2012

Representing Variable Source Credibility in Intelligence Analysis with Bayesian Networks

Ken McNaught
Cranfield University

Peter Sutovsky
Cranfield University

DOI: 10.4225/75/57a03050ac5cb
Originally published in the Proceedings of the 5th Australian Security and Intelligence Conference, Novotel Langley Hotel, Perth, Western Australia, 3rd-5th December, 2012

REPRESENTING VARIABLE SOURCE CREDIBILITY IN INTELLIGENCE ANALYSIS WITH BAYESIAN NETWORKS

Ken McNaught and Peter Sutovsky
Operational and Decision Analysis Group
Dept of Informatics and Systems Engineering
Cranfield University
Defence Academy of the UK, Shrivenham, Swindon, UK
k.r.mcnaught@cranfield.ac.uk, p.sutovsky@cranfield.ac.uk

Abstract

Assessing the credibility of an evidential source is an important part of intelligence analysis, particularly where human intelligence is concerned. Furthermore, it is frequently necessary to combine multiple items of evidence with varying source credibilities. Bayesian networks provide a powerful probabilistic approach to the fusion of information and are increasingly being applied in a wide variety of settings. In this paper we explore their application to intelligence analysis and provide a simple example concerning a potential attack on an infrastructure target. Our main focus is on the representation of source credibility. While we do not advocate the routine use of quantitative Bayesian networks for intelligence analysis, we do believe that their qualitative structure offers a useful framework for evidence marshalling. Furthermore, we believe that quantified Bayesian networks can also play a part in providing auxiliary models to explore particular situations within a coherent probabilistic framework. This process can generate fresh insights and help to stimulate new hypotheses and avenues of enquiry.

Keywords
Intelligence Analysis, Bayesian Network, Evidence Marshalling, Credibility Assessment

INTRODUCTION

The assessment of the credibility of a source of evidence is an important part of intelligence analysis (Joseph and Corkill 2011). This is perhaps most obvious in the case of human intelligence (HUMINT) although it is also true for other forms – for example, a clear image has more credibility than a fuzzy one and a crisp sound recording has more credibility than a muffled one. In many fields, including medicine, law and, of course, intelligence analysis, it is very often necessary to combine several pieces of evidence of different types to try and make sense of an uncertain and complex situation. Making that process more difficult is the variable credibility associated with the different evidential sources. Here, we use the term 'credibility' as a characteristic of the evidential source in a similar way to Schum et al. (2009), although the alternative term 'reliability' is often used in two-part classifications, such as that used by NATO.

In this paper, we show how a particular type of probabilistic graphical model, a Bayesian network (BN), provides a computational platform to support such a process of information fusion. However, we are not advocating the routine use of this approach in intelligence analysis, partly due to the difficulty of quantifying such models. Nonetheless, we believe that such models can be instructive and can help to shed light on various aspects of reasoning under uncertainty. Here we are particularly interested in asking if they can provide any insights regarding the combination of evidence from sources having variable credibility. An additional benefit is that they can help to address some of the severe problems and limitations encountered in the communication of uncertain information in the intelligence field as described, for example, by Weiss (2008).

BNs are being applied increasingly to a greater diversity of areas where uncertainty is prevalent and some kind of information fusion is required to make better sense of that uncertain environment. While the notion of applying probability and Bayes' theorem to intelligence analysis, at least in theory, has a long history, it is only since the advent of BNs that sufficient computational efficiency has been available to actually make it feasible for situations involving more than a handful of variables. However, the bottleneck typically then shifts to the problem of estimating or eliciting the necessary probability distributions to populate the model.

One of the main strengths of such formal models is that they help overcome the problem of cognitive biases which afflict human reasoning. Many such cognitive biases are known which negatively affect the mental processing of information, e.g. (Miller 1956). There is a rich psychological literature concerning them and their potential adverse effects on intelligence analysis was most famously expounded by Heuer (1999). Wastell

(2010) points out the need for formal models to complement an intelligence analyst's natural reasoning capabilities.

OVERVIEW OF BAYESIAN NETWORKS

Bayesian networks (Pearl 1988) belong to the family of probabilistic graphical models which exploit graphs to express dependencies between variables in a modelled domain. The graphs are easily readable and intuitive for humans and at the same time are convenient and efficient tools from an algorithmic perspective. In graphical models, variables are represented by nodes in the graph and dependencies between variables are represented by arcs. In a Bayesian network, the arcs are directed and the graph is both directed and acyclic (contains no directed loops). Hence, the qualitative structure of a BN is represented by a directed acyclic graph (DAG), portraying probabilistic dependencies and independencies within the domain. This contains a great deal of information, even before we consider any probability distributions. For example, consider the BN in Fig. 1. From the DAG, we can infer that variables A, C and F are all dependent, meaning that if we are uncertain about all of them but receive fresh evidence on one of them, our beliefs in the others will also be revised or updated. However, since there is no direct dependence between A and F, once we establish the true state of C, A and F then become independent since fresh evidence on either of those will not lead to a revision of our belief in the other. A and F are said to be conditionally independent given C. The chain ACF is said to form a series connection within the BN. Nodes D, F and G together form a so-called diverging connection. Again, these three nodes are dependent when they are all unknown but if we establish the true state of D, F and G then become independent. F and G are only dependent because they share a common parent, D, so are said to be conditionally independent given D. The final type of connection within a BN is a converging connection, e.g. C, D and F. This time C and D are independent of each other as long as we have no evidence about their common descendant F. Receiving evidence about F, however, induces probabilistic dependence between C and D. So now if additional evidence is received about C, belief in D will be updated and vice versa.

A fully specified BN, however, requires the construction of conditional probability tables (CPTs) for each node. For nodes with no arcs entering them, i.e. with no parent nodes, only a single prior distribution has to be specified. For nodes with a single parent, a conditional probability distribution will have to be specified for each possible state of the parent variable. Finally, for chance nodes with more than one parent, a conditional probability distribution is usually required for every possible combination of parent states. While at first sight, this may appear rather burdensome, there are often special cases where this requirement can be relaxed.

The ability of BNs to provide a flexible and powerful probabilistic modelling framework makes them suitable for applications in many fields. For example, risk modelling (Fenton and Neil 2012) and forensic analysis (Taroni et al. 2006) are two fields which share some commonalities with intelligence analysis and in which applications of BNs are increasing.

PROPOSED FRAMEWORK FOR EVIDENCE MARSHALLING

The importance of evidence marshalling for intelligence analysis and law enforcement applications has been highlighted by Schum (2001). Inferential reasoning networks are among various approaches considered by Schum for this task. How the qualitative structure of a BN provides an inferential reasoning network which can be used to provide a framework for evidence marshalling has been discussed by McNaught and Sutovsky (2012).

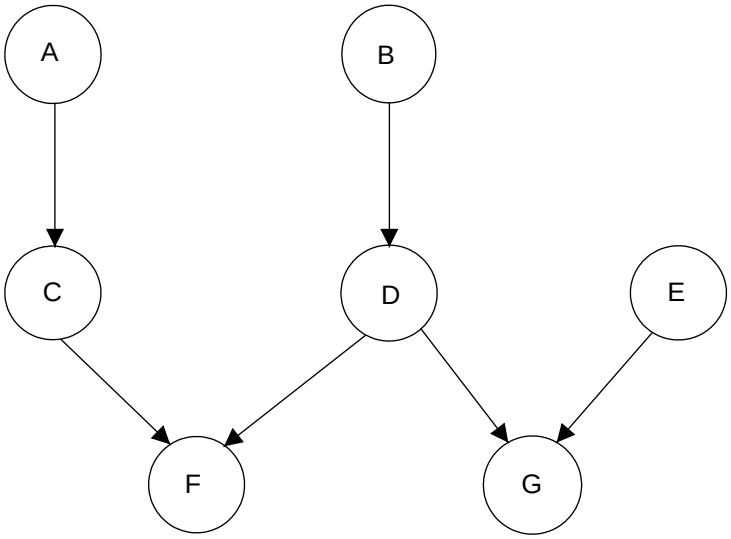

Fig.1. A DAG displaying the qualitative structure of an example BN.

Fig. 2 displays a generic network made up of four layers of nodes. Nodes correspond to propositions and the four different layers correspond to high-level hypotheses, ground truth variables and related activities implied by the parent hypotheses, observable indicators or potential items of evidence and, finally, nodes relating to the credibility assessments of the evidential sources.

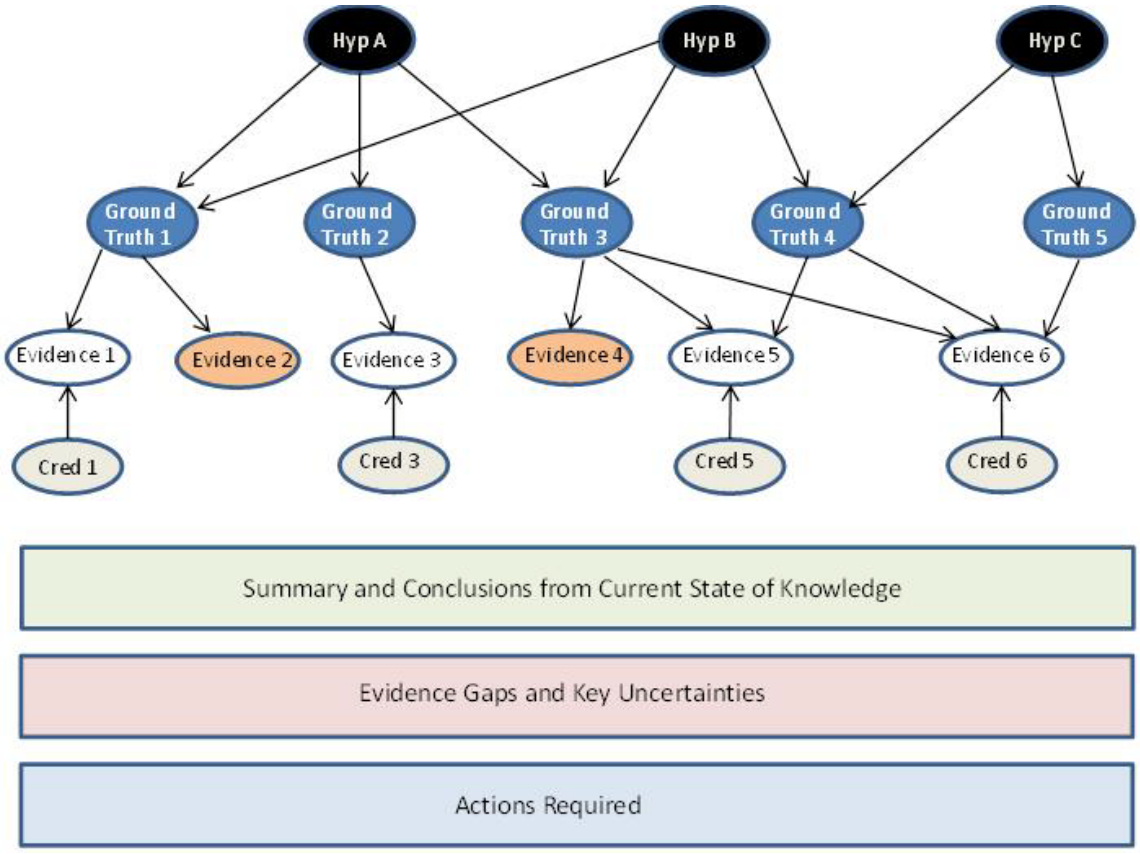

Fig. 2. Generic inferential reasoning network and associated workspace.

The workspace shown below the network represents the type of information that an analyst is prompted for by a prototype tool which we are developing. This is a free text area which allows beliefs to be recorded as time progresses and evidence unfolds. There are three main categories – a summary of the analyst's current understanding of the situation, an analysis of evidence gaps and key uncertainties, and finally a list of actions required. This is intended to encourage thoughtful reflection as the evidential picture unfolds, as well as the explicit recognition and analysis of evidence gaps, contradictions and uncertainties, including possible deception activities by the adversary in question. Finally, the investigator is invited to make a list of required actions such as requests for additional information, requests for resources, suggested new leads to investigate, current leads to drop, etc, based on the foregoing analysis. This helps to create an audit trail, which can be time-stamped, of what was done, when and why during the course of an investigation, clearly linking these decisions with the beliefs and possibilities being considered at the time and providing a logical justification for them. Such an approach also supports collaborative working, making it easier for co-workers and shift workers to understand each other and pick up where the other one has left off.

While we believe that a framework such as the above, based on BNs, is intuitive and encourages logical, coherent reasoning, we do not advocate routine quantification of such networks in order to calculate posterior probabilities of various hypotheses given what evidence has been observed. To facilitate this kind of probabilistic inference, it would be necessary to quantify each node with one or more probability distributions. This would require the elicitation of many probabilities and in the type of situation typically of interest to intelligence analysts, a good number of these would be very uncertain. While 'best estimates' might be used, the dangers of such an approach are clear. An alternative and less controversial use of a fully quantified BN in this context would be to employ it as an auxiliary model of the situation of interest. This would permit exploration of the problem, essentially a kind of 'what if' analysis, which might help to stimulate new hypotheses, questions and avenues of enquiry.

EXAMPLE BN FOR INVESTIGATION OF A POTENTIAL ATTACK

Fig 3 displays a fragment of a BN for a scenario concerning a potential attack on a critical infrastructure target. The full scenario and corresponding BN is discussed in McNaught and Sutovsky (2012). Here we consider just a fragment of that BN. This has been generated within the 'GeNIe' software which facilitates the construction of BNs and performs the subsequent probabilistic inference as evidence unfolds.

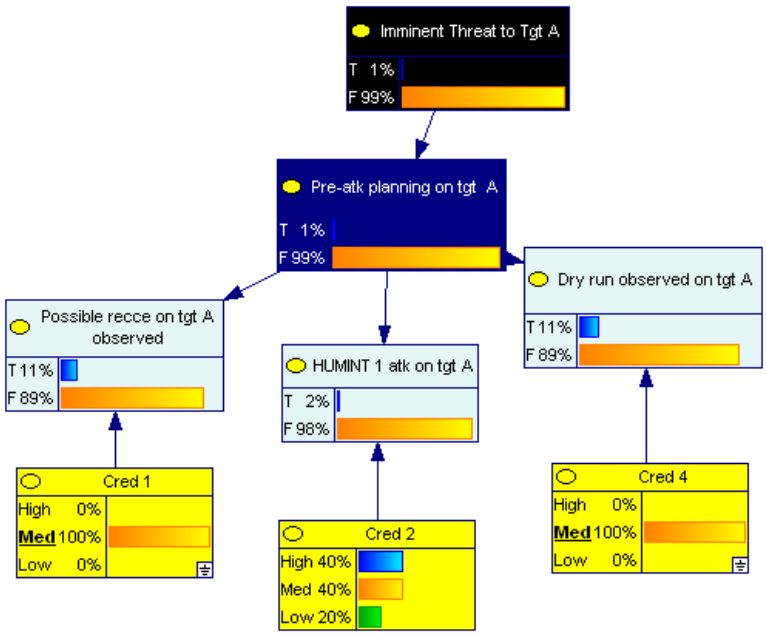

Fig.3. BN fragment relating to potential attack on an infrastructure target.

The high-level hypothesis at the top of the BN is the proposition that there is an imminent threat to a particular target, A. This then causally influences the proposition that pre-attack planning on Target A has taken place. The truth or falsity of this proposition then influences three observable indicators or potential items of evidence – human intelligence of such an attack, observation of a potential recce of the target, and observation of a potential 'dry run' or practise attack on the target. Fig 3 displays the initial probabilities associated with these

various propositions given the assumptions made about the scenario. These are then updated logically and coherently in accordance with probability theory (this process is automated within Genie) as evidence is observed (propositions are instantiated, often as either True or False in this simplified example).

Fig 4 shows how these probabilities change when evidence is entered reflecting the receipt of HUMINT that an attack is likely on Target A and, following that, an analysis of CCTV footage reveals a possible recce on the target. Given the probabilistic parameters assumed at the start, these two items of evidence together lead to our estimated probability of pre-attack planning having taken place rising to 90% and, in turn, our estimated probability of there being an imminent threat to Target A rising from 1% to 60% (it is not 90% because the pre-attack activity does not always lead to an actual attack). However, subsequent search for evidence of a dry run reveals nothing suspicious. This observation is entered in Fig 5 where its effect on the various unobserved propositions can be seen. In particular, the estimated probability of an imminent threat to Target A drops significantly to 34%. Again, this may be because a potential attack on this target has been abandoned or it was only ever a decoy target or the initial evidence pointing to Target A was mistaken. However, 34% is still an appreciable probability and would certainly justify heightened security around Target A and further investigation.

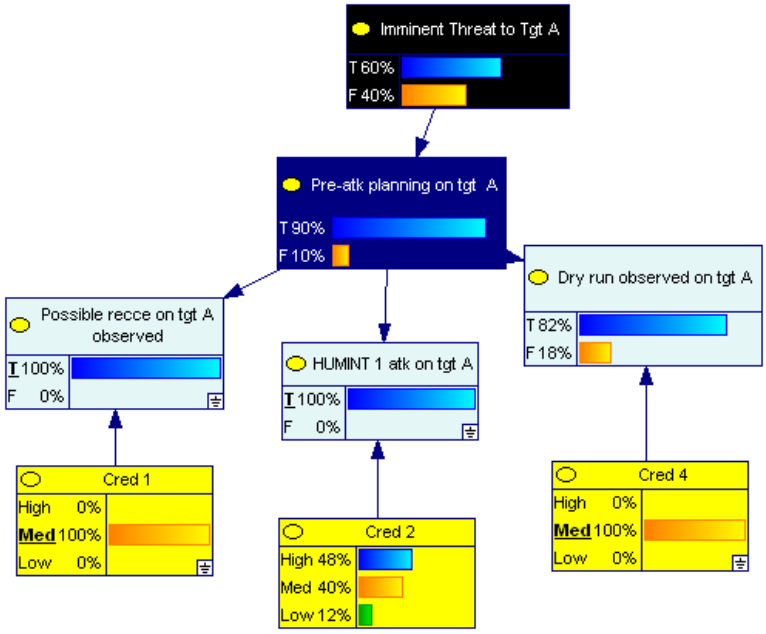

Fig. 4. Positive evidence of possible recce and HUMINT entered.

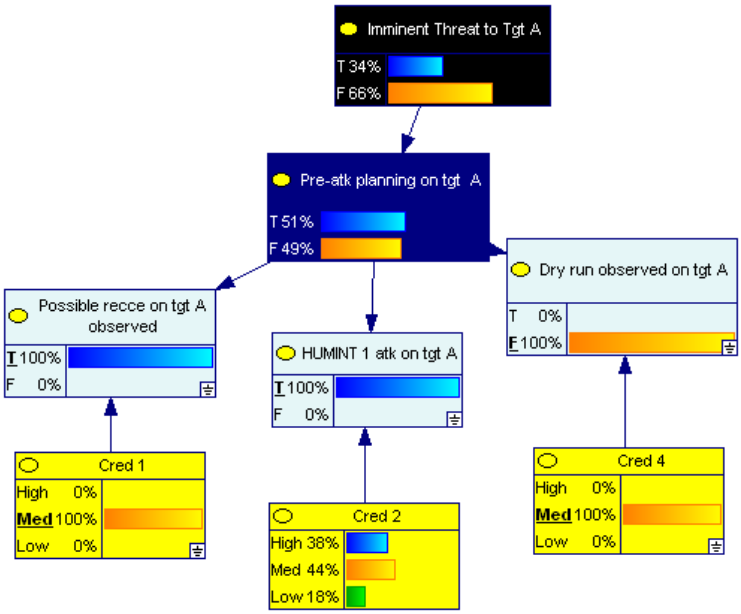

Fig. 5. Additional evidence of no 'dry run' detected is added.

CREDIBILITY ASSESSMENTS

An interesting thing to note in going from Fig 4 to Fig 5 is the update in the probability distribution associated with node 'Cred 2' which corresponds to our belief in the credibility of our HUMINT source. This is explained by the lack of dry run evidence entered in Fig 5 which shifts the credibility distribution of the HUMINT source towards lower levels. At first, an analyst may not expect a credibility assessment of a source to be affected by another item of evidence but this is entirely logical since the lack of 'dry run' evidence makes an attack on Target A less likely and so contradicts their story. Consequently, if we are unsure about it, then we should reduce the HUMINT source's credibility in such a situation. In this scenario, we began by assuming that the credibility of the HUMINT source was unknown but given what we knew about the source, their probabilities of having 'High', 'Medium' or 'Low' credibility was assessed to be 40%, 40% and 20%, respectively. This raises an interesting question of whether we should think of our source as having a probability distribution over their credibility or whether we should assign a fixed value, as we have done, for example, with the credibility of CCTV evidence. Nodes 'Cred 1' and 'Cred 4' relate to the credibility of CCTV evidence and have here been assessed as 'Medium', a fixed state. That reflects how much confidence we place in such a source and the ability to fix a definite value reflects long experience and familiarity with that source's strengths and weaknesses. It is debatable if we can therefore truly assign a fixed value of credibility, at whatever level, to a relatively recent human source who we may still have some doubts about. It may be more realistic to assign a distribution of credibility states, reflecting our lack of knowledge about the source, thus permitting this distribution to shift upwards to higher credibility states when evidence confirming the source's assertions is received or to shift downwards to lower credibility states when evidence contradicting the source is received.

In fact, the distribution of 'Cred 2' also changed in going from Fig 3 to Fig 4. On this occasion the distribution shifted to higher credibility states for our HUMINT source when evidence of a possible recce on Target A was received, providing support for their story.

Another way that belief in a source's credibility might change is when their previous statements can be checked. For example, a HUMINT source might previously have claimed that a certain individual, X, belongs to terrorist organization Z. Providing that this is not already common knowledge, this provides a potential check on the source's credibility. Fig 6 shows such a fragment before the additional information has been checked. Fig. 7 displays the situation after the claim has been verified, i.e. it becomes known that X does belong to Z. Fig. 8 on the other hand shows the effect on the source's credibility when the claim is disproved, i.e. it becomes known that X does not belong to Z. As can be seen in the figures, the credibility of the source rises and falls in line with the confirming and disconfirming evidence, respectively. This will have a knock-on effect on the extent to which subsequent evidence from that source influences beliefs in the remaining propositions.

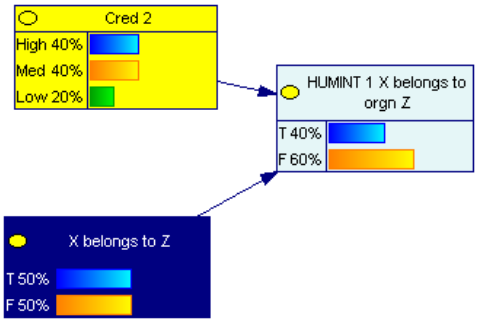

Fig. 6. BN fragment concerning the credibility of a HUMINT source.

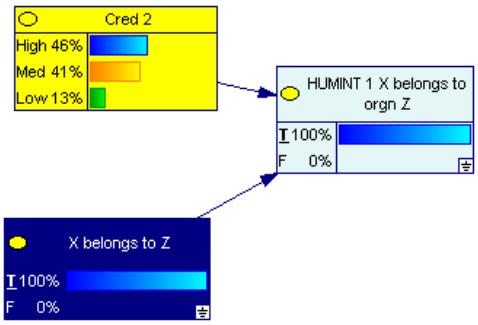

Fig. 7. Source credibility rises when their previous intel is confirmed.

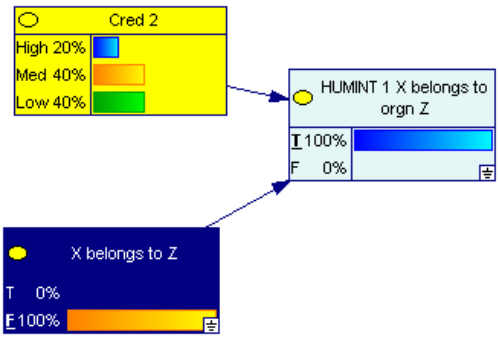

Fig. 8. Source credibility drops when their previous intel is found to be false.

CONCLUSION

In this paper, we have advocated the use of Bayesian networks as auxiliary models to explore complex, uncertain problem situations and to learn about the sensitivity of probabilistic inferences to assumptions made. However, we do not advocate their use to routinely calculate posterior probabilities of interest which would inevitably be misleading and inaccurate given the difficulties of estimating input parameters. Such use is only defensible when the input parameters are known with a reasonable degree of accuracy.

Nonetheless, in our view, familiarity with formal, coherent computational models of uncertain reasoning, particularly BNs which have probability theory as their axiomatic basis, could help analysts to overcome some of their innate cognitive biases and provide important insights regarding the combination of evidence and the sensitivity of inferences to source credibilities.

ACKNOWLEDGMENTS

The authors are grateful for funding provided by the UK's EPSRC under Grant Number EP/H023135/1. The models described in this paper were created using the GeNIe modeling environment developed by the Decision Systems Laboratory of the University of Pittsburgh and available at http://genie.sis.pitt.edu/.

REFERENCES

Fenton, N. and Neil, M.D., 2012. *Risk Assessment and Decision Analysis with Bayesian Networks*. Boca Raton, Florida: CRC Press.

Heuer, R., 1999. *The Psychology of Intelligence Analysis*. Washington DC: Center for the Study of Intelligence, CIA.

Joseph, J. and Corkill, J., 2011. Information evaluation: how one group of intelligence analysts go about the task. *Proceedings of the Australian Security and Intelligence Conference 2011*, pp 97-103.

McNaught, K.R. and Sutovsky, P., 2012. Evidence marshalling with inference networks: an application to homeland security. *Proceedings of the Defence and Homeland Security Simulation Workshop* 2012, pp 79-84.

Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.

Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Schum, D.A., 2001. Evidence marshaling for imaginative fact investigation. *Artificial Intelligence and Law*, 9, 165-188.

Schum, D.A., Tecuci, G. and Boicu, M., 2009. Analyzing evidence and its chain of custody: a mixed-initiative computational approach. *International Journal of Intelligence and Counter-Intelligence*, 22, 298-319.

Taroni, F., Aitken, C., Garbolino, P. and Biedermann, A., 2006. *Bayesian Networks and Probabilistic Inference in Forensic Science*. Chichester: Wiley.

Wastell, C.A., 2010. Cognitive predispositions and intelligence analyst reasoning. *International Journal of Intelligence and Counter-Intelligence*, 23, 449-460.

Weiss, C., 2008. Communicating uncertainty in intelligence and other professions. *International Journal of Intelligence and Counter-Intelligence*, 21, 57-85.