

# Situational Awareness Technologies for Disaster Response

Naveen Ashish, Rajesh Hegde, Charles Huyck, Dmitri Kalashnikov, Sharad Mehrotra, Padhraic Smyth and Nalini Venkatasubramanian

## 1. Introduction

Responding to natural or man-made disasters, in a timely and effective manner, can reduce deaths and injuries, contain or prevent secondary disasters, and reduce the resulting economic losses and social disruption. During a crisis, responding organizations confront grave uncertainties in making critical decisions. They need to gather situational information (e.g., state of the civil, transportation and information infrastructures), together with information about available resources (e.g., medical facilities, rescue and law enforcement units). There is a strong correlation between the accuracy, timeliness, and reliability of the information available to the decision-makers, and the quality of their decisions. Dramatic improvements in the speed and accuracy at which information about the crisis flows through the disaster response networks has the potential to revolutionize crisis response, saving human lives and property.

This chapter highlights some of the key information technology challenges being addressed in the Project RESCUE<sup>1</sup> [1], with a particular focus

---

<sup>1</sup> The work reported in this paper is part of the RESCUE Project (<http://www.itr-rescue.org>) whose goal is to radically transform the ability of organizations to gather, manage, analyze and disseminate information when responding to man-made and natural catastrophes. The 'Responding to Crises and Unexpected Events' (RESCUE) Project is funded in part by National Science Foundation (NSF) through its large ITR program under contract numbers 0331707 and 0331690.

on *situational awareness* technologies. Appreciating the IT challenges in improving crisis response requires a thorough understanding of how communication and control networks form among responding organizations, and how the response process is organized. Since the crisis domain might be new to a large number of readers, we begin by first briefly summarizing the crisis response process. This is done to set the stage for a discussion of the challenges being addressed by RESCUE and particularly the situational awareness thrust within it.

### **The Crisis Response Process**

Organized crisis response activities include measures undertaken to protect life and property immediately before (for disasters where there is at least some warning period), during, and immediately after disaster impact. Such activities may span a few hours to days or even months, depending upon the magnitude of the event. Depending upon the scale of the disaster, crisis response may be a large-scale, multi-organizational operation involving many layers of government, public authorities (such as state-managed utility companies), commercial entities, volunteer organizations, media organizations, and the public. In a crisis, these entities work together as a virtual organization to save lives, preserve infrastructure and community resources, and reestablish normalcy within the community. Depending upon the magnitude of the crisis, the operation of this virtual organization can span multiple levels. Field level operations such as evacuation, traffic management, triage, and provision of medical services are usually under the control of an on-site incident commander that reports back to a central Emergency Operations Center (EOC). In a large disaster, the management of area-wide resources requires a broader participation of government and industry. In large urban areas such as Los Angeles and New York, it is not uncommon for each city within a county to have its own EOC where representatives from fire, police, utility companies, Red Cross, and many other organizations participate in the response. Furthermore, each agency represented in the City EOC also has its own emergency operations center, usually in another location. In addition to these government-run centers, private industry (large businesses, NGO's, etc.) may also set up response centers that feed and receive information from government EOCs. While small disasters may be handled at the local level, the resources of local governments can become overwhelmed by the demands of larger events; in these cases, higher levels of the government become active participants in the response effort. Such a large-scale response may involve hundreds of autonomous organizations with different tasks and priorities. For ex-

ample, a county-wide disaster in the Los Angeles Area may mobilize emergency offices of State, County, and up to 88 different municipal authorities, along with a variety of other organizations (including fire departments, health services agencies, and NGOs such as the Red Cross). Each of these organizations may themselves represent a large consortium; for example, the health services organizations may consist of a variety of hospitals, triaging services, clinics, etc.

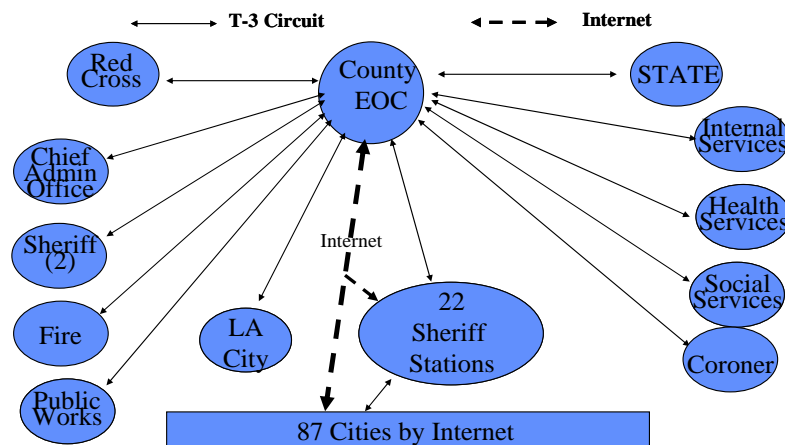
**The Response Cycle:** Irrespective of the nature and scale of the crisis and the organizations involved, crisis response activities can be viewed (at an abstract level) as consisting of four interrelated phases:

- *Damage assessment:* In this phase, disaster-related losses are identified on both incident-level and regional scales, and their magnitudes are assessed. Severely impacted areas, disruptions to critical infrastructure, situations where secondary hazards may develop if initial damage is not mitigated (e.g., earthquake-induced hazardous materials releases or dam failures), and other problems of high urgency are identified, and estimates of the time needed to restore disrupted systems are developed.
- *Needs assessment:* In this phase, incidents requiring some level of response are identified. For example, building collapses where victims are trapped may require search, rescue and medical resources, release of hazardous materials may require large-scale evacuation, etc. Operationally, these incidents are assigned a measure of urgency/priority, typically based on immediate threats to life safety.
- *Prioritization of Response Measures:* In this phase, incidents requiring response are matched with available resources. If the total demand is greater than the system's capacity to respond – as is invariably the case in large-scale disasters – decision-makers must establish priorities for response. Decision-makers must have an accurate assessment of the disaster situation and available resources in order to establish priorities.
- *Organizational Response:* In this phase, emergency resources are deployed and organizational decisions are disseminated to crisis-workers and the population at large. Ideally, response activities take place in accordance with pre-disaster planning. Decision-support systems are used to track key incidents and the progress of responding units, to optimize response activities, and to act as a mechanism for queuing on-going incidents.

Each aspect of the above process is part of an ongoing cycle, in which assessments, decisions, and interventions at one point in time produce implications for subsequent response activities. As the response proceeds and as more accurate information becomes available, new problems are identified, decisions are reassessed, and response activities may be reprioritized and sometimes even reversed. For this process to proceed effectively, government leaders, response personnel, and other actors must communicate rapidly with one another during each phase. The quality of the resulting decisions and the speed with which the process transitions through the four phases depends upon the timeliness and accuracy of information available to response workers.

### IT Challenges and RESCUE

In the RESCUE Project, our focus is to radically transform the speed and accuracy with which information flows through disaster response networks, networks that connect multitudes of response organizations as well



**Fig 1.** LA County Disaster Network

as the general public. We are working to develop information technology solutions that dynamically capture and store crisis-relevant data as it is generated, analyze this data in real-time, interpret it, and disseminate the resulting information to decision makers in the forms most appropriate for their various tasks. Challenges in realizing such IT solutions arise due to the scale and complexity of the problem domain, the diversity of data and data sources, the state of the communication and information infrastruc-

tures through which the information flows, and the diversity and dynamic nature of the responding organizations.

*Diversity of Information Sources:* Information relevant to decision making may be dispersed across a hierarchy of storage, communication, and processing units – from sensors (in-situ sensors, satellite imagery, remote sensing) where data is generated to heterogeneous databases belonging to autonomous organizations. Critical information may span various modalities – e.g., field-observations communicated via voice conversations among emergency workers, video data transmitted from cameras embedded in civil infrastructures, dispersed at crisis site, or carried by first responders, sensor data streams, or textual and relational information in databases. In some cases, information may even be embedded in the relationships among persons themselves; for instance, the migration and patterns of those fleeing an incident site may provide valuable clues as to the nature and exact location of the incident.

*Diversity of Information Users:* Information may need to be shared across diverse, loosely coupled, emergent multi-organizational networks in which different entities play different roles in response activities, have different needs and urgencies, have different cultures, and may have vastly different capabilities with respect to technology utilization. These organizations may or may not have policies in place regarding data sharing and collaboration. Furthermore, these organizational networks must rapidly reconfigure to adapt to the changing communication and control demands present during crisis events. Finally, different people/organizations have different needs and urgency levels regarding the same information. For example, while a field worker might require detailed information about the specific location of hazardous materials in a burning building, the monitoring and response team at a nearby command center may only need to know the number of hazardous-material locations in the vicinity of the catastrophe.

*State of the Infrastructure:* Information must be transferred across highly distributed, mobile infrastructure consisting of heterogeneous communication channels and systems that are prone to failures and vulnerable to attacks during a crisis. To address the above challenges in building an information infrastructure that provides reliable, timely, and integrated information, we are developing technology at two levels.

At the *systems* level our goal is to develop robust communications and systems infrastructure to ensure connectivity and availability of reliable

and efficient pipeline for information flow between different organizations and entities. For instance the *Robust Networking project* within RESCUE is focused on developing systems that provide computing, bi-directional communication, and higher layer services at a crisis site. The crisis site may lack electric power, fixed communication networks may be unpredictable, and responders might bring in heterogeneous mutually conflicting communication technologies. The goal is to develop a system that can operate under such extreme conditions by consolidating and enhancing available systems and seamlessly extending new capabilities to all end users and devices as communication services get incrementally restored. The *Information Dissemination project* (part of RESCUE) is exploring how customized warnings, alerts, advisories, and messages can be disseminated to the public at large under extreme situations when the infrastructure is failing and under surge demands. The approach developed exploits peer-based techniques to achieve fast, reliable, and scalable dissemination.

At the *informatics* level, the RESCUE project is exploring capabilities for information collection, sharing, fusion, and analysis. For instance, in the *situational awareness project* our objective is to design and develop technologies that can create actionable situational awareness from the avalanche of heterogeneous multi-modal data streams (audio, speech, text, video, etc.) including human-generated input (e.g., first responders' communications, field reports, etc.) during or after a disaster. In the *Information sharing* project we are developing a flexible, customizable, dynamic, robust, scalable, policy-driven architecture that ensures seamless information sharing amongst all entities involved in response with minimal manual human intervention and automated enforcement of information-sharing policies.

Finally, privacy concerns in infusing technology into real-world processes and activities arise for a variety of reasons, including unexpected usage and/or misuse for purposes for which the technology was not originally intended. These concerns are further exacerbated by the natural ability of modern information technology to record and make persistent information about entities (individuals, organizations, groups) and their interactions with technologies – information that can be exploited in the future against the interests of those entities. Such concerns, if unaddressed, constitute barriers to technology adoption or worse, result in adopted technology being misused to the detriment of the society. Another of our objectives thus is to understand privacy concerns in adopting technology from the social and cultural perspective, and design socio-technological solutions to alleviate such concerns. The challenge guiding our research is

whether Information Technologies can be designed with “knobs” that can be used to control disclosure of information amongst entities (individuals, organizations, government) with the objective of empowering technology adopters to fit the technology into existing (and possibly dynamically evolving) societal and cultural expectations with respect to privacy.

In this section we have introduced the RESCUE project and provided a brief overview of the various research thrusts centered on the IT challenges in improving disaster response. The following sections will focus on the situational awareness project, including our objectives, approach, research in various areas and also technology transition into artifacts.

## **2. Situational Awareness**

### **Response and Situational Awareness**

Each of the different sub-phases of response, as described in the introduction, as well as phases of a disaster cycle are decision-centric in which individuals and organizations make critical decisions that have implications on life and property. The quality of such decisions depends upon a variety of factors including social and cultural factors, level of expertise knowledge, and training, the established practice, etc. Among the most critical factors (of central interest to us in the situational awareness project) is the awareness of the situation (past, present, and future) and implications of actions or inactions. The situational information broadly consists of information about people (their vulnerabilities, location, demographics), resources (food, water, shelter) and progression of the event and activities (plume spread, storm track, evacuation progress). Our objective in the situational awareness project is to significantly enhance the situational awareness of decision-makers thereby improving their ability to make correct decisions. Specifically our goal is to develop advanced information processing and management technologies to turn the *available* information into *actionable* information for people and decision systems.

### **A Situational Awareness Driven Application**

As a motivating example, we present a real world application and process that we believe could benefit significantly (in terms of efficacy and

timeliness) from advanced situational awareness technologies. *Damage assessment* is an important task, immediately following a major disaster such as a hurricane or earthquake. An accurate and timely estimate of damage in the impacted and surrounding areas can help in several important decisions, such as resource allocation in response. The information, using which such assessments can be made, is collected in a variety of different ways: through use of aerial or satellite imagery; ground or field coverage; simulating the effects of the disaster through sophisticated loss or impact modeling tools – these tools are common for most hazards (e.g., hurricane, flooding, and earthquake); and finally use of information that becomes available on the internet – this may include blogs, internet surveys (e.g., the “did you feel it” system operated by the U.S. Geological Survey for earthquake events), reports by government or non-profit organizations.

The typical *modalities* for post-disaster information can include the following: aerial photos, field-based photos, satellite imagery (low, moderate and high-resolution optical data; radar or synthetic aperture radar data; lidar), video (with possibly audio commentaries), written reports, simulation results, etc. Data sources for much of this information come from government agencies (NOAA, USGS, FEMA, NASA, and local and state emergency response organizations); non-profit organizations, such as the Red Cross; industry (companies affiliated with or providing data to the insurance industry), and the public (e.g., through the internet)

All the above, raw, information today needs to be collected from the different sources above and then processed and integrated, mostly manually. Ultimately this information must be analyzed to convey the following kinds of information: a) areas of greatest impact, measured in terms on expected losses, lives lost, and social disruption; b) areas needing immediately response, either because of developing incidents, e.g., fire-following earthquake, or flooding after a storm; and c) areas that may be impacted by impending disasters, e.g., level breaks because of excessive storm surge.

The above problem is precisely of the nature that our situational awareness technologies aim to address i.e., if we can employ such technologies to make the information processing pipeline from raw disparate information to the end damage assessment analyses above, more automated, efficient, and timely. We now describe the technologies we are developing that can make such an assessment analysis application feasible.

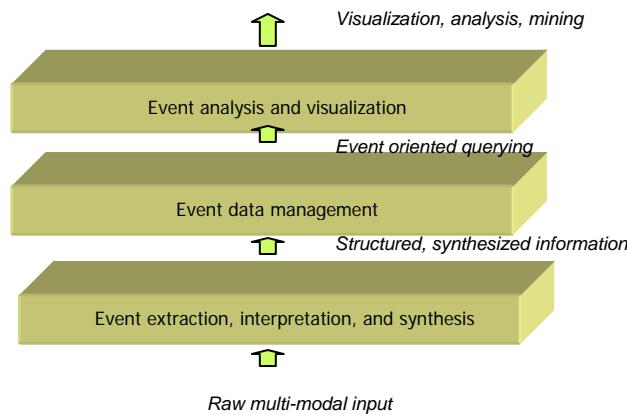
## **Our Approach**

Our approach to building situational awareness systems is an *event-centric* one i.e., one that is centered around events as fundamental units of

information. The motivation comes from our observation that situations (including disaster situations) are fundamentally comprised of events at various levels. For instance the out break of a fire, recovery of victims, dispatch of responders, collapse of a bridge are all events in a situation. It is information about such events that feeds into the analysis tools of the kinds described above. In very general terms an event is an occurrence of something of interest of a certain type, at a certain place and at/over a certain period of time. It is a semantic, *domain independent* concept which may have associated with it a set of different entities that play different roles and that may bear relationships to other events and/or to other entities in the real world. Events, in our view, provide a natural abstraction for modeling, representing, and reasoning about situations.

We have identified the following three key technical components in building situational awareness systems centered around events.

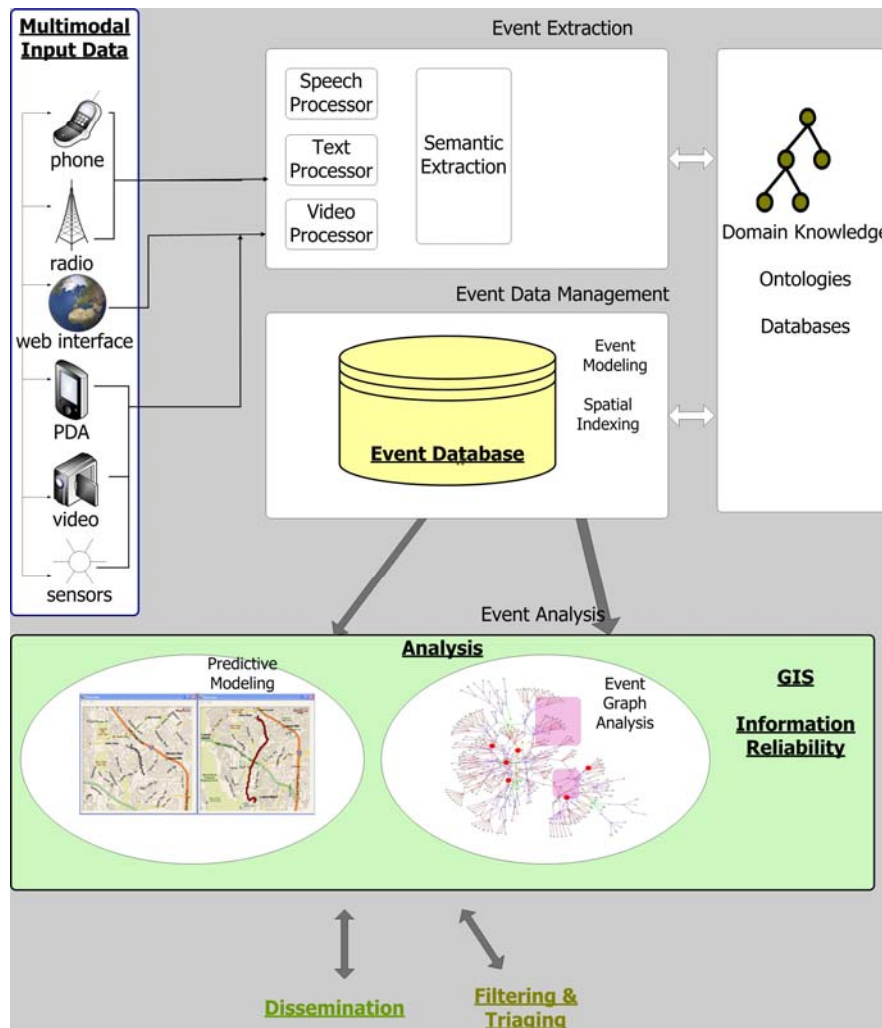
1. **Event Extraction, Interpretation, & Synthesis** - Multi-modal extraction, fusion and assimilation technology that enables the *extraction and synthesis* of higher level situational information from low level (signal) inputs. For instance in the damage assessment application above, we need automated capabilities to process the various pieces of information, such as images, audio and video feeds, text feeds, etc. and also integrated and fuse information coming from multiple disparate sources.



**Fig 2.** Situational Awareness System Tiers

2. **Event Data Management** – Technologies that support *modeling, storing, querying and indexing*, situational information. In

the damage assessment application for example, we need data management capabilities for storage and structured querying of the information collected, processed and integrated in the ingest phase. The information obtained about the unfolding situation and status should be queryable in a structured manner.



**Fig 3.** Situational Awareness System: Architecture

3. **Event Analysis and Visualization** – An environment that supports *end user analysis and visualization* tools over the situational information. The end goal is to obtain information useful for decision making, for instance assessments of damage, predicted areas still in danger, estimates of evacuation times etc. This aspect of situational awareness technologies is concerned with such analysis tools (model based simulations, estimation tools etc as an example) that perform such analysis over the collected information now in a situational database.

Event information thus flows through the three tiers described above. Events are first extracted from raw information, then stored and indexed in an event database, and finally end user analysis and visualization capabilities are provided over these extracted and stored events. An important point we wish to emphasize is that events in our view are independent of the media that they may be captured in and extracted from. The event is a higher level semantic concept corresponding to a real-world occurrence and abstracted from the media it is captured in.

We now describe our work and progress related to the three tiers.

### 3. Event Extraction

Event extraction is concerned with the extraction of high level events from raw streams, we are interested in event extraction from streams of all modalities i.e., text, audio, and video. Recall that we mentioned that human generated information is a key and often major source of information during disasters. A lot of such human generated information is available in forms such as audio (for instance phone calls, radio conversations between responders or responders and victims, etc.) or text (for instance operations command whiteboards, scribed text notes, etc.). In many instances the audio information can be converted to text, via a speech recognizer, for analysis. Analysis of text, particularly the extraction of events from text is thus a prime focus of our work in event extraction. In text extraction we have developed a system and platform for semantic extraction from text which starts from extraction of information from raw reports and documents and goes all the way to high quality extracted information. In audiovisual event extraction we have developed techniques for extraction of information by exploiting simultaneous audio and video streams, a multi-microphone based approach for robust speech recognition, and also an ap-

proach to video stream fusion from multiple sources and visual event detection.

#### *Semantic Extraction from Text*

Event information is presented in text streams in many forms. At real-time for instance, we may have textual information streaming in from a scribe at the disaster making observations and notes as the situation unfolds, or we may have transcribed (audio) conversations between say first responders and operations commanders. Examples of such scribed notes and transcribed conversations are shown in Fig 4.

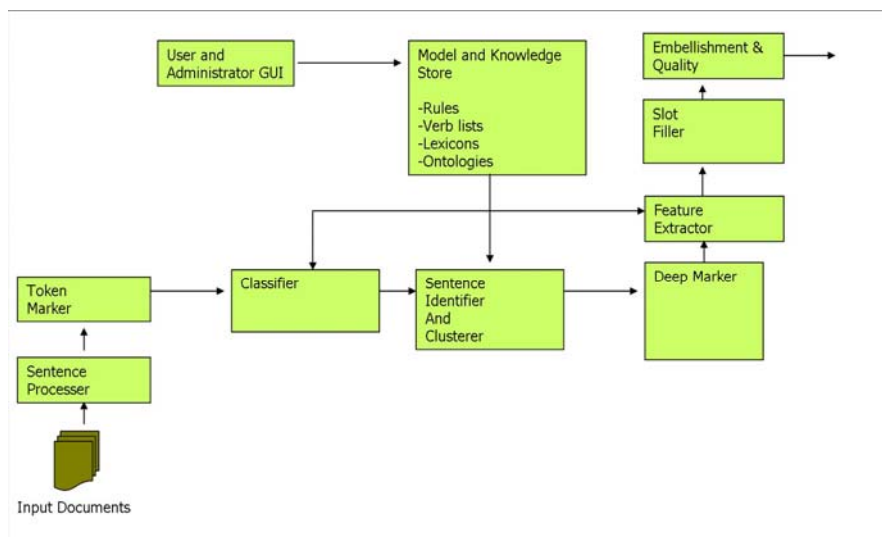
```
ST1: Sampling team still inside.  
  
COM1: Decon teams head out now.  
  
RESP1: Victim 1 shifted to hospital.  
  
RESP1: Sampling team, please provide status.  
  
ST1: Sampling team task completed, returning to  
base.
```

**Fig 4.** Transcribed (Responder) Conversations

There are interesting events described in the example text segment in Fig 4., for instance we have a report of the event of the Decon teams having left, the Sampling team having completed its task, etc. Then we also have information in text reports that become available *post* the disaster. For instance a large number of online news reports, postings in blogs, and message boards become available following a disaster. From the awareness perspective we would be interested in tasks such as extraction of facts from such reports (for instance extracting all reports of donation pledges for the affected victims) for further analysis.

We are developing a next generation extraction platform, called XAR (Fig 5), that performs all tasks in the extraction of information and events from text, starting with raw reports or documents and providing extracted information of high quality and reliability at the end. A key feature of XAR is that the architecture provides for exploiting many “off-the-shelf” and powerful text and language processing systems and tools that have recently become available to the research community; for instance text analysis tools such as GATE [9], and semantic (natural language) parsers such as Shalmaneser [10] and the Stanford Parser [11]. The XAR system

allows for building new extraction applications by specifying a ‘schema’ for the information or event to be extracted along with any *semantic* information that can be provided about the event (such as properties about the different event attributes or ‘slots’). The user may also further provide some processing rules which are essentially logical (Horn Clause based) deductive rules that state what event slot should be filled with what kind of entities. As an example one may define a schema for a donation pledge event to consist of the slots donor, and donor amount. Then we have a rule which states that for instance containing the verb “pledged”, any entity in the sentence which is a country or a relief organization and which appears before the verb pledged, is a value of the donor slot. The system employs the text processors provided to first identify such significant entities in the text and also their features (such as whether it is a country, its relative position in the sentence, etc.) Then, inductive rules are used for the extraction of the detailed information in the event. Such event extraction actually involves several phases such as first classifying which documents or reports could contain the event to be extracted in the first place, then identifying sentences (or groups thereof) that contain the event or are relevant to it, finally slot values are extracted from the sentences; these phases are illustrated in Fig 5 below. We have designed XAR to be able to exploit the semantic information (about the event to be extracted) to the maximum extent. This serves to a) Make the extraction rules simpler and b) Handle extraction cases that cannot be handled by pure linguistic analysis.



**Fig 5.** XAR Event Extraction System

Strictly speaking the above kinds of capabilities serve to extricate information i.e., extract no more than the explicitly stated information in text. This leaves us with limitations, for instance from the text in Fig 4. we can (at best) extract that the location that the victim has been moved to is “the hospital” but not be able to specify which hospital. This is a problem of disambiguation. In many cases, including extraction problems, it is possible to disambiguate such extracted entities. For instance in the above example we may have further contextual information that helps us infer that “the hospital” referred to is most likely the “UCI Hospital”. XAR includes the capability of improving the quality of extracted information by disambiguation. The disambiguation problem is one which we have originally investigated in the data management context, it arises because objects/entities, such as people, organizations, etc, in real world datasets are often referred to using object descriptions, e.g. person name, which might not serve as unique identifiers of those objects, leading to ambiguity. To solve this problem we have proposed a new domain-independent named-entity disambiguation framework [3,4]. The strength of the framework comes from its ability to analyze inter-object relationships, which the traditional techniques cannot do. The proposed approach views the underlying dataset as a graph of entities, represented as nodes, interconnected via relationships, represented as edges. To decide if two object representations co-refer, or refer to the same real-world entity, the approach analyzes all the relationships that exist between the two object representations by discovering all paths that exist in the graph between the two nodes corresponding to the two object representations. Therefore, the framework analyzes both: the direct relationships that exist among objects, that correspond to the direct edges among the objects, as well as indirect relationships among objects that correspond to paths of length more than one. The proposed disambiguation framework is capable of solving two common types of disambiguation challenges. These types are known under a variety of names in different research communities, we will refer to them as (fuzzy) lookup and (fuzzy) grouping. In the settings of the lookup problem the list of all possible objects in the dataset, along with their descriptions, is provided and the goal is for each reference to pinpoint the exact right object it refers to [3,4,22]. In grouping, there is no such list available and the goal is to simply correctly group references that co-refer [3]. Thus, lookup is a classification problem and grouping is a clustering problem.

To solve the lookup problem, the framework discovers the paths that exist among object and it then employs, what is called, a connection strength model to weigh those paths. The outcome of that process is a Nonlinear Programming Problem that needs to be resolved. While it can be solved

using a mathematical solver software, we proposed a simple iterative methods that scales better to large problem sizes. Once the problem is solved, the outcome is interpreted to produce the desired mapping of references to objects [4,22]. The framework solves the grouping problem by also measuring the connections strength and attribute similarity among objects. It then combines those measures and employs either merging or correlation clustering algorithm to produce the desired groups of references [3].

### *Audio-Visual Event Extraction*

This area of research focuses on developing robust technologies for extraction of information from all modalities and exploiting multimodality and contextual knowledge for the extraction of situational information. We are developing approaches for both extraction of information from independent single modality streams as well as extraction and fusion of information from multi-modal information streams about a situation. The research areas spanned by this effort on signal interpretation and analysis based on multiple modalities are Robust speech recognition, Speech enhancement, Auditory stream segregation, Audio Visual event detection and modeling, and cross integration areas like emotion detection from speech signals in enhanced QoS for emergency networks.

Speech Enhancement using multi microphone speech processing: We have developed techniques for more accurate identification and separation of audio information sources. This is done using a multi-microphone approach. We have developed a robust broadband adaptive *beamforming* algorithm (for identifying a particular audio signal source), which combines the robustness of the delay and sum (DS) beamforming in the look direction uncertainty with the high interference rejection capability of the conventional adaptive beamforming algorithm [6,7]. Such a broadband adaptive beamforming algorithm is robust to spatially spread sources. A handheld two-microphone array system has been prototyped and deployed in various RESCUE drills.

Video-based event detection for enhanced situational awareness: Work in this area is centered around *homography*, which in computer vision denotes a linear projective transformation that relates two different views of a scene [8]. This is a novel approach to track people and vehicles in crowded scenes using multiple cameras. Camera calibration is not needed in this approach to extract the world coordinates of objects. Object representation

and tracking in the homography domain are independent of detailed appearances of the objects, and are an efficient and robust method for handling visual events. Semantic understanding of visual events is achieved by spatio-temporal analysis of the objects in the multiple homography representations.

#### **4. Event Data Management**

Event data management is concerned with providing data management or “database like” capabilities for events. Our aim is to provide a data management system that treats events as first-class objects and provides storage, querying, retrieval, and indexing capabilities for the events. We are working on several issues in this regard. The first is the issue of event modeling where we have developed a semantic data model for events. Amongst the various attributes we associate with events, there is an important notion of uncertainty of information. Specifically spatial or location uncertainty is something that we have explored in depth. Below we described the work in event modeling as well as in handling spatial uncertainty in events.

##### *Event Modeling*

A semantic data model  $E$ , for events is described in [12]. There are some key notions related to events and the event centric approach to situational awareness. For instance a *report* is the fundamental information source containing event information. A report could be of any modality, for instance an (audio) phone call reporting the event, information in text reports such as text alerts or news stories, or audio-visual information from say a live TV coverage of a situation. A report is defined then as a physical atomic unit that describes one or more events. An *event* is an instance of an event type in space and time. So an instance of a bridge having collapsed is an example of an event. A situation comprises of a number of events. Events are extracted from reports. An *entity* is an object that occupies space and exists for an extended period of time. Events generally have entities, such as people, objects (such as say cars or planes), etc. associated with them. For instance a vehicle overturning event will have the particular vehicle overturned as one of the entities associated with that event. A *milieu* is the spatial, temporal or spatio-temporal context in which an event, an object or report is situated. Continuing with the bridge collapsing event example, the time and place where the incident occurred are milieus associated with the event.

Note that we have deliberately separated the notion of an event from a specific media, an event may certainly be related to a media (that it is captured in and extracted from) however the event itself is an event per se and not a “text event” or an “audio event”. We now describe the spatial uncertainty work.

### *Spatial Awareness and Indexing*

When reasoning about events and situations, *location* is a key attribute for understanding. To deal with the spatial aspects of events, we have proposed an approach to model and represent (potentially uncertain) event locations described by human reporters in the form of free text [13]. In addition, we have analyzed several types of spatial queries of interest in the context of awareness applications.

As a motivating example, consider the excerpts from two real reports filed by *Port Authority Police Department* (PAPD) Officers who participated in the events of September 11th, 2001:

"... the PAPD Mobile Command Post was located on West St. north of WTC and there was equipment being staged there ..."

"... a PAPD Command Truck parked on the west side of Broadway St. and north of Vesey St. ..."

These two reports refer to the same location of the same command post -- a point-location in the New York, Manhattan area. However, neither the reports specify the exact location of the events, nor do they mention the same street names. We would like to represent such reports in a way that it enables efficient evaluation of spatial queries and analyses. For instance, the representation must enable us to retrieve events in a given geographical region (e.g., around World Trade Center). Likewise, it should enable us to determine similarity between reports based on their spatial properties; e.g., we should be able to determine that the above reports might refer to the same location. Our goal is to represent uncertain locations specified in reports to allow for effective execution of analytical queries. Clearly, merely storing location in the database as free text is not sufficient either to answer spatial queries or to disambiguate reports based on spatial locations. For example, spatial query such as "retrieve events near WTC", based on keywords alone, can only retrieve the first report mentioned earlier. To support spatial analyses on free text reports, we need to project the spatial properties of the event described in the report onto the domain. We do so

by modeling uncertain event locations as random variables that have certain probability density functions (*pdfs*) associated with them. We have developed techniques to map free text onto the corresponding pdf defined over the domain. When the exact address of the event is available, there are many ready solutions that can allow to map it into the corresponding point inside the domain. We instead focus on another relatively frequent case where people report event locations based on certain *landmarks*. Landmarks correspond to significant spatial objects such as buildings, streets, intersections, regions, cities, areas, etc. embedded in the space in which the events or situation are immersed. Spatial location of events specified in those reports can be mapped into *spatial expressions* (s-expressions) that are, in turn, composed of a set of *spatial descriptors* (s-descriptors) (such as *near*, *behind*, *infrontof*, etc) described relative to landmarks. Usually, the set of landmarks, the ontology of spatial descriptors, and the precise interpretations of both are domain and context dependent. Table 1 shows excerpts of free text referring to event locations and the corresponding spatial expressions.

<i>free text</i>	<i>s-expression</i>
'near WTC'	near(WTC)
'on West St., north of WTC'	on(West St.) $\wedge$ north(WTC)

**Table 1.** Examples of s-expressions.

These expressions use WTC and West St. as landmarks. While the locations of landmarks are precise, spatial expressions are inherently uncertain: they usually do not provide enough information to identify the exact point-locations of the events. Our approach to representing uncertain locations described in free text consists of a two-step process. First, a location specified as a free-text is mapped into the corresponding s-expression, which in turn is mapped to its corresponding pdf representation. Given such a model, we have also developed techniques to represent, store and index pdfs to support spatial analysis and efficient query execution over the pdf representations [15,16]. This allows to have complex representations of (location) uncertainty but at the same time support complex queries over such representations efficiently.

## 5. Event Analysis and Visualization

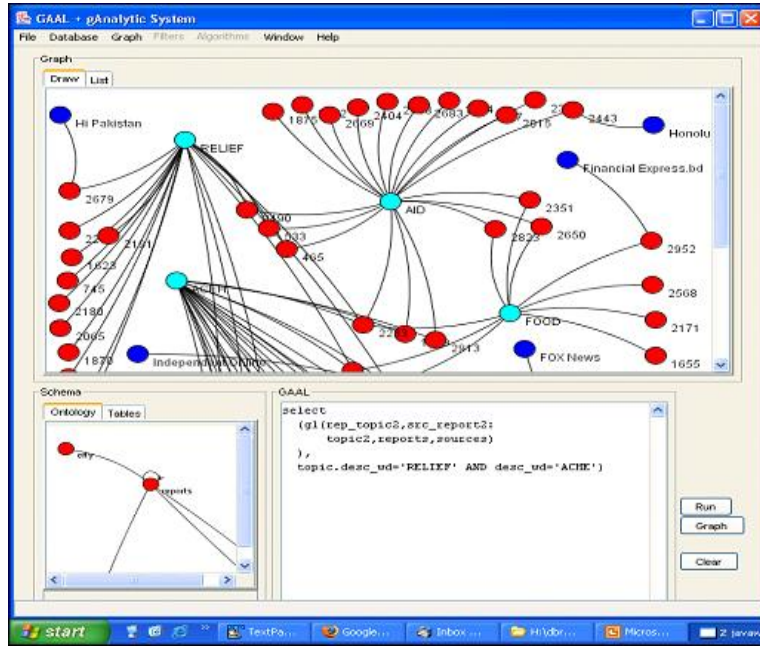
Analysis and visualization is concerned with providing intuitive and (possibly) visual analysis and querying capabilities for end users of situational information. It is reasonable to expect that end users of such systems, such as EOC managers or field commanders would finally like to see patterns and trends in the information collected and get intuitive and visual views of the information (say overlaid over maps etc.). In this regard we have developed tools such as a graph based query algebra and language over events which enables us to query and analyze events in a graphical manner. We have also developed techniques and systems in the area of predictive modeling, that we have applied to tasks such as people estimation and route prediction. We describe these below.

### *Semantic Graph Query and Analysis*

A large collection of text reports are generated during a disaster. In its original form, this data is of limited use since the user can apply only keyword based search. Although event/entity extraction can be employed to extract events, related entities and inter-event relationships to produce a more structured data, the extracted semi-structured data still does not fit the traditional tabular representation and hence most existing exploratory and analytical tools can not be readily applied (in most cases). Our *semantic graph query and analysis* research aims at developing techniques tailored to analyzing such data.

Once events and entities, their attributes and their relationships are extracted from various textual sources, they need to be represented using a generic data model that is amenable for more sophisticated query and analysis. Semantic graphs which capture data at the lowest semantic granule provide such a common representation. The nodes in such a graph correspond to uniquely identified events, entities, their attribute values and source reports. The edges correspond to a variety of semantic relationships among the nodes. In addition to extracted data, semantic graphs are also used to represent domain knowledge in the form of taxonomies and ontologies which can be used in the query and analysis of the data. This ability of semantic graphs to symmetrically represent both extracted data and domain knowledge makes them ideal for a variety of disaster report analysis tasks. For instance Fig 6 provides an illustration of the use of graphs for analyzing news reports (of a disaster). We have nodes (of different colors) representing entities such as news stories, news sources, news topics etc. and the graph edges depict relationships amongst them. The graphical rep-

resentation is more intuitive and powerful for end users, for instance one can visually see correlations between news stories and sources and news topics, correlations with and over certain periods of time etc.



**Fig 6.** Graph Analysis of Events

Below we discuss three ways in which our research is extending current technology to enable query and analysis on semantic graphs.

(1) *Query Algebra:* We have developed a query algebra, called *Graph Analytic Language (GAL)*, which offers a number of features to enable analysis of semantic graphs including graph-aware aggregation and grouping, well-defined integration of graph and set-based operations, operators to interface with ontology inference engines as well as a variety of bulk operators [17]. These operators enable us to pose analytical queries similar in spirit to OLAP queries (for example, over events at multiple levels of composition/ resolution).

(2) *Exploratory Analysis Operators*: We have developed a number of exploratory analysis operators (and their implementation algorithms), that enable the identification of taxonomy concepts that best *summarize* a set of reports, or best *differentiate* or *associate* a pair of report collections, for example from two disaster incidents [17,18].

(3) *Semantic Relationship Analysis*: We have developed a version of the OLAP cube operator (called Multi-Relational Iceberg-Cubes) that enables the user to detect entity attribute values that characterize prominent semantic relationships among them [17].

### *Predictive Modeling*

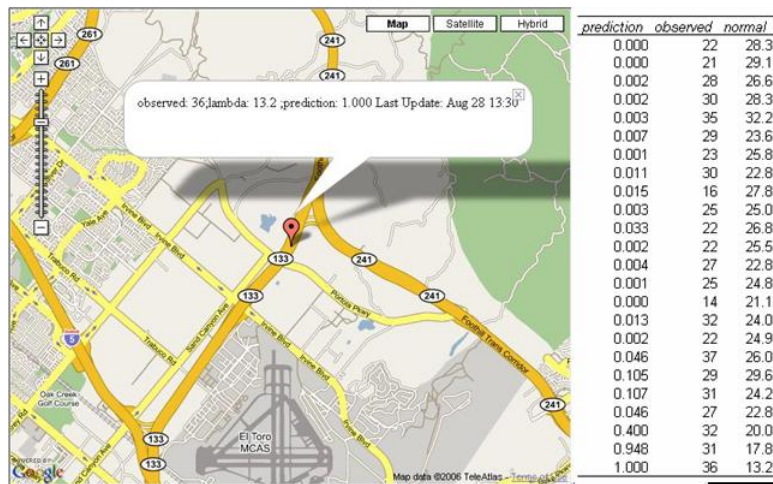
Another key aspect of analysis, particularly in disasters, is the ability to predict and estimate certain things. For instance, an estimation of the current location and number of people in a disaster area and a prediction of the future movement of those people could provide critical information to EOC (Emergency Operations Command) personnel responsible for rescue and evacuation, and also to members of the public looking for the best way to navigate to their desired destination. We are using statistical models and bayesian techniques to extract information about human behavior from a variety of “human detection” sensors such as loop sensors counting cars on a freeway, people counters at doors to buildings and GPS devices on cell-phones or cars. Our aim is to enable the creation of a “People Forecast” similar to a weather forecast that, instead of giving current weather conditions and a prediction of the weather throughout the rest of the day, gives an easily interpretable picture of the current condition of the population (such as the location and density of people in an area, indications of unusual behavior or unusual events taking place) and also a prediction of where people will be moving in the near future. We are working on two projects, one dealing with situational analysis and estimation of current conditions and the other dealing with predicting future movement of people.

### *Situational Assessment/Current conditions*

Occupancy analysis in real time is a difficult problem that has not been approached in a systematic way to our knowledge. Current approaches to large scale estimation of the number of people in a given area often use dated census information, perhaps with some simple adjustment for day-time versus nighttime predictions. No current approaches appear to use real-time measurements of human activity such as freeway traffic information given by loop sensors on highways to influence their predictions.

While occupancy analysis remains a difficult problem, we have identified smaller sub-problems that also give useful information, and we have had some success in extracting interesting and useful information.

The first problem we identified is detecting unusual activity or events from sensors that measure human activity in some form such as people counters or freeway sensors. Our algorithm learns (unsupervised) normal traffic patterns and detects "event" activity defined as short bursty periods of unusually high or low traffic in real time using a time varying Poisson process model. The output of the inference algorithm is a probability that an event is taking place at any given time, along with an estimate of the extra number of people or cars or whatever you are counting that are attributed to that event. Also, we are creating software for a real time event detection system for freeway traffic in southern California. Users will be able to log on to a website that will display a map of the area of interest that shows current event activity and a history of event activity. (see screen-shot below). As an example of how this could be used in an EOC center, imagine evacuation planning during a bomb threat to downtown LA; the real-time event detection system is brought up, events at off-ramps in the past 3 hours is selected, and all off ramps near Dodgers stadium indicate the presence of an event, furthermore an extra 20,000 cars are predicted to have used those off ramps above the normal traffic; this gives a quick indication that something is going on in Dodgers stadium as well as an idea of the popularity of the event.

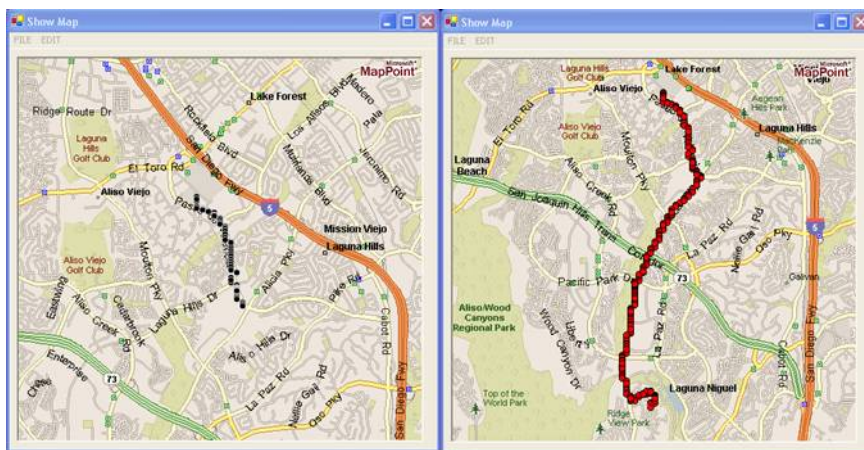


**Fig 7.** Anomalous Event Detection

Our second project deals with modeling car travel activities of individuals using GPS data. In particular, we have developed a probabilistic model that answers the following questions about individuals:

- 1) Given the current time, where the individual is most likely to be (e.g. his home, office or traveling).
- 2) Given the current time, where the individual is most likely to go (e.g. pick up his kids from the day-car center, grocery shopping etc).
- 3) Given that the person is traveling, which route the individual is most likely to take to the destination.

Our probabilistic model [19,20,21] is accurate almost 84% of the time in predicting the most likely destination/location of a typical individual and thus could provide valuable information to first responders and transportation planners in planning evacuation. Our hope is that even if a small sample of the traveling public agreed to collect their travel data and make that data publicly available, transportation management systems could significantly improve their operational efficiency. We also envision our probabilistic model to be used in devices like “*personal traffic assistants*” that run on the mobile devices which could help travelers re-plan their travel when the routes they typically use are impacted by failures in the system arising from accidents, terrorist attacks and natural disasters such as earthquakes.



**Fig 8. Route Prediction**

In the figure above, we show a typical execution of our system. On the left is the person's current location derived from the GPS device on his car/cell-phone and on the right is the route the person is most likely to take based on his past history (predicted by our model).

## 6. Artifacts

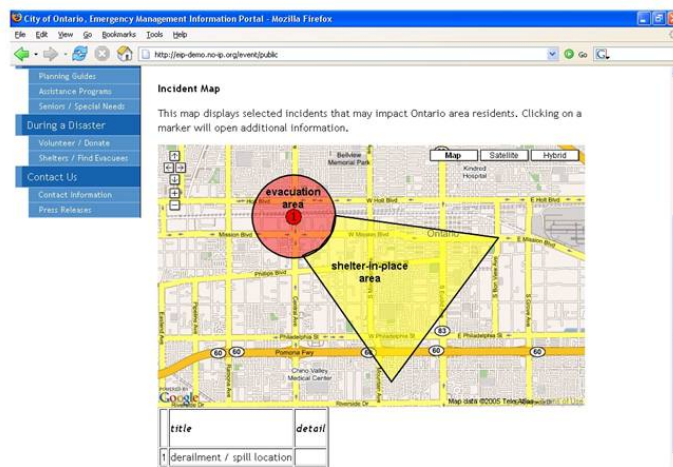
In addition to research, a key aspect of all RESCUE projects, including situational awareness, is the development of representative *artifacts* that are essentially practical demonstrations of our technology capabilities, and driven by first responder needs. In situational awareness we are pursuing the development of two such artifacts at the conceptual level. One is a conception of an "information dashboard" which is essentially an integrated dashboard of processed, integrated, multi-modal information that can be used for decision making and the other is a real-time multi-modal reconnaissance system that can be employed in real-time data collection and extraction. Here we describe our work on the situational dashboard.

### *The Situational Dashboard*

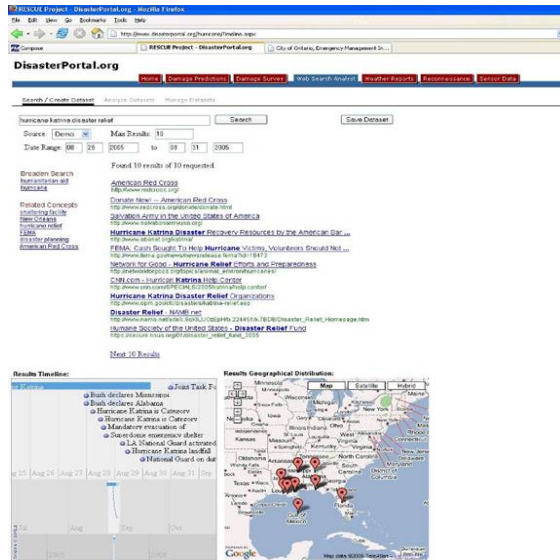
This artifact is aimed at facilitating awareness from multi-modal data and enabling end-users to apply situational information in decision-making and analysis. The situational dashboard provides integrated and fused information from multiple sources and provides analysis, querying, triaging and filtering functionality. Such a system can provide, say, an EOC with an awareness of events as they unfold. Emergency managers can use the information processed for planning, analysis and logistics tasks. For instance, information about traffic conditions and also the location of people or vehicles can be used for analysis such as evacuation and route planning. Clusters of information can be used for damage and impact assessment. It can be used to dispatch search and rescue and support logistical needs.

We are developing different actual manifestations of the dashboard concept. These include a set of disaster information portals that provide integrated online access to a variety of useful information, to both citizens as well as responders in a crisis. In conjunction with the City of Ontario, CA, we are developing a web-based portal (Fig 9) for dissemination of information to that community. This information includes real-time details of incidents and response, as well as the status of public infrastructure, such as roadways. During large-scale events the information to be shared can

include general announcements, maps, emergency shelter information, family reunification assistance, management of donations and volunteer resources, etc. Other capabilities we have developed include the ability for users to perform searches for the whereabouts of displaced family and friends, traffic flow monitoring and identification of anomalous events, and an automated system by which local businesses and organizations can sign up to receive customized information during an incident via a number of modalities including web, phone, or SMS.



**Fig 9.** Ontario Disaster Portal Interface



**Fig 10.** Hurricane Portal Interface

The other portal instantiation is a Hurricane Portal for large scale disaster. During such disasters it can be difficult for first responders groups as well as the public to extract a good level of situational awareness from the vast amount of information which is quickly generated by media as well as individuals. The RESCUE Hurricane Portal aims to vastly improve access to information of use to these groups, especially during the first several days after a hurricane makes landfall. The initial portal development (fig 10) is focused on developing tools to help users find pertinent data from the Web, eyewitnesses, and computer simulations, and aiding in transforming the found data into information which is actionable for the purpose of minimizing negative impact of the event to people and property. A key capability in this application is the development of multi-dimensional, *faceted* web search, analysis, and visualization capabilities which are customized to the disaster management subject domain.

We are aiming to release test versions of both these portals in the coming months, also we are taking active and regular guidance from responders and emergency response organization personnel affiliated with RESCUE in the development.

## 7. Conclusion

We described above our work in many areas towards a situational awareness system for disaster response. From the application perspective, our interactions with the first-responder community as well as our field work with the emergency organizations have led us to identify the key challenges in situational awareness. First, sometimes the information is almost entirely missing or so uncertain that it is impossible to meaningfully ascertain the situation. This is usually the case in the immediate aftermath of the disaster. An approach to improve situational awareness in such a case is better preparedness – e.g., deploying sensing equipment that can provide an early synopsis of the damage, etc. maintaining accurate information about resources, vulnerabilities, etc. so that accurate predictions can be made even in the absence of field level data. Second, situational information often arrives across multiple modalities at different times in an uncorrelated fashion from diverse sources. While there is an avalanche of data, much of it is underutilized since there do not exist any tools that support seamless access to useful actionable information in a timely fashion. The situational awareness technologies we have described in all the tiers i.e., event extraction, event data management and event analysis are aimed at addressing exactly this problem of converting available data to actionable information. Third, much of the information is human generated (e.g., by first responders, citizen journalists, etc.). While such information has the advantage of human interpretation, it could also suffer from cognitive bias and misinterpretation by humans. Technologies to leverage such human-generated information in a robust timely manner do not exist. Furthermore, even if such tools existed, the impact of taking actions based on such information is unclear. Also such information is usually “report-centric”. To make the information useful for decision making there is a need to make the information “situation-centric” through processing. The event-centric approach that we are developing to situational awareness serves to convert the report-centric information to information at the event and situation level of abstraction. Finally, information required to gain situational awareness often resides across different organizational boundaries with restrictions on their availability and policies for sharing. Tools for seamless sharing of information across information silos do not exist. We are addressing this problem as part of the policy based information sharing project described in the introduction.

The applicability of our research thus goes to the core of the advancements required for next-generation situational awareness systems in disaster response; our hope is to be able to advance the state-of-the-art in situ-

ational awareness technology as well as impact first responders in emergency organizations.

## Acknowledgements

We have described the work of several researchers and students working on the situational awareness technologies project in RESCUE. We wish to acknowledge these members who have made key contributions to the research as well as to the writing of this chapter. These include Rina Dechter, Ron Eguchi, Vibhav Gogate, Jon Hutchins, Charles Huyck, Yiming Ma, Sangho Park, Bhaskar Rao, Dawit Seid, Shankar Shivappa, Mohan Trivedi and Wenyi Zhang.

## References

- [1] Sharad Mehrotra, Carter Butts, Dmitri Kalashnikov, Nalini Venkatasubramanian, Ramesh Rao, Ganz Chockalingam, Ron Eguchi, Beverly Adams and Charles Huyck.  
Project Rescue: Challenges in Responding to the Unexpected  
IS&T/SPIE 16th Annual Symposium on Electronic Imaging. January 18–22 2004, San Jose, California, USA, 2003-12
- [2] Dmitri Kalashnikov, Dawit Seid, Yiming Ma, Naveen Ashish, Sharad Mehrotra and Nalini Venkatasubramanian  
Event Based Approach to Situational Awareness  
Calit2 Report, UC Irvine
- [3] Zhaoqi Chen, Dmitri V. Kalashnikov, Sharad Mehrotra: Exploiting relationships for object consolidation. IQIS 2005: 47-58
- [4] Dmitri V. Kalashnikov, Sharad Mehrotra, Zhaoqi Chen: Exploiting Relationships for Domain-Independent Data Cleaning. SDM 2005
- [5] Reynold Cheng, Dmitri V. Kalashnikov, and Sunil Prabhakar.  
Evaluating probabilistic queries over imprecise data.  
In Proc. of ACM SIGMOD Int'l Conf. on Management of Data (ACM SIGMOD), June 9-12, 2003.
- [6] Wenyi Zhang and Bhaskar Rao

Robust Broadband Beamformer with Diagonally Loaded Constraint Matrix and its Application to Speech Recognition, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-2006, Toulouse, France.

[7] Wenyi Zhang and Bhaskar Rao  
Robust Adaptive Beamformer with Feasibility Constraint on the Steering Vector, European Signal Processing Conference, EUSPICO-2006, Italy.

[8] S. Park, M. M. Trivedi  
A Track-based Human Movement Analysis and Privacy Protection System Adaptive to Environmental Contexts  
Proc. IEEE International Conference on Advanced Video and Signal based Surveillance, Sep2005

[9] H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan.  
GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications.  
Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02). Philadelphia, July 2002

[10] K. Erk and S. Pado  
Shalmaneser - a flexible toolbox for semantic role assignment  
Proceedings of LREC-06, Genoa.

[11] Dan Klein and Christopher D. Manning.  
Fast Exact Inference with a Factored Model for Natural Language Parsing.  
In Advances in Neural Information Processing Systems 15 (NIPS 2002), December 2002.

[12] Bo Gong, Utz Westermann, Srikanth Agaram and Ramesh Jain.  
Event Discovery in Multimedia Reconnaissance Data Using Event Clustering  
AAAI Workshop on Event Extraction and Synthesis, Boston MA, July 2006

[13] Dmitri V. Kalashnikov, Yiming Ma, Sharad Mehrotra, Ram Hariharan, Nalini Venkatasubramanian, and Naveen Ashish  
SAT: Spatial Awareness from Textual Input.  
In Proc. of Int'l Conf. on Extending Database Technology (EDBT), demo publication, March 26–30, 2006.

- [14] S. Mehrotra, C. Butts, D. Kalashnikov, N. Venkatasubramanian, K. Altintas, et al.  
CAMAS: A Citizen Awareness System for Crisis Mitigation.  
In Proc. of ACM SIGMOD Int'l Conf. on Management of Data (ACM SIGMOD), demo publication, June 13-18, 2004.
- [15] Dmitri V. Kalashnikov, Yiming Ma, Sharad Mehrotra, and Ram Hariharan.  
Index for Fast Retrieval of Uncertain Spatial Point Data.  
In Proc. of Int'l Symposium on Advances in Geographic Information Systems (ACM GIS), November 10-11, 2006.
- [16] Dmitri V. Kalashnikov, Yiming Ma, Sharad Mehrotra, and Ram Hariharan.  
Modeling and Querying Uncertain Spatial Information for Situational Awareness Applications.  
In Proc. of Int'l Symposium on Advances in Geographic Information Systems (ACM GIS), November 10-11, 2006.
- [17] Dawit Seid and Sharad Mehrotra  
Algebraic Support and Optimization for Semantic Queries  
RESCUE technical report, 2006.
- [18] Dawit Yimam Seid and Sharad Mehrotra  
Efficient Relationship Pattern Mining Using Multi-Relational Iceberg-Cubes.  
ICDM, 2004, pp. 515-518.
- [19] Vibhav Gogate, Rina Dechter, Bozhena Bidyuk, Craig Rindt and James Marca.  
Modeling Transportation Routines using Hybrid Dynamic Mixed Networks.  
In Uncertainty in Artificial Intelligence 2005.
- [20] Vibhav Gogate and Rina Dechter.  
Approximate Inference Algorithms for Hybrid Bayesian Networks with Discrete Constraints.  
In Uncertainty in Artificial Intelligence 2005.
- [21] Vibhav Gogate, Rina Dechter, Bozhena Bidyuk, Craig Rindt, and James Marca.

Modeling Travel and Activity Routines using Hybrid Dynamic Mixed Networks.

In 85th annual meeting of the Transportation Research Board, 2006.

[22] Dmitri V. Kalashnikov, Sharad Mehrotra

Domain-independent data cleaning via analysis of entity-relationship graph.

ACM Transactions on Database Systems 31(2): 716-767 (2006)