Panel discussion on Challenges in Higher Level Fusion: Unsolved, Difficult, and Misunderstood Problems/Approaches in Levels 2-4 Fusion Research

Organizer

Ivan Kadar, Interlink Systems Sciences, Inc., USA

Moderators

Ivan Kadar, Interlink Systems Sciences, Inc., USA
Per Svensson, Swedish Defense Reseach Agency, Sweden

Participants

Information fusion as enabling technology for network based defence
Per Svensson, Swedish Defence Research Agency, Sweden

Challenges in High Level Information Fusion
Joachim Biermann, FGAN, Germany

The curse of JDL
Henrik I. Christensen, Royal Institute of Technology, Sweden

Unification across JDL Data Fusion Levels 1 and 2
Alan Steinberg, Utah State University, USA

Situation Awareness: Issues and Challenges
Mitch M. Kokar, Northeastern University, USA

The Levels 2, 3, 4 Fusion Challenge: Fundamental Statistics
Ronald P. Mahler, Lockheed Martin Tactical Systems, USA

Issues and Challenges in Distributed Sensor Fusion: Connection-Resource Management Aspects
Ivan Kadar, Interlink Systems Sciences, Inc., USA

Reusability, Scalability, and Quality in Higher-Level Fusion Processing
James Llinas, University of Buffalo, USA

Information fusion as enabling technology for network based defence

Per Svensson
Research director
Swedish Defence Research Agency (FOI)
Division of Command and Control Systems
Department of Data and Information Fusion

My field of activity is information fusion (IF) in military and security applications, and I am going to discuss information fusion as an enabling technology for network based defence and security.

Our research group, then considerably smaller than it is now, got considerable experience in a subfield of *inductive* strategic IF (a. k. a. strategic intelligence analysis, data mining, knowledge discovery, pattern recognition ...) in the late 80's, when we developed and installed systems and methods for intelligence analysis on alleged foreign submarine intrusions [1–3]. However, our resources for developing and testing new methodology were quite limited at the time, and since this happened long ago and is by now at most of some historical interest, I will not discuss this area further here.

Today, we work mainly on *deductive* IF for real-time tactical applications, in particular for ground warfare scenarios. The need for such methods became evident in Sweden six years ago, when the first large-scale study on a future Swedish network-based defence was being carried out. When the study's contractor, the large US defence consultant company SAIC, proposed a so-called anchor desk of military intel analysts for creating useful intelligence information from the continuous data stream created by the large set of high-performance networked sensors also proposed by the study, we noted that the anchor desk would likely get swamped by the wealth of inflowing observations, at least unless the number of analysts at the desk was far greater than the Swedish defence could afford. Thus, lacking effective real-time information fusion automation, the initially proposed solution seemed likely to fail. Eventually, the contractor saw fit to include a requirement and generic architecture for such a system in their proposal.

We have chosen to consider this information fusion problem from the deductive viewpoint of process simulation [4]. The key problems to be addressed in these applications were ground target tracking, force aggregation, and multisensor perception management. In this work, we believe that we have made good progress, although we have a long way to go before the complete multisensor resource can be automatically or semi-automatically managed by the system. Also, our methods of force aggregation is based on uncertain a priori knowledge of enemy tactical doctrine, which may not be the most relevant problem in many future conflict scenarios. However, our simulation-based deductive approach has demonstrated its power in this perhaps somewhat limited application, although new, more powerful ways to manage the collection resource remain to be found and methods for threat assessment need to be devised and demonstrated.

Here also there is a credibility gap to be bridged. Our demonstrations seem to have been best received by tactical intelligence officers, while other groups of specialists, including, e.g., many sensor researchers, remain skeptical. And it must be admitted that our simulation is just a computer game with some moderately realistic input data. We do not yet know to what extent these methods could provide critically needed robust intel analysis support to future warfighters. To reach sufficient trust in this approach, large numbers of validation tests and exercises would be needed. At least, we believe that we now have a powerful framework in which to carry out such studies.

We are developing methods for *force aggregation* based on Dempster-Shafer clustering [5, 6], *ground target tracking* in terrain based on random sets and particle filtering [7, 8], *sensor resource management,* partly based on game theory [9–11], and *policy and behaviour recognition* based on dynamic Bayesian networks [12]. In particular, I want to emphasize the importance of *integrating* such methods in a common system framework or testbed, complete with an adequate tactical scenario simulation capability. This is another focus activity of our group, i.e., the *IFD03 demonstrator,* which will be presented in a lecture at this conference tomorrow [4]. In fact, while all the methodological problem areas mentioned above fall in the *unsolved* and/or *difficult* categories stipulated by this panel, I claim that key aspects of information fusion demonstrators are well qualified to go into the *misunderstood* category, even so among fusion researchers.

Here are some reasons why advanced IF demonstrators and testbed systems are key to the development of this technology:

- to explain to rightly sceptical prospective users and customers how, why, and when IF may work in tactical applications
- to make the crucial role of IF in future network-based tactical defence systems concrete and tangible
- network-based tactical IF is a sophisticated system-of-systems concept; to make the required subsystem parameter adaptations and resource allocations in real time, one needs *as part of the IF system itself* a functionally complete model of the IF system in its environment, including its own resource management subsystem, its environment, target systems and platforms, sensors, communications, platforms, and algorithms
- to quantitatively evaluate the benefits of tactical IF in various situations, a complete, easily adaptable system-of-systems model is required, which can be used in comprehensive scenario-based Monte-Carlo simulations.

A major issue when providing tactical IF capabilities for the network-based defence of the future, an issue which we are only beginning to address, is how to deal with the changing character of armed conflict: a major component of our current force aggregation and behaviour recognition methods is based on use of tactical doctrines, imperfectly known a priori. But what is known a priori about the structure and tactical behaviour of guerilla or terrorist groups? Waltz [13] has pointed out the need for interaction between the abductive-inductive data mining process and the deductive data fusion process in such situations.

With regard to the future impact of real-time tactical IF systems, I claim that this is largely dependent on the success of two approaches:

- current efforts to integrate new, network-based high-level structure and behaviour recognition, forecasting, and resource management methods in future real-time heterogeneous distributed multi-sensor C4ISR systems
- future efforts to create *real-time* structure and behaviour recognition methods for complex systems, capable of discovering, learning, and recognizing vague and dynamically changing relationship structure and behaviour of groups, or *swarms,* of agents.

Although we cannot yet say that we have provided a solution to the problem of tactical intel analysis for ground force operations, we now know much more about the stucture of such problems. In particular, we can acknowledge the need for developing interactive fusion processes for use in situations where a priori knowledge is low. How this should be done we do not yet know.

To conclude, a major challenge for information fusion research in coming years must be the issue of how to adapt to new, asymmetric threat situations, such as international peace enforcement operations and terrorist attacks, where a priori knowledge, in particular doctrinal knowledge, is low. Applying information fusion in civilian disaster situations, perhaps even in future warfare scenarios where opponents themselves use network-based dynamically organized forces, would seem to require similar capabilities.

References

[1] Ulla Bergsten and Johan Schubert. Dempster's rule for evidence ordered in a complete directed acyclic graph. *Int. J. Approx. Reason.* 9(1):37–73, 1993.

[2] Ulla Bergsten, Johan Schubert, and Per Svensson. Applying data mining and machine learning techniques to submarine intelligence analysis. In *Proc. Third Int. Conf. Knowledge Discovery and Data Mining,* pages 127–130, Newport Beach, USA, 14–17 August 1997. The AAAI Press, Menlo Park, CA, 1997.

[3] Per Svensson. Archipelagic ASW – a difficult enterprise in need of holistic approaches. In *Proc. Int. Conf. Multisource-Multisensor Information Fusion,* pages 410–417, Las Vegas, USA, 6–9 July 1998. CSREA Press, Athens, 1998.

[4] Simon Ahlberg, Pontus Hörling, Karsten Jöred, Christian Mårtenson, Göran Neider, Johan Schubert, Hedvig Sidenbladh, Pontus Svenson, Per Svensson, Katarina Undén, and Johan Walter. The IFD03 Information Fusion Demonstrator. This conference.

[5] Johan Schubert. Evidential force aggregation. In *Proc. Sixth Int. Conf. Information Fusion*, pages 1223–1229, Cairns, Queensland, Australia, 8–11 July 2003. Int. Soc. Information Fusion, Sunnyvale, CA, 2003.

[6] Johan Schubert. Clustering belief functions based on attracting and conflicting metalevel evidence. In Bernadette Bouchon-Meunier, Laurent Foulloy, and Ronald R. Yager, editors, *Intelligent Systems for Information Processing: From Representation to Applications*, pages 349–360, Elsevier Science, Amsterdam, 2003.

[7] Hedvig Sidenbladh. Multi-target particle filtering for the probability hypothesis density. In *Proc. Sixth Int. Conf. Information Fusion*, pages 800–806, Cairns, Queensland, Australia, 8–11 July 2003. Int. Soc. Information Fusion, Sunnyvale, CA, 2003.

[8] Hedvig Sidenbladh and Sven-Lennart Wirkander. Tracking random sets of vehicles in terrain. In *Proc. IEEE Int. Workshop Multi-Object Tracking,* Madison, USA, 21 June 2003.

[9] Ning Xiong and Per Svensson. Multi-sensor management for information fusion: issues and approaches. *Information Fusion* 3(2):163–186, 2002.

[10] Ronnie Johansson, Ning Xiong, and Henrik Christensen. A game theoretic model for management of mobile sensors. In *Proc. Sixth Int. Conf. Information Fusion*, pages 583–590, Cairns, Queensland, Australia, 8–11 July 2003. Int. Soc. Information Fusion, Sunnyvale, CA, 2003.

[11] Ronnie Johansson. *Information acquisition in data fusion systems.* Licentiate thesis, KTH NADA (TRITA-NA-0328), December 2003, Stockholm, Sweden.

[12] Robert Suzic. Representation and recognition of uncertain enemy's policies using statistical models. In *Proc. NATO/RTO Conf. Military Data and Information Fusion,* Prag 20–22 oktober, 2003.

[13] Edward Waltz. *Information warfare: principles and operations.* Artech House, Norwood, MA, 1998.

Challenges in High Level Information Fusion

Joachim Biermann
FGAN -FKIE / SDF
Neuenahrerstraße 20
53343 Wachtberg
Germany
j.biermann@fgan.de

Abstract - *A challenging question in high level fusion is to develop a sound understanding of the main principles of reasoning in the human dominated area of heuristic information processing and fusion. A profound and agreed-on theory of information fusion which really reflects the nature of the tasks to be solved is lacking. Unlike level 1 fusion there is neither a commonly accepted problem definition and task model nor are there proved methods for high level information fusion. Methods for the fusion of imperfect text information should be based on a deep understanding and thorough modelling of the behaviour of battlespace factions and the cognitive processes of the human operators who perform these tasks nowadays. The qualitative nature of such models will make it necessary to have non-numeric methods controlling the overall information fusion process on JDL levels 2 and 3.*

Keywords: Behaviour model, default reasoning, closed world representation, interactive fusion system.

1. Military aspects of high level fusion

High level fusion takes place at very different places in the military command and control cycle e.g. while producing situation awareness and deducing impact assessment from current and basic information about the battlespace. Up to now, in its essence, this has been a pure human, cognitive process applying heuristic procedures based on personnel skill and experience. The appraisal of the each piece of input information and of the deduced results is not done by numerical procedures based on technical measurements but by the subjective individual evaluation and assessment of the human operator. The analysis, integration and interpretation of information is performed to support the decision making of military leaders. They take responsibility for their decisions. If high level fusion is to be automated it has to be taken into account that this involves substituting a human operation which is contrary e.g. to tracking in sensor data fusion, a task which never was done by man himself but which always was done automatically. So every methodical approach to high level fusion has not only to be scientifically sound but also psychologically convincing to be accepted by the user. One way to do so, is to analyse the human cognitive processes and to adapt the human methodical principles for automated high level fusion.

2. Characteristics of high level fusion

Regarded from a sufficient level of abstraction there is a structural similarity between sensor data fusion and high level information fusion. In both cases we have imperfect input data about the real situation, coming from several sources, which is matched against a model of the world. The resulting output is taken to update the so far known picture of the real situation.

An important difference between these two similar structured data fusion tasks lies in the following two aspects:

- the kind of data to process
- the kind of model to match against.

In sensor data fusion numerical data from technical sensors is processed to get insight into the physical aspects of the battlespace elements, their type (object classification and identification) and their space-time behaviour (tracking). It answers the question "What kind of object is it and how is it acting actually?". The relevant attributes and features are determined by the physical nature and characteristics of the objects and their dynamic behaviour can be modelled appropriately by the rules of physics. Accordingly the matching of the input data against the physical models is done by mathematical methods (probability, statistics, optimisation, artificial neural nets,..). Theses models and methods are of normative and quantitative nature and they are regarded to be objective.

In high level fusion, rather than numerical data symbolic information, documents with text and pictures, has to be handled as input. The aim of intelligence processing is not to discover and quantify single objects but to understand the behaviour and intentions of the adversary. A most reliable picture of the battlefield has to be integrated from the imperfect information and the impact on the perceived development of the military situation has to be deduced. High level fusion answers the question "What does it mean?". In order to do this intelligence officers practise by default a method of heuristic reasoning which relies on their knowledge about the standards and rules of the adversary and the assumption that the opposing forces will act according to these regulations and constraints for the benefit of their mission. It is common military experience and expert knowledge, that the production of intelligence can be done by integrating current information based on the assumption of a default behaviour. Behaviour modelling, such as doctrinal templating, is a descriptive, qualitative method of knowledge representation. Complex battlespace elements are described in a general manner using text and graphics. The descriptions include, if available, their relations and dynamic behaviour as well as their military intensions and concepts, ethnic and or religious position, their moral concepts and political involvements.

For high level fusion systems two different kinds of models are needed:
i) behaviour models describing tactics of potential adversary factions and all necessary pre-conditions for the constituting activities; (used e.g. for conflicts of military type in low and high intensity or paramilitary conflicts)
ii) models defining the "normal" situation, the common and unsuspicious behaviour; (used e.g. in war against terrorism, to be able to detect deviation from the normal).

In the first case high level fusion is the task of detecting indicators for a pattern of activity defined by the behaviour model, in the second case the task is to detect changes in the normal pattern.

3. Challenges and difficulties

So essential challenges for the development of effective automated high level fusion systems are given by the following requirements:

- to define the high level information fusion problem on a sufficient level of abstraction and to develop appropriate task models
- to develop according suitable fusion methods.

Besides that, for a fusion system it is necessary to automatically obtain access to the semantic content of the input and background information. As high level fusion is mainly based on document input, it is very important to have automated extraction of the content of the single input document. The semantic meaning of a document has to be accessible for a subsequent automated context based information analysis and integration. It is necessary to have methods and procedures for an automated text extraction and text analysis. Efficient linguistic methods and adequate ontological models of each application area are needed to understand the semantic meaning of the document input, tools which are able to cope with the information workload in network-centric intelligence and other C2 activities. Basically this is not a fusion problem in itself but as it is connected to the modelling problem, e.g. by its ontological approach to semantic text analysis, it should be regarded as part of the up to now unsolved problem of high level fusion systems.

Knowledge based information fusion is focussing on the heuristic human evaluation process. Military information processing is modelled as a context dependent and template-based heuristic reasoning process and the real task is modelled by a closed world representation. This method of approach to high level information fusion is based on the above mentioned assumption that forces are organised in a structured manner and that they operate in a military reasonable and typical way. Respective rules, doctrines, or modus operandi can be used as a basis for matching templates for default reasoning. By using this approach, the analysis and fusion of incomplete and imperfect information of military reports and background knowledge can be supported substantially in an automated system. Such models can be developed by a thorough analysis of the real processes in intelligence and command and control. This has been shown to be true not only for classical war where doctrines are the basis for operations and training, but also for terrorist activities, like in Northern Ireland or in Spain and presumably even for the activities of groups like Al Qaeda.

Describing the battlespace by default behaviour models defines a “Closed World” representation of the application area knowledge. But the information context which is relevant for a sufficient understanding of the situation is not fixed, it is open. Especially in new missions with low background intelligence and little knowledge about the behaviour of the adversary side the initially available knowledge will not be sufficient and the models will have to be modified or new ones will have to be created.

Ontological models also define a "Closed World" representation of an application area. Any concept which is not covered by an ontology cannot be taken into account during semantic analysis. As a consequence, a fusion system based on such a module will not be able to consider the missing information context, there is a blind spot for the fusion system. A solution to the obvious learning problem might be the concept of an interactive high level fusion system. It supports the human operator by providing formally correct fusion hypotheses using every information which is available to the system but leaves knowledge maintenance and final decisions to him. Such interactive systems will be adaptable to new and evolving situations.

According to the descriptive and qualitative nature of the behaviour models which are represented by non-numerical methods and the fact that high level fusion has to deduce from incomplete information, it is likely to use qualitative non-numeric methods for the matching of the imperfect information with the qualitative models. Among these are methods of non-monotonic logic like default logic and close world assumption, which both correlate very much with the human cognitive processing. High level fusion will probably benefit from quantitative mathematical methods like fuzzy logic, evidential reasoning, and probability if a suitable abstraction of the problem is reached in order to be able to define detail problems which reasonably can be formulated as quantitative questions.

4. Conclusion

Up to now high level fusion has neither established a common understanding about the character of the battlespace and its elements nor developed a suitable approach to model this area. The development of models and associated information fusion methods should be based on a thorough analysis of the human cognitive processes and the information models used. The disadvantages of “Closed World” concepts of default reasoning could be met by interactive control of automated high level fusion systems, an approach which would allow to include additional human knowledge and leaves responsibility to the operator.

The curse of JDL

Henrik I Christensen
Kungl Tekniska Högskolan
Centre for Autonomous Systems, NADA
SE-100 44 Stockholm
Sweden
hic@kth.se

Abstract – *The JDL model provides a methodology for organisation of research in fusion, it is however, important to recognize that it does not provide an architectural framework for design of systems, and as such there may be complex interaction between levels in the JDL model, which is not directly captured by the model. A reductionistic approach to research in which each of the layers are considered independently might be a danger of not showing the real complexity of information fusion systems.*

Keywords: JDL, Systems Engineering, Panel, Architecture

1 Introduction

The JDL model [1] provides a most useful division of competencies involved in information fusion systems. In a reductionistic approach to research there is, however, a danger that the model be considered an architecture for information fusion systems. In such a model the layers are considered to exist in a pure hierarchy. I.e. signal refinement is performed and the data ase sent to object refinement, and subsequently situation analysis and threat assessment. This model is convenient as it provides well defined interfaces for each of the layers. One can make assumptions about the type of data that is received and the type of feedback to generate for control of lower layers in the systems. It is, however, important to recognize that in real systems there are much more complex interaction within the system. In a purely hierarchical system the generation of alarms would have to flow through all layers, while in reality the detection of a particular type of object might indicate a threat that can be signaled within any need for situation refinement. In the same spirit a particular threat or situation might provide valuable information that calls for a change in sensory parameters, and or signal processing algorithm which can be signaled directly (possible with the intervention of the process refinement layer (4)).

The division of information fusion research into the different levels is useful for separation of research issues, but it must also be carefully considered how the different levels interact. In computer vision a similar hierarchical model was proposed by the late David Marr [2]. For almost a decade the research was locked into a particular paradigm of hierarchical processing. Simple computational arguments can be used to demonstrate that such an approach is not very fruitful for construction of real-world systems, and it does not scale to complex scenarios in terms of real-time performance in which both reactive and deliberative processes have to be used for design of systems.

It is consequently important to consider not only component research on each of the layers according to the JDL model, but also the overall system design [3], where architectures play a critical role for design of efficient systems that can operate in the real-world for analysis of real information fusion problems. Systems engineerings is here required to carefully map out the interdependencies between different layers, and the overall operation of the system. Modern methods from software design and AI [4] provide a rich source of information for the design of such systems. The reactive (event driven) mode of operation is radically different from the deliberative model of control. At the same time the set of tasks for which the system is to be used will determine the required level of generality of an architecture.

If the systems issues are not considered as part of the information fusion research there is considerable risk that research in each of the component areas make unreasonable assumption about their role of a component and other components which it relies on. From a scientific point of view it is thus essential that the systems issues are considered, and that the interdependencies between different layers are made explicit. Within such a systems approach to design of applications there is a risk that the fusion research is locked into a purely hierarchical research paradigm as seen in vision. The JDL model would then become a curse rather than the fruitful structuring of the domain envisaged by the designers of the original JDL model.

Acknowledgements

This research is in part sponsored by the Swedish Defense Research Establishment (FOI) and the Swedish Foundation for Strategic Research (SSF) through its Centre for Autonomous Systems. The support is gratefully acknowledged. The issues raised has benefited significantly from discussions with Ronnie Johansson.

References

[1] Alan N. Steinberg and Christopher L. Bowman. Revisions to the JDL data fusion model. In David L. Hall and James

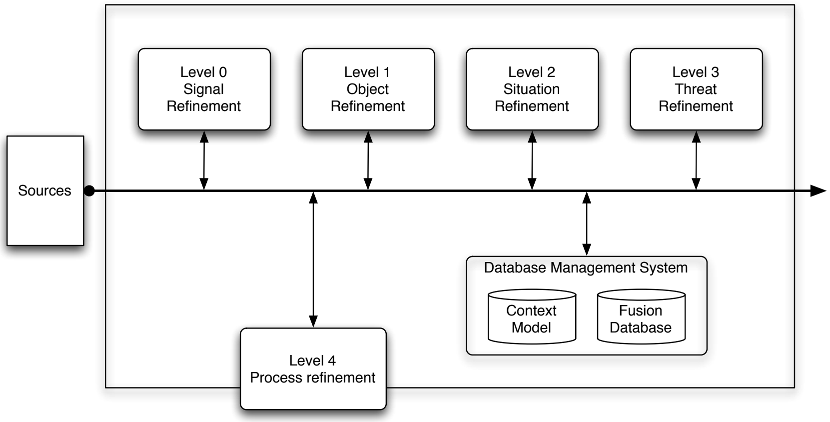

Fig. 1: The JDL Model

Llinas, editors, *Handbook of Multisensor Data Fusion*, pages 2–1 – 2–19. CRC Press LLC, Boca Raton, 2001.

[2] D. C. Marr. *Vision*. W.H. Freeman and Co., San Francisco, CA, 1982.

[3] A. Sloman and B.S. Logan. Building cognitively rich agents using the Sim_agent toolkit. *Communications of the Association for Computing Machinery*, 42(3):71–77, March 1999.

[4] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, second edition, 2003.

Unification across JDL Data Fusion Levels 1 and 2

Alan N. Steinberg
8151 Needwood Road, #T103
Derwood, MD 20855
USA
alan.steinberg@comcast.net

Abstract – *A Situation Logic model is adapted to the representation of problem in estimating and predicating states of entities and of situations. Situations are defined as sets of relationships, often with fuzzy membership. The possibility of asymmetric relations not inferable from individual target states complicates the problem of multitarget state estimation with unknown numbers of targets. This is amenable to a formulation in tern of random finite sets.*

Keywords: JDL data fusion model, functional model, situation assessment, knowledge representation, situation theory, context, combinatory logic, multi-target state estimation, random set.

1.0 Rethinking the JDL "Levels"

Recent work has occasioned a rethinking of the well-known JDL Data Fusion levels.[1] The suggested revised partitioning of data fusion functions is designed to capture the significant differences in the types of input data, models, outputs, and inferencing appropriate to broad classes of data fusion problems. In general, the recommended partitioning is based on different aspects of a situation for which the characterization is of interest to a system user.

The suggested data fusion functional levels are as follows:

- Level 0: Signal/Feature Assessment – estimation and prediction of signal or feature states;
- Level 1: Entity Assessment – estimation and prediction of entity attributive states (i.e. of entities considered as individuals);
- Level 2: Situation Assessment – estimation and prediction of the structures of parts of reality (i.e. of relations among entities & their effects on entities);
- Level 3: Impact Assessment – estimation and prediction of the utility/cost of outcome states;
- Level 4: Performance Assessment – estimation and prediction of a system's performance.

In the present paper, we concentrate on fusion levels 1 and 2 and their relation to one another.

2.0 Situations

Developments in Situation Theory [2] can be applied to estimation and prediction problems by

- Generalizing the notion of *situation* to include any structured part of reality: a single- or multi-target state, including attributes of and relations among entities;
- Using the structure of situations as means for inferring assigning prior and posterior statistics, both for "attributive" entity states and for "relational" states.

Attributes of individual entities are represented by one-place predicates, relations by $\geq$1-place predicates; i.e. attributes are 1-place relations.

Individuation of entities implies an inclusion function such as is found in such paradigmatic entities as biological organisms, purpose-made entities (e.g. an automobile or bird's nest), other discriminated spatially contiguous uniformities that persist over some practical time interval (a boulder, cloud, or swarm of gnats), or socially discriminated entities (e.g. a family or a nation).

Following Devlin, we define a *situation* as a *structured part of reality that is discriminated by some agent*. [2]

Depending on the way information is used; virtually any entity may be treated either as an individual or as a situation. For example, an automobile may be discussed and reasoned about as a single individual or as an assembly of constituent parts. The differentiation of parts is also subject to choices: we may disassemble the automobile into a handful of major assemblies (engine, frame, body, …) or into a large number of widgets or into a huge assembly of atoms.

Accordingly, the number of entities in a situation can be undecided. That is to say, the same situation can have an indeterminate number of entities, depending on the interests and focus of attention of agents reasoning about or experiencing the situation.

With such a broad definition, just about any aspect of the world, or any potential aspect thereof, can be considered a situation. We distinguish, in a natural way, between real situations (e.g. the Battle of Midway) and abstract situations or situation types (e.g. Naval Battles, 20th century Naval Battles, encounters at sea, movie subjects).

A *real situation* is a set of *facts*. *Abstract situations* are built of "pieces of information", called "infons" in Situation Theory. Infons have the general form $< R, x_1, \ldots,$

x_n,h,k,p>, where $x_1, \ldots, x_n$, are individuals (entities in our terminology), R is an m-place relation ($m \geq n$), h and k are a location and time (which may be points or extended regions) and p a polarity, i.e. a truth-value. A *fact* is simply an infon with polarity 1 in the situation constituted by the real world. Therefore, a *real situation* is a set of anchored infons (i.e. those with no unbound variables or parameters) with polarity=1 in the real world.

The symbol '¦ ' is used in representing inferences based on situational context (noting that this is not the same as Tarski's predicate-calculus implicature employing the same symbol). For a situation *s* and infon s , '*s*¦ *s*' is read as '*s* is true of *s*' or '*s* supports *s*'. Given a real situation *s*, the set of infons {*s* |*s*¦ s } is the corresponding abstract situation. [2]

Infon notation presents facts and other pieces of information as entities.[1] We can relate these informational entities to the states of constituent entities. Where $s = <R,x_1, \ldots, x_n,h,k,p>$, for *relation* R, etc., we can define a *relationship* R_s that maps from the composite state space $X^{(n)}$ to a space of relational states Y. With spatio-temporal conditioning, that would be $R_{s|h,k}$: $X^{(n)} \rightarrow Y$.

If p=1, $R_{s|h,k}(x_1, \ldots, x_n)$ maps to some relational state $y \in Y$. If p=0, $R_{s|h,k}(x_1, \ldots, x_n)$ maps to F .

If n=1, the one-dimensional relational space is attributive space; $R_{s|h,k}$: $X^{(1)} \rightarrow X$.

Some situations can be crisply defined; e.g., a chess game, of which the constituent entities and their relevant attributes and relationships are explicitly bounded. Other situations may have fuzzy boundaries. Fuzziness is present both in situation types (e.g. the concepts *economic recession* or *naval battle*) and of real-world situations (e.g. the 1930s, the Battle of Midway). Both can naturally be characterized via fuzzy membership functions. A fuzzy membership function *f* can be equated to a continuous distribution on infon polarities: If s $= <R,x_1, \ldots, x_n,h,k,p>$ and *s*¦ s, f_s(s)=*p*.

3.0 Attributes and Relationships

Level 1 fusion problems involve estimating or predicting infons headed by 1-place relations $R^{(1)}$ (or, possibly, *n*-place relations with *n-1* bound or parameterized variables). In the familiar data fusion problems, $R^{(1)}$ is often a multi-dimensional state vector of the familiar sort.

In level 2 fusion problems, context is relevant, so that an infon of interest may involve a multi-place relation.

[1] The notation of Situation Logic (as an application of Combinatorial Logic) is useful in allowing us to reason about relations as entities in their own right; e.g. in a higher-order predicate calculus, so that we can treat a relation an element of another relation. For example, in infon form, Real World¦ <illegal, cannibalism, USA, 21st century,1>. In including situations on its ontology, Situation Logic also lets us discuss and relate situations: Real World¦ <Worse, World War II, World War I, earth, 20th century,1>.

Relationships of interest in tactical military applications can include

- Relationships among objects in threat complex (deployment, kinetic interaction, organization role/subordination, comms, type similarity, etc.);
- Relationships among blue sensor & weapon platforms (spatio-temporal alignment, measurement calibration, confidence, communication/ coordination, etc.);
- Relationships between sensors & sensed entities (intervisibility, assignment/cueing, sensing, data association, countermeasures);
- Relationships between red & blue tactical entities (targeting, jamming, engaging, etc.);
- Relationships between entities of interest & other entities (terrain features; solar & atmospheric effects; weapon launch & impact points; etc.).

The state of a multi-target set X can be given in terms of infons involving a relation among its members. Because relational roles are not generally symmetrical, this must be given as a relation on some n-tuple of which the elements are the members of X;

$$R(\underline{X}) \stackrel{D}{=} R(x_1,..., x_n)\,;$$

for some n-placed relation R and n-tuple $\underline{X}$ of state vectors representing the state of n targets.

We would like to extend the finite random set formulation of multi-target state estimation to level-2 problems. Uncertainties in the truth of a proposition $R(\underline{X})$ can be expressed as distributions of multi-target states.

However, there are complications if the number of entities in a situation is unknown.

- A multi-target state of the sort of interest in situation assessment cannot in general be inferred from a set of single-target states $X = \{x_1,\ldots,x_n\}$. For example, we expect that '*x is married to y*' cannot be inferred from any set of statements of the forms '$P(x)$' and '$Q(y)$';
- Relationships of interest are often asymmetrical, such that ordering of elements is significant: infons are vectors. Therefore, a situational state estimate in which the number of targets is unknown needs the use of vectors in an unknown-dimensional space.

However, it is always possible to reduce a vector of any finite length to a set representation, as in the Wiener-Kuratowski definition of an n-tuple

$$< y_1,..., y_n > \stackrel{D}{=} \{Y \mid \exists i (i \le n \,\&\, Y = \{y_1,..., y_i\})\}; \quad (1)$$

e.g., $< y_1, y_2 > = \{\{y_1\},\{y_{1,2}\}\}$.

Thus, a relationship between two entities can be given by

$$R = <P^{(2)},\ x_1,x_2> = \{\{\ P^{(2)}\},\{\ P^{(2)},x_1\},\{\ P^{(2)},x_1,\ x_2\}\};$$

where $P^{(2)}$ randomly varies over binary relations and x_1,x_2 are familiar random entity state vectors.

The set of relationships constituting an uncertain situation $S = \{R_1,\ldots,R_m\}$ involving an unknown number of entities is one in which the R_j can differ in length; i.e., it is a set whose elements have the form given by (1). So S is a finite random set whose elements include both random entity state vectors x_i and random relational state vectors $<P, x_{1\ldots}, x_n>$. The battery of finite set statistic should apply in characterizing such level 2 inference problems. [4]

4.0 Situational Inferences

A situation can imply and can be implied by the states and relationships of constituent entities. Just as in Level 1 inferencing (i.e. with 1-place relations), we can write production rules based on logical, semantic, causal, customary or material (etc.) relationships among predicates of any length. The disposition of players and the configuration of the playing field are indicators that the situation is a baseball game, a bullfight, a chess match, an infantry skirmish, an algebra class, or a ballet recital: $P^{(n)}(X^{(n)}) \Rightarrow S$.

Often situational inferences can be given in the form of Boolean combinations of quantified expressions:

$$\exists n \exists x_1,\ldots,\exists x_n [P^{(n)}(x_1,\ldots,x_n)] \Rightarrow S;$$

or, more succinctly,

$$\exists n \exists \underline{X}^{(n)}[P^{(n)}(\underline{X}^{(n)})] \Rightarrow S.$$

To be sure, we are generally in short supply of situational ontologies that would enable us to write such rules. This is a job for expert system builders. For many types of situations, fuzzy rules will be required.

We can generalize from the familiar level-1 types of inferences, in which entity attributes are inferred from other attributes:

$$f<r^{(1)}(x)|Q^{(1)}(x),S); \qquad \text{(L1? L1 deduction)}$$

(e.g. single target likelihood functions or Markov transition densities). Relations between pairs of entities, or among n-tuples of entities, can similarly be inferred from other relations among them:

$$f(P^{(2)}(x,y) \mid Q^{(2)}(x,y),S) \qquad \text{(L2? L2 deduction)}$$
$$f(P^{(n)}(x_1,\ldots,x_n) \mid Q^{(n)}(x_1,\ldots,x_n),S)$$

This pattern of Level 2 ? Level 2 deduction would include, for example, multi-target likelihood functions or multi-target Markov transition densities familiar in finite random set statistics.

Cross-level inference patterns include the following:

- $f<r^{(2)}(x,y)|\, Q^{(1)}(x), R^{(1)}(y),S)$ (L1? L2 deduction);
- $f(\exists y[P^{(2)}(x,y)]|\, Q^{(1)}(x),S)$ (L1? L2 induction);
- $f(P^{(1)}(x)|\, Q^{(2)}(x,y),S)$ (L2? L1 deduction).

5.0 Multi-Target Estimation and Prediction

Thus armed, we may distinguish among problems in state estimation and tracking multiple targets.

We may distinguish the following types of motion models:

a) Independent Target Motion: each target's motion is not affected by that of any other entity; so multitarget prior probability density functions (pdfs) are simple products of the single target pdfs.

b) <u>Context-sensitive Single Target Motion</u>: at each time-step, a target x responds to the current situation, i.e. to the states of other entities, which may be dynamic but are assumed not to be affected by the state of x.

c) <u>Interacting Multiple Targets</u>: at each time-step, multiple entities respond to the current situation, which may be affected by the possibly dynamic state of other entities.[2]

We may similarly distinguish the following types of measurement models:

a) <u>Independent Measurements</u>: measurements of each target are not affected by those of any other entity; so multitarget posterior probability density functions are simple products of the single target posterior pdfs.

b) <u>Context-Sensitive Multiple Target Measurements</u>: measurements of one target may be affected by the state of other entities as in cases of additive signatures (e.g. multiple targets in a pixel), occluded or shadowed targets. Other cases involve induced effects; e.g. bistatic illumination, electromagnetic interference or disruption of the observing medium.[3]

References

[1] Steinberg, A. and Bowman, C., "Rethinking the JDL data fusion model", *Proc. 2004 MSS National Symposium on Sensor and Data Fusion,* 2004.

[2] Steinberg, A. and Bowman, C., "Revisions to the JDL data fusion model", Chapter 2 of *Handbook of Multisensor Data Fusion*, ed. David L. Hall and James Llinas, CRC Press, London, (2001).

[3] Devlin, K., *Logic and Information*, Press Syndicat e of the University of Cambridge, 1991.

[4] Mahler, R., "Random set theory for target tracking and identification", Chapter 14 of *Handbook of Multisensor Data Fusion*, ed. David L. Hall and James Llinas, CRC Press, London, (2001).

---

[2] Tracking targets with type (a) dynamics is clearly a level-1 (i.e. independent-target) problem. Type (b) and (c) dynamics are often encountered, but are generally treated using independent-target trackers; perhaps with a context-dependent selection of motion models, but assuming conditionally independence among target tracks. Type (c) cases, at least, suggest the value of trackers that explicitly model multi-target interactions.

[3] The absent category, Interacting Multiple Measurements, is actually a hybrid of the listed categories, in which entities affect one another's state and thereby affect measurements.

Situation Awareness: Issues and Challenges

Mieczyslaw M. Kokar
Department of Electrical & Computer Engineering
Northeastern University
360 Huntington Avenue
Boston, MA 02115
kokar@coe.neu.edu

1 Basic Situation Awareness Concepts

To explain what I mean by "situation awareness" I often use the example of watching the games like American football or baseball. Since I have never learned the rules and the strategies of these games, when watching them on TV, although I can clearly see where each player is and where the ball is, I still have no idea of what is going on. Clearly, in this case I cannot claim that I am aware of the situation.

The term "situation awareness" has been often interpreted in a somewhat simplistic way as merely the knowledge of all the objects in a specific area, and possibly their kinematic states. It is clear, however, that the meaning of this term implies more than that. For instance, in [1], awareness is explained as "AWARE implies vigilance in observing or alertness in drawing inferences from what one experiences." In other words, a subject is aware if the subject not only observes (experiences) the objects but also is capable of drawing conclusions from these observations.

The next question is about the meaning of being aware of a specific situation. The word "situation" is defined in [1] as: "1 a : the way in which something is placed in relation to its surroundings." So the essence of situation awareness is the knowledge of all the objects that are relevant to the subject and the knowledge of the relations that the subject is in with respect to these objects. In other words, awareness means knowing not only objects but also relations.

The notions of "awareness" and "situation" are not totally independent in the sense that they both have one common aspect, the aspect of "relation". In mathematics, a relation is a subset of the Cartesian product of a number of sets. For instance, the Cartesian product $A \times B$ of two sets A and B is the set of all ordered pairs $< a, b >$, i.e., $A \times B = \{< a, b >, a \in A, b \in B\}$. A relation R is then a subset of the Cartesian product, $R \subset A \times B$.

In logic, relations are treated as interpretations of predicates. While a predicate is a syntactic concept, a relation is a semantic counterpart of the predicate. Awareness requires the capability of drawing conclusions, where the process of drawing conclusions is also called inferencing or reasoning. In logic, reasoning is a process of applying inference rules to either the axioms of a given theory or to previously drawn conclusions (theorems) with the intent of deriving new theorems. The axioms and the theorems are expressed in terms of predicates. Since, as we pointed out earlier, the predicates are interpreted as relations, the inference process derives true statements about relations.

Logical reasoning viewed as the application of inference rules is a good model for the automation of the reasoning process. Once premises and axioms are declared and a conjecture is posted, the process of proving that a conjecture is true (i.e., that it is a theorem) can be viewed as a search in which inference rules are matched against the axioms and already derived conclusions until a newly derived conclusion is equal to the conjecture. Mathematical reasoning, on the other hand, relies more on the ingenuity of the human to derive proofs. In mathematical reasoning models, i.e., sets, functions and relations, play a more explicit role than they do in logical reasoning. In other words, mathematical reasoning can be viewed as a process of manipulating the semantic objects rather than the syntactic objects.

The purpose of this discussion was to stress the fact that while situation awareness clearly requires an inference capability, the inference process itself can be carried out using either a mathematical or a logical framework. As we mentioned above, the two frameworks are duals of each other and we wanted to make this fact explicit. The advantage of a logical framework is the flexibility that it provides, i.e., one has only declare axioms and premises and the reasoning process will proceed automatically. This approach is also termed as declarative programming. Automation is more difficult using traditional programming languages, but the advantage is that more complex theorems can be proven at a lower computational cost. This approach is termed procedural programming.

In addition to reasoning about relations, situation awareness involves one more aspect that has to do with the use of the concept of situation in real life. While a situation can be defined as a set of relations with other objects, both the objects and the relations change with both time and location. For instance, I am in different situations when I am driving home and when I am hiking in the mountains. To make use of situation awareness one must be able to recognize situations, assess their impact on one's goals, memorize situations, associate various properties with particular situations, communicate descriptions of situations to others. This leads to the requirement that situations be treated as objects, similarly like physical objects or conceptual objects.

The issue of situations as objects has been studied in logic. Here we refer only to the work of Barwise [2], but many others had contributed to the formalization of this term. Here is a quote from Barwise: "One of the starting points for situation semantics was the promotion of real situations from second class citizens to first class citizens. By a situation, then, we mean a part of reality that can be comprehended as a whole in its own right - one that interacts with other things. By interacting with other things we mean that they have properties or relate to other things."

The objects may be complex, incorporating both other objects as well as their properties and relations among the objects. It can even include other situations. In our work, we have developed an ontology for situation awareness - we call it Core SAW Ontology [3]. This ontology is available on our web site in the OWL representation. The discussion of this ontology can be found in [4]. We encourage the fusion community to use it as a starting point for developing more complex ontologies. This ontology is extendible to accommodate more complex domain ontologies. The use of the same ontology promotes interoperability among different systems.

2 Challenges

Situations as Objects: Situation awareness in Level 2 fusion is considered to be a computer-based process that can recognize and manipulate situations. The computer is also assumed to be responsible for making decisions depending on the situation. This is different from human-oriented situation awareness, where the human needs to be aware and then use this awareness for decision making. This kind of a requirement for a computer system creates many challenges that need to be addressed by the developers of such computer systems.

Relation Derivation Algorithms: How can a program determine whether a particular relation holds or not? For instance, in order to determine that two numbers, say 2 and 5, are in the "$\leq$" relation we need to have a procedure to make such a conclusion. Either we remember this as an atomic fact, or we apply a rule to establish such a fact. For instance, we may remember the ordering of all single-digit numbers and then compare the number of digits in a number. If one of the numbers has more digits and the first digit is not a "0", then it is larger than the other one. Obviously, this is not a complete rule, but it should be clear how such a rule could be specified. In general, for each relation a rule is needed to derive its validity. For any domain, a set of rules needs to be developed.

Relevance of Relations: The next question is which of the relations should be derived or monitored? Consider, for instance, that there are 100 objects in the area. Even if we consider only binary relations, the number of possible binary relations is equal to $2^{10,000}$. This is because the cardinality of the Cartesian product in this case is 100×100 and then there are as many relations as there are subsets of such a set. Typically, only a subset of these relations would be relevant. The challenge then is to develop methods for determining which of the possibly many relations are really relevant and thus should be monitored.

Complexity of Derivation Algorithms: In the previous section we mentioned two approaches to the derivation of relations - declarative (or logical) and procedural (or mathematical). Using the declarative approach it is relatively easy to specify all the facts that the derivation algorithm should consider. The derivation algorithms are (conceptually) simple search algorithms. However, such a search is exponentially complex. The challenge then is to find algorithms that can "understand" their own limitations and can assess their own abilities given the resources. On the other hand, the procedural approach would require the specification of domain-specific derivation procedures during the design of a situation awareness system. The burden then is on the designer to develop a relatively complete set of such procedures.
Certainty of Derived Relations Independently of which approach to relation derivation is chosen the issue of uncertainty of derived decisions is a great challenge. It is highly unlikely that the designer of a situation awareness system will provide a complete coverage of all relations that will potentially be needed by the user of such a system. Incorporating the computation of uncertainty of decisions adds to the complexity of this task. The logic-based solution is not a simple case, either. Although many attempts to combine, for instance, Bayesian reasoning with logical reasoning have been documented, this task is still a great challenge that needs to be addressed.

References

[1] Merriam-Webster Online, 2004. http://www.m-w.com/.

[2] J. Barwise. The situation in logic. CSLI Lecture Notes, 1989.

[3] SAW Core Ontology. http://vistology.com/onts/onts.html.

[4] C. J. Matheus, M. M. Kokar, and K. Baclawski. A Core Ontology for Situation Awareness. In Proceedings of the Sixth International Conference on Information Fusion, pages 545 –552, 2003.

The Levels 2, 3, 4 Fusion Challenge: Fundamental Statistics

Ronald Mahler
Lockheed Martin MS2 Tactical Systems
3333 Pilot Knob Road
Eagan MN 55121
USA
ronald.p.mahler@lmco.com

Abstract – *This is a position paper for the FUSION04 panel "Challenges in Higher Level Fusion: Unsolved, Difficult, and Misunderstood Problems/Approaches in Levels 2-4 Fusion Research." Multi-Sensor Integration (MSI or Level 1 fusion) is an increasingly mature field. However, the higher levels of information fusion—situation assessment (SA or Level 2), threat assessment (TA or Level 3), and sensor/resource management (SM or Level 4) present daunting theoretical and practical challenges. Furthermore, SM for SA and/or TA (SM-SA/TA) is an even more daunting challenge than SM for MSI (SM-MSI). We summarize the challenges associated with the higher levels of fusion. Then we argue that in the absence of unifying theoretical foundations based on fundamental multitarget statistics, the higher levels of fusion will remain indefinitely trapped in a "blindly groping for solutions" stage of maturity .*

Keywords: situation assessment, threat assessment, sensor management, resource management.

1 Introduction

The higher levels of information fusion—situation assessment (Level 2), threat assessment (Level 3), and resource/sensor management (Level 4)—present daunting theoretical and practical challenges. Multisensor integration (MSI), or Level 1 fusion, deals with the problem of estimating and predicting the number, positions, velocities, identities, etc. of individual targets. While MSI is an increasingly mature field, this is definitely not the case with Levels 2, 3, 4 fusion. In this position paper we argue that higher-level fusion will remain indefinitely trapped in a "blindly groping for solutions" stage of maturity unless they can be formulated in terms of a unifying statistical foundation.

In section 2 we summarize the challenges associated with Levels 2, 3, 4 fusion. In section 3 we summarize recent research conducted by us that, we believe, may lead to more systematic understanding of higher-level fusion.

2 The challenge of Levels 2, 3, 4 Fusion

• *Situation Assessment* (SA) or Level 2 fusion deals with the often ill-defined problem of estimating and predicting tactically significant relationships between "targets" that can be quite diverse in nature: individual targets with inherent tactical significance (e.g., mobile missile launchers); group targets (tank columns, armored infantry battalions); fixed emplacements (towns, bridges, airports, radar and television transmitters, command centers); and terrain features (forests, roads, towns, caverns, weather).

Much attention has been devoted to the specific SA problem of group target detection and tracking. But as for SA in the broader sense, it has been very unclear how to even formulate problems in general terms amenable to solution by algorithmics—let alone how to develop effective general algorithms. The few theoretical and other efforts that exist have typically addressed limited special cases, typically via case-by-case heuristics.

• *Threat Assessment* (TA) or Level 3 fusion. Whereas SA attempts to assemble tactically meaningful relationships from the information provided by MSI, TA attempts to infer level of threat, and proper tactical responses to such threat, from aggregated SA information. TA is even less mature as a field than SA. Once again, it is very unclear how to formulate problems amenable to algorithmics; and the few efforts that exist typically employ heuristics to deal with special cases.

• *Sensor Management for MSI* (SM-MSI). The purpose of SM-MSI is to "direct the right sensor on the right platform to the right target at the right time." It is clear that SM-MSI is inherently a problem in nonlinear optimal control theory, and much research has been conducted in recent years based on control-theoretic paradigms (though often not acknowledged as such). Two basic questions have hindered progress, however. First, how does one define control-theoretic objective functions that result in tactically desired allocative actions? Second, given the daunting combinatorics of multisensor-multitarget problems, what computationally tractable approximations will produce "optimal-enough" behavior?

• *Sensor Management for SA/TA* (SM-SA/TA). The purpose of SM-SA/TA should be to optimize knowledge about current or predicted tactically significant events, about their relative threat levels, and about possible responses to them. But this is not possible without, at minimum, a meaningful, algorithmically implementable specification of (1) the terms "tactically significant" or "threat"; and (2) objective functions that define what it means to optimize knowledge about tactical significance or threat.

3 Fundamental multitarget statistics

We summarize what, we believe, is required to develop a unifying fundamental statistics for higher-level fusion, based on existing fundamental statistics for MSI.

3.1 Multi-Source Integration (MSI)

Finite-set statistics (FISST) theoretically unifies the major aspects of MSI (including expert-systems theory and multisensor-multitarget statistics) under a single probabilistic umbrella [1, 5, 8, 10, 11], and is the subject of an invited plenary talk at this conference [9]. The FISST synthesis generalizes the familiar Bayesian "Statistics 101" engineering formalism, including formal Bayes modeling [4, 5, 8, 10], while avoiding the limitations of heuristic "plain-vanilla Bayesian" approaches [2, 4]. The fundamental statistics of a multitarget system are represented as multitarget posterior distributions, belief-mass functions, or probability generating functionals (p.g.fl.'s). FISST has been considerably extended and refined, especially in regard to the development of principled approximation strategies such as the probability hypothesis density (PHD) filter [6]. Such strategies depend on the FISST multitarget differential and integral calculus, especially the concept of a functional derivative of a p.g.fl.

3.2 Sensor management for MSI (SM-MSI)

SM-MSI is inherently a problem in nonlinear optimal control theory, but generalized to stochastic multi-object systems (targets, data, sensors, platforms). The FISST multitarget calculus provides a means of devising a unified control-theoretic approach, as well as principled approximations of it. Our approach is based on jointly approximated: (1) multitarget filters (the PHD and the multi-hypothesis correlator (MHC) filters); (2) "probabilistically natural" control-theoretic objective functions (the posterior expected number of targets of interest, or PENTI); and (3) optimization-hedging strategies to account for the uncertainties created by unknowable future observation-collections [7].

3.3 Situation/threat assessment (SA/TA)

We have shown that FISST can be extended to provide a theoretical foundation for group-target detection, tracking, and identification, and that the PHD filter of section 3.2 can be generalized to produce an approximate "group PHD" filter [3]. We have also shown how to incorporate concepts of target tactical significance into the fundamental statistics of multitarget systems, based on a "dynamic situational significance map" that mathematically specifies the meaning of "tactical significance" for a given theater at a given moment [7].

Even so, this is just a start. The approach in [3] should be generalized to include interacting multitarget motion models, which account for coordinated motion among units in group targets. Methods for assessing threat, potential threat, and intent must be devised.

3.4 Sensor management for SA/TA

We have proposed the elements of an approach for SM-SA/TA. It is based on three innovations: (1) the dynamic situational significance map of section 3.3; (2) a generalization of the PENTI objective function of section 3.2 to group targets and/or ordinary targets of tactical interest; and (3) the PHD and MHC approximate multitarget filters of section 3.3. Under this approach, sensors are directed to preferentially to collect observations from targets with actual or potential tactical status. As time progresses, the targets that remain visible are only those that are currently or potentially of active tactical importance.

References

[1] I. Goodman, R. Mahler, and H. Nguyen, *Mathematics of Data Fusion*, Kluwer, New York, 1997.

[2] R. Mahler, "Bayesian vs. 'plain-vanilla Bayesian' multitarget statistics," in I. Kadar (ed.), *Signal Processing, Sensor Fusion, and Target Recognition XIII*, SPIE Vol. 5429, to appear, 2004.

[3] R. Mahler (2002) "An extended first-order Bayes filter for force aggregation," in O. Drummond (ed.), *Signal and Data Processing of Small Targets 2002*, SPIE Vol. 4729, pp. 196-207

[4] R. Mahler, "Formal vs. heuristic modeling for multitarget Bayesian filtering," in O. Drummond (ed.), *Signal and Data Processing of Small Targets 2004*, SPIE Vol. 5428, to appear, 2004.

[5] R. Mahler, An Introduction to Multisource-Multitarget Statistics and Its Applications, Lockheed Martin Technical Monograph, Eagan MN, 2000.

[6] R. Mahler, "Multitarget filtering via first-order multitarget moments," *IEEE Trans. AES*, 39(4), ???, 2003.

[7] R. Mahler, "Multitarget sensor management of dispersed mobile sensors," in S. Butenko, R. Murphey, and P. Paralos (eds.), *Theory and Algorithms for Cooperative Systems*, Kluwer, New York, 2004.

[8] R. Mahler, "Random set theory for target tracking and identification," in D.L. Hall and J. Llinas (eds.), *Handbook of Multisensor Data Fusion*, CRC Press, Boca Raton FL, 2002.

[9] R. Mahler, "Random sets: Unification and Computation for information fusion—a retrospective assessment," *Proc. 7th Int'l Conf. on Information Fusion*, Stockholm, June 28-July 1, 2004.

[10] R. Mahler, "'Statistics 101' for multisensor, multitarget data fusion," *IEEE AES Mag., Part 2: Tutorials*, 19(1), 53-64, 2004.

[11] R. Mahler, "Unified Bayes multitarget fusion of ambiguous data sources," *Proc. Int'l Conf. on Integration of Knowledge Intensive Multi-Agent Systems*, 343-348, Cambridge MA, Sept. 30-Oct. 4, 2003.

Issues and Challenges in Distributed Sensor Fusion: Connection-Resource Management Aspects

Ivan Kadar
Interlink Systems Sciences, Inc.
1979 Marcus Avenue, Suite 210
Lake Success, NY 11042
USA
ikadar@systemssciences.com

Abstract - *Connection Management (CNM) plays a key role in spatially and geographically diverse distributed sensor networks. The role of Connection Management is to provide seamless demand-based resource-allocation and sharing of the information products. For optimum distributed information fusion performance, these systems must minimize communications delays and maximize message throughput, reduce or eliminate out-of-sequence measurements, take into account data pedigree and at the same time optimally allocate bandwidth resources and encode track data (sources of information) for optimum distributed estimation of target state. In order to achieve overall distributed "network" effectiveness, these systems must be adaptive, and be able distribute data on demand basis in real-time. While the requirements for these systems are known, research in this area has been fragmented. Challenges and reasons for lack of coherent research in this area are illustrated highlighting the need for a multi-disciplinary approach among communications, estimation and information theory, networking, optimization and fusion communities. A CNM concept is illustrated for optimum demand-based bandwidth scheduling.*

Keywords: Distributed sensor networks, distributed information fusion, connection management requirements and methods.

1 Introduction

It is well known that Connection Management (CNM) plays a key role in both distributed "Local" Network-Centric and "Globally" Connected Information-Centric [1] spatially and geographically diverse distributed sensor networks. The role of CNM is to provide seamless demand-based resource allocation sharing of the information products. For optimum distributed information fusion performance, these systems must minimize communications delays and maximize message throughput, reduce or eliminate out-of-sequence measurements [2], take into account data pedigree and at the same time optimally allocate routing and bandwidth resources, and optimally encode track/tracklet [3] data (i.e., to reduce number of bits required to represent sources of information) for optimum distributed estimation of target state. In order to achieve overall distributed "network" effectiveness, these systems must be adaptive, and be able distribute data on demand bases in real-time requiring an interface between the System/Sensors Resource Manager (SRM) and the CNM system. While the requirements for these systems are known, research in this area has been highly fragmented.

This problem, in part, can be attributed to the fact that CNM requires a multi-disciplinary approach, viz., understanding problems in estimation theory, all levels of fusion (predominantly Levels 1 and 4), communications, information theory/source coding, networking and optimization theory. In a practical sense this infers the need for close collaboration among practitioners in multiple disciplines.

This author has worked in most of the domains enumerated and thereby hopes to provide an objective assessment of the issues, highlight potential approaches and identify challenges in this important area of distributed fusion research.

2 Issues and Challenges

2.1 Issues

A key concern of multi-platform distributed data fusion [4-6] is maintaining and assessing the quality and utility of data shared among platforms. Constraints such as bandwidth limitations that restrict data flow and data generation necessitate the development of data encoding schemes and on-line adaptive demand driven scheduling that computes a metric to determine which platform can provide the best data for processing [7-10]. Several papers have addressed components of the CNM problem [7-9, 11-14] ranging from user data needs selection based on information theoreric and utility theory formulations for a SRM, covariance control for SRM, source coding with network constraints, and optimum channel allocation without explicitly addressing target state estimates.

The purported time-compression advantage of network-centric data fusion is known [1], that is, the ability of distributed fusion to combine spatially and geographically diverse sensors to improve fused track accuracy while reducing the time required over a stand-alone sensor. However, it should be noted at this point, that the "time-compression" advantage depicted does not take into account connection management (CNM) and data processing delays. As a matter of fact, the timely management and scheduling of bandwidth, in order to minimize delays and maximize throughput [13, 18], is one of the neglected ingredients of CNM.

2.2 Challenges

In order to achieve optimum connection resource allocation (i.e., CNM), the operational performance of a collection of networked sensors has to be examined subject to a number of interrelated coherent factors some of which may implicitly dependent on Level 1 algorithms and on interaction with the System/Sensors Resource Manager (SRM):

- the relative geometry among the sensors and the targets,
- the performance of the individual sensors,
- expected message time delays among the component sensors,
- communications bandwidth and networking limitations,
- distributed knowledge of the capabilities of the component sensors,
- the system tasking capabilities depending on available assets (an SRM function), and
- optimum or near optimum distributed estimation (fusion of sensors), subject to taking into account information pedigree in order to avoid double counting (and associated rumor propagation), elimination of out-of-sequence (OOS) measurements [2], data rate and network/channel bandwidth constraints [13, 14] .

In order to achieve the above goals, one needs at least to implement:

- robust distributed state estimation methods subject to information rate constraints, requiring
- optimum source coding schemes applied to the sensor track/tracklet data with respect to user defined measures-of-performance (MOP), such as rate distortion theory and associated Kullback-Leibler distance based methods [15-17] that minimize data rate (i.e., the number of bits required to represent track/tracklet data) and maximize information, and minimize the error in the target state estimate [14],
- and bandwidth allocation and routing schemes on a demand basis using optimum routing and channel allocation scheduling that maximizes throughput and minimizes delays by allocating data exchange among distributed sensors in the network [13, 18-19] and potentially reduce the need for source coding.

Questions that naturally arise include:

- How does data and fused estimate quality affected?
- What metrics are adequate and appropriate?
- Should relative sensors-to-target geometry be taken into account (while not effecting connectivity, it is indirectly effecting coverage and fused target state estimates)?
- Should data rate be unconstrained and optimally allocate channel/bandwidth resources vs. constrained on needed rates?
- What combination of approaches yield optimum fused target state estimate while minimizing overall network response time, maximizes throughput and minimizes inter-node delays (providing the operational benefit of spatially diverse distributed networked sensors over a real-time single sensor)?

In order to further highlight some of issues and challenges, a CNM system concept, shown in Fig.1, illustrates an optimum demand-based capacity allocation and information flow scheduling scheme. The CNM scheduling algorithm is subject to minimum delay and maximum throughput constraints, while taking into account the necessity of source coding methods at Level 1 and, interface with SRM to achieve optimum or near optimum target state estimates in real-time.

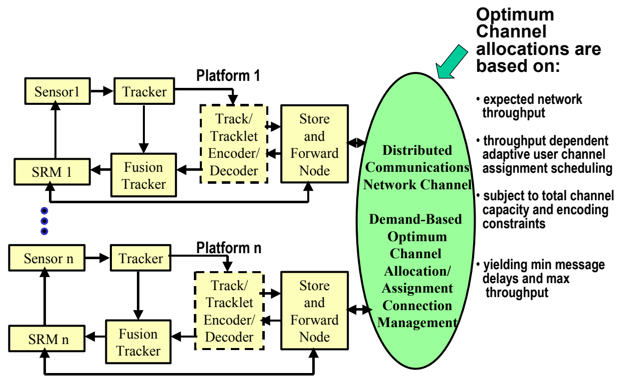

Fig. 1 A CNM Construct: Optimum Demand-Based Interacting Channel/bandwidth allocation

References

[1] D. S. Alberts, J. J. Garstka and F. P. Stein, Network Centric Warfare: Developing and Leveraging Information Superiority, *DoD C4ISR Cooperative Research Program Publications Series*, 1999.

[2] Yaakov Bar-Shalom, Update with Out-of-Sequence Measurements in Tracking: Exact Solution, to appear in the *IEEE Trans. on Aerospace and Electronic Systems.*

[3] Oliver Drummond, Tracklets and Hybrid Fusion with Process Noise, *Signal and Data Processing of Small Targets,* Oliver Drummond, Editor, SPIE Proceedings Vol. 3163, No. 6, 1997.

[4] Martin. Liggins, C.Y. Chong, I. Kadar, M. Alford, V. Vannicola and S. C. A. Thomopoulos, Distributed Fusion Architectures and Algorithms for Target Tracking, *Proc. IEEE*, *Vol. 88, No.1, January 1997.*

[5] Chee Y. Chong, S. Mori and K.C. Chang, Distributed Multitarget Tracking, *in Multitarget Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed., Artech House, 1990.

[6] Chee Y. Chong, E. Tse and S. Mori, Distributed Estimation in Networks, *Proc., American Control Conf.*, pp. 294-300, San Francisco, CA, 1983.

[7] P. Greenway, R.H. Deaves and D.H. Bull, Connection Management in Decentralized Data Fusion Systems, *Proc. of the 1996 IEEE International Conference on Multisensor and System Integration for Intelligent Systems*, Washington, D.C., pp. 8-11, December 1996.

[8] R. H. Deaves, D. Nicholson and P. Greenway, An Evaluation of Connection Management in a Decentralized Target Tracking System, *Proceedings Signal Processing, Sensor Fusion and Target Recognition VI*, Ivan Kadar, Editor., Proc. SPIE, Vol. 3068, pp. 21-24, April 1997, Orlando, Fl.

[9] R. H. Deaves et al., Simulation Based Evaluation of Communications Management Within Battlespace Scenarios, *Signal Processing, Sensor Fusion and Target Recognition VII*, Ivan Kadar, Editor, SPIE Proceedings Vol. 3374, 13-15 April, 1998, Orlando, FL.

[10] Ivan Kadar, Optimum Geometry Selection For Sensor Fusion, *Signal Processing, Sensor Fusion and Target Recognition VII*, Ivan Kadar, Editor, Proceedings of SPIE Vol. 3374, April 1998.

[11] Lucy Y. Pao and M. Kalandros , The Effects of Sensor Requests on Sensor Manager Systems, *Proceedings of the American Control Conference*, Albuquerque, NM June 1997.

[12] M. Kalandros and L. Y. Pao, Controlling Target Estimate Covariance in Centralized Multisensor Systems, *Proceedings of the American Control Conference*, Philadelphia, PA, June 1998.

[13] Ivan Kadar, Optmum Connection Management Scheduling, *Signal Processing, Sensor Fusion and Target Recognition IX,* Ivan Kadar, Editor, Proceeding SPIE Vol. 4052, April 2000.

[14] Xuezhi Wang, Rob Evans, Jonathan Legg, Distributed Sensor Fusion with Network Constraints, *Signal Processing, Sensor Fusion and Target Recognition IX,* Ivan Kadar, Editor, Proceeding SPIE Vol. 5429, April 2004.

[15] Samuel Kullback, Information Theory and Statistics, Dover Publications, New York, 1968.

[16] Thomas Cover and J. A. Thomas, *Elements of Information Theory,*Wiley Interscience, New York, NY 1991.

[17] Jack C. Kieffer, A Survey of the Theory of Source Coding with a Fidelity Criterion, *IEEE Trans. on Information Theory, Vol. 39, No. 5, Sept. 1993*.

[18] Ivan Kadar and R. E. Eaves, *Satellite Business Networks and Systems*, AIAA Selected Reprint Series, Vol. XX, American Institute of Aeronautics and Astronautics, New York 1985.

[19] Gregory J. Ponte, Fundemental Limits in Cooperative Sensor Networks, *Final Report 2000-2001, Project 00-0080,* Department of Electrical Engineering, Unversity of California, Los Angeles.

Reusability, Scalability, and Quality in Higher-Level Fusion Processing

James Llinas
Center for Multisource Information Fusion
University at Buffalo
Buffalo, NY 14260
USA
llinas@acsu.buffalo.edu

Abstract - *These panel remarks will center on two categories of issues: reuse and scalability, and three different but related notions of quality in fusion processing.*

1 Reuse and Scalability

Previous remarks by the author at other fusion conferences have asserted that one measure of the maturity of a field of study is an ability to (1) assess the nature of a problem in that field, and (2) given the problem characterization, the ability to nominate a class of solutions for it (i.e. specific algorithms in this case); that is, an ability to "map from problem space to solution space". Calibrating problems could be achieved via a type of "encoding", in which the parameters that influence the nature of the problem in various dimensions would be defined, along with their values. Not all of these parameters may necessarily have to do with the bounding of viable solutions (algorithms); it could be that problem characteristics that influence non-algorithmic aspects would be of interest, e.g. computational complexity. If this "mapping" knowledge is available, it is argued that it aids in more efficient solution generation, in objective comparisons across similar applications, and importantly in these days when a transitionability requirement is attached to every R&D effort, in gauging the extensibility of a research prototype to an operational application.

2 Quality

2.1 Quality 1: Formalized Performance

Most fusion researchers would agree that the root function of a fusion process is Data Association. Most would also agree that today's problems are of moderate to high dimensionality (e.g. many targets, many platforms, etc). Most would also agree that the inputs to a fusion process are random variables and that a fusion process is a measurable functional process, and so, no matter the nature of the fusion process, its output is also a random variable. If these three assertions are combined, it can be appreciated that methods for fusion process performance evaluation should consider the problem of "estimate-to-truth" association, the problem that arises when (a) the data association of the "system-under-test", ie the fusion process being examined, is imperfect (which is typical), and (b) when there are many objects involved (the high-dimensionality issue), so that, when desiring to compare fusion estimates to truth (as in a simulation), which estimate goes with which truth is unclear. Further, the very nature of an "estimate-to-truth" association for higher-level estimates has not been addressed even conceptually.

The performance evaluation methodology should ideally also consider inherent variances involved in these evaluations, and the possible use of experimental design techniques and analysis of variance methods as used in other communities when faced with these complexities. To date, the fusion community does not seem to have brought these well-known methods into its evaluation approaches; while their use adds some complexity and cost, under the right conditions, especially for basic research projects, the community should be thinking about improved formalization in performance evaluation. This issue has often undermined the viability of the transition of a promising prototype to an operational setting, as the user questions what the "Data Fusion Warranty" is.

2.2 Quality 2: Formal Mathematical/ Algorithmic Integrity

It is clear that all fusion researchers would desire that the mathematical or, more generally, the algorithmic integrity of their designed processes be assured. By integrity we mean formal correctness, e.g. consistency with assumptions, and with structure. In well-bounded applications this should (hopefully) not be difficult to achieve. But in not so well-bounded and especially in the distributed fusion cases, assuring mathematical/ algorithmic integrity is something to worry about. This issue extends further in complexity when multi-system interoperability is sought, i.e. in any environment that has not been "holistically" designed, i.e. designed with the other system in mind. Consider even the Level 1 problem, where the community has been able to nominate some viable solutions such as tracklets and covariance intersection etc. which were framed with these concerns in mind. But what about the problem of distributed object identification, what about distributed Level 2 and Level 3 processing and the assurance of integrity? These concerns also bear on the question of what is meant by inter-process conflict resolution (the term "adjudication" seems to have become standard for this).

2.3 Quality 3: Control of Local or Nodal Fusion Process Output Products at the “System” Level

Whether in a centralized framework or a distributed framework, there is a question that arises as to when to “post” the fused-product output to the “system”, being for example either the user or possibly a network. The question relates to both notions of (a) the temporally-and-completeness-based construction of a Level 2 or 3 product, which inherently consumes time and is inherently incomplete at the earliest time, and (b) control of both local demands for quality (within the local process) and quality for release of the product for “consumption” either by a user or a subsequent fusion node. Related to (a) is the question of the similarity of L2, L3 processing to traditional signal processing—e.g. one could ask: “when is a situation detected?”, or “when is a situation unfolding?”, or “when does the situation exist?”, in spite of the elements of the situation being (always) incomplete and uncertain. Related to (b) are such issues as Belief Revision or belief updating, and the issue of internal consistency and thus the meaning of consistency in L2, L3 processing.