ADA041970

A Multistage Model of Network Usage*

Bennet P. Lientz
Graduate School of Management
University of California, Los Angeles

ABSTRACT

The evolution of network service usage can be considered as being in three stages, constituting a usage life cycle. The first stage (inception) is that of initial low level usage. The second (expansion) is characterized by rapid growth in usage and is often unstable and uncontrolled. The third stage (stability) is a reaction to the second stage and imposes order on usage. Decision making, data collection, and measurement considerations are examined for this framework.

*This work was partially supported by the Information Systems Program, Office of Naval Research under contract N00014-75-C-0266 under project number NR 049-345.

1. Introduction

The usage of computer network and commercial time-sharing systems has grown rapidly in recent years (see for example [2], [5], [17]). Coincident with this growth has been a rise in concerns expressed about selection and usage of applications on network services. Security, privacy, and social problems have been cited by Enslow [8], [9]. Economic issues for the user's point of view has begun to be examined (e.g., Lientz [13]). A number of papers have addressed security issues (see for examples, Lientz and Weiss [10], [15]). Selection and evaluation has been considered in Lientz and Arnold [14]. Very little research has been directed toward the patterns of user behavior with respect to network usage. It has been cited, however, as a major area of inquiry [1], [5]).

The purpose here is to analyze usage of network services from the viewpoint of a user community within an organization. In section 2 a methodology and framework is developed for the evolution of network usage within three stages. This is similar to an approach by Nolan [16] in defining stages of evolution in information systems. Data collection and measurement are discussed. Section 3 considers the problems and possible actions in moving between two of the stages.

2. Model of Network Service Usage

In the past, network usage nas frequently been viewed as an uncontrollable, environmental factor to the network itself ([1], [8]). Usage is given. Then models attempt to optimize performance. It is important to note that such models do not consider the behavior from the user's point of view.

Consider a typical user group. Usage begins with one user beginning to use a particular service for a particular problem area. Vendors have developed thousands of such applications (see Datapro [2] or Association of Time Sharing Users [3]). We will refer to this initial period as stage one.

Stage one begins after the user has become aware of a network service. A selection process has occurred. The usage has frequently been cost justified on some grounds and approval has been obtained from legal, purchasing, and other relevant groups. In some organizations this is a difficult process. Network services are external, direct costs as opposed to internal computing services. Internal computer centers may resist attempts to go outside. Countering this are the availability of specialized software and data bases. These cannot be cost-effectively obtained and maintained internally. Some of the characteristics of stage one (inception) that often appear are:

- initial tight control
- usage based on original cost justification
- high visibility of usage due to review an approval process
- vendor personnel supply substantial user support
- scrutiny by management and potential users

Two modes of behavior have been observed. The first is a stabilization in usage. Controls continue and other potential users lose interest and do not become users.

The second mode in stage one is that usage grows. In this mode the user group moves into stage two (expansion). Several factors can contribute to this behavior including some or all of the following:

- Relaxation of controls (with little on-going cost justification) due to achievement of some benefits
- Growth in number of users of software due to ease of use of package
- Growth in number of users and/or usage of data bases
- Growth in number and type of applications
- Increased or continued vendor support to encourage more usage

It might be thought that cost considerations outweigh these factors. However, that is frequently not the case. As part of this research, a survey of users in four major organizations was conducted. It revealed that response time, ease of use, availability of services, and range of services were rated as more important than cost.

Charges for network services can increase by the nature of the charging procedure. Most timesharing networks charge on the basis of storage usage, processing time, and connect time. Storage usage tends to increase unless controls are exercised. The number of programs builds. The amount of data stores for later analysis increases. In two organizations it was observed that storage charges constituted over 50% of the total bill. Processing time reflects CPU and I/O usage. This can increase due to the volume of data being handled and the number and extent of reports. Connect time increases with the amount of output as well.

The growth in the expansion stage (stage two) can be characterized by increases in some or all of the following:

- number of users for tne same applications
- number and type of applications
- processing cost in some applications due to volume

It has been observed that some applications can be very efficient for low volume work and yet lose efficiency rapidly as volume increases.

To some this phenomena may seem appalling. After all, how could supposedly well managed organizations tolerate this behavior? There are several possible answers. Some are:

- The second stage of network usage occurs very rapidly. Costs escalate by a multiple over only a few months.
- Accounting procedures are weak. Costs are charged to a variety of centers with no aggregation.
- Management is convinced of the importance of the application despite rising costs.

To highlight the first point in one financial application, costs rose by a factor of 7 in one month from a base of $2500 per month.

The expansion stage (stage two) can be seen as unstable and rapidly growing. There is a point at which this becomes a management concern. Efforts are then made to change the situation. The third stage occurs as controls and methods make usage more stable. Usage may still increase, but at a slower more controlled rate.

There are several ways to traverse from stage two-- expansion to stage three-- stability. Some vendors recognize this problem. This is one of the reasons for vendor discounting based on volume usage. Some network service vendors also offer discounts for deferred or delayed processing. However, these vendor options can serve to merely mask the underlying problems.

The problem areas characterizing the latter part of the expansion stage can include:

- lack of monitoring and control
- absence of measurement and evaluation of usage
- use of network services for nonorganizational purposes
- overuse of network services for organizational purposes of peripheral interest

The available options faced by management are many. Several are:

- acquisition of a user controlled minicomputer
- migration of work to large internal computer center
- migration of work to another network service
- reduction in usage and use of deferred processing and volume discounts

Of these, the usual response is one of the last three. The acquisition of a minicomputer can be a major capital investment. Many more resources may be needed than with a network service. However, the other migration options can involve conversion and startup costs. The problem often occurs that there is insufficient time and manning for decision making. The next section (section 3) addresses measurement and decision making to alleviate this problem.

The third stage, stability, is characterized by the following:

- measurement tools in place
- management control and requirements for at least informal cost-justification
- reduction in costs through more efficient applications and management procedures to reduce storage costs
- heightened awareness of costs

In this stage, the number of users and type of applications are relatively constant. There still may be some growth in costs. However, it is due to increased vendor charges and modification of user demands.

To summarize, a model has been presented for the usage life cycle for network services. The model is composed of three stages:

1. inception - few applications, tight control
2. expansion - multiple applications and users, lack of control
3. stability - slow growth or stability in usage, control reimposed

The measurement and decision making during the second stage is of particular interest and is examined in section 3.

In the remainder of this section we present some examples of situations their behavior during the usage life cycle.

Example 1: A financial group began using a network service. The services included a financial planning "language". The systems used could be categorized as decision-support in nature. Costs rose in six months from less than a thousand dollars per month to seven thousand per month. The group decided to replace the network service with a desktop minicomputer. Development and conversion costs of the transfer were paid back in less than a year.

Example 2: An accounting group was using service X. In a four month period usage and costs had more than tripled to over $5000 per month. An investigation revealed that storage usage could not be substantially reduced. Storage costs consumed over 65% of the monthly charges. The group whose software was written in BASIC migrated to another service Y with a much lower storage cost.

Example 3: A mathematical modeling group was accustomed to using a network service organization. The services of the internal computer center were rejected due to difficulty in use for nonprofessional programmers, lack of availability, and response time. The group hired a junior programmer who was familiar with the internal time-sharing system. Since all requests were channeled through the programmer, the group was almost indifferent to the service used. The programmer migrated the applications to the internal computer center.

Example 4: A personnel department used a time-sharing network to handle reporting from a large data base. Reorganization occurred at the same time stage two costs were increasing. The use of the service was stopped completely.

The four examples, drawn from three companies, demonstrate the range of responses to the problems occurring in stage two. In all but one case there was no measurement and decision-making instrument in place. In these cases the crisis in cost-awareness and cost escalation caused major dislocations in the user groups. In the one case where a migration to an internal minicomputer was foreseen, the measurement and decision-making process was oriented to the timing of the migration and scheduling of applications to be converted. The methods presented in section 3 were employed in this case.

3. Trade-off Analysis in State Two

This section considers the problems associated with moving from stage two to three. The feasible alternatives for a given situation depend on a variety of organizational and system factors. The range of alternatives could include the following:

- acquire minicomputer and transfer systems
- transfer to an internal large/medium computer
- transfer to another network service
- reduce and/or change usage, but continue with the same network service

These alternatives are quite different in terms of costs, implications on system conversion, and other factors. This points up the need for trade-off analysis.

The definition and measurement questions involve identifying early in stage two (or one) the reasonable alternatives. This frequently is part of a feasibility study to justify the use of network services. Cost information is available. Estimates can be made on system performance and later conversion costs.

During the first two stages the following information can be obtained and recorded:

- o cost schedule of currently used service
- o usage in each monthly period by user identification and type of use for processing time, connect time, and storage of programs and data

- performance of the system in terms of response time, personnel and training support, and other factors
- support costs for terminals and telecommunications

The next step is to develop and apply a method for using this information for decision-support. The method we develop here is first posed as a stochastic process with an objective of analyzing first passage time and obtaining optimal stopping rules. This can be localized to Markov chains.

For convenience we define the following variables:

t - time period (usually a month)

j - index of alternative available

i - index denoting a single application (this can be localized to a particular user group for a single application)

u_{ijt} - usage of applications i on service j on period t

$C(u_{ijt}) = C_{ijt}(u)$ - cost of using application i, service j during time t with usage u_{ijt}

TC_{ijkt} - period amortized transfer cost for application i from service j to service k during time period t

As discussed earlier these quantities can be collected and accumulated from actual use or estimated using the earlier analysis and benchmark comparisons.

It can be seen that usage in a given period is a random variable with an unknown distribution. However, the distribution can be estimated using past patterns of usage and time series analysis. It can also be observed that the time intervals are discrete. Define $N_{it}(u)$ by

$$N_{it}(u) = k_0 \text{ where } k_0 \text{ gives the minimum of}$$

$$\sum_{\substack{s \\ s>t}} C_{ijs}(u) \text{ and } \inf_{\substack{k \\ k \neq j}} \sum_{\substack{s \\ s>t}} C_{iks}(u) + TC_{ijkt}$$

where j is the service in use at t-1. Intuitively, this is the service which gives the smallest future cost from time t onward.

Proposition 1: $\{N_{it}\}$ is a Markov chain.

Proof: It has been observed that the time intervals are discrete. It follows from the independence of usage between periods that the summations inside the infimum function are Markovian. The proposition then follows from Cinlar [4] and Doob [6].

Stopping rules can be developed for this process. This will be a sufficient measure for a single application or in the case where one application is dominant over all others.

For multiple applications we can define

$$M_t(\underline{u}) = k_0 \text{ where } k_0 \text{ is index minimizing}$$

$$\sum_{i}\sum_{\substack{s \\ s>t}} C_{ijs}(\underline{u}) \text{ and } \inf_{k} \sum_{i}\sum_{\substack{s \\ s>t}} C_{iks}(\underline{u}) + TC_{ijkt} \qquad (2)$$

where $\underline{u}$ is a vector $(u_i, \ldots, u_m)$ with m applications and u_i is the usage of application $i (u_i = u_{ijs} = u_{iks})$. In this case $M_t(\underline{u})$ is the service which minimizes total future costs from time t onward.

<u>Proposition 2</u>: $\{M_t\}$ is a Markov chain.

The proof is analogous to that for proposition 1 and is omitted.

The second definition for M_t is needed since functions of N_{it} are not meaningful unless all applications are approximately of the same usage and usage characteristics.

Alternative formulations which were considered and rejected included counting functions based on preference in a given period. The problem here is that the decision to move to alternative services must be based on cumulative usage and estimated cost.

For either Markov chain we can consider the classification of states. In the cases examples mentioned early data was collected and supported the hypothesis that the transition probabilities $\{p_{ij}\}$ for each chain were time independent.

By considering only feasible alternatives, we insure that there are no absorbing states, and that each state is recurrent and non-null. For periodicity we have

<u>Proposition 3</u>: If usage patterns are periodic, the processes $\{N_{it}\}$ and $\{M_t\}$ are periodic

<u>Proof:</u> This follows from the relationship between usage, the transformation to obtain the transition probability matrix, and the definition of the processes $\{N_{it}\}$ and $\{M_t\}$.

This case will, however, occur infrequently. The only observed case involved a financial planning model used four times per year heavily and very little during other months.

Having classified the states we can develop several stopping rules. Two that have been applied are 1) stop using j and use k if the conditional expectation of the end cumulative cost of j given current cumulative costs of j exceeds that of k given the current cumulative cost of j by a certain margin, 2) stop using j and use k if the expected number of visits of the processes to k exceeds that to j by a specified margin. It is easy to show that these give rise to stopping rules in the sense of Doob [6].

To conclude this section, we can discuss the implementation and results of using the procedures in the cases cited in section 2. Data collection and setting up of the models have been examined.

In applying the procedure for one organization, a beta distribution was employed for the distribution of usage in a given time interval and system. This was generalized to a Dirichlet distribution for usage of multiple systems. Standard and Bayesian approaches were employed. The Bayesian procedure was developed using Dickey and Lientz [6]. The Bayesian procedure was most useful when usage in stage two was accelerating. Non Bayesian methods can be used for steady usage and periodic behavior in stages one and three.

The results in using above stopping rules are of interest. In all cases transfer occurred under the second stopping rule first. In one case with periodic use at high levels, the first stopping rule provided better information and was more sensitive to costs. The reason is that at times of low usage the number of visits to states which represent optimal low usage services are visits during periods of correspondingly low costs.

4. Conclusions and Remarks

A model for network usage and its evolution has been presented. This model serves as a usage life cycle. Usage of productivity aids available from network service suppliers can substantially compress the time duration of the first stage. The second stage of expansion is seen as dynamic, only partially controlled growth in network usage. The third stage of stability represents a period of tighter user control and maturity. The model applies to time-sharing vendors as well as services offered by different vendors through a network. Data collection and measurement problems have been discussed. A mathematical approach has been formulated for the problem of moving between stages two and three. Much more remains to be done. Controls must be developed earlier in stage two. Providers of network services should consider means of helping users gain tighter control over their usage and perform trade-off analysis. The realization of the life cycle should be embedded in the evaluation of such application such as decision support systems.

References

[1] Abrams, M. D. "A New Approach to Performance Evaluation of Computer Networks," Computer Networks: Trends and Applications, 1974, 15-20.

[2] "All about Remote Computing Services," Datapro, v. 3, 1975.

[3] Association of Time Sharing Users, Newsletter, 1976-77.

[4] Cinlar, E. Stochastic Processes, Englewood Cliffs, N. J.: Prentice-Hall, 1975.

[5] Cotton, I. W. "Network Management Surveys," Natl. Bu. of Stds. Report, Feb. 1974.

[6] Dickey, J. M. and B. P. Lientz, "The Weighted Likelihood Ratio, Sharp Hypotheses about Chances, The Order of a Markov Chain," Annals of Math. Statist. 41, (1970) 214-226.

[7] Doob, J. L. Stochastic Processes, New York: J. Wiley & Sons, Inc., 1953.

[8] Enslow, P. H., Jr. "Network Viability: Economic, Legal and Social Considerations," Compcon 73, 1973, 7-9.

[9] Enslow, P. H., Jr. "Non-Technical Issues in Network Design: Economic, Legal, Social and Other Considerations," Computer, 6, August 1973, 20-30.

[10] Lientz, B. P. and I. R. Weiss, "The Vulnerability of Computer Auditing," CPA Journal, March, 1977.

[11] Lientz, B. P. "Management Evaluation of Network Performance with Security Measures," Computer Performance Evaluation (P. J. Kiviat, ed.) Chameleon Press, Ltd., London, Eng., 1976.

[12] Lientz, B. P. "A Comparative Evaluation of Versions of BASIC," Comm. ACM, 19, 1976, 175-181.

[13] Lientz, B. P. "Evaluation of Network Services," Proc. Assoc. for Comput. Mach. Meeting, 1976, 218-220.

[14] Lientz, B. P. and J. Arnold. "The Accountant and Computer Software Selection - Some Proposed Guidelines," Calif. CPA Quart., 1977 (in press).

[15] Lientz, B. P. and I. R. Weiss. "Effects of Security Measures on Network Performance," Computer Networks (Accepted for publication, in revision).

[16] Nolan, R. "Managing the Computer Resource: A Stage Hypothesis," Comm. ACM, 16, 7, July 1973, 399-405.

[17] Pyke, T. N., Jr. and R. P. Blanc. "Computer Network Technology: A State of the Art Review," Computer 6, August 1973, 12-19

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Multistage Model of Network Usage.		5. TYPE OF REPORT & PERIOD COVERED technical rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bennet P. Lientz		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0266
9. PERFORMING ORGANIZATION NAME AND ADDRESS Graduate School of Management University of California, Los Angeles		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 049-345
11. CONTROLLING OFFICE NAME AND ADDRESS Information Systems Program Office of Naval Research, Arlington, Va.		12. REPORT DATE June 1977
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) —— 18p.		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) distribution of this document is unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) ——		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) - computer network - usage life cycle - optimal stopping		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The evolution of network service usage can be considered as being in three stages, constituting a usage life cycle. The first stage (inception) is that of initial low level usage. The second (expansion) is characterized by rapid growth in usage and is often unstable and uncontrolled. The third stage (stability) is a reaction to the second stage and imposes order on usage. Decision making, data collection, and measurement considerations are examined for this framework.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

407436