Situational Awareness from Social Media

Brian Ulicny, Jakub Moskal
VIStology, Inc
Framingham, MA
{bulicny,jmoskal}@vistology.com

Mieczyslaw M. Kokar
Northeastern University
Boston, MA
m.kokar@neu.edu

***Abstract*—This paper describes VIStology's HADRian system for semantically integrating disparate information sources into a common operational picture (COP) for humanitarian assistance/disaster relief (HADR) operations. Here the system is applied to the task of determining where unexploded or additional bombs were being reported via Twitter in the hours immediately after the Boston Marathon bombing in April, 2013. We provide an evaluation of the results and discuss future directions.**

Keywords—social media, situational awareness, Boston Marathon bombing.

I. Introduction

The Homeland Security Act (2002) defines situational awareness as "information gathered from a variety of sources that, when communicated to emergency managers and decision makers, can form the basis for incident management decision-making" [1]. Incident commanders for humanitarian assistance/disaster relief (HA/DR) operations are better able to understand a situation and make appropriate decisions if they can view all of the relevant information in an integrated common operational picture (COP) in a way that allows them to make sense of the situation without being overwhelmed with information. However, HA/DR commanders should not be expected to know where all the relevant information is stored or how it is encoded. It would be better if a system would identify how to meet a commander's high-level information needs on the basis of previously annotated information stores that could be brought to bear in an emergency. In such dynamic situations, it would be desirable, too, if the system allowed an administrator to quickly annotate new information stores in order to make them answerable to the commander's needs and, secondly, provide enough annotation that the system knew how to query, transform, load and analyze data relevant to the commander's high level needs into the system.

In a large-scale emergency situation, such as the aftermath of the Boston Marathon bombings on April 15, 2013 [2], masses of people communicated information rapidly via social media and react to those messages, shaping the situation. Some were reporting what they were observing on the scene; others were not on the scene and merely commented or relayed information they received from elsewhere. While often dismissed as trivial, FEMA officials have testified that, "Social media is imperative to emergency management because the public uses these communication tools regularly…. With one click of the mouse, or one swipe on their smartphone's screen, a message is capable of being spread to thousands of people and have a tangible impact" [3].

In order for a commander to understand the situation and respond effectively, the commander must therefore have access to what people are saying on social media, and this must be presented in such a way that the commander can respond to it effectively. However, neither the commander, nor his or her staff, has time to read all of those messages and identify what is relevant in order to assess the situation. Semantic machine processing of the messages must provide the necessary insight into the relevance of particular messages and summarize their significance to the commander's information needs in a way that enables decisions and actions.

VIStology's HADRian project, our internal name for an AFRL SBIR Phase II project titled "Fusion, Management, and Visualization Tools for Predictive Battlespace Awareness and Decision Making", is focused on being able to quickly integrate disparate data sources into a COP by semantically annotating datastores using an ontology against which commander queries can be issued to determine relevant repositories, formulate the proper query to issue to the repositories, extract results, reason with the query results, filter them and display them. This project extends previous data virtualization work at VIStology sponsored by the Office of Naval Research for representing and reasoning about maritime track repositories annotated with an ontology; the current project, sponsored by AFRL, includes entities of a variety of types for use in HA/DR situations. In this paper, we examine the application of this technology to deriving situational awareness from social media.

II. HADRian Background and Concept of Operations

In the first phase of this project, we developed techniques for dealing with a range of object types and a variety of data representation formats as well as a different type of interface (RESTful web services, GPS track servers, among others). A guiding principle in this project is that HA/DR commanders cannot dictate where relevant information is uploaded by users. Our goal is to make it usable wherever content creators upload it, as long as it is online. Thus, we need to develop techniques for accessing it in various ways. It turns out that RESTful Web Services are very common for retrieving information produced by 'ad hoc sensor networks' and so we have focused on these. A proof-of-concept demo we developed reflects the retrieval and integration of information from disparate repositories into a single COP that are relevant to a scenario in which a plane crashes into a chemical factory. This scenario was drilled at Calamityville, a HA/DR training facility associated with the National Center for Medical Readiness at Wright State University on May 11, 2011. We used artifacts produced during this drill that exist in various repositories on the Web to illustrate our capabilities. We annotated the repositories that included them but do not modify the artifacts prior to incorporating them.

The Concept of Operations for our system is as follows:

1. A COP Administrator who manages the system **annotates repositories**, using an ontology, i.e. a formal representation of the conceptual domain.
2. The COP Administrator formulates High Level Query to **describe information needs** for current operation
3. The System **infers repositories** that may contain **relevant information** by **reasoning** over metadata that the repository has been **annotated** with.
    a. Information remains in place until it is needed. It is not initially all extracted, transformed and loaded (ETL).
    b. Users upload data wherever they usually upload it, not to a central repository.
4. The System issues appropriate **low level queries** to **repositories**
5. The System **filters out some** irrelevant data
6. The System **aggregates and displays data** in a COP
7. Users including the EOC (Emergency Operations Center) or Incident Commander and other operations center **interact with the data in the COP**.
8. The COP operator **pushes** elements of the displayed **information to users in the field** via their smartphone as needed.

In order to produce this demo, we developed:

1. Domain ontologies for representing repositories and queries, incorporating other ontologies as needed, such as UCore-SL [4] and a Distributed Interactive Simulation (DIS) Protocol Data Units (PDU) simulation data (for tracks) [11], to represent the conceptual and technical domain.

2. BaseVISor inference engine rules for reasoning about relevant repositories and rewriting query URLs in order to retrieve information elements from RESTful web interfaces and PDU sources that are relevant to this scenario. BaseVISor is VIStology's OWL 2 RL forward-chaining inference engine.

3. A novel technique for producing OWL representations of individual data items from the JSON output by RESTful web services. This allows us to generate OWL for reasoning without developing any custom software, on the basis of metadata and annotations alone.

4. Technology for integrating a variety of information types into the COP. We developed tools for integrating text, video, photos, and map overlays into a common COP based on Google Earth. We integrated Google Sketchup 3D facility models into the demo, and as well as GPS tracks, encoded as Distributed Interactive Simulation Protocol Data Unit binary data, as well as social media video, photos, and tweets in Phase I.

III. JIFX 13-4 Field Experiment

VIStology, Inc, recently conducted a field trial of its HADRian semantic information integration technology for Humanitarian Assistance/Disaster Relief operations at an invitation-only event sponsored by the Naval Postgraduate School held August 5-8, 2013, at McMillan Airfield, Camp Roberts, near Paso Robles, CA.

In the scenario that we pursued there, a commander needs to determine, on the basis of social media messages (here, only Twitter posts), where additional or unexploded bombs are being reported to be located (truly or falsely) in the aftermath of the Boston Marathon bombing in order to evaluate where to dispatch resources. In the immediate aftermath of the Marathon bombings, several locations were reported to have additional, unexploded bombs, all mistakenly as it turned out. Of course, it was not obvious at the time that the reports were false, and it was incumbent on public officials to maintain order and control at those sites if in fact they did contain a threat to public safety.

Our objective is to evaluate the feasibility of deriving situational awareness from a representative corpus of social media messages gathered immediately after the Boston Marathon bombing. The corpus consists of approximately 0.5 million messages that span the three hours following the bombing. In this experiment, information from social media users (here, Twitter users) was analyzed for answers to the high level query "Where are people reporting that additional or unexploded bombs have been found?"[1] Answers to this question were identified and presented in the COP in an appropriate way. The information included represented the following:

Where are additional/unexploded bombs being reported to exist?;

When were those messages propagated?;

How often have these messages been propagated (i.e. the amount of attention being directed to each location)?;

We were not able yet to represent, a future goal, answers to:

[1] This scenario was suggested to us by Desi Matel-Anderson, FEMA Innovation Advisor and Think Tank Strategic Vision Coordinator, at RELIEF 13-3.

How reliable and credible are the reports of a bomb at that location.

IV. System Description

The HADRian system can be thought of as having four functionalities that are relevant to this scenario:

A. Query Formulation and Repository Annotation
B. Relevance Reasoning and Repository Querying
C. Results Reasoning
D. Interactive Display

A. Query Formulation and Repository Annotation.

High level information needs are represented in our system ontology as instances of an OWL class called High Level Query (HLQ). In our system, an HLQ is not a query string in any particular query language, such as SQL or SPARQL. Rather, it is a description of one or more such queries, represented in OWL. That is, it should be possible to derive the OWL description of a query string by parsing and analyzing the query. We have made some attempts at translating SPARQL queries and even natural language queries into their OWL descriptions, automatically. However, at present, we rely on manually encoding HLQs in OWL directly.

A High Level Query is assigned various 'scopes' in the ontology: a Region Scope, a Time Scope, a Topic Scope, a Thing Scope and a Source Scope. Some of these scopes are related via annotation properties to classes or individuals in the ontology (in the case of Thing and Topic Scopes). An HLQ is related via an object property to individuals in the case of Time and Region Scopes. An HLQ essentially corresponds to an instance of a query of the form:

Find all instances of class T produced by instances of class S that are about instances of class U that existed in region R during temporal period P

Here, class T corresponds to the Thing Scope of the HLQ. A Thing Scope relates a query to the kind of thing that constitutes an answer to the query. For example, in English, "who" queries seek a Person or subclass of Person as an answer (e.g. Q: "Who can sign my timecard?" A: "Bill", "a manager"). A Topic Scope specifies what the specified 'things' from the Thing Scope are about: e.g. magazines about *Sports*. In the query template above, R corresponds to the Region Scope, which is an individual region in the ontology. P corresponds to the Time Scope, which is an individual temporal range in the ontology. The Source Scope S indicates that all of the things that satisfy the query must have been produced by an individual of class S or a subclass of S. The classes that are represented may be expressed with arbitrarily complex OWL class expressions.

Repositories are also a class in our ontology. Every repository also has a Thing, Topic, Region, Time and Source Scope. Thus, for example, a repository of tweets about traffic accidents in Paso Robles, CA, during 2012 from the Paso Robles (CA) Police Department would have the following scopes:

Thing Scope: StatusUpdate
Topic Scope: TrafficAccident
Region Scope: Paso Robles, CA
Time Scope: 2012
SourceScope: Paso Robles Police Department

HLQs and Repository Annotations are represented in an OWL ontology that incorporates the UCore-SL ontology [4] and aspects of the Dublin Core [5] and Geonames ontologies [6].

Any ontology editor can be used to annotate repositories and formulate queries. We currently use Protégé 4.x for this purpose, but any other OWL editor would do.

B. Relevance Reasoning and Repository Querying

Relevance Reasoning, in our system, is the process of identifying which repositories are relevant to a High Level Query based on its OWL annotations [8]. In HADRian, we do not examine the contents of the repository in identifying a relevant repository. The system only considers the metadata that has been assigned to it.

A Repository is inferred to be relevant to a HLQ if (but not only if) its scopes overlap with the Thing, Topic, Region and Time scopes of the HLQ. If a scope is specified in terms of a class, then a subclass or superclass overlaps with it. Regional and temporal overlaps are defined in the obvious way. A Topic Scope defined in terms of an individual coincides with any coreferential term.

A Repository, in our system, is a collection of items that could be represented in the COP. Repositories are a collection of items, and as such, they may be defined *extensionally* as pre-specified collection of things or *intentionally* as items that satisfy certain criteria, expressed as a query to a larger repository. For example, a collection of photos in some individual user's Flickr online photo album (flickr.com) represents a collection defined extensionally: the collection was defined by the user's selection of photos for that album. A Flickr query for photos taken in Yosemite Park on a particular date, however, is a repository that is determined *intensionally*. The set of photos that meet this criterion is not necessarily known in advance.

Each Repository must have a URL associated with it that enables the system to retrieve (extensional) or query (intensional) the data. Many of the repositories we deal with have RESTful interfaces. A query-defined repository for a RESTful interface may have parameters that are specified at run time based on the High Level Query. For example, a query for businesses listed in Yelp (yelp.com) may have a parameter for a zipcode that is filled at runtime by the zipcode corresponding to the area(s) that is (are) in the Region Scope of the HLQ.

For the Boston Marathon scenario, the HLQ has obvious Region (Boston, MA) and Time (April 15, 2013) scopes, but the Thing and Topic Scopes are not as obvious. The Thing Scope of the HLQ is defined as the class GeoFeaturesMentionedInStatusUpdates. This class is defined as a subclass of the intersection of the classes GeographicFeature (a UCore-SL class defined as "A PhysicalEntity whose (relatively) stable location in some GeospatialRegion can be described by location-specific data.") and the class of things are the subject of the mentionedIn object property with respect to some StatusUpdate. The class StatusUpdate is equivalent to the sioc:Post class, defined as "An article or message that can be posted to a Forum" [2].

The repository of tweets in this scenario thus has the Thing Scope StatusUpdate, but the HLQ has a Thing Scope of GeoFeaturesMentionedInStatusUpdates, which is neither a super- nor subclass of StatusUpdate. Therefore, it is not within the Thing Scope of the HLQ. A relevance reasoning rule, specified in BaseVISor rule language, states that if an HLQ has a Thing Scope that is a subclass of things mentionedIn some class C and a repository has a Thing Scope that is a subclass of C, then the repository is relevant to the HLQ.

BaseVISor is VIStology's customizable, forward-chaining OWL 2 RL inference engine. BaseVISor (vistology.com/basevisor) provides inference rules for the OWL 2 RL language profile, but it can be extended with custom rules. These rules may be augmented with user-supplied procedural attachments that perform custom functions in addition to default functionality for mathematical functions, string operations and the like [7].

In this case, the repository of tweets is pre-existent. Therefore, it is extensionally defined and does not require any run-time instantiation of lower level query parameters. We simply extract the contents of the repository and convert them to OWL, in order to do results reasoning.

The Topic Scope of the HLQ and the Repository both consist of the individual BostonMarathon2013 and the class UnexplodedBombs. Not every tweet in the repository is about UnexplodedBombs, although they are all presumed to be about the 2013 Boston Marathon. The class UnexplodedBombs is associated with a regular expression in the ontology that allows us to filter the query contents to only those tweets that are about both subjects.

C. Results Reasoning

After the relevant tweets are converted to OWL using a template that is part of the metadata annotation of the repository, BaseVISor is again used to reason about the results, in order to extract the required elements. Here a set of custom BaseVISor rules is used to identify locations mentioned in tweets about both unexploded bombs and the 2013 Boston Marathon. These rules produce a set of phrases that refer to locations. These location phrases are then mapped to known locations using a heuristic algorithm that chooses among the results of querying the Google Places and Google Maps Geocoding APIs, using the location phrase and a geographic region corresponding to Boston as the parameters of the search. This process associates locatable phrases with known locations and removes some phrases that are syntactically plausible but for which no identifiable location can be associated. For example, one of the extracted location phrases is 'BPD Commissioner Ed Davis', based on its context. This phrase corresponds to no known place by querying the Google APIs, so it is dropped from the output. Location phrases that do result in known places are collated. Several extracted phrases may coincide with the same known place, according to one or more of the Google APIs. A count of the number of tweets that are associated with each known place is kept. Various metadata elements associated with the known place are inserted into the KML document that is displayed as the result of the query.

[2] Semantically-Interlinked Online Communities (sioc-project.org)

D. Interactive Display

Finally, the KML is displayed in the COP as an answer to the High Level Query. Each placemark is labeled with one of the location phrases that produced it. A number in parentheses next to the placemark's title indicates the number of tweets that mentioned one of the location phrases mapping to this location. We emphasize this fact by rendering polygons underneath the placemarks that also correspond to the location volume in tweets: the higher and darker the color, the more frequently mentioned was the location. Clicking on the placemark reveals the phrases that produced the placemark, the type of place (according to Google), and the API source (Figure 1).

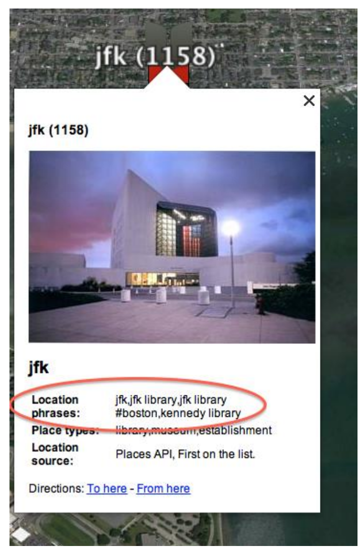

Figure 1 Expanded placemark shows location phrases that resulted in the placemark, number of tweets (1158), the type of place (library, museum) and the API source.

Each placemark can be removed from the COP by unchecking a widget in the list of placemarks on the left hand side of the COP (Figure 2). This set of placemarks can be viewed alongside other layers in Google Earth, such as baselayers presenting a photographic map of the various structures in the region as well as street names and other geographic features and attributes.

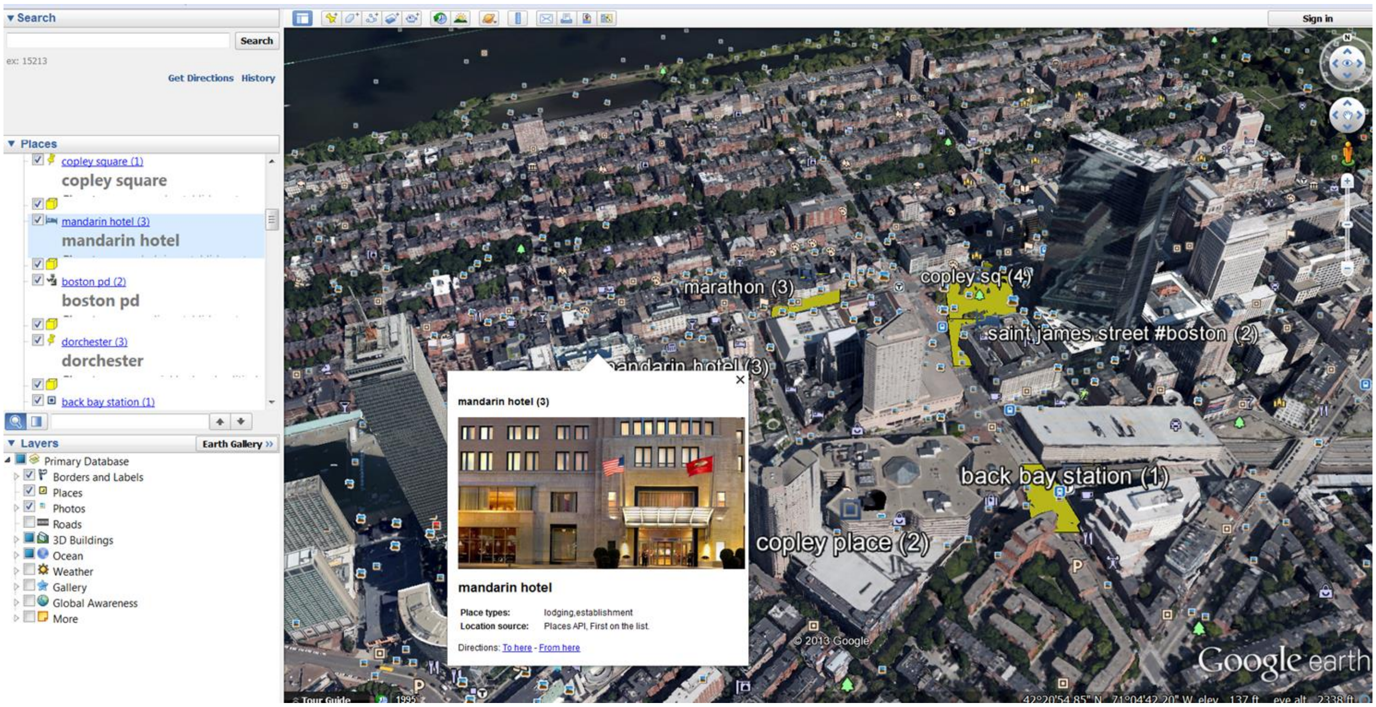

Figure 2 COP Indicating that three tweets about unexploded bombs mention the Mandarin Hotel, four mention Copley Square, one Back Bay Station and so on.

V. EVALUATION

In this exercise, we annotated a repository containing 509,795 twitter messages containing the hashtag #bostonmarathon between 4:06 PM and 7:04 PM on April 15, 2013, retrieved using Twitter APIs. The bombs are said to have exploded at 2:49 PM that day. The corpus was collected by Andrew Bauer and his colleagues at Syracuse University's School of Information Studies's NEXIS lab and made available on the Web as a CSV file.[3] The file contains the tweet ID number, text, creation time, associated latitude/longitude (if there is one) and user ID.

The latitude and longitude in the file represents the location of where the user sends the tweet from, not necessarily the location about which the user is reporting. Only 8,300 of the tweets had geocoded origins, or about 1.6% of the corpus. Generally, less than 1% of twitter users have enabled geotagging their locations using the location services on their smartphones or other devices [9][10]. In disaster relief datasets that we have examined, geotagged tweets approach 2% of the corpus. We were not concerned with the source location of tweets, but locations that were mentioned in the tweets, so we ignored these fields even when they were non-null. The repository was annotated in our ontology as described above.

We evaluated our processing by evaluating: the recall and precision of identifying tweets that mentioned unexploded bombs and the like; the recall and precision of identifying phrases specifying a location in the tweets; and the precision of associating a location phrase with a known place, using the Google APIs mentioned previously.

Precision in automatically identifying instances of a category is the ratio of true, positive identifications to positive identifications. Recall is the ratio of true, positive identifications to positive instances in the corpus as a whole. Finally, the F1-measure characterizes the accuracy of a categorization task as a whole by combining the recall and precision into a single metric, weighing each equally:

$$F_1 = 2 \cdot \frac{\mathrm{precision} \cdot \mathrm{recall}}{\mathrm{precision} + \mathrm{recall}}$$

To begin with, we did not evaluate the precision and recall of categorizing the corpus with respect to the topic of the Boston Marathon. We assume that all of the tweets in the corpus were about the 2013 Boston Marathon because of the time period in which they were sent in temporal proximity to the bombings. It is possible that some of the tweets in the corpus contain the hashtag #bostonmarathon but are in some sense not about the 2013 Boston Marathon. We have no way to evaluate the recall of this corpus. That is, we have no way to evaluate how many tweets were sent that were about the 2013 Boston Marathon but that did not contain this hashtag and were not collected in this corpus.

Of the tweets in this corpus, we identified 7,748 tweets that were about additional or unexploded bombs with a precision of 94.5%, based on a random sample of 200 tweets identified as such. That is, only 1.5% of the original corpus was identified as referring to additional bombs, using our pattern matching.

[3]https://www.dropbox.com/s/h8wezi2y6pzqfh4/041513_1606-1704_tweets.zip

Based on a random sample of 236 tweets from the original corpus, our recall (identification of tweets that discussed additional bombs) was determined to be 50%. That is, there were many more ways to refer to additional bombs than our rules considered. Thus, our F1 measure for accurately identifying tweets about additional bombs was 65%. Nevertheless, because of the volume of tweets, this did not affect the results appreciably.

Having thus reduced the corpus 98.5% in this way to only tweets that discussed unexploded bombs in addition to referring to the 2013 Boston Marathon, we now evaluate the precision and accuracy of identify location phrases. Location phrases were identified purely by means of generic pattern matching. We did not use any list of known places. Nor did we include any scenario-specific patterns. The precision with which we identified location phrases was 95%. That is, in 95% of the cases, when we identified a phrase as a location phrase, it actually did refer to a location in that context. Mistakes included temporal references and references to online sites. Our recall was only 51.3% if we counted uses of #BostonMarathon that were locative. (We mishandled hashtags with camel case.) Alternatively, since all of the tweets contained some variant of the hashtag #bostonmarathon, this is a somewhat uninformative location phrase. If we ignore this hashtag, then our recall was 79.2%. That is, of all the locations mentioned in tweets about additional bombs at the Boston Marathon, we identified 79.2% percent of the locations that were mentioned. Using the more lenient standard, our F1 measure for identifying location phrases in the text was 86.3%.

Our precision in associating tweets with known places via the Google APIs was 97.2%. Our precision in assigning unique location phrases to known places via Google APIs was 50%. That is, there were many location phrases that were repeated several times that we assigned correctly to a known place, but half of the unique phrase names that we extracted were not assigned correctly. Ten location phrases that were extracted corresponded to no known locations identified via the Google APIs. These included location phrases such as “#jfklibrary” and “BPD Commissioner Ed Davis”. The former is a phrase we would like to geolocate, but lowercase hashtags which concatenate several words are challenging. The latter is the sort of phrase that we expect would be rejected as non-geolocatable. See Table 1.

Table 1 Top 20 Identified Places with Number of Tweets

Known Place	**#Tweets**
JFK Library	1158
Boston	629
Boston Marathon	325
St Ignatius Catholic Church	47
PD	29
Boylston	8
CNN	5
Copley Sq	4
Huntington Ave	4
Iraq	3
Mandarin Hotel	3
Dorchester	3
Marathon	3
US Intelligence	3
Copley Place	2
Boston PD	2
BBC	2
Cambridge	2
John	2
St James Street #Boston	2

More qualitatively, the Twitter processing we described here resulted in 38 ranked places on the COP that were associated with additional or unexploded bombs. We compared these places with the places that were mentioned in the live blogs that were set up by CNN[4], the New York Times[5] and the Boston Globe[6] immediately following the bombings. These blog sites mentioned the following locations (only once, each)

Location [Source]: (# of Tweets Identified with That Location)

Boylston Street [Globe, CNN]: 8
Commonwealth Ave near Centre Street, Newton [Globe]: 0

Commonwealth Ave (Boston) [Globe]: 0

Copley Square [NYT]: 4

Harvard MBTA station [Globe]: 0

JFK Library [CNN, Globe, NYT]: 1158

Mass. General Hospital [Globe, NYT]: 0
(glass footbridge over) Huntington Ave near Copley place [Globe]: 4

Tufts New England Medical Center [NYT]: 0
Washington Square, Brookline [NYT]: 0

For three of these sites – Mass. General Hospital, Tufts Medical Center and Washington Square, Brookline, reports of unexploded bombs or suspicious packages occurred after the end of the tweet collection period, at 7:06 PM. Otherwise, the recall of our system was good, missing only the report of unexploded bombs at the Harvard MBTA station. A few tweets mentioning such a threat were in our corpus, but the

[4] http://news.blogs.cnn.com/2013/04/15/explosions-near-finish-of-boston-marathon/comment-page-18/
[5] http://thelede.blogs.nytimes.com/2013/04/15/live-updates-explosion-at-boston-marathon/
[6] http://live.boston.com/Event/Live_blog_Explosion_in_Copley_Square?Page=16

system failed to pick them up, either due to capitalization issues or unexpected use of hashtags.

Additionally, on average, tweets reflecting these locations were produced 11 minutes prior to their being reported on the sites mentioned. Thus, the tweet processing was more timely and more comprehensive than simply relying on a handful of news sites alone for situational awareness

I. Conclusion

In this paper, we described a system for integrating disparate information sources into a COP for Humanitarian Assistance/Disaster Relief operations by means of semantic annotations and queries, using a common ontology. We described the operation of the system and evaluated the results of an experiment in annotating and querying social media data streams in order to produce situational awareness. We applied our technology to a repository of tweets collected in the immediate aftermath of the Boston Marathon bombings in April, 2013, and demonstrated that a ranked set of places could be incorporated into the COP, showing the prominence of each site by tweet volume that was reported as being the site of an additional unexploded bomb or bombs. We evaluated the results formally and compared the results with the situational awareness that could be gleaned only from mainstream media blogs being updated at the same time. On average, the automatic processing would have had access to locations from tweets eleven minutes before these sites were mentioned on the mainstream media blogs. Additionally, sites that were prominent on Twitter (e.g. St Ignatius Church at Boston College or the Mandarin Oriental Hotel in Boston) were not mentioned on the news blog sites at all. We believe that these results show that this approach is a promising one for deriving situational awareness from social media going forward.

Acknowledgment

This work was performed under AFRL contract FA8650-13-C-6381 "Fusion, Management, and Visualization Tools for Predictive Battlespace Awareness and Decision Making". Thanks also to the JIFX 13-4 participants for helpful feedback.

References

[1] United States Code, 2010 Edition Title 6 - DOMESTIC SECURITY CHAPTER 1 - HOMELAND SECURITY ORGANIZATION SUBCHAPTER V - NATIONAL EMERGENCY MANAGEMENT Sec. 321d - National Operations Center 6 U.S.C. §321d(a)

[2] FEMA. Lessons Learned - Boston Marathon Bombings: The Positive Effects of Planning and Preparation on Response. August 2, 2013.

[3] Shayne Adamski. Written testimony of FEMA for a House Homeland Security Subcommittee on Emergency Preparedness, Response, and Communications hearing titled "Emergency MGMT 2.0: How #SocialMedia & New Tech are Transforming Preparedness, Response, & Recovery …" July 9, 2013

[4] Barry Smith, Lowell Vizenor and James Schoening, "Universal Core Semantic Layer", Ontology for the Intelligence Community, Proceedings of the Third OIC Conference, George Mason University, Fairfax, VA, October 2009, CEUR Workshop Proceedings, vol. 555.

[5] Dublin Core Metadata Initiative, http://dublincore.org

[6] GeoNames Ontology, http://www.geonames.org/ontology/

[7] C. Matheus, B. Dionne, D. Parent, K. Baclawski and M. Kokar. BaseVISor: A Forward-Chaining Inference Engine Optimized for RDF/OWL Triples. In Digital Proceedings of the 5th International Semantic Web Conference, ISWC 2006, Athens, GA, Nov. 2006.

[8] M. Kokar, B. Ulicny, and J. Moskal. Ontological structures for higher levels of distributed fusion in Distributed Data Fusion for Network-Centric Operations, D. Hall, C.-Y. Chong, J. Llinas, and M. Liggins II, Eds., ed: CRC Press, 2012, pp. 329-347.

[9] Cheng, Z.; Caverlee, J.; and Lee, K. 2010. You are where you tweet: A content-based approach to geo-locating twitter users. In CIKM '10.

[10] LEETARU, Kalev et al. Mapping the global Twitter heartbeat: The geography of Twitter. First Monday, [S.l.], apr. 2013. ISSN 13960466. Available at: <http://firstmonday.org/ojs/index.php/fm/article/view/4366>. Date accessed: 04 Sep. 2013. doi:10.5210/fm.v18i5.4366.

[11] McGregor, D., Brutzman, D., Armold, A., and Blais, C. L., "DIS-XML: Moving DIS to Open Data Exchange Standards," Paper 06S-SIW-132, Simulation Interoperability Standards Organization, 2006 Spring Simulation Interoperability Workshop, Huntsville, AL, April 2006