ldots + a_{n-1} t + 1 = 0 \tag{8}$$

for finding the Roots' Trajectories when $a_i \in \left[\underline{a}_i, \overline{a}_i\right]$ where $0 < \underline{a}_i < \overline{a}_i$ and $i = 1,2,\ldots,n-1$. That is equal to building the following equation's Root Locus for Eq.(9).

$$\left[\begin{array}{ll} t^n + 1 + a_1 \cdot t^{n-1} = 0 & \left| a_1 \in [\underline{a}_1;\overline{a}_1] \right. \\ P_1(t;a_1) + a_2 \cdot t^{n-2} = 0 & \left| a_i \in [\underline{a}_i;\overline{a}_i] \right. \; i = 1,2 \\ P_2(t;a_1;a_2) + a_3 \cdot t^{n-3} = 0 & \left| a_i \in [\underline{a}_i;\overline{a}_i] \right. \; i = 1,2,3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ P\,(t;a_1;\ldots;a_{n-2}) + a_{n-1} \cdot t = 0 & \left| a_i \in [\underline{a}_i;\overline{a}_i] \right. \; i = 1,2,\ldots,(n-1) \end{array}\right. \tag{9}$$

where $P_i(t;a_1;\ldots;a_i) = (((t^n + 1) + a_1 \cdot t^{n-1}) + a_2 \cdot t^{n-2}) + \ldots) + a_i \cdot t^{n-i}$ is n-th rate polinom of t variable and contains i changing parameters. In the set of Eq. (9) the first equation contains the one varied parameter, the second equation two and so on. The last equation from the set in the Eq. (9) contains $(n-1)$ changing parameters. We will have the Root Locus Family for all equations in the Eq. (9) set except the first one.

I use the Identical Transformation of the Polinom with the Complex Argument [2] that gives us the opportunity of finding the Root Locus equations and formulas for calculating the parameters in any point of trajectory. For example in the set of Eq. (9) the second root locus equation is:

$$[P_1(\eta;a_1)-\frac{\chi^2}{2!}\cdot P''_1(\eta;a_1)+\frac{\chi^4}{4!}\cdot P_1^{(4)}(\eta;a_1)-\ldots]\cdot[(n-2)\cdot\eta^{n-3}-$$
$$-\frac{\chi^2}{3!}\cdot\frac{(n-2)!}{(n-5)!}\cdot\eta^{n-5}+\frac{\chi^4}{5!}\cdot\frac{(n-2)!}{(n-7)!}\cdot\eta^{n-7}\ldots]-[P_1'(\eta;a_1)- \tag{10}$$
$$-\frac{\chi^4}{5!}\cdot P_1'''(\eta;a_1)+\frac{\chi^4}{5!}\cdot P_1^{(5)}(\eta;a_1)-\ldots]\cdot[\eta^{n-2}-\frac{\chi^2}{2!}\cdot\frac{(n-2)!}{(n-4)!}\cdot\eta^{n-4}+$$
$$+\frac{\chi^4}{4!}\cdot\frac{(n-2)!}{(n-6)!}\cdot\eta^{n-6}\ldots]$$

where η _ is the real root of Eq. (8).

χ _ is the imaginary root of Eq. (8)

Parameter a_2 in any $(\eta;\chi)$ point can be calculated by the Eq. (11) or Eq. (12):

$$a_2=-\frac{P_1(\eta;a_1)-\frac{\chi^2}{2!}\cdot P_1''(\eta;a_1)+\frac{\chi^4}{4!}\cdot P_1^{(4)}(\eta;a_1)+\ldots}{\eta^{n-2}-\frac{(n-2)!}{(n-4)!2!}\cdot\chi^2\cdot\eta^{n-4}+\frac{(n-2)!}{(n-6)!4!}\cdot\chi^4\cdot\eta^{n-6}\ldots} \tag{11}$$

or

$$a_2=-\frac{P_1'(\eta;a_1)-\frac{\chi^2}{3!}\cdot P_1'''(\eta;a_1)+\frac{\chi^4}{5!}\cdot P_1^{(5)}(\eta;a_1)+\ldots}{(n-2)\eta^{n-3}-\frac{(n-2)!}{(n-5)!3!}\cdot\chi^2\cdot\eta^{n-5}+\frac{(n-2)!}{(n-7)!5!}\cdot\chi^4\cdot\eta^{n-7}} \tag{12}$$

We may find a_2 coefficient using Eq. (12) only in the complex points of the Root Locus.

To solve Eq. (12) system I use the program on the Borland Pascal and the scheme portrait of the Root Locus. This method makes possible to build the Root Locus with any frequency of the points.

For example third rate reduced standard Algebraic Equation

$$t^3+a_1t^2+a_2t+1=0 \tag{13}$$

The Root Locus of this equation are showed at the fig. 1 that is one of the results of the program. I can use this program for building the Root Locus for Eq. (13).

I also use Kharitonov's Strong Theorem [3] for solving the problem of Robust Analysis. For four Kharitonov's Polinoms we find root distribution using the Root Locus. In this case it is not necessary to find their complete picture. The process of finding Kharitonov's Polinoms becomes much more easier. I may also determine the system's performance: transient response, overshoot and oscillation.

I also consider the Regulator's and Correction Unit's synthesis. For solving this problem we have to find area of the characteristics equations roots' plane. This area is determined by using the given values of the System's Dynamics Performance.

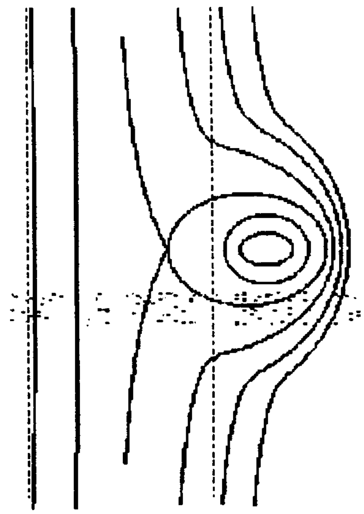


Figure 1

After that we can change the Systems Adjustment Parameters so, that Kharitonov's Roots with the worst dynamics locate in the given area. If this is possible then the problem of Parameters Robust Synthesis has the solution. In the other situation we have to use the Adaptation Principe.

Conclusions. The Root Locus method is one of the modern parts of Automatic Control Theory that is not learned and developed well enough. In this work I have tried to answer a lot of questions; for example, when and where to use the Root Locus Method, is it possible to use this method in any situation or not and so on. I also processed the program that is used in practice in Georgian Technical University.

References:

1. Uderman E. G. Root Locus Method in the Automatic Control Theory (1963). Gosenergoithdat. 50-66.
2. Bendricov G.A., Teodorchic K.F. Analytical Theory for Building the Root Locus. (1959). Avtomatica i Telemechanica. N3. 75-80.
3. Kharitonov V.L. Stableness of the Differential Equations' System's Family. Differential Equations.(1978). 53-57.
4. Bessecersky V.A., Popov E.P. Automatic Control Systems' Theory (1972). Nauka. 18-29.

NUMERICAL SIMULATION OF STOCHASTIC CONTROL SYSTEMS

Oleg Yu. Kulchitski, Dmitri F. Kuznetsov
St. Petersburg State Technical University
Department of Mechanics and Control Processes
29, Polytekhnicheskaja street, St. Petersburg
E-mail: controll@citadel.stu.neva.ru

Abstract. This paper is devoted to the problem of correct numerical simulation of control systems, described by stochastic differential equations. Numerical schemes for these equations being constructed in this paper on the basis of *recurrence* Taylor-Ito expansion, which is an original rezult of this paper. Considering one method for approximation of repeated stochastic integrals, which are used for numerical realization of constructed schemes.

Keywords. stochastic control system, simulation, recurrence Taylor-Ito expantion, repeated stochastic integral, approximation.

1. Mathematical model of stochastic control systems. We shall consider following control systems:

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{a}(\mathbf{x}_t, t) + B\mathbf{u}(t) + \Sigma(\mathbf{x}_t, t)\xi_t, \tag{1}$$

where $\mathbf{x}_t = \mathbf{x}(t, \omega)$–is a solution of equation (1); $\mathbf{a}(\mathbf{x}, t) \in R_n$; $\Sigma(\mathbf{x}, t) \in R_{n\times m}$–are some matrix functions; $\xi_t \in R_m$–is a gaussian process with independent components, zero expectation and finite standard deviation; $B \in R_{n\times l}$; $\mathbf{u}(t) \in R_l$–is a control function.

Stratonovich in [1] considered following modification of system (1):

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{a}^*(\mathbf{x}_t, t) + \Sigma(\mathbf{x}_t, t)\frac{1}{\mu}\xi_{\frac{t}{\mu^2}}; \ \mathbf{a}^*(\mathbf{x}, t) = \mathbf{a}(\mathbf{x}, t) + B\mathbf{u}(t), \tag{2}$$

where μ–is a small positive parameter; $\xi_{\frac{t}{\mu^2}} \in R_m$– is a gaussian process with independent components, zero expectation and finite standard deviation for which: $\int_{-\infty}^{\infty} M\left\{\xi_t\xi_{t+\tau}\right\} d\tau = 1$. Stratonovich in [1] prooved, that following Stratonovich stochastic differential equation (further SDE):

$$d\mathbf{x}_t = \mathbf{a}^*(\mathbf{x}_t, t)dt + \Sigma(\mathbf{x}_t, t)d\mathbf{f}_t, \tag{3}$$

where $\mathbf{f}_t \in R_m$–is a Wiener process with independent components, is a limit form as $\mu \to 0$ (on distribution) of equation (2). From mathematical point of view Ito equation is more convenient, then Stratonovich equation, therefore we shall proceed from (3) to the appropriate Ito equation:

$$d\mathbf{x}_t = \mathbf{a}^{**}(\mathbf{x}_t, t)dt + \Sigma(\mathbf{x}_t, t)d\mathbf{f}_t;\ \mathbf{a}^{**}(\mathbf{x}, t) = \mathbf{a}^{*}(\mathbf{x}, t) + \frac{1}{2}\Sigma(\mathbf{x}, t)\Sigma_f^2\left(\frac{\partial^{\top}}{\partial \mathbf{x}}\Sigma(\mathbf{x}, t)\right)^{\top}, \quad (4)$$

where

$$\Sigma_f^2 = diag\{\sigma_{f_i}^2\}_{i=1}^m;\ \frac{\partial^{\top}}{\partial \mathbf{x}} = \left[\frac{\partial}{\partial \mathbf{x}^{(1)}}, \ldots, \frac{\partial}{\partial \mathbf{x}^{(n)}}\right],$$

and $\sigma_{f_i}^2 < \infty$ - is an intensity of Wiener process component $\mathbf{f}_t^{(i)}$. Thus by virtue of connection of the equations (3) and (4) we shall consider the equation (4) in Ito form, omitting $**$, as a mathematical model of stochastic control systems.

2. Problem of numerical simulation of SDE. We shall consider the problem of numerical simulation of Ito SDE:

$$d\mathbf{x}_t = \mathbf{a}(\mathbf{x}_t, t)dt + \Sigma(\mathbf{x}_t, t)d\mathbf{f}_t;\ \mathbf{x}_0 = \mathbf{x}(0), \quad (5)$$

where $\mathbf{x}_t = \mathbf{x}(t, \omega) \in R_n$ is a process, being a solution of the system (5); $\mathbf{a}(\mathbf{x}, t) \in R_n$ and $\Sigma(\mathbf{x}, t) \in R_{n\times m}$ are nonrandom functions, which are multiply differentiated on collection of arguments and satisfying to conditions of existence and uniqueness of solution of the system (5).

The G.N. Milstein approach in [2] to the problem of numerical simulation of SDE give the idea to solution of this problem, which is based on expansion of the process $\mathbf{x}_s$ in a vicinity of a fixed instant t with using of a Ito formula. G.N. Milstein [2], and also E. Platen and W. Wagner [3], G. Liske, E. Platen and W. Wagner [4] obtained, on the basis of this approach, the Taylor-Ito expansion of the process $\mathbf{x}_s$ in a vicinity of a fixed point t up to smalls s-t about 3/2 order in mean square. The integration numerical method of the system (5) with a local mean square error of 3/2 order of a smallness was constructed in these works on the basis of a Taylor-Ito expansion up to smalls $O(\Delta^{3/2})$ (Δ is a step of numerical method). In G.N. Milstein monography [5], and also in [6,7] we can see the Taylor-Ito expansion up to smalls of second order. The numerical method with a local mean square error about $O(\Delta^2)$ is also constructed in these papers. The Taylor-Ito expansion up to smalls s-t about a smallness of 5/2 order and numerical method for the system (1) with a local mean square errors about $O(\Delta^{5/2})$ were obtained by P.E. Kloeden, I.W. Wright and E. Platen in [8], and also D.F. Kusnetsov in [9].

Thus, now in a literature the Taylor-Ito expansion of the process $\mathbf{x}_s$ in a vicinity of a fixed point t with a mean square error about $O((s-t)^{3/2})$, $O((s-t)^2)$ and $O((s-t)^{5/2})$ are known. Now the exacter Taylor-Ito expansions are unknown in a literature.

The authors of the present paper in [10] offered the approach to the solution of the problem of the recurrence Taylor-Ito expansion of the process $\eta_s = R(\mathbf{x}_s, s)$, which has allowed by a uniform image to receive the known and unknown in a literature numerical schemes for SDE. This approach is based on property of integration order replasement in repeated stochastic integrals.

3.The integration order replacement in repeated stochastic integrals.

Lemma 1 Let $\psi(t)$ is a non-random finite function and ϕ_t is a non-look-ahead random process for a Wiener process and $\mathbf{M}\{\phi_\tau^2\} < \infty$. Let constants $C_i < \infty, \gamma_i > 0$ such, that $\mathbf{M}\{(\phi_\tau - \phi_t)^2\} \leq C_1|t-\tau|^{\gamma_1}$ and $((\psi(t)-\psi(\tau))^2 \leq C_2|t-\tau|^{\gamma_2}$; $i = 1,2$ are exists for each monents $t, \tau \in [a,b]$. Then:

1. The mean-square limit:

$$J_{ab} = l.i.m. \sum_{j=0}^{N-1} \phi_{\tau_j} \left(W_{\tau_{j+1}}^{(i)} - W_{\tau_j}^{(i)}\right) \int_{\tau_{j+1}}^{b} \psi(t_1) dW_{t_1}^{(l)} \ \ (\text{as } max \mid \tau_{j+1} - \tau_j \mid \to 0)$$

is exists and by definition:

$$J_{ab} = \int_a^b \phi_t dW_t^{(i)} \int_t^b \psi(t_1) dW_{t_1}^{(l)};$$

where $i, l = 0, 1, \ldots, m$; $W_t^{(q)} = f_t^{(q)}$ for $q = 1, \ldots, m$ and $W_t^{(q)} = t$ for $q = 0$.

2. With probability 1:

$$\int_a^b \phi_t dW_t^{(i)} \int_t^b \psi(t_1) dW_{t_1}^{(l)} = \int_a^b \psi(t) \int_a^t \phi_{t_1} dW_{t_1}^{(i)} dW_t^{(l)}.$$

The following theorem is a generalization of the lemma 1 rezults.

Theorem 1 If $\psi_l(t)$; $l = 1, \ldots, k-1$ and ϕ_t are satisfy to conditions of lemma 1 for $\psi(t)$ and ϕ_t correspondingly, then:

1.The mean-square limit:

$$J_{ab}^{(k-1)} = l.i.m. \sum_{j=0}^{N-1} \phi_{\tau_j} \left(W_{\tau_{j+1}}^{(i)} - W_{\tau_j}^{(i)}\right) \int_{\tau_{j+1}}^{b} \psi_1(t_2) dW_{t_2}^{(l_1)} \ldots \int_{t_{k-1}}^{b} \psi_{k-1}(t_k) dW_{t_k}^{(l_{k-1})}$$

$$(\text{as } max \mid \tau_{j+1} - \tau_j \mid \to 0)$$

is exists and by definition:

$$J_{ab}^{(k-1)} = \int_a^b \phi_{t_1} dW_{t_1}^{(i)} \int_{t_1}^b \psi_1(t_2) dW_{t_2}^{(l_1)} \ldots \int_{t_{k-1}}^b \psi_{k-1}(t_k) dW_{t_k}^{(l_{k-1})}.$$

2.With probubility 1:

$$J_{ab}^{(k-1)} = \int_a^b \psi_{k-1}(t_k)... \int_a^{t_3} \psi_1(t_2) \int_a^{t_2} \phi_{t_1} dW_{t_1}^{(i)} dW_{t_2}^{(l_1)} \dots W_{t_k}^{(l_{k-1})};$$

where $i, l_1 \dots l_{k-1} = 0, 1, \dots, m;\ k \in N;\ W_t^{(q)} = f_t^{(q)}$ for $q = 1, \dots, m$ and $W_t^{(q)} = t$ for $q = 0$.

The proofs of lemma 1 and theorem 1 we can see in [15].

4. Recurrence fractional power stochastic (FPS) Taylor-Ito expantion of processes, which are generated by solution of Ito SDE. In this item we shall formulate the theorem, which is prooved in [10,12,14] (Kulchitski, Kuznetsov), of the stochastic prosess $\eta_s = R(\mathbf{x}_s, s)$ recurrence FPS Taylor-Ito expansion with a residual term in integrated form.

Theorem 2. Let Ito process $\eta_s = R(\mathbf{x}_s, s)$, which is generated by Ito SDE solution, is r+1 times differentiated in the sense of Ito (definition of differentiability in the sense of Ito see in [12]) on trajectories of the Ito SDE (5). Then it is decomposed in a vicinity of a fixed instant $t \in [0, T]$ in to the following recurrence FPS Taylor-Ito series with a residual term in integrated form for $s \geq t$:

$$\eta_s = \sum_{q=0}^{r} \sum_{k,j,l_1,\dots,l_k \in D_q} \frac{(s-t)^j}{j!} {}^{(k)}L^j G_{l_1} \dots G_{l_k} \{\eta_t\} \overset{k}{\bullet} {}^{(k)}J_{l_1 \dots l_{k_{s,t}}} + H_{r+1_{s,t}}, \tag{6}$$

where equality is fair with probability 1 and right part of (6) is exists in mean square and

$$H_{r+1_{s,t}} = \sum_{k,j,l_1,\dots,l_k \in U_r} \frac{(s-t)^j}{j!} {}^{(k)}L^j G_{l_1} \dots G_{l_k} \{\eta_t\} \overset{k}{\bullet} {}^{(k)}J_{l_1 \dots l_{k_{s,t}}} + D_{r+1_{s,t}};$$

$$D_{r+1_{s,t}} = \sum_{k,j,l_1,\dots,l_k \in A_r} \left(\int_t^s \frac{(s-\tau)^j}{j!} {}^{(k)}L^{j+1} G_{l_1} \dots G_{l_k} \{\eta_\tau\} \overset{k}{\bullet} {}^{(k)}J_{l_1 \dots l_{k_{s,\tau}}} d\tau + \right.$$
$$\left. + \int_t^s \frac{(s-\tau)^j}{j!} \left({}^{(k+1)}G_0 L^j G_{l_1} \dots G_{l_k} \{\eta_\tau\} \overset{k}{\bullet} {}^{(k)}J_{l_1 \dots l_{k_{s,\tau}}} \right) \overset{1}{\bullet} d\mathbf{f}_\tau \right);$$

$$\sqrt{M\left\{H_{r+1_{s,t}}^2\right\}} \leq C(s-t)^{\frac{r+1}{2}};\ C = const < \infty;$$

$${}^{(k)}L^j G_{l_1} \dots G_{l_k} \{\cdot\} = \left\| \underbrace{L\{\dots\{L}_{j}\{G_{l_1}^{(i_1)}\{\dots\{G_{l_k}^{(i_k)}\{\cdot\}\}\dots\}\}\}\dots\} \right\|_{i_1,\dots,i_k=1}^{m},$$

$${}^{(k+1)}G_0 L^j G_{l_1} \dots G_{l_k} \{\cdot\} = \left\| G_0^{(i_{k+1})} \underbrace{\{L\{\dots\{L}_{j}\{G_{l_1}^{(i_1)}\{\dots\{G_{l_k}^{(i_k)}\{\cdot\}\}\dots\}\}\}\dots\}\} \right\|_{i_1,\dots,i_{k+1}=1}^{m},$$

$$ {}^{(k)}L^{j+1}G_{l_1}\dots G_{l_k}\{\cdot\} = \left\| \underbrace{L\{\dots\{L}_{j+1}\{G_{l_1}^{(i_1)}\{\dots\{G_{l_k}^{(i_k)}\{\cdot\}\}\dots\}\}\}\dots\} \right\|_{i_1,\dots,i_k=1}^{m}, $$

$$ G_k^{(i)} = \frac{1}{k}\left(G_{k-1}^{(i)}\{L\{\cdot\}\} - L\{G_{k-1}^{(i)}\{\cdot\}\}\right);\ k=1,\ 2,\ \dots;\ i=1,\dots,m; $$

$$ {}^{(k)}J_{l_1\dots l_{k\,s,t}} = \left\| J_{l_1\dots l_{k\,s,t}}^{(i_1\dots i_k)} \right\|_{i_1,\dots,i_k=1}^{m}; $$

where

$$ J_{l_1\dots l_{k\,s,t}}^{(i_1\dots i_k)} = \begin{cases} \int_t^s (s-\tau_1)^{l_k}\int_t^{\tau_1}(s-\tau_2)^{l_{k-1}}\dots\int_t^{\tau_{k-1}}(s-\tau_k)^{l_1}d\mathbf{f}_{\tau_k}^{(i_1)}\dots d\mathbf{f}_{\tau_2}^{(i_{k-1})}d\mathbf{f}_{\tau_1}^{(i_k)} \\ \text{for } k>0 \\ 1 \text{ for } k=0 \end{cases}; $$

$$ A_q = \{(k,j,l_1,\dots,l_k):\ k+j+l_1+\dots+l_k = q\}; $$

$$ D_q = \{(k,j,l_1,\dots,l_k):\ k+2(j+l_1+\dots+l_k) = q\}; $$

$$ U_r = \{(k,j,l_1,\dots,l_k):\ k+2(j+l_1+\dots+l_k) \ge r+1; $$

$$ k+j+l_1+\dots+l_k \le r\};\ k,j,l_1,\dots,l_k = 0,1,\dots; $$

operation $\overset{l}{\bullet}$ —is:

$$ {}^{(k+l)}A \overset{l}{\bullet} {}^{(l)}B = \left\| \sum_{i_{k+1},\dots,i_{k+l}=1}^{m} A^{(i_1\dots i_{k+l})}B^{(i_1\dots i_l)} \right\|_{i_1,\dots,i_k=1}^{m} $$

where:

$$ {}^{(p)}C \overset{def}{=} \left\| \sum_{i_1,\dots,i_l=1}^{m} C^{(i_1\dots i_p)} \right\|_{i_1,\dots,i_p=1}^{m}; $$

$L\{\cdot\}$, $G_0^{(i)}$; $i=1,\dots,m$—are differential operators of drift and diffuzion from Ito formula.

Corollary. Let $R(\mathbf{x},t) \equiv \mathbf{x}$; $s=(p+1)\Delta$; $t=p\Delta; \Delta>0$; $p=0,\ 1,\dots$. Then from theorem 2 we have:

$$ \mathbf{x}_{(p+1)\Delta} = \sum_{q=0}^{r}\ \sum_{k,j,l_1,\dots,l_k\in D_q} \frac{\Delta^j}{j!}\,{}^{(k)}L^jG_{l_1}\dots G_{l_k}\{\mathbf{x}_{p\Delta}\} \overset{k}{\bullet}\,{}^{(k)}J_{l_1\dots l_{k(p+1)\Delta,p\Delta}} + H_{r+1_{(p+1)\Delta,p\Delta}}, \quad (7) $$

where

$$ \sqrt{M\left\{H^2_{r+1_{(p+1)\Delta,p\Delta}}\right\}} \le C\Delta^{\frac{r+1}{2}};\ C = const < \infty; $$

5. Numerical schemes for Ito SDE. From (7) follows the general formula for numerical scheme for Ito SDE with a local mean square error not more, than $C\Delta^{\frac{r+1}{2}}$ $(r=0,\ 1,\dots)$:

$$ \mathbf{x}_{p+1} = \sum_{q=0}^{r}\ \sum_{k,j,l_1,\dots,l_k\in D_q} \frac{\Delta^j}{j!}\,{}^{(k)}L^jG_{l_1}\dots G_{l_k}\{\mathbf{x}_p\} \overset{k}{\bullet}\,{}^{(k)}J_{l_1\dots l_{k(p+1)\Delta,p\Delta}}. \quad (8) $$

Consider some numerical schemes, which are follows from (8).

5.1 Euler scheme:

$$\mathbf{x}_{p+1} = \mathbf{x}_p + \Delta^{1/2} b_{1/2_{p+1,p}} + \mathbf{a}(\mathbf{x}_p, p\Delta); \quad b_{1/2_{p+1,p}} = \sum_{i_1=1}^{m} \sigma_{f_{i_1}} G_0^{(i_1)}\{\mathbf{x}_p\} I_{0_{p+1)\Delta,p\Delta}}^{(i_1)}.$$

5.2 Milstein scheme:

$$\mathbf{x}_{p+1} = \mathbf{x}_p + \Delta^{1/2} b_{1/2_{p+1,p}} + \Delta b_{1_{p+1,p}},$$

where

$$b_{1_{p+1,p}} = \mathbf{a}(\mathbf{x}_p, p\Delta) + \frac{1}{\sqrt{2}} \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \sigma_{f_{i_1}} \sigma_{f_{i_2}} G_0^{(i_2)} G_0^{(i_1)}\{\mathbf{x}_p\} I_{00_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)}.$$

5.3 Scheme with a local mean square error $O(\Delta^2)$:

$$\mathbf{x}_{p+1} = \mathbf{x}_p + \Delta^{1/2} b_{1/2_{p+1,p}} + \Delta b_{1_{p+1,p}} + \Delta^{3/2} b_{3/2_{p+1,p}},$$

where

$$b_{3/2_{p+1,p}} = \sum_{i_1=1}^{m} \sigma_{f_{i_1}} \left(\left(G_0^{(i_1)} L\{\mathbf{x}_p\} - L G_0^{(i_1)}\{\mathbf{x}_p\} \right) \frac{I_{1_{(p+1)\Delta,p\Delta}}^{(i_1)}}{\sqrt{3}} + \right.$$
$$\left. + L G_0^{(i_1)}\{\mathbf{x}_p\} I_{0_{(p+1)\Delta,p\Delta}}^{(i_1)} \right) +$$
$$+ \frac{1}{\sqrt{6}} \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \sum_{i_3=1}^{m} \sigma_{f_{i_1}} \sigma_{f_{i_2}} \sigma_{f_{i_3}} G_0^{(i_3)} G_0^{(i_2)} G_0^{(i_1)}\{\mathbf{x}_p\} I_{000_{(p+1)\Delta,p\Delta}}^{(i_3 i_2 i_1)}.$$

5.4 Scheme with a local mean square error $O\left(\Delta^{5/2}\right)$:

$$\mathbf{x}_{p+1} = \mathbf{x}_p + \Delta^{1/2} b_{1/2_{p+1,p}} + \Delta b_{1_{p+1,p}} + \Delta^{3/2} b_{3/2_{p+1,p}} + \Delta^2 b_{2_{p+1,p}},$$

where

$$b_{2_{p+1,p}} = \frac{L^2\{\mathbf{x}_p\}}{2} + \frac{1}{2} \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \sigma_{f_{i_1}} \sigma_{f_{i_2}} \left(G_0^{(i_2)} G_0^{(i_1)} L\{\mathbf{x}_p\} \frac{I_{01_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)}}{\sqrt{3}} \right.$$
$$+ G_0^{(i_2)} L G_0^{(i_1)}\{\mathbf{x}_p\} \left(I_{10_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)} - \frac{I_{01_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)}}{\sqrt{3}} \right) +$$
$$\left. + L G_0^{(i_2)} G_0^{(i_1)}\{\mathbf{x}_p\} \left(\frac{1}{\sqrt{2}} I_{00_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)} - I_{10_{(p+1)\Delta,p\Delta}}^{(i_2 i_1)} \right) \right)$$
$$+ \frac{1}{2\sqrt{6}} \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \sum_{i_3=1}^{m} \sum_{i_4=1}^{m} \sigma_{f_{i_1}} \sigma_{f_{i_2}} \sigma_{f_{i_3}} \sigma_{f_{i_4}} G_0^{(i_4)} G_0^{(i_3)} G_0^{(i_2)} G_0^{(i_1)}\{\mathbf{x}_p\} I_{0000_{(p+1)\Delta,p\Delta}}^{(i_4 i_3 i_2 i_1)}.$$

In schemes 5.1-5.4:

$$J_{l_k \ldots l_1 s,t}^{(i_1 \ldots i_k)} = \frac{\sigma_{f_{i_1}} \ldots \sigma_{f_{i_k}} \; I_{l_k \ldots l_1 s,t}^{(i_1 \ldots i_k)} \; \Delta^{k/2 + l_1 + \ldots + l_k}}{\sqrt{(2l_k + 1)(2(l_k + l_{k-1}) + 2) \ldots (2(l_k + \ldots + l_1) + k)}} \text{ with } M\left\{ \left(I_{l_k \ldots l_1 s,t}^{(i_1 \ldots i_k)} \right)^2 \right\} = 1.$$

Finite-difference modifications of schemes 5.1-5.4 and other numerical schemes see in [13].

6. Approximation of repeated stochastic integrals. For numerical realization of schemes from item 4 we must to approximate the repeated stochastic Ito integrals (further: Ito RSI) : $J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$. This is a very difficult problem. In [5,8] there is one method of Stratonovich RSI approximation. We shall consider more general method, which is contained in [11]. According to this method we shall represent Stratonovich RSI $(s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$ in form

$$(s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t} = \underbrace{\int_t^s \int_t^s \ldots \int_t^s}_{k} K(\tau_1,\tau_2,\ldots,\tau_k)d\mathbf{f}^{(i_k)}_{\tau_1}d\mathbf{f}^{(i_{k-1})}_{\tau_2}\ldots d\mathbf{f}^{(i_1)}_{\tau_k}, \tag{9}$$

where $K(\tau_1,\tau_2,\ldots,\tau_k) = (s-\tau_1)^{l_k}\ldots(s-\tau_k)^{l_1}$ for $\tau_k \le \tau_{k-1} \le \ldots \le \tau_1$ and $K(\tau_1,\tau_2,\ldots,\tau_k) = 0$ another. We shall expand $K(\tau_1,\tau_2,\ldots,\tau_k)$ on $\underbrace{[s,t]\times\ldots[s,t]}_{k}$ with helping of full ortochonal system $\{\varphi_l(\tau)\}_{l=0}^{\infty}$ on $[s,t]$ for which: $\int_t^s \varphi_l^2(\tau)d\tau = 1$. From this expansion and representation (9) we have:

$$(s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t} = \sum_{j_1,\ldots,j_k=0}^{\infty} C_{j_1\ldots j_k}\prod_{l=1}^{k}\int_t^s \varphi_{j_l}(\tau_l)d\mathbf{f}^{(i_{k-l+1})}_{\tau_l},$$

where $\sigma^{-2}_{f_{i_{k-l+1}}}\int_t^s \varphi_{j_l}(\tau_l)d\mathbf{f}^{(i_{k-l+1})}_{\tau_l}$ are independent gaussian random values with zero expectation and standard deviation 1. Let $(s)J^{*(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$ is an approximation of Stratonovich SRI $(s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$:

$$(s)J^{*(i_1\ldots i_k)}_{l_k\ldots l_1 s,t} = \sum_{j_1,\ldots,j_k=0}^{p} C_{j_1\ldots j_k}\prod_{l=1}^{k}\int_t^s \varphi_{j_l}(\tau_l)d\mathbf{f}^{(i_{k-l+1})}_{\tau_l},$$

where $p(p<\infty)$ is choosed from condition:

$$\sqrt{M\left\{\left((s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t} - (s)J^{*(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}\right)^2\right\}} \le \varepsilon.$$

Further we have an approximation of Ito RSI $J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$ from connection between Stratonovich and Ito RSI: $(s)J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$, $J^{(i_1\ldots i_k)}_{l_k\ldots l_1 s,t}$. For example:

$$J^{(i_1 i_2)}_{00_{s,t}} = \begin{cases} (s)J^{(i_1 i_2)}_{00_{s,t}},\ i_1 \ne i_2 \\ \frac{1}{2}\left(J^{(i_1)}_{0_{s,t}}\right)^2 - \frac{1}{2}\sigma^2_{f_{i_1}}(s-t),\ i_1 = i_2; \end{cases}$$

An approximation of integral $(s)J^{(i_1 i_2)}_{00_{s,t}}$ with helping of trigonometric system has an aspect:

$$(s)J^{*(i_1 i_2)}_{00_{s,t}} = \sigma_{f_{i_1}}\sigma_{f_{i_2}}\frac{(s-t)}{\sqrt{2}}\left(\xi^{(i_1)}_{0_{s,t}}\xi^{(i_2)}_{0_{s,t}} + \frac{1}{2\pi}\sum_{r=1}^{p}(\xi^{(i_1)}_{2r_{s,t}}\xi^{(i_2)}_{2r-1_{s,t}} -\right.$$

$$\left. -\xi^{(i_1)}_{2r-1_{s,t}}\xi^{(i_2)}_{2r_{s,t}} + \sqrt{2}(\xi^{(i_1)}_{2r-1_{s,t}}\xi^{(i_2)}_{0_{s,t}} - \xi^{(i_1)}_{0_{s,t}}\xi^{(i_2)}_{2r-1_{s,t}})\right).$$

where p is selected from condition for a mean square error of approximation, which has an aspect:

$$\frac{\sqrt{6}}{4\pi}\sigma_{f_{i_1}}\sigma_{f_{i_2}}\sqrt{\frac{\pi^2}{6}-\sum_{r=1}^{p}\frac{1}{r^2}} \le \varepsilon$$

and $\xi_i^{(j)}$ –are independent for difference i and j gaussian random values with zero expectation and standard deviation 1.

References

1. Stratonovich R.L. (1966) The conditional Markov processes and their application to theory of optimal control. Moscow University.

2. Milstein G.N. (1974) Approximate integration of stochastic differential equations. Theor. Prob. Appl., 19, 557-562.

3. Platen E., Wagner W. (1981) On a Taylor formula for a class of processes. Lietuvos Mathem. Rink., 21, 121-133

4. Liske H., Platen E., Wagner W. (1982) About mixed multiple Wiener integrals. Preprint P.-Math.- 23/82, J Math., Akad. Der. Wiss. Der. DDR, Berlin.

5. Milstein G.N. (1988) A Numerical integration of the stochastic differential equations. Ural university, Sverdlovsk.

6. Kloeden P.E., Platen E. (1989) A survey of numerical methods for stochastic differential equation. J. Stoch. Hydrol. Hydraul., 3, 155-178.

7. Pettersson R. (1992) Stratonovich-Taylor expansion and numerical methods. Stoch. Anal. And Applications, 10 (5), 603-612.

8. Kloeden P.E., Platen E., Wright I.W. (1992) The Approximation of multiple stochastic integrals. Stoch. Anal. And Applications, 10 (4), 431-441.

9. Kusnetsov D.F. (1996) Methods of numerical simulation of solutions of systems of the stochastic differential Ito equations in mechanics tasks. Dissertation paper, St.-Petersburg State Technical university (in Russian).

10. Kulchitski O.Yu., Kusnetsov D.F. (1993) Taylor-Ito expansion of a Ito process in a vicinity of a fixed instant. VINITI, N2637-V93 (in Russian).

11. Kulchitski O.Yu., Kusnetsov D.F. (1994) The Approximation of multiple stochastic Ito integrals. VINITI, N1678-V94 (in Russian).

12. Kulchitski O.Yu., Kusnetsov D.F. (1996) "Generalization of Taylor expansion for a class of processes, which are generated by solution of Ito SDE. VINITI N3507-V96 (in Russian).

13. Kusnetsov D.F. (1996) Finite-difference approximation of Taylor-Ito expansion and finite-difference methods of numerical integration of Ito SDE. VINITI, N3509-V96 (in Russian).

14. Kulchitski O.Yu., Kusnetsov D.F. (1996) The Taylor-Ito expansion of Ito processes,

wich are generated by a solution of differential Ito equation. The Fist International Scientific and Practical Conference. Differential Equations and Applications. St.Petersburg, 135-136.

15. Kulchitski O.Yu., Kuznetsov D.F. (1996) Repeated stochastic integrals and their properties. VINITI N3506-V96 (in Russian).

MODELING OF ELECTRICAL NONLINEAR CIRCUITS IN AUTOMATIC CONTROL SYSTEMS

Makarova M. V.

Higher Naval Engineering Academy

1, Kadetskiy bul., Pushkin,

St. Petersburg, Russia

e-mail: nba@hq.math.lgu.spb.su

Abstract. The cubic Lienard equation applied to the description of nonlinear electrical circuits is considered. Conditions of the presence of stable limit cycle are given. For example, an opportunity of modeling auto-oscillations in real electrical circuits is shown.

Keywords. Nonlinear electrical circuits, differential equations, Lienard equation, limit cycle, auto-oscillations.

Introduction. Dynamics of nonlinear electrical circuits is often described by ordinary differential equations of the second order [1]. A great part of these equations is resulted in the cubic Lienard equation after using polynomials of second [2] or third [2,3] degree for approximation of the characteristics of nonlinear elements and proper change of variables. In offered work the conditions for existence of a stable limit cycle will be given for this class of Lienard equations. The established fact allows to simulate auto-oscillations in electrical circuits.

Environmental Impact. Mathematical model, describing auto-oscillations in some nonlinear electrical circuits is suggested and analyzed.

We shall consider the Lienard equation

$$\ddot{x} + f(x)\dot{x} + g(x) = 0,$$

which by replacement

$$\dot{x} + F(x) = y, \quad F(x) = \int_0^x f(s)ds$$

is resulted in system of the differential equations

$$\begin{cases} \dot{x} = y - F(x), \\ \dot{y} = -g(x). \end{cases} \quad (1)$$

The point designates derivative with respect to time t.

The system (1) will be studied for

$$F(x) = a_1 x + a_2 x^2 + a_3 x^3,$$

$$g(x) = b_1 x + b_2 x^2 + b_3 x^3.$$

In work [4] conditions for existence of a limit cycle, surrounding three singular points, were receired for system (1) in general case.

Theorem. Let the following conditions are carried out:

1) $g(\alpha_{-1}) = g(0) = g(\alpha_1) = 0$, $\alpha_{-1} < 0$, $\alpha_1 > 0$, $xg(x) > 0 \forall x \notin [\alpha_{-1}, \alpha_1]$;

$xg(x) < 0 \forall x \in]\alpha_{-1}, 0[\cup]0, \alpha_1[$;

2) $(F(x))' = f(x) < 0 \forall x \in]\beta_{-1}, \beta_1[$, $\beta_{-1} < \alpha_{-1}$, $\beta_1 > \alpha_1$;

3) $\int_0^{\beta_i} g(x)dx = G(\beta_i) > 0$, $i = -1, 1$;

4) $\exists \varphi(x) \in C^1$ such, that:

a) $\varphi(0) = 0$; $\psi(x) = (\varphi(x) - F(x))' < 0 \forall x \notin [\delta_{-1}, \delta_1], \delta_{-1} < \alpha_{-1}, \delta_1 > \alpha_1$;

b) $\Phi(x) = \varphi(x) + \dfrac{g(x)}{\psi(x)}$ or is bounded above at $x \le -r$, or is bounded below at $x \ge r$,

where $r > max(|\delta_{-1}|, \delta_1)$;

5) $\overline{\lim_{x \to +\infty}} \varphi(x) = +\infty$

6) $\underline{\lim_{x \to -\infty}} \varphi(x) = -\infty$

Then the system (1) has at least one stable limit cycle, surrounding three singular points.

In our case system (1) signs a kind

$$\begin{cases} \dot{x} = y - a_1 x - a_2 x^2 - a_3 x^3, \\ \dot{y} = -b_1 x - b_2 x^2 - b_3 x^3. \end{cases}$$

In addition we shall introduce functions:

$$f(x) = 3a_3x^2 + 2a_2x + a_1$$

$$G(x) = \frac{b_3}{4}x^4 + \frac{b_2}{3}x^3 + \frac{b_1}{2}x^2.$$

The requirements on these function, at which all conditions of the theorem will be satisfied:

1) Let the function $F(x)$ has two nonzero roots of different sings: $x_2 < 0 < x_1$,

$$x_{1,2} = \frac{-a_2 \pm \sqrt{a_2^2 - 4 a_1 a_3}}{2a_3}$$

2) Let the function *g(x)* has two nonzero roots of different sings: $\tilde{x}_2 < 0 < \tilde{x}_1$

$$\tilde{x}_{1,2} = \frac{-b_2 \pm \sqrt{b_2^2 - 4 \cdot b_1 b_3}}{2 \cdot b_3}$$

3) We shall consider: $a_1 < 0$, $a_3 > 0$, $b_1 < 0$, $b_3 > 0$. The case $a_1 > 0$, $a_3 < 0$ is resulted in considered case, if to change y on -y, t on -t in system (2).

4) The function $F(x)$ has extremum in points $\hat{x}_1; \hat{x}_2$ which are the roots of the equation $f(x) = F'(x) = 0$.

$$\hat{x}_{1,2} = \frac{-a_2 \pm \sqrt{a_2^2 - 3a_1a_3}}{3a_3}$$

5) Let the roots x_r, x_l of function $G(x)$

$$x_{r,l} = \frac{-b_2 \pm \sqrt{b_2^2 - \frac{9}{2}b_1b_3}}{\frac{3}{2}b_3}$$

satisfy the unequality: $x_r < \hat{x}_1$, $x_l > \hat{x}_2$

Statement.

At fulfilment of the requirements 1) - 5), if

$$x_2 < \hat{x}_2 < x_l < \tilde{x}_2 < 0 < \tilde{x}_1 < x_r < \hat{x}_1 < x_1$$

conditions of the theorem are satisfied, and, hence, the system (2) has at least one stable limit cycle surrounding three singular points $(x_2, F(x_2)); (0,0); (x_1, F(x_1))$.

Proof.

Let's assume $\alpha_{-1} = \tilde{x}_2$, $\alpha_1 = \tilde{x}_1$, $\beta_{-1} = \hat{x}_2$, $\beta_1 = \hat{x}_1$

Let the requirements 1) - 4) are carried out. Then the function $g(x)$ satisfies to a condition 1) of the theorem.

The function $F(x)$ decreases at $x \in [\beta_{-1}, \beta_1]$. Hence, $f(x) < 0$ on this interval. Thus, the condition 2) of the theorem is carried out.

The requirement 5) provides fulfilment of a condition 3).

We shall show the validity of condition 4) of the theorem.

We shall choose $\varphi(x) = \frac{1}{2}F(x)$.

The condition 4a) is carried out, since

$\varphi(0) = 0$, $\psi(x) = (\varphi(x) - F(x))' = -\frac{1}{2}(F(x))' = -\frac{1}{2}f(x) < 0$ at

$x \notin [\delta_{-1}, \delta_1]$, where $\delta_{-1} = \beta_{-1} < \alpha_{-1}$, $\delta_1 = \beta_1 > \alpha_1$

and $\psi(\delta_{-1}) = -\frac{1}{2}f(\beta_{-1}) = 0$; $\psi(\delta_1) = -\frac{1}{2}f(\beta_1) = 0$

since β_{-1}, β_1, are the extremum points $f(x) = (F(x))'$.

The condition 4b) of the theorem is carried out too.

Really, let us consider function $\Phi(x) = \varphi(x) + \dfrac{g(x)}{\psi(x)}$.

In our case $\Phi(x) = \frac{1}{2}F(x) - \dfrac{2g(x)}{f(x)}$.

We are to choose number $r > \max(|\delta_1|, \delta_1)$, let $r = x_0 > x_1$.

At $x_0 \geq x_1$ it can be shown, that $\Phi(x)$ is bounded below.

At the choice of function $\varphi(x) = \frac{1}{2}F(x)$ since $F(x) > 0$ at $x \in (x_1, +\infty), F(x) < 0$ at $x \in (-\infty, x_2)$ we conclude, that conditions 5) and 6) of the theorem are carried out:

$$\overline{\lim_{x \to +\infty}} \varphi(x) = \frac{1}{2} \overline{\lim_{x \to +\infty}} F(x) = +\infty;$$

$$\underline{\lim_{x \to -\infty}} \varphi(x) = \frac{1}{2} \underline{\lim_{x \to -\infty}} F(x) = -\infty.$$

We have shown, that all conditions of the theorem, are carried out ,hence, its conclusion is valid. Thus, at our assumptions the system (2) has at least one stable limit cycle, surrounding three singular points.

The statement is proved.

Example.

$$\begin{cases} \dot{x} = y + 15x - 2x^2 - x^3, \\ \dot{y} = 2x - x^2 - x^3. \end{cases}$$

$$F(x) = x^3 + 2x^2 - 15x;$$

$$g(x) = x^3 + x^2 - 2x;$$

$$x_1 = 3; \tilde{x}_1 = 1; \hat{x}_1 = \frac{5}{3}; x_r = 1{,}44;$$

$$x_2 = -5; \tilde{x}_2 = -2; \hat{x}_2 = -3; x_l = -2{,}77.$$

The singular points:

$O_1(1,-12), O_2(-2,30)$ - unstable nodes, $O\ (0,0)$ - saddle.

The system has a stable limit cycle, surrounding three singular points O_1, O, O_2

Hence, it is possible to simulate an electrical circuit in a mode of auto-oscillations by choosing the appropriate values of parameters in system (2).

Conclusions. The conditions of auto-oscillation regime in nonlinear electrical circuits are obtained by using qualitative theory of ordinary differential equations.

References

1. Danilov L.V., P. N. Mathanov, E. S. Filippov (1990). The theory of nonlinear electrical circuits. L.: Energoatomizdat.

2. Flatto L. (1964). Limit cycle studies for circuils containing one Esaki diode. J. of Math. Annal. and Appl. V. 9, pp.360-383.

3. Grishkina N.A., V. I. Korolev (1971). Dynamics of the circuit on tunnel diodes. High School's news. Radiophysics, V. 14, no. 7, pp. 1078-1083.

4. Jilevich L.I. (1994). About existence of a limit cycle of one autonomus system. Diff. eq., V.30, no. 3, pp. 377-380.

5. Sidorov A.S. (1971). The theory and projection of the nonlinear impulse circuits with tunnel diodes. M.: Sov. Radio.

NON-LINEAR MODEL OF TOKAMAK PLASMA SHAPE STABILIZATION

Boris A. Misenov, Alexander D. Ovsyannikov,
Dmitri A. Ovsyannikov, Evgenii I. Veremei, Alexei P. Zhabko
St. Petersburg State University
Institute of Computational Mathematics and Control Processes
2, Bibliotechnaya sq.
Petrodvorets, St. Petersburg, 198904 Russia
e-mail: mis@vrm.apmath.spb.su

Abstract. One of the most important problems of modern natural science is controlled thermonuclear fusion. A few approaches to this problem solution are known — among them nuclear reactor-tokamak occupies the main position. The first question connected with this type of thermonuclear reactors is the plasma shape control problem. In the present publication the results of plasma shape stabilization problem investigation in the main element of thermonuclear reactor-TOKAMAK are presented. The development of the complex of appropriate mathematical models and methods is an integral ITER project constituent of demonstrational and operating TNR versions construction.

Keywords. Tokamak, plasma control, non-linear control, modal control, LQG-optimization

One of the central problems to be investigated is plasma control in the toroidal magnetic chamber. It is well-known that plasma as an object of control is a mobile non-stable dynamic system with extremely high values of physical parameters, temperature as high as millions centigrades, first of all. This fact rises very high the requirement for electromagnetic capture system control accuracy. The special attention is paid to the stabilization of the plasma pinch to chamber walls gap in some fixed control points. Therefore the high requirements for dynamic characteristics of the control process arise: duration, over-regulation, dispersion, oscillations etc. The mathematical model describing the plasma shape control dynamics [1] is an ordinary differential equation system

$$\frac{d\Psi}{dt} + RI = V, \tag{1}$$

here Ψ is magnetic flux vector, I is current through active and passive magnetic coils vector, V is vector of voltages applied to coils, R is diagonalized resistance matrix. System (1) is supplemented with non-linear equations

$$\Psi = \Psi\left(I, I_p\right), \ \mathbf{g} = \mathbf{g}\left(I, I_p\right), \tag{2}$$

where $\mathbf{g}$ — regulated gaps vector, I_p — nominal plasma pinch current.

Non-linear system (1), (2) can be linearized in the operating mode vicinity, this operation leads to the linear differential equation system with constant coefficients

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} = \mathbf{C}\mathbf{x}, \end{cases} \tag{3}$$

where $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{u} \in \mathbf{R}^m$, $\mathbf{y} \in \mathbf{R}^m$ ($n = 22$, $m = 7$) are deviations of current, voltage and gaps from the nominal values.
In linear system (3) analysis the following facts are important.
1. Matrix **A** has first eight zero columns, this fact is the sequence of the super-conductivity of the active coils and leads to zero eigenvalues with multiplicity eight.
2. Therefore, system (3) is not totally controllable in Kalman's sense. So, the dimensionality of the non-controlled part is unity and its eigenvalue is zero.
3. Among the matrix **A** eigenvalues one lays in the open right semi-plane, that is determined with the non-stable plasma shape in the vertical direction.

The central problem of the present paper is the analytical determination of controls

$$\mathbf{u} = \mathbf{W}(p)\mathbf{y}, \tag{4}$$

stabilizing system (3). Here $p \equiv d/dt$, $\mathbf{W}(p)$ — the transfer matrix of the controller (4). We investigate the possibility of various approaches to $\mathbf{W}(p)$ determination together with the developed methods adaptation to the specific features of the system under consideration.
Amongst other approaches to controller (4) synthesis a special attention is paid to the use of the following methods:

— modal control theory;
— means-square optimization theory.

*1. **Modal approach.*** Two problems are investigated in the frameworks of this approach. First is related to the synthesis of the stabilizing controllers of PD-structure

$$\mathbf{u} = \mathbf{K}\mathbf{y} + \mathbf{P}\dot{\mathbf{y}}, \tag{5}$$

here **K** and **P** are constant matrices.
Remark that in spite of the simplicity, the choice of controller matrices is difficult because in this situation it is impossible to pre-describe arbitrary characteristic polynomials roots for the closed system (3), (5).
Various kinds of methods are proposed as a general concept of matrices **K** and **P** search for auxiliary problems solution, like

$$I = I(\mathbf{K},\mathbf{P}) \to \min_{(\mathbf{K},\mathbf{P})\in\Omega}, \tag{6}$$

here $I(\mathbf{K},\mathbf{P})$ — some functional, characterizing the matrices **K** and **P** choice quality in the limits of the feasible set Ω.
The simplest variant of problem (6) is the problem of closed system stability degree maximization.
As «stability degree» we shall designate the function

$$\alpha = \alpha(\mathbf{K},\mathbf{P}) = -\max_{i=\underline{1,n}} \operatorname{Re}\delta_i(\mathbf{K},\mathbf{P}),$$

where $\delta_i(\mathbf{K},\mathbf{P})$ are the characteristic polynomial roots:

$$\Delta(s,\mathbf{K},\mathbf{P}) = \det(\mathbf{E}s - \mathbf{A} - \mathbf{BF}),$$

$$\mathbf{F} = (\mathbf{E}_m - \mathbf{PCB})^{-1}(\mathbf{KC} + \mathbf{PCA}).$$

Then the auxiliary problem (6) is

$$I(\mathbf{K},\mathbf{P}) = -\alpha(\mathbf{K},\mathbf{P}) \to \min_{(\mathbf{K},\mathbf{P})\in\Omega},$$

here $\Omega = \mathbf{E}^{m^2} \times \mathbf{E}^{m^2}$.
Remark that the solution of this problem can be obtained with any numeric method that does not use the function gradient as such, because of its indifferentiability.
The transient process supported with the best controller, synthesized on the (6) basis is characterized with 8 s duration as the controlling impacts lie in the linear domain and the over-regulation value P=5%.

The second problem of this class is related to the construction of a controller (4) as a regulator with asymptotic observer (Kalman's filter for example)

$$\begin{cases} \dot{\mathbf{z}} = \mathbf{Az} + \mathbf{Bu} + \mathbf{H}(\mathbf{x} - \mathbf{z}), \\ \mathbf{u} = \mathbf{Kz}, \end{cases} \tag{7}$$

where $\mathbf{z} \in \mathbf{R}^{22}$ — observer state vector, $\mathbf{H}$ — a given coefficient discrepancy matrix. For (7) structure controller we have an equation

$$I = I(\mathbf{K}) \to \min_{\mathbf{K}\in\Omega_k}, \tag{8}$$

of over-regulation minimization on the set

$$\Omega_k = \{\mathbf{K} : \delta_i(\mathbf{K}) \in \Omega_c,\ \Delta(\delta_i(\mathbf{K})) = 0,\ \Delta(s) = \det(\mathbf{E}s - \mathbf{A} - \mathbf{BK})\}, \tag{9}$$

where Ω_c — given area at the root plane, $i = \overline{1,22}$. A method of its solution with reduction to an unconditional extremum problem is proposed.

2. ***Mean-square optimization.*** According to this approach some disturbances and white noises are introduced into the mathematical model of a controlled object and instead of (3) the system

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} + \varphi(t), \\ \mathbf{y} = \mathbf{Cx} + \psi(t), \end{cases} \tag{10}$$

is investigated, where $\varphi(t)$, $\psi(t)$ are the white noises with constant spectral density matrices Φ and Ψ accordingly. It is known [2], that LQG-optimization problem consists of the construction of the controller

$$\begin{cases} \dot{\mathbf{z}} = \mathbf{Az} + \mathbf{Bu} + \mathbf{H}(\mathbf{x} - \mathbf{z}), \\ \mathbf{u} = \mathbf{Kz}. \end{cases} \tag{11}$$

Then matrix $\mathbf{K}$ is found as a solution of the control determination problem

$$\mathbf{u} = \mathbf{Kx}, \tag{12}$$

minimizing the quadratic functional

$$I = I(\mathbf{K}) = \int_0^\infty \left(\mathbf{y}^T \mathbf{R}\mathbf{y} + c_0 \mathbf{u}^T Q \mathbf{u}\right) dt \tag{13}$$

in the Hurwitz characteristic polynomial area.
Matrix **H** is constructed as a solution of Kalman's filtration problem [2] for a functional

$$J = J(\mathbf{H}) = \left\langle (\mathbf{x}(t) - \mathbf{z}(t))^T (\mathbf{x}(t) - \mathbf{z}(t)) \right\rangle. \tag{14}$$

If the controls have some constraints like saturation

$$u_i(t) = \begin{cases} -u_{i0}, & u_i(t) < u_{i0,} \\ u_i, & -u_{i0} \le u_i(t) \le u_{i0,} \\ u_{i0}, & u_i(t) > u_{i0,} \end{cases} \tag{15}$$

a new method for LQG optimization is proposed with weight coefficient c_0 choise in functional (13). The idea lies in the construction of the mean-square accuracy of stabilization dependency on the energy expenditure for a non-linear system with the further search for its minimum, determining the value $c_0 = c_{0p}$, assumed as the operating point. The transient process, supported with the synthesized LQG-optimal regulator for c_{0p} is characterized with duration T=6 s during the control impact overcome of the linear region and the over-regulation value P=10%.
The comparison of the modal methods and LQG-methods application shows that in spite of better speed of response, LQG-regulator has a fault of structure complexity, economy and over-regulation value.
In the frameworks of the mean-square synthesis, along with LQG-approach, some other optimization methods in $\mathbf{H}_2$ and $\mathbf{H}_\infty$ spaces that are intensively developed nowadays can be used.
Unlike LQG-synthesis here the coloured disturbances and noises in the measurements are investigated for mathematical model

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \varphi(t), \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \psi(t), \end{cases} \tag{16}$$

here $\varphi(t)$, $\psi(t)$ — the stationary stochastic processes with the given spectral density matrices $\Phi(\omega)$ and $\Psi(\omega)$ accordingly, the components of them being the even rational-fractional functions.
The essence of this approach is the construction of a controller

$$\mathbf{u} = \mathbf{W}(p)\mathbf{y},$$

supporting the mean-square functional minimum

$$I = I(\mathbf{W}) = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \left(\mathbf{y}^T \mathbf{R}\mathbf{y} + c_0 \mathbf{u}^T Q \mathbf{u}\right) dt \tag{18}$$

in Hurwitz characteristic polynomial for the closed system (16), (17).
If spectral density matrices are given, the problem can be solved with new spectral methods of $\mathbf{H}_2$-optimization.
However, the formulation of the problem in the guaranteed variant is possible, as there are some uncertainties in the spectral densities assignment. So, such uncertainty can be described with the condition

$$\Phi(\omega) \in \mathbf{P} = \left\{\Phi(\omega) : \|\Phi(\omega)\|_{\mathbf{H}_2} = 1\right\}. \tag{19}$$

In this case the search for a guaranteeing evaluation I_0 of (18) and guaranteeing controller $\mathbf{u} = \mathbf{W}^0(p)\mathbf{y}$ can be presented as

$$I_0 = \inf_{\mathbf{W}\in\Omega} \sup_{\Phi\in\mathbf{P}} I(\mathbf{W}, \Phi), \tag{20}$$

$$\mathbf{W}^0 = \arg \inf_{\mathbf{W}\in\Omega} \sup_{\Phi\in\mathbf{P}} I(\mathbf{W}, \Phi), \tag{21}$$

To solve (20), (21) the optimization methods in $\mathbf{H}_\infty$-space can be applied [3]. A close links of optimization in $\mathbf{H}_2$ and $\mathbf{H}_\infty$ norms determines the corresponding new spectral approaches to the guaranteeing controller synthesis. This approach is adapted to the plasma shape stabilizing controller synthesis.

3. ***Optimal stabilization curves.*** Let consider the weight coefficient c_0 problem. We minimize the functional (13) which consists of two addenda. The first of them is

$$I_x = \int_0^\infty \mathbf{x}^T \mathbf{C}^T \mathbf{R} \mathbf{C} \mathbf{x}\, dt$$

and gives the quadratic error of control. The second,

$$I_u = \int_0^\infty \mathbf{u}^T Q \mathbf{u}\, dt$$

gives the integral quadratic control power. Therefore we can represent our functional (13) as

$$I(c_0, \mathbf{K}) = I_x + c_0 I_u$$

and obtain parametric dependencies $I_x = I_x(c_0)$ and $I_u = I_u(c_0)$, i. e. parametric dependence $I_x = I_x(I_u)$. These values can be obtained for a given initial conditions $x(0) = x_0$ and two variants of the system model. The first model is linear (without control saturation) and the corresponding curve will be decrease monotone function. The second model takes into account the control saturation so the curve is not monotone and it has minimum. It is reasonable to choose the value of variable parameter c_0 near this minimum.
The results of computations for particular model are presented in the Fig. 1. Two curves are depicted. The lower curve is the optimal stabilization curve for the linear model (without

control saturation).The two upper curves correspond two various saturation models. The chosen points of parameter c_0 value are denoted as R.

References

1. Albanese R., V. Coccorese, G. Rubinacci (1989). Plasma modelling for the control of vertical instabilities in Tokamaks. Nuclear Fusion, 29, no. 6.
2. Kwakernaak H., R. Sivan (1977). Linear optimal control system. Mir, Moscow.
3. Bokova Ya. M., E. I. Veremei (1995). Computational aspects of $\mathbf{H}_2$-optimal synthesis spectral method. Izvestiia RAN.Teoriia i sistemy upravleniia, 4, 88–96.

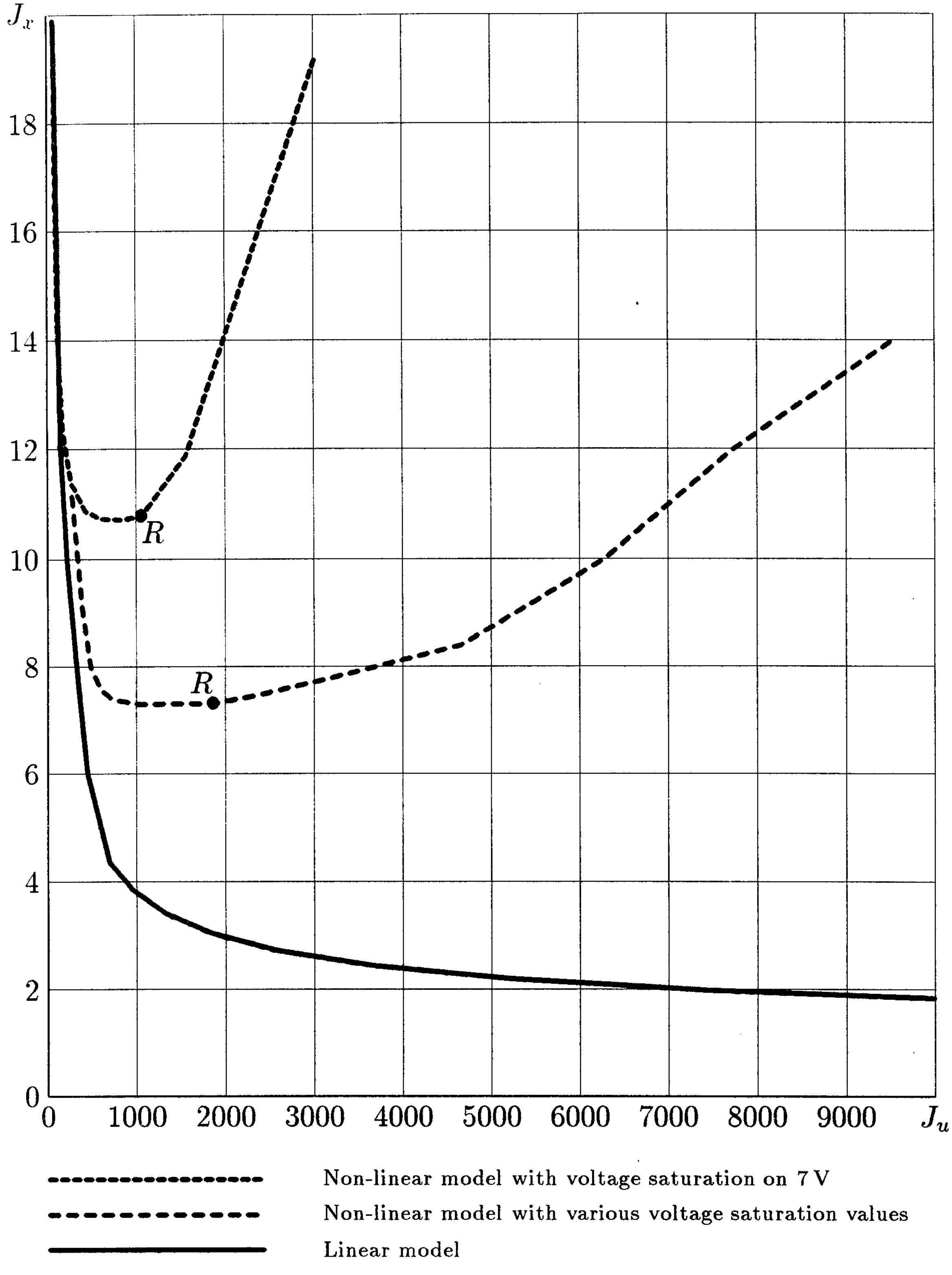


Fig. 1. Optimal stabilization curves for linear and non-linear methods

DYNAMICS OF NONLINEAR HYSTERESIS CONTROL SYSTEMS UNDER LARGE OSCILLATIONS OF LIFTING BODIES

R.A.Nelepin, A.M.Kamachkin
St.Petersburg State University
Department of Applied Mathematics
2, Bibliotetchnaya sq., Petrodvorets
St. Petersburg , RUSSIA
e-mail: postmaster@num.apmath.lgu.spb.su

Abstract. The existence of periodic modes in nonlinear control systems, together with exact construction of these modes, obtaining their properties, parameters and phase space configurations constitute the fundamental problems of nonlinear oscillation theory. Another no less challenging problem arising concurrently is posed by design of control in the nonlinear system to provide the onset of oscillation processes with desired properties. This paper is concerned with analytic solution of these problems for certain classes of nonlinear controlled systems containing bang-bang hysteresis nonlinearities and special forms of periodic external disturbance affecting the controlled entity. The problem is solved by two techniques - method of fitting and method of sections developed by one of the authors in nineteen-sixties.

Keywords: Automatic Control; Nonlinear System; Periodic Oscillations; External Disturbance; Dynamic Behavior of Control System; Bang-bang Hysteresis.

Introduction. Mathematical models under consideration may arise in control of spatial motion of a system of solid bodies, for instance, a guided satellite whose structural features allow to treat it as a solid body moving under bang-bang control and an external force varying according to harmonic law. The same model could be used to determine the effect of periodic disturbing force on the performance of shipborne automatic systems in case when the sea is regular. System under consideration

$$\dot{\overline{X}} = A_0\overline{X} + B_0F(\sigma) + K_0\overline{F}(t),\ \sigma = (C, \overline{X}), \tag{1}$$

where $\overline{X} = (x_1, ..., x_n)'$ - n-dimensional state vector of the system (' - transposition symbol) ; $A_0 = \{a^0_{ks}\}$ - n-dimensional constant real matrix; $B_o = (b^0_1, ..., b^0_n)$, $C = (c^0_1, ..., c^0_n)'$ - n-dimensional constant real column matrices; nonlinear function $F(\sigma)$ is given as follows : $F(\sigma) = m_1$, for $-\infty < \sigma < l_2$, $F(\sigma) = m_2$, for $l_1 < \sigma < +\infty$, where m_1, m_2, l_1, l_2 - real constants, with $m_1 < m_2$, $l_1 < l_2$. Function $u = F(\sigma)$ describes a bang-bang relay with hysteresis with threshold numbers l_1, l_2 and two output values m_1, m_2, hysteresis loop is run counterclockwise in coordinates (F,σ). Moreover, it is assumed that the following conditions are satisfied :

$$-(C,A_0^{-1}B_0 m_1) > l_2,\ -(C,A_0^{-1}B_0 m_2) < l_1.$$

Given scalar time function $\overline{F}(t)$ describes the external disturbance of the controlled entity, generally

$$\overline{F}(t) = F_0 + e^{\alpha t} \sum_{\rho=1}^{n} F_\rho \sin(\rho\omega t+\varphi_\rho), \tag{2}$$

where F_0, F_ρ, φ_ρ ($\rho = \overline{1,n}$) - real constants, $T = 2\pi/\omega$ - period of function $\sum_{\rho=1}^{n} F_\rho \sin(\rho\omega t+\varphi_\rho)$;

ω - circular frequency.

Problem Formulation. Investigation Techniques. The objective of analysis includes the following: first, we propose exact analytic techniques for investigation of nonautonomous systems of form (1), (2), which enable to ascertain the existence of steady-state oscillations from the system's coefficients, and use these techniques to develop algorithms enabling to obtain the parameters of steady-state oscillations; second, identify the domain corresponding to harmonic modes within the coefficient space of systems (1), (2), and thus to determine the safe range of coefficient selection precluding the occurrence such undesirable, application-wise, modes as beats and strange attractors. In other words, we need to construct bifurcation surfaces in the coefficient space of system (1), (2), the crossing of which will either change the number of periodic solutions or altogether alter the pattern of system's dynamic behavior.

The formulated problems are first solved by constructing a set of transcendental equations in moments and points where periodic solution intersects with hyperplanes of form $(C, \overline{X}) = l_i$ $(i = 1, 2)$, thus determining the breakdown of function $F(\sigma)$ into linear sections. If there exists such $t = T_b$ so that $\overline{X}(0) = \overline{X}(T_b)$, where $\overline{X}(t)$ - Cauchy-form solution of set (1), (2); then it follows that the initial vector $\overline{X}_0 = \overline{X}(0)$ can be written as

$$\overline{X}_0 = \left[E - e^{A_0 T_b}\right]^{-1} \int_0^{T_b} e^{A_0 (T_b - t)} B_0 \left[F(s) + K_0 \overline{F}(t)\right] dt \tag{3}$$

where E — identity ($n \times n$) matrix. Therefore, T_b —periodic solution of set (1) and (2) can defined formally as follows

$$\overline{X}(t) = e^{A_0 t}\left[E - e^{A_0 T_b}\right]^{-1} \int_0^{T_b} e^{A_0 (T_b - \tau)} B_0 \left[F(\sigma(t)) + K_0 \overline{F}(\tau)\right] d\tau + \\ + \int_0^t e^{A_0 (t-\tau)} B_0 \left[F(\sigma(\tau)) + K_0 \overline{F}(t)\right] d\tau \tag{4}$$

The properties of functions σ(t) and $\overline{F}$ should be known.

We shall apply solution (4) to obtain transcendental equations enabling to determine the parameters of periodic solution describing the forced vibration of system with bang-bang hysteresis described by function F(σ). Periodic solutions are assumed to belong to the class of periodic solutions with two switch points at hyperplanes (C, $\overline{X}$) = l_i (i = 1, 2) and period T_b = n·T, where n - positive integer.

Let points $\overline{X}_1$ and $\overline{X}_2$ lie on periodic trajectory and (C, $\overline{X}_1$) =l_1, (C, $\overline{X}_2$) =l_2. The representative point will return to its initial position in time [0,T_b], then the following equalities will hold

$$\overline{X}_1 = \left[E - e^{A_0 T_b}\right]^{-1} \left\{ e^{A_0 \tau_1} \int_{\tau_1}^{T_b} e^{A_0 (T_b - \tau)} B_0 \left[m_1 + K_0 \overline{F}(\tau)\right] d\tau + \right. \\ \left. + \int_0^{\tau_1} e^{A_0 (\tau_1 - \tau)} B_0 \left[m_2 + K_0 \overline{F}(\tau)\right] d\tau \right\} = \left[E - e^{A_0 T_b}\right]^{-1} Q_1, \tag{5}$$

$$\overline{X}_2 = \left[E - e^{A_0 T_b}\right]^{-1} \left\{ e^{A_0 (T_b - \tau_1)} \int_{\tau_1}^{T_b} e^{A_0 (\tau_1 - \tau)} B_0 \left[m_2 + K_0 \overline{F}(\tau)\right] d\tau + \right. \\ \left. + \int_{\tau_1}^{T_b} e^{A_0 (T_b - \tau_1)} B_0 \left[m_1 + K_0 \overline{F}(t)\right] d\tau \right\} = \left[E - e^{A_0 T_b}\right]^{-1} Q_2, \tag{6}$$

Switch conditions and equalities (5) and (6) yield the set of transcendental equations to solve for switch times τ_1 and $\tau_2 = T_b - \tau_1$:

$$l_1 = (C, \left[E - e^{A_0 T_b}\right]^{-1} Q_1);$$

$$l_2 = (C, \left[E - e^{A_0 T_b}\right]^{-1} Q_2); \tag{7}$$

Generally, such a system could not be treated by analytical approach, therefore the transformation $\overline{X}$ = PX is applied to initial automatic control system in the case when the characteristic equation of system (1), (2) has distinct roots, the transformation converts the system to an equivalent one [1]:

$$\dot{X} = AX + BF(\sigma) + K\overline{F}(t),\ \sigma = (\Gamma, X), \tag{8}$$

where X = $(x_1, ..., x_n)'$ (' - transposition symbol), A - diagonal NxN matrix, B = (1, ...,1)', K - certain n-dimensional vector; $\Gamma = (\gamma_1, ..., \gamma_n)$, with

$$\gamma_i = (D'(\lambda_i))^{-1} \sum_{k=1}^{n} c_k N_k(\lambda_i),\ i = \overline{1,n} \tag{9}$$

where λ_i — roots of equation D(p) = det $[A_0 - E_p]$ = 0, D'(p) = dD(p)/dp, c_k — components of vector C; $N_k(p) = \sum_{i=1}^{n} b_{ik}^0 D_{ik}(p)$, b_{ik}^0 - components of vector B_0, $D_{ik}(p)$ - cofactor of (i, k) element of determinant D(p). If there are repeated roots among the solutions of equation D(p) = 0, the applied transformation $\overline{X}$ = PX is written in a somewhat different form [1].

If (n-2) roots of equation D(p) = 0 are identical to (n-2) roots of equation $\sum_{k=1}^{n} c_k N_k(p) = 0$; then (n-2) values γ_i , determined by formulae (9), vanish [2]. In this case function $\sigma(t)$ is obtained from the set of second-order equations, with the rest of $x_i(t)$ variables obtained from linear nonhomogeneous first-order equations. If, for example, matrix A is a 2x2 matrix, function $\sigma(t)$ is determined from a first-order equation. In the same way, it may be assumed that $\gamma_i = 0$ (i = $\overline{1, n-1}$) in case of system (7); then function $\sigma(t)$ is determined from a first-order differential equation, with the rest of x_i variables obtained by successive integration of (n-1) nonhomogeneous linear differential equations.

Hence, in the described cases the solution of transcendental equation set (7) for n-dimensional system (1), (2) is reduced to solution of transcendental equation set for a two-dimensional or one-dimensional system.

Provided that values τ_1 and τ_2 and vectors $\overline{X}_1$ and $\overline{X}_2$ satisfy the initial system (1), (2), conditions imposed on parameters of transcendental equation set (7) to guarantee the existence of solutions τ_1 and τ_2 and, therefore, solutions $\overline{X}_1$ and $\overline{X}_2$, will be the sufficient conditions for parameters of initial system (1), (2) to guarantee the existence of periodic solutions with period $T_b = n\cdot T$ and two switchpoints.

Algorithm for determination of periodic modes and bifurcation surfaces. **Since the problem of transcendental equation set for n-dimensional system (1), (2) has been reduced to solution of two transcendental equations of form (7), consider system (1) for n=2 and the following conditions: characteristic equation has two distinct real roots λ_1 and λ_2, while the external disturbance is a pure sine, i.e. $\overline{F}(t) = K\sin(\omega t+\varphi)$. If the linear transformation of the initial system was chosen so as to provide $\gamma_1 = 0$, $\gamma_2 = c_2$, and $T_b = n\cdot T$, the transcendental equation system takes the following form:**

$$l_1 c_2^{-1}(1-e^{\lambda_2 T_b}) = \lambda_2^{-1}(m_2 - m_1)\, e^{\lambda_2\tau_1} + m_1\lambda_2^{-1}e^{\lambda_2 T_b} - m_2\lambda_2^{-1} + $$
$$+K(e^{\lambda_2 T_b}-1)\left[\frac{\lambda_2}{\lambda_2^2+\omega^2}\sin(\omega\tau_1+\varphi)+\frac{\omega}{\lambda_2^2+\omega^2}\cos(\omega\tau_1+\varphi)\right], \qquad (10)$$

$$l_2 c_2^{-1}(1-e^{\lambda_2 T_b}) = (m_2 - m_1)\lambda_2^{-1}e^{\lambda_2(T_b-\tau_1)} + m_2\lambda_2^{-1}e^{\lambda_2 T_b} - m_1\lambda_2^{-1} + $$
$$+K(e^{\lambda_2 T_b}-1)\left[\frac{\lambda_2}{\lambda_2^2+\omega^2}\sin\varphi+\frac{\omega}{\lambda_2^2+\omega^2}\cos\varphi\right], \qquad (11)$$

Some easy manipulations of system (10), (11) yield equations used to obtain τ_1:

$$\Phi(T_b) = \sin[\omega\tau_1 + (\varphi + \mathrm{arctg}(\omega/\lambda_2))], \qquad (12)$$

where $\Phi(T_b)$ — constant calculated from

$$\Phi(T_b) = K^{-1}(e^{\lambda_2 T_b}-1)^{-1}[l_1 c_2^{-1}(1-e^{\lambda_2 T_b}) - \lambda_2^{-1}(m_2-m_1)e^{\lambda_2\tau_1} - m_1\lambda_2^{-1} + m_2\lambda_1^{-1}]. \qquad (13)$$

Value $e^{\lambda_2\tau_1}$ determined from equation (11) is substituted in formula (13), to wit,

$$e^{\lambda_2\tau_1} = (-m_2 l_2^{-l} e^{l_2 T_b} + m_l l_2\, e^{\lambda_2 T_b})*$$
$$*\left\{\frac{l_2}{c_2}(1-e^{\lambda_2 T_b}) - (e^{\lambda_2 T_b}-1)\left[\frac{K\lambda_2}{\lambda_2^2+\omega^2}\sin\varphi + \frac{K\omega}{\lambda_2^2+\omega^2}\cos\varphi\right] - \frac{m_2}{\lambda_2}e^{\lambda_2 T_b} + \frac{m_1}{\lambda_2}\right\}^{-1} \qquad (14)$$

Denominator on the right-hand side of formula (14) must not vanish.

Equation (12) generally has an infinite number of positive solutions τ_1 provided that $|\Phi(T_b)| \leq 1$, and has no solutions if $|\Phi(T_b)| > 1$. In a particular case when $T_b = T$, where T — period of external sinusoidal disturbance of system (1), (2), equation (12) have no more than two solutions τ_1' and τ_1'', therefore the initial system (1), (2) have no more than two distinct periodic solutions with two switch points.

Equation of form

$$|\Phi(T_b)| = 1 \tag{15}$$

may be treated as a bifurcation surface equation in parameter space ($\lambda_1, \lambda_2, c_1, c_2, K, m_1, m_2, l_1, l_2$). For instance, in the particular case when $T_b = T$ the realization of equality (15) means that two distinct periodic modes converges into one which disappears altogether if the parameters of system (1), (2) subsequently vary so that inequality $|\Phi(T_b)| > 1$ is realized. If the external disturbance of system (1), (2) takes the form $\overline{F} = K \cdot e^{\alpha t} \cdot \sin(\omega t + \varphi)$, we have the equation of the following form instead of (12)

$$Q(T_b)\, e^{-\alpha\tau_1} = \sin[\omega\tau_1 + (\varphi + \mathrm{arctg}\frac{\omega}{\lambda_2})], \tag{16}$$

where $Q(T_b)$ - constant which could be calculated from known parameters $\lambda_1, \lambda_2, c_1, c_2, K, m_1, m_2, l_1, l_2, T_b$.

This paper presents a simple and instructive case of construction of bifurcation surface equation in multidimensional parameter space.

The equation is obtained from condition of transcendental equation set (7) solvability for τ_1 and τ_2. When τ_1 is found, the switch points of periodic mode are calculated from formulae (5) and (6).

Conclusions. As opposed to well-known technique of shuttle iterations [2] the proposed approach does not require the Hurwitz property of matrix A_0 and allow to consider cases when $T_b = n \cdot T$ for $n > 1$.

References

1. Nelepin, R.A. (1967) Exact Analytical Methods in Theory of Nonlinear Automatic Systems. Sudostroyeniye, Leningrad. (in Russian).

2. Pokrovskiy, A.V. (1986) Existence and Calculation of Stable Modes in bang-bang Systems. Avtomatika i Telemekhanika, 1986, No.4, 16-23. (in Russian).

DISCRETE ROBUST STABILIZABILITY OF LINEAR MINIMUM-PHASE SAMPLED SYSTEMS WITH DELAY

Georgi Penev, Elena Ivanova
St.Petersburg Institute for Informatics and Automation of RAS
39, 14th Line, St.Petersburg, 199178
e-mail: penev@mail.iias.spb.su

Abstract. Linear minimum-phase sampled systems with time-delay and unknown order and system matrices are considered. Problems of robust regulators design and determination of robust stabilizability conditions are solved. For any compact subset of the set of minimum-phase systems is defined a set of discrete regulators stabilizing each system of the subset. The constructed regulators application need only the determination of the system degree and leading coefficient of the numerator polynomial of the continuous system transfer function.

Keywords. Robust control, stability, discrete system, system with time delay.

Introduction. Robust control synthesis is usually based on the stability testing of interval polynomials [1-3]. But many important problems have not been solved because of the complicated nonlinear dependence of the coefficients of polynomials that describe the closed-loop systems. One of them is the problem of robust discrete control design for minimum-phase linear sampled systems. This problem has been solved in [4,5] for "small enough" values of the sampling period and regulators, depending only on the system degree and leading coefficient of the numerator polynomial of the system transfer function have been designed. Robust regulators for a class of linear sampled systems with delay have been investigated in [6]. But there exists no such common interval of available values of the sampling period that regulators suggested in [4-6] could stabilize any system of the considered set of systems. Thus the whole class of mimimum-phase systems can not be said to be robustly stabilizable.

In this paper we treat both the problems of design of robust regulators of minimum-phase systems with delay and determination of robustly stabilizable sets, i.e. such subsets of the set of systems under consideration that there exists a common interval of available values of the sampling period for the entire subset.

Design of robust regulators. Consider linear delay control systems $\boldsymbol{S}$ described in continuous time by the equations

$$\frac{dx(t')}{dt'} = Ax(t') + Bu(t'-\tau), \qquad (1)$$

$$y(t') = Cx(t'), \qquad (2)$$

where $u(t'), x(t'), y(t')$ are the vectors of control, system state and output at the instant of time $t' \in [0,+\infty)$, $u \in R^l, x \in R^n, y \in R^p, \tau$ is the delay, $\tau > 0$, A, B, C are constant matrices of appropriate dimensions, $A \in R^{n\times n}, B \in R^{n\times l}, C \in R^{p\times n}$.

For any piecewise-constant control with sampling period $\delta > 0$,

$$u(t') = u_t, t' \in [t\delta, (t+1)\delta), \; t = 0,1,... \tag{3}$$

system (1), (2) is determined in discrete time, t=0,1, , by the finite-difference "input-output" equation of the form

$$a(h,\delta)y_t = b(h,\delta)u_t \,, \tag{4}$$

In (4) h is the back shift operator, $h\,y_t \overset{def}{=} y_{t-1}\,,\; y_t \overset{def}{=} y(t\delta)\,,$ a and b are polynomials in h,

$$a(h,\delta) = \sum_{i=0}^{n} a_i(\delta)h^i, a_i \in R, a_0 \overset{def}{=} 1 \,,$$

$$b(h,\delta) = \sum_{j=0}^{n} b_j(\delta)\, h^j, b_j \in R^{p\times l} \,.$$

In [4] has been shown that for any system $\boldsymbol{S}$ there exists such number $m \geq 1$ that

$$\beta_0 \overset{def}{=} CA^{m-1}B \neq 0, CA^kB = 0, k = \overline{0, m-2} \,.$$

The number m is said to be degree of system $\boldsymbol{S}$. Another definition of the system degree is based on the "input-output" description of system (1),(2) by the differential equation of the form

$$y^{(n)}(t') + \alpha_1 y^{(n-1)}(t') + ... + \alpha_n y(t') = \beta_0 u^{(s)}(t'-\tau) + ... + \beta_s u(t'-\tau) \,, \tag{5}$$

where $\alpha_i \in R, i = \overline{1,n}, \beta_j$ are $p \times l$ matrices, $\beta_j \in R^{p\times l}, j = \overline{0,s}\, \beta_0 \neq 0$. The degree of system (5) is $m \overset{def}{=} n - s$.

Robust discrete regulators that stabilize system (4) are designed below under the following conditions of uncertainty: the order n and elements of matrices A,B,C are unknown. It is only assumed that the degree m, delay τ and coefficient β_0 are given. The class of such systems $\boldsymbol{S}$ is denoted by $\sum_n(m,\tau,\beta_0)$. Control (3) is said to be robust if it stabilizes any system $\boldsymbol{S}$ of a given subset $\tilde{\Sigma}$ of set $\sum_n(m,\tau,\beta_0)$ and control (3) is specified by any equations with time-invariant parameters.

The design of the suggested regulators is based on the use of the simple model defined by equation

$$\nabla^m y_{t+1} = \tilde{\beta}\, r(h)u_t \,, \tag{6}$$

where

$$\nabla \overset{def}{=} 1-h, \widetilde{\beta} = \beta_0 \frac{\delta^m}{m!}, r(h) = \sum_{i=0}^{m} r_i h^i,$$

$$r_j = \sum_{k=0}^{j} (-1)^k C_{m+1}^k (j-k+1-\sigma)^m, \; j = \overline{0,m}, \; \sigma \overset{def}{=} \frac{\tau}{\delta}.$$

For SISO systems control (3) is determined by the equality

$$\widetilde{\beta}\,\gamma(h)u_t = \mu(h)y_t \quad (7)$$

where $\gamma(0) \overset{def}{=} 1$ Polynomials $\gamma(h), \mu(h)$ in (7) have to stabilize the closed-loop system (6), (7),

$$d(h)y_t = 0 \quad (8)$$

The characteristic polynomial d in (8) has the form

$$d(h) = (1-h)^m \gamma(h) - hr(h)\mu(h) \quad (9)$$

Theorem 1. *Assume that in (1) τ has the form $\tau = \sigma\delta$ with $\sigma \in [0,1)$ independent on δ and γ, μ are any polynomials satisfying equation (9) with any stable polynomial d. Then for any minimum-phase SISO-system $S \in \sum_n (m, \tau, \beta_0)$, $1 \le m \le n$, there exists such $\delta_* > 0$ that for any $\delta \in (0, \delta_*)$ the closed-loop system (4), (6) is asymptotically stable.*

For MIMO-systems we have the following theorem.

Theorem 2. *Assume that $S \in \sum_n (m, \tau, \beta_0)$ is any minimum-phase system with delay τ satisfying the condition of theorem 1, the inequality $\dim\{y\} \le \dim\{u\}$ is valid, the rang of matrix β_0 is full, $rang\,\beta_0 = p$, and $H \in R^{l \times p}$ is any matrix such, that $rang\,H = p$ and matrix $\beta_0 H \in R^{p \times p}$ is invertible (for example $H = \beta_0^T$). Define control (3) by the equalities*

$$u_t = H(\widetilde{\beta} H)^{-1} v_t, v_t \in R^p, \quad (10)$$

$$\gamma(h)v_t = \mu(h)y_t, \quad (11)$$

where $\widetilde{\beta} = \frac{\delta^m}{m!}\beta_0$ and polynomials γ, μ with scalar coefficients satisfy (7), (9) for any stable polynomial d. Then there exists such $\delta_* > o$ that for any $\delta \in (0, \delta_*)$ the linear discrete system (4), (10), (11) is asymptotically stable.

Note that system (6) is completely controllable since $\nabla^m = (1-h)^m$ and $r(1) \neq 0$ Thus equation (9) is solvable for each polynomial d, in particular, when it is stable [7]. Setting

$$\gamma(h) = 1 + \gamma_1 h + \ldots + \gamma_m h^m,$$

$$\mu(h) = \mu_0 + \mu_1 h + \ldots + \mu_{m-1} h^{m-1},$$

$$d(h) = 1 + d_1 h + \ldots + d_{2m} h^{2m},$$

one can determine the unknown coefficients of polynomials γ, μ by solving $2m$-ordered linear system of algebraic equations that corresponds to the polynomial equation (9).

Note that equations (10),(11) are equivalent with equation (7) if the values of u_t, x_t are scalar. Thus theorem 1 can be considered as a special case of theorem 2. On this basis only the case of MIMO-system is investigated below.

Similar theorems as formulated above are valid for the problem of output tracking. As in [4] in this case is assumed that the desired output $y_k^d, k = 0,1,\ldots$, satisfies the following condition:

c)The sequence $y_k^d, k = 0,1,\ldots,$ *is determined by the equality* $y_k^d = \tilde{y}^d(k\delta)$, *where function* $\tilde{y} \in C^{(n)}(T)$ *and it's derivatives* $\tilde{y}^{(i)}, i = \overline{1,n}$, *are continuous and bounded for all* $t \in T$

In the case under consideration the regulator can be defined as follows

$$u_t = H(\tilde{\beta} H)^{-1} v_t, v_t \in R^p, \tag{12}$$

$$\gamma(h) v_t = \mu(h) z_t, z_t \overset{def}{=} y_t - y_t^d, \tag{13}$$

Regulator described by (12), (13) ensures the "final" stability of the desired motion, i.e. for any $\varepsilon > 0$ there exist such $\tilde{\delta}(\varepsilon) > 0$ and $\tilde{t}(\varepsilon)$ that inequality $\|y_t - y_t^d\| \leq \varepsilon$ is valid for any $\delta \in (0, \tilde{\delta}(\varepsilon))$ and $t \geq \tilde{t}$.

Theorem 3. *Assume that the conditions of theorem 2 are fulfilled ,the desired motion* $y_t^d \in R^p$ *satisfies condition a) and the control* u *is defined by (12),(13). Then for any minimum-phase system* $S \in \sum_n(m, \tau, \beta_0)$, *described in continuous time by equations (1),(2), the desired motion* y_t^d *for discrete system (4) is finally stable with respect to the sampling period* δ.

The proofs of theorems 1-3 are quite similar to the analogous results of the paper [4] for systems without delay.

Robust stabilizability of compact sets. In [8] has been proved that any compact subset of the considered here class of systems is robustly stabilizable when $\tau = 0$. It turns out that similar assertion is available if $\tau \neq 0$ Note that any subset of set $\sum_n(m,\tau,\beta_0)$ is said to be compact [8] if the corresponding subset $\Xi = \{\xi\}$ of vectors $\xi \in R^n$ which components are the values of $\alpha_1,\ldots,\alpha_n$ and elements of matrices $\beta_1,\ldots,\beta_s$ in (3), is compact in Euclidean space $R^N, N \overset{def}{=} spl + n$

Theorem 4. *Assume that* $\hat{\Sigma}(m,\tau,\beta_0) = \bigcup_{n=1}^{N_1} \hat{\Sigma}_n(m,\tau,\beta_0)$ *with unknown* N_1 *is a compact subset of the set* $\sum_n(m,\tau,\beta_0)$ *of minimum-phase systems (5) satisfying the conditions of theorem 2. Then for any polynomials* $\gamma(h),\mu(h)$ *satisfying (9) with any stable polynomial* $d(h)$ *there exists such* $\delta_* > 0$ *that for any* S $\hat{\Sigma}(m,\tau,\beta_0)$ *and* $\delta \in (0,\delta_*)$ *system (4), (10), (11) is asymptotically stable.*

Proof. From theorem 1 follows the existence of such $\delta_1 = \delta_1(\xi)$, that for any $\delta \in (0,\delta_1)$ the closed-loop system (4), (10), (11) is asymptotically stable. Assume that there exists no a common value $\delta > 0$ that regulator (10), (11) stabilizes all the systems (4) from $\hat{\Sigma}(m,\tau,\beta_0)$. Then there exists such infinite sequence $\{\xi\}_{k=1}^{\infty}$ that $\lim_{k\to\infty} \delta_1(\xi) = 0$ Let now Ξ be the set of vectors ξ of parameters of systems from $\hat{\Sigma}(m,\tau,\beta_0)$ .From the compactness of this set follows the existence of such subsequence ξ_{k_i} that is converging to a vector $\xi_* \in \Xi$ and $\lim_{i\to\infty} \delta_1(\xi_{k_i}) = 0$ Therefor we have $\delta_1(\xi_*) = 0$, but $\delta_1(\xi) > 0$ for any $\xi \in \Xi$ Hence the assumption is not true, and theorem 4 is thus proved.

Quite similarly on the basis of theorem 3 can be proved the following theorem about the robustness of compact sets for the problem of output tracking.

Theorem 5. *Assume that the conditions of theorem 3 are satisfied and* $\sum_n(m,\tau,\beta_0)$, $\hat{\Sigma}(m,\tau,\beta_0)$ *are the sets defined as in theorem 4. Then there exists such value* $\delta_*(\xi) > 0$ *that the desired output* y_t^d *of any system (4) corresponding to a system from* $\hat{\Sigma}(m,\tau,\beta_0)$ *with regulator (12), (13) and* $\delta \in (0,\delta_*)$ *is finally stable with respect to the sampling period* $\delta \in (0,\delta_*)$.

Evidently theorem 2 can be applied when the set Ξ is a parallelepiped in R^{n+s} in the case of SISO-systems. In this case it is easy to verify that systems from $\hat{\Sigma}(m,\tau,\beta_0)$ are minimum-phase by using Kharitonov theorem [1]. Theorem 4 can be applied also in more complicated cases when the set $\{(\beta_1,\ldots,\beta_s)\}$ is a parallelepiped in R^s and the set $\{(\alpha_1,\ldots,\alpha_n)\}$ is any compact set in R^n

References

1. Kharitonov,V.L. (1979). The Routh-Hurwitz problem for families of polynomials and quazipolynomials. Mathematical Physics, no.26. Nauk. dumka, Kiev. (in Russian).
2. Zhabko,A.P., V.L. Kharitonov (1993). Methods of linear algebra in control problems. St.Petersburg University, St.Petersburg. (in Russian).

3. Barmish,B.R. (1993). New tools for robustness of linear systems. Macmillan Publishing Company, New York.
4. Penev, G.D. (1992). Robust control of linear systems. Questions of mechanics and control processes. Issue 15. St.Petersburg University, St.Petersburg, pp. 134--145 (in Russian).
5. Penev, G.D. (1995). Robust and adaptive control of nonlinear and quazi-linear systems. Prepr. of the Nonlinear Control Systems Design Symposium, NOLCOS'95, v.1. Tahoe City,California, pp. 331--335.
6. Ivanova,E.G. (1994).Robust Control of Linear Delay Systems. Vestnik of St.Petersburg University: Mathematics, v.1, no.1, pp.1--7 (in Russian).
7. Fomin, V.N., A.L.Fradkov, V.A.Jacubovich (1981).Adaptive control of dynamic objects. Nauka, Moscow (in Russian).
8. Penev,G.D. (1997). Discrete robust stabilizability of linear minimum-phase continuous systems. Vestnik of St.Petersburg University . Ser.1, Issue 1 (no. 1), pp. 30-35 (in Russian).

AN INVESTIGATION OF THE NONLINEAR CONTROL SYSTEM BY THE SECTIONS INTO PARAMETER SPACE METHOD

Vladimir N. Shamberov
St.Petersburg State Marine Technical University
Department of Automation and Measurement
3 Lotsmanskaya Street
St.Petersburg 190008, RUSSIA
e-mail: kuz@mtu-ic.spb.su

Abstract. An example of a forth-order nonlinear control system containing non-Coulomb model of dry friction is under investigation. For that the method of sections into parameter space is used.

Keywords. Control system, nonlinearity, parameter (coefficient) space, persistent oscillations, limit cycle, phase portrait, bifurcation surface.

Introduction. Nonlinear control systems containing functions which are not analytical at all points are most complicated for a rigorous research.

Of all known investigation techniques let us take the sections into parameter space method introduced by R.A. Nelepin for the nonlinear tasks of the automatic control theory [1, 2]. While investigating nonlinear high-order systems this method makes it possible to use analytically precise results of an investigation into low-order sub-systems.

Nonlinear automatic system. Let us consider the control system dynamics equation.

$$\left.\begin{aligned} &\dot{\eta}_1 = \eta_2 ; \\ &\dot{\eta}_2 = \eta_3 ; \\ &\dot{\eta}_3 = -\mathbf{a}\eta_3 - \mathbf{b}\eta_2 - \mathbf{c}\eta_1 + \xi ; \\ &\dot{\xi} = \mathbf{f}(\sigma) ; \\ &\sigma = -\mathbf{c}_1\eta_1 - \mathbf{c}_2\eta_2 - \mathbf{c}_3\eta_3 - \mathbf{r}\xi . \end{aligned}\right\} \quad (1)$$

Where : $\eta_1, \eta_2, \eta_3, \sigma, \xi$ - variables ; $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{r}$ - real coefficients ; $\mathbf{f}(\sigma)$ -nonliear function ; the dot denotes differentiation with respect to time **(t)**. The linear path (three first equations) of the system is totally controlled by ξ.

The structural mathematical model of this control system is illustrated in Fig. 1.

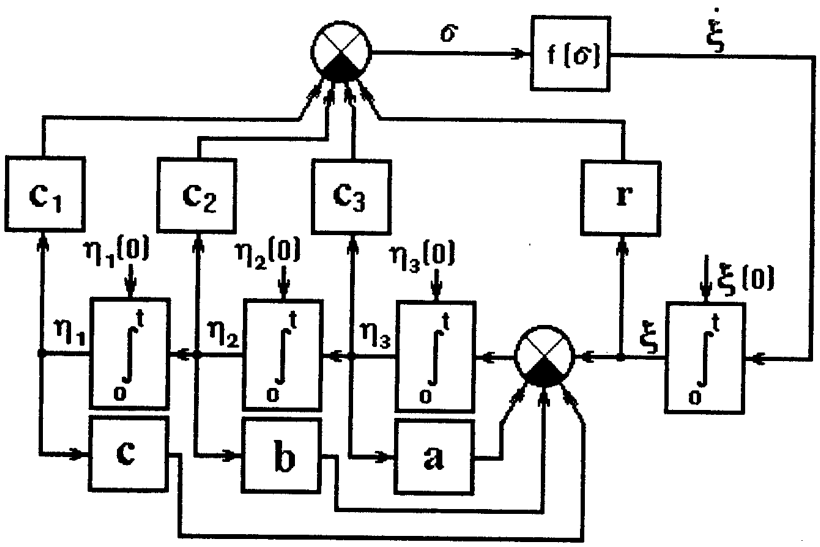


Fig. 1 Structural mathematical model of the initial system

The given system belongs to the power-operated control systems class, introduced by A.I.Lurie [3] for the automatic control theory tasks :

$$\dot{\eta}_k = \sum_{\alpha=1}^{n-1} b_{k\alpha}\eta_\alpha + h_k\xi, \quad \dot{\xi} = f(\sigma), \quad \sigma = \sum j_s\eta_s - r\xi \tag{2}$$

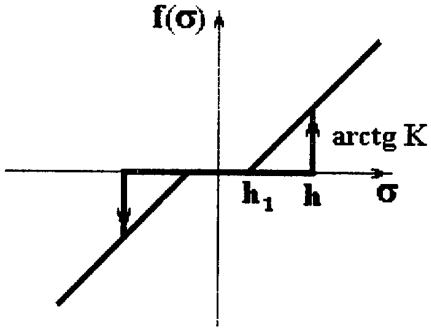
when :

$$n = 4;\ k = 1,2,3;\ b_{11} = b_{13} = b_{21} = b_{22} = 0;\ h_1 = h_2 = 0;\ b_{12} = b_{23} = 1;$$
$$h_3 = 1;\ b_{31} = -c;\ b_{32} = -b;\ b_{33} = -a;\ j_1 = -c_1;\ j_2 = -c_2;\ j_3 = -c_3.$$

Non-Coulomb dry friction model leeds to non-analytic function (the nonlinearity) are obtained as

$$f(\sigma) = \begin{cases} K\sigma - Kh\mathrm{Sign}(\sigma), & \text{when } |b| > h_1 \ \&\ f_0 \neq 0\ ; \\ 0, & \text{when } |b| < h \ \&\ f_0 = 0\ ; \\ K(h - h_1)\mathrm{Sign}(\sigma), & \text{when } |b| = h \ \&\ f_0 = 0\ ; \end{cases} \tag{3}$$

here f_0 - the previous value of the function ; $h,\ h_1,\ K$ - real positive coefficients (parameters of the nonlinearity), where $h > h_1$.



It is a piecewise-linear multivalued function shown in Fig. 2 representing the combination of nonlinearities : « hysteresis », « dead zone », « jump ». Such nonlinearities are brought about by simulation of dry friction in control actuators.

Fig. 2 Nonlinearity diagram

Let's consider that coefficients **a, b, c** are given, and coefficients $\mathbf{c_1, c_2, c_3, r, h, h_1, K}$ are parameters that must be defined to make the dynamic behavior which is needed: stability, static and dynamic accuracy, etc.

Linear transformation [2]

$$\eta_k = -\sum_{i=1}^{n-1} \frac{N_k(\lambda_i)}{\lambda_i D_1'(\lambda_i)} x_i - \frac{N_k(0)}{D_1(0)} \xi , \qquad (4)$$

where :

$n = 4;\ k = 1,2,3;\ D_1'(\lambda_i) = 3\lambda_i^2 + 2a\lambda_i + b;\ N_1(\lambda_i) = 1;\ N_2(\lambda_i) = \lambda_i;\ N_3(\lambda_i) = \lambda_i^2;$
$N_1(0) = N_3(0) = 0;\ D_1(0) = c;\ \lambda_1,\ \lambda_2,\ \lambda_3 -$ equation roots of :

$\lambda^3 + a\lambda^2 + b\lambda + c = 0;\ x_1,\ x_2,\ x_3$ - new variables the initial system can be put in form :

$$\left.\begin{aligned}
&\dot{x}_1 = \lambda_1 x_1 + f(\sigma)\ ;\\
&\dot{x}_2 = \lambda_2 x_2 + f(\sigma)\ ;\\
&\dot{x}_3 = \lambda_3 x_3 + f(\sigma)\ ;\\
&\dot{\xi} = f(\sigma)\ ;\\
&\sigma = -\gamma_1 x_1 - \gamma_2 x_2 - \gamma_3 x_3 - \gamma_4 \xi\ ;
\end{aligned}\right\} \qquad (5)$$

for all that :

$$\gamma_1 = \frac{c_1 + c_2\lambda_1 + c_3\lambda_1^2}{\lambda_1 D_1'(\lambda_1)}\ ;\ \ \gamma_2 = \frac{c_1 + c_2\lambda_2 + c_3\lambda_2^2}{\lambda_2 D_1'(\lambda_2)}\ ;\ \ \gamma_3 = \frac{c_1 + c_2\lambda_3 + c_3\lambda_3^2}{\lambda_3 D_1'(\lambda_3)}\ ;\ \ \gamma_4 = \frac{c_1}{c} + r\,.$$

Structural mathematical model of the system is shown in Fig. 3

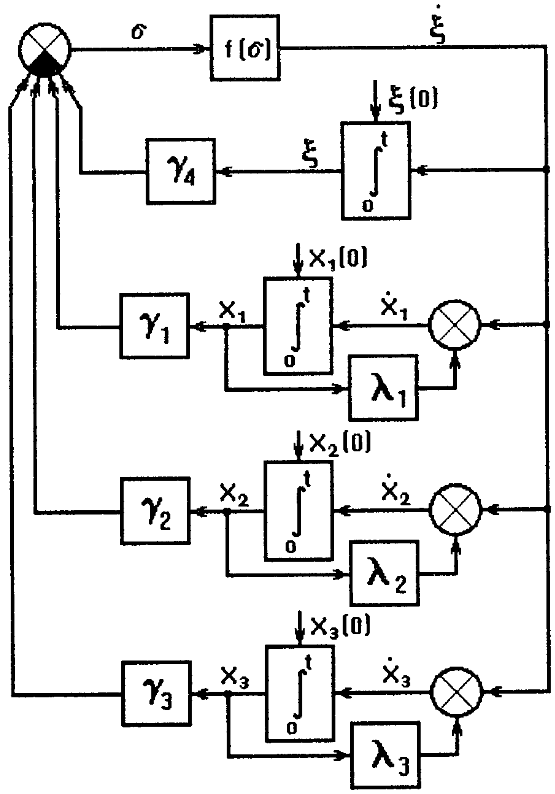


Fig. 3 Structural mathematical model of the transformed system

The linear transformation won't be special, if the roots λ_1, λ_2, λ_3 are simple roots [2]. It means that the initial system as well as the transformed system is interchangeable in the process of investigation, that is the stability (instability) of the initial system leads to the stability (instability) of the transformed one, etc.

Parameters (coefficients) space system structure. The system (5) has the following sections [4], inside which the fourth-order system dynamical behaviour is similar to the second-order closed systems and the first-order systems dynamical behaviour :

a) if roots are real and negative - there are three sections :

$$\left.\begin{aligned} c_1 + c_2\lambda_1 + c_3\lambda_1^2 = 0\ ; \\ c_1 + c_2\lambda_2 + c_3\lambda_2^2 = 0\ ; \end{aligned}\right\} \quad (6)$$

$$\left.\begin{aligned} c_1 + c_2\lambda_1 + c_3\lambda_1^2 = 0\ ; \\ c_1 + c_2\lambda_3 + c_3\lambda_3^2 = 0\ ; \end{aligned}\right\} \quad (7)$$

$$\left.\begin{aligned} c_1 + c_2\lambda_2 + c_3\lambda_2^2 = 0\ ; \\ c_1 + c_2\lambda_3 + c_3\lambda_3^2 = 0\ ; \end{aligned}\right\} \quad (8)$$

b) if one root is real and positive (for instance λ_1), and two others (for instance λ_2 and λ_3) are real and negative - there is one section :

$$\left.\begin{aligned} c_1 + c_2\lambda_2 + c_3\lambda_2^2 = 0 ; \\ c_1 + c_2\lambda_3 + c_3\lambda_3^2 = 0 ; \end{aligned}\right\} \quad (9)$$

c) if one root is real, and two others are complex conjugate with a negative real part - there is one section :

$$\left.\begin{aligned} c_2 + 2\mu c_3 = 0 ; \\ c_1 - c_3\left(\mu^2 + \nu^2\right) = 0 . \end{aligned}\right\} \quad (10)$$

where μ and ν - correspondingly real and imaginary parts of the roots.

In the parameters space c_1, c_2, c_3 sections look like straight lines, coming out from the origin of the coordinates (Fig. 6).

Inside the sections under investigation variables x_i, x_{i+1} (Fig. 4) naturally attenuate at $t > 0$, since real parts of the roots λ_i and λ_{i+1} are always negative.

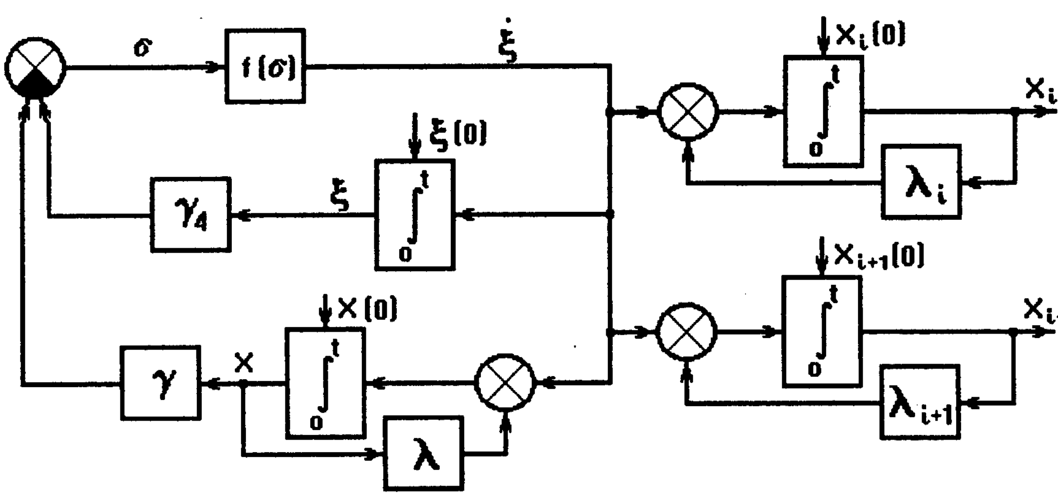


Fig. 4 Structural mathematical model of the transformed system into the section

The nature of changes of σ - variable (or ξ and x - variables) is defined by the second-order closed system dynamical behaviour.

$$\ddot{\sigma} + g_1\dot{\sigma} = -d_1\dot{f}(\sigma) - d_2 f(\sigma), \quad (11)$$

when : $g_1 = -\lambda;\ d_1 = \gamma + \gamma_4;\ d_2 = -\lambda\gamma_4$.

For all that, if the breaking of the second-order system's phase space is known, it is possible to gain the breaking of the fourth-order initial system's phase space, being considered into the sections.
The dynamical system (11) was investigated by the point transfer method in work [5]. This investigation resulted in breaking of its coefficients space into 10 areas with qualitatively different dynamics (Fig. 5).

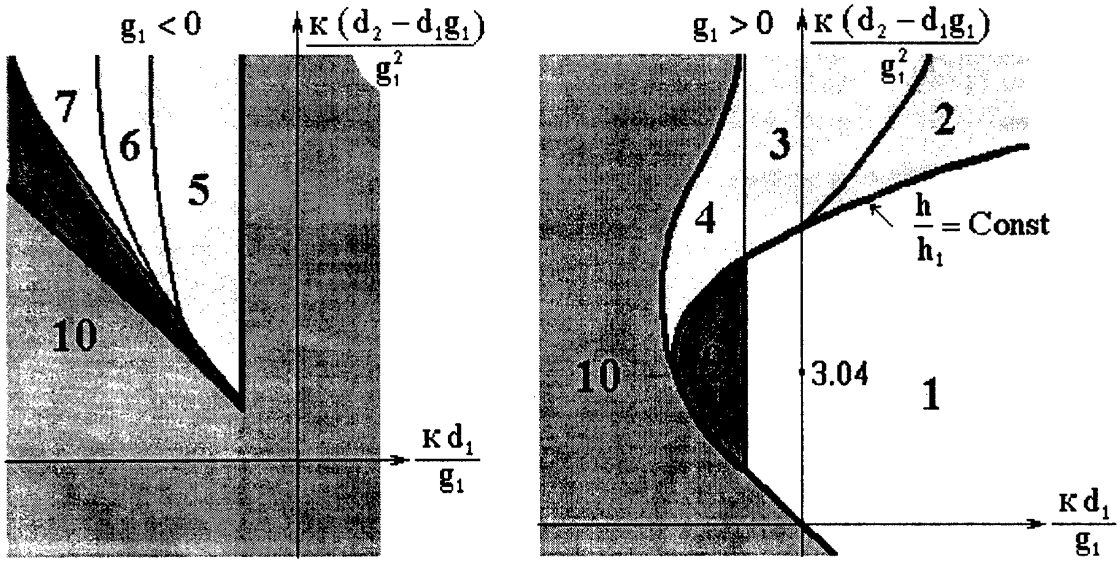


Fig. 5 Structure of breaking down the space of coefficients

Area 1 - the system is globally stable ; areas 2, 3, 5, 6, 7 - the system makes persistent oscillations (there's a single stable limit cycle in the phase portrait of the system) ; area 4 - the system either makes persistent oscillations or remains unstable (there are two limit cycles, it being known that the inside cycle is stable, while the outside one is unstable); area 8 - the system's behaviour is similar to strange attractor ; area 9 - the system is conditionally stable (there's a single unstable limit cycle in the phase portrait) ; area 10 - the system is unstable. The persistent oscillations of areas 2 - 7 qualitatively differ in their nature and (or) conditions of their appearance.
Bifurcation surfaces of all these areasare defined precisely. For global stability of the system the execution of the following conditions is necessary and sufficient :

$$\mathbf{A > 1 \;\; \& \;\; D < A, \text{ or}}$$

$$\mathbf{A > 1, \;\; A < D + 1 \;\; \& \;\; (A + 1 - D)Q > (D - 1)\exp\left(-\frac{A}{A-1}\ln\frac{D-1}{D-A}\right), \text{ or}}$$

$$\mathbf{A_1 > 0, \;\; D \le 1 \;\; \& \;\; (2 - D)Q > \sqrt{(1-D)^2 + A_1^2}\,\exp\left[-\frac{1}{A_1}\left(\mathrm{arctg}\frac{1-D}{A_1} + \frac{\pi}{2}\right)\right], \text{ or}}$$

$$\mathbf{A_1 > 0, \;\; 1 < D < 2 \;\; \& \;\; (2 - D)Q > \sqrt{(D-1)^2 + A_1^2}\,\exp\left(-\frac{1}{A_1}\mathrm{arctg}\frac{A_1}{D' - 1}\right),} \qquad (12)$$

where the parameters **A, $\mathbf{A_1}$, D** are defined as $\mathbf{A = \alpha_1/\alpha;\; A_1 = -\beta/\alpha;\; D = -d_1K/\alpha}$, α and β - correspondingly the real and imaginary parts of the complex conjugate roots in the characteristic equation $\mathbf{p^2 + (g_1 + d_1K)p + d_2K = 0}$, α and α_1 ($\alpha > \alpha_1$) - are real roots of this equation. The parameter $\mathbf{Q = (h + h_1)/(h - h_1)}$ - characterizes the given nonlinearity.

The results of this investigation being extended over the sections and their neighborhood, it is not difficult « to build » the whole bifurcation surface in the space we are interested in, with the help of other methods(the computer simulation method, for example).

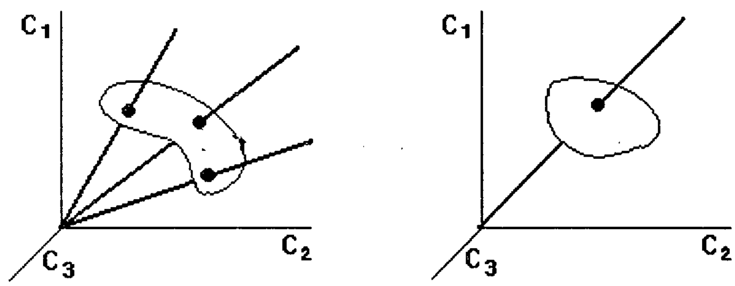


Fig. 6 The bifurcation surface in the parameters space

Three (or even one) plausible points of this surface obtained in advance, make the use of other methods much more efficient and reliable.

Conclusions. The author is placing hopes that the problem under discussing may convince someone interested that this method is a handy way of investigation into nonlinear high-order control systems.

References

1. Nelepin, R.A.(1964) About the research into nonlinear automatic systems by the sections into parameter space method. Academy of Sciences of the USSR, Engineering cybernetics, № 6.

2. Nelepin, R.A.(1967) Exact analytical methods in nonlinear automatic systems theory. Shipbuilding Press. Leningrad.

3. Lurie, A.I. (1951) Some nonlinear problems in automatic control theory, State Technical & Science Press. Moscow - Leningrad.

4. Nelepin, R.A., Kamachkin A.M., Turkin I.I., Shamberov V.N. (1990). In R.A.Nelepin (ed .), Algorithmical synthesis of nonlinear control systems'. Leningrad University Press, Leningrad.

5. Shamberov, V.N. (1987) The research of the coefficients (parameters) space of nonlinear dynamical system', Proc. Analysis & Synthesis of Control Systems, N10, Leningrad University Press, Leningrad, pp.162-8.

OPTIMAL CHOICE OF OBSERVED VALUES FOR C^2 CLASS FUNCTIONS

Valentina.V. Tchashnikova
St.Peterburg University
Department of Applied Mathematics
Universitetskaja naberegnaja 7/9
St.Peterburg , Russia

Abstract. Conditions for the optimal basis of observed values of nonlinear functions from class C^2 are considered .

Keywords. Approximation, basis, gradient, linear, observing, optimal, remainder .

We shall deal with R^m - Euclid space and $\{ f_i \}$, i=1,... n - a set of scalar functions, defined at a neighbourhood of a point $x_0 \in R^m$. We start from assumption that :

1) f_i - are C^2 class functions ;
2) $f_i(x_0) = 0$ for i=1,...n ;
3) $n > m$;
4) every set $\{ f_{ik}(x_0) \}$ k=1,...m defines a function $F : X \to Y$, $Y \in R^m$;

and there is at least one function F(x) on a set $K = \{ y ; |y_{ik}| \leq 1, k=1,..m\}$ for which the reverse function $F^{-1}(y)$ is defined .Also region $\gamma = F^{-1}(K)$ and class $\Phi = \{F(x) \mid \exists F^{-1}(y)\}$are defined.

Definition 1 . Let us call a function $F^*(x)$ the optimal one when

$$\inf_{F \in \Phi} \sup_{y \in k} |c(x_0 - F^{-1}(y))| ,$$ where c - is a certain vector .

The corresponding collection of scalar functions $\{ f_{ik} \}$ k=1, . . . m , will be called an optimal observations basis.

In a plain case, when f_i - are linear, the problem of optimal basis finding is reduced to the linear programing problem [1] .

Suppose $\hat{x} \in X$ is a certain fixed vector , that $\|\hat{x} - x_0\| \leq \delta$, $\delta = const$.

Every function F(x) corresponds to its own matrix $A(\hat{x})$ of gradient values of functions calculated in the point $\hat{x}$.

Let us consider the set A of these matrixes .

Definition 2 . We say, that $F^*(x)$ is linear-approximatly-optimal, if the function $G^*(x) = A^*(\hat{x})(x - x_0)$ is optimal for a corresponding matrix A^* according to definition1.

Now consider conditions with which the optimum of basis according to definition 2 follows from the optimum of basis according to definition 1. It needs corrections conditioned by:

1) remainder of Taylor's series for F ,
2) using the point $\hat{x}$ instead of x_0 in calculating of gradients .

The Taylor's series of function $f_{ik}(x)$ at the point (x_0) taking account $f_{ik}(x_0)=0$ will be

$$f_{ik}(x) = A_{ik} \Delta x + 1/2\, \Delta x^T B_{ik} \Delta x , \qquad \Delta x = x - x_0 ,\ k=1, \ldots m . \qquad (1)$$

There is A_{ik} - gradient of function f_{ik} , calculated in x_0 . Without loss of generality suppose , that $A_{ik} \Delta x \geq 0$. The second addend in (1) is remainder in Lagrange's form , namely quadratic form , where B_{ik} is a matrix of second partial derivatives of functions

f_{ik} , calculated in x_Θ , such as $\| x_0 - x_\Theta \| \le \Delta x$.
In every equations of system (1) vector A_{ik} is expressed as

$$A_{ik} = \hat{A}_{ik} + \Delta A_{ik} \quad (2)$$

where $\hat{A}_{ik}$ s gradient of f_{ik} , calculated in $\hat{x}$.
Substituting (2) into (1) we'll have

$$f_{ik}(x) - y_{0ik} = \hat{A}_{ik}\Delta x + A_{ik}\Delta x + 1/2\ \Delta x^T B_{ik}\Delta x \quad (3)$$

Let Θ_{ikj} j-column of matrix B_{ik} , then $| 1/2\ \Delta x^T \Theta_{ikj} | \le \delta/2 \| \Theta_{ikj} \|$.
Every component of vector ΔA_{ik} can be expressed in Lagrange's form as :

$$\Theta^T_{ikj}(x_0 - \hat{x}) = (\partial^2 f_{ik}/\partial x_i \partial x_j |_{x\theta}, .., \partial^2 f_{ik}/\partial x_j \partial x_m |_{x\theta})^T (x_0 - \hat{x}),$$

$$\Theta^T_{ikj}(x_0 - \hat{x}) \le \| \Theta^T_{ikj} \| * \delta .$$

Denote $\Delta A_{ik}\Delta x + 1/2\Delta x^T B_{ik}\Delta x = \Delta_{ik} * \Delta x$. Every component of Δ_{ikj} is restricted by modulus :

$$| \Delta_{ikj} | \le \delta * (\| \Theta^T_{ikj} \| + 1/2 \| \Theta^T_{ikj} \|) = M_{ikj} ,\ k=1,..m ,\ j=1,..m \quad (4)$$

Thus the inequality is fairly :

$$(\hat{A}_{ik} - M_{ik}\alpha^0)\Delta x \le f_{ik}(x) - y^0_{ik} \le (A_{ik} + M_{ik}\alpha^0)\Delta x \quad (5)$$

where $\alpha^0 = \text{diag}(\text{sign}\ \Delta x_1 ,.. \text{sign}\ \Delta x_m)$.
Denote $H(\alpha) = M\alpha$, where M is a matrix, composed from vektors M_{ik} , k=1, ...m ; $\alpha \in \Omega$;

$$\Omega = \{ \alpha = \text{diag}(\text{sign}\ \alpha_1 ,.. \text{sign}\ \alpha_m),\quad \alpha_j \in [-1,1], j=1,..m\}.$$

Consider a set of matrixes $\Gamma = \{ \hat{A} + H(\alpha) \}$, $\alpha \in \Omega$.

Lemma. Let A - matrix of gradients of functions { $f_{ik}(x)$ k=1,... m , culculated in $\hat{x}$, let the restrictions (4) are found , and let rang $(\hat{A} + H(\alpha))$=m for whole set Ω .
Then for everyone $y \in K$, that y=F(x) there is matrix $\tilde{A} \in \Gamma$, such as
$\tilde{A}\ \Delta x = \tilde{y}$, where $\Delta x = x - x_0$.
This fact is derived from inequality (5) and from properties of continuity of linear function and continuity of reverse function for nonsingular matrixes .

Let λ - minimal singular value of marix A, ν - maximal one of matrixes $H(\alpha)$.
Suppose that $\lambda - \nu > 0$. (6)
Let restriction (4) are found for all functions of set { f_i }, i=1,... n , and assume the first m of these functions form the optimal basis of linear approach .
Remark : henceforth in order to avoid unnecessary denotations instead of $\hat{A}$, M , H we'll take, dependig of contents , either matrixes of optimal basis , or extended matrixes $n \times m$.
We'll call the first rows of matrix $\tilde{A} \in \Gamma$ - basic vectors , others - unbasic ones .
Let us find upper bounds for coefficients of decomposition of vector C and all unbasic vector from Γ on basic vectors . Denote :

$\hat{C}$ - vector of coefficients of decomposition C on rows of matrix $\hat{A}$ (on basic gradients) ,
$\hat{b}_j$ - vector of coefficients of decomposition of j - unbasic gradient on basic gradients ,
$\hat{C} + \Delta C$ - vector of coefficients of decomposition vector C on basic vectors from Γ ,
$\hat{b}_j + \Delta b_j$ - vector of coefficients of decomposition of j - unbasic vector from Γ on basic vectors from Γ , j=m+1, ... n ,

$$\beta_j = \max_i | \hat{b}_{ij} | ,\quad \omega_j = \max_i | h_{ij} | .$$

Thus it is fair : $(\hat{C} + \Delta C)(\hat{A} + \Delta H) = C$; $\Delta C = - \hat{C} M\alpha(\hat{A} + H)^{-1}$.
Without loss of generality , we'll take , that $\| C \| = 1$.
Every component of ΔC is restricted by modulus :

$$\Delta C_i \le \nu(\lambda - \nu)^{-1} \quad (7)$$

Further consider $(\hat{b}_j + \Delta b_j)(\hat{A} + H) = \hat{A}_j + H_j$, j=m+1, ... n.

$$\Delta b_j = (H_j - \hat{b}_j H)(\hat{A} + H)^{-1}, \quad \sum_{i=1}^{m} \Delta b_{ij} \le \nu * m(\omega_j + \nu\beta_j)/(\lambda - \nu) \qquad (8)$$

Theorem. If the first m functions of set $\{f_i\}$, i=1,... n, form vector-function F(x), which is optimal by linear approach and the restriction (4) are found and conditions (6) are valid and for all of i=1,..m , j=m+1, . . . n inequalities

$$\hat{C}_i > \nu(\lambda - \nu)^{-1}, \qquad (9)$$

$$\Sigma\, \hat{b}_{ji} < 1 - \nu * m(\omega_j + \nu\beta_j)/(\lambda - \nu) \qquad (10)$$

are fair, then this vector-function is optimal according to definition 1.

Proof. The condition (6) is more powerful than demand of lemma, hence for any $x \in \gamma$ there is $\tilde{A} \in \Gamma$, that $\tilde{A}(x - x_0) = F(x)$.

We should only proof that for any extended matrix $\tilde{A} \in \Gamma$ the first m rows form optimal basis.

If we apply geometric interpretation of simplex-method, then vector C or its continuation crosses the right face of convex casing of vectors $\hat{A}_k$, k=1,. . .n, which is stretched on vectors $\hat{A}_i$, i=1,. . .m, corresponding to optimal basis. In a process of parameters modification the optimal basis can be changed in two cases :

1) vector C stopped to cross the pointed face,
2) any of basic vectors stopped to belong to convex casing.

In the first case there is k, $1 \le k \le m$, such as $\hat{C}_k + \Delta C_k < 0$, but it can't be happened because of inequalities (7) and (9) are fair.

In the second case there is k, $1 \le k \le m$, such as $\hat{C}_k + \Delta C_k < 0$, but it can't be happened because of inequalities (8) and (10) are fair.

Therefore the first m rows form optimal basis for all matrixes from set Γ, and hence, on the basis of lemma, the first m functions from set $\{f_i\}$, i=1,... n, form optimal basis according to definition 1.

Conclusion.

This theorem defines a sufficient conditions of linear programming method application for optimal choice of basis of observed values for nonlinear C^2 class functions.

Reference

1. Elyasberg P.E. Determination of motion by measuring's results (Russian)
   Moscow : Nauka, 1976.

CONSTRUCTIVE METHOD OF NONLINEAR ANALYSIS OF CONTROLLED SYSTEMS*

Stanislav Vassilyev
Irkutsk Computing Centre
of Siberian Branch of Russian Academy of Science
134, Lermontov street, Irkutsk, 664033, RU
e-mail: vas@icc.ru

Abstract. A new method of analysis of nonlinear controlled systems with discrete time has been proposed. It is based on some comparison theorems in terms of homomorphisms of models. Constructive algorithm for searching for these homomorphic models which are easier for analysis has been proposed. In the capacity of dynamic properties which are studied in the mathematical models the complicated properties of reachability under phase constraints have been considered. Some illustrative example in the form of control problem for a neural network has been described.

Keywords. Nonlinear analysis, controlled systems, difference equations, comparison method.

Introduction. The main subject of this paper is to propose constructive method of nonlinear analysis and synthesis of controlled systems with discrete time under phase constraints and perturbations. Besides, we would like to show that the basic mathematical theorems for this method can be obtained automatically on the basis of SG-method of automatic reasoning which has been described in [1, 2].

These basic theorems will be obtained in terms of homomorphic mapping of mathematical models and allow to reduce the complexity of studied models. These mapping play the role of vector Lyapunov functions in stability theory. We shall propose efficient algorithms of their searching for.

Finally, we illustrate our results by a discrete controlled system in the form of neural network.

Basic theorems. Let us consider the synthesis of a comparison theorem on the property of reachability of the abstract dynamical system $F : T \times X \to X$ w.r.t. the sets $X^0 \subseteq X$, $X^1 \subseteq X$, where T means a set of time moments t, X signifies a phase space of system F.

Let define the reachability as the dynamical property:

$$P = \forall x_0(x_0 \in X^0 \to \exists t(t \in T \& \exists x(x = F(t, x_0) \& x \in X^1))). \tag{1}$$

Given comparison system $G : T \times Y \to Y, Y^0 \subseteq Y, Y^1 \subseteq Y$ and auxiliary function $v : X \to Y$, where Y is partially ordered by a relation $\leq$, it is naturally to ask: under what kind of conditions W from the facts that 1) G has the property $R = \forall y_0(y_0 \in Y^0 \to \exists t_1(t_1 \in T \& \exists y(y = G(t_1, y_0) \& y \in Y^1)))$ and 2) the majorizing condition holds

$$M = \forall x'_0(x'_0 \in X^0 \to \forall y'_0(y'_0 \in Y^0 \& v(x'_0) \leq y'_0 \to$$
$$\forall t'(t' \in T \to \forall x'(x' = F(t', x'_0) \to \forall y'(y' = G(t', y'_0) \to v(x') \leq y'))))),$$

the origin system F will have the property P?

*This paper is immediate continuation of the papers [1, 2]. This research is conducted within the research grants No 96-01-00144, 95-07-20058 from RFBR.

Let apply SG-process [2] for synthesizing nontrivial W. It is possible to check that SG-process for

$$\forall{:}\mathbf{T}\,\exists{:}\mathbf{T}\{M, R, \forall{:}\mathbf{T}\,\exists x_0{:}x_0 \in X^0\,\forall tx{:}t \in T,\ x = F(t, x_0),\ x \in X^1\ \exists : \mathbf{F}\}$$

is infinite (we understand here the formulas M, R as $\forall$-formulas). After at least of 3-d application of α with intermediate ω's and on the basis of β, instantiation and other transformation rules [2, 3] we can obtain the formulas which will be equivalent to the following formulas:

$$\forall x_0(x_0 \in X^0 \to \exists y_0''(y_0'' \in Y^0 \& v(x_0) \leq y_0'')), \tag{2}$$

$$\forall x_0(x_0 \in X^0 \to \forall t''(t'' \in T \to \exists x''(x'' = F(t'', x_0)))), \tag{3}$$

$$\forall t''(t'' \in T \to \forall y''(y'' \in Y^1 \to \forall x''(x'' \in X \& v(x'') \leq y'' \to x'' \in X^1))). \tag{4}$$

It is some estimation of v w.r.t. the sets X^1, Y^1. The meaning of (2), (3): (2) means also estimation of v w.r.t. X^0, Y^0; (3) requires a prolongability of trajectories $F(t, x_0)$.

Thus, D =(2) & (3) & (4) is desirable solution. This gives in system dynamics the new comparison theorem $D \to (M\&R \to P)$.

Let us complicate our mathematical model and studied dynamic property. Let this model be described in the form of difference equation

$$x(t+1) = F(x(t), u(t), p(t)), \tag{5}$$

where $t \in T^\tau = \{1, 2, \ldots, \tau\}, \tau \in N = \{1, 2, \ldots,\}$, or $t \in N = T^\infty$, $\mathrm{x} = x(t) \in X$, $\mathrm{u} = u(t) \in U$, $\mathrm{p} = p(t) \in P$, X, U, P are abstract sets, $F : X \times U \times P \to X$.

By $\mathcal{U}$ and $\mathcal{P}$ we shall denote some classes of admissible controls $u : T^{\tau_u} \to U$ and perturbations $p : T^{\tau_u} \to P$. Pair (u, p) forms a multi-input sequence $(u, p)_t = (u(t), p(t)), t \in T(u, p)$, i.e. $(u, p) : T(u, p) \to U \times P$, where $T(u, p) = domu \cap domp$. We will assume that $T(u, p) = T^\tau$ for all u, p.

For every $\mathrm{x}_0 \in X$ and every input sequence $(u, p) : T^\tau \to U \times P, u \in \mathcal{U}, p \in \mathcal{P}$, the next-state function F of automaton (5) determines the corresponding process $x(\cdot, \mathrm{x}_0, u, p)$: $T^\tau \to X$. In initial time moment 1, $x(1, \mathrm{x}_0, u, p) = \mathrm{x}_0$.

For example we consider the following class of dynamic properties of the system (5). Their definition $\mathcal{S}$ are representable as follows

$$\mathcal{S} = \{K_1 \mathrm{x}_0 \in X^0,\ \ K_2 u \in \mathcal{U},\ \ K_3 p \in \mathcal{P},\ \exists x(\cdot, \mathrm{x}_0, u, p),\ \ K_4 t_1 \in N(\alpha),$$
$$K_5 t \in N(\tau, t_1),\ K_6 \mathrm{x}_* \in X^*\}\ [(x(t) = \mathrm{x}_*)\ \ l\ \ (K_7 t_2 \in N(\tau, t_1, t)\ x(t_2) \in X^1)] \tag{6}$$

where $X^0 \subseteq X, X^* \subset X, X^1 \subseteq X$, values of symbols K_i of all quantifiers belong to the set $\{\forall, \exists\}$ and can be chosen independently; an order of quantifiers in the brackets $\{,\}$ is arbitrary fixed, but the quantifier $\exists x(\cdot, \mathrm{x}_0, u, p)$ has to occur in the definition after the quantifiers with variables x_0, u, p, as well as the quantifier $K_5 t$ after $K_4 t_1$; $l \in \{\&, \vee\}, \tau \in N$ or $\tau = \infty, N(\alpha), N(\tau, t_1), N(\tau, t_1, t)$ are some subsets of T^τ which are generally speaking dependent on either α or τ, t_1 or τ, t_1, t respectively, α is some parameter, $\alpha \in N$. Accounting only the variety of the quantifier order under its restrictions above and the variety of values of symbols K_i, l, we obtain that formally this class consists of 135 168 definitions (some of them are logically equivalent), e.g., the property

$$\forall \mathrm{x}_0 \in X^0\ \ \exists u \in \mathcal{U}\ \ \forall p \in \mathcal{P}\ \ \exists x(\cdot, \mathrm{x}_0, u, p)\ \ \exists t_1 \in N\ \ \forall t \in N : t \geq t_1$$
$$\exists \mathrm{x}_* \in X^*\ (x(t) = \mathrm{x}_*\ \ \&\ \ \forall t_2 \in T^{t_1}\ \ x(t_2) \in X^1) \tag{7}$$

which is said to be *strong* (X^0, X^*) – *reachability with right-side invariance of goal set* X^* *under state restrictions* X^1 *and perturbations.*

Let consider second automaton

$$y(t+1) = F'(y(t), u'(t), p'(t)), \quad (8)$$

$t \in T^\tau, \mathrm{y} = y(t) \in Y, \mathrm{u}' = u'(t) \in U', \mathrm{p}' = p'(t) \in P', F' : Y \times U' \times P' \to Y, u' \in \mathcal{U}', p' \in \mathcal{P}'$ (we assume that output function is identical and $y(1, \mathrm{y}_0, u', p') = \mathrm{y}_0$). In contrast to common comparison method [4, 5] and as further developing of [6, 7] we will use a mapping $V = (v, s, r), v : X \to Y,\ s : \mathcal{U} \to \mathcal{U}',\ r : \mathcal{P} \to \mathcal{P}'$, which is a *homomorphism* of the automaton F in the automaton F', i.e. for all $t \in N, \mathrm{x} \in X, u \in \mathcal{U}, p \in \mathcal{P}$

$$v(F(\mathrm{x}, u(t), p(t))) = F'(v(\mathrm{x}), s(u)(t), r(p)(t)). \quad (9)$$

In the automaton F' we will consider the class of dynamic properties $\mathcal{S}'$ analogous to (6)

$$\mathcal{S}' = \{K_1 \mathrm{y}_0 \in Y^0,\ \ K_2 u' \in \mathcal{U}',\ \ K_3 p' \in \mathcal{P}',\ \exists y(\cdot, \mathrm{y}_0, u', p'),\ K_4 t_1' \in N(\alpha),$$

$$K_5 t' \in N(\tau, t_1'),\ K_6 \mathrm{y}_* \in Y^*\}\ [(y(t') = \mathrm{y}_*)\ \ l\ \ (K_7 t_2' \in N(\tau, t_1', t')\ \ y(t_2') \in Y^1)]. \quad (10)$$

We will assume that the values of the corresponding quantifiers from $\mathcal{S}, \mathcal{S}'$ are identical, e.g., $K_1 = \forall \in \mathcal{S}$ iff $K_1 = \forall \in \mathcal{S}'$, etc. Particularly, the class (10) containes the property, analogous to (7) which is denoted by (7′).

We denote by $\bar{a}$, where a is a type-quantifier, the result of replacing symbol $\forall$ or $\exists$ by dial one. Besides that $(ab)^k = ab$ if $k = \forall$, otherwise $(ab)^k = \bar{b}\bar{a}$. Let $\bar{A}$ be a set $X \backslash A$. By representing the expression (9)&(10) $\to$ (6) in L [1] and applying of SG-process [2] we will obtain an infinite process [3]. After application of α with intermediate ω's, on the basis of β and transformation rules [2, 3, 7] we obtain some solutions of logical equations W&(9)&(10) $\to$ (6), and after interpretation of W the following theorem.

Theorem 1. The automaton (5) possesses a property $\mathcal{S}$ of the class (6) if there exist an automaton (8) with the analogous property $\mathcal{S}'$ and a homomorphism V of F in F' such that:

1) $(\forall \mathrm{x}_0 \in X^0\ \ \exists \mathrm{y}_0 \in Y^0)^{K_1}\ \ v(\mathrm{x}_0) = \mathrm{y}_0$, 2)$(\forall u \in \mathcal{U}\ \ \exists u' \in \mathcal{U}')^{K_2}\ s(u) = u'$,

3) $(\forall p \in \mathcal{P}\ \ \exists p' \in \mathcal{P}')^{K_3}\ r(p) = p'$, 4)$v(\bar{X}^1) \subseteq \bar{Y}^1$,

5) if $X^* = \{\mathrm{x}_*\} - singl$, then $X_c^* = \{\mathrm{y}_*\} - singl$, $\ v^{-1}(\mathrm{y}_*) = \{\mathrm{x}_*\}$, otherwise:

5.1) if $K_6 = \exists$, then 5.1.1) if the scope of K_6 does not contain $\forall \mathrm{x}_0, \forall u, \forall p, \forall t_1, \forall t$ ($\forall$-condition), then $v^{-1}(Y^*) \subseteq X^*$, otherwise $v(X^*) = Y^*$ and v^{-1} is one-to-one mapping on Y^*,

5.2) if $K_6 = \forall$, then 5.2.1) if the scope of K_6 does not contain $\exists \mathrm{x}_0, \exists u, \exists p, \exists t_1, \exists t$ ($\exists$-condition), then $v^{-1}(Y^*) \neq \emptyset \to Y^* \neq singl$, otherwise $v(X^*) \subseteq Y^*$ and v^{-1} is one-to-one mapping on $v(X^*)$.

Corollary 1. If there exists a homomorphism V of the F in F' and 1)$v(X^0) \subseteq Y^0$, 2) $s(\mathcal{U}) = \mathcal{U}'$, 3) $r(\mathcal{P}) \subseteq \mathcal{P}'$, 4) $v(\bar{X}^1) \subseteq \bar{Y}^1$, 5) $v^{-1}(Y^*) \subseteq X^*$, then (7′) $\to$ (7).

Consider a partition Π of the state set X of automaton, where the Π-equivalence class containing state x is denoted $[\mathrm{x}]_\Pi$. In literature, the subsets $[\mathrm{x}]_\Pi$ are called also the blocks. Then the partition Π is said to have the *substitution property* (SP) if, for every pair of Π-equivalent states x_1 and x_2, the successor of x_1 is Π-equivalent to the successor of x_2 under each input (u, p). A partition that possesses the SP is called an *SP-partition* (two trivial SP-partitions are contained in every automaton [8]).

A set $A \subset X$ is said to be *representable* (or *has representation) by a (nontrivial) partition* Π if A is equal to union of some blocks of Π. If all the blocks of the union are single-point sets, we say that the representation *is constructed by single-point blocks* of Π. If the union consists of more than one block, we say that the representation *is constructed by many blocks* of Π. A set A is said to be *representable* if there exists a partition on X by which A is representable.

If $A \subseteq X$, Π is a SP-partition on X, we will denote $\Pi(A) = \{y : \exists x \in A, y = [x]_\Pi\}$. Let define the mapping $v : X \to Y$, $Y = \Pi(X), v(x) = [x]_\Pi$, and comparison automaton $F' = \Pi(F)$ with state set $Y = \Pi(X), U' = U, P' = P$ and next state function $F'(v(x), u, p) = v(F(x, u, p))$. It is easy to see that the v is homomorphism of F in F' and the following is true.

Corollary 2. Let for the automaton F there exists a nontrivial SP-partition Π such that the sets X^*, X^1 are representable by Π and when $X^* \neq singl$ for considered dynamic property $\mathcal{S}$ of the class (6) the following conditions are satisfied:

1) if $\forall$-, $\exists$-condition are not valid (when $K_6 = \exists$ or $K_6 = \forall$, resp.), then the representation of X^* is constructed by single-point blocks of Π;

2) if in the case $K_6 = \forall$ the $\exists$-condition is valid, then the representation of X^* is constructed by many blocks of Π.

Then the F possesses the property $\mathcal{S}$ iff the automaton $\Pi(F)$ possesses the analogous property $\mathcal{S}_c$ w.r.t. the sets $Y^0 = \Pi(X^0), Y^* = \Pi(X^*), Y^1 = \Pi(X^1), \mathcal{U}' = \mathcal{U}, \mathcal{P}' = \mathcal{P}$.

Corollary 3. If for the automaton F there exists a nontrivial SP-partition such that the sets X^*, X^1 from (7) are representable by Π, then (7) is valid iff the automaton $\Pi(F)$ possesses the property (7′), where $Y^0 = \Pi(X^0), Y^* = \Pi(X^*), Y^1 = \Pi(X^1), \mathcal{U}' = \mathcal{U}, \mathcal{P}' = \mathcal{P}$.

For example, let consider the property (7) of the automaton in $X = E^3, E = \{0, 1\}$,

$$\begin{cases} x_1(t+1) = \bar{x}_1 x_2 u \vee x_1 \bar{x}_2 \bar{x}_3 u \vee \bar{x}_1 x_3 \bar{p} u \vee x_1 \bar{x}_2 p u, \\ x_2(t+1) = \bar{x}_2 p u, \\ x_3(t+1) = \bar{x}_1 \bar{x}_3 \bar{u} \vee x_1 x_2 x_3 \vee \bar{x}_2 \bar{x}_3 \bar{u} \vee x_3 p u \vee \bar{x}_1 \bar{p} \bar{u} \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{p} \vee \bar{x}_1 \bar{x}_2 x_3 p \end{cases} \tag{11}$$

under arbitrary control inputs $u : N \to E$, and perturbations $p : N \to E$. There are two nontrivial SP-partitions in all: $\Pi_1 = \{\{1,2\},\{3,4\},\{5,6\},\{7,8\}\}$, $\Pi_2 = \{\{2,3,6,7\},\{1,4,5,8\}\}$, where by 1,2,..., 8 we code the states 001, 010, 011, 100, 101, 110, 111, 000, resp. If $X^0 = \{3\}, X^* = \{1,2\}, X^1 = \{1,2,3,4,7,8\}$, then X^*, X^1 are representable by $\Pi_1, \Pi_1(X^0) = \{\{3,4\}\}, \Pi_1(X^*) = \{\{1,2\}\}, \Pi_1(X^1) = \{\{1,2\},\{3,4\},\{7,8\}\}$. Let code the blocks of Π_1 by 00, 01, 11, 10, resp. We obtain the homomorphic automaton F' in E^2

$$\begin{cases} y_1(t+1) = y_2(y_1 \vee p), \\ y_2(t+1) = \bar{y}_1(y_1 \vee u). \end{cases} \tag{12}$$

Let $Y^0 = \Pi_1(X^0) = \{01\}, Y^* = \Pi_1(X^*) = \{00\}, Y^1 = \Pi_1(X^1) = \{00, 01, 10\}$.

It easy to see that the automaton (12) possesses the property (7′) with the control $u(t) \equiv 0$, and consequently in the (11) the (7) is valid with the same control.

Finally, consider for the automaton (5) when X is finite an algorithm of varifying the properties $\mathcal{S}$ of the class (6) by generating homomorphisms $v(x) = [x]_\Pi$ and homomorphic automata $F' = \Pi(F)$, where Π are partitions representing X^*, X^1. In the part of generating the partitions Π, this algorithm modifies a known procedure for generating all SP-partitions [8]. Before giving the algorithm, the following definitions and remarks are first presented.

For partitions Π_1, Π_2 on X we write $\Pi_1 \leq \Pi_2$ iff $\forall B \in \Pi_1\ \forall (x_1, x_2) \in B^2\ \exists \mathcal{D} \in \Pi_2\ (x_1, x_2) \in \mathcal{D}^2$ and Π is said to be *smaller* than (or equal to) Π_2 [8]. If X^* is representable by some two partitions Π_1, Π_2, X^1 is representable by the same Π_1, Π_2 too, $\Pi_1 \leq \Pi_2, \Pi_1 \neq \Pi_2$, then the number of states of the automaton $\Pi_2(F)$ is less than that for $\Pi_1(F)$. From this point of view $\Pi_2(F)$ is more preferable for using in the capacity of homomorphic automaton F'.

The *sum* $\Pi_1 + \Pi_2$ is defined as partition on X obtained by merging those blocks of Π_1 and Π_2 which have at least one common state [8]. The set of all SP-partitions is closed under the "+" operation. If X^*, X^1 are representable by Π_1, Π_2, then they are representable by $\Pi_1 + \Pi_2$.

A pair of states is said to be *identified w.r.t.* X^*, X^1 when both states appear in the same block of a partition Π by which X^* and X^1 are representable.

The Algorithm.

1. For every pair of different states $(\mathrm{x}, z) \in \chi = ((X^* \times X^*) \cup (\bar{X}^* \times \bar{X}^*)) \cap ((X^1 \times X^1) \cup (\bar{X}^1 \times \bar{X}^1))$, compute the smallest SP-partition $\Pi_{\mathrm{x},z}$ which identifies the pair w.r.t. X^*, X^1. This is performed by first placing the states x, z in a block. Then for each input (u, p), find the next-states for x, z; these next-states must also be identified w.r.t. X^*, X^1 and the transitive law applied where possible. If some generated block of Π has to contain a pair from $\bar{\chi}$, then the generating partition for the pair (x, z) failes and the pair (x, z) is not considered further.

2. Find the maximal partitions among all possible sums of partitions $\Pi_{\mathrm{x},z}$ generated in step 1.

3. For the maximal nontrivial partitions Π construct corresponding automata $F' = \Pi(F)$.

4. Let for automaton (5), its property $\mathcal{S}$ of the class (6), and some automaton F' constructed in step 3 the conditions 1), 2) of Cor. 2 are valid. If in the F' there exists the analogous property $\mathcal{S}_c$, then the automaton (5) possesses the property $\mathcal{S}$; moreover, $\mathcal{S} \rightarrow \mathcal{S}_c$.

Remark. To reduce the number of sums that must be performed in step 2 it is possible to use some modification of algorithm for generating a lattice of SP-partitions [8].

Let consider the application of the Algorithm to the neural network model

$$x(t+1) = \pi(Ax(t) - b),\ x(t) \in E^5,$$

where $\pi(x) = 1$ if $x \geq 0$, $\pi(x) = 0$ if $x < 0$,

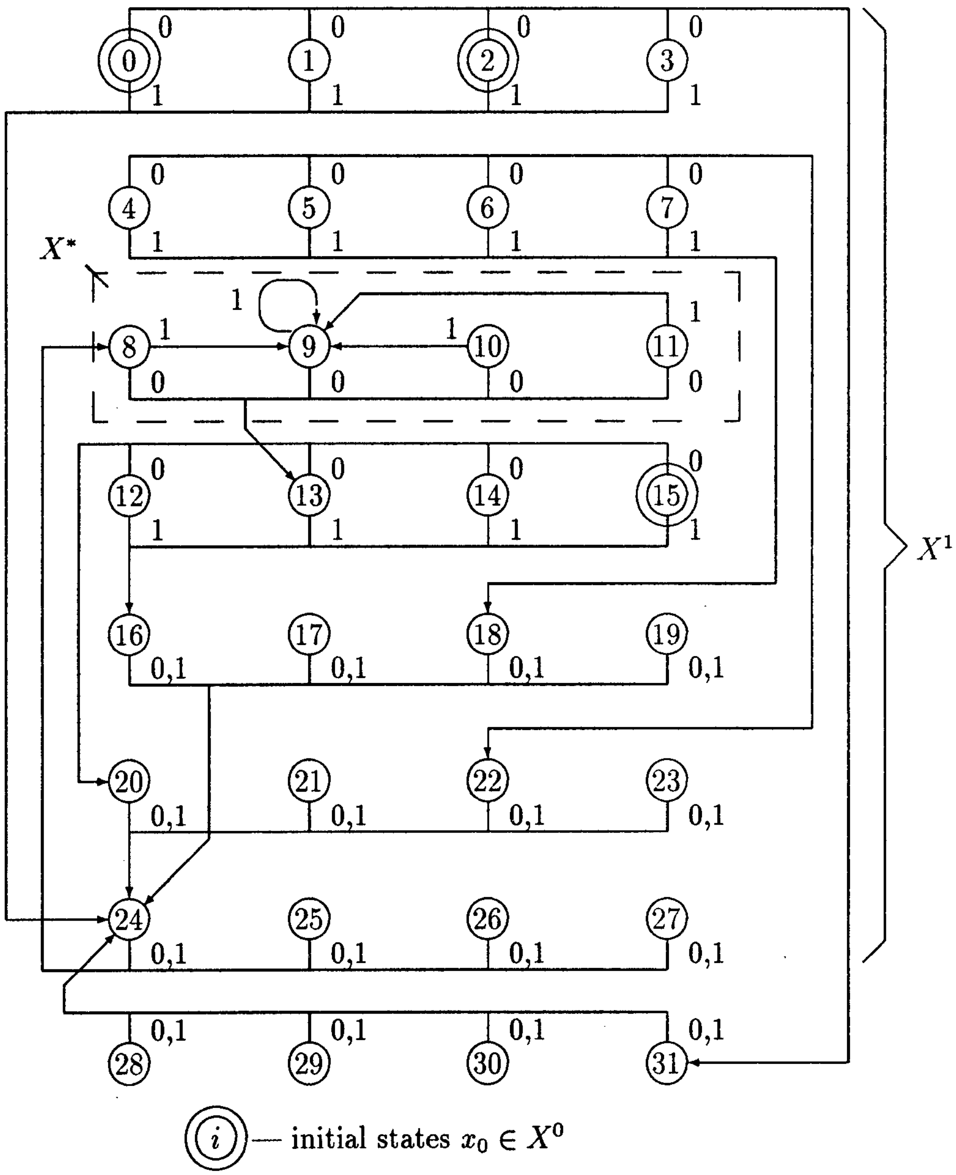
$$A = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \end{pmatrix},\quad b = (0, 0, \mathrm{u}(t), 0, 0),\quad \mathrm{u}(t) \in E.$$

The problem is to find $u(t)$ satisfying the property (7) with $X^0 = \{0, 2, 15\}, X^* = \{8 - 11\}, X^1 = \{0 - 27\}$ in coding of binary states $00000, 00001, \ldots, 11111$ by corresponding decimal numbers $0, 1, \ldots, 31$. Figure 1 illustrates the dynamics of this model. The answer is $u(t) \equiv 1$, because the algorithm generates the nontrivial SP-partition $\Pi = \{\{0-3\}, \{4-7, 12-15\}, \{8-11\}, \{16-23\}, \{24-27\}, \{28-31\}\}$ and automaton F' (with 6 states) which is described by Fig. 2 and possesses the property (7′) under control $u(t) \equiv 1$. It should be noted that the considered neural network does not belong to the class of symmetric neural networks which is mainly investigated in literature and from the view point of very narrow class of dynamic properties [9].

Analogously we can consider as controls not only the threshold vector b, but also the values a_{ij} of interaction between the neurons as well as more general case when b_i, a_{ij} are real numbers and some of them can be considered as perturbations and another ones as controls.

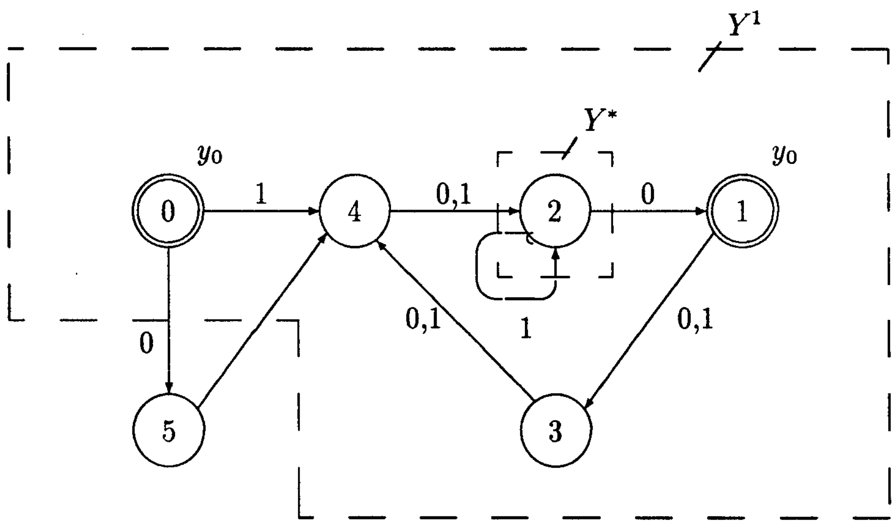
Our algorithm of synthesis the control for providing in neural networks arbitrary property of the class (6) can be easily expanded for other dynamic properties using SG-method and comparison automata, particularly with several initial state sets X_i^0, goal sets X_i^* and state restriction sets $X_i^1, i = 2, 3 \ldots$, that is important for many applications.

Conclusion. On the basis of regular method of hypothesizing from [1, 2] some basic marhematical theorems for reducing the complexity of analysis of properties of the class of



X^* — goal set,
X^1 — phase-restriction set.

Figure 1: Transition diagram of NN (synchronous iteration graph or next state function F of the sequential machine)



$v : X \to Y,\ v(F(x,u)) = F'(v(x),u);$
$v : (\{0-3\}) = \{0\},$
$v : (\{4-7\} \cup \{12-15\}) = \{1\},$
$v : (\{8-11\}) = \{2\},$
$v : (\{16-23\}) = \{3\},$
$v : (\{24-27\}) = \{4\},$
$v : (\{28-31\}) = \{5\},$

$Y^0 = v(X^0) = \{0,1\},$
$Y^* = v(X^*) = \{2\},$
$Y^1 = v(X^1) = \{0-4\},$
$u = u(t) \equiv 1$ — is desirable control.

Figure 2: Transition diagram of next-state function F' of homomorphic comparison automaton

reachability for difference equations with controls and perturbations have been obtained. These theorems in terms of homomorphisms have been formulated. Some constructive algorithm of searching for these homomorphism and homomorphic models has been proposed. After solving the problem of searching for a desirable control this control provides the desirable dynamic property in the origin model.

References

1. Vassilyev, S. N. (1997) Theory and application of intelligent control. In the Proceedings of the Conference.
2. Vassilyev, S. N. (1997) New method of reasoning: deduction + hypothesizing. In the Proceedings of the Conference.
3. Vassilyev, S. N. (1996) Modelling logical derivation and hypothesis generation. Proc. Symp. "Modelling, Analysis and Simulation" of CESA'96 IMACS Multiconference, 1, Lille, pp. 148-153.
4. Matrosov, V. M. (1973) Comparison method in system's dynamics. Equat. Different. et functionnelles non linear, Herman, Paris, pp. 407-445.
5. Lakshmikantham, S., V. M. Matrosov, S. Sivasundaram (1991) Vector Lyapunov functions and stability analysis of nonlinear systems. Kluwer Acad. Publ.
6. Vassilyev, S. N. (1981, 1982) Comparison method in system analysis, I - IV. Diff. uravnenia (Diff. Equations), v.17, 9, pp. 1562-1573, v.17, 11, pp. 1945-1954 (1981), v.18, 2, pp. 197-205, v.18, 6, pp. 938-947 (1982).
7. Vassilyev, S. N. (1990) Machine synthesis of mathematical theorems. J. Logic Programming, 9, $N2\&3$, pp. 235-266.
8. Lee, S. C. (1978) Modern switching theory and digital design. Prentice-hall, Inc.
9. Goles, E., S. Martinez (1990) Neural and Automata Networks. Dynamic Behavior and Applications, Kluwer Acad. Press.