DESIGN OF A TEMPORAL GEOSOCIAL SEMANTIC WEB FOR EMERGENCY MANAGEMENT OPERATIONS

Bhavani Thuraisingham, Latifur Khan, Murat Kantarcioglu
The University of Texas at Dallas

POSITION PAPER

1. Introduction

Since 9/11 and Katrina in August 2005, the United States federal, state and local governments have worked well together to share information for emergency management operations for natural disasters and for terrorism. Furthermore, the South East Asia Tsunami of December 2004 has also improved information sharing between inter-governmental agencies and the NGOs. However there is still a lot to be done before we can have successful emergency management operations. We need to migrate from the need to know paradigm to the need to share paradigm. In fact this was one of the recommendations of the 9/11 commission report. This is true not only to combat terrorism, but also for emergency operations as well as for military and non military reconstruction and stabilization operations. Such operations are needed also after events like 9/11, Katrina and the South East Asian Tsunami as well as the wars in Iraq and Afghanistan. Studies have shown that security, power and jobs are ingredients for success during reconstruction and stabilization. It is important to give power to some of the players during emergency management and some other players during reconstruction. It is critical that investments are made to stimulate the local economies. The studies have also analyzed the various technologies that are needed for successfully carrying out emergency management operations which includes sensors, robotics and information management. We are focusing on the information management component for Emergency Management. As stated in the work by the Naval Postgraduate School, we need to determine the social, political and economic relationships between the stakeholders as well as understand the interactions between the governor's office, the law enforcement and homeland security communities. This work has also identified the 5Ws (*Who, When, What, Where* and *Why*) and the H (*How*).

To address the key technical challenges for emergency management, we are developing a system that can work for Emergency Management Operations as well as for Stabilization and Reconstruction Operations. This system is based on a Temporal Geosocial Service Oriented Architecture System (TGS-SOA) that utilizes Temporal Geosocial Semantic Web (TGS-SW) technologies. The data that has to be managed for emergency operations as well as the relationships between the data have to be represented using semantic web technologies. Reasoning techniques are used to extract the social and geospatial relationships between the individuals. Various ontologies are aligned and these networks are then analyzed to extract the nuggets to help the emergency management officials. We are developing techniques for representing temporal geosocial information and relationships, integrating such information and relationships, querying such information and relationships and finally reasoning about such information and relationships so that the governor and other indicuiduals of power can answer questions related to the 5Ws and H. Furthermore, the nuggets to be extracted to support the analysts and the officials have to satisfy the 5 Rights. That is, we need the *right* nuggets at the *right* place at the right time to the *right* people to make the *right* decisions. Therefore real-time operations are also critical for emergency management.

2. Our Approach

Architecture: Our system is based on services and we are designing and prototyping technologies for a Temporal Geosocial Semantic Web (TGS-SW). A system based on TGS-SW will capture the semantics of information and relationships. Therefore we are developing semantic web technologies for representing, managing and deducing temporal geosocial relationships. Our current work in geospatial information management and social networks is exploring the use of FOAF, GRDF (Geospatial Resource

Description Framework) and SNRDF (Social Networks RDF). Therefore we will incorporate the temporal element into these representations and subsequently develop appropriate representation schemes based on RDF and OWL. Our system will first integrate the information from various data source (both geospatial and text as well as structured data). This is the geospatial information integration component. Then the integrated information will be analyzed to extract the nodes and links in a social network. Finally the social network relationships will be integrated with the geospatial relationships to build a temporal geososocial information system. The information managed by such a system will be analyzed to extract the nuggets to support the emergency preparedness officer. We will describe the major components of the system.

Geospatial Information Management: Geospatial data could be extracted for numerous data sources including web pages, google maps and blogs. We have built an information integration system that integrated the geospatial data from various data sources (both geospatial and otherwise). We are examining *scalable* integration techniques for handling heterogeneous geosocial data utilizing our techniques on ontology matching and aligning. Moreover, to ensure data/information from multiple heterogeneous sources can be integrated smoothly, we are exploring data/information conversion and transformation rules, identify redundancy and inconsistence. In our recent research on geospatial information integration, we have developed knowledge discovery methods that resolve semantic disparities among distinct ontologies by considering instance alignment techniques. Each ontological concept is associated with a set of instances, and using these, one concept from each ontology is compared for similarity. We examine the instance values of each concept and apply a widely popular matching strategy utilizing N-grams present in the data. However, this method often fails because it relies on shared syntactical data to determine semantic similarity. Our approach resolves these issues by leveraging K-medoid clustering and a semantic distance measure applied to distinct keywords gleaned from the instances, resulting in distinct semantic clusters. We claim that our algorithm outperforms N-gram matching over large classes of data. We have justified this with a series of experimental results that demonstrate the efficacy of our algorithm on highly variable data. We are exploiting not only instance matching but also name and structural matching in our project for geosocial data that evolves over time.

Temporal Social Network Relationships: Temporal social networks model, represent and reason about social networks that evolve over time. Note that in countries like Iraq and Afghanistan the social and political relationships may be continually changing due to security, power and jobs. Therefore it is important to capture the evolution of the relationships. In the first phase, our design and development of temporal social networks is focusing on two major topics, namely, semantic modeling of temporal social networks and fundamental social network analysis. For semantic modeling of temporal social networks, we are extending the existing semantic web social networking technologies such as Friend-of-a-Friend (FOAF) ontology to include various important aspects such as relationship history that is not represented in current social network ontologies. For example, we are including features to model the strength and trust of relationships among individuals based on the frequency and the context of the relationship. In addition, we are including features to model relationship history (e.g., when the relationship has started and how the relationship has evolved over time) and the relationship roles (e.g., in the leader/follower relationship, there is one individual playing the role of the leader and one individual playing to role of the follower). In essence, by this modeling we intend to create an advance version of social network such as Face book specifically designed for SARO objectives. Note that there are XML based languages for representing social network data (SNML) and temporal data (TML). However we need a semantic language based on RDF and OWL for representing semantic relationships between the various individuals. We are exploring RDF for repenting social relationships (SNRDF and extended FOAF). Representing temporal relationships (When) is an area that needs investigation for RDF and OWL based social networks.

We are using the social network analysis to identify important properties about the underlying network to address some of the *5W (who, what, when, where and why) and 1H (how) queries*. To address queries for determining who to communicate with, we are using various centrality measures such as degree centrality and betweeness centrality to measure the importance of a certain individuals in a given social network. Based on such measures that are developed for social network analysis, we can test which of the centrality measures could be more appropriate in finding influential individuals in social networks. Especially, we plan to test these measures on available social networks such as Facebook (it is possible to download information about individuals in your own network on Facebook). For example, if a centrality measure is a good indicator, than we may expect individuals with high centrality value to have more posts on their Walls on Facebook or tagged in more number of pictures.

To answer the queries for determining *what information is needed*, we are examining the use of relational naïve Bayes models to predict which attributes of the individual is a more reliable indicator for predicting their friendliness to the military. Since we do not have any open data on such military data, we are using Facebook data to predict attributes that are more important indicators for individual's political affiliation to test our relational naïve Bayes models. To address queries for determining *when to approach them*, we are using various domain knowledge rules to first address when *not* to approach them. For example, in Iraq, it may not be a good idea to approach Muslim individuals during Friday pray. Later on, we will try to build knowledge discovery models to predict the best times for approaching certain individuals based on their profile features (i.e., their religion, social affiliation and etc.). In addition, we plan to use similar knowledge discovery models to answer the queries for understanding how those individuals may interact with the military personnel. In order to test our knowledge discovery models, we are analyzing various e-mail logs of our group members to see whether our model could predict best times to send an e-mail to an individual to get the shortest response time and the positive answer. To address queries for determining *why certain individuals' support is vital*, we are examining community structure mining techniques especially cluster analysis to see which group a certain individual belongs to and how the homophily between individuals in the group affects the link structures. Again, Facebook data could be used to test some of these hypotheses.

Temporal Geosocial Semantic Web: We are integrating our work on modeling and reasoning about social network information with the geospatial information. First we are utilizing concepts from concepts from SNML (social network markup language), GML (geospatial markup language) and TML (temporal markup language). However to capture and semantics and make meaningful inferences we need something beyond syntactic representation. As we have stated, we have developed GRDF (Geospatial RDF) that basically integrates RDF and GML. One approach is to integrate SNRDF (Social network RDF that we are examining) with GRDF and also incorporate the temporal element. Another option is to make extensions to FOAF to represent temporal geosocial information. Next we are integrating the geosocial information across multiple sites so that the commanders and allied forces as well as partners in the local communities can form a common picture. We have developed tools for integrated social relationships as well as geospatial data using ontology matching and alignment. However, we are exploring ways of extending our tools to handle possibly heterogeneous geosocial data in databases across multiple sites.

We are also exploring appropriate query languages to query geosocial data. There are query languages such as SPARQL and RQL being developed for RDF databases. We have adapted SPARQL to query geospatial databases. We are also exploring query languages for social network data. Query languages for temporal databases have been developed. Therefore our challenge in this subtask is to determine the constructs that are needed to extend languages like SPARQL to query geosocial data across multiple sites. It is also crucial to reason about the information and relationships to extract useful nuggets. Here we are developing ontologies for temporal geospatial information. We have developed ontologies and reasoning tools for geospatial data and social network relationships based on OWL and OWL-S. We are incorporating social relationships and temporal data and reason about the data to uncover new information. For example, if some event occurs at a particular time at a particular location for 2 consecutive days in a row and involves the same group of people, then it will like occur on the third day at the same time and same location involving the same group of people. We can go on to explore why such an event has occurred. Perhaps this group of people belong to a cult and have to carry out activities in such a manner. Therefore we are developing reasoning tools for geosocial data and relationships.

3. Summary and Directions

This paper has described the challenges for developing a system based on temporal geosocial semantic web for emergency management as well as for military stabilization and reconstruction operations. A system based on geosocial semantic web technologies can be utilized by emergency preparedness officials, military personnel, decision makers, and local/government personnel during/after an emergency or combat operation. There are several areas that need to be included in our research. One is security and privacy. We need to develop approaches for enforcing appropriate policies for emergency management. These policies may include confidentiality policies, privacy policies and trust policies. For example, certain geospatial as well as social relationships may be visible to certain parties. Furthermore, the privacy of the individuals involved have to be protected, Different parties may place different levels of trust on each other. We believe that building a system based on geosocial techologies is a challenge. Furthermore, incorporating security will make it even more complex. Nevertheless security has to be considered at the beginning of the design and not as an afterthought. Our future work will also include handling dynamic situations where some parties may be trustworthy at one time and may be less trustworthy at another time.