

THE INTERNATIONAL
C2 JOURNAL

VOLUME 5, NUMBER 1, 2011

SPECIAL ISSUE

*Beyond Command and Control:
Sense Making under Large World Uncertainty*

GUEST EDITOR

Jason K. Levy
Virginia Commonwealth University

A Simple Heuristics-Based Model for
Threat Prediction to Support Decision-Making

Kellyn Rein

THE INTERNATIONAL C2 JOURNAL

David S. Alberts, Chairman of the Editorial Board, *OASD-NII, CCRP*

The Editorial Board

Éloi Bossé (CAN), *Defence Research and Development Canada*

Berndt Brehmer (SWE), *Swedish National Defence College*

Lorraine Dodd (GBR), *Cranfield University*

Reiner Huber (DEU), *Universitaet der Bundeswehr Muenchen*

William Mitchell (DNK), *Royal Danish Defence College*

Sandeep Mulgund (USA), *The MITRE Corporation*

Mark Nissen (USA), *Naval Postgraduate School*

Mink Spaans (NLD), *TNO Defence, Security and Safety*

Andreas Tolk (USA), *Old Dominion University*

About the Journal

The International C2 Journal was created in 2006 at the urging of an international group of command and control professionals including individuals from academia, industry, government, and the military. The Command and Control Research Program (CCRP, of the U.S. Office of the Assistant Secretary of Defense for Networks and Information Integration, or OASD-NII) responded to this need by bringing together interested professionals to shape the purpose and guide the execution of such a journal. Today, the Journal is overseen by an Editorial Board comprising representatives from many nations.

Opinions, conclusions, and recommendations expressed or implied within are solely those of the authors. They do not necessarily represent the views of the Department of Defense, or any other U.S. Government agency.

Rights and Permissions: All articles published in the International C2 Journal remain the intellectual property of the authors and may not be distributed or sold without the express written consent of the authors.

For more information

Visit us online at: www.dodccrp.org

Contact our staff at: publications@dodccrp.org



A Simple Heuristics-Based Model for Threat Prediction to Support Decision-Making

Kellyn Rein (Fraunhofer FKIE, DE)

Abstract

During military endeavors a large amount of information floods in from a variety of sources, both human sources and non-human sources such as autonomous (robotic) vehicles, sensors, etc. Timely evaluation of intelligence with background information is necessary for effective operations, but the sheer volume of incoming data poses a tremendous challenge. Automatically fusing data derived from multiple diverse sources into recognizable patterns of potentially threatening behavior can provide a distinct operational advantage, allowing commanders to react with agile, appropriate responses to an ever-changing situation. This advantage would be even more marked if the model is flexible enough to respond to changing patterns of behavior, and if the time needed for processing was close to real-time.

In this article we present an easily modifiable model for threats which use Battle Management Language (BML) as the underlying standardized language for representation and fusion. The model is simple and based upon heuristics, its strength lies in the analysis and quantification of uncertainties in the fusion process. And most importantly, it provides a near real-time solution for first-pass processing of inflowing information to provide early warning of developing threats.

Introduction

Military operations today, be they warfare, anti-terrorism, peacekeeping or disaster relief, are increasingly complex. As opposed to the past, in which during combat operations the enemy was recognizable, today's threats are often hidden and only indirectly identifiable. Furthermore, in non-combat operations there is a need for coordination and cooperation between coalition nations on the military level, as well as with non-military organizations as various as international aid agencies, non-government agencies, local government, and tribal leaders (Hayes 2007). The number of players complicates the reporting and communications structures. At the same time there is an increasingly overwhelming amount of data available about the area of operations: environmental information gathered by a variety of sources, both human and non-human (e.g., devices such as sensors), gleaned from media sources such as web pages, newspaper and television, and available in databases and ontologies. Local information may come from refugees or prisoners of war, in addition to normal HUMINT sources. The capacity to identify situations and patterns of action out of this sea of data, to evaluate these based upon knowledge at hand, and to identify and successfully exploit transient opportunities can support swift, appropriate responses to developing situations (Alberts and Hayes 2003).

In this ever more complex world of military operations knowledge is without question power (Alberts and Hayes 2003). Achieving and maintaining an advantage over the enemy requires the rapid acquisition, synthesis and analysis of information in the theater of operations.

The ability to recognize developing situations and patterns of action make it possible to successfully exploit transient opportunities to deliver a swift, appropriate response or to avert potential attacks—in other words, to be agile. *Agility* as defined in (Alberts 2007) is “the critical capability that organizations need to meet the challenges of complexity and uncertainty.” In particular, the ability to forecast

potentially dangerous developing situations would provide a significant advantage over the enemy. In the case of disaster relief or peacekeeping, timely reaction to the situation is even clearer and unambiguous: lives may be depending upon it.

Each source of data or information provides another piece of a complex mosaic. Sifting and sorting through this wealth of data, to turn it into useful information determines the advantage in a war situation and saves time, energy, and resources in non-war activities; as noted in the Allied Joint Intelligence Counter Intelligence and Security Doctrine: “Information is of great value when a deduction of some sort can be drawn from it. This may occur as a result of its association with some other information already received.” (AJP 2.0 2003)

A human analyst will sort and combine information to build a coherent picture from varied pieces, intuitively analyzing discrepancies and evaluating uncertainties in the process. However, the volume of information streaming in quickly overwhelms even the best analyst. Evaluating and correlating this huge amount of data today also requires a different approach than in the past; no longer can a handful of analysts deal with the massive volume of incoming information. It is increasingly clear that the only way to cope is to automatically pre-process data received from various sources in order to detect developing threats and situations. Fusing information derived from multiple sources into recognizable models can build a far more accurate picture than information provided by a single source or single source type (e.g., sensors), alone.

Any information fusion system, in order to be an effective support tool, must contain mechanisms for a respective structuring of the available information, for correlating potentially related information, and for providing a reasonable assessment of the uncertainties connected with the fusion results. In an ideal world, the information which has been gathered would be complete, unambiguous, and

true. We also would have reliable, well-tested models to support the fusion of information. This fusion then would result in consistent and trustworthy predictions.

There has been much work in developing Bayesian models for associating ideal information on various actors and events in the area of interest and delivering results concerning the uncertainties in the connections evaluated to build reliable models. This, however, can be complex and suffers from problems such as lack of recursion, etc. DaCosta and Laskey (2005) therefore offered an alternative strategy which solves a number of these problems by using multi-entity Bayesian networks.

Unfortunately, however, in the real world the information we gather is not ideal but incomplete and ambiguous, and thus not totally reliable. The enemy tries to conceal his activities from our view. We will see or hear of some but not all enemy activities, much remains hidden. As a result, the observable patterns of behavior upon which our models are based represent just the tip of the iceberg, resulting in “connect-the-dots” models. Further, our underlying data is generally neither complete nor unimpeachable—sensors return results within a range of reliability or fail, humans misinterpret, distort or deliberately lie.

Thus, uncertainty creeps in at all levels within the information gathering, analysis, and fusion process. Understanding and evaluating this uncertainty is necessary for the appropriate quantification of that uncertainty, which is vital to support decision-making.

In this article we will look at information fusion; present Battle Management Language and discuss its use as a basis for information fusion; discuss a heuristics-based method for modeling threats based upon BML; examine issues surrounding the different types of uncertainty in the fusion process, and show how these are represented within the model for fusion which integrates calculation of and demonstrate how these uncertainties may be quantified.

The Information Fusion Process

While there are the numerous definitions for data and/or information fusion, the one which best encapsulates the focus of our work comes from the University of Skövde's Information Fusion Website: “[the] combining, or fusing, of information from different sources in order to facilitate understanding or provide knowledge that is not evident from individual sources.” There are several steps in the fusion process (Kruger et al. 2008), all of which, with the exception of the first bullet, are further examined in this article:

- collection of data and information from various diverse sources,
- where necessary, the conversion of data from its original form to a standardized format in preparation for pre-processing,
- selection and correlation of potentially related individual pieces of information,
- mapping of individual pieces of data to existing threat models,
- evaluation of the credibility of the results of the correlation and mapping process, and
- assessment of the accuracy of models used as the basis for fusion.

Here one must also note that the choice of wording (“information fusion”) is also with intention: Alberts et al. (2001), define *data* as “a representation of individual facts, concepts or instructions”, and *information* as “various points along the information spectrum between data and knowledge.” Since, as will be discussed in the following section, the representation of the results of intelligence collection using BML statements preserves semantic information and context, we feel BML statements represent more than “individual facts” and hence are more than “data.”

Battle Management Language As A Basis For Fusion

Information which has been obtained from a variety of sources is generally in a variety of different formats. This can pose a significant hurdle for automatic fusion of the different pieces of information. Converting available information into a standardized format would greatly support fusion.

Originally designed for commanding simulated units, BML is a standardized language for military communication which has been developed under the aegis of SISO and the NATO MSG-048 "Coalition BML." It has been expanded to communicate not only orders (and requests) but also reports (Pullen et al. 2009). There are also expansions to BML being developed such as AVCL (Autonomous Vehicle Command Language) (Davis et al. 2006), which will facilitate communications (tasking and reporting) with autonomous vehicles. Ordering of and reporting from individual and groups of robots is currently being developed (Remmersmann et al. 2010).

BML is based upon the Joint Consultation, Command and Control Information Exchange Data Model (JC3IEDM), which is used by all participating NATO partners. As NATO standard, JC3IEDM defines terms for elements necessary for military operations, whether wartime or non-war, and is sufficiently expressive to formulate both military and non-military communications for a variety of different deployment types (Figure 1). It also provides a basis for standardized reporting among NATO coalition partners.

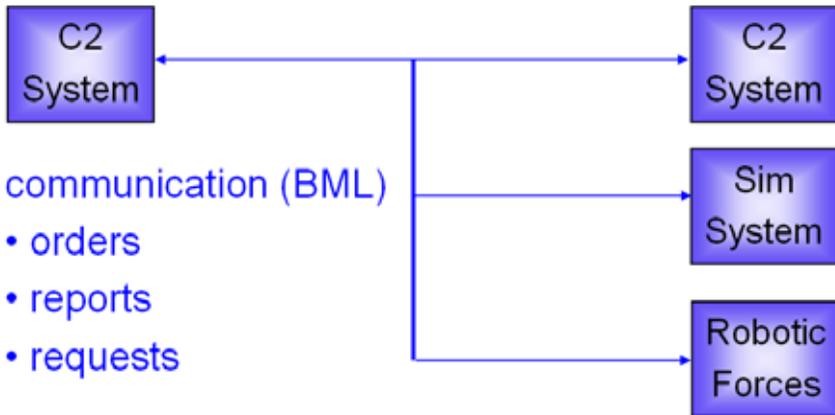


Figure 1. BML is a formal language for military communications such as orders, reports, and requests, which provides a common format for information.

BML has been designed as a controlled language (Schade and Hieb 2006) based on a formal grammar. This grammar is modeled after one of the most prominent grammars from the field of computational linguistics, Lexical Functional Grammar (LFG) (Bresnan 2001). This renders BML an unambiguous language which can easily be processed automatically.

As described by Schade and Hieb (2007), a basic report in BML is a single (atomic) statement which delivers a “fact” about an individual task, event, or status. A task report is about a military action either observed or undertaken. An event report contains information on non-military, “non-perpetrator” events such as flooding, earthquake, political demonstrations, or traffic accidents which may be important background information for a particular threat; for example, a traffic accident may be the precursor of a pending IED detonation. Status reports provide information on personnel, materiel, facilities, etc., whether own, enemy, or civilian—such as number of injured, amount of ammunition available, condition of an airfield or bridge.

There are several important ways in which BML basic reports support the fusion process. First is the fact that each BML “report” is a statement representing a single (atomic) statement. This atomicity is essential for the fusion process; each statement of a more complex report may be processed individually.

Secondly, each basic report has its own values representing source and content reliability. Third is that each report also has a reference label to its origination so that the context is maintained for later use by an analyst. Natural language text sources such as HUMINT reports usually contain multiple statements. Some of these statements may be declarative (“three men on foot heading toward the village”), other statements may be speculative (“possibly armed”). While an analyst may assign a complex HUMINT communication an overall rating (e.g., using the familiar “A1”-“F6” system from the JC3IEDM), individual statements contained therein have varying credibility. Therefore the conversion to BML assigns first the global rating, but adjusts each individual statement according to the uncertainty in its formulation, e.g., on the basis of modality term analysis.

$R_{eventreport-1234773102096}$	=	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 2px 10px;">event</td> <td style="padding: 2px 10px;">explosion</td> </tr> <tr> <td style="padding: 2px 10px;">affected</td> <td style="padding: 2px 10px;">Old Market</td> </tr> <tr> <td style="padding: 2px 10px;">at-where</td> <td style="padding: 2px 10px;">XYCity</td> </tr> <tr> <td style="padding: 2px 10px;">when</td> <td style="padding: 2px 10px;">160931ZFEB09</td> </tr> <tr> <td style="padding: 2px 10px;">source</td> <td style="padding: 2px 10px;">eyeball</td> </tr> <tr> <td style="padding: 2px 10px;">reliability</td> <td style="padding: 2px 10px;">completely reliable</td> </tr> <tr> <td style="padding: 2px 10px;">credibility</td> <td style="padding: 2px 10px;">RPTFCT</td> </tr> </table>	event	explosion	affected	Old Market	at-where	XYCity	when	160931ZFEB09	source	eyeball	reliability	completely reliable	credibility	RPTFCT
event	explosion															
affected	Old Market															
at-where	XYCity															
when	160931ZFEB09															
source	eyeball															
reliability	completely reliable															
credibility	RPTFCT															

Figure 2. Feature-value matrix for explosion report

Of particular interest is that each BML statement can be easily parsed and its elements stored in the form of a feature-value matrix (Figure 2). This is important for two reasons: (1) this format preserves important semantic information, i.e., the context in which a person, event or location is mentioned; and (2) it allows fusion of individual communications through unification, a standard algorithm in

computational linguistics (Shieber 1987). Additionally, operational and background information from the deployment area which is stored in databases or ontologies is essentially also represented as feature-value matrices, thus providing the common format necessary for fusion (Jenge and Frey 2008).

Conversion to BML

In coalition endeavors, there may be numerous languages used by not only the forces involved but also open source and other information in yet further languages such as Dari used in the area of deployment. Converting available HUMINT, regardless of language, into a common format for fusion is necessary.

For conversion of natural language HUMINT texts we utilize the method described in Jenge, Kawaletz, and Schade (2009). This process of pre-analyzing natural language reports starts with information extraction (IE) based on the work of Hecking (2003), who applied IE techniques to the analysis of battlefield and HUMINT reports, and uses the freely available open-source tool GATE (<http://gate.ac.uk/>) to run the texts data through the standard IE processing pipeline. The method uses shallow information extraction techniques based on GATE. We alleviate the disadvantages of the shallow approach by using ontological knowledge about verbs and their frames. The verbs and frames we consider are taken from the HUMINT domain. The frame information attached to a verb constrains the semantic roles that can be assigned to the sentence's constituents.

The method presented for report analysis can be a component of larger systems, e.g., machine translation systems that translate reports into all languages being used in a complex combined operation. For example, at present we are working on a prototype which converts reports from German into BML for further processing.

Naturally, some reported data like sensor data first has to be dealt with by means of data fusion (Hall et al. 2001; Biermann et al. 2004) to create information that can be structured by thematic roles.

Once all underlying data and information is converted, we can map what we glean to models which describe potential threats. In the next section we will look at a simple model for representing threats.

A Simple Model For Threats

Military and security subject matter experts who are tasked with examining and evaluating intelligence information for potential threats have checklists (“rules of thumb”) of events or conditions that they watch out for as signs of potential developing threats or situations. For example, the checklist of the factors which might constitute forewarning of a potential bomb attack on a camp would include such things as the camp appearing to be under surveillance, reports that a local militant group may have acquired bomb materiel, and a direct tip from an informant concerning an attack. Many of these factors may be further broken down into more detail, the matching of which “triggers” the activation of the factor. For example, the acquisition of blasting caps would activate the trigger “hardware” in the branch “materiel” of the threat “bomb attack.” We can represent the hierarchy defined within the checklist as a simple tree-like structure as shown in Figure 3. A quite comprehensive list of precursors of potential terrorist attacks can be found in Thompson (2007).

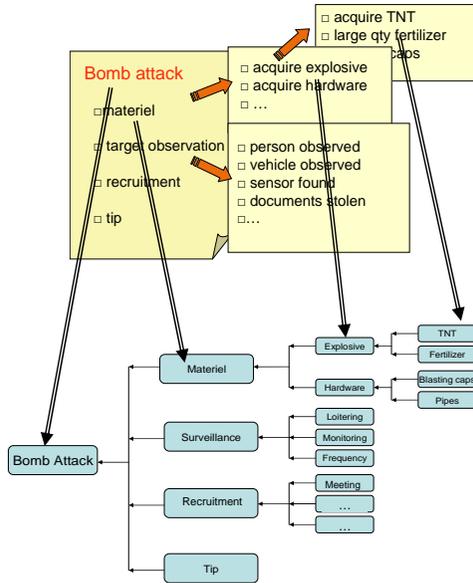


Figure 3. Checklist converted to a tree structure

However, individual observed events are not necessarily unambiguous indicators of a single specific threat. For example, in a given area of deployment the purchase of a large quantity of chemical fertilizer by an individual may be a signal of two potential threats—the construction of a bomb (a direct threat), or opium cultivation (an indirect threat). Additionally, the various elements contained within the checklist need to be related to one another in a way that makes sense. At the time the checklist is created, two things need to be done.

1. Each piece of information which triggers a specific element of one or more checklists is contained in a lookup table, that indicates which threats it triggers;
2. We define within the checklist model itself how various elements must be related in order to be relevant.

The trigger lookup table contains specific patterns (usually a BML verb plus one or more elements of the BML statement which must match). For example, using the checklist appearing in Figure 3, the lookup table would contain the BML entries “procure TNT,” “procure fertilizer,” and “procure blasting caps.” In this case, the conversion from natural language to BML would have standardized various natural language forms indicating “gained possession in some method” (i.e., bought, stole, etc.) to “procure.” If we receive a message with “procure goat” and we are not interested in tracking goats, the statement will be ignored.

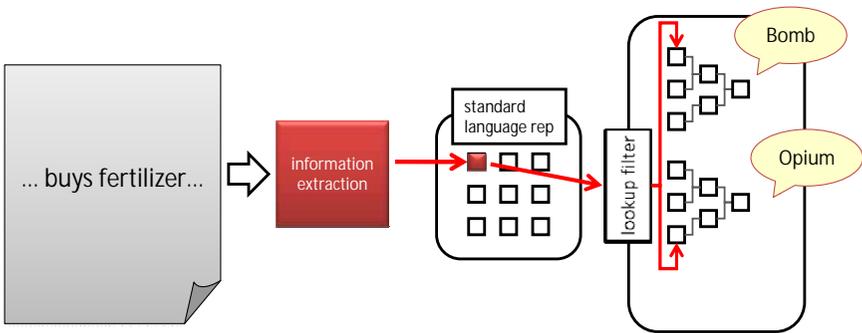


Figure 4. Mapping new information into the model

When a pattern is matched to an entry in the lookup filter, it will be passed to one or more checklists for further processing (Figure 4). An example of multiple matching would be “procure fertilizer;” in our area of activity, large quantities of fertilizer would trigger activation of the checklists of both threat of a homemade bomb or the cultivation of opium. In the case of multiple triggering, weights based upon heuristics will be assigned by the designer to each instance in the lookup table. For example, if our experience leads us to believe that acquisition of large quantities of fertilizer is much more likely to indicate opium cultivation than an IED, we would weight the lookup table entries according. This weighting is discussed further in the following section on uncertainty.

Uncertainty In The Fusion Process

Ideally, the information which comes in from the field would be reliable and unambiguous in its content but also derives from unimpeachable sources. In an ideal world, it would also be unambiguously clear which messages can be clustered together, and which messages corroborate or contradict each other. And in an ideal world, the underlying threat models to which these messages are mapped are accurate mirrors of reality.

However, the world in which we live is a messy one, far from the ideal. Sources provide information which may be ambiguous, misleading, or even contradictory. Sorting the huge amount of information and correlating various pieces of information into useful clusters poses problems through ambiguity or indirect relatedness. Incomplete information may still be significant for decision-making; however it must be acknowledged as less certain than more complete knowledge. And finally, despite our best attempts to model actions and situations, such models are seldom completely accurate. In other words, at all stages of the information gathering, analysis and fusion process uncertainty creeps in. In (Kruger 2008) three levels of uncertainty in the information fusion process were described: *data*, *fusion* and *model*.

- *Data level*, which consists of uncertainties surrounding the source and content of incoming data and information;
- *Fusion level*, concerned with the correlation of individual elements in the model and how strongly or weakly each points to a given threat;
- *Model level*, which concerns the reliability of the model itself.

Data level

At the data level, uncertainties arise from the perceived competence of the source of the information as well as the perceived truth of the information delivered by the source. The uncertainties at this level are the only uncertainties which are dynamic during processing; fusion and model level weightings are built into the model by its creator. At the data level we attempt to evaluate the quality of the information we are presented with. The uncertainties at this level revolve around the reporter's (or analyst's) belief in the credibility of the source of the report as well as the validity of the content. As described early in this paper, the assignment of reliability is based upon the original analyst evaluation (A1-F6), adjusted where necessary by linguistic analysis based upon signals such as modal expressions. For device-based information, we may have known statistics available concerning the reliability of the device which we can use as a measuring stick.

In general, it should be noted, it is next to impossible to independently rate the perceived reliability of source and content. As former CIA analyst Richards J. Heuer, Jr. (2005) notes, "Sources are more likely to be considered reliable when they provide information that fits what we already think we know." Further, we humans tend to put greater trust in the information that is delivered by a source that we trust and, of course, the converse; we mistrust the information from a source we are suspicious of. Other factors play a role in the perception of truth. Nisbett and Ross (1980) offer a thorough discussion of the fallibility of human perception.

Fusion level

At the fusion level, uncertainty arises as to how individual pieces of information are identified as belonging together (correlation-based uncertainty) and as to how clearly a given piece of information indicates the presence of (i.e., provides evidence for) a potential developing threat (evidential uncertainty).

Fusion level uncertainty provides an estimation of how strongly different pieces of information are related to one another (correlation uncertainty), as well as the likelihood that they are evidence of a particular threat (evidential uncertainty). Each of these types of uncertainty is evaluated at two levels: correlations are evaluated at the trigger and the factor level. Evidential uncertainty is evaluated at a local and at a global level.

For the purposes of fusion it is important to be able to clarify how elements should be related to each other and how closely. Clearly when, for example, the same individual's name appears in two reports, there is an indisputable connection—however, when the connection is two individuals who have a common friend, the connection is weaker. Further discussion of how various elements may be correlated is discussed in Kruger (2008).

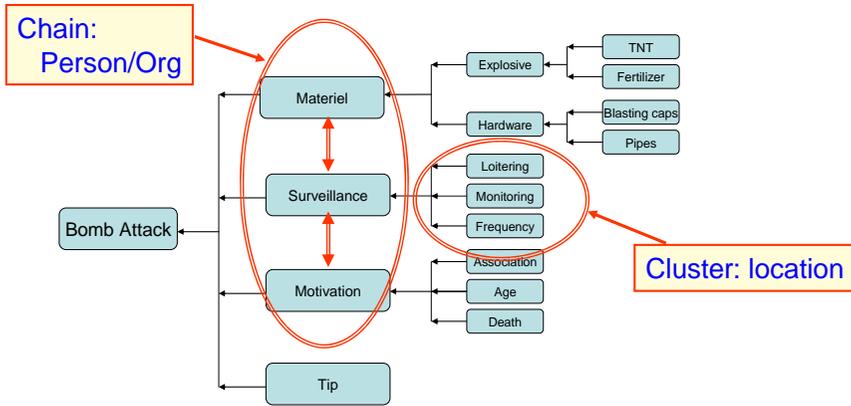


Figure 5. Representation of correlations in threat model

Additionally, we also need to determine which attributes need to be correlated and which can or should be ignored. There may be clustering based upon a common set of features for the triggers, but a different set of attributes between factors. Within the model the correlating attributes are identified at different levels. For example, the connection between surveillance activities and materiel acquisition is based upon persons or organizations, while reports concerning surveillance are clustered around a given location or facility (Figure 5). (Since the enemy will not build the bomb intended for the location at the location itself, we connect “surveillance” and “bomb” by correlating individuals or organizations.)

Evidential uncertainty determines how strongly or weakly a given trigger or factor signals a particular threat. This evidential uncertainty can be subdivided into two levels: global and local.

Local weighting for evidential uncertainty within a given structure reflects how different elements are weighted as to how significant an indicator of the threat they are. For example, while “fertilizer” may be a trigger for bomb materiel, it may not be as strong an indicator as, say, “commercial explosive,” and would therefore have a

relatively low weighting. Likewise, the fact that the same car has been observed several times parked across from the gate of the camp may be a weak warning of surveillance; the presence of an unknown monitoring device such as microphones in the facility is a clear and unambiguous sign of hostile activity (Figure 6).

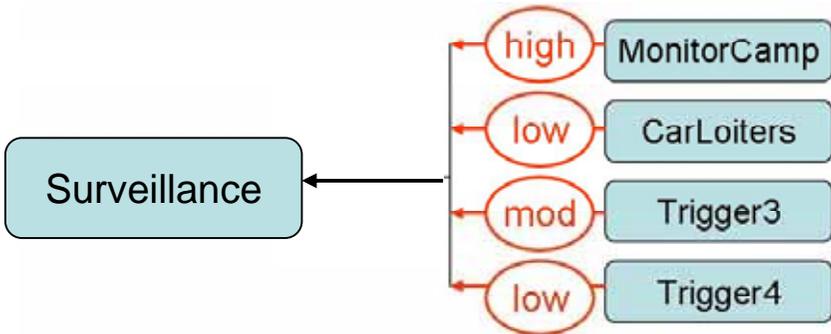


Figure 6. Local evidential weighting in threat model with weights represented semantically

A given threat factor may have multiple triggers, the sum of whose weights may be greater or less than one hundred percent (Figure 6). Our simple model at the moment foresees that they simply reinforce one another. For example, it can happen that hostile sensors are found near the military facility, a car has been observed parking in the vicinity of the front gate, and building plans have been stolen. Both the first and third trigger carry the weight “high,” as each on its own should contribute significantly to the threat likelihood. Each additional report of suspicious behavior simply increases our certainty that a given threat is building—we do not anticipate that the full list of possible triggers will be observed.

Global evidential uncertainty assigns values that reflect the probability of this trigger indicating a particular threat out of the threats in which it features. Weights according to the (heuristically determined) relative frequency of each are built into each model by the designer,

e.g., fertilizer is determined to be a weaker indicator of a bomb than of opium cultivation (for a more in-depth discussion see Kruger et al. 2008; Kruger 2008). In Figure 7, for example, it is three times more likely in this particular environment that the procurement of a large quantity of fertilizer indicates opium cultivation rather than bomb construction.

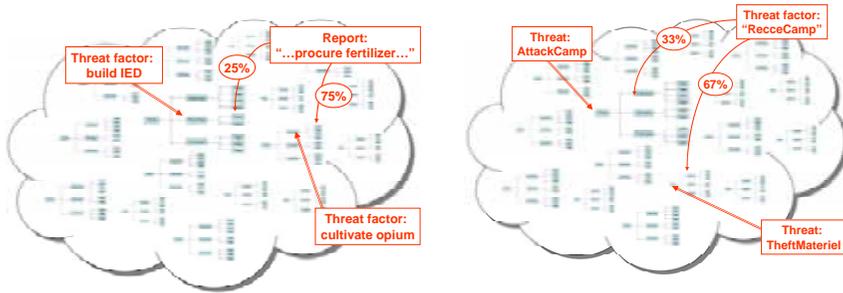


Figure 7. Global evidential weighting. Left: A single trigger (*procure fertilizer*) may activate several different threat factors. Right: A single threat factor may play a role in more than one threat model.

At the global level weights indicate relative probabilities based upon heuristics determined by the designer; threats are considered ultimately mutually exclusive from the global point of view.

Model Level Uncertainty

Up to this point, the evaluation relative weights for evidence have been assigned under the assumption that the observed event is hostile. However, in real life, there may be innocent explanations for the events observed. For example, the procurement of a large amount of fertilizer may have a completely innocent explanation—cultivation of wheat in the agrarian world in which our forces are deployed as depicted in Figure 8.

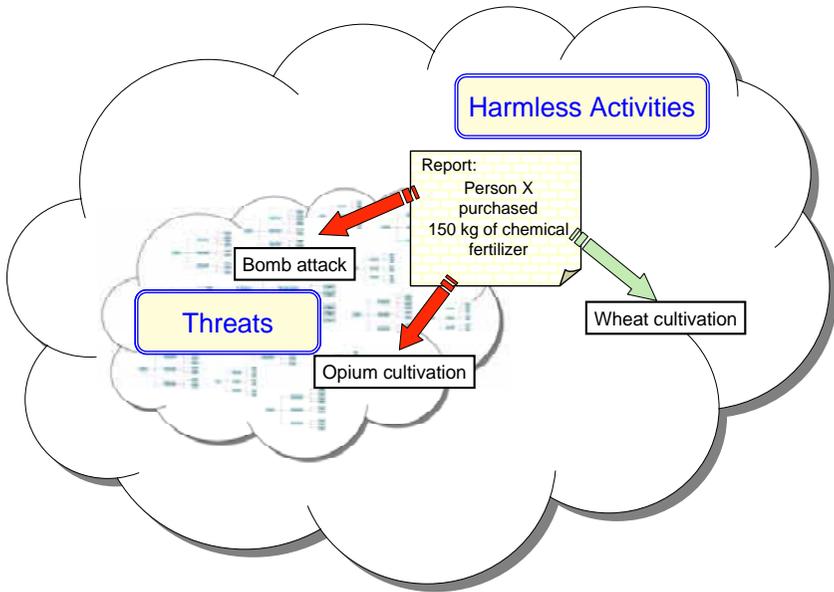


Figure 8. Observed events may not necessarily represent threats

As we are operating with incomplete knowledge of the world, how certain can we be, even when all indications appear to point to a specific danger, that this prediction is accurate (model uncertainty)—in other words, a “reality check.” The enemy tries as best possible to hide their activities, and we try to deduce from what we can observe of the enemy’s activities what is being planned. Sometimes what we think we see is in fact not there. We may learn from experience that a particular constellation of indicators results in a real threat only 25% of the time, and this should be reflected in the overall evaluation.

Using The Model For Decision-Making

A single BML statement activates a trigger. BML thus provides appropriate granularity. Activation of a single trigger activates the threat factor to which that trigger belongs.

As individual pieces of data flow into the system, these are scanned and, via lookup tables, either flow into existing (previously triggered) instances of threats or instantiate new threats. The model at this time uses a simple additive algorithm for accumulating weights for the individual pieces in each instance. This means as each new piece is added to an existing instance, the cumulative total increases, either slowly (low applicability, low correlation, low evidential weight, etc.) or rapidly (high correlation, high evidential value, etc.).

The various weights—the credibility (source, content) of the initial information, the evidential weighting between and within threat instances—interact to assure that there is a certain amount of checks and balances; unreliable information (assigned a low credibility) may trigger a strong indicator for a threat, but doubt is covered through the balance of the weights (Figure 9).

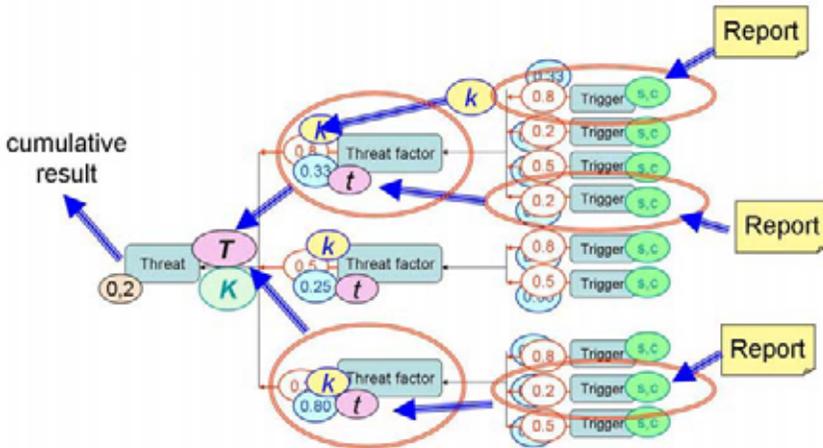


Figure 9. Flow of calculation through a checklist tree

As more information flows into a model instance, the cumulative result increases and eventually reaches a predefined threshold, at which point an analyst is informed of the potential impending threat. Since all of the underlying information will be available to

the analyst (attached through the instance), he/she will be able to evaluate the warning, examine where its strengths and weaknesses are (e.g., high credibility of source and information, overall moderate correlation, but in one area very weak correlation) and decide whether or not to modify the automatically generated evaluation, and whether or not to advise decision-makers of this potential threat.

The advantage of this type of model is that it is relatively easy to implement, in that it simply reflects the “rules of thumb” that soldiers, analysts, and others use to evaluate incoming intelligence information. It is also extremely easy to modify as one learns new patterns of behavior by the enemy, and it is easy for the user to see where the information used for the evaluation of the model comes from. In other words, it is an agile model.

It is important to keep in mind that this system is not intended to replace human decision-making, but rather to support it. The results which are produced should be viewed only as pattern detection to identify possible “what-if” scenarios. These results are passed on to a decision-maker who reviews and ultimately decides on the validity of the connections made by the fusion system. It is therefore important to consider in what form the results are presented to the decision-maker.

One possibility is that a single value is produced, the higher this value, the greater the likelihood that the threat may be building. As more supporting evidence comes in, this value would be increased, eventually bubbling up to a critical value for action. This would have the advantage for the decision-maker that relatively weakly-built instances of threats would remain active in the background, but still be available for further concretization (or deleted based upon a predetermined set of criteria). The disadvantage of the single value model would be that the decision-maker would have no idea where the strengths and weaknesses of the prediction lie.

Yet another, and perhaps the most optimal, possibility would be a hybrid of both of the above. In this variant, there would be a single numerical value for overall uncertainty but with the additional information concerning the various types of uncertainty (data, fusion, model) given in semantic form. Particularly in those deployments where lives may be at stake, the user should be able to request some more detailed information concerning how the rating was derived. This would allow the decision-maker, where appropriate, to modify or enhance the risk evaluation of a particular schema based upon diverse factors such as personal knowledge, experience, or the well-known “gut feeling.”

Conclusions

Our model provides a simple format for describing threats to support automatic fusion of incoming information. Using a standardized language such as BML as the basis for representing individual occurrences or states, we can automatically preprocess and match incoming information to simple hierarchically constructed threat models to provide early warning of potential developing threats.

The purpose of this model is to support, not replace, human decision-making, providing a rapid “first pass” evaluation of incoming information. It is not designed not to find new information or learn previously unknown connections but rather to use basic pattern recognition to quickly identify possible areas of trouble. It does not learn, but rather *models* human heuristics and experience.

The simplicity of the model contributes to its being quite fast. This model is easy to create as it depends on structures similar to the “rules of thumb” used by soldiers and analysts. Modification of the model to conform to new patterns of enemy behavior is likewise easy by adding new triggers or factors to existing models. Further, it has the advantage that decision-makers and analysts are able to

track the information which contributed to results of the system, and to easily identify strengths and weaknesses of any given instance created by the system.

References

Allied Joint Intelligence, Counter Intelligence and Security Doctrine (AJP 2.0), 2003.

Alberts, D.S., J.J. Garstka, R.E. Hayes, and D.A. Signori. 2001. *Understanding Information Age Warfare*. Washington, DC: CCRP Publications.

Alberts, D.S. and R.E. Hayes. 2003. *Power to the Edge*. Washington, DC: CCRP Publications.

Alberts, D.S. and R.E. Hayes. 2006. *Understanding Command and Control*. Washington, DC: CCRP Publications.

Alberts, D.S. 2007. Agility, Focus, and Convergence: The Future of Command and Control. *The International C2 Journal*, Special Issue, The Future of C2: 1-30.

Biermann, J., L. de Chantal, R. Korsnes, J. Rohmer, and C. Ünderger. 2004. From Unstructured to Structured Information in Military Intelligence: Some Steps to Improve Information Fusion. *NATO RTO SCI-158 Panel Symposium on Systems, Concepts and Integration (SCI) Methods and Technologies for Defence Against Terrorism*, October 25 - 27, London, United Kingdom.

Bresnan, J. 2001. *Lexical-Functional Syntax*. Malden, MA: Blackwell.

Clark, R.M. 2007. *Intelligence Analysis: A Target-Centric Approach*, 2nd edition revised. Washington, DC: CQ Press.

Costa, Paulo C.G. and Kathryn B. Laskey. 2005. Multi-Entity Bayesian Networks without Multi-Tears. Available from World Wide Web: <http://ite.gmu.edu/~klaskey/papers/Costa_Laskey_MEBN_wo_Tears.pdf>

Davis, D., C. Blais, and D. Brutzman. 2006. Autonomous Vehicle Command Language for Simulated and Real Robotic Forces. Paper (06F-SIW-004) presented at the *2006 Fall Simulation Interoperability Workshop*, September 10-15, Orlando, FL.

General Architecture for Text Engineering, GATE. <http://gate.ac.uk/>

Hall, D. and J. Llinas (eds.). 2001. *Handbook of Multisensor Data Fusion*. Boca Raton, FL: CRC Press.

Hayes, R.E. 2007. It's an Endeavor, Not a Force. *The International C2 Journal*, Special Issue, The Future of C2: 145-176.

Hecking, M. 2003. Information Extraction from Battlefield Reports. Proceedings of the *8th International Command and Control Research and Technology Symposium (ICCRTS)*, Washington, DC.

Heuer, R.J., Jr. 2005. Limits of Intelligence Analysis. *Seton Hall Journal of Diplomacy and International Relations* 49(1): 75-94.

Huijsen, W. O. 1998. Controlled Language – An Introduction. Proceedings of the *Second International Workshop on Controlled Language Applications (CLAW 98)*, May 21-22, Pittsburgh, PA: 1-15.

- Jenge, C. and M. Frey. 2008. Ontologies in Automated Threat Recognition. Paper presented at the *Military Communications and Information Systems Conference (MCC 2008)*, September 23-24, Cracow, Poland.
- Jenge, C., S. Kawaletz, and U. Schade. 2009. Combining Different NLP Methods for HUMINT Report Analysis. Presented at *NATO RTO Information Systems Technology (IST) Panel Symposium*, October 19-20, Stockholm, Sweden.
- Kruger, K. 2008. Two 'Maybes', One 'Probably' and One 'Confirmed' Equals What? Evaluating Uncertainty in Information Fusion for Threat Recognition. Paper presented at *MCC 2008*, September 23-24, Cracow, Poland.
- Kruger, K., U. Schade, and J. Ziegler. 2008. Uncertainty in the fusion of information from multiple diverse sources for situation awareness. Presented at *Fusion 2008*, June 30 - July 3, Cologne, Germany.
- Nisbett, R.E. and L.D. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Pullen, M., Corner, D., Singapogo, S.S., Clark, N., Cordonnier, N., Menane, M., Khimeche, L., Mevassvik, O.M., Alstad, A., Schade, U., Frey, M., de Reus, N., de Krom, P., LeGrand, N., and Brook, A. 2009. Adding Reports to Coalition Battle Management Language for NATO MSG-048. Paper (09E-SIW-003) presented at the *IEEE 2009 European Simulation Interoperability Workshop*, July 13-16, Istanbul, Turkey. <<http://netlab.gmu.edu/pubs/09E-SIW-003.pdf>>
- Remmersmann, T., B. Brüggemann, and M. Frey. 2010. Robots to the Ground. Proceedings of *Military Communications and Information Systems Conference (MCC 2010)*, September 27-28, Wroclaw, Poland.

Schade, U. and M.R. Hieb. 2006. Development of Formal Grammars to Support Coalition Command and Control: A Battle Management Language for Orders, Requests, and Reports. Paper presented at the *11th ICCRTS*, September 26-28, Cambridge, United Kingdom.

Schade, U. and M.R. Hieb. 2007. Battle Management Language: A Grammar for Specifying Reports. Paper 07S-SIW-036 presented at the *2007 Spring Simulation Interoperability Workshop*, March 25-30, Norfolk, VA.

Shieber, S.M. 1987. An Introduction to Unification-Based Approaches to Grammar (Volume 4 of *CSLI Lecture Notes Series*). Stanford, CA: Center for the Study of Language and Information.

Thompson, J. 2005. Precursors of Hostile Intention: Signs of a Potential Terrorist Attack. Available from World Wide Web: <<http://www.mackenzieinstitute.com/2005/precursors-attack.htm>>