FOI is an assignment-based authority under the Ministry of Defence. The core activities are research, method and technology development, as well as studies for the use of defence and security. The organization employs around 1350 people of whom around 950 are researchers. This makes FOI the largest research institute in Sweden. FOI provides its customers with leading expertise in a large number of fields such as security-policy studies and analyses in defence and security, assessment of different types of threats, systems for control and management of crises, protection against and management of hazardous substances, IT-security an the potential of new sensors.

FOI
Defence Research Agency
Command and Control Systems
P.O. Box 1165
SE-581 11 Linköping

Phone: +46 13 37 80 00
Fax: +46 13 37 81 00

www.foi.se

A survey of some Information Fusion research fields applicable in Operations Other Than War

PONTUS HÖRLING (ED), CHRISTIAN MÅRTENSON,
JOHAN SCHUBERT, ROBERT SUZIC, PONTUS SVENSON

FOI-R-- 2000 --SE
ISSN 1650-1942

User report
June 2006

Command and Control Systems

Pontus Hörling (Ed), Christian Mårtenson, Johan Schubert, Robert Suzic, Pontus Svenson

A survey of some Information Fusion research fields applicable in Operations Other Than War

FOI-R--2000--SE
ISSN 1650-1942

User report
June 2006

Command and Control Systems

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Command and Control Systems P.O. Box 1165 SE-581 11 Linköping	FOI-R--2000--SE	User report
	Research area code 4. C4ISTAR	
	Month year June 2006	**Project no.** E7097
	Sub area code 41 C4I	
	Sub area code 2	
Author/s (editor/s) Pontus Hörling (Ed) Christian Mårtenson Johan Schubert Robert Suzic Pontus Svenson	**Project manager** Pontus Hörling **Approved by** Martin Rantzer **Sponsoring agency** FM **Scientifically and technically responsible** Pontus Hörling	

Report title

A survey of some Information Fusion research fields applicable in Operations Other Than War

Abstract

During a decade, information fusion within the military domain has focussed on putting together a situation picture from the "battle field" where an enemy, known regarding organisation and doctrine, is moving ahead. Clustering and association of observation reports, target tracking and management of sensors and other information collecting resources were prioritised.

Today's situation where Swedish military personnel are expected to act abroad in peace supporting and peace enforcing missions (OOTW – Operations Other Than War) means that the information that is gathered often have a more police than military type of character. It is important too keep contact with the local police and inhabitants, as well as being able to handle riots if needed. Intelligence gathered by humans (Humint) will often be more important than intelligence gathered with sensors. An important intelligence source will be the observations from the soldier's daily patrolling. The adversary is more diffuse regarding organisation, resources and intentions. The information that is available when the mission is commenced, as well as that collected during an ongoing mission must be structured and made searchable in an efficient way.

This report briefly describes the current state-of-the-art within selected areas of information fusion related research that has a potential to support OOTW. Various approaches to text mining, information structuring using wiki technology, uncertainty handling in ontology design, new aspects on user-centric situation awareness, riot simulation and riot control are described.

Keywords

Information fusion, OOTW, Text mining, Wiki, Uncertain ontologies, Situation assessment, Decision support, Crowd control

Further bibliographic information	Language English
See also the group website **www.foi.se/infofusion**	
ISSN 1650-1942	**Pages** 28 p.
	Price acc. to pricelist

<table>
<tr><td rowspan="5">Utgivare
FOI - Totalförsvarets forskningsinstitut
Ledningssystem
Box 1165
581 11 Linköping</td><td>Rapportnummer, ISRN
FOI-R--2000--SE</td><td>Klassificering
Användarrapport</td></tr>
<tr><td colspan="2">Forskningsområde
4. Ledning, informationsteknik och sensorer</td></tr>
<tr><td>Månad, år
Juni 2006</td><td>Projektnummer
E7097</td></tr>
<tr><td colspan="2">Delområde
41 Ledning med samband och telekom och IT-system</td></tr>
<tr><td colspan="2">Delområde 2</td></tr>
<tr><td rowspan="4">Författare/redaktör
Pontus Hörling (red)
Christian Mårtenson
Johan Schubert
Robert Suzic
Pontus Svenson</td><td colspan="2">Projektledare
Pontus Hörling</td></tr>
<tr><td colspan="2">Godkänd av
Martin Rantzer</td></tr>
<tr><td colspan="2">Uppdragsgivare/kundbeteckning
FM</td></tr>
<tr><td colspan="2">Tekniskt och/eller vetenskapligt ansvarig
Pontus Hörling</td></tr>
<tr><td colspan="3">Rapportens titel
En översikt av några områden inom informationsfusionsforskningen tillämpbara i operationer andra än krig</td></tr>
<tr><td colspan="3">Sammanfattning
Informationsfusion inom klassisk krigföring har under ett årtionde fokuserat på sammanställning av lägesinformation från ”slagfältet” där en organisatoriskt och doktrinärt känd motståndare rycker fram. Klustring och association av observationsrapporter, målföljning, samt styrning av sensor- och andra informationsinsamlingsresurser prioriterades.

Dagens situation där svensk militär personal framför allt antas agera i fredsuppehållande och fredsframtvingande operationer (OOTW - Operations Other Than War) utomlands innebär att den information som samlas in om motståndaren är av ofta mer polisiär än militär karaktär. Att upprätthålla kontakter med lokal polis och civilbefolkning samt att hantera upploppssituationer är viktigt. Mänskligt insamlad underrättelseinformation (Humint) får ofta större tyngd än sensorinformation. Soldaters dagliga patrullering och informationsinsamling blir en viktig källa. Motståndaren är mer diffus: Vilken organisation har de, vilka resurser kan de nyttja, vilka allierade har de, vilka intentioner har de? Den information som finns tillgänglig vid uppdragets början samt den som samlas in under pågående uppdrag måste på ett systematiskt och effektivt sätt kunna struktureras och göras sökbar.

Rapporten beskriver kortfattat det rådande läget inom utvalda områden av informationsfusionsforskningen som har potential att kunna stödja OOTW. Det rör olika varianter av textutvinning, strukturering av information med wikiteknologi, osäkra ontologier, nya aspekter på användarcentrerad lägesförståelse samt upploppssimulering och upploppshantering.</td></tr>
<tr><td colspan="3">Nyckelord
Informationsfusion, OOTW, textutvinning, Wiki, Osäkerhet i ontologier, Situationsbedömning, Beslutsstöd, Upploppshantering</td></tr>
<tr><td>Övriga bibliografiska uppgifter
Se även gruppens webplats www.foi.se/infofusion</td><td colspan="2">Språk Engelska</td></tr>
<tr><td>ISSN 1650-1942</td><td colspan="2">Antal sidor: 28 s.</td></tr>
<tr><td>Distribution enligt missiv</td><td colspan="2">Pris: Enligt prislista</td></tr>
</table>

Contents

1 Introduction 6
2 Text mining and structuring 7
2.1 Keyword search 7
2.2 Semantic search 7
2.3 Statistics-based search 8
2.4 Some examples of research results and products 9
2.5 Text structuring: Wiki 10
2.6 Handling uncertainty in ontologies 12
3 State of the art in user-centric situational awareness 13
3.1 Mixed-initiative approaches 15
3.2 Bridging fusion, long term prediction and dynamic sensor management ... 17
3.3 Predictive battlespace awareness 17
4 Crowd simulation and computerized analysis of crowd control 19
4.1 Genetic algorithms for intelligent agent behaviour, optimization and crowd control 19
4.2 Modeling and simulation of riots 20
5 Conclusions 23
6 References 23

1 Introduction

The ongoing transformation of the Swedish defence from an organization that was mainly aimed at coping with an invasion by a mechanized adversary, towards an organization that is supposed to primarily handle international Peace Support and Peace Enforcement Operations requires new types of information fusion systems. Systems designed for gathering and fusing positional and type information of military objects such as armoured battle vehicles moving according to an invasion plan could certainly be usable in the older scenarios, and was also developed at different military research centres, Rasmussen (1998), Ahlberg et al (2004), Ahlberg et al (2006). Similar scenarios could still be observed, such as during the war in Iraq, but are then an issue for larger military powers than Sweden, e.g. the US or NATO. Sweden, today, has a primary focus on coalition operations within NATO/PfP and EU. An EU initiative is that of the Battle Groups, i.e., EU coalition Rapid Reaction Forces that should be deployable within crisis areas far outside Europe. As Framework Nation, Sweden is expected to engage in the 2008 and 2011 Battle Groups organized by the Nordic countries, called the Nordic Battle Groups 08 and 11 (NBG08 and NBG11). The main tasks are expected to be Peace Support and Peace Enforcement Operations.

What information will be important for the success of such missions? Civilians are very involved as actors and/or victims. We are often dealing with ethnical or religious conflicts that sometimes have roots far back in time. People today more often live in cities or villages which require knowledge of conflict handling in urban and suburban areas with its entire infrastructure. Swedish troops may face warlords and irregular armies, sometimes remnants of the military organizations of one or several collapsed nations. Weapons of different lethality are widespread among people. Lawless societies make criminal activities abundant, activities which often finance warlords, clan leaders and other parties. Non-Governmental organizations, such as the Red Cross or Amnesty International, are often present and pursue parallel aid actions or information collection. Furthermore, Swedish troops should cooperate with forces from other nations within coalitions, sometimes also with well-behaved regional military troops and police forces. Extensive Rules of Engagement (RoE) have to be followed and searched for guidance when facing specific situations. The ever greater interest in Effect Based Approaches to Operations (EBAO) opens new information dimensions, where political, economic, social and infrastructural aspects are interwoven with the military goals in order to not only “win the war”, but also “win the peace”. The ability to handle all this new information is crucial for EBAO missions to be successful.

In the project “Techniques, Methods and Demonstration systems for Information fusion” (TMDI) at the Swedish Defence Research Agency (FOI), we are currently studying a few more or less mature areas where output typically ranges from commercial products to preliminary research results, many of them that could be of help in the above mentioned situations. Some of the areas selected might seem a bit unrelated to others, but we regard them all important in typical Operations Other Than War (OOTW). The intention with this User Report is to provide a sampling of state-of-the-art research within these fields. These are mainly related to text mining, structuring disparate information using the Wiki technology, some comments on uncertainty in ontologies, computer tools for user-centric situational awareness, and simulation of crowds and riots in order to analyze methods to control them. Mainly, the subjects are in various ways described by reviewing and commenting upon a set of relevant publications from scientific journals, technical magazines, conferences, and workshops.

2 Text mining and structuring

When preparing a peace-keeping or peace-enforcing mission in a certain area, so-called Intelligence Preparation of the Battlespace (IPB) before the main troop arrive there could render the operation more successful. To understand the cause of the conflict, the different parties in it, the culture, religion etc often means to find and study information sources in text format, such as from the internet, encyclopaedias, newspapers and strategic intelligence reports. This could give an initial “societal” Situation Picture (SP), not obtainable using sensors, which of course will contribute to the SP by giving information on topography, vegetation, infrastructure and other properties detectable by various sensor assets. During the mission, information will be collected by the troops by observations, interviews and interrogations, much of it disseminated as text. By continuously adding this information to the initial SP, the SP could be kept up-to-date. Of course it is necessary to have an effective information collection strategy, but it must also be possible to effectively search, collate and associate the growing amount of information collected in order not to “get lost” among all documentation. This means that effective tools for text mining must be present, as well as the ability to store the documents generated during the mission, and ability to access the databases and networks where other relevant information is available.

2.1 *Keyword search*

The simplest examples of using software to retrieve unstructured information are free text search tools, i.e., keyword search. The simplest example of such a tool is the built-in search facility in most modern operating systems, such as Windows or Mac OS X. Other examples of such tools are most Web search engines, such as Google or Yahoo, where different Boolean keyword combinations are used as search criteria. A comprehensive list of search tools for web sites and intranets can be found at http://www.searchtools.com/index.html.

2.2 *Semantic search*

Another approach to searching documents is to *tag* all documents in the repository with relevant tags which give a semantic meaning to the document, certain paragraphs, or sentences in it. Searches can then be made based on these tags, and the user can be presented with, for instance, all documents that contain all of a given set of tags. Most such techniques rely on taking tags from an ontology that specifies “all that is known” about the subject. A problem with this approach is that all users must know and agree to use the ontology. It can also be difficult to adapt the ontology to new information. There are several approaches that can be made so solve these problems. Different ontologies might describe the same or more or less overlapping areas, and advantages can therefore be won by trying to integrate overlapping parts of ontologies to one single ontology. A review of methods for ontology integration can be found in Eklöf and Mårtenson (2006).

Yet another approach is to allow users to use any word or phrase when they tag documents; tags are then searched using an ordinary, text-based search engine. This is the approach chosen for several web-sites that rely on collaborative filtering to add meta-data to its contents, such as http://www.flickr.com or http://www.deli.cio.us.

2.3 Statistics-based search

This is a way to analyze the text in documents based on statistical ideas. Tools based on this principle can for instance cluster or group documents that have many occurrences of a certain word or certain combination of words, such as two or more words often occurring close to each other in the text (directly connected, in the same sentence, paragraph or document). Documents that are similar in these respects can be grouped together. Documents where certain groups of words often occur might have a higher probability to contain other interesting groups of words etc. In principle, there might be many different types of patterns that can be searched for, and the limits are typically set by the implementations of the tools themselves.

2.3.1 Example: Autonomy

Autonomy (see http://www.autonomy.com) is an example of a software application that does more processing than just simple text search. The software provides an integrated framework that allows users to access information through advanced searching capabilities. Instead of relying on meta-data that is stored in each document, Autonomy uses the full text of the stored documents in searching. It can also search simultaneously on the web, and present results from both the internal database and the web to the user.

An important part of Autonomy is the reliance on *concepts* instead of on specific words in searches. This means that a user does not need to list all possible synonyms of the word they are looking for. It also means that documents can be retrieved even if they do not contain the word the user entered, provided that they do contain a word that represents the concept for which the user searched.

This use of concepts is one of the main marketing points of Autonomy. In their publicity material, they do not reveal how they implement this functionality. It is very likely that the technique used to implement this is based on so-called Latent Semantic Indexing (LSI), see for instance Weiss et al (2004), which is commonly used in text analysis and mining software. LSI relies on having a large repository of documents which are used to find sets of words that are in some sense similar. The similar words are found by applying a transformation to the representations of the documents, making an approximation in the transformation, and then transforming back into the original representation. The similar words can be taken to represent a concept. It is important to note that the concepts found by this method do not need to represent a concept that makes any sense to humans. This is because the approximation that was performed in the LSI is a purely mathematical operation, which does not use any knowledge about the subject area of the documents.

Users of Autonomy can also benefit from seeing other peoples searches, and can get information on what documents other users who performed similar searches chose to view. Automatic agents that perform searches at regular times can be set up, so that users can be alerted to the presence of new material that is relevant to them. Documents in the repository can be clustered, *i.e.,* those documents that are similar to each other (as determined by the mathematical representation of the document) are shown close to each other in a cluster map. Summaries for each cluster can be displayed, allowing the user to quickly get an overview of the information content of the repository. The clustering is performed using standard methods from text mining.

Autonomy provides functionality that allows it to automatically transcribe TV or video. The transcripts can then be used to represent the videos in the repository, so that a search can return relevant media clips. An important use of this for command and control could be to directly import relevant information from, e.g., CNN into MilWikiKB (for MilWikiKB, see section 2.5). This is something that could conceivably be tested in the “Demo 06 Höst”-demonstration. Relevant video to import could be taken from the media reports that were produced as part of the PSYOPS training during the Viking 05 exercise.

Autonomy also has a solution for handling access rights. Before a user is granted access to a document, a check is performed to see whether the user should be allowed to view it. It is possible to configure the system both so that the user sees that the document exists but cannot access it, or so that the user is not made aware of the document’s existence. For intelligence and military applications, this is an extremely important feature.

There is quite a lot of hype surrounding the Autonomy suite of software. Much of the company’s marketing hinges on the alleged advanced text analysis features of the software, in particular, its use of concept-based search and using information theory to determine the most important terms/concepts in a document. The methods used seem, however, to be rather standard applications of text mining. What is quite amazing about Autonomy is instead the framework that the company has built around the text mining methods. The design of the software is completely modular, and it is possible to write both import modules that allow it to read new kinds of documents as well as tools that use its capabilities, thus providing new kinds of analysis.

2.4 Some examples of research results and products

A short and highly readable note by Grimes (2003) provides a brief overview of text mining. It discusses the problem of searching based on manual classification using tags from an ontology: there is too much data to classify, and it is too complicated to learn a general ontology. Instead, by using text mining to cluster similar documents and display summaries of them, it is possible to navigate quickly through large repositories of documents. The paper discusses the differences in approach between, for instance, Autonomy, which uses a completely statistical approach to the data, and other software that combines the statistical approach of ”deriving” taxonomies with using high-level, manually produced ontologies.

It is worth pointing out that much of the criticisms against using manual classification and ontologies are more applicable to general systems that handle information from the web rather than the type of specialized information that is most interesting for the military. In specific, small domains where it is possible for humans to learn an ontology, and where the number of documents is either sufficiently small so that they can be manually classified or sufficiently simple so that automatic classification works, it will always be better to use semantic search based on an ontology rather than ”brute force” text mining.

A paper by Noble (2004) discusses some problems inherent to mining the internet for open source intelligence information. Such mining can be an important complement to other sources of information. Of particular interest is the large number of newspapers that also provide web access to their articles. The paper notes that most assessments of the credibility of Humint[1] relies completely on the judgment of a human analyst, and briefly discusses possible ways of automating parts of this process. The author suggests using three principles when assessing the reliability of a source: the historical reliability of the source, the consistency of the current report with known facts, and the consistency between this report and other reports. The paper also discusses how Commercial Off-The-Shelf (COTS) tools can be used to gather information from the internet. The paper is interesting, but does not go very far in providing information on how to actually implement automatic web mining systems.

A more recent paper by Noble (2005), which can be seen as a continuation of the preceding one, discusses in some more detail how to obtain information from the web and what it might look like. It describes a framework based on COTS software that can be used to extract information from the web. The process is semi-automated, and ”mixed-initiative” in that the computer is used to harvest data from the web and perform simple linguistic analysis, while the human operator has the main responsibility for assessing reliabilities and fusing information from different sources. The paper provides enough information that it would, in principle and provided that similar COTS tools are available, be possible to reproduce the analysis.

There are several toolboxes and software environments available on the market or as open-source products that address this field. Unstructured Information Management Architecture (UIMA), IBM (2006), is an open-source framework for handling unstructured information. UIMA is a product of IBM research that has been made available as open-source software. The framework can be compared to the (proprietary) IDOL framework of Autonomy and allows users to combine modules written for different purposes to analyze unstructured information. The framework relies on having a common representation of the data. By writing appropriate plug-ins, it is possible to analyze different kinds of data, such as text, voice, video, or images. The framework also provides the means to perform semantic searches on the data. It has so far been used for applications ranging from analyzing medical texts to a system for answering natural language questions, PIQUANT.

2.5 Text structuring: Wiki

The TMDI project decided to study the Wiki technology as one promising idea to collectively structure disparate information as web documents and resources. It was tested as one of the information structuring components (called MilWikiKB, Military Wiki Knowledge Base) at the “Demo 06 Vår” demonstration in Enköping in March 2006, Mårtenson and Brännström (2006). Its main use is supposed to be the entering of open background information about a crisis before arriving there, and continuous updating of information during the operation. An intranet between different staffs, intelligence centres and information processing nodes is assumed to be present[2]. Text mining and link analysis tools could further be added to facilitate the structuring work, even though that has not yet been the case with the MilWikiKB.

[1] Humint – Intelligence originating in information collected and provided by humans.

[2] Because of the aim of wikis is the collaborative entering, structuring and agreement upon information, this technology might be of less use in a stand-alone system for handling secret information not allowed to be available on an intranet.

The first wiki system, WikiWikiWeb, was created by Ward Cunningham in 1995 (http://c2.com/cgi/wiki). Today many implementations exist, varying from simple hacks to full featured content management systems.

In a recent study by Giles (2006), a selection of Wikipedia articles was compared to their counterparts in the highly respected Encyclopaedia Britannica. The average number of errors in the Wikipedia articles was found to be four, to be compared with three in Britannica (however, Britannica has disputed upon the way the investigation was conducted and interpreted). This shows that the constant review by thousands of users can be compared to that of a few paid experts. However, even if the Wikipedia comes close to a traditional encyclopaedia regarding some quality aspects, its major benefit lies in the speed with which articles are updated and new knowledge is included.

2.5.1 Semantic wikis

The increasing popularity of the wiki concept proves its usefulness as a tool for collaboratively creating web documents. However, since the product of wiki collaboration is text in natural language, it is only suitable for other people to read; support for machine interpretation is poor. To deal with this problem, a number of ongoing research and development projects aim at introducing explicit semantics in wikis. This is in line with current efforts in developing the Semantic Web (see, for instance, http://www.w3.org/2001/sw), an extension of the World Wide Web where the content of web-pages are given meaning through the use of standards, mark-up languages and related processing tools.

The first step towards creating a semantic wiki is to give meaning to the links between pages. Consider a page describing the life of Shakespeare. The page might contain a link to a page about Stratford-upon-Avon, Shakespeare's place of birth, and to a page describing the play Hamlet. In a normal wiki there is no difference in interpretation between the two links, they simply connect the different pages. In a semantic wiki the links are labelled, often using Resource Description Framework (RDF), a language for expressing triplet relations in the form of subject, predicate and object. The subjects and objects in a wiki are the interlinked pages and the predicates are expressions describing the kind of relations. In the above example, the subject would be the Shakespeare page, the object either the page of Stratford-upon-Avon or Hamlet, and the predicates are links of the type bornIn or authorOf respectively.

This kind of tagging enables a computer to process the added link information and lays the foundation for more advanced querying of the wiki content. For example, it will be possible to ask the wiki for a list of all persons (included in the wiki) born in Stratford-upon-Avon, or when reading the Hamlet page, to ask for other works by Shakespeare. Pages containing lists with dynamic content can be updated automatically instead of manually as is done today in, for instance, Wikipedia.

In the context of a command and control knowledge base, the semantic features can enhance situational awareness. Advanced queries will be based on semantic concepts rather than keywords, improving the precision of the information access. For example, in a military scenario, it will be possible to ask for all friendly units participating in a certain mission in a certain region. In an intelligence scenario, one can perform queries like "give me all pages related to a certain person or his/her family".

One of the major points of criticism of the Semantic Web is the question of how to make people and companies tag their web pages. Even if the mark-up process can be made relatively simple, the amount of time to perform the tagging will be considerable. Without any incentives, there is a risk that the glories of the Semantic Web will remain a dream. However, the designers of the Semantic Wikipedia believe that the large and devoted community of Wikipedia authors will manage the task of tagging the Wikipedia, see Völkel et al (2006). Similar arguments can be found in Wagner et al (2006), which discusses the possibility of semantic management of very large governmental web sites using wiki technology. People in certain positions and with certain skills are organized into communities of interests, each of which will be responsible for continuously tagging a designated part of the site. Many authors sharing the burden of administrating the information will be the mechanism enabling large-scale semantic knowledge management.

2.5.2 Different wiki-implementations

There are many different implementations of the wiki concept. The most widely used is MediaWiki, which was also the choice for implementing MilWikiKB. This implementation is used in http://www.wikipedia.org, and is described in detail at http://www.mediawiki.org/wiki/MediaWiki.

An interesting and very light-weight implementation is TiddlyWiki (http://www.tiddlywiki.com/). This implementation actually does not need a web server to run on, since it is implemented completely in JavaScript.

MoinMoin (http://moinmoin.wikiwikiweb.de/) is an experimental Wiki implementation that is particularly noteworthy because of its extensibility. It is possible to change several of its functionalities without changing the source code. It also has some support for WYSIWYG-editing, and provides a simple mark-up-language for editing pages in the wiki.

There are many more interesting wiki implementations. A listing of many of them can be found at http://www.wikimatrix.org/. This site also makes it possible to compare several different implementations and to search the implementations based on what features one is interested in.

2.6 *Handling uncertainty in ontologies*

Uncertainties in ontologies could have a large impact on tools that rely on them. Therefore, it is important to be able to cope with such uncertainties. Uncertainty here could for instance mean that there are uncertain relations between concepts (is this really a meaningful relation between these two classes), uncertainty of which class is a subclass of which, if a class should be a separate class or (perhaps) be merged with another one and so on. These are uncertainties that are inherent to the ontology. When using the ontology in practice, we normally populate it with instances of the concepts/classes. Uncertainties about to which class a certain instance belongs is another kind of uncertainty that has to be handled. One example of a practical application, relevant to the TMDI project, where such uncertainties could be important is social network analysis, Svenson (a) (2006).

This is an area where not much has been published. The subject has generated a lot of discussion at several conferences and workshops, but these have not been documented in writing. Much of the ideas are about applying Bayesian networks to handle the uncertainties.

An interesting two-page summary of the area can be found in an unpublished manuscript by Kokar and Matheus. It begins by discussing what probability should mean in an ontology. Should it be possible to add uncertainty to the presence of classes in the ontology or just to relations between classes? In the latter case, should a probabilistic weight on the relation be interpreted to imply that the subclass "sometimes" (with a certain probability) implies the superclass, or the other way around?

There have been some attempts to use Bayesian networks to represent uncertainty in ontologies. A criticism of this that is raised in the manuscript is that while Bayesian networks are used to encode causal relations, an ontology primarily encodes taxonomical relations. In particular, look at the conditional probability P(A|B). In a Bayesian network, this is interpreted as

> "there exists a causal relation between *A* and *B* meaning that *B* influences *A* in that whenever the value of *B* changes, so does the value of *A*."

while an ontological relation that B is a subclass of A is interpreted as

> "any element *b* of *B* is also an element of *A*.

It is not clear that attaching probabilities to these two statements mean the same thing. The root cause of the difficulties is that causality is not equivalent to logical implication. Consider the logical statement "A implies B". This statement can be true even if there is no causal relation whatsoever between A and B: it might just happen that they are both true. An example used by Kokar and Matheus will clarify: they consider the statement "If George Bush was president of the U.S. in 2004, then Germany is in Europe." This statement is true, but there is no causal relation in it.

A paper by Cesar et al (2005) presents an extension of the OWL language to handle uncertainty in ontologies. The extension uses fragmentary multi-entity Bayesian networks to handle the uncertainty. Such a fragment would replace an ordinary class in the ontology, and defines both a random variable and its probability distribution. The solution is interesting, but seems not to be general enough to sufficiently represent uncertainty in relations.

A short working note by Noy et al (2006) from the W3C contains a suggestion for how to handle n-ary relations (i.e., relations between n different entities) in an ontology. The proposed way is to introduce a new class for each such relation, where the new class has links to each of the n classes that should originally interact. For example, a relation such as "Christine has breast cancer with high probability" would be implemented by adding a new class "Diagnosis relation". The object "Christine" would be linked to this class, which would contain links to objects "breast cancer" and "high". The emphasis in the note is on other important uses of n-ary relations. It is not clear whether a simple solution such as the one proposed here is enough to handle all interesting cases of uncertainty in ontologies. The proposed solution would require the addition of a large number of extra classes to represent the relations with might be associated with uncertainty.

3 State of the art in user-centric situational awareness

In a sense, almost all research about information fusion deals with user-centric situational awareness, since the goal is most often to provide a user with enough information so that they can achieve their objectives. Information fusion is commonly partitioned into level 2 (situation assessment), level 3 (impact assessment) and 4 (adaptation)[3]. The work performed within the "User-centric situation awareness" part of the TMDI project this year relates to all three levels, Svenson and Mårtenson (2006), Svenson (b) (2006). In this section, we report on a number of recently published papers that are relevant to our work. Quite recently, suggestions have been raised to add a fifth level to the JDL model. This level is often called "user refinement" and can be seen as adding a human in the loop to the JDL model. Some of the ideas suggested in the papers mentioned below can be seen as implementations of this fifth level.

Salerno et al (a) (2005) argues for the necessity of having automated situation awareness in order to achieve information superiority on the battlefield. They define situation awareness as knowing what the opponent is doing and why they are doing it. This is consistent with including level 3 fusion as a special-case of level 2 fusion. The paper presents two different applications, one based on the build-up to the Kuwait war in 1991, and the other a network intrusion example. In both cases, the authors break down the situation into events that the system should look for. For example, an observation of troops moving closer to the border increases the level of aggression that the situational awareness systems believes that Iraq has against Kuwait.

An earlier paper, Salerno et al (2004), by the same authors describes the development of the tool in the previous paper and gives some more details about the Kuwait application. In comparison to the previous paper, this paper discusses the differences and similarities between the JDL model and Endsley's model for situation awareness (see, e.g., Endsley et al (2003)) in more detail. It tries to overlap the differences between, on one hand the information fusion community, and on the other hand the cognitive community, which starts with the human's need instead of bottom-up from the sensors. The ability to make forecasts regarding future events on the battlefield (i.e., to have predictive situation awareness) is a sign of the highest level of situation awareness. This is the highest level of situation awareness in Endsley's original model. A later extension of this also includes a level called "resolution". Users at this level of situation awareness are not only aware of the enemy's plans and intentions, but also have a clear idea of what they should do in order to achieve the best possible outcome. The framework described in the paper is designed in order to achieve this level of situation awareness. In the framework, analysts determine the question of interest. Information and data collectors then collect and process data that is relevant to the query. Data, however, is not enough: users cannot look at all available data and reach conclusions. For this reason, the framework also includes knowledge discovery tools that mine the data for interesting information. In fusion terminology, this corresponds to finding relations between the objects that were found in the level 1 processing.

In Salerno et al (b) (2005) possible ways of evaluating the results of the tool described in the previous papers are discussed. The paper extends evaluation metrics used for level 1 fusion to higher level fusion. Level 2 fusion is distinguished from level 1 by the introduction of "artificial" objects, that represent several level 1 objects that were determined by the system to belong together. In order to quantify the degree to which a level 2 fusion system succeeds in this, the authors propose to use the "data to

[3] Level 0 (signal processing) and 1 (object refinement; normally some kind of object detection, discernment, tracking and/or identification) are often referred to as "sensor fusion" in contrast to "information fusion" for the higher fusion levels.

information ratio", which is the ratio between the number of observations and the number of complex entities in the situation picture. This metric tells us how much the cognitive load of a human analyst would be reduced by the fusion system.
The other metrics used are more or less standard extensions of previously known metrics. For instance, recall (the number of correct situations divided by the number of situations in ground truth) and precision (number of correctly detected entities divided by the number of detected objects) are both standard metrics from machine learning and data mining. Perhaps the most interesting discussion in the paper regards the need for taking account of the costs of different errors in the fusion algorithms. All parts of the situation picture are not always equally important. Consider two competing fusion systems, one that misses a very important enemy convoy that it approaching a target, and the other that misses a neutral convoy that is receding from the area of operations. A straight-forward application of the metrics discussed above would punish both fusion systems equally, whereas it might be more useful to differentiate between them based on the relative significance of the errors they made. Since different parts of the situation picture are important to different users, it might be best to use different fusion systems (with different strengths and weaknesses) for different users, depending on their need.

A paper by Matheus (2005) introduces the SAWA architecture for situation awareness. It discusses the need to keep track of a large number of relations among entities in order to provide a good enough situation picture to users. Perhaps the simplest example of such a relation is given by the force aggregation problem, where we determine which objects belong together in platoons, and recursively which platoons belong together in companies. There are however many other types of relations that are important for situation awareness. For example, the relation "threatens" between enemy objects and friendly or neutral objects is important. This paper describes the structure of SAWA and applies it to a logistics problem. The paper describes the ontology developed and used for this domain. Neither the ontology nor SAWA currently handle uncertainties. The logistics planning process described in the paper does not use mixed-initiative, it would however be possible to expand SAWA to use that. Mixed-initiative approaches to the field are described in the next section.

3.1 *Mixed-initiative approaches*[4]

An article by Allen (1999), which is part of a special section devoted entirely to mixed-initiative interactions, gives a readable overview of the subject. Mixed-initiative interaction is a recently introduced paradigm for systems where several agents interact in a flexible manner, i.e., where there is no fixed structure that determines whose turn it is to "control" the reasoning process. Most often, such systems consist of two agents: a human operator and a computer system that assists the human in some task. In principle, the same methods could be used also for systems consisting of many computer tools that interact with each other and/or with humans.

In mixed-initiative interaction, each agent does what it does best. For instance, a human could supply intuition and experience to a planning process, while the computer uses its computing power to calculate, e.g., the time needed to complete a

[4] For more information regarding this and the state of the art in mixed-initiative interaction, we refer the interested reader to a forthcoming FOI-report devoted exclusively to the subject.

certain combination of tasks. In some parts of the planning process, it will be the human who has the initiative and control, while in others the computer will take over and show results to the user. This is in contrast to traditional computer tools for planning, which most often rely on the user to specify the problem and then presents a complete solution to the user. By combining the strengths of the user and the computer and allowing them to reason in dialogue with each other, better results are achieved. Many mixed-initiative systems have been implemented for planning and logistical processes. One example is the following paper by Smith et al.

Smith et al (2005) presents a system for resource management that is based on mixed-initiative interaction. The system as presented in the paper allows users to specify tasks that should be performed. The application for which the system is intended is planning for special operations forces, although the principles and methods introduced by the authors could be easily extended to other application areas. The system uses an ontology that describes the resources needed to carry out particular missions, and allows users to specify constraints on the order in which different tasks should be performed. Constraints regarding where a particular resource should be used in order to achieve the desired effect are also included in the system. A standard algorithm is used to solve the constraint-satisfaction problem that ensues. The user can control the process both by specifying missions that should be performed and by editing the capabilities and resources that are present. What-if type analyses are also possible to perform, by changing the problem characteristics. Similar systems could be developed for the Swedish Armed Force’s needs in, for instance, international operations.

Planning of sensor movement and scheduling is a field where mixed-initiative approaches have been tested. Mullen et al (2006) discusses the need to control sensors based on the needs of the users. It discusses an architecture for information fusion which uses a market-oriented approach to determining which requests for information should be fulfilled first. The authors argue that in future service-based command and control systems, traditional optimization algorithms are of less use than the agent and market-based approach presented. The paper presents an architecture for implementing a market-based sensor control system on a single platform. The architecture uses a sensor manager that receives requests for information/sensor coverage from “consumers” (humans or intelligent agents). The requests also include a “bid”, stating how much the agent would be willing to pay the broker for the information. The sensor manager matches the requests to the available sensor resources and performs an auction to determine which requests get satisfied.
The paper concludes with the results from several experiments, which (unsurprisingly) show that it is important to include both the gain in information and the utility which the information gives to the mission are important.

An issue of Internet Computing contains a special section on sensor networks. While most of the articles in the issue do not use a “user-centric” perspective, the short essay by Miller (2006) discusses the important question of how the results from the sensor networks should be presented to the user. The article enumerates several important things to think about when designing applications, and emphasizes that any successful use of a sensor network will require that the analysis and display programs are constructed based on what the users of that particular application want to obtain.

3.2 *Bridging fusion, long term prediction and dynamic sensor management*

Kadar et al (2004) describes a system that predicts a future enemy path and uses this to determine figures-of merit for dynamic sensor planning. The paper assumes that modules for long-term prediction of enemy movements are available, and uses them to simulate the results of different own sensor allocation schemes. This could be used as a help to a user that must plan resource allocation. A major drawback of the approach as presented in the paper is that no consideration is given to the inherent uncertainty of predicting the future. The system only considers the most likely future location of the enemy, instead of taking all the different possibilities into account using some sort of averaging scheme (as in Mårtensson and Svenson (2005)). For the application and time-scales considered in the paper, this might not be a severe limitation. However, if one tries to extend the ideas beyond the simple application of tracking vehicles on a road network on the time-scale of minutes, it becomes necessary to include uncertainty handling into the system.

3.3 *Predictive battlespace awareness*

C2 systems must offer well designed tools for making predictions of the situation picture. Phister et al (2003) discusses the need for Predictive Battlespace Awareness (PBA). Commanders are required to look into the future and plan for what might happen. High-level commanders need to see longer into the future than low-level commanders need to do, which means that they require different sorts of tools to help them. The paper talks about the three different types of command:

- Command by direction is when a commander has access to the entire battlefield and can make personal decisions regarding everything.
- Command by planning, which is perhaps the dominant method today, relies on staff's preparing plans that cover all eventualities and on the troops to carry out the plan without delays and difficulties. Since plans can go awry not only because of enemy actions (which with some degree of confidence can be predicted) but also because of the inherent frictions of war, this method can not always work. It requires that we have information superiority, so that we know what the enemy is about to do.
- Command by guidance, in contrast to the other methods, relies on the capabilities and initiatives of all participants of the forces. Here, the commander decides and communicates the outline and minimum goals of an effort. The method places the initiative at the low-level commanders, who must use their local situation awareness to out-think their opponents.

As is seen, regardless of command method chosen, it is necessary to predict what the enemy will do in the future. The authors point out that it might also be important to include what neutral actors of the battlefield will do when making plans. The authors present a table taken from a report on predictive battlespace awareness (PBA) by the scientific advisory board of the US Air Force. The table, which is reproduced here in figure 1, contrasts PBA with the traditional IPB (intelligence preparation of the battlefield) process. After this review, the paper goes on to introduce the concept of Joint Synthetic Battlespace (JSB) as a simulation environment where commanders can train using real equipment. The JSB is also meant to be used online, during a mission. In this use, the authors envision a system that simulates possible futures and allows commander's to test various courses of actions to see which one is the best. The paper concludes by listing a number of technological challenges that must be overcome before a JSB can be realized.

IPB	PBA
Intel-centric	Commander-centric
Product-centric	Process-centric
Structured, iterative process	Dynamic Process
Focuses on red as an independent actor	Emphasizes red-blue-gray interaction and interdependencies
Describes adversary's courses of action	Emphasizes commander's anticipation and pre-emption of the adversary through decisive effects
Focuses on courses of action initiated after inception of conflict/crisis	Enables the commander to shape the pre-post-conflict environment to his advantage; reduces uncertainty during conflict/crisis
Generally delimited by specific geographical boundaries (operational and/or tactical levels)	Provides continuity of awareness from the strategic to the tactical level of operation
Focuses on target identification / definition	Focuses on commander's decisions to produce decisive effects in the battlespace
Stovepiped processes and information-flow	Horizontally, vertically integrated processes; ubiquitous information through the publish-subscribe-broker information architecture

Figure 1 This figure (from Phister et al) shows the conceptual differences between IPB and PBA.

A paper by Piccerillo and Brumbaugh (2004) deals with the relationship between EBAO and Intelligence, Surveillance and Reconnaissance (ISR) operations. It argues that one way of creating a link between EABO and ISR is to consider ISR operations to support anticipatory (or predictive) battlefield awareness. Instead of reacting to the enemy's actions and planning ISR operations in order to get more information on what the enemy is doing right now, we need to predict the enemy's next moves, so that we can direct our ISR resources to the places where they will be, rather than to the places where they are. This means that it is vital to make as good predictions of enemy courses of actions as possible. The approach of the paper can thus be seen as emphasizing level 3 information fusion over level 2, and in particular to allow level 3 needs to dominate those of level 2 when it comes to sensor adaptation. Predictive battlefield awareness is defined by the authors as the state of awareness achieved and maintained by the commander allowing him to correctly anticipate future conditions, focus ISR assets so that their capabilities are used in an optimal way, and shape the battle space, i.e., force the opponent to take those actions which we prefer them to take. Predictive battlefield awareness is a “continuous process providing visualization, exploitation, collaboration, re-tasking, processing and dissemination of information to the appropriate level in an intensely dynamic environment”. The paper emphasizes that some ISR is synchronized with the tempo of the battle, while some are not. Using as examples the application of Predator and Global Hawk in Iraq, the authors note that switching between these two types of missions can lead to reduced efficiency. To alleviate this problem, the authors propose that operators and commanders must work closely together in both the planning and execution phases of ISR. Replanning must be possible. This leads to high demands for the visualization of the ISR operations, so that commander's without detailed expert knowledge can take part in the process.

4 Crowd simulation and computerized analysis of crowd control

In typical missions today, Swedish troops must be able to handle riots and be successful in crowd control. Therefore, the TMDI project decided as one of its activities to use an already developed crowd simulator to study if it is possible from crowd simulations to learn how to control a crowd in a certain city environment.

The functionality of the simulation here is aimed to generate a data base over simulated riot cases. This is done in beforehand for a specific city or suburb with its street and block network, before an assumed critical situation might arise. The military tactical commanders, responsible for keeping security, use the generated data base with the most similar cases for the on-going situation, and study the predictive situation picture given alternative courses of actions.

The problems that we have to solve in order to fulfil this simulation functionality are identified as the following ones:

- An appropriate computer representation of decisions for the genetic algorithm
- Agent simulation models
- Level of simulation complexity

A novelty in our approach is that we combine agent based simulation and genetic learning to generate optimal ways to control a crowd:

4.1 Genetic algorithms for intelligent agent behaviour, optimization and crowd control

4.1.1 What are Genetic Algorithms?

A genetic algorithm is a method within the field of artificial intelligence for solving optimization problems that is based on natural selection, the process that drives evolution. Starting from a population of random individual problem solutions, the genetic algorithm evolves the population by repeated incremental modifications. At each step the genetic algorithm selects at random two individuals from the population as parents and uses them to create an offspring for the next generation. Successively over time the population evolves towards an optimal solution to the problem at hand. This biologically inspired process is a very robust optimization tool suitable for difficult optimization problems.

4.1.2 Optimization and Crowd control

In the last several years there has been a substantial amount of research on how to model crowds and give them complex behaviours using intelligent agents. In spite of all this research there is very little work on the subject of crowd control. There is some work on using genetic algorithms for adaptation of the behaviour of intelligent agents. For instance, Canova and Tyler (1998) describe how the web can provide an environment for modelling the adaptive behaviour of humans using genetic algorithms, and Mitchell et al (1996) used a genetic algorithm to evolve cellular automata rule sets to perform particular tasks. These rule sets guide the behaviour of cellular automata and the genetic algorithm was used to adopt the behaviour when the problem changed.

Several scientists involved in the “Project Albert” (lead by the US Marine Corps, see http://www.projectalbert.org) have also used genetic algorithms for optimization of agent-based behaviour. In a project report Graves et al (2000) demonstrated how to use genetic algorithms to improve the rule bases that defines when individual agents take various actions in agent-based simulation, for example, when advancing towards enemy lines or shooting at enemy soldiers. Dixon and Reynolds (2003) used genetic algorithms for peacekeeping scenarios on their Behaviour Action Simulation Platform (BASP) for cases where it is anticipated that the model will evolve from one or more preceding models.

At the Center for Technology-Enhanced Learning (University of Missouri-Rolla) Chaloupek (2003) used genetic algorithms for crowd control in disaster relief. The focus is on evolving strategies used by the Police or the Fire Brigade for managing the situation when a crowd of people quickly has to evacuate a building. The system developed is intended both for evaluation of current methods, as well as to evolve new methods for managing the crowd during evacuation. The crowd of people is simulated by having a number of intelligent agents, modelling the victims, run around in the building. In the building there are sources of danger such as fires that should be avoided. The Police or the Fire Brigade use strategies such as putting up barricades or directing people away from the scene, to help guide the victims out of the building. These are the strategies that are optimized by the genetic algorithm relative the building and the situation at large in order to manage the crowd and guide it safely out of the building.

If we want to extend the method with replanning capabilities we might use methods similar to the sensor scheduling optimization used by Schrage and Gonsalves (2003) and combine that with genetic algorithms.

At FOI genetic algorithms have previously been used in evolving rule bases for prediction and tactics. Schubert (1996) and Bergsten et al (1997) used genetic algorithms for prediction of future events in connection with antisubmarine warfare. The prediction is aimed at the next few hours and was intended for the tactical commander. They develop a method that can recognize a sequence of incoming intelligence reports. The method is based on a comparison with historical data material, where prediction rules are evolved by the genetic algorithm using the historical data material. Schubert and Jöred (1999) used genetic algorithms in a pilot support project to develop flight tactics in beyond visual range (BVR) air-to-air combat.

4.2 Modeling and simulation of riots

In this part we summarize some of the recent most distinguishing research findings in *agent based approach* for modeling crowds and riots. An agent is everything that can perceive using its sensors and act using its effectors. An agent can be an individual or a group of individuals, Russel and Norvig (1995). In a simulation, a software agent is a representation of real world agent.

Goldstone and Janssen (2005) state: “Agent-based models describe interactions among individual agents and their environment, and provide a process-oriented alternative to descriptive mathematical models”. Considering the role of agents we need different models of agents that are run by the simulation process. According to Jager et al (2001) rioters are seen as a *non-homogeneous group*. In their study, approximately one per cent of the participants are classed as “hard core”. These members frequently survey their surroundings to detect favourable conditions for their actions. The second group, the "hangers on", consists of approximately ten per cent of the total crowd and is similar to the hard core members except that they survey their surroundings less actively and frequently. The third and final group, the bystanders, make up the rest of the crowd and are present mainly as spectators who are interested in what is happening, see also Grieger (2003). The level of cognition and complexity of the agent model description should be conditioned by the role of a real world agent. We think that the main effort should be spent on the modeling riot leaders (called "hard core" by Jager et al (2001)).

4.2.1 Conceptual Agent model

An agent model in crowd simulations could have following properties:

- Crowd dynamic part (model of pedestrian movement)
- Interaction part (interactions between protesters, interactions between protesters and police forces and environment)
- Statistical sampling part (setting initial properties of agents from statistical distributions)

The first part is the crowd dynamics part. The research of professor Dirk Helbing, Helbing and Molnar (1995), has been a guideline for describing how a group of agents moves through an environment and interact with each other. According to him, the motion of pedestrians can be described as if they would be subject to `social forces'. The social force is according to this work as well as Helbing et al (2000) *a sum of inertial, repelling and attracting forces*.

Helbing and Molnar (1995) proposed a model based on physics and socio-psychological forces in order to describe the human crowd behaviour in panic situations. His model describes which forces influence an agent's movement. The resulting force (as well as the change of the agent's velocity) is a sum of those forces according to the formula:

Resulting force on agent = inertia of agent and desired direction during a certain time interval $+ \sum f_{ij} + \sum f_{iw}$.

According to Helbing each agent *i* in the crowd has got a predefined speed**,** i.e. the desired velocity in a certain direction and to which it tends to adapt its instantaneous velocity within a certain time interval (1:st term of the formula), for symbolic representation see Braun et al (2003). Simultaneously, the agents try to keep a velocity-dependent distance from other agents *j* and walls *w* using interaction forces f_{ij} and f_{iw} (2:nd and 3:rd term of the formula), respectively.

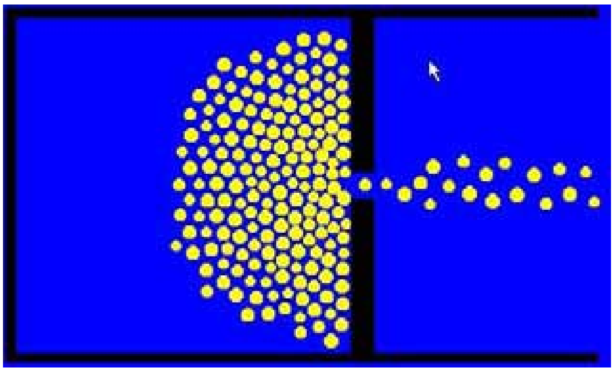

Figure 2: An image of Helbing´s Model Simulation.

Helbing's model generates realistic phenomena, see Figure 2. For example, it describes the formation of arcs at the exit and the increasing evacuation time with increasing desired velocity as described by Helbing, and Molnar (1995).

Inspired by this work Braun et al (2003) add an "altruism" force in their agent model and run comparative simulations. The generalization of Helbing's model described in that work permits simulation of several different populations simultaneously.

The attraction force in Suzić (a) (2006) is dependent on the distance to the agent's places of interest. Those places include meeting points, key buildings and places where own (police or military) forces are located. Inertial forces can also be used. Helbing's computer simulations of interacting pedestrians show that the social force model is capable of describing the self-organization of several observed collective effects of pedestrian behaviour very realistically. Helbings results have been submitted to verification and are *proven* to have high accuracy in crowd modelling.

The second part of the behaviour model is about interaction between different agents and objects. In Suzić (a) (2006) agent models were modelled mainly in a simple and heuristic manner and additional studies are needed. The idea was that out of many small interactions a complex behaviour emerges like in a swarming case, see Bonabeau et al (1999). The interaction between rioters and own force is a matter of force balance and performance in Suzić (a), (2006). By measuring health status of the own force (e.g. heart pulse), rioters level of exhaustion and using performance modifier functions, see Cornwell et al (2002), we could obtain a guess on how leaders of troublemakers would act.

An agent should apply what it knows about the environment and about the own capabilities, desires, and beliefs of other agents when it chooses the (inter)actions that it expects will maximally achieve its own effects. Intuitively, an agent is attempting to quantify how much better off it would be in a state resulting from having performed a given action. In a multi-agent setting, however, an agent usually cannot anticipate future states of the world unless it can hypothesize about the actions of other agents, see Gmytrasiewicz and Durfee (2000), Brynielsson (2002). This estimation of the effects given different actions can be performed by the use of embedded simulations.

The cognitive agent's models require response curves, see Suzić (b) (2006), that are Performance Modifier Functions (PMF) which describe the influence of stress and emotions on decision making. Studies by Cornwell and Silverman (2002), for example, show how the presence of chanting and music effects changes response of protesters to own actions. PMF could be a guideline for what a group of rioters can sustain or not depending on their motivation and emotions.

The third and final part is the statistical sampling part; see Suzić and Wallenius (2005). Given an initial distribution of demonstrators' positive-negative emotions, parameter values for each agent are sampled. Different behaviors of the agents are manifested depending on, for example, anger level. This implies that different behaviors are triggered and eventually different outcomes (effects) occur at each turn of the embedded simulation. Moreover, change of emotional state in a real world agent can lead to a change of its role or to a change that greatly affects its behavior.

4.2.2 Crowd simulation federate

Petty *et. al.* (2004) is developing a crowd federate model. An additional requirement on a crowd federate model is that it has to be compatible (interoperable) with other military systems. According to these authors, "In spite of the military challenges and risks imposed by crowds, models of crowds are essentially absent from current production for military simulations". In a white paper, McKenzie et al (2005) call for developing a crowd simulation capable of generating crowds of non-combatant civilians that exhibit a variety of *realistic* individual and group behaviours at differing levels of fidelity. They provided a summary of crowd modelling from a psychological perspective with special attention to military scenarios. The second perspective of their summary was focused on modelling crowds.

5 Conclusions

The reviewed material shows a large span in maturity of the different fields. Software for text mining and Wikis have been along for quite a while and commercial products (mainly text mining) and open source solutions (Wiki servers) are available, each of them with their cons and pros regarding the type of information mining and structuring they focus on. A through analysis of one's needs should be done before selecting solution. The study of uncertainties in ontologies is a current research filed where not very much has been produced. Regarding user-centric situation awareness, much has been done in various aspects of the field. In this report we put some emphasis on mixed-initiative collaboration for sensor management. Finally, we note that simulations of the dynamics of human crowds have been going on within the simulation and agent modeling scientific communities for quite a while. In our report, we briefly showed a new approach for machine learning via genetic algorithms of the most suitable ways to control the crowd behaviour.

6 References

Ahlberg et al (2004). Ahlberg, S., Hörling, P., Jöred, K., Mårtenson, C., Neider, G., Schubert, J., Sidenbladh, H., Svenson, P., Svensson, P., Undén, K. Walter, J. **The IFD03 information fusion demonstrator**. In *Proc. 7:th Int. Conf. on Information Fusion (FUSION 2004),* P. Svensson and J. Schubert (Eds.), Stockholm, Sweden, 2004.

Ahlberg et al (2006). Ahlberg, S., Hörling, P., Johansson, K., Jöred, K., Kjellström, H., Mårtenson, C., Neider, G., Schubert, J., Svenson, P., Svensson, P., Undén, K., Walter, J. **An information fusion demonstrator for tactical intelligence processing**

in network-based defense. To appear in *Information Fusion*. http://dx.doi.org/10.1016/j.inffus.2005.11.002

Allen (1999). Allen, J.F. **Mixed-initiative Interaction.** *IEEE Intelligent Systems* 14(5), sep-oct 1999.

Bergsten et al (1997). Bergsten, U., Schubert, J., and Svensson, P. (1997). **Applying Data Mining and Machine Learning Techniques to Submarine Intelligence Analysis.** In *Proc. 3:rd Int. Conf. on Knowledge Discovery and Data Mining* (KDD'97), pp. 127−130, US, 1997.

Bonabeau et al (1999). Bonabeau, E. Dorigo, M. Theraulaz, G. **Swarm Intelligence: From Natural to Artificial Systems.** *Oxford University Press*, ISBN 0-19-513159-2, NY, US 1999.

Braun et al (2003). Braun, A., Musse, S.R., de Oliveira, L.P.L., Bodmann, B.E.J. **Modeling Individual Behaviors in Crowd Simulation.** In *Proc. 16:th Int. Conf. on Computer Animation and Social Agents* (CASA 2003) pp 143, Publisher IEEE Computer Society ISBN 0-7695-1934-2, Washington, US, 2003.

Brynielsson (2002). Brynielsson, J. **A Decision-Theoretic Framework Using Rational Agency.** In *Proc. 11th Conf. on Computer-Generated Forces and Behavioral Representation,* number 02-CGF-047, pages 459- 463, Orlando, Florida, US, 2002.

Canova and Tyler (1998). Canova, B.S., and Tyler., J.G. **An adaptive distributed architecture for near real-time genetic algorithm execution.** In *Proc. 1998 Int. Conf. on Web-Based Modeling & Simulation*, P.A. Fishwick, D.R.C. Hill, R. Smith (Eds.), pp. 96–101. The Society for Computer Simulation International, San Diego, CA, USA, 1998.

Cesar et al (2005). Cesar, P., da Costa, G., Laskey, K.B., Laskey, K.J. **PR-OWL: A Bayesian Ontology Language for the Semantic Web.** ISWC 2005 Workshop on Uncertainty Reasoning for the Semantic Web. http://ite.gmu.edu/~klaskey/papers/URSW05_PR-OWL.pdf

Chaloupek (2003). Chaloupek, J. **Improved Crowd Control Utilizing a Distributed Genetic Algorithm.** Center for Technology-Enhanced Learning, University of Missouri-Rolla, USA, 2003.

Cornwell et al (2002). Cornwell, J.B., Silverman, B.G., O'Brien, K., Johns M.A. **Demonstration of the PMFExtraction Approach: Modeling the Effects of Sound on Crowd Behavior.** *Proc. 11:th Conf. on Computer Generated Forces and Behavior Representation,* , pp. 107-113, Orlando FL, US, May 7-9 2002.

Dixon and Reynolds (2003). Dixon D.S., and Reynolds, W.N. **The BASP Agent-Based Modeling Framework: Applications, Scenarios and Lessons Learned.** In *Proc. 36th Hawaii Int. Conf. on System Sciences*, Track 3, Vol. 3, p. 93.3, Hawaii, USA, 6−9 January 2003. IEEE Computer Society, Washington, DC, USA, 2003.

Eklöf and Mårtenson (2006). Eklöf, M., Mårtenson, C. **Ontological Interoperability.** *FOI-R—1943—SE*, January 2006.

Endsley et al (2003). Endsley, M.R., Bolté, B., Jones, D.G. **Designing for situation awareness.** *Taylor & Francis publishers* (ISBN 0-7484-0966-1).

Giles (2006). *Nature,* December 2005. http://www.nature.com/news/2005/051212/full/438900a.html)

Gmytrasiewicz and Durfee (2000). Gmytrasiewicz, P.J. and Durfee, E.H. **Rational coordination in multi-agent environments.** *Autonomous Agents and Multi-Agent Systems,* 3(4):319-350, 2002.

Goldstone and Janssen (2005). Goldstone, R.L., Janssen, M.A. *Computational models of collective behaviour Trends in Cognitive Sciences,* Vol. 9(9), pp. 424-430, September 2005.

Graves et al (2000). Graves, T., Picard, R., and Upton, S. **Improving Rule Bases for Agent Based Simulations.** Project Albert, Marine Corps Warfighting Laboratory, United States Marine Corps Combat Development Command, Quantico, VA, USA, 2000.

Grieger (2003). Grieger, D. **An Overview of Crowd Control Theory and Considerations for the Employment of Non- Lethal Weapons.** DSTO System Sciences, Technical report DSTO-GD-0373, Laboratory, Edinburgh, AU, 2003.

Grimes (2003). Grimes, S. **The word on Text Mining** *Intelligent Enterprise* Vol. 6(19) 2003. http://www.intelligententerprise.com/031210/619decision1_1.jhtml;jsessionid=WQAPKGLEFGKI4QSNDBGCKH0CJUMEKJVN

Helbing, and Molnar (1995). Helbing, D. and Molnar, P. **Social force model for pedestrian dynamics.** *Phys. Rev. E* 51, pp. 4282-4286, 1995.

Helbing et al (2000). Helbing, D., Farkas, I.., Vicsek, T. **Simulating Dynamical Features of Escape Panic.** *Nature* 407, pp. 487-490, 2000.

IBM (2006). http://www.research.ibm.com/UIMA

Jager et al (2001). Jager, W., Popping, R., van de Sande, H. **Clustering and Fighting in Two-party Crowds: Simulating the Approach-avoidance Conflict.** *Journal of Artificial Societies and Social Simulation,* Vol 4(3), 2001.

Kadar et al (2004). Kadar, I., Chu Chang, K., O'Connor, K., Liggins, M. **Figures-of-merit to bridge fusion, long term prediction and dynamic sensor management.** In *Proc. SPIE* 5429, p 343, 2004.

Kokar and Matheus. Kokar, M.M., Matheus, C.J. **Adding Uncertainty to Ontologies – Some Questions.** (unpublished manuscript).

Matheus et al (2005). Matheus, C.J., Kokar, M.M., Baclawski, K., Letkowski, J.A., Call, C., Hinman, M., Salerno, J., Boulware, D. **SAWA: An Assistant for Higher-Level Fusion and Situation Awareness.** In *Proc. SPIE*, 5813, 75, pp. 75-85, 2005.

McKenzie et al (2005) McKenzie, F.D., Petty, M.D., Kruszewski, P.A., Gaskins, R.C., Nguyen, Q.-A. H., Seevinck, J., Weisel, E. W. Crowd **Modeling for Military Simulations using Game Technology.** *White paper*, http://www.engenuitytech.com/communities/whitepapers_form.php?wref=3, Canada, 2005.

Miller (2006). Miller, J. **"Happy together: Sensors and users"** *Internet Computing*, March-April 2006.

Mitchell et al (1996). Mitchell, M., Crutchfield, J.P., and Das, R.. **Evolving cellular automata with genetic algorithms: A review of recent work.** In *Proc. 1st Int. Conf. on Evolutionary Computation and its Applications* (EvCA'96). Russian Academy of Sciences, Russia, 1996.

Mullen et al (2006). Mullen, T., Avasarala, V., Hall, D.L. **Customer-Driven Sensor Management.** *IEEE Intelligent Systems* march-april 2006 http://csdl2.computer.org/dl/mags/ex/2006/02/x2041.pdf

Mårtenson and Brännström (2006). Mårtenson, C., Brännström, M. **Enhancing situational awareness by exploiting wiki technology.** In *Proc. CIMI*, paper S3.2, Enköping, Sweden, may 16 – 18, 2006 (ISBN 91-89410-03-03, also as FOI-S—2160—SE).

Mårtensson and Svenson (2005). Mårtenson, C., Svenson, P. **Evaluating sensor allocations using equivalence classes of multi-target paths.** In *Proc. 8:th Int. Conf. on Information Fusion.* Philadelphia, USA, 25-29 July 2005. Paper B9-1, pp. 1-8. (Also in FOI-S--1817--SE, Swedish Defence Research Agency, 2005)

Noble (2004). Noble, D. F. **Assessing the Reliability of Open Source Information.** In *Proc. 7:th Int. Conf. on Information Fusion,* Stockholm, Sweden, 2004. http://www.fusion2004.foi.se/papers/IF04-1172.pdf

Noble (2005). Noble, D. F. Fusion of Open Source Information. In *Proc. 8:th Int. Conf. on Information Fusion*, Philadelphia, U.S., 2005. **http://www.ebrinc.com/files/Noble_Fusion.pdf**

Noy et al (2006). Noy, N., Rector, A., Hayes, P., Welty, C. **Defining n-ary relations on the semantic web.** Working Group Note 12 April 2006. http://www.w3.org/TR/2006/NOTE-swbp-n-aryRelations-20060412/

Petty et al (2004). Petty, M.D., McKenzie, F.D., Gaskins, R.C., Weisel, E.W. **Developing a Crowd Federate for Military Simulation.** In *Proc. Spring 2004 Simulation Interoperability Workshop*, Arlington VA, US, April 18-23 2004, pp. 483-493, 2004.

Phister et al (2003). Phister, P.W., Busch, T., Plonisch, I.G. **Joint synthetic battlespace: Cornerstone for predictive battlespace awareness.** In *Proc. CCRTS*, 2003. http://www.dodccrp.org/events/2003/8th_ICCRTS/pdf/005.pdf

Piccerillo and Brumbaugh (2004). Piccerillo, R.A., Brumbaugh, D.A. **Predictive Battlespace Awareness: Linking Intelligence, Surveillance and Reconnaissance Operations to Effects Based Operations.** In *Proc. CCRTS*, 2004. http://www.dodccrp.org/events/2004/CCRTS_San_Diego/CD/papers/132.pdf

Rasmussen (1998). Rasmussen, P.H. **Data Fusion Demonstrator, Projektstatus med udgangen af 1997.** *FOFT F-18/1998,* Forsvarets forskningstjenste, Köpenhamn, Danmark,1998.

Russell and Norvig (1995). Russell, S.J., Norvig, P. **Artificial Intelligence.** ISBN 0-13-103805-2, Prentice-Hall, New Jersey 1995.

Salerno et al (2004). Salerno, J., Hinman, M., Boulware, D. **Building A Framework For Situation Awareness.** In *Proc. 7:th Int. Conf. on Information Fusion*, Stockholm, Sweden, pp 219-226, 2004.

Salnerno et al (a) (2005). Salerno, J., Hinman, M., Boulware, D. **A Situation Awareness Model Applied To Multiple Domains.** In *Proc. SPIE* 5813, 65, 2005.

Salnerno et al (b) (2005). Salerno, J., Blasch, E., Hinman, M. Boulware, D. **Evaluating Algorithmic Techniques in Supporting Situation Awareness.** In *Proc. SPIE* 5813, 2005.

Schrage and Gonsalves (2003). Schrage, D., Gonsalves, P.G. **Sensor Scheduling Using Ant Colony Optimization.** In *Proc. 6:th Int. Conf. on Information Fusion*, pp. 379-385, Cairns, Australia, 2003.

Schubert (1996). Schubert, J. **Prediction of foreign submarines future position.** *FOA-R—96 00296-3.4--SE*, 1996.

Schubert, J., and Jöred K., (1999). Unpublished presentation.

Smith et al (2005). Smith, S.F., Hildum, D.W., Crimm, D.R. **Comirem: An Intelligent Form for Resource Management.** *IEEE Intelligent Systems* 20(2), 2005. http://ieeexplore.ieee.org/iel5/9670/30619/01413167.pdf?arnumber=1413167

Suzić and Wallenius (2005). Suzić, R. and Wallenius, K**. Effects Based Decision Support for Riot Control: Employing Influence Diagrams and Embedded Simulation.** In *Proc. of Workshop on Situation Management,* SIMA, US, 2005.

Suzić (a) (2006). Suzić, R. **Embedded Simulation Systems for Network Based Defense.** *FOI-R--1828--SE,* Swedish Defence Research Agency, Stockholm, Sweden, Stockholm, Sweden, 2006.

Suzić (b) (2006). Suzić, R. **A generic model of plan recognition using embedded simulations, microeconomics and behavior models.** Accepted for publication in BRIMS Conference May 15-18, Baltimore, MD, US, 2006.

Svenson and Mårtenson (2006). **SB-Plan: Simulation-based support for resource allocation and mission planning.** In *Proc. CIMI*, paper S3.6, Enköping, Sweden, may 16 – 18, 2006 (ISBN 91-89410-03-03, also as FOI-S—2161--SE).

Svenson (a) (2006). **Social Network Analysis And Information Fusion For Anti-Terrorism.** In *Proc. CIMI*, paper S3.1, Enköping, Sweden, may 16 – 18, 2006 (ISBN 91-89410-03-03, also as FOI-S—2159—SE).

Svenson (b) (2006). **Force aggregation for OOTW: classification based on capabilities.** In *Proc. CIMI*, paper S1.2, Enköping, Sweden, may 16 – 18, 2006 (ISBN 91-89410-03-03, also as FOI-S—2162—SE).

Völkel et al (2006). Völkel, M., Krötzsch, M., Vrandecic, D., Haller, H., Studer, R. **Semantic Wikipedia.** Submitted to WWW2006. (Accessible june 8:th 2006 at http://www2006.org/programme/files/pdf/4039.pdf)

Wagner et al (2006). Wagner, C., Cheung, K. S. K., Ip, R.K.F. **Building Semantic Webs for e-government with Wiki technology.** *Electronic Government,* Vol. 3(1), 2006. http://colab.cim3.net/file/work/SICoP/2006-02-09/EGov%20Wiki.pdf

Weiss et al (2004). Weiss, S., Indurkhya, N., Zhang, T., Damerau , F. **Text Mining: Predictive Methods for Analyzing Unstructured Information.** *Springer, 1:st ed.* 2004. (ISBN 0387954333).