Military Ontologies for Information Dissemination at the Tactical Edge

Emily C. LeBlanc, Duc N. Nguyen, Marcello Balduccini, William C. Regli
Applied Informatics Group
Drexel University
Philadelphia, PA

Joseph B. Kopena, Thomas A. Wambold
Bellerophon Mobile
Philadelphia, PA

Abstract

In this paper we present our methodology for constructing the Tactical Heterogenous Ontology Representation (THOR), a doctrine-based military ontology that models numerous aspects of the military domain as described by field manuals, joint publications, and scenarios informed by subject matter experts. Developed for DARPA's Content-Based Mobile Edge Network (CBMEN) program, the ontology provides the conceptual vocabulary needed to describe and request the warfighter-generated content. CBMEN leverages smart device technologies alongside concepts from content-based networking and knowledge representation in order to enable intelligent reasoning over dynamic mission data and the efficient transmission of critical mission content at the tactical edge.

Introduction

Warfighters at the tactical edge are generally required to communicate with base facilities in order to report and retrieve new information about their locations, missions, and potential threats. With current communications technology, this critical operational content is not always immediately accessible nor distributable for warfighters who are outside of the range of communication of their bases. Lacking new information may compromise missions and possibly the safety of active units. There is a need for reliable and efficient content distribution and retrieval of battlefield content at the tactical edge.

DARPA's Content-Based Mobile Edge Networking (CBMEN) program aims to enable ad hoc communication among smart devices which are in range of one another. The smart-device equipped warfighters may publish about or request information about mission content such as observation reports, updated orders, or mission-time imagery. Each piece of content in the system is tagged with semantically rich metadata that can be queried and reasoned over in an intelligent way as to return the most relevant available results. In order to tag mission content in a meaningful way as described, we require an intuitive and well-structured representation of the military domain. In this paper we present the methodology used to develop the Tactical Heterogenous Ontology Representation (THOR). THOR's ontology is a doctrine-base military domain representation that models a wide landscape of essential concepts and relationships described in field manuals, joint publications, and scenarios developed with subject matter experts. The ontology provides the terms, concepts, and relationships necessary to describe critical mission content, thus enabling intelligent reasoning over dynamic mission data in order to increase situational awareness for warfighters.

The section that follows discusses background information about the work's motivation and technologies of the CBMEN program. Next, we will describe our methodology for creating the ontology and its coverage of the military domain. Following that, we discuss in greater detail the role of rich metadata for the effective dissemination of battlefield content along with some examples of class definitions and instances. The next section gives examples of competency questions for determining the usefulness of the ontology in the military scenarios. In the following section, we give an overview of the complete application developed by Drexel University and Bellerophon for the CBMEN program. Finally, we present a scenario that illustrates the use of the ontology and reasoning in a simulated mission.

Background and Motivation

As a motivation for our work, consider the following scenario in which a platoon of dismounted warfighters are carrying out patrolling operations in a remote region. A fireteam leader of a squad uses his connected mobile device to snap a photo of a suspicious object and identifies it as a potential explosive device outside a remote village and attempts to share this photo with members of the entire platoon and commanders and intelligence teams at command posts.

These devices utilize tactical mobile ad hoc networks (MANETs) to exchange this content and information amongst platoon and squad members where fixed network infrastructures (e.g. satelite or mobile cellular networks) do not exist. MANET environments utilize infrastructureless ad hoc wireless networks (ad hoc WiFi), optimized routing (e.g. OLSR (Clausen and Jacquet 2003)), caching and opportunistic forward approaches (e.g. delay-tolerant networking (Cerf et al. 2007)), and additional middleware to perform content retrieval, storage, and forwarding in lieu of a standard fixed infrastructure network (Sivakami and Nawaz

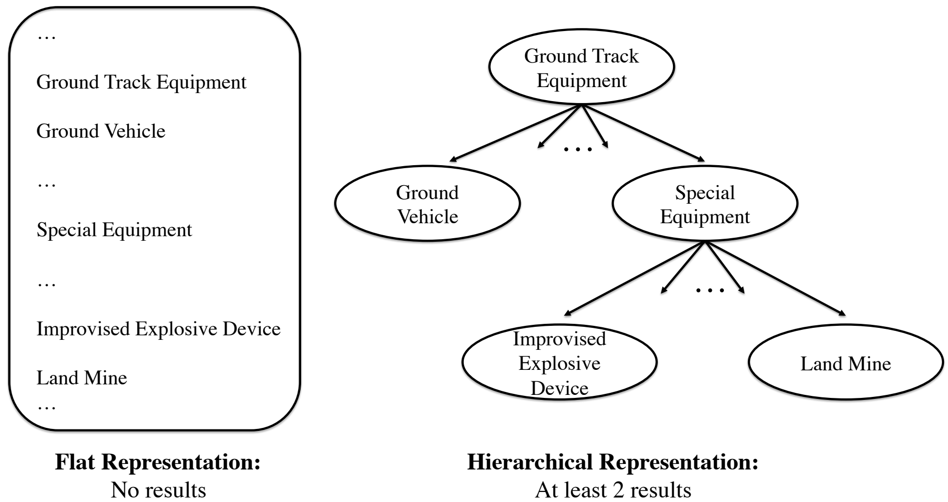

Figure 1: Flat and Hierarchical Representation

2011).

DARPA's CBMEN program leverages concepts from content-based networking, knowledge representation, and reasoning in order to enable information dissemination at the tactical edge. The program aims to maximize warfighter situational awareness by facilitating reliable and efficient access of battlefield content in tactical edge networks where reliable connectivity is not guaranteed. Mobile devices operating in this tactical edge network context must utilize local secondary storage and processing capabilities to perform many tasks. Modern tactical edge networking is therefore based on disruption tolerant and information-centric techniques, utilizing extensive caching and comparatively expressive addressing.

In content-based networking (Jacobson et al. 2009), the files (or content) of the system are addressed by name rather than by their host machine location. Content for the mobile-device-equipped warfighter is in the form of files such as imagery, map tiles, orders, and documents. Rather than exchanging the content itself, the nodes within CBMEN will exchange metadata describing available content. If the metadata description of content is relevant to the warfighter, then CBMEN system will deliver the binary files to their device for consumption.

The ramifications of exchanging meaningful metadata instead of the actual content in an edge-network are wide-ranging, but its primary effect is in the reduction of the amount of data flooding an already resource-constrained mobile edge network, enabling exchange of higher priority content when the need arises.

The CBMEN system is composed of five primary components: applications, the content naming registrar, the Autogen interface, the content management service, content datastore, and content distribution service. The content management service attaches unique identifiers to content for use by the registrar, datastore and distribution services. The content datastore holds content created by the device and cached content from other devices. The distribution service efficiently publishes and receives content to and from the network. The content management, content datastore, and content distribution service are the key components of the content-access middleware charged with disseminating the metadata and content efficiently across the tactical edge network. It is beyond the scope of this work and is described in detail in (Strayer et al. 2013).

The ontology forms the conceptual backbone of the system with the content metadata serving as instance data over which the registrar will reason. The developed ontology constructs utilize basic class, subclass relationships. They are written in OWL and restricted to OWL-Lite. A detailed description of the rationale and methodology for the ontology and its development is described in the following section.

Ontology

The THOR ontology is composed of a number of sub-ontologies modeling a number of military sub-domain concepts, properties, and relationships. Each sub-ontology is organized by resource and contains extensive taxonomies of the military warfighter domain. It extends the work of the command and control ontology (Nguyen et al. 2010), incorporating knowledge from resources such as US Army FM 5-0 Operations Process (U.S. Department of Army 2007), MIL-STD 2525 Military Symbology (U.S. Department of Defense 2008) and other official doctrine. Using these resources to create the domain representation ensures that military users of all ranks will be familiar with the vocabulary and basic structures of the representation. Other Department of Defense data models such as UCore [1], and NIEM [2] have also been consulted and concepts have been incorporated into the THOR ontology as needed.

[1] http://ise.gov/universal-core-ucore

[2] https://www.niem.gov/Pages/default.aspx

In addition to the doctrinal representations, the ontology incorporates generic communication entities and domain-independent features capturing basic message structure, geo-temporal marking, and provenance. The ontology is encoded using RDF Schema (RDFS)[3], and the Web Ontology Language (OWL)[4], version 2, $\mathcal{EL}$ profile for tractable expressive description logic (Baader, Brand, and Lutz 2005), as well as Terse RDF Triple Language[5] for encoding instances of ontology classes. The languages provide a sufficiently constrained set of constructs for modeling the domain, enabling a great deal of knowledge to be captured in forms that are easily computable by smaller mobile devices. The ontology employs the languages to construct formal conceptualizations of military domain aspects relevant to dismounted warfighters, CBMEN's primary intended users.

The ontology provides a substantial basis for integrating a wide range of military applications into the system. Moreover, it is designed for future expansion and evolution. Other domains and application specializations may be incorporated in the future without modification of the core ontology.

Methodology

An iterative process was used to develop the THOR ontology. The process has been roughly divided into two tracks, corresponding to structured knowledge contained within the set of military doctrine and domain-independent documentation and for military-specific development that falls outside of what is immediately available from the doctrine.

The essential steps for the doctrinal track include:

1. *Conceptual Deconstruction:* Studying source material, such as military doctrine and sample application data, and identifying the fundamental concepts and relationships therein. The main products of this step are documents listing terms and properties and representing natural connections between them.
2. *Concept Analysis and Organization:* Organizing the concepts and relationships of the previous step into precisely defined structures. The output of this step is a collection of spreadsheets capturing these structures by category.
3. *Encoding:* Transcribing the structures of the previous step into RDF/OWL documents representing the ontology.
4. *Evaluation:* Testing the resulting iteration of the ontology to determine its ability to capture the required data and support applications in the domain, and iterating over the development process as necessary until the ontology meets the requirements.

For the military-specific development that falls outside of enumeration by field manuals, realistic mission scenarios have been deconstructed using the process shown in Figure 2:

1. Identifying all top-level concepts among the data objects manipulated and background knowledge used in this scenario;
2. Determining class–subclass relationships among them;
3. Elaborating and identifying properties of those classes;
4. Specifying data domains and ranges for those properties;
5. Checking this deconstruction against the scenario elements and iterating as necessary.

The scenario deconstruction is then utilized to evaluate the capabilities of the THOR ontology and drive further development until all military-specific elements and relationships introduced by the scenario are represented by the ontology.

Evaluation of the ontology has been conducted using the three following mechanisms:

- Reviewing with subject matter experts (SMEs).
- Theoretical application to a motivating scenario. For example, given a mission scenario, is the current ontology suitable to describe all necessary aspects of the situation?
- Collection of and testing against a set of competency questions (described in a later section of this paper).

Coverage

Following from these iterative development processes and mechanisms, the THOR ontology supports describing and querying content associated with a wide swatch of operationally relevant applications and data. An overview of the major topic areas is as follows:

- Data Specifications, e.g., file formats and languages
- Force Structure, e.g., battalion, company, platoon, squad
- Unit and Personnel Roles, e.g., unit identity, size
- Equipment, e.g., ground vehicles, ships, aircraft
- Mission Types, e.g., Patrol, raid, attack, defend
- Operations, e.g., movement descriptions and routes, tasks and maneuvers
- Geography and Environment Conditions, e.g., MGRS, Lat/Lon
- Military Message Types, e.g., spot reports, PLI, orders, priorities
- Provenance, e.g., authorship, time-stamping, and versioning

Rich Metadata for Battlefield Content

THOR's approach to addressing and requesting objects has several distinct advantages over other schemes used in content-based networking. Most of these are based on essentially opaque string descriptors, which are generally both human-oriented and flat. These schemes present a number of shortcomings, including errors in human data entry, an inability to seamlessly scale specificity/generality, the necessity of publishers and consumers using exactly the same descriptors for a connection to be realized, and unnecessary verbosity.

[3] `http://www.w3.org/TR/rdf-schema/`
[4] `http://www.w3.org/standards/techs/owl`
[5] `http://www.w3.org/TR/turtle/`

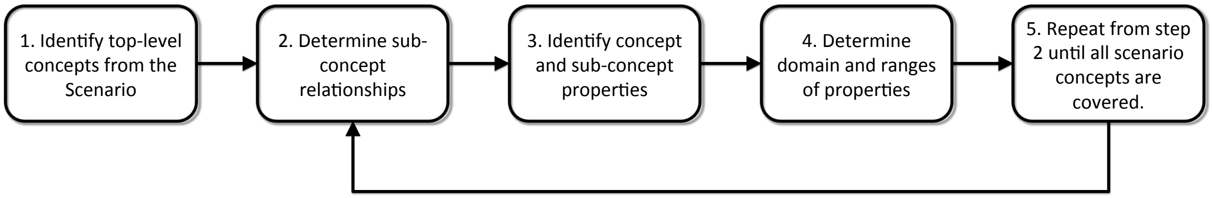

Figure 2: Process for deconstructing a motivating scenario.

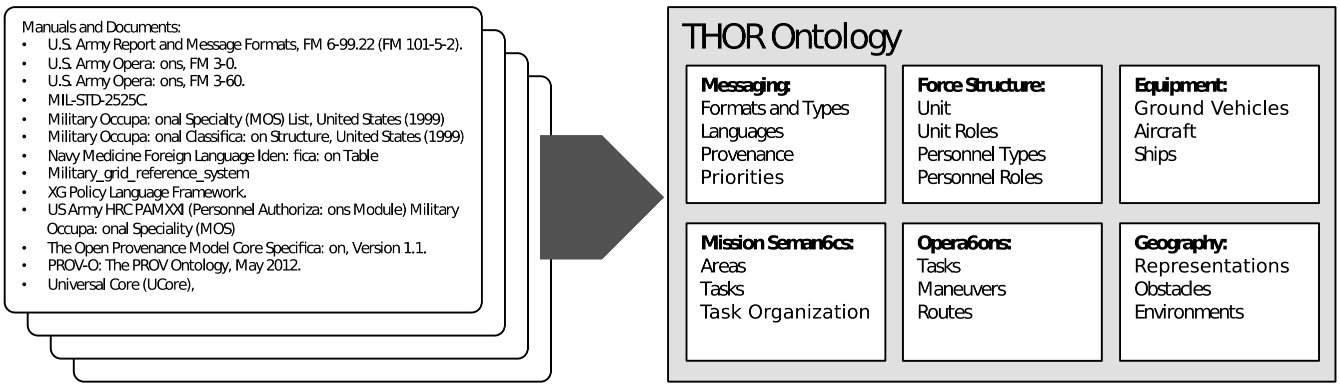

Figure 3: THOR ontology sources and key topic areas.

THOR's ontology-backed metadata addresses these concerns. The structured but schema-free metadata model enables content to be expressively described, enabling applications and users to capture whatever aspects they deem potentially relevant. Extensive taxonomies and automated super/sub-class inferences further enable searches and subscriptions to be as general or specific as necessary while maintaining comparative terseness.

Figure 1 illustrates how hierarchical organization of terms facilitates this better than flat labels, enabling the system to return the most fitting set of information for a query. For example, Warfighter A observes an improvised explosive device (IED) on their patrol. They publish a Spot Report about the IED, using the term "equipment:ImprovisedExplosiveDevice". Warfighter B is moving into the same area looking for suspicious equipment and subscribes to Spot Reports about "equipment:GroundTrackEquipment". Because "equipment:IED" is a subclass of "equipment:GroundTrackEquipment", THOR will recognize Warfighter A's report of an IED sighting as content that should be delivered to Warfighter B.

Here we will provide some examples of how one creates instances of the classes in the ontology. First is the encoding of the domain concept "GroundTrackEquipment" in Turtle.

```
<owl:Class ref:ID=GroundTrackEquipment>

<rdfs:label xml:lang=en>
Ground Track Equipment</rdfs:label>

<rdfs:subClassOf>
<owl:Class ref:ID=Equipment/>
<rdfs:subClassOf>

</owl:Class>
```

The following is a sample instance of the class "GroundTrackEquipment" named "e234".

```
:e234 a equipment:GroundTrackEquipment

equipment:mil2525_affiliation [a roles:Hostile];
```

The following is a sample instance of a Spot Report named M2 about an instance of ground track equipment. M2 contains an image of the equipment being reported, and the report is created by the Rifleman. In the interest of space, the prefix "messages" is abbreviated to "m", and the class "GroundTrackEquipment" to "GTE".

```
:M2 a       msg:SpotReport ;

m:contentType [ a provenance:ImageEntity ] ;
m:equipment [ a equipment:GTE ] ;
m:originator [ a roles:Rifleman ] ;
```

Competency Questions

Another technique used in developing and evaluating the THOR ontology is the use of competency questions - specific use cases applying the ontology to automated reasoning, defining a set of queries the ontology must be able to answer. A list of these questions in natural language was developed through discussion with SMEs and studying motivating scenarios. Each question was then modeled as sample data and a query using the ontology, with refinements and additions being prompted if that proved impossible.

The following example represents a class of simple search queries:

Retrieve all spot reports observing ground vehicles.

This request may be encoded as a query in the THOR ontology as follows:

```
SELECT ?id ?lat ?lon ?equ  ?reportUnit
       ?reportUnittype ?equtype

WHERE {
   ?id a messages:SpotReport .
   ?id messages:latitude ?lat .
   ?id messages:longitude ?lon .
   ?id messages:equipment ?equ .
   ?id messages:priority ?reportUnit .
   ?reportUnit rdf:type ?reportUnittype .
   ?equ a equipment:GroundVehicle .
}
```

With the sample data conceived from motivating scenarios, matching this query requires applying a taxonomy of vehicles, e.g., that a Jeep and a tank are both ground vehicles.

Another example asks for a more specific data value:

Return the collection of available spot reports about hostile people and their affiliation.

The ontology permits capturing this in SPARQL as follows:

```
SELECT ?id ?lat ?lon ?equ ?afftype ?equtype

WHERE {
   ?id a messages:SpotReport .
   ?id messages:latitude ?lat .
   ?id messages:longitude ?lon .
   ?id messages:equipment ?equ .
   ?equ a equipment:SpecialEquipment .
   ?equ equipment:mil2525_affiliation
        [ a roles:Hostile ] .
   ?equ equipment:mil2525_affiliation ?aff .
   ?aff a ?afftype .
      MINUS { ?aff rdf:type ?subtype .
              ?subtype rdfs:subClassOf
                ?afftype .
      }
}
```

This query includes a complex construct selecting the most specific affiliation, rather than all its taxonomic antecedents, based on the MIL2525 labels. Correctly matching this query requires both a taxonomy of affiliations in the ontology, and the ability of the reasoner to process the *MINUS* operator used in this query to exclude redundant results. The following example question requires testing literal data:

Return the collection of available reports that fall within a specified area.

This may be rendered in SPARQL as a query using the ontology:

```
SELECT ?id ?lat ?lon ?equ  ?reportUnit
       ?reportUnittype ?equtype

WHERE {
   ?id a messages:SpotReport .
   ?id messages:latitude ?lat .
   ?id messages:longitude ?lon .
   ?id messages:equipment ?equ .
   ?id messages:priority ?reportUnit .
   ?reportUnit rdf:type ?reportUnittype .
   ?equ a equipment:GroundVehicle .

   FILTER (?lat < 3818 && ?lat > 3817)
   FILTER (?lon > -7730 && ?lon < -7729)
}
```

This question demonstrates that content may be associated with particular positions, and that it is possible to reason over those defining regions of interest.

The ontology defines a rich representation of military content using terms and relationships derived from subject matter experts and military doctrine precisely to enable this sort of inference. Using it, warfighters may specify exactly what they are publishing and what content they wish to receive, being as specific or general as they wish with the system making automated connections as appropriate. In this way they may better manage the extensive volume and diverse variety of content that will be present on the battlefield.

Registrar

The purpose of this section is to give an overview of the complete application developed by Drexel University and Bellerophon for the CBMEN program. THOR's Content Naming Service is CBMEN's registrar, implementing its search and subscription capabilities. This component consists of a thin shell interfacing with the larger CBMEN system that passes data to and from Masterchief. Developed for the CBMEN program, Masterchief is a general purpose Semantic Web knowledge base designed specifically to support applications on mobile devices, and the applications are both intended to be installed on the device of every warfighter. In CBMEN it stores content metadata and applies it to searches and subscriptions enacted by applications.

Masterchief

Inside Masterchief there are three major pathways, implementing an RDF triple store, SPARQL queries, and subscriptions and OWL reasoning. All of these are rooted on an SQLite database which provides the underlying data storage

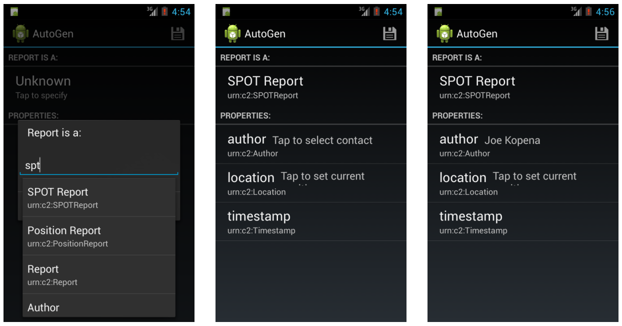

Figure 4: The Autogen interface reporting options and populating fields.

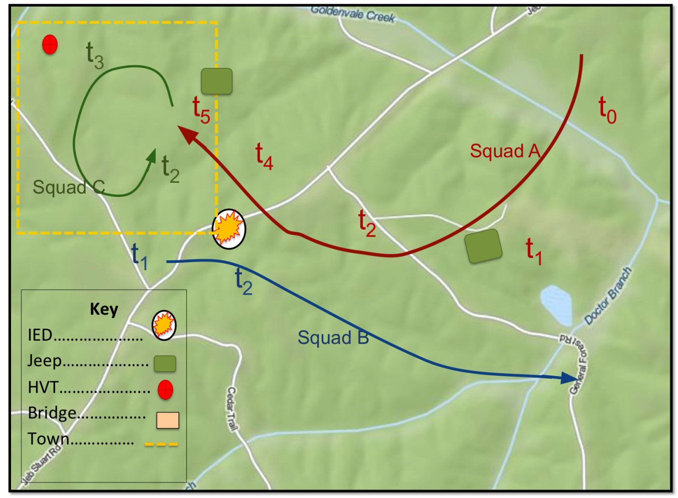

Figure 5: Events and Movement from Motivating Scenario

and manipulation. A major design point here is the assumption that modern mobile devices have limited primary memory but relatively plentiful secondary memory. Masterchief thus pushes nearly all logic and storage into the database, rendering its primary memory consumption very low and actually constant in the absence of a few particular query operators. This approach also enables persistence and ready rehydration of the registrar across system reboots.

Searches are executed in Masterchief by transforming the specified SPARQL query into SQL over the triple store's partition scheme. A recursive traversal of the SPARQL structure generates a single SQL statement capturing the entire output of the query.

Autogen

THOR's metadata approach to labeling and querying content provides for sophisticated filtering and discovery. Creating appropriate metadata that is as richly descriptive as possible though is a significant challenge. Users must be presented with an intuitive interface and the ability to rapidly navigate and utilize the extensive ontological structure. Further, this interface must be populated from the ontology itself as hardcoding this interface will render it likely to become unsynchronized with ontology updates over time.

Constructing and maintaining such an interface is a significant burden to place on CBMEN application developers. Appropriately interrogating the ontology requires a fair amount of familiarity with knowledge representation and specifically RDFS/OWL. A significant amount of work may also be entailed to connect the ontology to an interface.

These concerns are addressed in CBMEN via the Autogen component. Autogen is a free-standing Android activity that any application may invoke to have the user populate metadata. While applications are free to generate metadata in any way they wish, this provides an easy mechanism to populate descriptions for publishing with content.

Autogen provides an interface for the user to attach metadata in the form of instance data to content. For each piece of content, Autogen will query Masterchief for an initial set of concepts for which to generate and attach instance data. Once the entry is complete, the constructed object is returned to the invoking application where it may be readily serialized to RDF and published with the associated content. In this way Autogen provides applications a dynamically created interface for generating metadata based on the THOR ontology that is easy to use for both application developers and users.

Scenario

Several operational scenarios have been used to motivate and evaluate the THOR ontology during development. These have been created alongside and shared with the larger evaluation and demonstration efforts of the entire CBMEN system. The following example illustrates how CBMEN would be used to support dismounted warfighters exchanging content in an area of operations (AO).

Dismounts from three squads are equipped with CBMEN-enabled mobile devices. The squads each have a different mission: Squads A and B are conducting patrol missions nearby a town where Squad C, a quick reactionary force (QRF), is deployed to capture a known high-value individual. Prior to the mission, the CBMEN devices are configured to subscribe to relevant content publications based on the dismount's role and mission. Once configured and ready to deploy, Squads A and B carry out their missions recording and publishing reports, including observations of potential Improvised Explosive Devices (IEDs) and IED factories, persons of interest, road obstructions, and other alerts. Squad C identifies and captures the target high-value individual and publishes a spot report with an image of them. The forward operating base receives the report of the HVI capture and image and other relevant reports of IEDs and other priority reports. Upon mission completion, each squad returns to their forward operating base for debriefing as well as aggregating and archiving all relevant mission data. Figure 5 shows the activities of each squad in the AO.

Now we observe the scenario in more detail, looking at the events of each time step. Each Squad has three devices, held by the Squad Leader, a Fire Team Leader, and the Fire Team's Automatic Rifleman. The set of possible subscriptions contains spot reports about Special Equipment with

Unit Role	Squad A	Squad B	Squad C
Squad Leader	Special Equipment	Special Equipment	Special Equipment
Fire Team Leader	Special Equipment, HVI	Special Equipment	Special Equipment
Rifleman	Special Equipment	IED, Obstacles	Special Equipment

Table 1: Unit role content interests for each squad.

Time step	Squad A	Squad B	Squad C
t_1	Observes and Reports: Jeep	Observes and Reports: IED	-
t_2	Receives: IED	-	Receives: IED
t_3	Observes and Reports: Bridge	-	Observes and Reports: HVI
t_4	Receives: HVI	-	-

Table 2: Timeline of publications and satisfied subscriptions.

Hostile affiliation, High Value Target(HVT) with Suspicious affiliation, and Ground Track Vehicles. Squad subscriptions are detailed in Table 1. Note that there is no subscription for spot reports about Ground Track Vehicles.

At time t1, Squad A sees a seemingly suspicious jeep driving outside of the perimeter of the AO. At the same time step, Squad B observes an unexploded IED. Squads A and B cross paths at time t2, and Squads A and C receive Bs spot report about IEDs because they are both interested in Special Equipment (B does not receive a duplicate of its own spot report). At time t3, Squad C observes what might be a known HVT ducking between buildings in the village. In the last time step t4, Squad A receives C's spot report about the potential HVT because it is interested in HVTs. No Squad receives the spot report about the Jeep because no one has a subscription for spot reports about Ground Track Vehicles. Table 2 gives a breaks down each squad's event by time step.

In this scenario, we have shown that squads can subscribe to reports about high level concepts in the hierarchy (Special Equipment) and can receive reports about concepts subsumed (IED) by the higher level class. We have also demonstrated that even if a report is available to a warfighter by proximity (Jeep), she will not receive the message because there is no subscription or active query looking for that class or any of its ancestors.

Conclusion

In this paper, we discussed the necessity of a doctrine-based ontology as part of a larger solution to the issue of disseminating critical information at the tactical edge. We described the methodologies and sources used to build the ontology, and provided samples of class instances as well as some complex SPARQL queries that can be used to reason for relevant information indirectly. For context, we described the reasoning system, Masterchief, and the ontology interface, Autogen. Finally, we gave a sample scenario in which highlighted some of the less obvious capabilities of the ontology and reasoning system.

Acknowledgement

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) and SPAWAR Systems Center Pacific (SSC Pacific) under Contract No. N66001-12-C-4052. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of DARPA or SSC Pacific.

References

Baader, F.; Brand, S.; and Lutz, C. 2005. Pushing the $\mathcal{EL}$ envelope. In *In Proc. of IJCAI 2005*, 364–369. Morgan-Kaufmann Publishers.

Cerf, V.; Burleigh, S.; Hooke, A.; Torgerson, L.; Durst, R.; Scott, K.; Fall, K.; and Weiss, H. 2007. Delay-tolerant networking architecture. Technical report, RFC 4838, April.

Clausen, T., and Jacquet, P. 2003. Optimized Link State Routing Protocol (OLSR). RFC 3626, Internet Engineering Task Force.

Jacobson, V.; Smetters, D. K.; Thornton, J. D.; Plass, M. F.; Briggs, N. H.; and Braynard, R. L. 2009. Networking named content. In *CoNEXT '09*.

Nguyen, D. N.; Kopena, J. B.; Loo, B. T.; and Regli, W. C. 2010. Ontologies for distributed command and control messaging. *Formal Ontologies in Information Systems*.

Sivakami, R., and Nawaz, G. 2011. Secured communication for manets in military. In *Computer, Communication and Electrical Technology (ICCCET), 2011 International Conference on*, 146–151.

Strayer, T.; Kawadia, V.; Caro, A.; Nelson, S.; Ryder, D.; Clark, C.; Sadeghi, K.; Tedesco, B.; and DeRosa, O. 2013. CASCADE: Content access system for the combat-agile distributed environment. In *Military Communications Conference, IEEE*, 1518–1523.

U.S. Department of Army. 2007. The operations process. Technical Report FM 5-0, U.S. Department of Army.

U.S. Department of Defense. 2008. Common warfighting symbology. Technical Report MIL-STD-2525C, U.S. Department of Defense.