

Swedish-Singapore studies of Bayesian Modelling techniques for tactical Intelligence analysis

Pontus Svenson, Robert Forsgren, Birgitta
Kylesten, Peter Berggren

FOI Swedish Defence Research Agency
Sweden

pontus.svenson@foi.se

robert.forsgren@foi.se

birgitta.kylesten@foi.se

peter.berggren@foi.se

Wong Rong Fah, Magdalene Selina Choo,
John Kho Yew Hann

Defence Science and Technology Agency
Singapore

wrongfah@dsta.gov.sg

magdalene_CHOO@nscs.gov.sg

kyewhann@dsta.gov.sg

Abstract - In this paper we present the results of an experiment, which combined the Impactorium tool with the RAHS System. We describe the intelligence analysis process supported by the tools and give a brief introduction to the tools. Results from a Swedish study that influenced the version of Impactorium tested in Singapore are reported. We conclude with some suggestions for future studies of Bayesian belief modelling for intelligence analysis.

Keywords: intelligence analysis, situation monitoring, threat assessment, bayesian belief networks.

1 Introduction

In this paper, we considered the problem of situation monitoring and threat prediction by intelligence analysts in a 6-month terrorist attack scenario, and report the results of a joint experiment undertaken by Sweden and Singapore in 2007. The experiment aimed at testing the situation awareness of the analysts and their decision-making.

Intelligence analysis is a difficult problem, which can benefit from computer support tools, provided that these are constructed in a way that takes into account how humans work. While some minor components (e.g., entity extraction in text) of the intelligence analysis process are very suitable for automation, the analysis of the output is mainly a human task. In this paper, we describe some of the experimental studies that have been made with the Impactorium tool developed at FOI [1],[2]. The main focus is on the experiment performed in Singapore in 2007.

The paper is outlined as follows. Section 2 briefly describes intelligence analysis and decision-making and is followed by a description of how Bayesian belief networks could be used to help intelligence analysts achieve situation awareness and improve their decision-making. Next, the RAHS and Impactorium systems are briefly described. This is followed by a section on the experiment in Singapore and a brief discussion.

2 Background: Intelligence Analysis

Intelligence analysis is used to provide adequate situation awareness for decision-making, but also involves decision-making itself. Situation awareness [3] is the perception, comprehension and near-future projection of the environmental elements making up a situation. In terms of information fusion, it is mainly related to situation assessment, level 2 of the JDL model [4]. For intelligence analysis, the meaning of a situation varies from setting to setting. Generally, it is important to have an understanding of the basic entities making up a situation as well as their relations. For many applications, it is the network of connections between events, persons, and objects that are of interest. Computer tools for helping people keep track of such networks are often very useful.

Decision making is the process of maintaining a mental representation and the navigation procedures of the problem space (i.e. the current state of affairs, the goal state, and availability of resources), while taking a leap of faith in making a choice to embark on a particular course of action to solve the problem. Algorithmic methods of problem solving and the utility theory of decision-making suggest that people approach this process systematically: they would consider most, if not, all aspects of the problem, make choices between alternatives via weighing the benefits and costs of decisions and selecting the decision that maximises the outcome benefits and minimises the costs [5].

However, research done with expert military personnel, fire fighters [6] and design engineers [7] showed that experts made decisions by comparing the current situation to previously experienced situations. They identified appropriate courses of action via mentally simulating the typical course of action for similar situation to evaluate the particular solution's viability in the current situation [7]. Experience with multitudes of similar types of problems and their associated solutions exposed the experts to various facets of the problem and thus enabled them to develop elaborate schemas of the problems, which prove

useful as reference points for future decisions. This method of decision making seems to facilitate time-constrained decisions with ill-defined goals, and ambiguous data [5], [6].

3 Bayesian modelling techniques

Algorithm and utility based decision support systems, of which the Bayesian technique modelling tools is an example [6], have been shown to be effective in augmenting humans when they are making decisions that have stable problem spaces, requires less expertise, under less time pressure, and affects many stakeholders [8].

Bayesian technique modelling is a comprehensive representation of cause and effect or correlational relationships between factors depending on the user's representation of the environment. This technique has been used to identify enemy units and installations [9], generate alternative hypotheses [10], in order to facilitate the development of courses of action and centre of gravity by commanders. For example, it has been used to represent causal relationships in the military domain for the prediction of mine impact and enemy troops. Under the Mine Burial prediction program, the Office of Naval Research, USA, developed the Mine Burial Expert System Model (MBESM) which is a Bayesian network model to simulate the impact of different types of mines in particular deployment sites [11]. While the Centre of Gravity Network Effects Tool (COGNET) is a modelling framework developed by Defence Science and Technology Organisation, Australia, to determine enemy and friendly COG. The model is a Bayesian representation of potential factors that estimate an enemy's COG [12].

On the other hand, Bayesian technique modelling has yielded mixed feedback in the intelligence domain. In a study comparing intuitive versus Bayesian probability, six analysts' were told to determine the likelihood that the USSR would attempt to destroy China's growing nuclear capabilities within the month [13]. In the intuitive probability condition, analysts would simply generate a figure of likelihood of the event occurring, while in the Bayesian probability condition, analysts would list the indicators of the events, its probability of occurrence and its conditional probability on the event's occurrence, but allowed the mathematics to determine the figure of likelihood of the event occurring. Findings showed that the likelihood figure determined through Bayesian probability was lower than the intuitive probability figure and was thus more accurate in estimating the likelihood of the event occurring as USSR did not attempt to destroy China's growing nuclear capabilities within the month. A similar exercise was conducted to evaluate the probability of various critical consequences occurring in the event of failed peace talks between Egypt and Israel. Yet, with re-evaluation thirty years on, the researchers are skeptical about the accuracy of the conditional probabilities that were generated by the analysts: they were not evaluating the probabilities of outcomes occurring conditional on the

success or failure of the peace talks instead they were evaluating the casual impact of the success or failure of the talks [14] suggesting that the conditional probabilities generated did not address the issue at hand. Critics have also cautioned that the Bayesian technique interface of displaying all the components of a situation might obscure the big picture understanding required by experts to make sense of the situation [7].

The present research aims to evaluate the effectiveness of the use of Bayesian technique on analysts' quality of situation awareness and decision making when performing threat assessment.

4 RAHS and Impactorium

4.1 RAHS System

The RAHS system is built on Service Oriented Architecture (SOA), which simplifies the integration of data and tools from different agencies as web services. There are three major capability blocks in the RAHS system. The research and analysis tools allow analysts to closely examine signals collected from environmental scanning. The perspective-sharing tools allow analysts to offer insights on pieces of information they receive, and for these perspectives to be analyzed for convergence and divergence, and for possible outliers or even wild cards. The scenario building tools allow analysts to build system maps and ranking models, and to examine various scenario and strategy options.

The RAHS system empowers analysts with a suite of tools, to help them process large amounts of data, which can either be unstructured text obtained from the Internet, or reports uploaded by the analysts. The system gives the analyst the flexibility to apply these analytical tools in any order, in support of the analytic process. These tools include namely, search, summary, entity and temporal analysis, and clustering. The automated workflow which acts as an orchestration layer enables on-the-fly workflow configuration and execution. The Orchestrator enables service discovery and sharing of data, algorithms and visualization in the RAHS environment.

Data organising capabilities that enable collaborative modelling efforts by the analysts are incorporated into the RAHS system, thus allowing analysts to connect across silos and challenge previous thinking assumptions. Capabilities for model monitoring are also provided, to enable matching of models with incoming data streams, and to allow the human team to explore the interpretations and implications of these data.

4.2 FOI Impactorium

Impactorium is a continuously evolving set of information fusion tools developed at FOI. While the tool-set started as a simple situation monitoring application [1], it has since grown to include more modelling and fusion components and its range of application now includes forensic analysis,

situation monitoring for military and civil infrastructure applications, and information structuring and analysis for intelligence applications.

The initial versions of Impactorium were developed in a FOI project that aimed at delivering tools for situation and impact assessment for Battle Groups, the European Unions rapid deployment forces. Interviews with Swedish Armed Forces personnel who had participated in earlier international missions were performed [15]. The aim of the interviews was to determine what information needs they have and what tools they used to spread information. As expected, the need for having the right information at the right time and place was stressed, as was the need for processed intelligence analyses rather than access to raw information.

The Impactorium tool was developed to be used by staff officers primarily in the J2 and J3 functions and to help ensure that they could process the information quicker and provide field officers with updated analyses of the current situation.

The Impactorium tool as used in the experiment has three components: a modelling tool where the user constructs a Bayesian belief network corresponding to events of interests and their indicators; an indicator activation component, where the user associates reports with indicators; a visualisation component, where probabilities of different events are shown and the user can sort and filter the available reports based on events they are associated with.

The tool is based on information structuring with the help of *indicators* and Bayesian belief networks. Each report that is read by the analyst is tagged with relevant indicators selected from a previously defined set. The indicators can also have values and are used to update the leaf nodes in the Bayesian belief network.

The visual interface of Impactorium consists of a matrix that has four fields ({low probability, low impact}, {low probability, high impact}, {high probability, low impact}, and {high probability, high impact}) denoting the a priori probability and impact of events. An event is described as a combination of a variable (e.g. threat) and its value (e.g. high, medium and low level). For each event in the matrix, a number of indicators are listed. The indicators are examples of observations that can be made which indicate that a particular event is about to or has already occurred. Figure 1 shows the main visual display of Impactorium.

By specifying the dependence relationship between indicators and events in the Bayesian belief networks, the number of parameters to consider when determining the probabilities of specific events occurring can be reduced. In the Impactorium, it is possible to determine whether the events are connected via the noisy-OR-principle or the noisy-AND-principle. An example Bayesian belief network is shown in Figure 2.

Impactorium was integrated with the RAHS system as a web service, enabling RAHS users to view the list of Impactorium indicators by indicator groups, update indicators observed, and update the Impactorium

Display events of interest

(colour indicates risk for event to occur)

List of observation and intelligence reports

Map positions of reports

Details of selected reports

Details of selected events

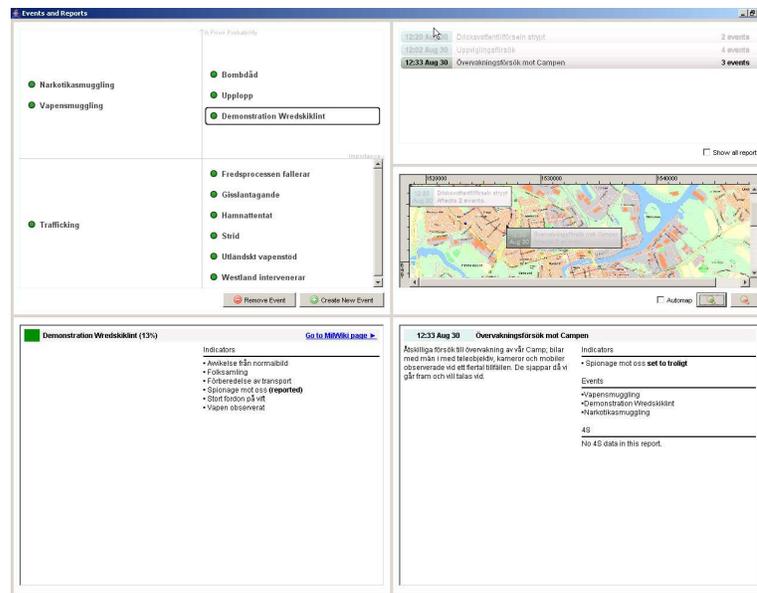


Figure 1. The user interface of version 1 of Impactorium

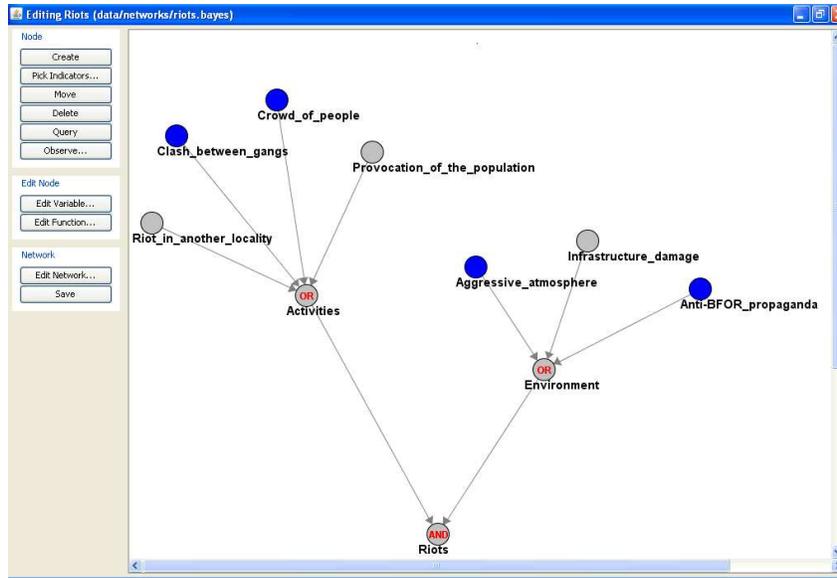


Figure 2. Bayesian network giving a threat model for riots.

situation picture. The Orchestration Layer was used to create monitoring workflows to search for and sort incoming newspaper and intelligence reports via the RAHS system information acquisition capability, into relevant user folders periodically. Users then analyzed the reports in these folders to choose interesting articles to read in detail or to perform research and analysis functions. A tagging interface was added to the RAHS

article viewer to allow the user to specify the indicator(s) referred to by the article he just read, and to send these indicators to the Impactorium web service, which then updated the common Impactorium situation information. Thus, the Impactorium situation picture is constantly updated to reflect the input of all users based on the latest reports. The workflow process used is shown in Figure 3.

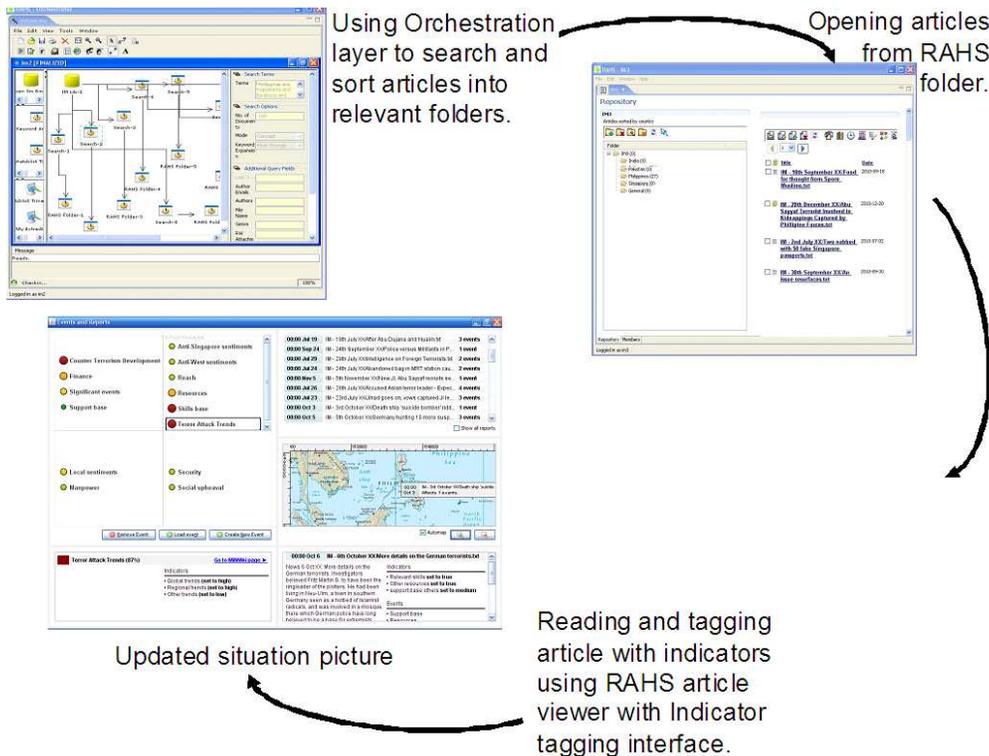


Figure 3. Workflow process combining RAHS System and Impactorium

4.3 Design influences on Impactorium

Initial Swedish studies of Impactorium were conducted with users before and after their deployment to Kosovo, and with students at the Swedish National Defence College. In both cases, a scenario based on a standard Swedish Armed Forces exercise framework scenario was used. The setting was that the users were responsible for keeping order in the city of Norrköping. A two-day scenario was used with a number of reports sent to the users for analysis and processing. Among the reports, clues were planted that indicated several coming events that would disturb the order in Norrköping. Several noise reports were also included, in order to make sure that the task of detecting the events of interest was not too easy.

Participants were instructed to work together in planning for a peace meeting which will take place in the city. After a brief presentation of the purpose of the experiment, the test began with a training session, so that participants could become familiar with Impactorium. Participants were observed during the experiment and interviewed after completing the task. The interview was conducted with a number of questions based on previous research in decision making.

A follow-up study was conducted after the users had returned from international deployment.

The results from the combined interviews indicate that the tool can help give users a better understanding of the on-going situation, which would help them in their decision-making. The users were, understandably, a bit sceptical towards the automatic calculation of probabilities of future events, but embraced the idea of being able to sort and filter reports based on the indicators.

The users also agreed that the tool should be used by intelligence analysts rather than decision-makers themselves.

A complementary study was performed with students at the National Defence College. Here, a class of 12 persons was divided into three groups

- one group which was given access to the complete Impactorium tool
- one group which was given access to a computer tool that listed all available reports and could display more information about them
- a single user was also presented with no computer tool at all, but was instead given all the reports in the scenario on paper, one at a time.

The purpose of the exercise was to demonstrate to the students the possibilities of automatic information fusion tools. After the scenario was completed, the results were discussed with the students, but no structured interviews were made. The exercise came at the end of a 24 class hour introductory course on information fusion and decision support tools.

The results of the exercise indicated the relevance of information fusion tools. The users who were given only

the report viewing tool were able to detect only one of the antagonistic events hidden in the scenario. This demonstrates the bias of humans to stop looking for alternate hypotheses when we have selected a good enough hypothesis about the future. In contrast, the users who had access to the complete Impactorium tool were able to detect all three impending hostile events. This shows the usefulness of the probability calculations, which would indicate likely events in red, thus alerting the users to them. The single user who had no computer support tool at all was able to detect one of the events, but did not have time to read all the reports. The users all commented on the usefulness of Impactorium.

5 Singapore Experiment

5.1 Experiment design

Two groups of analysts were engaged in this research to use either the baseline RAHS system with Microsoft Office (RAHS++ group) or the RAHS system with Impactorium tool (RAHS + IM group). The effectiveness of the Impactorium tool in augmenting the analyst in performing threat assessment was measured by comparing the levels of situation awareness and the quality of decisions made by analysts using the different decision support systems.

Questionnaires were developed to elicit the level of analysts' situation awareness through the three main components of situation awareness: perception, comprehension and projection of future states. This is similar in structure and concept to Situation Awareness Global Awareness Technique (SAGAT) [16], [17]. The implementation of the SAGAT methodology included a modification: the questionnaires were administered at natural breaks in the session (for example: lunch break and end of session) instead of freezing the experiment at random points. But following SAGAT methodology, all information presented just before the point of questionnaire administration was tested. One questionnaire was used during the practice session and two were used in the trial session. Questions in the perception category queried information relevant to the pre-event indicators; questions in the assessment category elicited information with regards to the meaning of the information perceived; and the questions from the projection category obtained information about possible future events. Information apropos to the accuracy and timeliness of decisions made by analysts was elicited through the appropriate and accurate delivery or non-delivery of updates of critical events to higher management.

5.2 Practice session scenario

The scenario used in the practice session consisted of three main themes: intentions of the terrorists, capabilities of the terrorists, and opportunities for strike. In the scenario, the terrorists had no intentions to attack

scenario time. Interviews with individual participants were conducted after the completion of the questionnaires.

6 Results of the Singapore experiment

Scores for the practice session were not analysed.

6.1 Situation awareness

Two independent raters scored the situation awareness questionnaires, and the inter-rater reliability was $r = 0.852$. Figure 5 shows that there was no difference between the scores in both phases of the research suggesting that learning effect was minimal in performing the tasks and answering the questionnaires. Thus, scores were collapsed across the phase of research for further analysis.

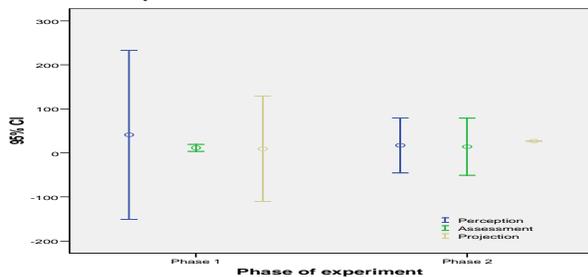


Figure 5. Confidence intervals for situation awareness scores of both phases

Figure 6 shows the comparison of scores across the type of software used. A breakdown of the scores shows that there is a trend towards better perception and projection for users using the RAHS + IM software.

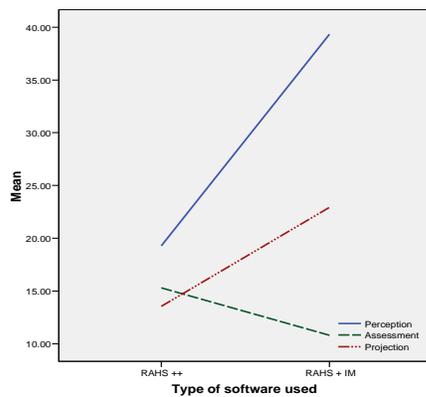


Figure 6. Mean situation awareness scores for both phases of the trial session.

The use of RAHS and Impactorium has mixed effects on analysts' quality of situation awareness when performing threat assessment. Perception and projection scores were consistently better for users of RAHS + IM, which suggest that the use of Impactorium enabled users to see more of the elements in their situation as well as better anticipate future occurrences than users of the RAHS++

system. However, assessment of the environment was not facilitated with the use of Impactorium.

6.2 Decision making

Quality of decision making was operationalised as the timely delivery of an accurate report to emphasize the occurrence of an incident of critical importance. Timely delivery is defined as in time to prevent the incident from occurring, specifically two weeks to a month before the critical incident, while accuracy of the report was based on the number of simulated facts reported and how the story is weaved together.

In phase 1 of the trial, the RAHS + IM users were correct in not issuing a report however the RAHS++ user incorrectly stated in their report that there will an incident occurring in July. In phase 2 of the trial, the RAHS++ user correctly stated in their report that there will be an incident occurring around Christmas, along with the supporting evidence of infiltration while the RAHS + IM users did not submit a complete report.

The use of RAHS and Impactorium has mixed effects on analysts' quality of decision making when performing threat assessment. The quality of decisions made did not differ when using either RAHS ++ or RAHS + IM. This could be because decision making measures were not sensitive enough to measure a decision point, and it might have sufficed to provide an avenue for verbal updates to a superior. On the other hand, this finding is aligned with research that suggests that experts are more inclined to use schemas and heuristics in making decisions than systematic analysis and that formal analysis might impede judgements [7].

6.3 Reflections

During the after-session interview, RAHS + IM users acknowledged the usefulness of the Bayesian technique software, which enabled them to test the hypothesis of the situation, and to verify their mental models of the situation. However, they stated that they would like to see some improvements made to the software. Specifically, the inability to differentiate between the activation of the indicators by different users, track who made the changes and whether the group agreed to the assessment meant that the users were unable to fully understand changes in the situation picture.

The interviews highlighted that Impactorium is a functional tool for testing one's understanding of the situation. Although collaboration facilities need to be in place to draw a balance between full access rights to modify all models and view its outputs and access rights to view outputs only. This finding is congruent with suggestions that decision support systems should enhance team identity via allowing tracking of team mates task and progress, and enhance team cognition via providing a common situation picture while making allowances for different perspectives [7].

An important constraint with this research is that the project was conducted in a simulated environment with hypothetical elements thus it might have contained an artificiality that violated the flow of terrorist strikes and reporting procedures. To minimise the effects of a contrived environment, the scenario was based closely on an actual incident while the task was based on standard operating procedures.

7 Conclusion

Bayesian technique modelling is relatively effective in improving situation awareness, however, its effectiveness in enabling timely and accurate decision making is ambiguous. Further research needs to be done on the conditions for its effectiveness. Findings from the project suggest that the concept of use is an important issue for consideration when implementing the software. Bayesian technique software could be used in facilitating alternative hypothesis generation and evaluation, as opposed to providing answers for questions in general. This is in line with research, which has identified conditions for the enhanced use of Bayesian technique in the intelligence domain [9]. Firstly, binary questions, for example, "Is an attack more likely or unlikely?", are more effectively answered than ambiguous questions. Secondly, periodic review of the a priori standardised weights would ensure that the relationships between factors are still relevant with respects to the current situation. In addition, revisiting the relationships to verify the accuracy of the initial state probability (which assumes that given all things equal, relationships between factors will have particular effects) would be optimal because in actual situations, the probability of effects between relationships of factors might be more complex and thus reflect convoluted relationships. Thirdly, the acknowledgment that a true zero is inapplicable in a real intelligence situation and thus the use of continuous scales in the specification of probability is artificial. Finally, the appreciation of the challenge in representing non-events as interesting patterns in Bayesian technique.

In future work, we plan to study how the use of Bayesian belief networks could be integrated into the strategic intelligence analysis process in more detail.

Acknowledgements: This work was supported by Singapore's National Security and Coordination Centre; and by the FOI research projects "Situation and impact assessment for battle groups" and "Tools for

information management and analysis", which are funded by the R&D programme of the Swedish Armed Forces.

References

- [1] Svenson, P., Berg, T., Hörling, P., Malm, M., Mårtensson, C., Using the impact matrix for predictive situational awareness, In Proceedings of the 10th International Conference on Information Fusion (FUSION 2007), 2007
- [2] Forsgren, R., Kaati, L., Mårtensson, C., Svenson, P., Tjörnhammar, E., An Overview of the Impactarium Tools 2008, In Proceedings of the Second Skövde Workshop on Information Fusion Topics (SWIFT 2008) (2008).
- [3] Endsley, M. R. (1997). The role of situation awareness in naturalistic decision making. In Zsombok, C. E. & G. Klein (Eds.), *Naturalistic decision making* (pp. 269–283). Mahwah, NJ: LEA.
- [4] Liggett, M.E., Hall, D.L., Llinas, J. (eds), *Handbook of Multisensor Data Fusion: Theory and Practice*, Second Edition, CRC Press 2008
- [5] Elliott, T. (2005). Expert decision making in naturalistic environments: A summary of research. Australia: Defence Science and Technology Organisation. Pp. 1-59.
- [6] Klein, G. & Klinger, D. (1991). Naturalistic decision making. *Human Systems IAC*, 11, 16-19.
- [7] Klein, G. (1997). Implications of the naturalistic decision making framework for information dominance. United States: U.S. Air Force Armstrong Laboratories. Pp. 1-52.
- [8] Klein, G. (1997). Making decisions in natural environments. United States: U.S. Army Research Institute for the Behavioural and Social Sciences. Pp. 1-19.
- [9] Schweitzer, N. (1994). Bayesian Analysis for Intelligence: Some Focus on the Middle East. *Studies in Intelligence*, 20, 31-44.
- [10] Das, B. (1999). Representing uncertainties using Bayesian networks. Australia: Defence Science and Technology Organisation. Pp. 1-57
- [11] Rennie, S. & Brandt, A. (2002). An expert system approach for predicting mine burial. Proceeding of the 5th International Symposium on Technology and the Mine Problem. Naval Postgraduate School, Monterey, CA, 21-21.
- [12] Falzon, L. & Priest, J. (2004). The Center of Gravity Network Effects Tool: Probabilistic modeling for operational planning. Australia: Defence Science and Technology Organisation. Pp. 1-45.
- [13] Fisk, Charles E. (1972). The Sino-Soviet Border Dispute: A Comparison of the Conventional and Bayesian Methods for Intelligence Warning. *Studies in Intelligence*, 16(2), 53-62.
- [14] Lanir, Z. & Kahneman, D. (2006). Speaking to Policymakers: An Experiment in Decision Analysis in Israel in 1975. *Studies in Intelligence*, 50(4), 11-19.
- [15] Johansson, K. & Kylesten, B. (2006). "Hur ser verkligheten ut under internationella missioner? – en explorativ beskrivning för ökad domänförståelse" (in Swedish, English title: "Reality during international missions – an explorative description for increased domain understanding"), FOI Report FOI-R—1953—SE.
- [16] Endsley, M.R. (1995). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*, 37, 65-84.
- [17] Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). Proceedings of the National Aerospace and Electronics Conference (NAECON), 789-795. New York: IEEE