Project Rescue: Challenges in Responding to the Unexpected

S. Mehrotra, C. Butts, D. Kalashnikov, N. Venkatasubramanian
University of California, Irvine[1]

R. Rao, G. Chockalingam
University of California, San Diego[2]

R. Eguchi, B. Adams, C. Huyck
ImageCat, Inc.[3]

ABSTRACT

This paper provides an overview of Project RESCUE, which aims to enhance the mitigation capabilities of first responders in the event of a crisis by dramatically transforming their ability to collect, store, analyze, interpret, share and disseminate data. The multidisciplinary research agenda incorporates a variety of information technologies: networks; distributed systems; databases; image and video processing; and machine learning, together with subjective information obtained through social science. While the IT challenges focus on systems and algorithms to get the right information to the right person at the right time, social science provides the right context. Besides providing an overview of the nature of RESCUE research activities the paper highlights challenges of particular interest to the internet imaging community.

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). Clearly, 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. The 'Responding to Crises and Unexpected Events' (RESCUE) Project was recently conceived with the objective of radically transforming the ability of organizations to gather, manage, analyze and disseminate information when responding to man-made and natural catastrophes. 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. Project RESCUE involves researchers from six universities (University of California, Irvine; University of California, San Diego; University of Colorado, Boulder; University of Illinois, Urbana Champaign; University of Maryland, College Park; and BYU) and a Long Beach based advanced technology company, ImageCat, Inc. The project is funded through an NSF Large Information Technology Grant and has active participation and collaboration with Government agencies from the Cities of Irvine, San Diego, Los Angeles as well as the County of Los Angeles and State of California

The present paper highlights some of the key information technology challenges being addressed in Project RESCUE. Appreciating the IT challenges in improving crisis response requires a thorough understanding of how communication and control networks form among responding organizations, and of how the response process is organized. Since the

[1] ITR-Rescue, 5251 California Avenue, Suite 210, Irvine CA 92612-2815, Phone: 949-824-1147 http://www.itr-rescue.org
[2] Cal-(IT)², EBU1, 7th Floor, 9500 Gilman Drive, La Jolla, CA 92093-0405 Phone: 858-822-1189, info@calit2.net, http://www.calit2.net
[3] Union Bank Building, 400 Oceangate, Suite 1050, Long Beach CA 90802 Phone: 562-628-1675, http://imagecatinc.com

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 the RESCUE Project.

1.1. 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 example, 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 ongoing incidents.

As indicated in Figure 2, 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.

2. OVERVIEW OF THE RESCUE PROJECT

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 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 infrastructures 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.

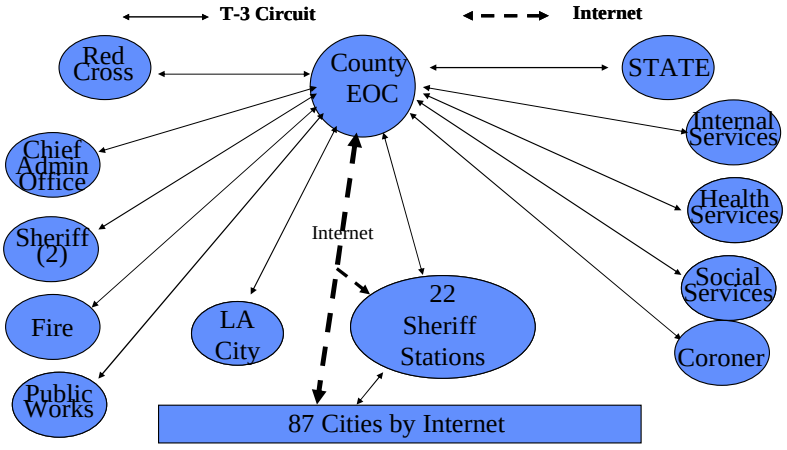

Figure 1: LA County Disaster Network

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.

1.2. Summary of Research Agenda

Overcoming the challenges described above requires an integrated research program in information technology and social science. While IT solutions can facilitate rapid and seamless access to, and dissemination of, information, the social science research on understanding the distinctive nature of dynamic virtual organizations, their information needs, and the social and cultural aspects of information sharing across organizations and individuals can provide the right context. To explore such a multidisciplinary approach, our research (summarized in Figure 2) is organized along four inter-related activities that together capture the information flow process during crisis response.

Information Collection: The objective of information collection is to gather relevant crisis-related information from a variety of information sources in a timely and efficient manner. Our research addresses how large volumes of highly dynamic multimodal information generated by various information sources can be effectively collected and transmitted over networks that might be unreliable and/or insecure. We explore integrated distributed systems, data management, and networking systems that enable information to seamlessly flow in real-time from information sources to collection points. In designing such systems, we focus on two novel aspects of the information collection problem. First, we explore technological solutions to enable humans to act as vital information sources in a crisis. We believe that eyewitnesses' and first responders' observations and interpretations from the scene (given the benefit of human cognizance) can provide a vital adjunct to instrumented sensors as tools for situation assessment. Leveraged properly, human sources can result in much more accurate situation awareness and hence better response. The importance of human input is recognized among response workers[1], yet existing crisis response systems do not systematically exploit such information sources. We will therefore seek scalable and robust IT solutions that enable crisis workers (as well as other observers) to play the role of ***human as sensor***. A second novel aspect of our research is the exploration of privacy in the context of information collection, with the objective of developing customizable solutions that maintain critical functionality while minimizing privacy loss. Spaces instrumented with video cameras, sensors, tracking systems based on RFID or similar technologies, while facilitating surveillance and situation monitoring, leave traces of individual activities without the capacity for "opt-out" on the part of those being monitored. Such systems are potentially prone to abuse (e.g., misuse of data by law enforcement or private entities) and may in some cases violate legal constraints on government surveillance activities.). We envision data collection mechanisms that are privacy-aware and adaptive. Using principles of data reduction, minimal collection, need-to-know provision, and statistical obfuscation, our research aims to develop IT solutions which collect and provide critical information to decision makers without falling prey to the "record everything and store forever" approach. Such mechanisms will empower the users, to the extent possible, to control the acceptable loss of privacy, and adapt to the varying needs of the situation.

Information Analysis: a set of related research activities that focus on extracting useful information (from the perspective of crisis response) from raw data (sensor data, text, voice, and video streams) and assimilating or fusing data across multiple modalities to develop higher-level understanding. The information analysis component bridges the gap between raw data and semantically richer representations that are useful to humans in the context of their tasks (damage assessment, planning, situational awareness). A fundamental abstraction that permeates various aspects of information analysis is that of an "*event*". We define an event as a significant phenomenon or occurrence embedded in space and time. Events can occur at multiple spatial and temporal scales. For a field worker (e.g., a firefighter requiring fine-grained situation awareness), change in location of his team members at a given time might be an event. At a higher layer, relevant events may correspond to more aggregated information. For example, incident commanders monitoring the response at the city EOC may choose to focus on the total number of deployed firefighter or police units at a given region at a given time instead of detailed information about each individual/resource. Our research on information analysis focuses on extracting meaningful events from multimodal data streams (e.g., transcribed voice conversations in the context of human sensors, video feeds, various types of instrumented sensor data), and techniques to assimilate events, fuse information, and integrate it with (possibly static) information -- e.g., maps of the crisis site.

Information Sharing: a set of related research activities that facilitate seamless sharing of information amongst decision makers, especially when such communication crosses organizational boundaries. The crisis response environment is characterized by Dynamic and Evolving Virtual Organizations **(*DEVOs),*** aggregate entities in which participating organizations collaborate to perform a variety of tasks related to disaster mitigation (e.g., rescue and evacuation, maintaining law and order). Seamless mechanisms for inter-organizational information sharing can revolutionize how such emergent collaborations are established and maintained, resulting in dramatic improvements to

crisis response. Such information may have been dynamically collected and analyzed, or preexisting in organizational knowledge and databases. Challenges in information sharing across DEVOs arise due to frequent structural and functional changes (e.g. expansion, extension) within organizations, emergence of complex inter-organizational relationships, *lack of centralized control*, and an element of the surprise resulting in unexpected inter-organizational relationships and data needs[6,40,10,42]. Disaster response networks are characterized by heterogeneity in the nature of network relationships (e.g., direction and control vs. voluntary coordination, formal or contractual vs. informal relationships) and shifting composition, as new organizational entities join the network in response to changing conditions and disaster-related demands[32]. To address the information sharing challenges in such virtual organizations, our work represents an integrated approach to the assessment of information sharing needs and optimization of organizational structures in the context of response, coupled with technological innovations that facilitate their formation and functioning. From the IT perspective, our challenge is to develop an open, distributed system with adaptive and scalable mechanisms for rapid and secure sharing of information. Such a system must incorporate mechanisms for dynamic trust negotiation between organizations, as well as techniques for "translating" initially incompatible data formats and storage practices to permit data transfer.

Information Dissemination: a set of related research activities that attempt to determine how information can be dispersed quickly and efficiently to various entities and organizations. Our research focuses on how various forms of social behavior emerge during disasters, and how this knowledge can help customize and personalize information delivery to save lives and property. This research area will focus on the challenges associated with the timely dissemination of information to entities participating in disaster response activities, to other organizations (e.g., mass media organizations), and to the general public. For over five decades, empirical social science research has focused on issues related to the dissemination of hazard-related information in both pre- and post-disaster contexts. We now have a good understanding of how crisis-relevant information can be effectively disseminated, as well as a wide range of challenges and difficulties associated with communicating in turbulent, rapidly-evolving, and uncertain environments[29,31,38]. This research has conceptualized information dissemination as a communication process consisting of four elements: *origin of information; messages; dissemination channels; and receivers of information*[41]. Effective information dissemination strategies (that ensure timely delivery to recipients in the form most useful to them) require a deeper understanding of each of these elements and their interactions. Our focus in this research is two-fold: (1) developing a structured approach to understand interactions between the above listed communication elements in the context of crisis response (2) designing advanced IT solutions that customize both dissemination strategies and message content to meet needs of diverse organizational actors and segments of the public.

1.3. RESCUE Project Research Testbeds

Three testbeds are being employed to deploy the technologies developed by the research team and determine the extent to which the information- and organizationally-based strategies result in demonstrable improvements in response effectiveness. These testbeds have been chosen to simulate a variety of disaster situations that pose varying requirements on the IT infrastructure. The testbeds provide the team with an experimental platform to field-test and refine research on information collection, analysis, sharing, and dissemination in controlled yet realistic settings, significantly enhancing our research capability.

1.3.1. Mobile Incidence Level Response (MILLR) Testbed

The MILLR testbed consists of a rapidly deployable mobile networking, computing, and geo-localization infrastructure in the context of incident-level response to spatially localized disasters such as the World Trade Center attack. The testbed focuses on situations where the crisis-site either does not have an existing infrastructure or, alternatively, where the infrastructure is severely damaged. This testbed focuses on supporting basic services essential to the first responders that can be transported to crisis sites for rapid deployment. Such services include communication among the first responders, accurate geo-localization both inside and outside of buildings in urban as well as rural areas, computation infrastructure, incidence level command center, and technology to support information flow from/to crisis sites to/from regional emergency centers. For this purpose, four UCSD 35 passenger shuttles equipped with mobile wireless technologies (802.11b wireless coverage with a 3G cellular data backhaul), compute, location, and multiple display resources on board are being used as mobile command and control centers. The primary demonstration study of the MILLR testbed is in the context of localized incident management in the Gas Lamp district in downtown San Diego. **The Gas Lamp Quarter** is being instrumented in partnership with the civic authorities and the San Diego Police

Department. This 10 block historic region adjoins the San Diego convention center, sports stadiums as well as city, county and federal facilities and is described as Southern California's premier dining, shopping and entertainment district. Working with the San Diego Police, the RESCUE research team will deploy the MILLR testbed in the GLQ area to facilitate situation awareness during events organized in GLQ.

An important component of the MILLR testbed is a multimodal notification system that can disseminate information uniformly across various networks and devices. A preliminary version of this system has been deployed and tested in various scenarios ranging from homeland security to dissemination of travel related information to the public. The system is capable of alerting/notifying multiple parties on a massive scale via voice calls and SMS. It has been used in conjunction with a prototype geo-fencing system (consisting of camera based technology to automatically detect intrusions) around the Coronado Bridge to alert responsible security agencies (via a voice call and a text message) when a vessel enters the geo-fence. It was also used during Super Bowl 2003 (working in collaboration with the San Diego Police department) to provide anywhere, anytime access to real time traffic information on the San Diego freeways, through cell phones and to allow the police to communicate with the fleet that was responsible for shuttling the fans from to the Qualcomm stadium to the parking lots which were a few miles from the stadium. Using our notification system, the police were dynamically able to change the prompts of the traffic notifications system so that the drivers of the fleet were aware of the latest congestions and closures. The traffic notification has continuously operational since Super Bowl serving the San Diego commuters via a phone call (866) 500 0977 and the web http://traffic.calit2.net providing them with up to the minute traffic information.

1.3.2. Crisis Assessment, Mitigation, and Analysis System (CAMAS) Testbed

In CAMAS our focus is on leveraging the existing fixed infrastructure for mitigating incident-level crisis in instrumented spaces such as airports, malls, nuclear facilities, research facilities, etc. Such instrumented spaces may have an existing surveillance and network infrastructure that may be partially/completely operational and can be used by first-responders for crisis mitigation. To simulate such an environment, we will instrument a significant portion of the UCI campus to create smart corridors and open spaces. The UCI campus consists of two rings of building built around a central park. Various schools in the campus (e.g., School of Information and Computer Science, School of Engineering, School of Social Science) consist of a set of geographically clustered buildings spread over the two rings with large open areas (called quads). We will instrument the inner ring and the quad areas of the campus with wireless access points to provide wireless connectivity, and install video cameras to monitor events in these regions. The instrumented space corresponds to the heart of most campus level events and gatherings (e.g., for students during Graduation Day Ceremony) and provides us with an ideal experimental testbed for many of the information technology solutions (e.g., collecting multimodal data in dynamic settings, capturing live events and testing multimodal event extraction algorithms, studying collaboration among user groups in dynamic settings in the context of specific events, privacy issues in information gathering, real-time voice analysis, models of reliability of human input, etc.) we are developing as part of the RESCUE project. Such an infrastructure also allows us to simulate a variety of network conditions, ranging from full connectivity to situations when the network is mostly inoperable, and to study the impact of network conditions on data collection, analysis and processing techniques for situational assessment. Over the course of the RESCUE Project, the UCSD Cyber Shuttle along with mobile networking and geo-location infrastructure, will be brought to the UCI campus and will serve as an Incident Command Center to control part of the ground activities during a few special events (e.g., Graduation Day at UCI). In addition to these elements, CAMAS also incorporates passive information collection from human sensors, in the form of an integrated problem reporting system which translates raw text messages (and, ultimately, speech) into information regarding problem events on campus. Research currently underway includes an automated data analysis system to convert such event reports into probabilistic information regarding unknown problem states. This system incorporates informant and problem attributes into the analysis, thereby allowing for automatic downweighing of suspect information and robustness to denial of service attacks. The information resulting from the analysis can then be utilized by the automatic alert dissemination system to notify on-campus responders of potential problems, while minimizing false positives.

1.3.3. Advanced Traffic Rerouting for Unplanned Events (TRUE) Testbed

Our final testbed simulates a geographically dispersed disaster (e.g., an earthquake) and studies the role of information technology in coordinating large-scale activities such as evacuations and forestalling cascades. After an unplanned event, emergency responders need the ability to minimize loss and disruption through efficient traffic re-routing.

Although transportation models are widely used in public policy, these methods have not been adopted for regional emergency response. Faster computers, widespread use of handhelds and wireless networks now make it possible for responders to detect impediments to traffic flow in real time and calculate routing alternatives in near-real time. We are currently developing an Internet Mapping System (IMS) that models the impact of unplanned events with the help of user input and generates a suite of evolving routes and evacuation alternatives so as to minimize traffic disruption. Such a system will use detailed, building-level activity data to disaggregate Origin-Destination matrices and estimate evacuation demand. Using earthquakes as testbeds, damage to the network is being estimated through loss estimation tools, such as EPEDAT. This allows for both simulations and real-time damage estimates using USGS ShakeMaps. For other unplanned events, users will be able to define custom scenarios by specifying locations, regions, or transportation corridors associated with hazards. Alternatively, users will have the option to import avoidance areas from tools such as plume models (ALOHA, CAMEO), or flood models (HAZUS-MH). Given network capacity, the system will establish route alternatives so that the aggregated time to traverse the route (linear detour), or the time exposed to the event (evacuation) is minimized.

While the scope of research in the RESCUE project is broad, below we focus on a few of the challenges we are addressing of interest to the internet imaging community in the following three sections.

3. DYNAMIC PRIVACY PRESERVING MULTIMODAL DATA COLLECTION

Our objective in the context of developing a data collection framework is two fold: (1) We wish to design a robust end-to-end data collection infrastructure that provides seamless access to dynamic multimodal information from heterogeneous sources connected via potentially damaged or partially available networks. (2) We wish to explore issues related to privacy in data collection especially in the crisis context with the objective of developing customizable solutions that explore a tradeoff between functionality and the (loss of) privacy. We discuss each in turn.

Quality-Aware Multimodal Information Collection: Our approach to information collection is based on the following two observations. First, applications may be able to tolerate bounded amounts of inaccuracy in the data collected. For example, approximate loop sensor data might suffice to build traffic models accurate enough to support traffic planning. Similarly, a low-resolution image might provide a required level of situation awareness for surveillance monitoring. A collection framework can exploit the applications' tolerance to data quality in order to reduce the data capture and transmission overhead. The second observation is that data producers (sensors, video cameras, etc.) may have (limited) storage and computation capabilities that can be exploited to reduce overhead. For example, data captured at a sensor can be filtered, temporarily stored, compressed and then transmitted to reduce overall cost. Similarly, simple video processing can be performed at the camera to detect events that need to be transmitted instead of transmission of the raw video. Exploiting the capability of the data producers in order to minimize resource usage requires the collection framework to be tightly integrated with the data analysis and decision support system.

Developing an integrated end-to-end data collection infrastructure requires a judicious choice of strategies at the data management, middleware and networking layers. The traditional approach is to isolate functionality at each layer; networking protocols ensure connectivity and middleware mechanisms ensure real-time collection in the presence of faults. Built on top of this substrate is a data management system that holds periodic information snapshots from data sources in a (logically) centralized database server[39,30,46,11]. The drawback with such a layered approach is that by generalizing the collection process it fails to capture variations in required data quality from disparate sources. This can result in wasted resources (e.g. due to collection from irrelevant sources) and degraded information quality (e.g. delayed/lost reports due to fluctuations in network availability), both of which can significantly compromise information flow in a crisis environment[19]. What is required is an integrated end-to-end quality-aware distributed data collection architecture that exploits application quality requirements to drive the data management, middleware and networking layers. Much in the spirit of end-to-end QoS based distributed multimedia systems; we envision a data collection system where each layer supports mechanisms, interfaces, languages to translate application quality requirements into resource provisioning policies. Designing such a quality-aware data collection system, especially in failure-prone insecure environments is challenging because application quality has many dimensions - e.g. priority, timeliness, security, result quality. To enable effective data collection, these application quality metrics must be translated to data quality metrics that can be used to drive the collection process. Secondly, devising effective collection

protocols requires us to understand how the different elements of quality (priority, security etc.) compete with each other. For example, encryption protocols for secure collection may interfere with timeliness requirements. Thirdly, effective utilization of resources while achieving application quality is key; for this, application quality requirements must be incorporated in making resource allocation decisions. For example, network and storage resources can be dynamically reallocated for fine-grained monitoring of high priority data sources or sources exhibiting rapid change.

We have, in our prior work, explored an end-to-end sensor data collection framework (QUASAR)[24,27,25, 26,33,34] and a collection framework for adaptive network management (the AutoSeC[22,20,21] system). These systems have demonstrated the feasibility of the approach in supporting specific applications. For instance, we have effectively developed (a) quality-sensitive protocols for target tracking of mobile hosts using acoustic sensors[47] and (b) optimized collection of system, network and host information for resource provisioning of multimedia requests in mobile networks. In the RESCUE project, our primary challenge is to explore such a quality-based framework for multimodal data collection.

Privacy in Data Collection: Our focus here is to design data collection systems that are "privacy aware". Spaces instrumented with sensors, cameras, tracking systems based on RFID or similar technologies, network traffic monitors, and related mechanisms can significantly facilitate information collection and situation monitoring during crisis and emergencies; however, such systems infringe upon one of the most basic liberties –right to privacy and are subject to misuse. Our goal is to explore privacy-aware adaptive data collection solutions that empower the subjects, to the extent possible, to control the acceptable loss of privacy, and adapt to the varying needs of the situation.

To make our ideas more concrete, let us consider an example of an airport in the not too distant future that has been partially instrumented with a video surveillance system. This system can be used to monitor the airport personnel/ employees (porters, security personnel, food suppliers, pilots) as they go about their regular tasks in secure areas within the airport. A surveillance system such as the one anticipated in the workplace, might be acceptable (or perhaps welcome) during times when the nation is on high alert but would make most people wary under normal situations. Let us consider a more privacy aware approach that achieves a similar goal. Coupled with the video surveillance system is a RFID based tracking system that identifies employees as a member of a group but maintains their anonymity within the group. The RFID tag can also serve as a key that can be used to encrypt the region of the video where the employee appears. The privacy of the employee is fully preserved as long as they do not violate the access privileges associated with the group. Entry into a region requiring additional clearance will cause the system to trigger a video capture. Similarly, an entry by a person other than an employee without a required RFID tag would be detected by the video would cause an alarm. Note that the approach above maintains anonymity (as long as the subject does not violate any of his privileges), it is adaptive (preserving different levels of privacy under different circumstances), and it empowers the user to control the privacy level (a user can identify himself as belonging to any group he is a member off – the degree of privileges of course differs based on group membership).

The above example represents an instance of the privacy preserving approach to data collection, specifically for video. In practice, information about people/organizations resides in multiple media types and across correlated databases. Current statistical data and inference control techniques are primarily motivated by the need to protect individual identities during the massive production of computerized statistics by public agencies (e.g. Census Bureau)[45]. Statistical data protection mechanisms such as k-anonymity ensure that a tuple in a database cannot be distinguished from at least k other tuples[36]. New multimedia inference control techniques are required to prevent users from composing classified information from multiple media types or derive classified information from pieces of less classified information[13]. In the context of the RESCUE project, we are exploring techniques for combining traditional role-based access control with dynamic inference control techniques to generate a privacy preserving multimedia data collection framework.

4. MULTIMODAL EVENT EXTRACTION AND ANALYSIS

One of the integral components of our research is an end-to-end data analysis system that captures and analyzes multimodal data (e.g., voice and video input from in-field officers and cameras, GPS, sensor data), extracts meaningful events/information from transcriptions, populates key databases, and uses this information in real-time as input into a damage and impact assessment system. Such an analysis system is at the very core of the human-as-sensor concept,

whereby input from the response and rescue teams based on their personal observations, local "eye-witnesses," and personal communications from others on the ground is used for situation assessment. Such input from first responders or informants with the additional benefit of human interpretation, has the potential to provide more robust information gathering capability under adverse circumstances, resulting in more effective and timely response. For example, the envisioned analysis system can be used to enhance the current 911 service during times of disasters, when system resources are under extreme stress. The service can be used to prioritize and/or filter calls from humans, preserving both network and human bandwidth.
The figure above shows various components of the data analysis system we are building. Below we focus on only two of the components - event extraction and analysis. Besides the extraction and analysis framework, the system consists of various other components whose functionality is described below.

- *Data Mediation Infrastructure* – a system that enables dynamic access to heterogeneous databases in the context of diverse tasks and varying access control policies to data. Such a mediation infrastructure can provide access to the knowledge/database useful in information extraction as discussed below.
- *Event Data Management System* -- an XML based model for representing, querying, and analyzing spatio-temporal events.
- *Event Visualization System* -- a GIS based framework that supports spatio-temporal event monitoring in the context of specific tasks such as activity planning, activity monitoring, etc using multi projector displays.
- *Event Mining system* -- that looks for unusual patterns and correlations among events.

Event Extraction from Multimodal Data Streams. One of the first challenges in designing the data analysis system we envision is to accurately extract meaningful events from the multimodal input. Since transcribed voice input is among the most valuable sources of information, we focus on event extraction from text. Information extraction from unstructured text - even in restricted domains - is notoriously difficult[37,9]. In crisis-specific event extraction, these difficulties are compounded due to the inherent noise in voice transcriptions. The approach we are exploring is to first develop a taxonomy of event types with each of which an event model is associated. The disaster response experts may initially specify such event models. Classification approaches based on the event models are used to determine the type of events. Over time, the system will come across event types that cannot be classified automatically. In such cases, the event taxonomy as well as the corresponding models will evolve over time possibly with the help of an analysts/users.

Once an event is appropriately classified, the set of properties associated with the event are extracted. (Almost) all event types share some properties (such as spatial and temporal location of the event). Other properties might be event specific. For example, with a report about a suspicious group of people collected near a facility, one may associate properties such as number and noise level of the people in the group, the approximate age and description. Data extraction accuracy can be significantly improved by exploiting multiple information sources. For example, textual data can be supplemented with location information (from GPS), video data, redundant observations across other text streams, as well as knowledge existing in current knowledge/data bases (available via the mediation component); the context provided by this additional information can be used to restrict the space of possible events and their properties improving the accuracy of the extractor. While context and multimodal data, has the potential of dramatically improving text extraction accuracy, presence of textual data can also benefit multimedia analysis. In particular, video scene analysis can benefit from other information sources such as maps, aerial images, localization information, video from fixed and calibrated cameras and analyzed text. Such input can be systematically incorporated for landmark-based registration of the video acquired by low cost small field of view cameras carried by first responders. Such an approach will improve the interpretability of the video as well as help in improving its perceptual quality[43,23].

Event Analysis: Once basic events have been extracted, they are subjected to variety of analysis tasks (see figure above). For example, a same event may be reported multiple times (e.g., due to multiple observers making and reporting the same event). Alternatively, two events might be linked or correlated in some ways. Similarity matching and data mining techniques properly enhanced to handle spatio-temporal events can be used for this purpose. Furthermore, multiple tightly coupled events might correspond to a larger fused event[28]. These analysis tasks are vital in proper interpretation, ranking, prioritization, and triaging of information to decision makers. We emphasize below two such that highlights the synergistic fusion of our research in social science and information technology – the essence of our multidisciplinary approach.

The first such task is ascertaining the *reliability of human reports*. While human input can play a vital role in crisis response, it must be used with care since informant accounts are known to be highly distorted in a number of ways[4] due to a multitude of factors. Exacerbating circumstances include intentional abuse (whereby a certain category of people may try to negatively affect the system on purpose), confusion and errors in recall, linguistic ambiguity, biases due to prior expectancy, etc. Due to the flow of information through social networks, individual errors in perception may be compounded by factors such as rumors, exaggerations, and differential information availability[7]. Informant reliability can be gauged based on knowledge about the informant, such as a group he/she belongs to, the history and prior reliability reports of similar/related events, the informant's reporting on events known to be exaggerated or caused by rumor, etc. Although sophisticated methods exist for integrating multiple, erroneous informant reports[3], much of the work has not incorporated models of social context in which information is generated. To cope with these conditions, information collection and analysis procedures must be developed which can not only model cognitive and perceptual biases, but also social structural effects.

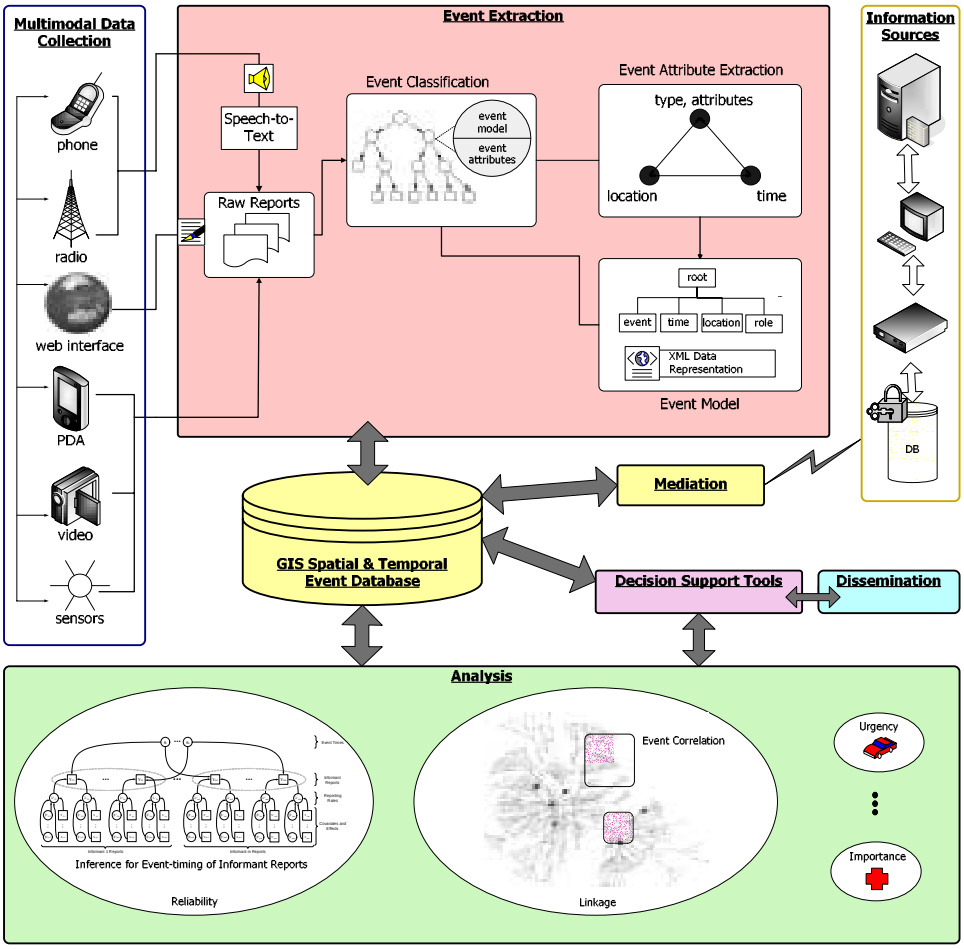

Figure 2: Multimodal Event Extraction and Analysis

Another related task is determining i*mportance and urgency* of an event. Importance and urgency are two tightly coupled factors that play a crucial role in information triaging/filtering. Events could be important but not urgent, or vice versa. Urgency and importance depend upon factors such as the type of event, its attributes, potential impact, the vulnerability of the affected population, etc. These factors ultimately combine to determine utility/loss functions for intertemporal choices regarding crisis interventions.

Within the RESCUE project, we are currently developing an inferential modeling framework for analysis of personal accounts in order to simultaneously infer the true state of a focal event; the social network of information flow among informants; and determinants of informant accuracy, urgency and importance. Some preliminary work in this context has recently been done in the context of network inference[8] and is being built upon here. In conjunction with these inferential models, we are in the process of devising active sampling methods to be deployed by crisis response organizations (e.g., via field agents) for the rapid collection of data for subsequent analysis. By deploying the inferential models and sampling strategies in a variety of event contexts, we can obtain background information regarding informant error processes and initial predictors of informant network structure. Such analyses will help in designing information technology solutions for information triage to responding organizations, and will ultimately be incorporated into decision support systems for organizations involved in crisis response.

5. MULTIMODAL SCENE SYNTHESIS AND VISUALIZATION FOR SITUATION AWARENESS

In any major disaster – whether natural or human generated – the first priority in response is to comprehend as quickly and precisely as possible the nature and scope of the disaster. Situation awareness becomes key in prioritizing the allocation of limited resources, as well as preventing cascading events that can exacerbate the initial effects of the disaster. Situational awareness in crisis has evolved considerably in the last several years. Major factors in this development have been the emergence of a new suite of technologies that allow near real-time imaging of disasters and the integration of these technologies with sophisticated decision-making software tools.

As an event unfolds, data are generated at a rapid pace from a multitude of sources. E-911 calls, on-site video, media reports, remotely-sensed data, and reconnaissance information all contribute to an understanding of what has happened. Until these data are fully synthesized, life and property losses can mount because of delayed or ineffective response. Visualization of data *as it is generated* can be invaluable in helping to make the right decision at the right time. Through an online geographic information system (GIS), we can enable emergency managers to monitor the progress of key incidents and thus, help optimize the deployment of resources. The widespread acceptance of GIS Internet Map Servers (IMS) and the recent introduction of remote sensing into emergency management have made it possible to deploy such systems.

In the past decade, GIS has evolved from a mainframe to a desktop-computing environment. Desktop GIS programs have been widely used in emergency management agencies to create maps that help coordinate response and recovery efforts. Recently, GIS tools have further evolved to the online environment, where data collected in real time can be effectively disseminated to decision-makers. The widespread use of tools like MapQuest has also made the general population more comfortable with online mapping programs. As government organizations begin to make their data available online through IMS, these spatial databases will provide an important context for other data that are collected. In the near future, online GIS systems will allow emergency responders to share spatial data on-line through a web browser. A base map that consists of pre-event satellite imagery; elevation maps; and databases for streets, schools, hospitals, lifelines, and political boundaries, will provide contextual information for an event.

Remotely-sensed data from satellite, manned, and unmanned aerial vehicles can provide an overview of natural and manmade disasters that supercedes the schematic view provided by vector GIS data. In the immediate aftermath of a disaster, commercial satellite imagery can provide important information on what areas have been impacted by the disaster. At a regional scale, comparative analysis of high-resolution optical imagery acquired before and after the event can provide a 'quick-look' damage assessment, distinguishing areas of catastrophic damage. An example of this type of use is presented in Figure 4 for the May 21, 2003 Algerian Earthquake. Following this event, the Quickbird satellite was used to identify areas of collapsed apartment buildings in the northern town of Boumerdes. Similar uses have been documented for wildland-urban interface fires (such as the ones that occurred in October and November of 2003), and after the World Trade Center attack[16].

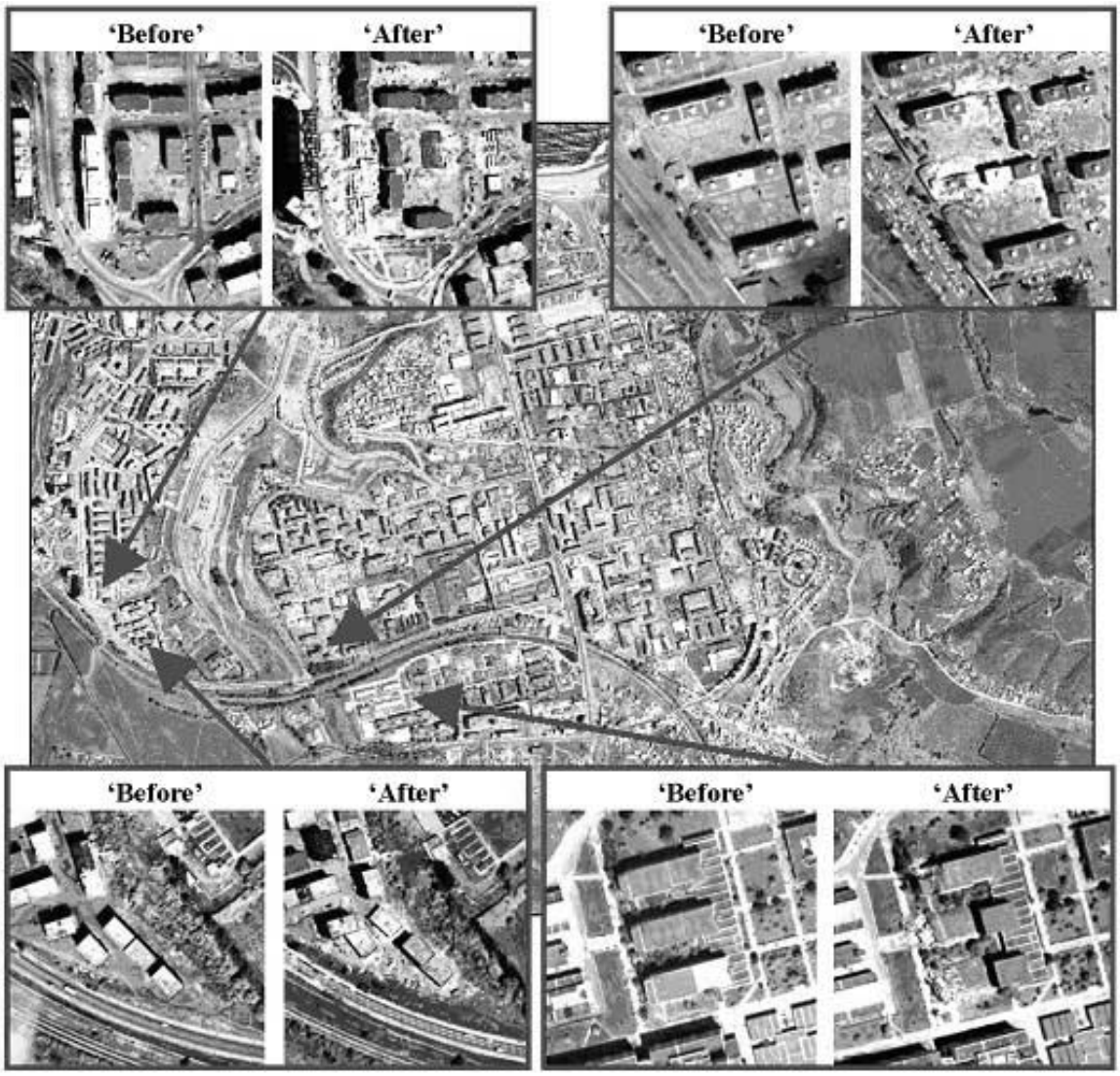

Figure 3. Examples of building damage in the city of Boumerdes, identified by visual inspection of pan-sharpened Quickbird imagery, acquired before and after the 5/21/03 earthquake[2]. Images courtesy of DigitalGlobe, http://www.digitalglobe.com

Although remote sensing technologies and GIS are currently used with great success in emergency management, future online mapping efforts will transcend the boundaries of GIS data, into image processing software and 3D rendering tools. In the RESCUE project, we are exploring mechanisms for seamless transfer of data between databases, GIS, CAD, and image processing programs -- this will significantly improve the efficiency of mapping operations. Where an incident is reported outside a building, for example, an IMS will be able to provide the link to detailed CAD drawings critical for search and rescue, or for responding to other incidents such as terrorist threats. With the advent of E911 cell-phone location and digital imagery captured from cell-phones, spatially located video and photographs can be hyperlinked to points in online maps, so that as an event unfolds, analysts will be able to examine a specific incident from various angles. In our system, loss estimates, post-event imagery, observations of damage, photos, and video linked to GPS data will be uploaded to a server and easily viewed online in the context of a base map. The end-user will be able to view this data as a map interface in a web browser, or on a PDA, enabling informed decisions.

In this context, key areas of research we are exploring include: a) fusing of imagery from different sources, scales and sensors; b) change detection methodologies to quantify local and regional damage; c) rendering of incidents using on-line, real-time 3-D visualization tools; d) rapid geo-referencing of remote sensed data; e) extracting the "built" environment from raw imagery; and e) more effective integration of imagery with decision-making methodologies

6. ACKNOWLEDGEMENTS

We would like to acknowledge all the members of the RESCUE project research team (http://www.itr-rescue.org) and our government partners from the cities of Irvine, Los Angeles, and San Diego, the County of Los Angeles, and the State of California. Many of the ideas presented in this paper have evolved from discussions with them. We would also like to acknowledge the help of Lynn Harris in the preparation of this document. Finally, the research described in this paper is funded through NSF Grant UCI-0331707.

7. REFERENCES

1. E. Auf der Heide, Disaster Response: Principles of Preparation and Coordination (electronic edition). http://216.202.128.19/dr/static.htm, 1989.
2. B. J. Adams, Huyck, C. and Eguchi, R. "The Boumerdes (Algeria) Earthquake of May 21, 2003: Preliminary Reconnaissance Using Remotely Sensed Data", http://mceer.buffalo.edu/research/ boumerdes/default.asp, 2003
3. W. Batchelder, and Romney, A. K., "Test theory without an answer key", Psychometrika, **53(1)**, 1988.
4. H. R. Bernard, Killworth, P., Kronenfeld, D., and Sailer, L. "The problem of informant accuracy: the validity of retrospective data". Annual Review of Anthropology, **13**, 1984.
5. Dan Boneh and Matthew Franklin, "Identity-Based Encryption from the Weil Pairing", In *Proceedings of Crypto 2001*, 2001.
6. J. R. Brouillette and E. L. Quarantelli. "Types of patterned variations in bureaucratic adaptation to organizational stress." Sociological Quarterly, **41**, pp 39-46, 1971.
7. C. T. Butts, "A Bayesian model of panic in belief", Computational and Mathematical Organization Theory, **4(4)**, 1998.
8. C. T. Butts, "Network inference, error, and informant (in)accuracy: a Bayesian approach", Social Networks, forthcoming, 2003.
9. Mary Elain Califf and Raymond Mooney, "Relational Learning of Pattern-Match Rules for Information Extraction", *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99)*, pp. 328-334, 1999
10. L. Comfort, "Shared Risk: Complex Systems in Seismic Response", Amsterdam: Pergamon/Elsevier Science, 1999.
11. P. A. Dinda and D. R. O'Hallaron, "An Extensible Toolkit for Resource Prediction in Distributed Systems", Technical Report CMU-CS-99-138, CMU, 1999.
12. Xuhua Ding and Gene Tsudik, "Simple Identity-Based Cryptography with Mediated RSA", *In Proceedings of RSA Conference: Cryptographer's Track*, April 2003.
13. T. D. Garvey, The inference Problem for Computer Security, SRI International, 1992.
14. Einhorn, H.J. and Hogarth, R.M. "Confidence in Judgment: Persistence of the Illusion of Validity." Psychological Review, **85**, pp. 395-416, 1978.
15. Csilla Farkas and Sushil Jajodia, "The Inference Problem: A survey", *SIGKDD Explorations*
16. C. Huyck and Adams, B.J. "Emergency response in the wake of the World Trade Center attack: The remote sensing perspective", MCEER Special Report Series, **3**, 2002.
17. Hakan Hacigumus, Balakrishna R. Iyer, Chen Li, Sharad Mehrotra, "Executing SQL over encrypted data in the database-service-provider model", In *Proceedings of SIGMOD conference*, pp. : 216-227, 2002.
18. Hakan Hacigumus,Sharad Mehrotra, Balakrishna R. Iyer. "Providing Database as a Service", In *Proceedings of ICDE Conference*, 2002.
19. Q. Han and S. Gutierrez-Nolasco and N. Venkatasubramanian, "Reflective Middleware for Integrating Network Monitoring with Adaptive Object Messaging", Technical Report, University of California-Irvine, 2003.
20. Q. Han and N. Venkatasubramanian. "Aggregation Based Information Collection for Mobile Environments", Journal of High Speed Networking, To appear, 2002.
21. Q. Han and N. Venkatasubramanian, "A Cost Driven Approach to Information Collection for Dynamic Environments", *Proceedings of IEEE MWCN*, 2002.

22. Q. Han and N. Venkatasubramanian, "AutoSeC: An Integrated Middleware Framework for Dynamic Service Brokering", IEEE Distributed Systems Online, 2(7), 2001.
23. K. Huang, and M. Trivedi, "Video Arrays for Real-Time Tracking of Persons, Head and Face in an Intelligent Room," Machine Vision and Applications, Omni-directional Vision, 2002.
24. Iosif Lazaridis and Sharad Mehrotra "Progressive Approximate Aggregate Queries with a Multi-Resolution Tree Structure," *ACM SIGMOD Conference on Management of Data*, May, 2001.
25. I. Lazaridis, K. Porkaew, S. Mehrotra, "Dynamic Queries in Visualization Environments," *In Proceedings of EDBT Conference*, 2002.
26. I. Lazaridis, S. Mehrotra, K. Porkaew, and R. Winkler, "Database Support for Situational Awareness," *Book chapter in the Handbook on Research in Computer Science, Army Research Laboratory*, 2001.
27. I. Lazaridis and S. Mehrotra, "Capturing Sensor-Generated Time Series with Quality Guarantees", *Proceedings of IEEE ICDE*, 2002.
28. Andrew Moore, "Rule-Based Anomaly Pattern Detection for Detecting Disease Outbreaks", *Proceedings of the 35th Symposium on the Interface of Computing and Statistics*, 2003.
29. D. S. Mileti and J. H. Sorensen, "Communication of Emergency Public Warnings". Oak Ridge, TN: Oak Ridge National Laboratory. Report No. ORNL-6609, 1990.
30. N. Miller and P. Steenkiste, "Collecting Network Status Information for Netowrk-Aware Applications", *Proceedings of IEEE InfoCom*, 1999.
31. J. M. Nigg, "Risk communication and warning systems", Natural Risk and Civil Protection, pp. 369-382, 1995.
32. J. M. Podolny and K. L. Page, "Network forms of organization", Annual Review of Sociology **24**, pp. 57-76, 1998.
33. K. Porkaew, I. Lazaridis, S. Mehrotra, "Indexing and Query Processing in Mobile Object Databases," *International Conference on Spatial and Temporal Databases*, 2001.
34. Kriengkrai Porkaew, Iosif Lazaridis and Sharad Mehrotra, "Dynamic Queries over Mobile Objects", *Proc. of EDBT Conference*, 2002.
35. Latanya Sweeney, "K-anonymity: a model for protecting privacy", International journal on Uncertainty, Fuzziness and Knowledge-based Systems, **10(5)**, 557-570, 2002.
36. Pierangela Samarati and Latanya Sweeney , "Protecting Privacy when disclosing Information: k-anonymity and its enforcement through Generalization and suppression".
37. M. Shinozuka, "On-line Damage Identification of Water Delivery Systems", *13th ASCE Engineering Mechanics Conference*, 1999.
38. Sorensen, J. H., "Hazard warning systems: Review of 20 years of progress." Natural Hazards Review **1**, pp. 119-125, 2000.
39. M. Stemm and R. Katz and S. Seshan, "A Network Measurement Architecture for Adaptive Applications", *Proceedings of IEEE InfoCom*, 2000.
40. R. A. Stallings and E. L. Quarantelli, "Emergent citizen groups and emergency management." Public Administration Review, **45**, pp. 93-100, 1985.
41. K. Tierney, Trinet Studies & Planning Activities in Real-Time Earthquake Early Warning: Task 2: Lessons and Guidance from the Literature on Warning Response and Warning Systems. Report prepared by ABS Consulting/EQE International, UCLA Center for Public Health and Disasters, and the University of Delaware Disaster Research Center. Irvine, CA: ABS Consulting., 2000.
42. K. J. Tierney, M. K. Lindell, and R. W. Perry. Facing the Unexpected: Disaster Preparedness and Response in the United States. Washington, DC, Joseph Henry Press, 2001.
43. M. M. Trivedi, I. Mikic, S. Bhonsle: "Active Camera Networks and Semantic Event Databases for Intelligent Environments", *Proc. IEEE Workshop on Human Modeling, Analysis and Synthesis*, 2000.
44. B. Thuraisingham, "The use of conceptual structures for handling the inference problem", In Database Security, **V**, pp. 333-362.
45. L. Willenborg and T De Waal, "Elements of statistical disclosure control", New York: Springer-Verlag, 2001.
46. R. Wolski and N. T. Spring and J. Hayes, "The network weather service: a distributed resource performance forecasting service for metacomputing", Future Generation Computer Systems, **15(5-6)**, 1999
47. X. Yu and K. Niyogi and S. Mehrotra and N. Venkatasubramanian, "An Adaptive Information Collection Middleware for Sensor Network Applications", *ACM/IFIP/USENIX International Middleware Conference*, June, 2003