High-Level Fusion using Bayesian Networks: Applications in Command and Control

P. Bladon, P.S. Day, T. Hughes and P. Stanley
BAE Systems Advanced Technology Centre
Bristol, BS34 7QW
UK

peter.bladon@baesystems.com / pete.day@baesystems.com
tim.j.hughes@baesystems.com / philip.stanley@baesystems.com

ABSTRACT

In this paper, we discuss how Bayesian networks can be used to develop automated situation-assessment tools suitable for use as decision aids in a command and control system. Inevitably, the introduction of a new technology raises a number of validation, systems integration and human-factors questions. Those issues pertinent to Bayesian network decision aids are identified and their implications discussed. We then describe in detail the implementation of such a system capable of providing Combat-ID and Threat Assessment advisories in the naval anti-air warfare role and its assessment within a realistic (synthetic) human-in-the-loop experiment. We discuss the experimental system, the experimental design and protocol and the experimental results. In a controlled experiment using 14 subjects with relevant military experience we found that the Bayes' net decision aid system was preferred by the majority of the experimental subjects and led to a number of operator performance improvements which could directly contribute to improved operational effectiveness.

1.0 INTRODUCTION

Achieving shared situation awareness and self-synchronisation is a stated objective of Network Centric Warfare and its UK equivalent, Network Enabled Capability [1,2]. This requires networking of sensors and users across the battlespace, which will undoubtedly increase the amount of information available to decision makers. However, increasing the volume of information may not necessarily lead to better decisions. Operators may become overwhelmed with the task of filtering and assessing data from different sources. Therefore it may be necessary to deploy tools for decision support, to help operators assimilate information and make better decisions, faster.

Decision making involves processing uncertain information. It requires the assessment of the current state of the world (which is itself uncertain) and the projection of that state forward in time, so that the risks and rewards of different courses of action can be assessed.

In this paper, we consider tools to assist this "situation assessment" task, which we define as the process of placing observations in context with existing (i.e. prior) knowledge. Such a tool must be capable of integrating multiple sources of information of different types (sensor data, historical data, intelligence reports, etc) and dealing with the uncertainty inherent in all data sources. It must also be capable of incorporating relevant prior knowledge into the process. In recent years Bayesian probability theory, which provides a consistent framework for reasoning with uncertain data [4], has been advocated as a suitable tool for building this type of system, and a number of concept demonstrations have been constructed [14,15,16,17,18].

Bladon, P.; Day, P.S.; Hughes, T.; Stanley, P. (2006) High-Level Fusion using Bayesian Networks: Applications in Command and Control. In *Information Fusion for Command Support* (pp. 4-1 – 4-18). Meeting Proceedings RTO-MP-IST-055, Paper 4. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 DEC 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
High-Level Fusion using Bayesian Networks: Applications in Command and Control	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **BAE Systems Advanced Technology Centre Bristol, BS34 7QW UK**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002031., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**54**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

However, there is a large technology gap between a concept demonstration and a deployable system. Bridging the gap requires the development of detailed concepts of use; understanding the requirements of users, their modes of interaction with the system; and addressing the issues of system validation, verification and integration.

One could attempt to study these issues in the abstract. However, it is our belief that only by building and experimenting with a representative system can the true systems development and operational issues be identified and potential solutions to these issues found more quickly. In this paper, we report how we have constructed a decision support tool for situation assessment and conducted an experimental assessment of its utility in a realistic (though simulated) setting. In Section 2 we describe how Bayesian networks can be used for the situation assessment task; in Section 3 we highlight specific system implementation issues and identify some of their potential solutions; and in Section 4 we describe the experiment, its design and the results. We summarise our progress and successes and discuss where future effort should be deployed in Section 5.

2.0 SITUATION ASSESSMENT USING BAYESIAN NETWORKS

BAE Systems have been developing decision support systems designed to help an operator achieve "situation awareness" by rapidly and consistently assimilating uncertain data from multiple sources i.e. by producing an automated "situation assessment" for an operator to use. As noted above, this process must deal with the uncertainty inherent in information sources arising from data collection and allow prior knowledge (e.g. encyclopaedic knowledge of platform capabilities, expert knowledge of threat, etc) to be incorporated into the decision process.

Bayesian probability theory provides a consistent mathematical framework for representing and manipulating uncertainty, and allows prior knowledge to be utilised in inference [4]. Hence we believe it is a natural choice for modelling the situation assessment process.

Bayesian networks provide a computationally tractable method of implementing Bayesian probability theory [5, 6]. In recent years, they have become the de-facto language for expressing problems in probability and statistics. A Bayesian network uses a graph to specify relationships between different variables, which can be exploited to provide inference efficient algorithms. Bayesian networks can be used to express and generalise many common inference tasks. For example Figure 1 shows the structure of networks commonly used in level 1 data-fusion applications [12, 13]: (a) Naïve Bayes classifier (used in classification and data mining); (b) a Kalman filter (used in tracking and control); and (c) a feed-forward neural network (used for regression and classification).

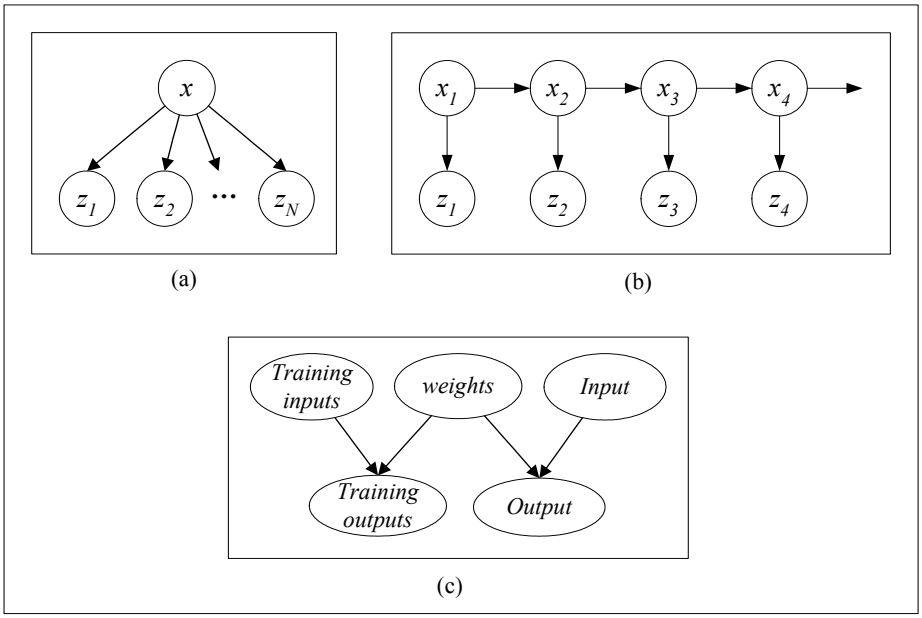


Figure 1: The structure of Bayesian networks for (a) a naïve Bayes classifier, used to classify a variable 'x' using a set of observations $z_1,\ldots,z_N$, (b) a Kalman filter, used to estimate a state variable x over time, using a sequence of observations z and (c) a feedforward neural network, showing the relation of training inputs and outputs to the weights.

Bayesian networks can also be applied to more sophisticated or "high-level" inference or fusion tasks. By exploiting the structure of the graph as a cause and effect model of a domain, it is possible to build a "knowledge base" that can be used as a (probabilistic) expert system. In the civil domain, Bayesian networks have been applied to a wide range of applications that are effectively civil analogues of the military "situation assessment" problem. These include medical diagnosis [7], fault finding [8, 9], data mining [10] and network intrusion detection [11]. Following the successful application of these techniques in the civil domain, they are also beginning to be exploited in military applications that include situation assessment and threat assessment [14 – 17] and effects based operations [18].

High-level fusion applications necessarily require domain specific models. In our experience (and in the experience of others [21]), constructing useful domain specific models is a significant challenge. We have found it most useful to build models of individual entities in the battlespace. This is because the domain is relatively small (the number of variables to describe an entity is small compared to the number to describe a whole battlespace), closed (i.e. each entity uses a fixed number of variables); and contains a manageable number of hypotheses (for example hypotheses about Hostility, ID and Threat). However, it is a domain containing non-trivial inference, for example inferring ID and Threat using data present in a command and control system. It also provides a stepping-stone to more sophisticated applications, for example inference about the behaviours of groups of entities.

Figure 2 shows a simplified model for assessing Hostility, ID and Threat by placing sensor data in context with historical data and expert knowledge. The model is a simplified version of those used in the experiments, described in Section 4. The model combines:

- **Track Data** extracted from a tracking system;
- **ID Data** such as Automatic Target Recognition (ATR) and Electronic Surveillance measures (ESM) information that can be fused within the Bayesian paradigm;

- **Intelligence Information** about intent. Other intelligence such as known force dispositions can also be encoded;
- **Historical Data:** knowledge of the performance envelope of a platform performance is used to provide weak evidence for ID;
- **Spatial Context Information:** knowledge of air lanes (and potential exclusion zones) is used to inform hostility;
- **Expert knowledge**: is used to combine different factors into a threat assessment.

Note that the model contains a concept that is inherently abstract: Threat. In this respect, the model captures a domain expert's subjective knowledge of how attack profile and ID should be interpreted in a situation assessment. Note also that the model has an underlying "tree-like" structure at its heart, reminiscent of the Naïve Bayes Classifier of Figure 1(a). This is to be expected: "sensor" readings (ESM, ATR, Speed, etc) will only be correlated via the ID or class of the object under inspection. In more realistic models, sensor performance may be correlated via other contextual nodes relating to environmental factors (e.g. prevailing weather conditions that effect multiple sensors in different ways). This tree-like structure allows individual sub-components of the network (for example ATR and ID Platform) to be designed and verified independently of the rest of the network.

Models of this type can be constructed using Subject Matter Experts (SMEs) in an iterative knowledge capture process. Typically one or more domain experts construct a cause and effect model of the domain (i.e. a model where the conditioning is interpreted causally) which is then parameterised using data (where available) or further knowledge elicitation. Once constructed, data can be used to refine estimates of parameters or even structure.

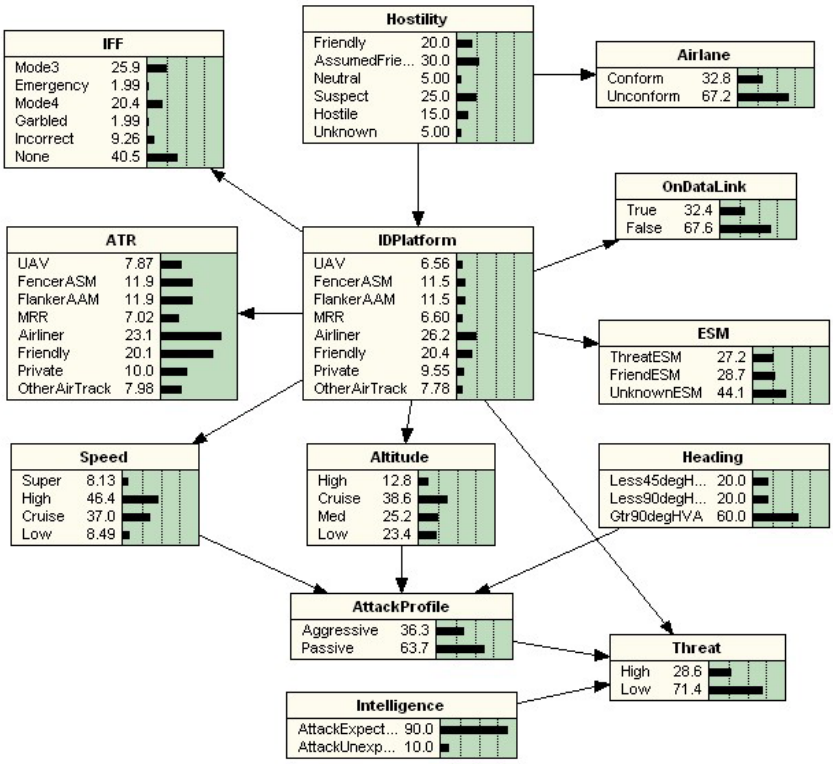


Figure 2: A simplified model of a Bayesian network used for Combat ID and threat assessment in our experiments

The model can be used in a number of ways to extract (inferred) information about the Hostility, ID Platform and Threat variables. Since, in a Bayesian network, there is no concept of "inputs" or "outputs", inferences can always be made regardless of the number of observations available at the time. This is done by conditioning the model with available evidence (i.e. observation and background data) and running inference algorithms of different types. For example, for the model shown in Figure 2, evidence for Speed, Altitude and Heading can be extracted from an analysis of track data; evidence for (Identification Friend or foe (IFF), ATR and ESM can be obtained from appropriate sensors systems; evidence for Airlane can be extracted from a combination of track data, Air Combat Order (ACO) and Air Tasking Order (ATO) data. There is a wide choice of inference algorithms that allow different types of question to be answered by a model. For example, for the model shown in Figure 2, we may be interested in inferring:

- The probability of any ID or Hostility hypothesis, given the evidence (e.g. P (ID Platform=UAV | Evidence) = 0.95);
- The most likely ID or Hostility of the track (e.g. ID Platform=UAV);
- The most important piece of evidence supporting ID hypothesis (e.g. ATR is the most important piece of evidence – the one that makes the largest difference to the data likelihood);
- The measurement that will maximise the discriminative power of the model (e.g. use ATR system).

Which inference algorithm to use is dependent on the role of the model in a decision support task, which in turn depends on the concept of use of the tool. In summary, we believe Bayesian networks have an important role to play in building tools for situation assessment, as they provide a consistent framework for handling uncertainty, can be used to constructed rich domain models that can incorporate prior (expert) knowledge, and they can be used to consistently infer different types of information (e.g. P(ID), most likely ID, etc) from a common knowledge base. However, the usefulness of such tools will always depend on the quality of the underlying model, it suitability for the task, its concept of use, and the way in which it is integrated into the system.

2.1 System Issues

In moving from a bench-top technology demonstration or concept to a more realistic system demonstration a number of implementation issues must be addressed, which we discuss in this section. These include:

- Validation and verification of models;
- Integration with existing information systems and network technologies; and
- Interaction with human operators and existing military processes and doctrines.

2.1.1 Validation and verification of models

Validating any knowledge based system is challenging. However, the causal structure of a Bayesian network can be exploited in the validation process. Groups of nodes can be validated independently once the overall structure has been determined, providing data is available for all nodes in a group. For example, consider the ATR and ID Platform nodes of the model in Figure 2. These two nodes constitute a probabilistic ID system, which can be objectively validated independently of the rest of the model, providing suitable data are available. Of course, the parameters of the model can also be learned from data, using standard probabilistic learning algorithms.

To validate the model as a whole we would generally need access to three types of data: sensor data; encyclopaedic data and expert derived data. Sensor models can be incorporated directly into the model, where validated models are available. Otherwise they can be validated given sufficient statistical data. Encyclopaedic data, such as the performance envelopes of platforms, can be used, for example to validate

the relationships between ID and observed behaviours. However, 'expert derived' data presents a slightly more difficult task. Typically we might want to infer a subjective quantity such as the level of threat a platform presents. In this case, we would advocate isolating the smallest group of variables that influence the subjective measure (i.e. the node's Markov blanket), and generate a series of samples that can be reviewed, and if necessary updated, by a quorum of experts.

2.1.2 Integration with information systems and network technologies

Current state of the art network technologies such as C2IEDM [19] and Link 16 [20] encode limited notions of uncertainty, particularly with respect to ID hypotheses. Data in the system is presumed to be true, which manifests itself in procedures for manual conflict resolution.

A probabilistic approach would allow more sophisticated management of the uncertainty associated with each piece of information. We would argue that this represents a truer representation of the state of knowledge we have about the world which, if properly communicated, may prevent decisions being taken on information that is incorrect (e.g. potentially reducing Blue on Blue incidents).

Clearly, there is a requirement to understand how probabilistic information is best used, but also how it can operate along side, or be integrated with legacy information systems that have no notion of uncertainty.

2.1.3 Interaction with the human operators

The level of interaction between the human and the decision support system depends on the operator's requirements, his current ability to assimilate information and the level of automation provided by the system. These are dependent on the task being undertaken.

Using a probabilistic approach to decision support, there are a number of alternatives for presenting information, including:

- Present only the most likely hypothesis. This has the advantage that only minimal changes will be required to existing operator interfaces. However, the danger exists that the information will be treated as absolute truth.

- Present probabilistic information on relevant hypotheses (e.g. hostility and identity). This would require operators to be trained to interpret this information. It may lead to large amounts of data being displayed and increase the potential that the operator is distracted.

- Present probabilistic information on relevant hypotheses and allow the operator to interact with the model. For example, the most likely hypothesis could be presented with a list of corroborative evidence to review using a "what-if" analysis. Again, this would require sufficient operator training, and the development of sophisticated interfaces which would represent a large departure from current practise.

3.0 EXPERIMENTAL INVESTIGATION

3.1 Overview

An experimental investigation, funded by a BAE Systems internal research and development programme, was undertaken to allow some of the system issues to be explored in a realistic environment, and the maturity of the technology to be assessed. The study (referred to as the 'Shared Situation Awareness'

(SSA) experiment) was aimed at assessing the potential operational advantages to be accrued from the use of a Bayesian network (BN) based decision support system in a naval command and control task.

The philosophy of operation for the decision support system was that it should act as an adjunct to the presentation of the existing data within a Combat Management System (CMS). The system was intended to act in an advisory rather than fully automated role to assist the human operator to perform his task, not to replace him.

The decision support system used was based on a Combat-ID / Threat Assessment network similar to example shown in Figure 2 but which contained a 'richer model', with a different structure, and more variables. The model was configured using expert knowledge gathered from personnel with Naval and Air-Force air picture compilation expertise together with extensive scenario analysis based on the DoDAF framework.

3.2 Implementation of Concept

The BN model is only a small component of the experimental system. The other components including the Human machine interface (HMI), scenario generation, data marshalling and processing are shown schematically in Figure 3.

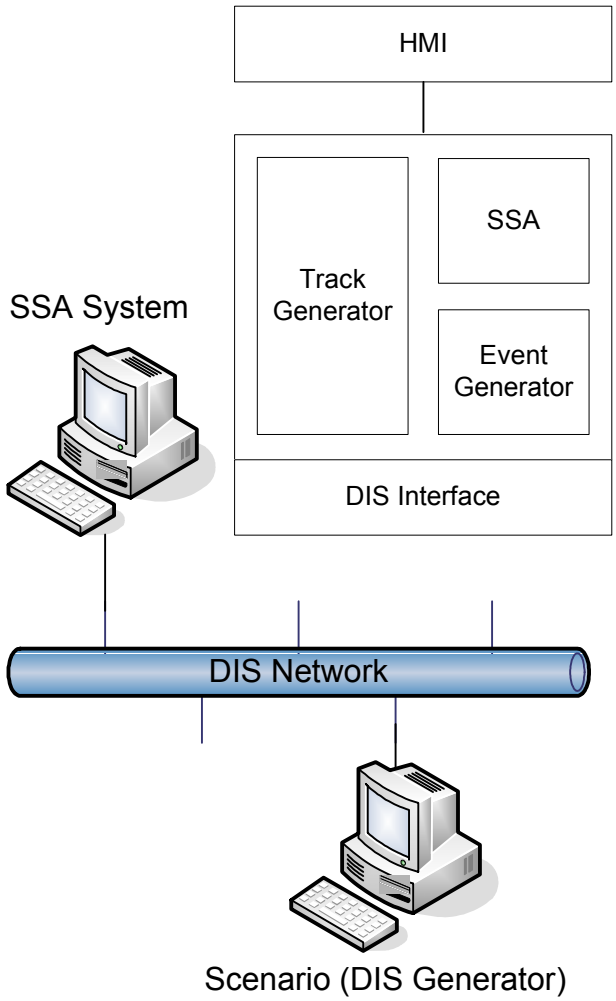


Figure 3 Experimental Configuration

Scenario information was propagated using a DIS protocol network. Scenarios were generated using VR Forces™, and logged for later replay.

The tracking module reads ground truth from the DIS interface and simulates the tracking of entities, according to the distribution of platforms and associated sensors in the scenario.

The event generator simulates the generation of all non-track data used in the scenario (i.e. intelligence reports, analysis of ACO / ATO data (e.g. air-lanes), etc). It works by reading ground truth data from the DIS network, and applying multiple filters, one for each element of the non-track data in the system. For example, an "Airlane event" would be triggered by analysis of the position of each entity and all air-lanes, resulting in a distribution over air lane occupation for each track. This mechanism was also used to inject all scenario specific events via a configuration file.

The SSA component contains a track management module that manages all track data, the event data and their correlation. This is used to build an "evidence" vector for each entity which is input to a generic BN inference module that is configured via an XML file. The SSA module also manages the interface with the HMI, a specially modified version of a Naval Combat Management System (CMS) developed by BAE Systems Insyte organisation, known as the "Common Naval System" (CNS) see figure 4. All state information was stored within the SSA module, to minimise the modifications necessary to the CNS.

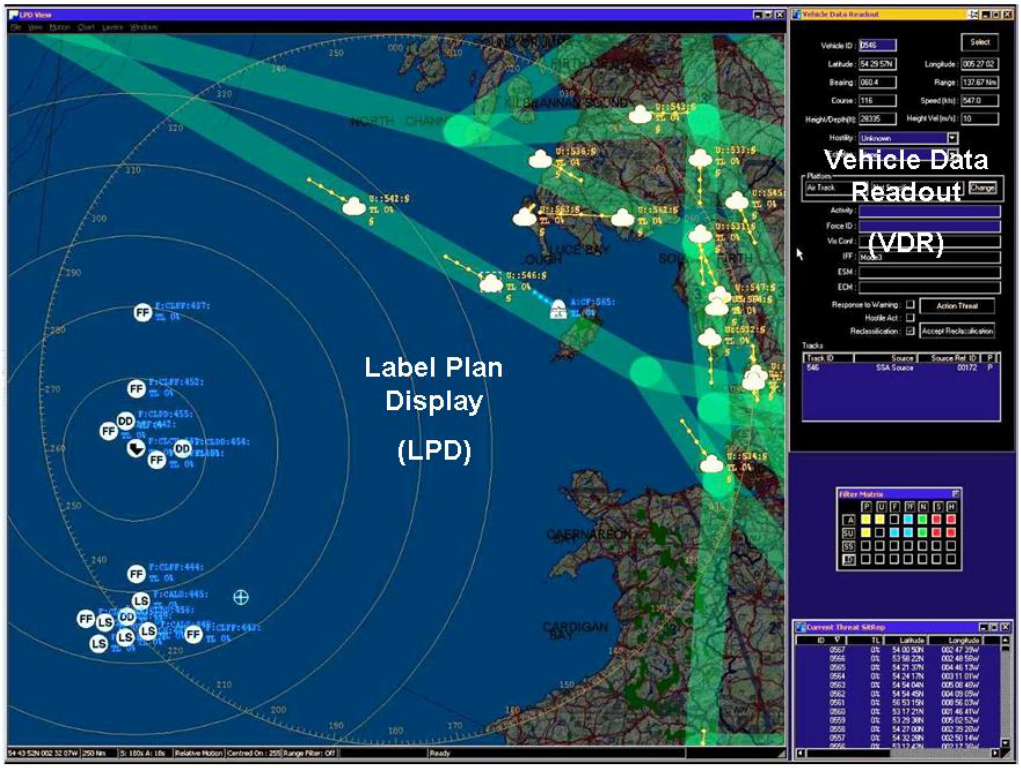


Figure 4: Experimental Interface

3.2.1 Human Machine Interface

The existing human machine interface (HMI) of the CNS command system was modified to facilitate the communication between the human operator and the SSA system. The primary means of communication was via the implementation of a split symbology convention. Information was also communicated via the track label, and the main tote display known as the Vehicle Data Readout (VDR).

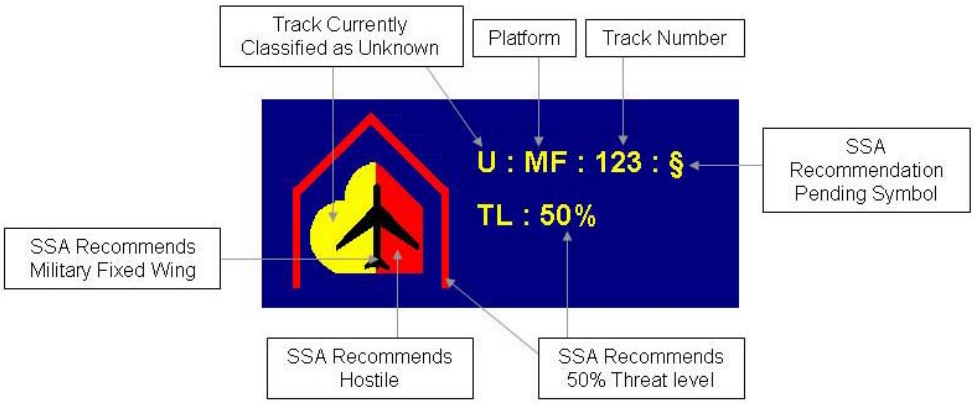


Figure 5: Split Symbology and Track Labels

The symbology, illustrated in Figure 5, is an adaptation of the Mil Std 2525B symbol set, which was specifically chosen because of its symmetry. The left half of the symbol shows the status, in terms of the hostility level that the track currently possess within the command system. The right half of the symbol indicates a suggested change to that status indicated by the BN decision support system (known as SSA). In the case shown in Figure 5, the system currently has the entity classified as an unknown air track, (normally a full yellow cloud). The SSA is suggesting to the user that it should be reclassified as a hostile (red "house" shape) based on its characteristics and behaviour.

Recommendations for platform type (the ID Platform variable in the model) are communicated by the icon in the centre of the symbol. In the example shown the BN's suggestion is that the track should be reclassified as a military fixed wing aircraft. Finally, an additional field was added to the track label (the 'SSA Recommendation Pending' symbol) to indicate that the SSA had made an assessment of the track which the operator had not yet responded to.

The operator was able to directly accept or reject the SSA's hostility and track ID recommendations via various means built into the HMI. In the event of an operator rejection the system would not re-advise him until a further change of state in the classification was evident.

Figure 6 shows the threat symbology. The red chevrons around the track symbol shown indicate the BN's assessment of how much of a threat the track represents to the operator's own ship. The BN is capable of assessing threat level on a probability scale from 0 to 1, however to simplify the interpretation and presentation of this probabilistic data to the operator the threat levels were quantized into 3 levels according to the scheme shown below.

- Low Threat = 0 to 49%
- Medium Threat = 50 to 74%
- High Threat = 75 to 100%

Low threat tracks (0 - 49%) were displayed without 'threat chevrons', medium threat tracks (50 – 74%) had one 'threat chevron' and high threat tracks (75% or greater) had two 'threat chevrons' (see Figure 6). Therefore, more chevrons mean a greater level of threat. The quantized percentage level for each track was also displayed in the track label. See Figure 5.

Figure 6: Threat Symbology

3.3 Experimental Task

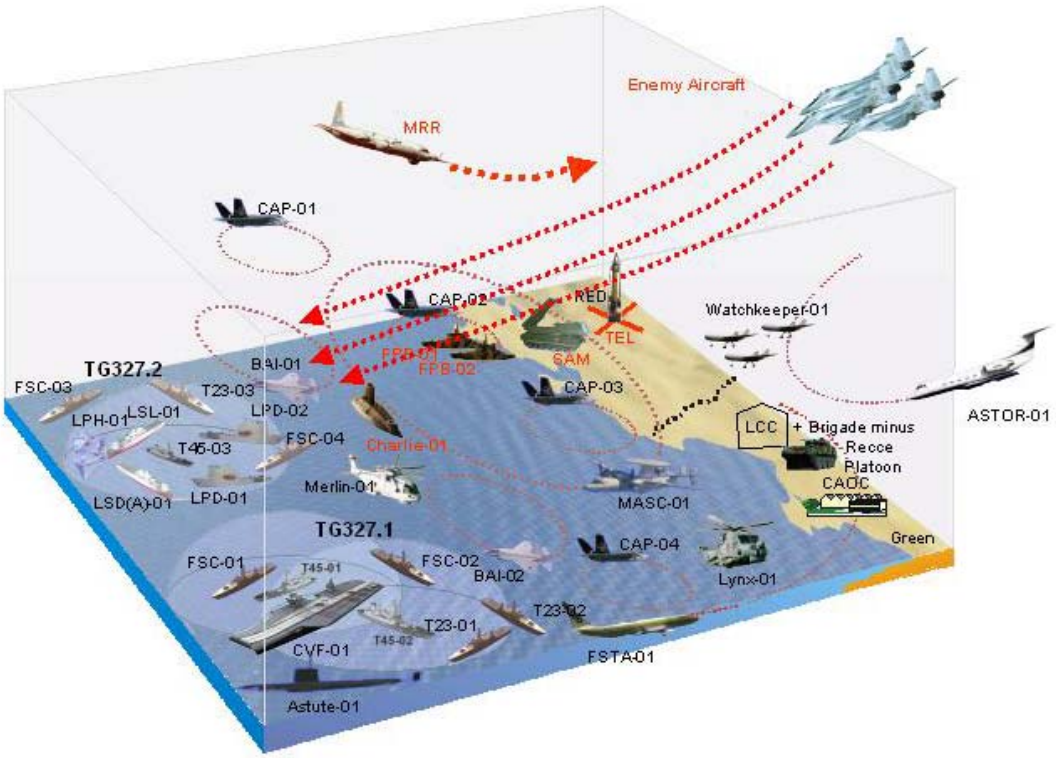


Figure 7: Schematic of Global Operational Context

The experimental task was for an operator to differentiate and classify air tracks within an air/land/sea littoral battlespace, on behalf of a Task Force Commander[1]. Using the identification criteria provided, for each air track the operator was asked to establish its: ID allegiance or hostility classification, its platform ID (type of platform), and the level of threat to the task force flagship posed by the track.

Figure 7 shows a schematic of the global operational context of the scenarios within which the operators were asked to participate during the experiment. It included up to 180 potential enemy air tracks which had to be differentiated and classified by the operators. To prevent learning between the conditions of the experiment two versions of the experimental scenario based on the same operational context were produced, each with differing timelines to prevent the operators predicting the events. The experimental

[1] Other tasks that are typically performed by operators in this role such as track management and communications were not simulated in this experiment.

scenario was created using a proprietary synthetic environment tool (VR Forces™) and this was used to drive the command system simulation with which the operators interacted as described in section 3.2.

3.4 Experimental Design

The experiment adopted a fully counterbalanced, within subjects design in which the only independent variable manipulated was whether the operator had the assistance from the BN decision support system (SSA) or not. A total of 14 experienced operators who have performed, or are currently performing air picture compilation tasks took part in the experiment. Of these 11 had a Royal Navy (RN) background and 3 had a Royal Air Force (RAF) maritime patrol background.

Each subject participated in the experiment alone, and on their arrival they were given a full briefing and practice session lasting a minimum of 2 hours to minimise any learning effects. This was followed by two experimental runs each of 2 hours in duration separated by a break for lunch. The participants completed the experimental task both with and without the assistance of the BN decision support system. The order of presentation of the conditions and scenarios was fully counterbalanced across the subjects. Following the experimental runs each subject was debriefed and asked to complete a short questionnaire concerning their experiences with the technology.

Objective Performance Measures included:

- Time to correctly identify each track's hostility (ID allegiance)
- Time to correctly identify each track's type (ID Platform)
- Time taken to respond to (take action against) tracks which posed a threat to the Taskforce flagship
- The number of tracks validated

Subjective data gathered included:

- Situation Awareness Ratings (Crew Awareness Rating Scale (CARS))
- Workload ratings (Instantaneous Self Assessment (ISA))
- Questionnaire responses

The whole process was also video recorded with a sound track to permit post experiment analysis of interesting events.

3.5 Results

The relatively large number of experimental participants generated a large amount of data. Limited space prohibits a full presentation of the results of the experiment. The following sections contain a selection of the most noteworthy findings. The results are presented under the headings Objective and Subjective Measures respectively.

3.5.1 Objective Measures

The data collected obtained on the time taken for the operators to respond to a threat. i.e. identify a track as a potential threat to the task force flag ship is shown in Figure 8. These data did indicate a potential for enhanced operator performance. When using the BN decision support the operators were on average almost two minutes (1 min 50 secs) quicker at determining whether an incoming track posed a threat to their ship. This represents an approximately 14% reduction in the time to take action against a potential threat compared to when the operators' interface was not augmented by the BN. We believe this time saving could translate into a significant operational advantage in defensive operations.

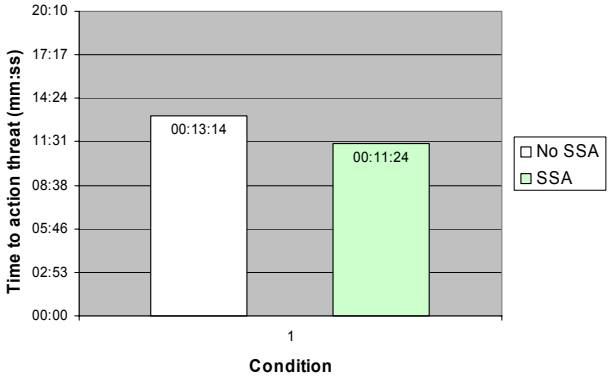


Figure 8: Average Time taken to Respond to Threats

However, when considering the average time for the operators to first correctly identify a track's hostility partitioned by time of day as shown in figure 9, we can clearly see that there was a large experimental effect due to learning.

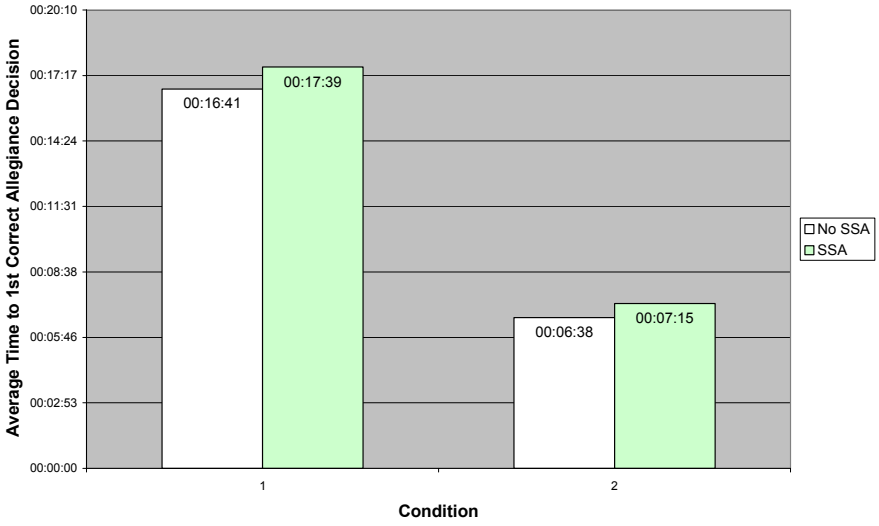


Figure 9: Average Time to Operator's First Correct Identification of ID Allegiance (by time of day)

Subjects were around two and half times faster to achieve correct hostility classifications in the afternoon compared to the morning. This was by far the biggest effect seen in the experiment. Somewhat unexpectedly these data also indicate that the operators were on average generally slower to achieve a correct hostility assessment with the BN decision aid than without it.

There are a number of hypotheses (which we are yet to fully explore) which might explain these data. For example it seems clear that despite our best efforts the time allowed for training and practise was probably insufficient. Subjects were clearly still learning the task throughout the morning experimental session. Perhaps in future more time needs to be allowed to ensure the subjects are at the top of the learning curve.

Also, we think that despite efforts to give the subjects lots to do by having a large number of tracks, the overall workload in the task was too low. This is borne out by the overall workload ratings collected from the subjects (see section 3.5.2). In the absence of ancillary tasks such as track and communications

management, the subjects seemed well able to cope with the density of tracks presented, even without a decision support system. In fact the BN decision aid may have given them more to think about and do.

A further hypothesis is that the BN decision aid technology may prove to be more operationally advantageous when the operators are required to work harder. In this case they may be more reliant on the decision support in order to free up cognitive resources for other tasks. It would be interesting to repeat the experiment and manipulate the level of workload to see the effects on performance.

Finally it was recognised that there were aspects of both the implementation of the SSA HMI and the underlying BN which were less than optimal and may have introduced additional tasks.

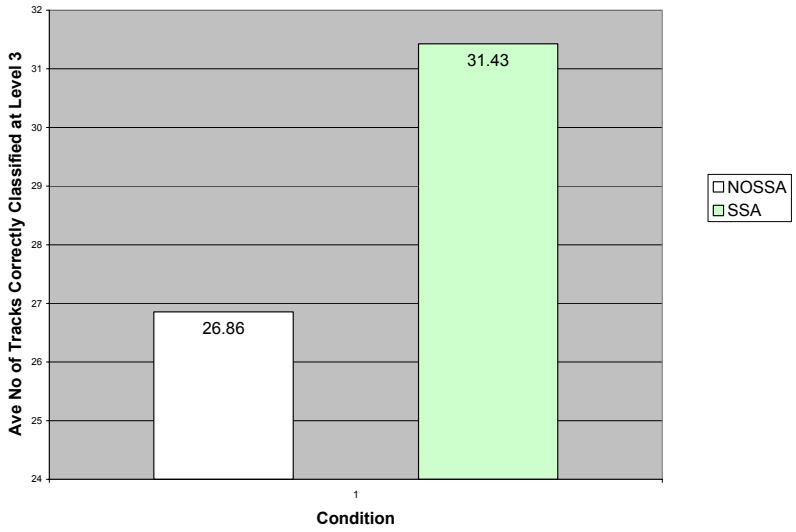


Figure 10: Number of Correct Platform ID Classifications

The final quantitative data to consider shows the average number of tracks for which operators were able to provide a level 3 or higher platform classification. Level 1 is the basic platform identification which identifies the track as an air or surface track. Level 2 delineates between military, civilian and missile tracks. At level 3 the operator makes decisions concerning the type of air vehicle represented by the track such as whether it's a fixed or rotary wing aircraft. Level 4 is concerned with the role of the aircraft such as Bomber or fighter etc. In our experiment the subjects were able to correctly classify approx 14.5% more tracks to level 3 or above when using the BN decision support compared to the unaided condition. It seems therefore that the operators may have a more detailed or deeper awareness of the real state of the world when using the BN decision aid technology compared to the unaided condition.

3.6 Subjective Measures

Considering the instantaneous Self Assessment (ISA) workload scores there were no discernable differences between the experimental conditions in terms of median workload score. The overall task workload was shown to be very low. Both conditions were rated on average as 2 on a scale of 0 to 5. When the workload profiles over time for the two conditions are we can see that they were highly correlated (r=0.84, N=41).

Similar results were found for operators' ratings of their situation awareness (SA) using the Crew Awareness Rating Scale (CARS). There was only one minor discernable difference in the SA profiles for the two experimental conditions. From these data we can perhaps conclude that at least the introduction of

the BN decision support technology seemed to have no measurable negative impact in terms of operator workload or SA.

3.6.1 Extracts from Questionnaire Results

The first finding relates to the operators' views on how the BN came up with its recommendations. Figure 11 illustrates that the majority of subjects (64%) reported that they understood how the BN was making its decisions. This in turn seemed to reinforce the trust that the operators were prepared to invest in the systems' recommendations.

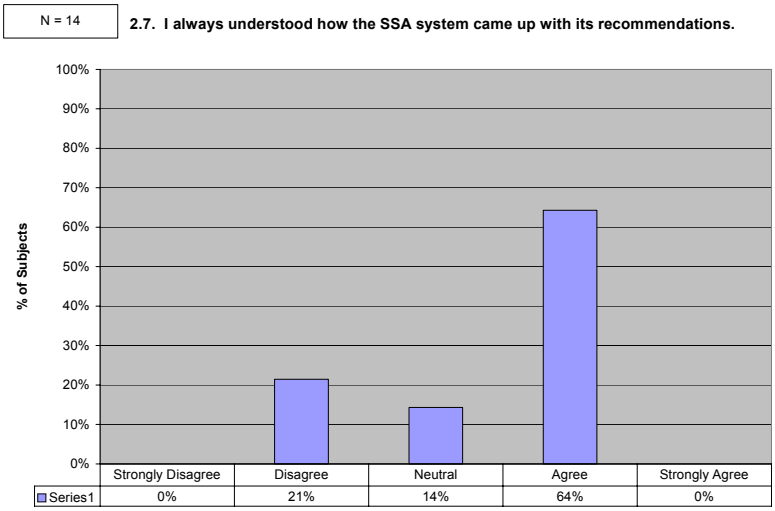


Figure 11: "I Always understood the SSA recommendations"

When required to make a choice, 79% (11 out of 14) of our subjects said they preferred the BN decision aid system over the non-augmented interface (see Figure 12).

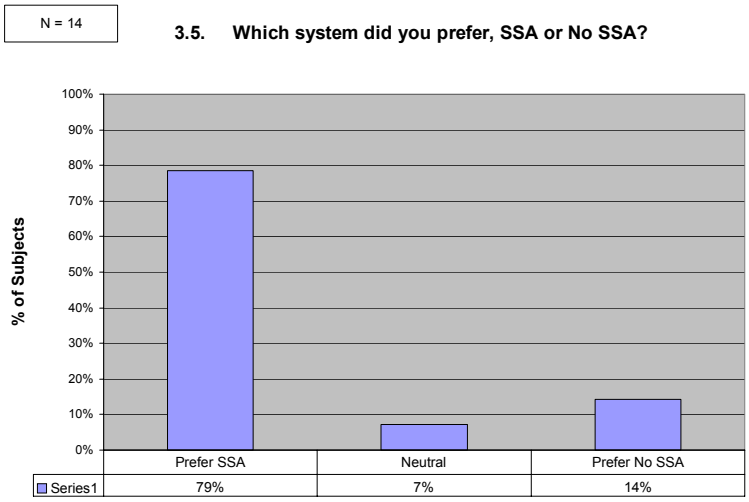


Figure 12 "Which system did you prefer?"

Overall the operators' opinions of the augmented system were generally positive. Examples of operators' comments included:

- *"…Platform change suggestions + threat recommendations were all very useful. SSA has good potential."*
- *"Less thought required (with SSA). More time to concentrate on the tactical environment, i.e. Track(s) release to data links, etc."*
- *"It was good to have a computer that thinks for you and is there to help, not just do!"*
- *"Allows you to leave less important track investigation till later. Gives immediate heads-up on possible hostiles."*
- *"In many aspects, it is better to be presented with a solution which can be vetoed, rather than trying to find the answers.*
- *"Having advice based on objective criteria provided reassurance about classification decisions and made the decisions quicker to take…"*

3.7 Summary of Experimental Findings

Whilst not all measures showed a benefit resulting from the use of the BN decision support system, (SSA) the results obtained did indicate a clear potential for enhanced operator performance, in terms of a 14% reduction in the time required to take action against a possible threat when the operators' interface was augmented by the SSA. This could translate into improved operational effectiveness in defensive operations.

The 'time to first correct allegiance decision' data showed no performance advantage from the addition of the BN decision aid technology. The operators were on average slightly slower to achieve a correct ID allegiance with SSA than when it was not available to them. This could be due to the requirement to verify the BN's suggestions prior to making a response, which is an addition to the operator's task when compared to the unaugmented condition. The operators' lack of familiarity with the system may exacerbate this. Alternatively, operators may have been more content to leave the system alone when a BN generated advisory was displayed.

There is also evidence that operators had a more complete picture with respect to platform type classification, whilst using the BN decision aid. Without it, on average, operators were able to identify fewer tracks in terms of platform type. With it, they were more likely to correctly identify platform and achieve a more detailed level of platform identification. This indicates that the subjects may have a deeper awareness of the real state of the world when using the SSA system which could be operationally significant.

The analysis of subjective workload data indicated no differences between the conditions in the experiment. Hence the introduction of the SSA did not result in a perceived increase in workload by the operators as may have been expected. The operators' perceived level of situation awareness also showed no difference between the conditions. These findings are possibly a function of the low workload imposed by the operational task, which was so low that the subjects were easily "on top of the operational situation" even without assistance from the SSA. This low task workload may have also accounted for the lack of the expected advantage for the SSA technology in terms of time to correctly identify hostility.

The post-experiment questionnaire revealed that the operators' opinions about the technology were generally positive. The majority of participants understood the decisions made by the SSA, however, they did not necessarily agree with them. This may be because the SSA did not account for historical track

behaviour. Further work is recommended to address this issue. The majority of subjects did approve of the SSA's means of presentation, and 79% preferred the SSA system compared to the un augmented system.

A significant body of opinion about the design and operation of the CNS, the SSA tool and the HMI implementation was also gathered which will be valuable in any future development of these technologies.

In summary, the subjects thought that the BN decision aid assisted their performance on the task. They acknowledged some problems with the implementation of the system, but notwithstanding these issues they overwhelmingly preferred it and recognised the potential benefits to be accrued from its use following further refinement of the BN and HMI.

4.0 SUMMARY

The future battlespace implies an increasing need for decision support technologies. BAE Systems has developed high level fusion technology based on Bayesian Networks that can support human decision processes in a NEC environment. We have described an experiment in which we assessed a Bayes Net decision support application in a naval command and control task. The experiment described was conducted in a realistic, simulated, naval scenario and has demonstrated some potential performance benefits for this technology in operational effectiveness terms.

This exercise although specific to the naval domain has illuminated many challenging systems engineering and implementation issues which remain to be addressed prior to wider acceptance of this technology in military command and control applications. Issues include understanding the role of the human and his/her means of interaction with the system. An effective solution to this challenge was implemented for the experiment described, however it is apparent that for the user to gain the full benefits which can accrue from the use of such systems, interface designs which will allow the operator to access, assimilate, understand and act upon the full richness of the probabilistic information available need to be further investigated. In conjunction, techniques need to be developed to allow the underlying BN to take account of the historical pattern of information relating to tracks and to present this to the user in an appropriate manner. Other challenges discussed include the validation and verification of probabilistic inference systems and the integration of such systems with information systems and network technologies which currently have little or any notion of uncertain data.

5.0 ACKNOWLEDGEMENTS

We would like to thank the following people for their involvement and support in this work:

BAE Systems Naval Ships	Geoff Slater, Paul Vangasse
BAE Systems Air Systems	Herman Claesen, Charles Patchett, Peter Wilkinson
BAE Systems C4ISR	Colin Smedley, Paul Wilmot, Ben Haigh, Ian Marshall
AMS (New Malden)	Dave Woof, Richard Bowker
BAE Systems ATC	Andy Wright, Joe Senior, Bill Wessel, Andy Page, Peter Mantell

6.0 REFERENCES

[1] Network Centric Warfare, Department of Defense Report to Congress, 27th July 2001, http://www.c3i.osd.mil/NCW

[2] Network Enabled Capability, UK MoD, JSP 777 Edn 1.

[3] M. Endsley, Towards a Theory of Situation Awareness in Dynamic Systems, Human Factors, **37,** 32-64, 1995.

[4] R.T. Cox, Probability, Frequency and Reasonable Expectation, Am. J. Phys., **14**, 1-13, 1946.

[5] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufman, 1988.

[6] F. V. Jensen, Bayesian Networks and Decision Graphs, Springer, 2001

[7] David E. Heckerman. An empirical comparison of three inference methods. In Proceedings of the Conference on Uncertainty in Artificial Intelligence, pages 158-169, San Francisco, CA, Morgan Kaufmann. 1988

[8] Eric Horvitz and Matthew Barry. Display of information for time-critical decision-making. In P. Besnard and S. Hanks, editors, Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, (UAI-95), pages 296-305, San Francisco, CA, 1995. Morgan Kaufmann.

[9] David Heckerman, John S. Breese, and Koos Rommelse. Decision-theoretic troubleshooting. Communications of the ACM, 38(3): 49-57, 1995.

[10] Bayesia. 6, rue Lonard de Vinci - BP0102 53001 Laval Cedex, France. http://www.bayesia.com.

[11] Data mining of large customer databases (Hewlett-Packard). http://crl.research.compaq.com/vision/multimedia/dm/default.htm.

[12] U.S. Department of Defense, Data Fusion Subpanel of the Joint Directors of Laboratories, Technical Panel for C3, "Data fusion lexicon," 1991.

[13] Multi-Sensor Kinematic and Attribute Tracking Using a Bayesian Belief Network Framework, In NATO Symposium on Military Data and Information Fusion (MP-IST-040), Prague, 20-22nd October 2003.

[14] A Bayesian Network for Combat Identification, S. P. van Gosliga and H. Jansen, In NATO Symposium on Military Data and Information Fusion (MP-IST-040), Prague, 20-22nd October 2003.

[15] S. Musman and P. Lehner, Real-time scheduling under uncertainty for ship self defense, IEEE Expert, Special Issue on Real-Time Intelligent Systems, 1999.

[16] P. Bladon and R.J. Hall. Information fusion using Bayesian multi-nets. In SPIE Defense and Security 2004: Image and Signal Acquisition and Processing, 2004.

[17] E. Santos Jr. Constructing adversarial models for threat/enemy intent prediction and inferencing. In SPIE Defense and Security 2004: Displays, Navigation and Simulation, 2004.

[18] J. Caroli, D. Fayette, N Koziarz, and T Stedman. Tools for effects based course of action development and assessment. In Proceedings of the Command and Control Research and Technology Symposium: The Power of Information Age Concepts and Technologies, San Diego, CA, 2004.

[19] Multilateral Interoperability Programme, LC2IEDM, http://www.mip-site.org

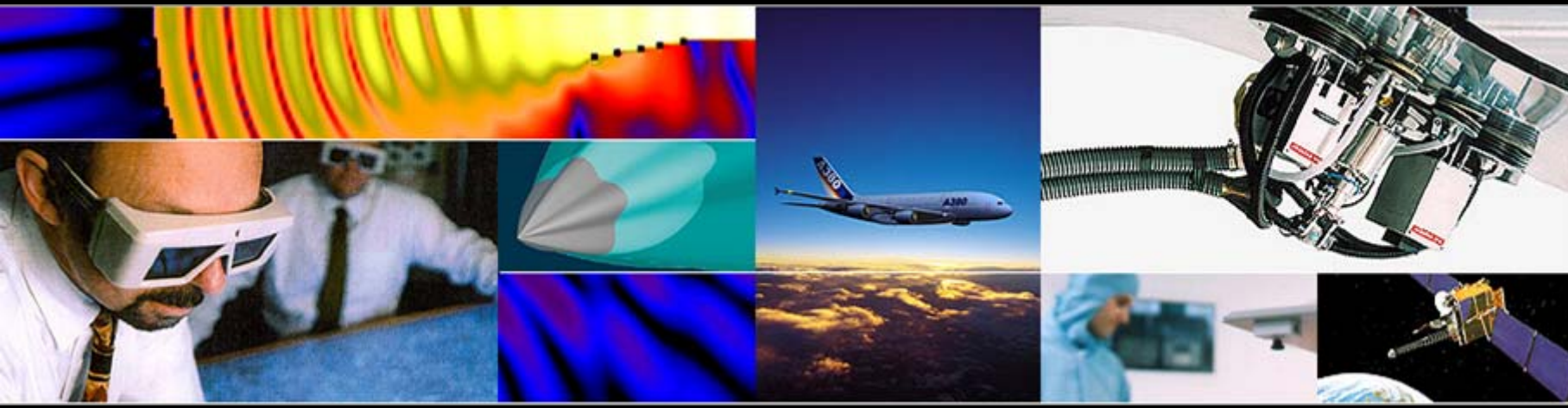
[20] NATO Standardization Agreement : Tactical Data Exchange – Link 16, STANAG 5516 Edition 3

[21] K.B. Korb and A.E. Nicholson, Bayesian Artificial Intelligence, Chapman & Hall / CRC Boca Raton, 2004.

BAE SYSTEMS

High-level fusion using Bayesian Networks: Applications in Command and Control

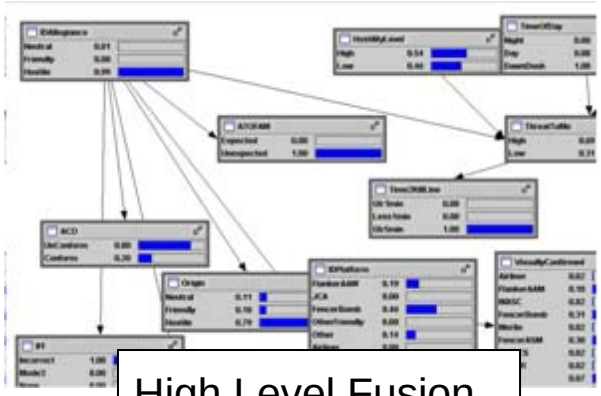
Peter Bladon, Peter,S. Day, Tim Hughes, and Philip Stanley



Overview

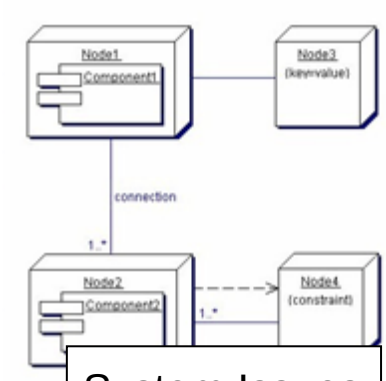
BAE SYSTEMS

The Future Battlespace…..



High Level Fusion

Experimental investigation

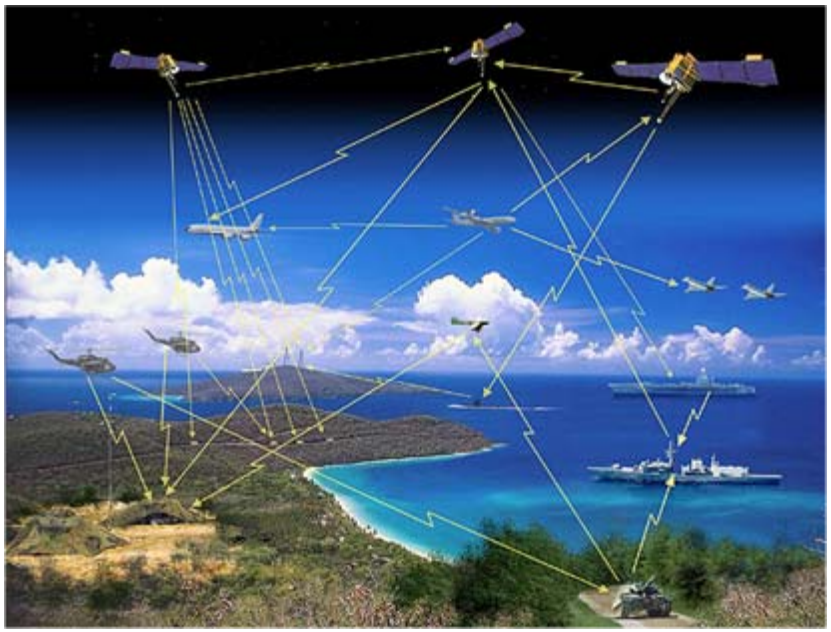


System Issues

Challenges of NEC

BAE SYSTEMS

Networking the battlespace provides the opportunity for the tempo and effectiveness of military operations to be increased



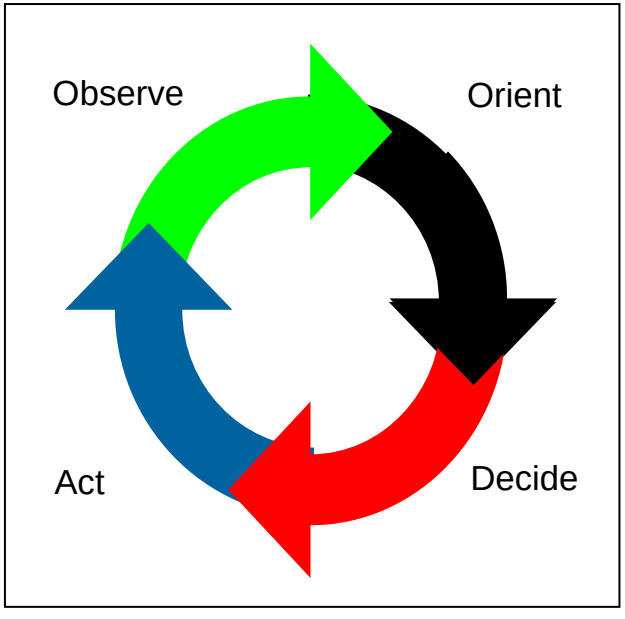
- *"Networked integration of sensors, decision-makers and effectors."*
- *"Timely provision and exploitation of information and intelligence to enable effective decision making and agile actions."*
- *Information Superiority*
- *Decision Superiority*
- *Shared Situation Awareness*

NEC - JSP 777 -Edn 1

Challenges of NEC

BAE SYSTEMS

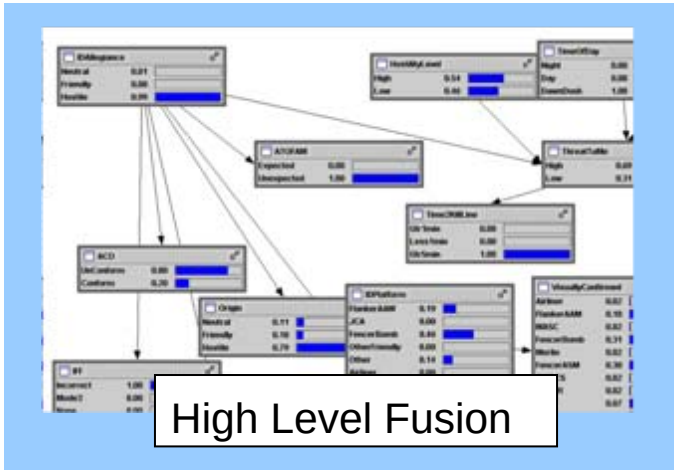
- **How do we increase the tempo of decision making?**
  - Provide technology to support decision making process
- **What technology?**
  - **High-level fusion** to support **Orient** and **Decide** elements of OODA
- **But:**
  - What are requirements of technology?
  - How should it be used?
  - What is its impact on other systems and processes?



Overview

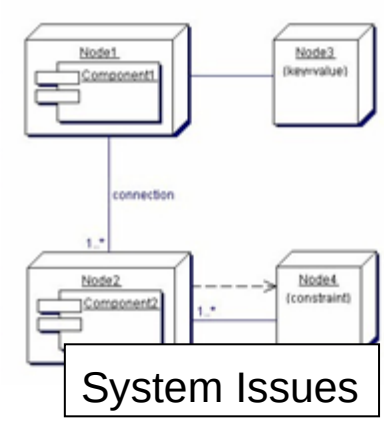
BAE SYSTEMS

The Future Battlespace…..



High Level Fusion

Experimental investigation



System Issues

High Level Fusion

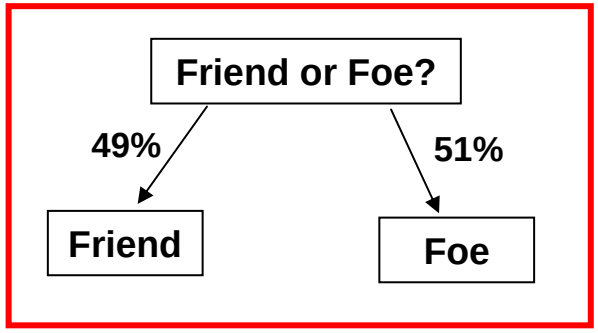
BAE SYSTEMS

- What is ‘high-level fusion’?
  - the process of placing observations in context with existing knowledge
- Information sources can be many and varied:
  - Data from sensors of different types
  - Intelligence information of varying trustworthiness
  - Background / historical data stored in databases
  - Expert knowledge
- All contain inherent **uncertainty**

High Level Fusion

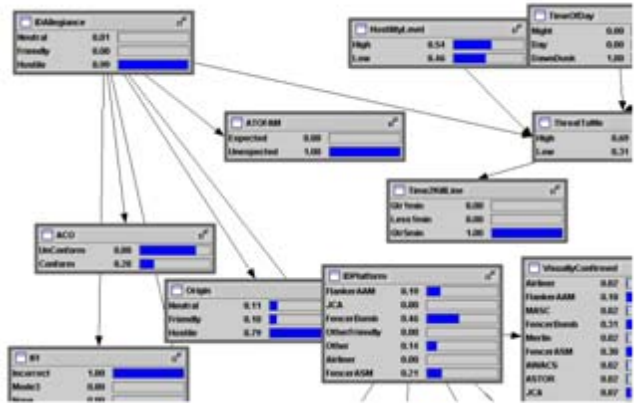
BAE SYSTEMS

- Decision making
  - must take **uncertainty** into account from **input** through to **decision**
  - as making **hard decisions** (too soon) can be dangerous



- What technologies are there?
  - to represent and manipulate uncertainty
  - to model complex domains
  - to encode prior knowledge

Bayesian networks

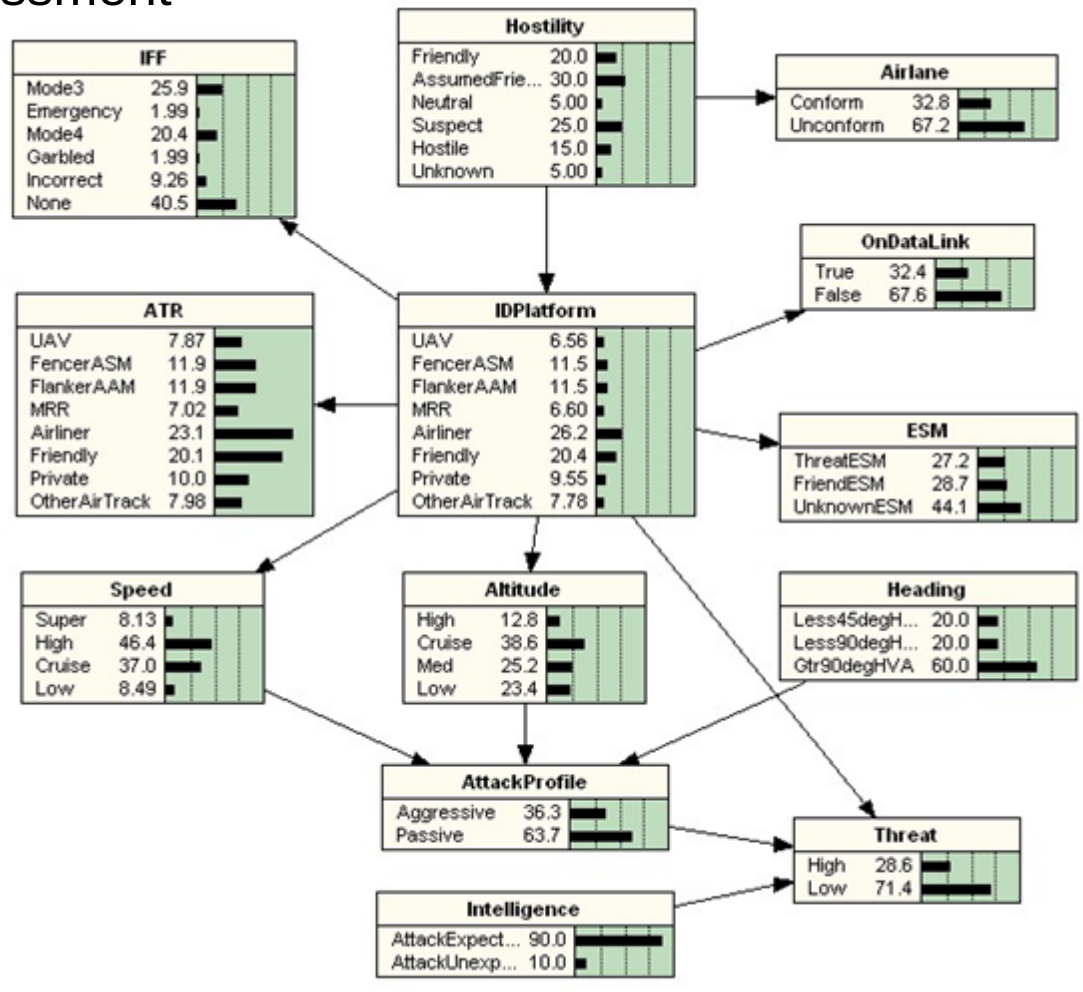


Bayesian Network Example

BAE SYSTEMS

Combat ID and Threat Assessment

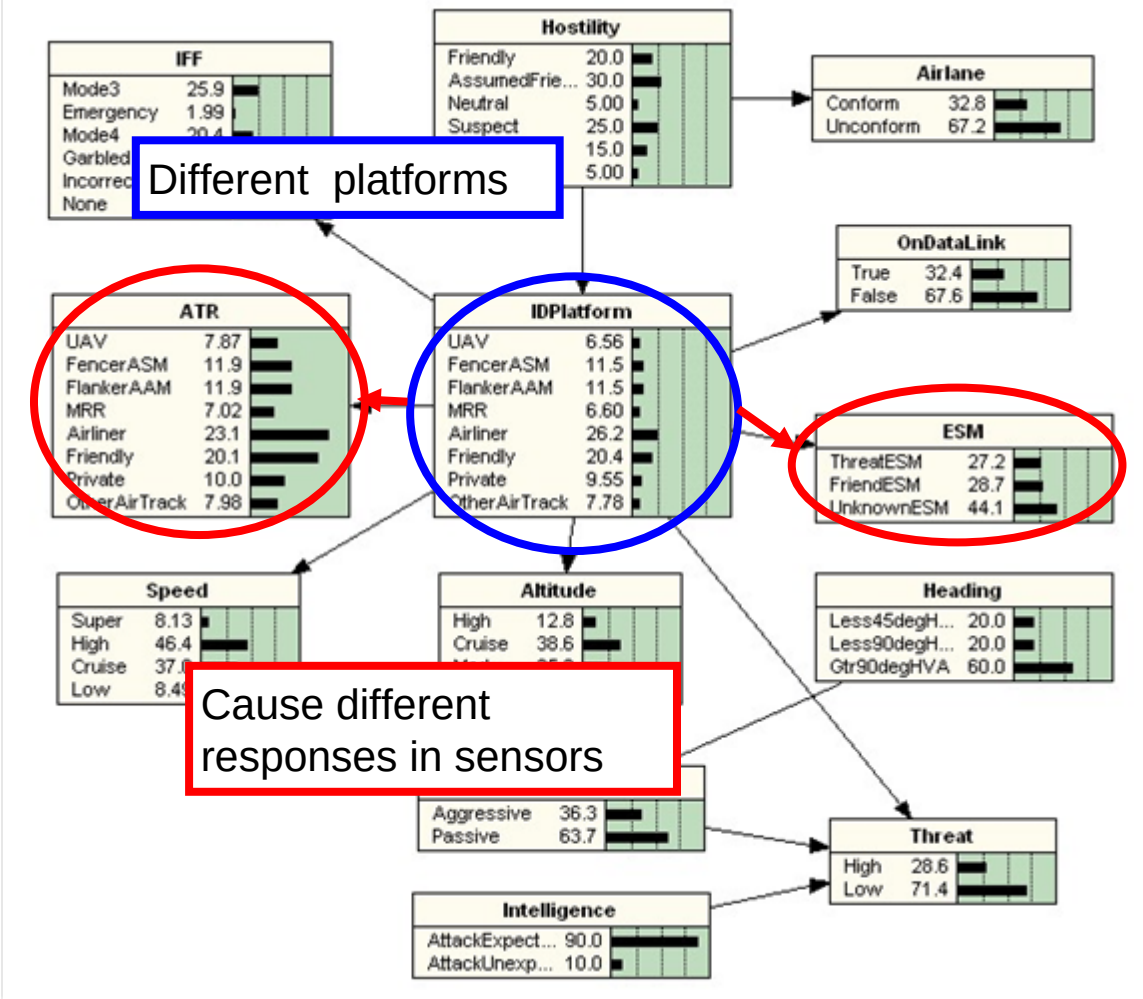
- Model includes
  - The behaviour of each track with respect to a defended asset
  - background information about the scenario and platform capabilities
  - other ‘contextual’ information sources



Model Construction

BAE SYSTEMS

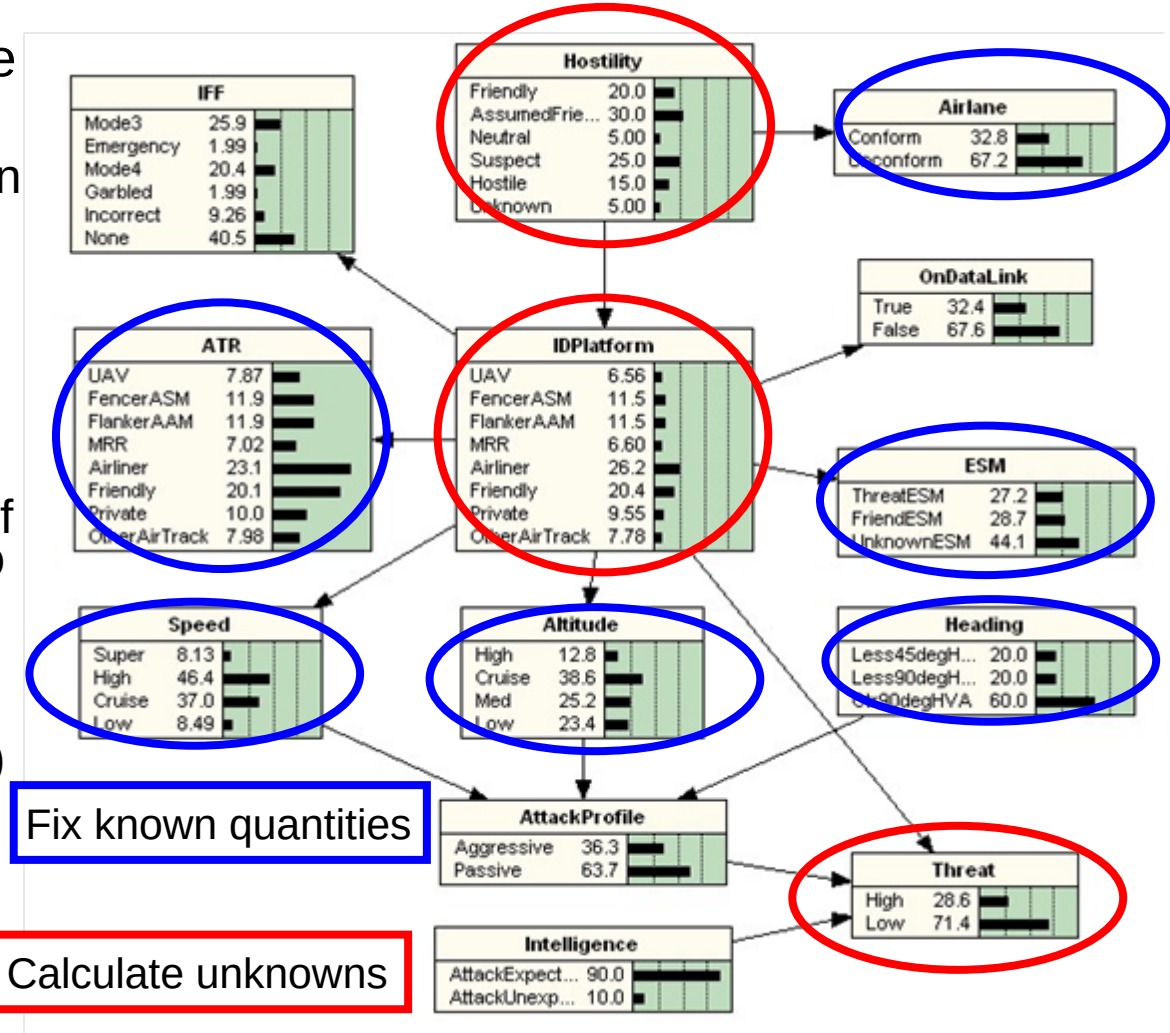
- Knowledge Elicitation
  - 'Cause and effect' models can be built using domain experts
  - Allows definition of abstract quantities such as Threat
- Machine Learning
  - Models can be learned from data
  - Models can be refined using data



Using Models

BAE SYSTEMS

- Enter available evidence
  - Sensor readings
  - Background information
- Use ‘inference’ for:
  - Probability of ID given evidence
  - Most likely ID identity
  - Most important piece of evidence supporting ID hypothesis
  - What next best measurement I can make to increase P(ID)
  - How sensitive are conclusions of P(ID) to evidence



Advantages and Challenges

BAE SYSTEMS

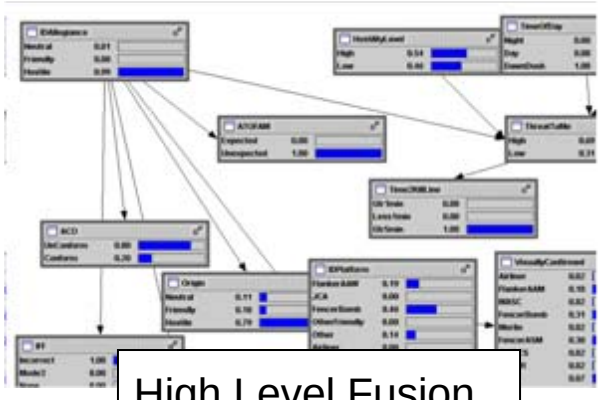
- Advantages
  - Generality of modelling approach
  - Ease of integration with existing probabilistic systems:
    - Tracking
    - ATR
  - Flexibility of inference algorithms
    - Current state of knowledge can be used for sensitivity analysis and system control
  - Fine grained auditability

- Challenges:
  - Model construction
  - Deployment and use

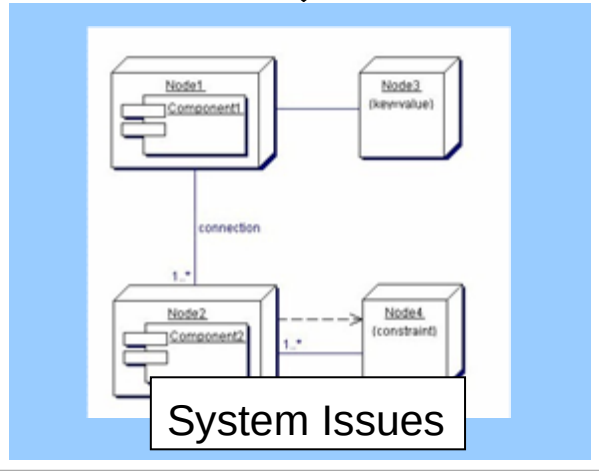
Overview

BAE SYSTEMS

The Future Battlespace…..



High Level Fusion



System Issues

Experimental investigation

System Issues

BAE SYSTEMS

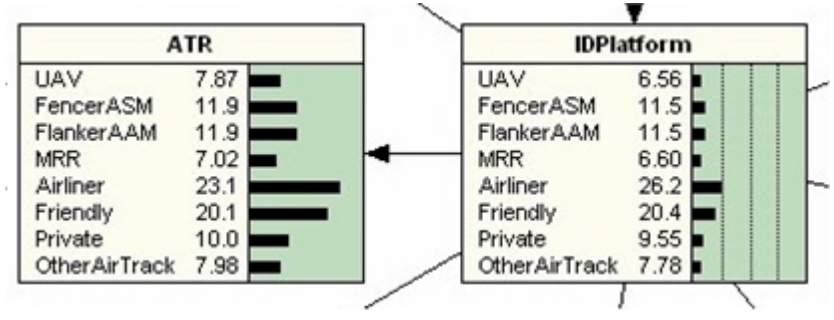
- Validation and verification
- Integration with information systems and network technologies
- Interaction with the human operators
- Effect on processes and doctrine

System Issues

BAE SYSTEMS

- Validation and verification
  - Models and model components (i.e. sub-systems) can be verified where data is available

ATR Sub-system



  - **'Expert derived'** components (e.g. Threat) can only be verified by analysing System **Measures of Effectiveness**

- Integration with information systems and network technologies
- Interaction with the human operators
- Effect on processes and doctrine

System Issues

BAE SYSTEMS

- Validation and verification
- Integration with information systems and network technologies
  - Data models (e.g. C2IEDM) / data link protocols (Link 16) / displays typically have no (or limited) notion of uncertainty
  - Everything in system is presumed to be true
  - Reflected in need for manual conflict resolution procedures
  - How do we modify legacy systems to incorporate uncertainty?
- Interaction with the human operators
- Effect on processes and doctrine

System Issues

BAE SYSTEMS

- Validation and verification
- Integration with information systems and network technologies
- Interaction with the human operators
  - Do the operators need to know about ‘uncertainty’ of information?
  - If so what do they need to know?
  - How should interfaces be designed to communicate decision support information?
  - How should operators be trained to understand and use this information?
- Effect on processes and doctrine

System Issues

BAE SYSTEMS

- Validation and verification
- Integration with information systems and network technologies
- Interaction with the human operators
- Effect on processes and doctrine

  - Current processes / doctrine is based on capabilities of existing systems

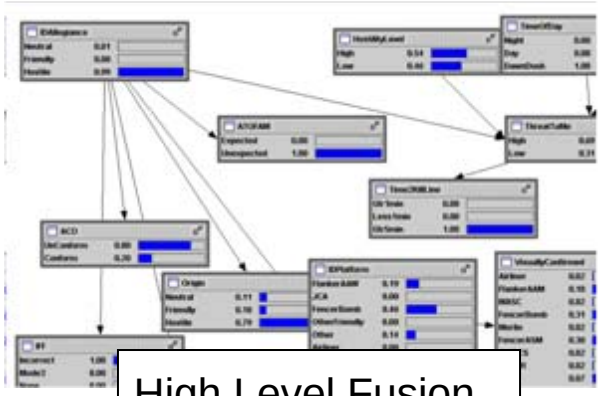
  - New technologies allow new processes and team roles

  - What should they be, and how do we derive them?

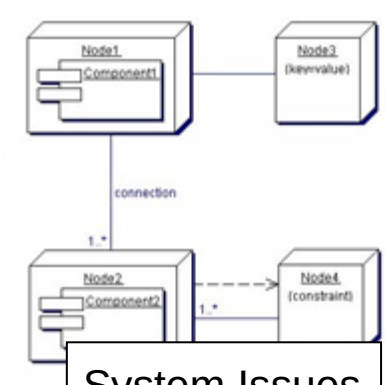
Overview

BAE SYSTEMS

The Future Battlespace…..



High Level Fusion



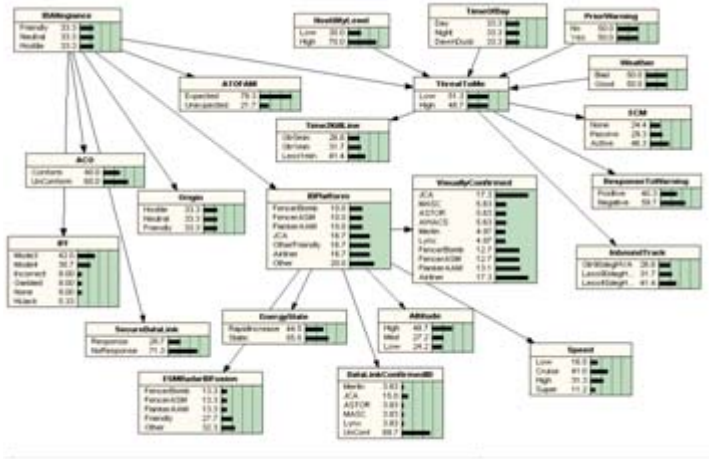
System Issues

Experimental investigation

Experimental Investigation

BAE SYSTEMS

- An experimental investigation undertaken with the BAE SYSTEMS Delivering NEC programme has allowed exploration of some of these system issues
- Referred to as ‘Shared Situation Awareness’ (SSA) experiment
- Based on Combat-ID / Threat Assessment task
  - Similar to example shown earlier
  - But with a ‘richer model’

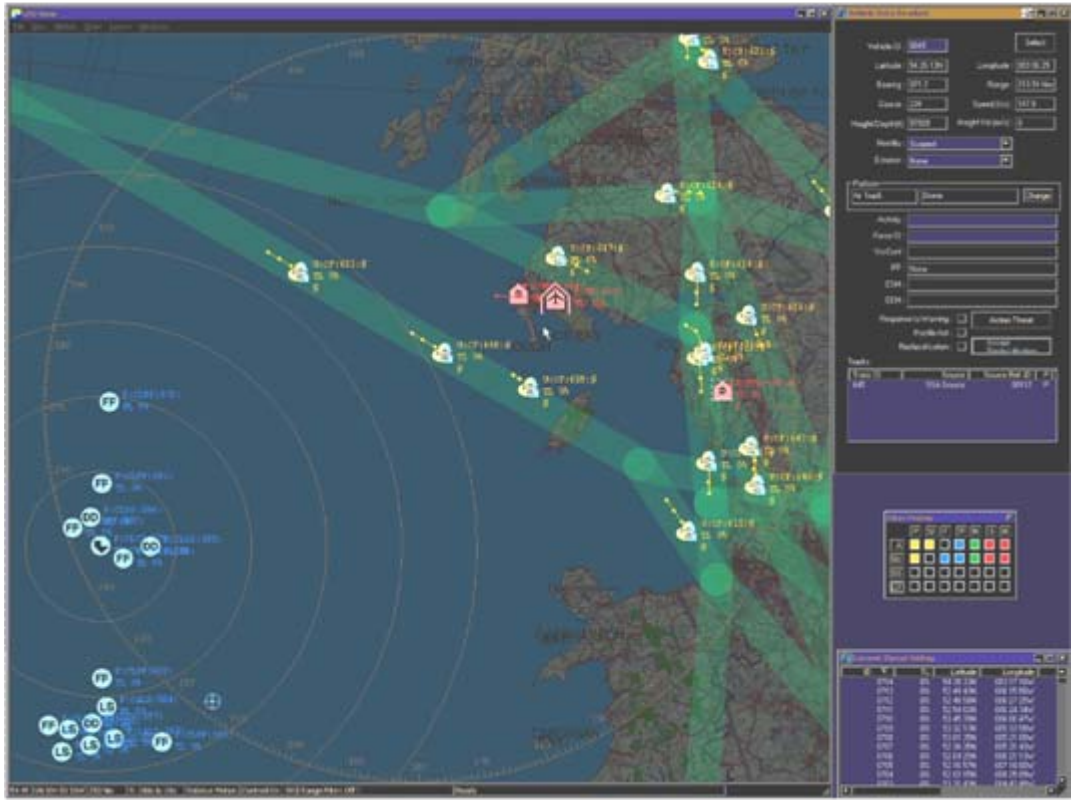


- Experimental goal:
  - assess operator benefit of tool in a naval picture compilation task
- Model designed using
  - Expert knowledge from domain experts with Naval and Air-Force picture compilation expertise
  - Extensive scenario analysis based on DoDAF framework

Implementation Concept

BAE SYSTEMS

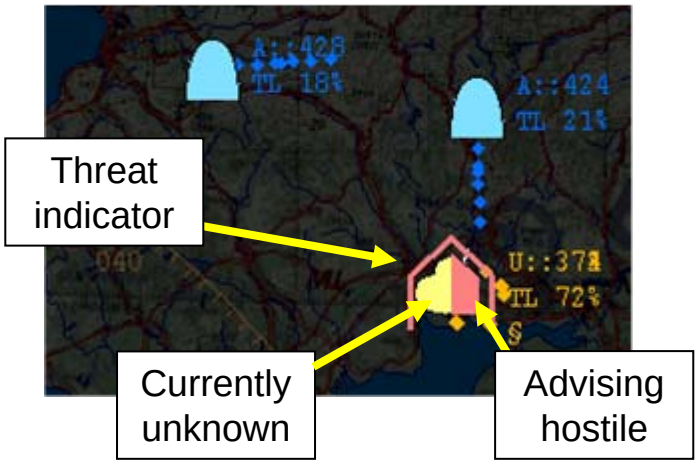
- Bayes Net (BN) used as a decision aid incorporated into a modified naval command system
- Advisory, not automated system
- Automatically updated recommendations from the BN presented to the operator for assessment
- An adjunct to the presentation of the underlying data (altitude, velocity, IFF etc)
- Display of probabilistic data and Bayes Net reasoning kept to minimum in this implementation



Human Machine Interface

BAE SYSTEMS

- Recommendations from SSA appear in symbol, track label, VDR



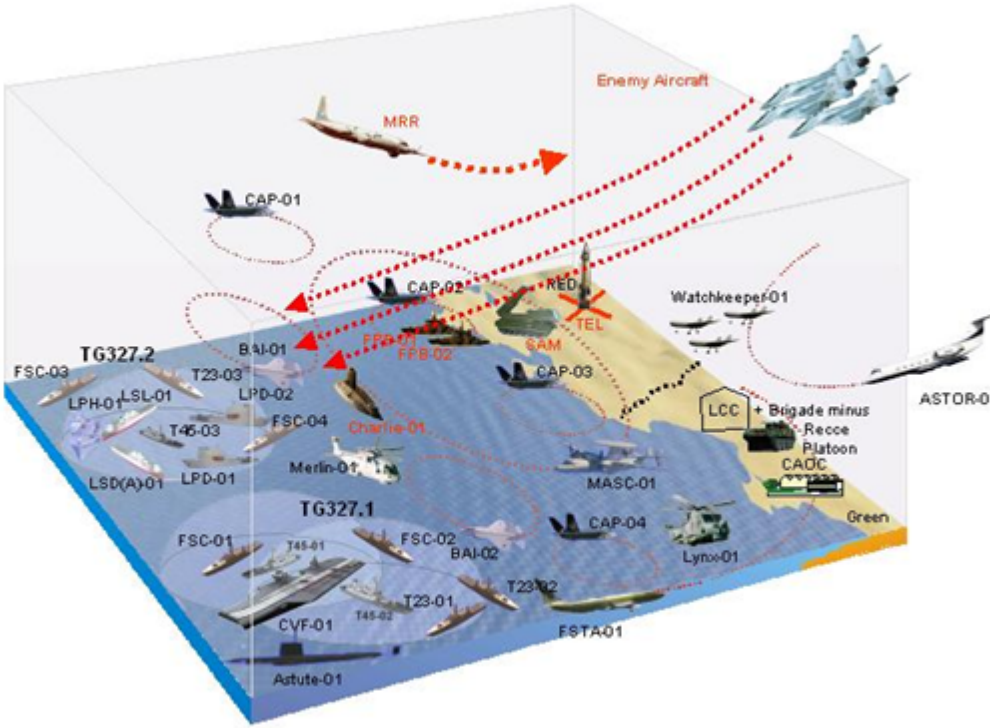
Detail of modified
MIL-STD-2525 symbology

- Additional interface components permitted direct acceptance/rejection of SSA advisories

Experimental Task

BAE SYSTEMS

- Compile the Task Force Air picture
- For each command system track, establish:
  - Its ID allegiance/hostility
  - Its Platform ID (Type of Platform)
  - Its level of threat
- Other tasks e.g. track management, not simulated



Experimental Design

BAE SYSTEMS

Design

– Subjects
  - 14 operators with picture compilation experience (11 RN, 3 RAF)

– Factors
  - Two timelines, two conditions (with and without SSA)
  - Experimental manipulation within subjects and between timelines in balanced order

Procedures

– Briefing and Training (2 hours)
  - HMI
  - Task
  - Scenario

– Experimental test (4 hours)
  - Two experimental sessions

– De-brief (1/2 hour)

Measures

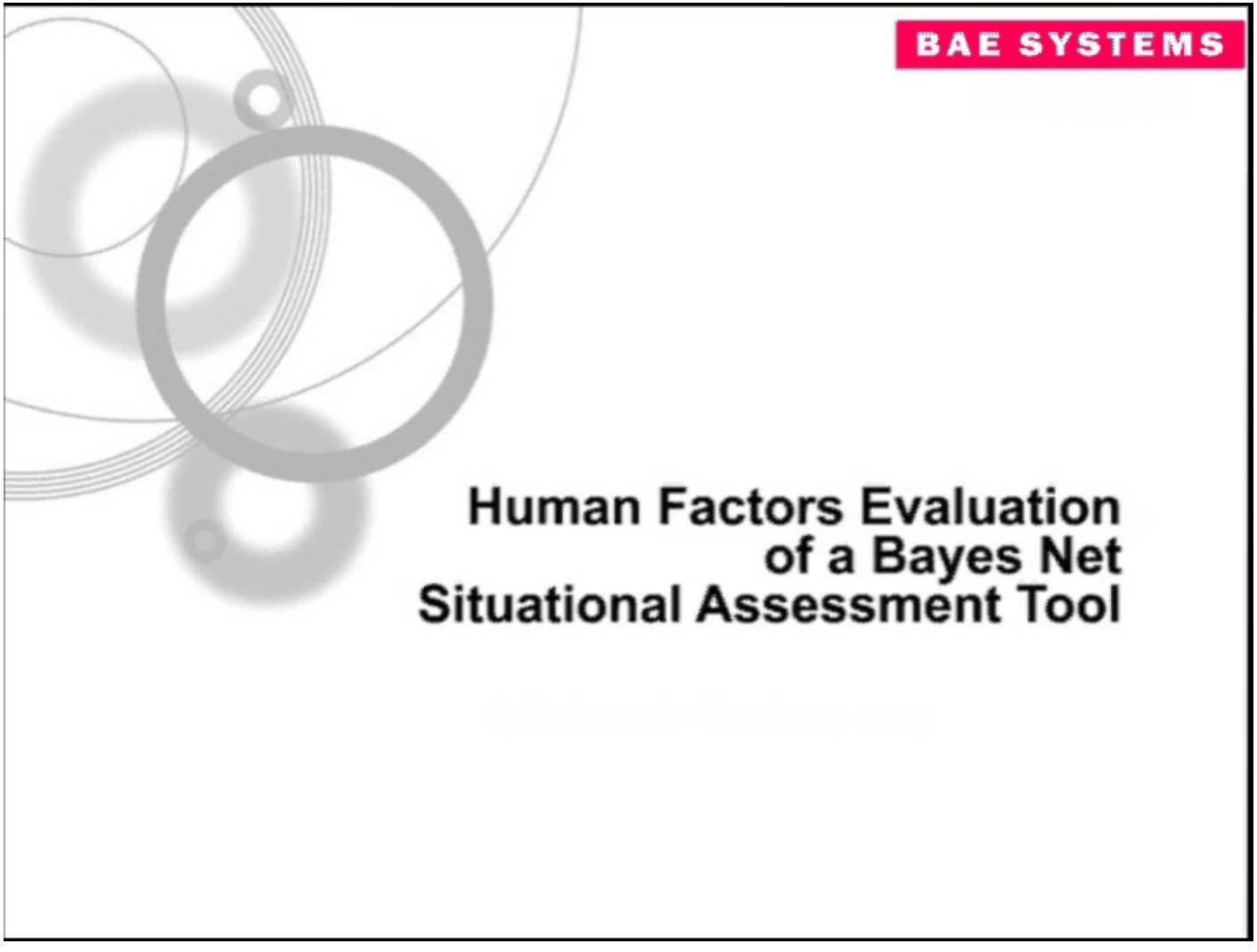
– Objective e.g.
  - Time to correct ID Allegiance
  - Time to correct ID Platform
  - Time to Action Threat
  - Number of tracks validated

– Subjective
  - Questionnaire
  - SA and Workload ratings

– Video records

Video Summary

BAE SYSTEMS



Results

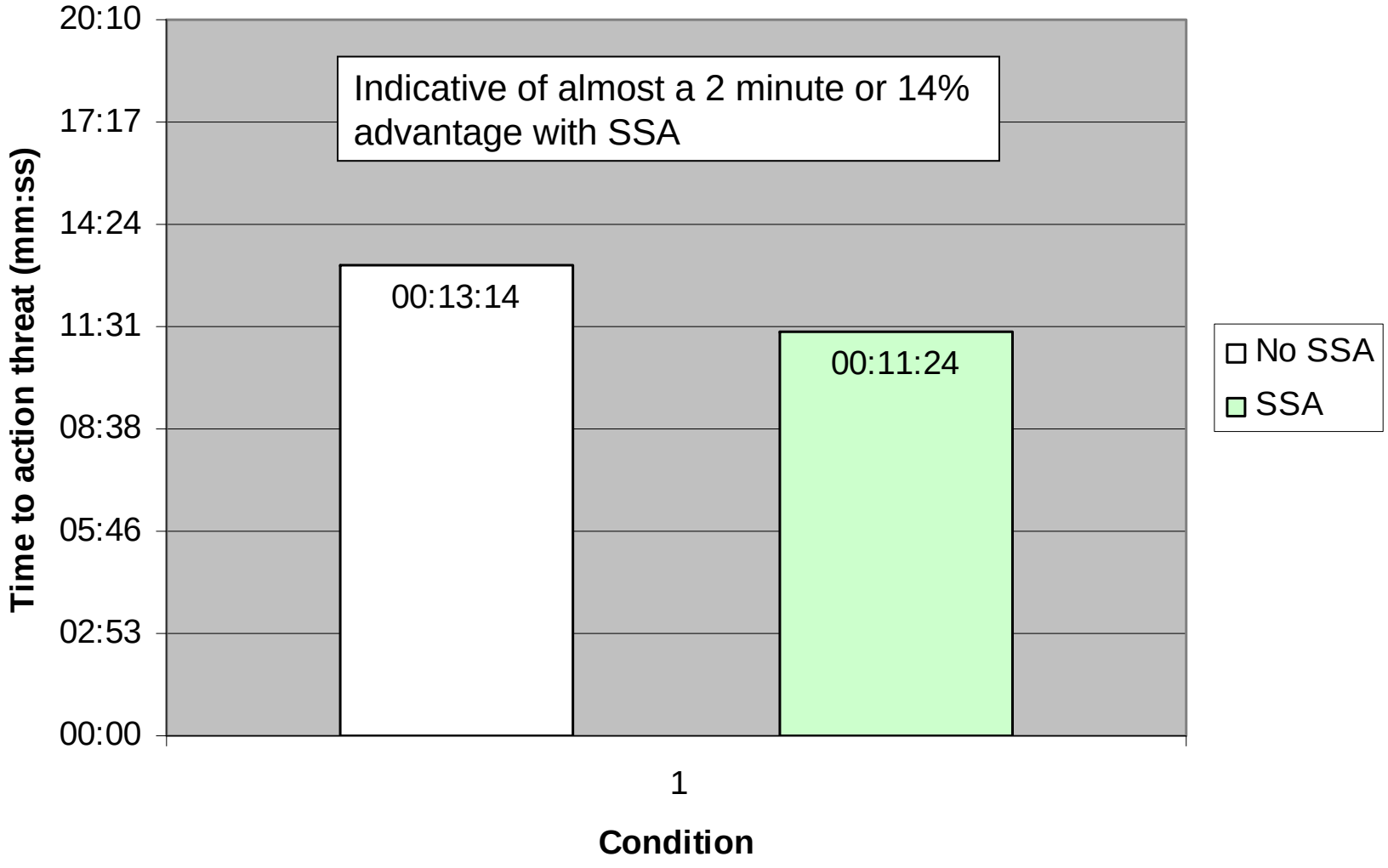
BAE SYSTEMS

- Quantitative Results
  - Time to respond to threat
  - Time to ID
  - Number correct IDs
  - Workload

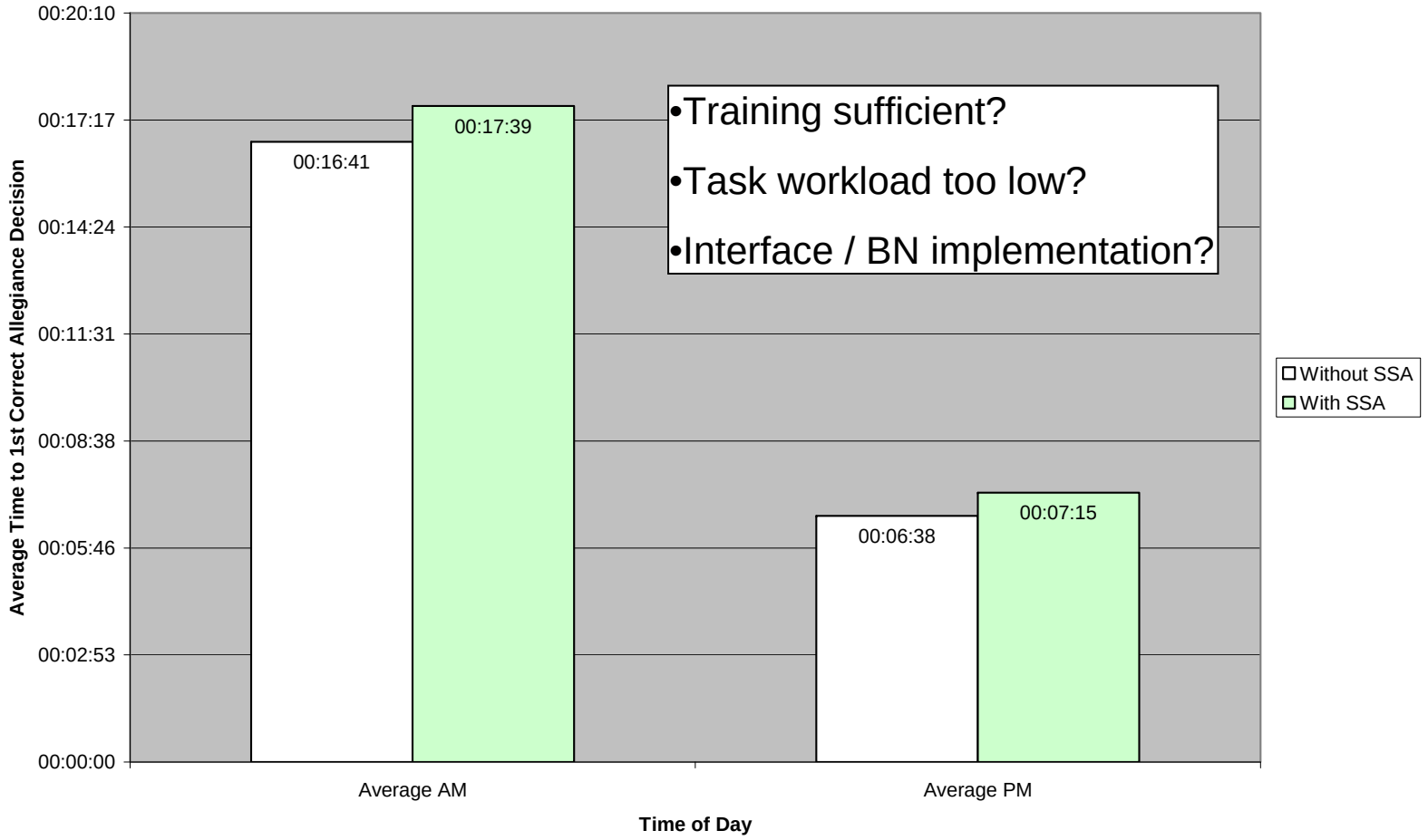
- Qualitative
  - Subjective measures extracted from Questionnaire

Average Time Taken to Respond to Threats

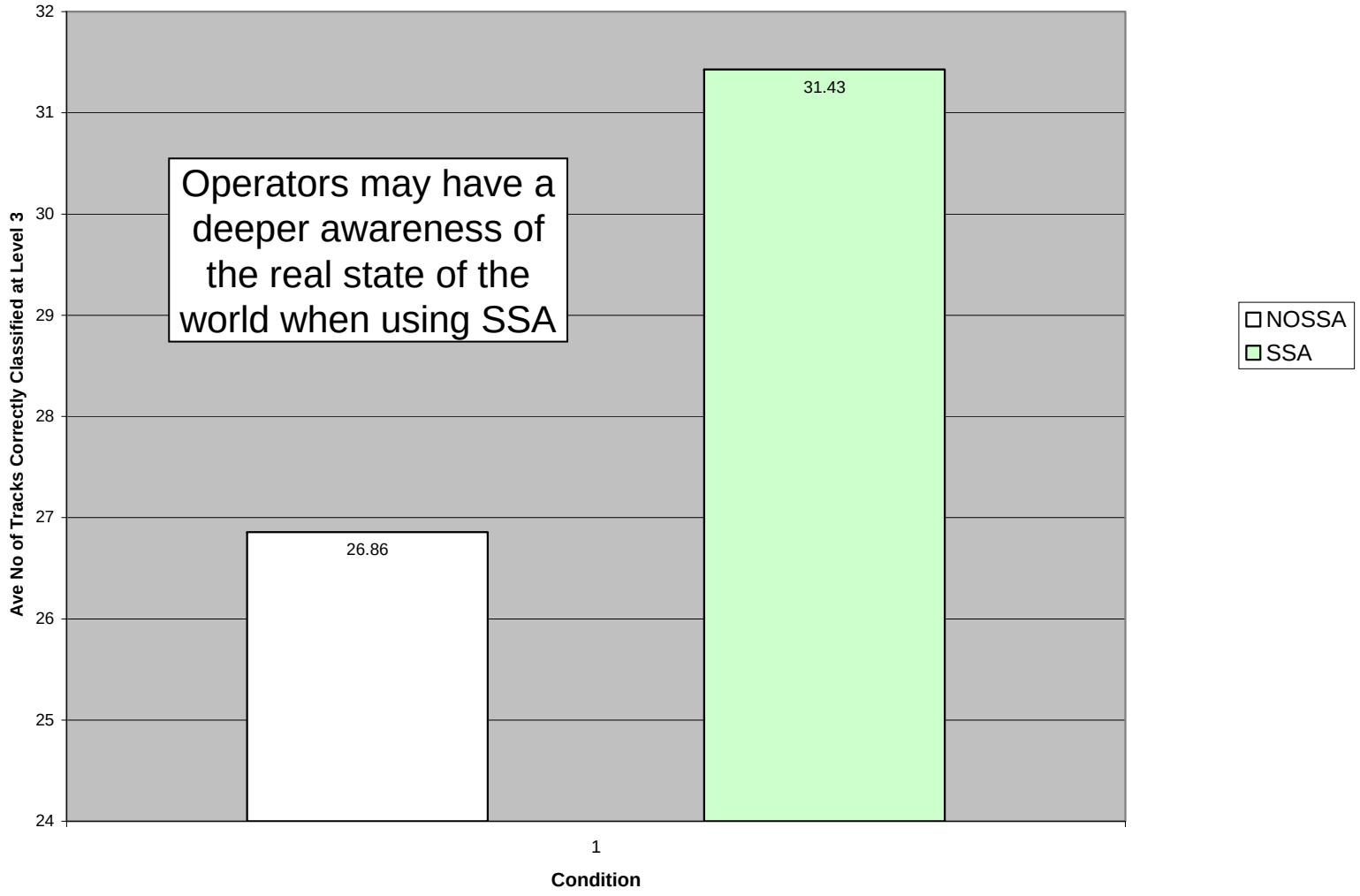
BAE SYSTEMS



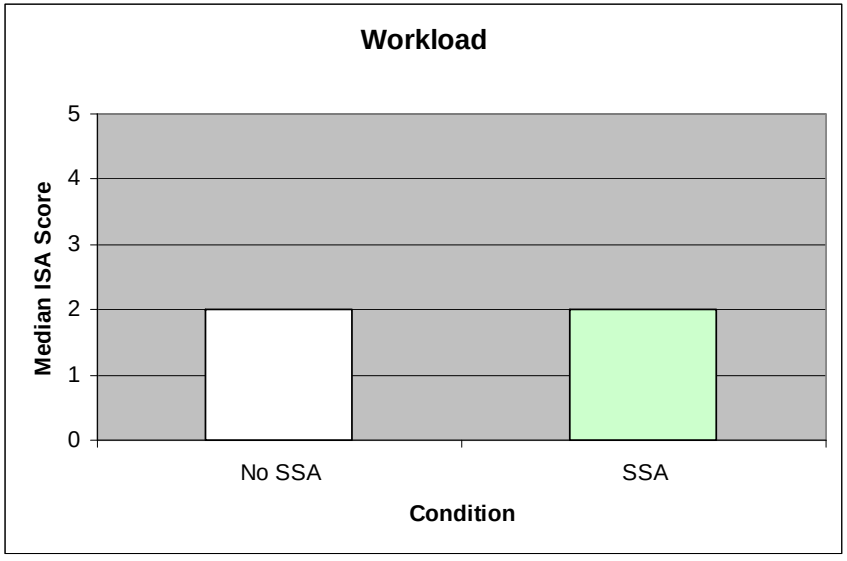
Average Time to Operator's First Correct Identification of ID Allegiance



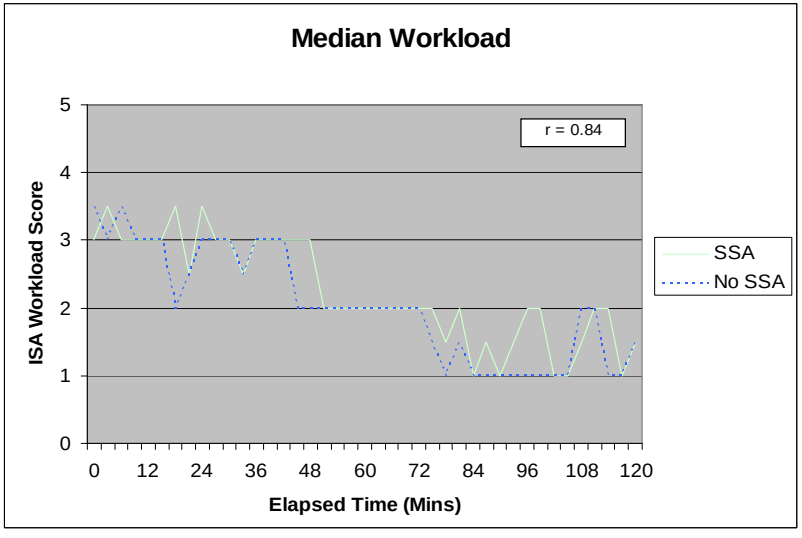
Number of Correct Platform ID Classifications



Operator Workload



Introduction of SSA had no negative impact in terms of operator workload

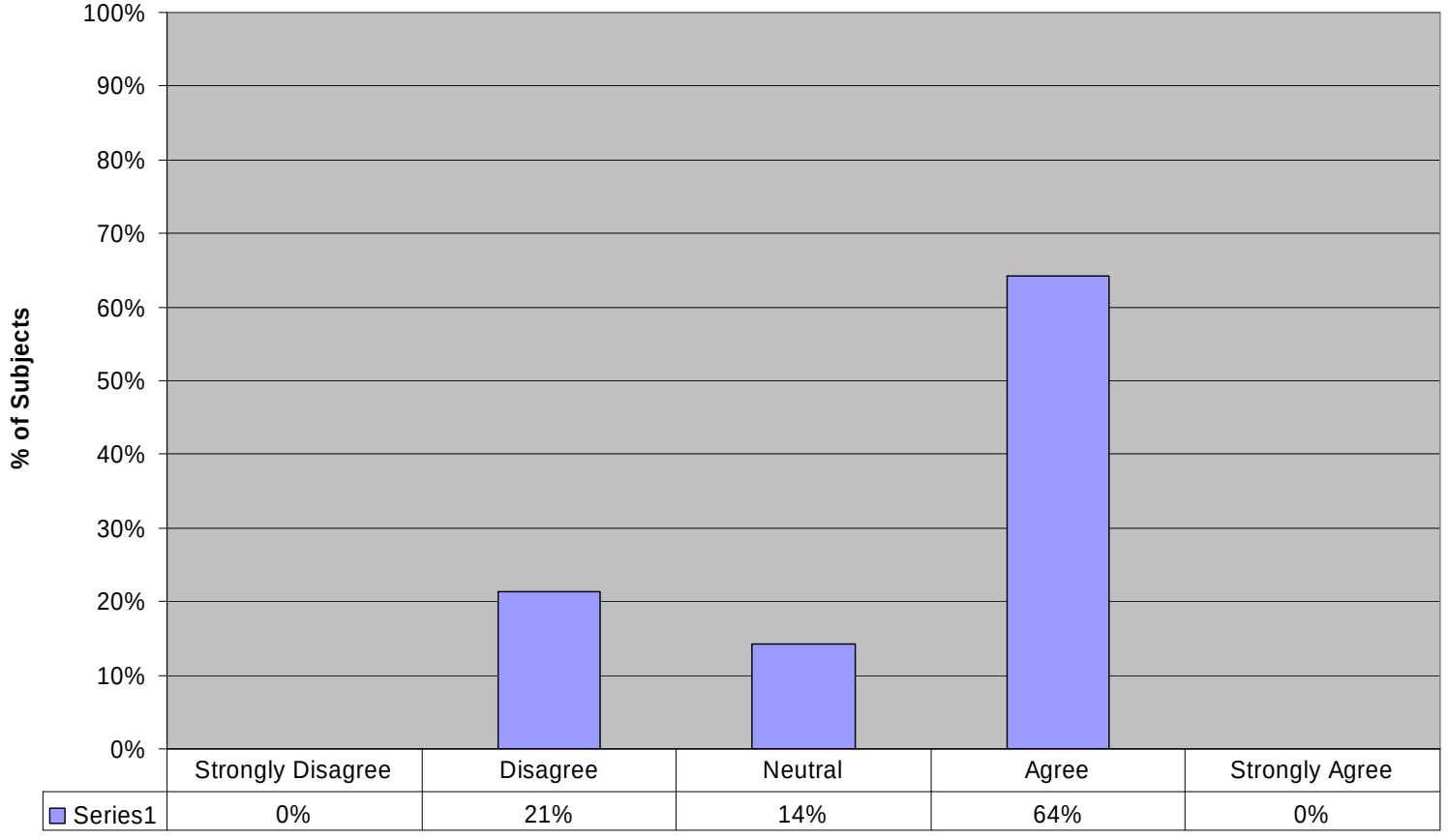


Extract from Questionnaire Results

BAE SYSTEMS

N = 14

2.7. I always understood how the SSA system came up with its recommendations.

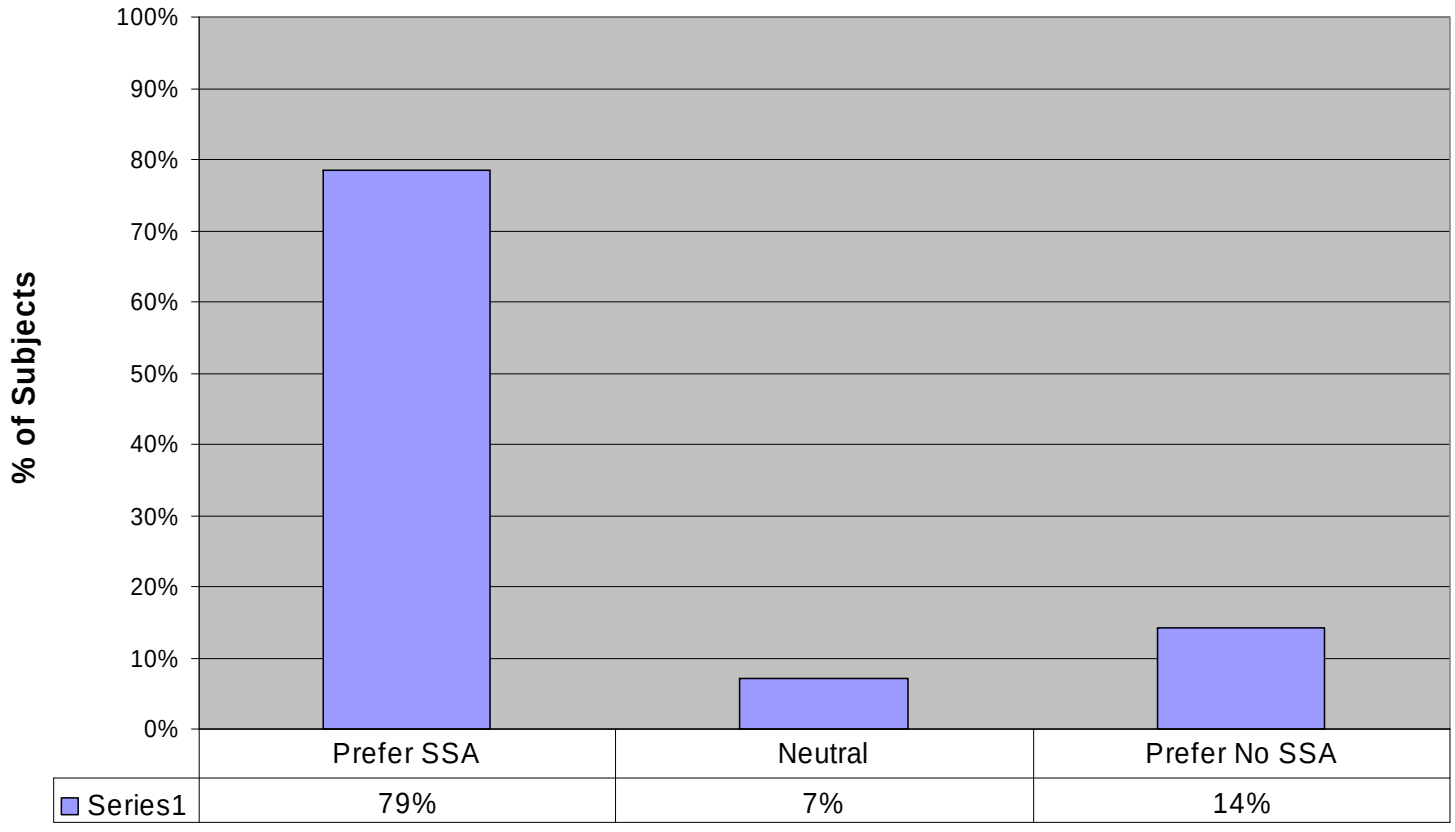


Extract from Questionnaire Results

BAE SYSTEMS

N = 14

3.5. Which system did you prefer, SSA or No SSA?



Operator Comments About the SSA Technology

BAE SYSTEMS

- *"…Platform change suggestions + threat recommendations were all very useful. SSA has good potential."*
- *"Less thought required (with SSA). More time to concentrate on the tactical environment, i.e. Track(s) release to data links, etc."*
- *"It was good to have a computer that thinks for you and is there to help, not just do!"*
- *"Allows you to leave less important track investigation till later. Gives immediate heads-up on possible hostiles.".*
- *"Faster classification of contacts allowed concentration on suspects more quickly. Clear indication of hostility recommendations and threat levels."*
- *"In many aspects, it is better to be presented with a solution which can be vetoed, rather than trying to find the answers.*
- *"Having advice based on objective criteria provided reassurance about classification decisions and made the decisions quicker to take…"*

Summary of Experimental Findings

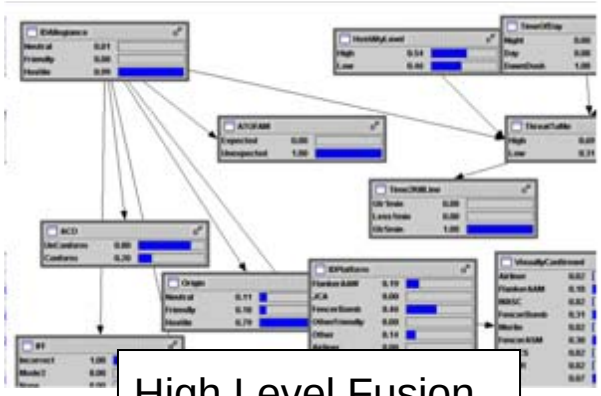
BAE SYSTEMS

- Potential improvements in operational effectiveness demonstrated
  - SSA provided a 14% time advantage for threat identification
  - 14.5% more tracks identified to a lower level with SSA
- Operators thought that the SSA system assisted their performance on the task
- Some problems with the implementation of the system were acknowledged but the majority of operators understood the decisions made by the technology and approved of its means of presentation
- Operators overwhelmingly preferred the SSA system
- Operators could identify potential benefits following further refinement of the BN and HMI

Overview

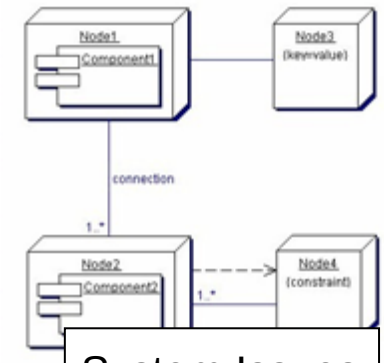
BAE SYSTEMS

The Future Battlespace…..



High Level Fusion

Experimental investigation



System Issues

Summary

BAE SYSTEMS

- The future battlespace implies an increasing need for decision support technologies
- BAE Systems has a high level fusion technology based on Bayesian Networks that can support human decision processes in a NEC environment
- Potential performance benefits for this technology have been demonstrated in a realistic, simulated, naval scenario
- Challenging systems engineering and implementation issues remain

Acknowledgements

Programme:

- Delivering NEC – BAE SYSTEMS ICP – Programme Manager - Geoff Slater

Experimental Team:

- Scenario development - Paul Wilmot & Charles Patchett
- SMEs for knowledge elicitation - Paul Wilmot & Charles Patchett
- Provision and modification of CNS - Dave Woof & Richard Bowker
- HMI & experimental design - Tim Hughes, Bill Wessel & Philip Stanley
- SSA Tool - Peter Bladon, Joe Senior, Andy Page, Peter Mantell
- Integration - Peter Bladon, Joe Senior, Andy Page, Tim Hughes & Bill Wessel
- Experimental facilities - Ben Haigh, Ian Marshall, Colin Smedley, Dave Woof, Richard Bowker, Alan Collinson
- Execution of the experiment - Tim Hughes, Bill Wessel, Philip Stanley, Colin Smedley & Paul Wilmot