

**Sixth-Year RESCUE Annual Report –
10/01/08 to 09/30/09**

Responding to Crisis and Unexpected Events

ITR

Collaborative Research: Responding to the Unexpected

National Science Foundation Award Numbers:

IIS-0331707, University of California, Irvine

IIS-0331690, University of California, San Diego

1 OVERALL PROGRESS IN YEAR 6 – EXECUTIVE SUMMARY

RESCUE - *Responding to Crises and Unexpected Events* – is a multi-year research project (with two NSF-granted project extensions) funded under NSF’s Large Information Technology program. This project has involved seven major research institutions located throughout the U.S.; over 40 faculty and senior research staff members with backgrounds in computer science, social science, and engineering; over 60 graduate and undergraduate students; and over 30 government and industry partners.

Year 6 has been primarily a *transitional* year for RESCUE where research projects were either moved toward completion or to other platforms where they could be leveraged with other ongoing research initiatives. Because of the strong desire to leave a suite of legacy products, Year 6 – by design – has focused on three major goals: development of user-friendly artifacts, increased interaction with government and industry groups, and leveraging research findings in order to create new course curricula for both undergraduate and graduate programs. In total, RESCUE has produced 12 large and consequential artifacts; over a dozen formal or informal working agreements with government and industry; and 14 new courses where utilization of information technologies is presented as a key factor in effective disaster response. Based on these achievements, we are confident that the RESCUE project has significantly affected the way in which information technologies are viewed as important tools in responding to crises.

One noteworthy development which occurred in Year 6 was to subsume the RESCUE project under the CERT (Center for Emergency Response Technologies) program at UCI. CERT is designed to be an interdisciplinary effort that brings computer scientists, engineers, social scientists, and disaster science experts together to explore technological innovations in order to deliver the right information to the right people at the right time during a crisis. This development has served to ensure that the mission of RESCUE continues beyond the final year of this project.

This report provides a summary of progress and achievements in Year 6. The report is comprised of four major sections: Executive Summary, Research Progress in Year 6, and Broader Impacts, and Future Research Directions.

1.1 Year 6 Progress

Year 6 of the RESCUE project has been filled with many significant achievements. As in past research summaries, we discuss progress along the lines of six overarching strategies. These strategies were selected because they apply across all research areas, they underscore the importance of our mission, i.e., radically transforming the ways in which emergency response organizations and the public gather, process, manage, use and disseminate information during catastrophes, and finally, they ensure that interdisciplinary education is promoted to the largest extent possible. Progress under each of these strategies is discussed below:

- 1 Structure RESCUE research to focus on a small set of problem-focused, multidisciplinary research projects that are driven by end-user needs and that also offer significant opportunities for scientific contributions.

[Progress in Yr 6: Many of the projects that were introduced at the beginning of this study are now completed. Those that have not been completed have been integrated into other non-RESCUE projects. For example, much research has been done on technologies for situational awareness. Through a grant from the Department of Homeland Security, this research continues with an

emphasis on technologies for fire response. Key challenges or problems that have been addressed by the RESCUE program are listed below along with the research innovation.

- Representation and reasoning with uncertainty during disasters: information extraction and synthesis by exploiting multi-modal data and semantic knowledge (Situational Awareness from Multimodal Input (SAMI)).
- Family reunification during and after disasters: technologies for entity disambiguation and extraction from text (SAMI).
- Robust networking at a crisis site: the San Diego Science Festival (Expo Day) was an opportunity to bring research technologies to practitioners in a large-scale networking deployment with first responders (CalMesh).
- Information sharing architectures: technology for a flexible and customizable policy-driven architecture for information sharing that ensures the right information flows to the right person at the right time with minimal manual human intervention (Simple Web Authentication (SAW)).
- Effective information dissemination: delivering timely and accurate information to the public, first responders and those who are actually at risk or likely to be at risk (MICS, CCD).
- Crisis modeling tools: simulation tools developed to test the efficacy of new and emerging information technologies within the context of natural and manmade disasters (MetaSIM).

2 Create a set of living laboratories and simulation tools that serve as testbeds which mimic “real-world” conditions for regional and incident-level crises and that reflect RESCUE’s mission and objectives.

[Progress in Yr 6: This continues to be a real strength of the RESCUE project. The project continues to operate several major testbeds: the transportation testbed that includes a major portion of the southern California highway network; the UCI Responsphere testbed (www.responsphere.org) that creates a campus-level pervasive environment that supports a variety of networking and sensing capabilities; the San Diego GLQ/Extreme Networking testbed that serves as a mobile living laboratory for deploying and testing a variety of communication technologies in various locations and drill sites, including exercises with the UCSD Campus, San Diego Police and Regional Metropolitan Medical Strike Team; and the Champaign, Illinois testbed where local emergency response officials are actively involved in evaluating information sharing technologies as used in simulated disaster scenarios. In addition to these testbeds, a fifth testbed was introduced in Year 6 that focuses on the safety of firefighters. In this testbed, two different environments are addressed: fires in buildings and wildfires. We continue to demonstrate the usability and value of these testbeds by expanding the number of outside organizations involved with these activities. Notable examples include: a) working with Caltrans on a project to characterize the spatio-temporal signature of traffic accidents using loop sensor data; b) working with local firefighters in Orange County during two exercises in 2008 and 2009 to test out a newly-developed situational awareness system to aid fire department incident commanders during large-scale fires; c) working with the City of Inglewood and the State of California in the Great Southern California Shakeout exercise that involved over 5.1 million participants (see www.shareout.org); d) demonstrating RESCUE projects and the UCSD Responsphere at the 2009 San Diego Science Festival and Expo Day which attracted over 50,000 people; e) developing the first Apple iPhone “app” using the RESCUE automated peer-to-peer traffic system (<http://traffic.calit2.net>); and f) working with local city officials to examine the efficacy of various information-sharing policies based on a threat caused by the derailment of a train carrying hazardous materials through Campaign, Illinois. In addition, the RESCUE project held an important workshop in September of last year with government and school officials to examine the need for more effective emergency information dissemination strategies for schools. This workshop was attended by over 40 key stakeholders in the community.]

3 Develop integrative artifacts that will serve as a legacy for the RESCUE project, thus ensuring that the broader impacts of this five-year research program are realized.

[Progress in Yr 6: Twelve large, consequential artifacts have been developed under the RESCUE project. They are:

1. *CallMesh Networking System:* an affordable mesh networking solution enabling Internet access and team communication where the infrastructure has been compromised or damaged.
2. *Peer-to-Peer Adaptive Information Collection System:* a fully automated peer-to-peer system in San Diego, Los Angeles and the Bay Area (in northern California) that collects and relays highway incident information to the general public and to first responders.
3. *Rich Feeds System/Optiportable:* a system that demonstrates how unconventional and emergent data feeds can be captured, preserved, integrated and exposed either in real-time or after the fact.
4. *Disaster Portal:* an easily customizable web portal that can be (and has been) used by first-responders to provide the public with real-time access to information related to disasters and emergency situations in their community.
5. *SAW:* an authentication technology that eliminates the need for users to remember passwords.
6. *Trustbuilder2:* a flexible framework designed to allow researchers to quickly prototype and experiment with different trust negotiation approaches and processes.
7. *Clouseau:* a system that compiles trust negotiation authorization policies into patterns, translates credentials into abstracted objects, and leverages efficient pattern matching algorithms developed by the artificial intelligence communities to determine all satisfying sets of credentials for a particular policy.
8. *SATware:* a multi-sensor data stream querying, analysis, and transformation middleware that serves as a platform for a sensor-based observation system that addresses situational awareness and privacy.
9. *Crisis Alert – an artifact for customized dissemination:* a system that serves as a research tool to respond to issues identified in the hazard warning literature dealing with the under- or over-response to crises.
10. *MetaSim:* a web-based collection of simulation tools developed to test the efficacy of new and emerging information technologies within the context of natural and manmade disasters, where the level of effectiveness is determined based on how much the response improves with the technology in place.
11. *Responsphere:* IT infrastructure housed within the UCI campus that tests the efficacy of a number of RESCUE technologies, including localization techniques.
12. *Data Repository:* a repository of disaster-related data sets used to evaluate the efficacy of RESCUE technologies.

4 Actively engage end-user community throughout the life of the project to validate the efficacy of the research and to serve as early adopters or testers of research products generated from RESCUE.

[Progress in Yr 6: The RESCUE project has worked closely with its Community Advisory Board to ensure that 1) end-users are aware of the products and artifacts that are being created by the RESCUE research program; 2) that each major project has a government partner to help identify research needs that will lead to implementation opportunities; 3) that RESCUE artifacts will have a home after completion of the current RESCUE project; 4) that proper tests and evaluations are conducted to ensure the efficacy of the research and its products; and 5) that partnerships developed during the current RESCUE project continue beyond Year 6. Some of the key organizations which

have played a major role in this project are: the State of California; the County of Orange; the Cities of Los Angeles, San Diego, Irvine, Ontario and Champaign; the police, environment, health, and safety, and security departments at the University of California (Irvine and San Diego Campuses), and the U.S. Geological Survey. In addition, a number of industry partners have donated their time and equipment to the RESCUE project. As part of the continuity of Rescue and its artifacts, we have created a Center for Emergency Response Technologies (CERT) at UCI. The mission of CERT is to organize and direct research within the emergency response domain and to facilitate further technology adoption of Rescue technologies.]

- 5 Address the social, organizational, and cultural contexts in which technological solutions are adopted and implemented in order to better understand how appropriate technologies can be developed and transferred to users. Create awareness of issues in scientific and industrial communities through workshops, focus groups, panels and open testbeds.

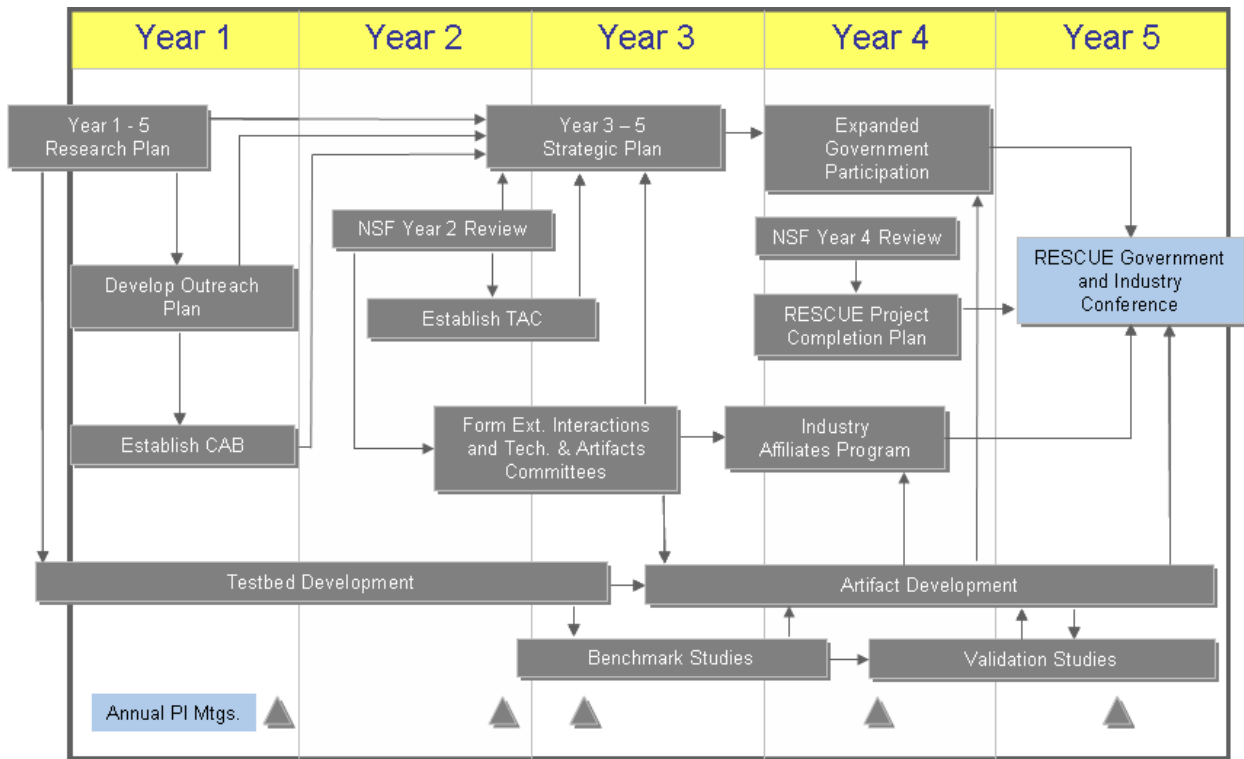
[**Progress in Yr 6:** The RESCUE program has had a core element that focuses on social, organizational, and cultural aspects of crisis response. This effort, led by Kathleen Tierney at the University of Colorado, has concentrated on the use of information and communication technologies among members of the public during disasters. Events that have been studied include the 2001 World Trade Center attacks, the 2004 Indian Ocean earthquake and tsunami, Hurricane Katrina, and the 2007 Virginia Tech shootings. Peer-to-peer communications behaviors in these events highlight the extent to which information and communication technologies are revolutionizing risk communication, information sharing, and collective sense-making with the public during extreme events.]

- 6 Actively engage a broad range of student populations through a multi-course interdisciplinary series on emergency response and focused research projects for graduate, undergraduate students. Create a concerted K-12 outreach effort through demos, lectures and internships. Leverage campus-level programs for underrepresented groups (e.g. California Alliance for Minority Participation and Women-in-CS) to actively recruit minority students.

[**Progress in Yr 6:** RESCUE continues to have a major impact on course curriculum throughout all the universities involved with the project. So far, 14 separate courses on crisis response topics have been developed and taught since the beginning of the RESCUE project. Some example classes include: System Artifacts geared towards First Responders; Special Topics in Information Technology for Homeland Security; and Issues in Crisis Response. In Year 6, the RESCUE Project had nine students graduate with either a PhD or Masters degree. In addition, RESCUE researchers have continued to reach out to the K-12 community by sponsoring high school interns and participating in campus events for high school students.]

1.2 Progress towards Project Milestones

Figure 1 shows the project plan with key milestones. With the exception of the RESCUE Government and Industry Conference, all critical milestones have been met. Significant progress has been made in artifact development and validation, as well as in user implementation. Because of the added emphasis placed in developing and testing project artifacts, and in actively involving more end-users, the Executive Committee of the RESCUE Project decided to postpone the conference to after Year 6. Our current plan is to hold a more limited workshop during the month of November of 2009. CERT – the umbrella that oversees the RESCUE project – has received funding from the Department of Homeland Security to hold a two-day workshop to focus on Information Technologies and their role Crisis Response. We plan on using this workshop both as platform for showcasing RESCUE artifacts as well as identifying new and important research directions.



Progress in each research project and artifact development effort is described in Sections 2.1 through 2.3 of this report. Key highlights include: 1) the emergence of cohesive multidisciplinary teams for each integrative research project; 2) the establishment of key partnerships with the user community for all the RESCUE artifacts; 3) the application of technologies in real-world environments, e.g., Disaster Portal in 2007 southern California wildfires, participation in the Great Southern California Shakeout exercise in 2008, and Peer-to-Peer traffic information dissemination testbed deployed in San Diego, Los Angeles/Orange Counties, and Bay Area; and 4) validation of research results using simulation methods, feedback from end-users, and application in actual events.

RESCUE is also continuing its push to engage industrial partnerships in the project. Numerous discussions, presentations, and demonstrations had been held with a number of companies including Boeing, Apani Network, IBM, Convera, Sun Microsystems, Printonix, Motorola, Ether2, IBM, Conexant, Yahoo, SBC, Asvaco, VDT, CDWG, Psion Teklogic and AMD, resulting in substantial equipment donations, new strategic alliances, and partnership discussions and/or non-disclosure agreements.

Efforts to increase participation of the user community in RESCUE include establishment of key partnerships with response organizations in the context of artifacts (e.g., City of Ontario for the Disaster Portal (www.DisasterPortal.org/Ontario) and expansion of the Community Advisory Board (e.g., addition of new members representing the Orange County Fire Authority & the Boeing Company). In consultation with Ellis Stanley, the Chair of the Community Advisory Board, the project has initiated a strategic expansion of CAB to increase the likelihood of national-level impact and project continuity beyond Year 6. Key researchers within RESCUE are in discussions with representatives from the Centers for Disease Control, the Federal

Emergency Management Agency, the Office of Homeland Security for California, the Department of Homeland Security – Emergency Management, Los Angeles County – Department of Health Services/Bio-Terrorism/Preparedness and Response with regard to actively engaging their members in the project's CAB. RESCUE researchers have also deployed their technologies during several drills, including full-scale emergency response exercises.

Further efforts resulting in a strengthening of the industry/academic strategic partnerships include the creation of a school-based Center for Emergency Response Technologies (CERT). This Center, housed in the Donald Bren School of Information and Computer Sciences, was created as a vehicle for Rescue continuity and to facilitate our efforts to transfer technologies from the research laboratory and testbeds to the first responders and communities. Through CERT, we are pursuing licensing opportunities of Disaster portal technologies, most notably, with the city of Champaign Illinois and Rancho Cucamonga California.

2 RESEARCH PROGRESS IN YEAR 6

2.1 Research Projects

Project 1: Situational Awareness from Multimodal Input (SAMI)

Project Summary

The SAMI project is focused on research and technology development to realize the next-generation of situational awareness systems. Our ultimate goal is to develop an end-to-end situational awareness “engine” that can be used for particular situational awareness applications, primarily in the area of disaster response. Situational awareness, in the context of disaster response can be broadly viewed as consisting of the *past* (knowledge), *present* (comprehension), and *future* (prediction) state of the resources, infrastructure, entities (people), and the incident. Our research is aimed at addressing technical challenges in three key areas, namely information extraction and synthesis from raw sensory data, situational data management, and analysis and visualization technologies for decision making and support. The key technical approach we have investigated in extraction and synthesis is exploitation of (a) multimodality, and (b) semantic knowledge in extracting and interpreting data. The key challenges in situational data management addressed include representation and reasoning with uncertainty. In the context of data analysis our focus has been on understanding patterns of human behavior over time. Examples include analysis and understanding of Web access logs, event detection and prediction with vehicular traffic and accident data, and classifying human activities from low-cost observation modalities used for ubiquitous sensing such as RFID, video, et cetera.

Our work in the last several years has significantly increased our ability to extract and synthesize vital information from raw multi-modal data streams that are often available during disasters. The use of semantics has resulted in pioneering approaches that help to address traditionally complex data extraction problems. We have developed probabilistic models that learn patterns of human behavior that are hidden in time series of count data.

Many of the research outcomes of this project have been incorporated into RESCUE artifacts with Disaster Portal being the most prominent. For instance, technologies for entity disambiguation and extraction from text are the driving force behind the family reunification component of the portal. Likewise, predictive modeling techniques are the driving force behind occupancy forecasting.

Activities and Findings

Overall, and in Year 6, we have made significant research progress in the many different areas in SAMI. Many of the different projects are now at a sufficiently mature stage in the research. Below we provide a summary of our progress in the areas of the Disaster Portal, The fire incident command board (FICB), Sensor data visualization, Localization technologies, Situational awareness information integration, Semantic information extraction from text, Situational awareness from text, and Automated event detection.

Major findings for Year 6 include:

- Completely developed approach to integrating integrity constraints in uncertain databases.
- WeST toolkit developed for data disambiguation.
- Developed toolkit for automated image tagging.

- Developed next generation techniques for person localization by combining multiple location detection techniques.
- Developed next generation information integration system for SA applications.

Products and Contributions

FICB

The Fire Incident Command Board (FICB) is a situational awareness system intended to aid fire department incident commanders during emergency response activities. It accomplishes this by providing integration of a variety of data streams into a single, easy to use dashboard. The data provided via the FICB includes data collected in real time from diverse sensors (both fixed and mobile) deployed at the incident scene (e.g. video cameras, speech / other audio, physiological sensing, location sensing), as well as precompiled data (e.g. GIS / maps, building floor plans, hazmat inventories, facility contact information). The FICB provides the ability to monitor, query, and store the data from these diverse sensors in a user friendly manner.

A prototype implementation of the FICB has been created. The prototype has been implemented by combining elements of some existing systems developed by RESCUE (e.g. SATware streams system) with new components (EBox prototype, computer aided dispatch system). The SGS-Darkstar toolkit has been used as an integration platform in order to implement the FICB incident model which is comprised of the elements of the firefighting domain such as personnel, equipment, physical infrastructure, etc. FICB merges the data streams appropriately so that they may be represented with the relevant portions of this model in the user interfaces in order to provide a view of the overall situation to the incident commander in real-time.

We have performed several assessments of the FICB. These include a situational awareness assessment using the SAGAT methodology conducted during an exercise held at UCI on May 12th. In the simulated hazmat incident, one IC had access to the SAFIRE system while the other relied on more traditional technologies (radio). The results of this experiment are being analyzed for inclusion in an article or technical report. A SAFIRE usability study was conducted at the May 17th SAFIRE firefighter forum as part of a tabletop exercise, in order to evaluate improvements in decision making due to enhanced situational awareness provided by the SAFIRE system. Results indicate a high degree of both usability as well as decision making impact (by virtue of increased information and enhanced situational awareness) in those respondents with Incident Command experience. Qualitative feedback was also captured in the study.

SATViewer

SATViewer is a system for visualizing information captured in the SATware database which stores data collected from multiple sensors in the Responsphere infrastructure. The purpose of this project is to provide an interface for saving sensor data and visualizing it after the saving session. The system is implemented for SATware middleware and uses installed sensors for such middleware. The key challenge in designing a visualization tool for such a pervasive system is that of information overload - limitations in user perception and in available display sizes prevent easy assimilation of information from massive databases of stored sensor data. For instance, in the Responsphere setting, there are over 200 camera sensors deployed at two buildings; even a very simple query for monitoring these buildings will have to visualize 400 streams (audio/video) for any given time.

This work attempts to address the information overload problem using two key strategies:

- (i) Ranking relevant sensor streams
- (ii) Summarization of selected sensor streams

In particular the focus is on the capability to 'link' multimedia data to a spatial region, and to a specific time, as well to synchronize diverse sensor streams so as to visualize them effectively. The application allows us to record data from sensors using a map or a list of the sensors. Moreover, it allows for querying saved sensor data by specifying sensors of interest and the time interval. Finally it allows adding new sensors in the system.

Localization framework

The problem we address is the definition of a general framework in which any location detection technique can fit, being modeled as a generic location component. The main purpose of such a framework would be to answer location queries with the best trade-off between accuracy and precision, choosing the fittest location technology or the best combination of technologies to solve each query. The following steps were taken to address the problem:

- Definition of a localization component interface, which is a model for a generic localization technology. The component is modeled as a black-box which is able to provide a non deterministic location prediction modeled as a probability mass function (PMF) on a set of locations.
- Definition of a taxonomy of location queries which best applies to the most common localization problems. All types of queries were then formalized inside the framework.
- Definition of an aggregation algorithm capable of elaborating answers coming from one or more localization components and aggregate them. Answers from different components are sorted by their relevancy as regards the current query. The answers are then progressively aggregated into one single PMF using Bayesian inference. The algorithm detect when an answer does not bring improvement to the global PMF, and it discards it.

The framework has been implemented for SATware middleware. A Nokia N95 smart-phone was used to provide information to the implemented components. Several components and the aggregation algorithm were incorporated in a number of SATware mobile agents. SATware middleware is suitable to host the defined localization system, because the modularity of the framework, defined formally, is preserved. Several localization techniques have been adapted to fit in the framework (i.e. to provide probabilistic answers). Components involving the following localization techniques were implemented:

- Wi-Fi fingerprinting: a database matching technique based on wireless LAN. . This technique involves a nearest-neighbor search on a data space of previously collected signal strength readings (fingerprints). Distances from the fingerprints in the data space were used to calculate a probability for each location.
- GPS: the coordinates provided by a GPS receiver were used to build a Rayleigh distribution based on the accuracy value provided by the receiver itself.
- Bluetooth: Bluetooth technology was used to implement a simple anchor-based proximity localization system. This component outputs a uniform truncated PMF around fixed Bluetooth anchors.
- Speech: a simple natural language parser was written to extract location information from recognized speech. These information are used to retrieve PMFs which were previously written and which are stored in a database
- Historic: this component uses previously calculated PMFs as a prior. Movement information coming from an accelerometer is also used to better exploit location information from the past.

EBox

Our work on the FICB system as a decision support system has motivated our work in progress on the “Software EBox”. Essentially, the EBox is an information integration system targeted towards situational awareness (SA) applications. In literally any SA system, including one for fire response, one requires access to a variety of different data of different types from different data sources. For instance in the context of FICB it is beneficial to have (integrated) access to information such as maps of the area, floor plans of various buildings, knowledge of building entrances and exits, knowledge of the presence of hazardous materials and chemicals, key personnel at the site and their contact information, etc., Besides many urban sites these days typically have buildings or other structures instrumented with sensors, such as say surveillance cameras, that can also be exploited for real-time situational information.

Information integration technology indeed addresses some of the problems in achieving the kind of information integration envisioned above, however (i) The integration assembly can be slow and requiring a high level of expertise, (ii) There is no mechanism to incorporate real-time sensors, and (iii) Problems of access control and authorization in such a data sharing environment have not been addressed. In the commercial space there are several vendors offering custom information assembly solutions, however the cost of such solutions is upwards of \$2M per site.

We are developing the EBox technology which is a general purpose information integration system for SA applications. The novel features of the EBox framework include: (i) A new data integration approach for rapidly building new applications without requiring much assembly expertise, (ii) An approach for integrated access to real-time sensors via the SATWare framework and (iii) Mechanisms for access control in data sharing.

Semantic Extraction from Text

We have continued our work on techniques for automated extraction and synthesis of information of text. In prior work in RESCUE we developed an approach for information extraction using the framework of uncertain databases. Motivated by the application of information extraction, we have further explored the incorporation of semantic information – in the form of database integrity constraints, for improving the quality of extraction. The incorporation of integrity constraints into uncertain databases is a problem of general interest, as the additional knowledge in the constraints can be used to reduce the uncertainty in the original database – thus improving the accuracy of retrieval over the database. The problem of incorporating even very simple constraints into uncertain databases however is intractable as we identified. We have developed an approach for the efficient and effective incorporation of integrity constraints into uncertain databases using approximation techniques While addressing the general problem, we have evaluated this approach in the specific task of information extraction and demonstrated it to be very effective in ultimately improving the quality of extracted data.

Situational Awareness from Text

Situational awareness (SA) applications monitor the real world and the entities therein to support tasks such as rapid decision-making, reasoning, and analysis. Raw input about unfolding events may arrive from variety of sources in the form of sensor data, video streams, human observations, and so on, from which events of interest are extracted. Location is one of the most important attributes of events, useful for a variety of SA tasks. We considered the problem of reaching situation awareness from textual input. We proposed an approach to probabilistically model and represent (potentially uncertain) event locations described by human reporters in the form of free text. We analyzed several types of spatial queries of interest in SA applications. We designed techniques to store and index the uncertain locations, to support the

efficient processing of queries. Our extensive experimental evaluation over real and synthetic datasets demonstrated the effectiveness and efficiency of our approach.

Entity Resolution

Entity Resolution (ER) is an important real world problem that has attracted significant research interest over the past few years. It deals with determining which object descriptions co-refer in a dataset. Due to its practical significance for data mining and data analysis tasks many different ER approaches has been developed to address the ER challenge.

We proposed a new ER Ensemble framework. The task of ER Ensemble is to combine the results of multiple base-level ER systems into a single solution with the goal of increasing the quality of ER. The framework we proposed leverages the observation that often no single ER method always performs the best, consistently outperforming other ER techniques in terms of quality. Instead, different ER solutions perform better in different contexts. The framework employs two novel combining approaches, which are based on supervised learning. The two approaches learn a mapping of the clustering decisions of the base-level ER systems, together with the local context, into a combined clustering decision. We empirically study the framework by applying it to different domains. The experiments demonstrated that the proposed framework achieved significantly higher disambiguation quality compared to the current state of the art solutions.

Web People Search

Currently, searches for webpages of a person with a given name constitute a notable fraction of queries to web search engines. Such a query would normally return webpages related to several namesakes, who happened to have the queried name, leaving the burden of disambiguating and collecting pages relevant to a particular person (from among the namesakes) on the user. To address this challenge, we developed a Web People Search approach that clusters web pages based on their association to different people. Our method exploits a variety of semantic information extracted from web pages, such as named entities and hyperlinks, to disambiguate among namesakes referred to on the web pages. We demonstrated the effectiveness of our approach by testing the efficacy of the disambiguation algorithms and its impact on person search.

We also have investigated a new server-side WePS approach. It is based on collecting co-occurrence information from the Web and thus it uses the Web as an external data source. A skyline-based classification technique is developed for classifying the collected co-occurrence information in order to make clustering decisions. The clustering technique is specifically designed to (a) handle the dominance that exists in data and (b) to adapt to a given clustering quality measure. These properties allow the framework to get a major advantage in terms of result quality over all the 18 methods covered in the recent WePS competition.

Speech Processing

Associating textual annotations/tags with multimedia content is among the most effective approaches to organize and to support search over digital images and multimedia databases. Despite advances in multimedia analysis, effective tagging remains largely a manual process wherein users add descriptive tags by hand, usually when uploading or browsing the collection, much after the pictures have been taken. This approach, however, is not convenient in all situations or for many applications, e.g., when users would like to publish and share pictures with others in real-time. An alternate approach is to instead utilize a speech interface using which users may specify image tags that can be transcribed into textual annotations by employing automated speech recognizers. Such a speech-based approach has all the benefits

of human tagging without the cumbersomeness and impracticality typically associated with human tagging in real-time. The key challenge in such an approach is the potential low recognition quality of the state of the art recognizers, especially in noisy environments. In our work we explored how semantic knowledge in the form of co-occurrence between image tags can be exploited to boost the quality of speech recognition. We postulate the problem of speech annotation as that of disambiguating among multiple alternatives offered by the recognizer. An empirical evaluation has been conducted over both real speech recognizer's output as well as synthetic data sets. The results demonstrate significant advantages of the proposed approach compared to the recognizer's output under varying conditions.

In addition we have studying the retrieval aspect of the problem. Modern data processing techniques such as information extraction, entity resolution, data cleaning, and automated tagging often produce result consisting of objects whose attributes may contain uncertainty. This uncertainty is frequently captured in the form of a set of multiple mutually exclusive value choices for each uncertain attribute along with a measure of probability for each alternative value. In our work we study the problem of retrieval on top of such a probabilistic representation. Unlike existing approaches on probabilistic databases (Trio, MayBMS, etc.) that provide a probabilistic answer to user queries, our goal is to offer a deterministic answer that maximizes an application specific quality measure. In particular, we explore application metrics such as precision, recall, and F_a in the context of retrieval queries. Deterministic answering of retrieval queries over probabilistic representation is a problem similar in nature to that of index term selection studied extensively in the context of document collections -- it poses similar tradeoffs between measures such as precision and recall. We formalize the resulting optimization problem, and develop techniques for optimal answering of queries. The empirical evaluation over three different domains demonstrates the significant advantage of the proposed solution over the state of the art techniques.

Event Detection

In this last year we began a collaboration with a group of transportation engineers from the Institute of Transportation Studies (ITS: <http://www.its.uci.edu/>) at UC, Irvine. Together, we are partnering with the California Department of Transportation (Caltrans) on a project to characterize the spatiotemporal signature of traffic accidents using loop sensor data. We have extended the algorithms developed during the last several years of the Rescue project to find normal traffic patterns and detect and characterize unusual traffic conditions using additional measurements. The extended model uses both a flow measurement (a count of vehicles passing over the sensor) and an occupancy measurement (a measure of the fraction of time the sensor is covered by a vehicle). We have applied this new model to a large group of sensors on several southern California freeways. This extended model is sensitive to smaller changes in traffic flow, and has led to interesting analysis of delay due to traffic incidents. We are currently consolidating our findings into a report that will be submitted to a transportation journal.

We also started a second project this last year where we attempt to predict the flow profile of freeway on and off ramps using census information. There are many stretches of highway where there are no functional loop sensors. Large scale problems such as dynamic population density estimation require inference over this missing data. We imported census data into a Google maps application, created boundaries around the geo-locations of the ramp sensors, and extracted information relative to each sensor. We have developed a model of the ramp profile given information about the area surrounding the sensor. We are currently in the analysis phase of this project.

We are also working on another extension to the model where spatial links are modeled explicitly, and will soon be applying the event detection models on web traffic data.

Sensor Data Collection Scheduling

A distributed camera network allows for many compelling applications such as large-scale tracking or event detection. In most practical systems, resources are constrained. Although one would like to probe every camera at every time instant and store every frame, this is simply not feasible. Constraints arise from network bandwidth restrictions, I/O and disk usage from writing images, and CPU usage needed to extract features from the images. Assume that, due to resource constraints, only a subset of sensors can be probed at any given time unit. This paper examines the problem of selecting the “best” subset of sensors to probe under some user-specified objective - e.g., detecting as much motion as possible. With this objective, we would like to probe a camera when we expect motion, but would not like to waste resources on a non-active camera. The main idea behind our approach is the use of sensor semantics to guide the scheduling of resources. We learn a dynamic probabilistic model of motion correlations between cameras, and use the model to guide resource allocation for our sensor network.

Although previous work has leveraged probabilistic models for sensor-scheduling, our work is distinct in its focus on real-time building-monitoring using a camera network. We validate our approach on a sensor network of a dozen cameras spread throughout a university building, recording measurements of unscripted human activity over a two week period. We automatically learnt a semantic model of typical behaviors, and show that one can significantly improve efficiency of resource allocation by exploiting this model.

Project 2: Robust Networking and Information Collection

Project Summary

The main objective of this project is to provide research solutions that enable the restoration of computing, communication, and higher-layer services at a crisis site in a manner that focuses on the needs and opportunities that arise proximate to the crisis (in both time and space dimensions). Commercial systems are often based on assumptions that fall apart during a crisis when large-scale loss of power and destruction of antenna masts and servers are common. In addition, self-contained relief organizations that arrive at a crisis site often carry communication equipment that fail to interoperate, are inadequate for the needs at the scene, and may even interfere with each other making the task of forming an ad-hoc organization harder. In summary, the challenge is to compose a set of research solutions to assist in crisis response that is designed to serve the dynamically-evolving situation at the crisis site.

Activities and Findings

The highlight for the UCSD RESCUE team this past year was our participation in the San Diego Science Festival's Expo Day in April, which included the largest deployment of CalMesh ever. In addition, technical progress was made in a number of areas and refinements of previous technologies continued.

San Diego Science Festival - Expo Day

Researchers from the UCSD ResponSphere and RESCUE projects participated in the grand finale of the month-long San Diego Science Festival: Expo Day at Balboa Park. More than 50,000 people attended the event held on Saturday, April, 4, 2009, which featured 200+ exhibition booths. Organizers called it "the largest one-day science gathering ever in the United States."

More than three dozen researchers (PIs, faculty, staff, postdocs, graduate and undergraduate students) were on-hand at Balboa Park to run demonstrations, provide information to visitors, and manage the wireless network and the many experiments. A wide variety of measurements were taken, both on the day of the event and as preparation and reference samples. The huge crowds and changing network environment over the day-long event served as an excellent live real-time testbed for evaluation how our wireless technologies performed in areas with heavy pedestrian and cellular traffic.

This activity has resulted in at least one new collaboration. The Balboa Park Online Collaborative (BPOC) has asked the Robust Networking team to help them wirelessly connect museums in Balboa Park to help create a more unified management and online data sharing and collaboration systems.

The Calit2 San Diego Science Festival Photo Gallery is available here:
http://projects.calit2.net/gallery/main.php?q2_itemId=3436.

Products and Contributions

1. CalMesh

At the request of the San Diego Science festival organizers, we set up a CalMesh ad-hoc wireless network covering most of the exhibit areas along the Prado (around 40+ booths), many of which would otherwise have not had WiFi connectivity or only spotty access to the Internet, much less high-speed web access. The coverage area was along the eastern part of the Prado from the Lilly Pond to the Fountain. This was the largest deployment of CalMesh ever. See Figure 1, below.

The deployment also provided an excellent opportunity to collect data, to further our research on communications in cases of emergency. Many experiments were conducted and measurements taken on the network and surrounding environment. The CalMesh network was linked to the Internet via the High Performance Wireless Research and Education Network (HPWREN) and a commercial provider, Sky River, both had access points on the Natural History Museum's roof. An access bonding solution from Mushroom Network Inc. aggregated the two channels together. Combined, they provided 45 megabits per second of bandwidth.

For Expo Day, 10 nodes total were deployed: 8 Mesh Nodes (one open access point [AP]) and 2 Stand-alone gateway [GW] nodes (3G access and inside the Natural History Museum). Each box (node) is a WiFi Access point, supporting data, voice and video applications with a 15-hour battery and GPS. The cost is about \$1000 per node. CalMesh uses a variety of ways to connect to the Internet: Wired (DSL, Ethernet, Cable), WiFi Hotspots, Cellular (3G), and Satellite access. CalMesh has multiple interfaces:

Two WiFi interfaces: 1) One for Mesh connectivity only where the actual routing protocol runs. 2) One for Access Point only, where legacy WiFi clients can connect.

Two Ethernet interfaces: 1) Internet connectivity (gateway) and 2) Wired access to servers.

One Cellular interface: Currently USB based cellular data devices are supported



Figure 1. CalMesh deployment in Balboa Park (approximate node locations). The lily pond is where CalMesh GW4 is located (on the left of the photo), the fountain is on the right, surrounded by a circle of wide concrete. The Prado is the area in between.

Measurements: The internet access providers (SkyRiver and HPWREN) stored traffic logs at their respective network operations centers (NOCs) that are available to us. Each CalMesh node stored traffic traces by the use of tcpdump where the IP addresses and type of traffic of each user can be recorded. In addition, each node stored the number of received and sent packets, sampled at programmable time intervals. In addition, some of the access points were open for the general public to study how such traffic impacted the network performance. Signal to noise ratio measurements on the move were made along the node placement route by walking up and down the Prado collecting location-aware data. Monitoring of the aggregated traffic through the backhauls was also performed to give us an idea of total bitrate for event, quality and possibly type of traffic. Data analysis is ongoing.

2. Gizmo

The big hit among the crowds of families in attendance was Gizmo, Rescue's family of autonomous multi-radio devices that serve as adaptable and reliable research platforms on wheels to deploy different technologies and gather sensor data in real time. They are designed to transport cameras, other sensors, and wireless access points to and around disaster sites in order to get communications going again in an emergency. They are configured with 2.4GHz WiFi communications and control, 75MHz PCM control only, Video pan tilt 640x480 60fps, and Audio ADPCM.

This year a mobile touchscreen kiosk based on the Gizmo technology (the Nokia Siemens Networks Gizmo operator) was developed and made its public demo debut in the Calit2 Expo day booth. Gizmoi was upgraded for use in rough terrains. Also, a "Backpack Cam" video feed device, which could be worn while walking around the festival, was created from a Gizmo-based webcam.

3. Condor

The project WiFli CalMesh Condor was originated with the idea of expanding and improving our network deployment capabilities. As it is often the case, in emergency response environment, not every area is accessible nor is every terrain smooth. Therefore, deployment becomes really

complicated and time consuming. Gizmo and MOP projects helped in facilitating the deployment of the CalMesh network. Even though the newly upgraded Gizmo can go through rough terrains, it will always be limited by possible obstacles. The WiFli CalMesh Condor offers a faster and dependable system.

4. BlueMap

Bluetooth Malware Analysis and Prevention (BlueMap). We deployed 15 Bluetooth sensor nodes within the Expo area to record Bluetooth devices in the vicinity of each sensor node and store this data in a central database that can later be used to infer, among other things, the potential spread of mobile phone virus during a highly populated event. Analysis of malware spread between mobile phones over short-range wireless links (Bluetooth, WiFi etc.) helps gain a better understanding of the potential for proximity driven malware spread and develops means to both detect and prevent such outbreaks. We saw 440 unique Bluetooth devices during the measurement period between 3 p.m.–6 p.m.

5. CogNet Monitoring

On Expo day we conducted network monitoring using CogNet. We observed all the channels in the 2.4GHz and 5.2GHz spectrums and sampled network traffic from all the channels. We measured the channels used for CalMesh continuous capture. The results from this measurement are mainly focused on studying the network behavior as well as studying the wireless environment behavior. In addition, we visualized the network traffic environment on a portable visualization display live at the event.

Testbed and Data Stats: Four locations were used for the traffic monitoring with a visualization node in the Calit2 booth. Two types of devices: 1) CalNodes for sampling traffic across 11 channels in 2.4 GHz (1:11 ratio) and 2) CogAP for visualization of Live traffic and Network state prediction using Neural Network based Cognition engine. The experimental time period was 8:30 a.m. to 6:00 p.m. on Festival Day and noon to 1:15 p.m. on reference days. The sampling CalNodes collected 7,076,853 packets total. The complete capture CalNodes collected 15,012,407 packets on selected channels (1, 6, 10, 11).

Results Summary: Significant impact on the cyber world on Expo day when compared to reference sampling prior to and afterward, due to the physical world activity, was observed across several parameters (amount of traffic, number of clients, packet length) which confirmed our past observations made in previous drills. *Key observations made:* Time line of active client association to CalMesh nodes and Spatio-temporal traffic characterization of 802.11 b/g channels (11 channels in 2.4 GHz spectrum).

6. Cellular Network Monitoring

EVDO Cellular Network Logging. Cellular network monitoring was conducted before, during, and after the event using the Qualcomm CAIT tool. We logged physical, MAC, and upper layer information such as signal strengths, power control, frame error rates, RLP retransmission, data rates, handoff status to observe differences in the mobile users perceived performance between light load and high load conditions and to observe effects of excess number of users in an EVDO cellular network from mobile point of view.

Observations: Forward Link: 1) Received power is significantly lower during the science fair (Active sector selection is dynamic) and 2) Mobile switches to more distant sectors due to load in the nearby sectors. Reverse Link: 1) Increased transmit power and 2) Mobile needs to

transmit with more power due to interference in the loaded sector. Also, the mobile's battery will drain faster to maintain similar QoS on crowded day as compared to a normal day

7. Laser Scanning & Wireless Network/Coverage Simulation

During the past year, the laser scanning equipment and processing continue to be extremely important and are used frequently. They are currently, or have been, used in about 10 different projects (with multiple subprojects), ranging in purpose from assessment of wireless signal transfer, to cultural heritage, art history, archaeological exploration, and structural analysis of historic buildings and locations. The data across projects has been used for the development of new, rapid meshing techniques.

While cellular networks are deployed based on an extensive design process, in emergency scenarios it is impossible to have a long design/performance analysis phase before the rapid deployment. The main idea for this experiment was to define and create a process to use a 3-D laser scanner and quickly scan the deployment environment, input it into the wireless simulation tool and find the optimum locations for temporary deployment of CalMesh nodes. We used the point-cloud exemplar to model the wireless propagation environment to give a view of theoretical behavior of the system.

Experiment design: The MAPTEK I-Site 4400 3-D scanner was used before the event, and the data input into the EDX Wireless simulation tool to determine the best locations for CalMesh nodes to maximize coverage and capacity for wireless network during the event. During the experiment the inter-operability between 3-D scanner (formats) and EDX tool needs to be tested. Also, real signal strength and operability of the network during event must be observed to validate accuracy of design.

Outcome: The main parking area scanned successfully and relevant data collected. The size of data was around 500 Mb which was more than other experiments. This was due to extended area we intended to simulate during this event. The output format was ESRI/Arc View ASCII grid format. Since color accuracy was not important for our experiment the total scanning process took less than a day. During scanning process, 6 different points set and scanned before final interpolation to extract the actual data. The data converted to EDX .201 format using EDXCV conversion tool before feeding to EDX. The outcome of simulation indicates optimum locations for our CalMesh nodes during the event.

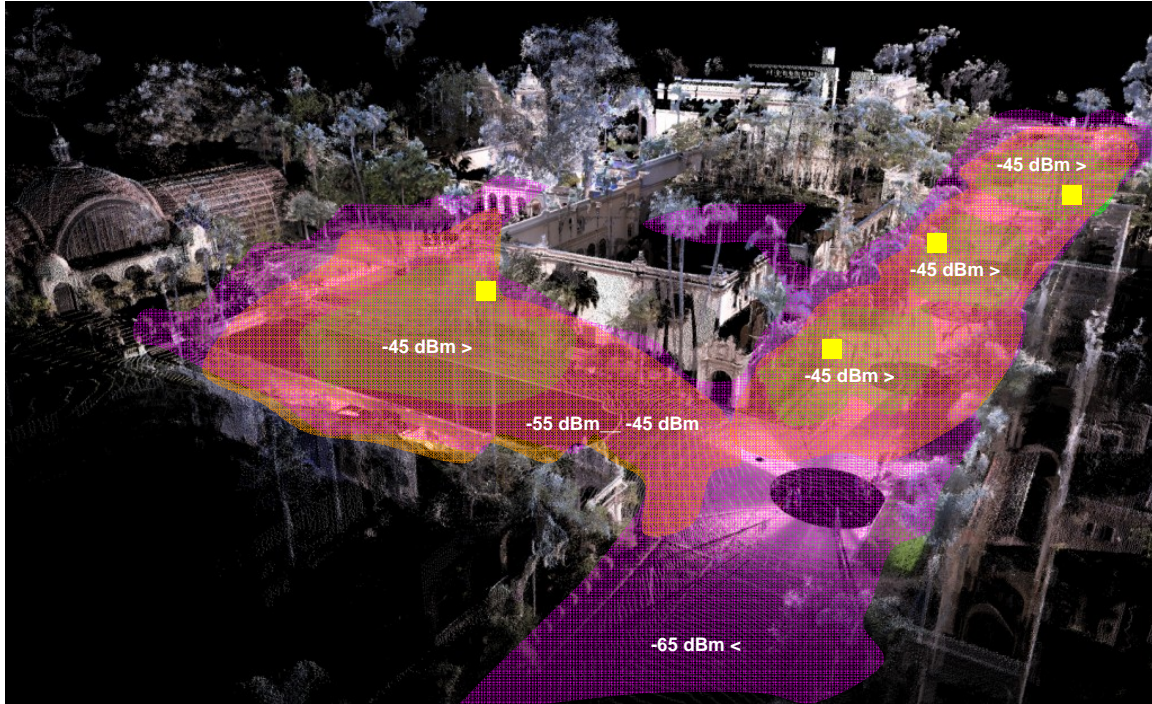


Figure 2. Overall coverage area

Observations and Analysis: As it is shown in Figure 2, most of the area had a good signal strength (>-45 dBm) and at the corners there were areas with moderate signal strength (<-55 dBm). Based on this simulation results, we were confident that CalMesh locations are optimum locations to provide the best coverage for all exhibits in the central areas of exhibition as it was identified before. Lack of color information for our simulation caused a huge reduction in scanning time; six points of scanning and each about 45 minutes and smaller file size; 500 mega versus terra bytes. EDX was able to consider interference effect of the nodes on each other during simulation.

Conclusion: The value of a two-steps design and deployment for wireless networks in emergency cases was shown during the drill. Although after any disaster or unexpected events there is no time for a systematic wireless deployment, but a real-time simulation/deployment approach can be useful to maximize the coverage and reliability in these types of deployments. We did not use a real-time scan/simulation/deployment approach for this, however, different components tested separately and their interoperability tested in a non real-time scenario.

8. Wireless Mesh Network Electromagnetic Interference Analysis

We have deployed dozens of wireless mesh network nodes at various locations and activities throughout San Diego, CA, including Expo Day this year. At the festival, we had complete control over our own wireless mesh network nodes (Cal_Mesh_Demo), and had access to the Balboa Park WiFi access points and cooperation back at UCSD. In this way, we could supplement the connectivity of our mesh network with the UCSD WiFi infrastructure. Previously, nearby WiFi access points presented a significant source of interference since they would consume air-time capacity without sharing their frequency spectrum.

Detailed measurements of electromagnetic wireless spectrum were performed during and before these activities to improve frequency and spatial network planning. This activity included

on site measurements around museum venues and rooftops with interviews of museum personnel who wish to use and deploy wireless networks. Most of these museums are operated by city commissions and private foundations with a strong interest in educational outreach. Most of the museums utilize 2.4GHz WiFi indoor communications, WLAN 10Mb/s indoor CAT5 networks, with ADSL backhauls up to 1.544Mb/s. These medium-band wireless communications systems do not interfere with interfere with our mesh network node within a distance of 10m or more by virtue of their low transmit power (23-DBm) that are within the dynamic range of our low noise amplifiers in our receivers. Surprisingly, the most significant communication degradation is the channel fading caused by the moist Benjamin Focus trees that obstruct wireless communications from rooftop to ground within 30m of ground based mesh network nodes. The solution is to site the mesh network node antennas at distance of 3m or more from the trees so that they remain away from the scattering leaves. The exceptions are numerous Lemon Eucalyptus trees that are transparent to our wireless mesh network nodes. We employed sufficiently directional antennas that are directed in corridors down the outdoor promenade to provide wireless network coverage to the science exhibitor's tents. We did not use our wide area coverage nodes with our 10 meter tall antenna caddy since we had adequate roof access. Vehicles at the festival with the most broad band communications available are TV news crew vans. The vehicles may be equipped with Ethernet routers with 1Gb/s total capacity, but local on-site outbound communication is limited to a 6MHz bandwidth channel. The TV news crew vans can become a significant source of short term source of electromagnetic interference to our mesh network nodes, so we have learned to route around these vans.

9. Rich Feeds Integration and NUTSO/Optiportable

Rich Feeds was used to display a variety of live and real-time data (with and without terrain overlay); clicking on the indicators (tacks) revealed the data and/or detail. The data was also archived for future analysis. The following technologies were integrated: 1) maps of node locations, 2) radio frequency (RF) spectrum sampling points, and 3) GPS location-based tracking technology in vehicles showing the location of the Calit2 vehicles at the scene.

Optiportable was deployed, using both CGLX and XDMX to show off the system's capabilities, in the booth in a 5x3 configuration (5 30"-screens across, 3 down). Various webpages, both static and dynamic (including live CogNet monitoring data graphs) were displayed, as well as the "Backpack Cam" video feed, on the XDMX side. Like NUTSO, Optiportable is portable visualization system consisting of fifteen 30-inch displays. It runs the same core system and software that NUTSO does (Rocks 5.1 x86_64, Viz 5.0, latest Hiper [CGLX])

10. Peer to Peer Traffic System

Our automated peer to peer traffic system (<http://traffic.calit2.net>) has been further disseminated with an iPhone app: commuters in California equipped with the Apple iPhone can now get personalized traffic information via the "California Traffic Report," the first iPhone application from Calit2 at UCSD. In the first ten days since the app became available through Apple's App Store on Feb. 7, roughly 2,650 people have downloaded the application, and downloads continue to run at a clip of roughly 250 per day. The California Traffic Report made it into the first page of "Top Free" apps in the Travel section of the App Store. (ref: <http://www.calit2.net/newsroom/release.php?id=1471>)

Project 3: Policy-Driven Information Sharing (PISA)

Project Summary

The objective of PISA is to understand data sharing and privacy policies of organizations and individuals; and devise scalable IT solutions to represent and enforce such policies to enable seamless information sharing across all entities involved in a disaster. We are working to design, develop, and evaluate a flexible, customizable, dynamic, robust, scalable, policy-driven architecture for information sharing that ensures the right information flows to the right person at the right time with minimal manual human intervention and automated enforcement of information-sharing policies, all in the context of a particular disaster scenario: a derailment with chemical spill, fire, and threat of explosion in Champaign.

Activities and Findings

During year 6, we integrated our simple web authentication (SAW) approach with RESCUE's friends and family reunification portal to demonstrate a new mechanism for easily and safely sharing personal information on the portal, so that only acquaintances are able to access certain information. For use with the friends and family reunification portal, we also studied how to improve query performance of fuzzy text search using list-compression techniques.

Simple Authentication for the Web (SAW) is our user-friendly alternative that eliminates passwords and their associated management headaches by leveraging popular messaging services, including email, text messages, pagers, and instant messaging. SAW (i) removes the setup and management costs of passwords at sites that use email-based password reset; (ii) provides single sign-on without a specialized identity provider; (iii) thwarts passive attacks and raises the bar for active attacks; (iv) enables easy, secure sharing and collaboration without passwords; (v) provides intuitive delegation and revocation of authority; and (vi) facilitates client-side auditing of interactions. SAW can potentially be used to simplify web logins at all web sites that currently use email to reset passwords. Additional server-side support can be used to integrate SAW with web technology (blogs, wikis, web servers) and browser toolbars for Firefox and Internet Explorer. We have also shown how a user can demonstrate ownership of an email address without allowing another party (such as a phishing web site) to learn the user's password or to conduct a dictionary attack to learn the user's password.

With SAW, the identities of those authorized to gain access must be known in advance. In some situations, only the attributes of those authorized to gain access to a resource are known in advance – e.g., fire chief, police chief, city manager. In such a situation, we can avoid the management headaches and insecurity associated with accounts and passwords by adopting trust negotiation, a novel approach to authorization in open distributed systems. Under trust negotiation, every resource in the open system is protected by a policy describing the attributes of those authorized for access. At run time, users present digital credentials to prove that they possess the required attributes. .

During the past year, we concentrated our efforts on Wireless Authentication using Remote Passwords (WARP). Current single sign-on techniques, including our own SAW, require a user to directly contact a third party during authentication. These approaches are unsuitable for wireless access, since the user does not have the network access necessary to contact a third party. WARP is a new in-band protocol that allows a user to prove to a wireless access point that she knows her password, without the access point gaining access to her password or to data that can be used to launch an off-line attack on the password. WARP has the potential to be used beyond wireless access protocols, as well.

To demonstrate the potential of WARP, we created an advanced authentication prototype that allows a user to demonstrate ownership of an email address without disclosing enough

information to an attacker (such as a phishing web site) for the attacker to receive the user's password or to conduct a dictionary attack to learn the user's password. We have developed one approach that strengthens existing client/server authentications on the web. A second approach serves as a single sign-on mechanism that allows the user to prove that she knows her password at a third party, such as her email provider, without leaking information to an attacker. This second approach works for web logins as well as wireless access.

Products and Contributions

Unless otherwise mentioned, each of these software packages is available at <http://isrl.cs.byu.edu>:

TrustBuilder2

A Framework for trust negotiation, discussed above. Available from <http://dais.cs.uiuc.edu/dais/security/tb2/>

Hidden Credentials

A credentialing system for protecting credentials, policies, and resource requests. Hidden credentials allow a service provider to send an encrypted message to a user in such a way that the user can only access the information with the proper credentials. Similarly, users can encrypt sensitive information disclosed to a service provider in the request for service. Policy concealment is accomplished through a secret splitting scheme that only leaks the parts of the policy that are satisfied. Hidden credentials may have relevance in crises involving ultra sensitive resources. They may also be able to play a role in situations where organizations are extremely reluctant to open up their systems to outsiders, especially when the information can be abused before an emergency even occurs. We have observed on the UCI campus that some buildings have lock boxes that are available to emergency personnel during a crisis. The management of physical keys is a significant problem. Hidden credentials have the potential to support digital lockboxes that store critical data to be used in a crisis. The private key used to access this information during a crisis may never have to be issued until the crisis occurs, limiting the risk of unauthorized access until the crisis occurs.

LogCrypt

A Tamper-evident log file system based on hash chaining. This system provides a service similar to TripWire, except that it is targeted for log files that are being modified. Often, an attacker breaks into a system and deletes the evidence of the break-in from an audit logs. The goal of LogCrypt is to make it possible to detect an unauthorized deletion or modification to a log file. Previous systems supporting this feature have incorporated symmetric encryption and an HMAC. LogCrypt also supports a public key variant that allows anyone to verify the log file. This means that the verifier does not need to be trusted. For the public key variant, if the original private key used to create the file is deleted, then it is impossible for anyone, even system administrators, to go back and modify the contents of the log file without being detected. During this past year, we completed experiments to measure the relative performance of available public key algorithms to demonstrate that a public key variant is practical. This variant has particular relevance in circumstances where the public trusts government authorities to behave correctly, and also benefits authorities by giving them a stronger basis for defending against claims of misbehavior. This technology may allow more secure auditing during a crisis.

Nym

A practical Pseudonymity for Anonymous Networks. Nym is an extremely simple way to allow pseudonymous access to Internet services via anonymizing networks like Tor, without losing the

ability to limit vandalism using popular techniques such as blocking owners of offending IP or email addresses. Nym uses a very straightforward application of blind signatures to create a pseudonymity system with extremely low barriers to adoption. Clients use an entirely browser-based application to pseudonymously obtain a blinded token which can be anonymously exchanged for an ordinary TLS client certificate. We designed and implemented a Javascript application and the necessary patch to use client certificates in the popular web application MediaWiki, which powers the popular free encyclopedia Wikipedia. Thus, Nym is a complete solution, able to be deployed with a bare minimum of time and infrastructure support.

Thor Credential Repository

Thor is a repository for storing and managing digital credentials, trusted root keys, passwords, and policies that is suitable for mobile environments. A user can download the security information that a device needs to perform sensitive transactions. The goals are ease of use and robustness.

SACRED

An implementation of IETF SACRED (Securely Available Credentials) protocol.

SAW

Simple Authentication for the Web. Discussed above.

Friends and Family Reunification Portal:

<http://fr.ics.uci.edu/> and

<http://www.disasterportal.org/Ontario/home.htm;jsessionid=727B73686605B2A304F65FF802696EF8> At the latter URL, the reunification portal has been incorporated into the Disaster Portal for the City of Ontario.

Clouseau Policy Compliance Checker

To help make trust negotiation practical for use in situations such as disaster response, we designed, built, evaluated, and released the Clouseau policy compliance checker, which uses a novel approach to determine whether a set of credentials satisfies an authorization policy. That is, given some authorization policy p and a set C of credentials, determine all unique minimal subsets of C that can be used to satisfy p . Finding all such satisfying sets of credentials is important, as it enables the design of trust establishment strategies that can be guaranteed to be complete: that is, they will establish trust if at all possible. Previous solutions to this problem have relied on theorem provers, which are quite slow in practice. We have reformulated the policy compliance problem as a pattern-matching problem and embodied the resulting solution in Clouseau, which is roughly ten times faster than a traditional theorem prover. We have also shown that existing policy languages can be compiled into the intermediate policy language that Clouseau uses, so that Clouseau is a general solution to this important problem.

Project 4: Customized Dissemination in the Large

Project Summary

This project focuses on information that is disseminated to the public at large specifically to encourage self-protective actions, such as evacuation from dangerous areas, sheltering-in-place, and other actions designed to reduce exposure to natural and human-induced threats. Specifically, we have developed an understanding of the key factors in effective dissemination to the public and designed technology innovations to convey accurate and timely information to

those who are actually at risk (or likely to be), while providing reassuring information to those who are not at risk and therefore do not need to take self-protective action. There are three key factors that pose significant challenges (social and technological) to effective information dissemination in crisis situations – variation in warning times, determining specificity of warning information to effectively communicate to different populations, and customization of the delivery process to reach the targeted populations in time over possibly failing infrastructures. Our approach to address these challenges is a focused multidisciplinary effort that 1) understands and utilizes the context in which the dissemination of information occurs to determine sources, recipients, channels of targeted messages and 2) develop technological solutions that can deliver appropriate and accessible information to the public rapidly. The ultimate objective is a set of next generation warning systems that can bring about an appropriate response, rather than an under- or over-response.

In the year 2008-2009, we made significant research progress in both the delivery and content layers of the dissemination project. In the delivery layer, we developed extremely fast and reliable protocols for delivery of short warning messages (e.g. an earthquake early warning). Much progress was also achieved in the design of protocols for fast, reliable dissemination over multiple wireless networks and mixed wired/wireless networks. At the content level, our focus was on enabling customizability, reliability and scalability and efficiency in the distributed publish/subscribe paradigm. While much of the research was conducted by postdocs and Ph.D. students, M.S. class projects were designed to implement, evaluate and test many of the proposed protocols exposing the larger student population to issues in effective information dissemination..

Activities and Findings

Crisis Workshop for Schools: In September 2008, we hosted a one day workshop on Emergency Information Dissemination in Schools held at the UC Irvine campus. This workshop provided us with perspective on interesting and important issues in enabling fast and effective crisis information dissemination to the very diverse K-12 community in our state and nation. In this workshop we brought together school, city, county and state representatives to discuss the current processes, systems and challenges in disseminating emergency warnings and alerts in schools. In particular, input was sought from stakeholders to help understand the extent to which recent and upcoming information and communication technologies can be customized to help the dissemination process in our schools to achieve the desired level of response. Attendees to the workshop included academia (social scientists, IT), stakeholders from multiple school districts (with varying socio-economic distributions) and K-12 levels (elementary, middle schools, high schools, university) and various levels of government (state, local, county). We were able to collect information about existing warning systems, processes, procedures for emergency warnings and alerts to schools that span different communities, populations and socioeconomic categories via panels and discussion sessions. The second half of the workshop focused on technologies and processes to help educational institutions better prepare for disasters and effectively respond in real-time to emergencies. As an outcome of this workshop we plan to move forward in the development of a short notice warning system (effective in the case of earthquakes, tornadoes, fires and flash floods) and specifically customize them for the school scenario. We also obtained willingness of institutions and our K-12 community to participate in our longer term goals towards pilot studies for testing these processes and systems.

Research contributions in Dissemination over Wired/Wireless/Mixed Networks: This year, we focused on designing a specialized protocol for very rapid and reliable dissemination of

small content (e.g. an earthquake early warning). We designed and evaluated FaReCast - a Fast, Reliable Application-Layer Multicast Protocol for Flash Dissemination in wired networks. To disseminate a message from a single source to a large number of targeted receivers, a natural approach is the tree-based application layer multicast (ALM). However, in timeconstrained flash dissemination scenarios, e.g. earthquake early warning, where time is of the essence, the tree-based ALM having single point of failure is unsuited, because the typical solution to enhance the reliability via ack-based failure recovery protocols such as TCP takes too much time. In this paper, we exploit path diversity, i.e. exploit the use of multiple data paths, to achieve fast and reliable data dissemination. First, we design a forest-based M2M (Multiple parents-To- Multiple children) ALM structure where every node has multiple children and multiple parents. The intuition is to enable lower dissemination latency through multiple children, while enabling higher reliability through multiple parents. Second, we design a multidirectional multicasting algorithms that effectively utilize the multiple data paths in the M2M ALM structure. A key aspect of our reliable dissemination mechanism is that nodes, in addition to communicating the data to children, also selectively disseminate the data to parents and siblings. As compared to trees using traditional multicasting algorithm, we observe an 80% improvement in reliability under 20% of nodes failed with no significant increase in latency for over 99% of nodes. This work has been submitted for publication to a top-tier conference.

Our research in information dissemination over wireless networks can broadly be classified into two categories – (a) instantaneous dissemination in connected networks and (b) delay-tolerant dissemination when network endpoints may be disconnected or intermittently connected. Specifically, we have worked on providing reliability guarantees for broadcast of data with rich content over connected wireless ad hoc networks. Our motivation stems from multiple emergency response scenarios, where at rescue sites mission-critical contents need to be disseminated to all practitioners instantaneously to keep their situational awareness in sync. However, existing approaches to dissemination reliability have their limitations -- they often assume network size knowledge, or that receivers know about the dissemination in advance, and that applications have uniform reliability needs. Without making similar assumptions, we propose a distinct approach which accommodates the varying reliability needs of dissemination applications. We developed the RADcast (Reliable Application Data broadcast) protocol as an integration of two components: (a) Peddler, which ensures that receivers obtain the dissemination metadata, and (b) Pryer, which delivers the actual data to dissemination-aware receivers. We indicated how reliability guarantees/performance tradeoffs can be achieved by a careful instantiation of Peddler and Pryer. We have implemented RADcast on mobile devices inside a middleware and determined its feasibility. Furthermore, through extensive simulations, we showed that RADcast guarantees desired reliability in all cases, and performs consistently under varying network conditions and device mobilities. As compared to existing approaches, RADcast either incurs significantly lower latency/message overhead, or reduces latency by 50% with a tradeoff in message overhead. As an extension to our previous work on instant disseminations in connected networks, we have further studied optimizing the performance of communications in emergency response networks through the smart deployment of a wireless mesh backbone and its gateway. While building the mesh backbone boosts the scalability, properly designating a gateway mesh router (to be connected to the external backhaul) enhances the timeliness of communications. We hence look into the gateway designation problem, which concerns, given the topology of a mesh backbone, determining which mesh router is to serve as the gateway, so that the communication latency between the gateway and other mesh routers is minimized. We mathematically model the problem using the notion of centrality from graph theory. Through building a solution framework and making optimizations in solving each sub-problem, we propose a distributed algorithm (FACE) for locating the gateway. FACE is an approximate algorithm that works in a time- and cost-efficient manner without

compromising on the optimality of solutions. A thorough performance evaluation shows that FACE effectively reduces the transmission overhead and time overhead by up to 71%, while the gateway designated by FACE can cut down on communication latency by up to 92%. This work was published and presented at the prestigious IEEE Infocom 2009 conference.

Continuing our research on information dissemination in infrastructure-less wireless networks, we have studied delay-tolerant disseminations in disconnected networks. Different from instant disseminations in connected networks, such disseminations need to accommodate the intermittent connectivity between mobile nodes while achieving the performance goals. In particular, we have looked at spatial dissemination in disconnected networks. Spatial dissemination is a specific form of information dissemination that enables mobile users to send information to other mobile users who are or will appear at a specific location (a user-defined region). Such geo-messaging services are on the rise; they typically are built upon centralized solutions and require users to have reliable access to a stable backend infrastructure for storing and communicating content. In order to offer disruption-tolerant services in scenarios where infrastructure access is unavailable, in our work, we have developed a distributed solution to spatial dissemination that can work without the need for an infrastructure. Our solution utilizes the concepts from disruption-tolerant networking to build a flexible service that leverages the intermittent ad-hoc connectivity between users. We proposed Sticker, a spatial dissemination protocol that aims to maximize delivery reliability without incurring significant storage/transmission overheads. Sticker employs the store-carry-and-forward model, and strives to optimize dissemination performance by addressing three sub-problems - replication, forwarding and purging. Our experiments show that, Sticker achieves delivery ratios that are close to the maximum possible values; as compared to existing techniques, it either cuts down storage/transmission overheads by over 50%, or greatly enhances both delivery reliability and storage efficiency. Sticker hence offers a beneficial complement to infrastructure-based centralized solutions for providing pervasive spatial dissemination services. This work has been submitted for publication to a premier networking conference.

In line with our theoretical research work on disseminations in disconnected networks, we design and prototype PassItOn, a fully distributed opportunistic messaging application on off-the-shelf mobile handheld devices. Our goal is to build up a proof-of-concept platform on real mobile devices, and thus show the feasibility and potentials of utilizing human movements for dissemination applications. Meanwhile, we seek to shed lights on the design, implementation and deployment issues in building such systems, and thus stimulate new ideas and perspectives on addressing these issues. Moreover, we aim to offer a real testbed on which new mechanisms, protocols and use cases can be tested and evaluated. The PassItOn application accommodates multiple forms of communications in disruption-tolerant networks, including one-to-one unicast, one-to-many group-cast, one-to-all broadcast and content-based publish/subscribe. Moreover, it addresses the spatial and temporal aspects of disseminations -- users can specify the location where and the time frame during which the content/message they created are relevant. In addition, PassItOn takes into account user preferences, thus users are able to specify how much resource they are willing to share and for whom. Based on that, an incentive mechanism has also been implemented. PassItOn was demonstrated (and a demo paper was published) at the IEEE Percom 2009 conference.

Research Contributions in Customized Content Dissemination: Our work at the content layer built upon our prior efforts at creating a publish/subscribe based architecture for personalized content dissemination. The framework enables users to subscribe for notifications of interest – the subscription mechanism allowed end users to not only state what

information they wished to receive, but also how, i.e., in which format and over what kind of channel (Internet, telephony, wireless messages etc.) they wished to receive notifications in. Enabling such customized notifications efficiently and reliably over large populations (i.e., scalably) was the main goal of our research efforts.

We addressed the design of “efficient” publish-subscribe systems by enabling better subscription management in the infrastructure. Specifically, we developed MICS, An Efficient Content Space Representation Model for Publish/Subscribe Systems. One of the main challenges faced by content-based publish/subscribe systems is handling large amount of dynamic subscriptions and publications in a multidimensional content space. To reduce subscription forwarding load and speed up content matching, we developed MICS (Multidimensional Indexing for Content Space), a model for efficient representation and processing of large numbers of subscriptions and publications. MICS creates a one dimensional representation for publications and subscriptions using Hilbert space filling curves. Based on this representation, we propose novel content matching and subscription management (covering, subsumption and merging) algorithms. Our experimental evaluation indicates that the proposed approach significantly speeds up subscription management operations compared to the naive linear approach. To further improve the efficiency of subscription management in content-based pub/sub systems, we developed techniques for efficient subscription subsumption evaluation. Using a novel approach based on negative space representation for subsumption checking, we provide efficient algorithms for subscription forwarding in a dynamic pub/sub environment. Heuristics were also designed for approximate subsumption checking that greatly enhance the performance without compromising the correct execution of the system and only adding incremental cost in terms of extra computation in brokers. Results of this work were published and presented at the prestigious ACM Middleware 2008 conference and at the ACM DEBS 2009 (Distributed Event Based Systems) conference – a prime venue for advances in pub-sub systems.

Managing subscriptions efficiently is one aspect; we next addressed the need for creating customized notifications for individual users based on their format needs and capabilities. We designed and developed *CCD, an Efficient Customized Content Dissemination framework*, i.e., a content-based publish/subscribe framework that delivers matching content to subscribers in their desired format. Such a framework enables the pub/sub system to accommodate richer content formats including multimedia publications with image and video content. In our proposed framework, users (consumers) in addition to specifying their information needs (subscription queries), also specify their profile which includes information about their receiving context -- characteristics of the device used to receive the content (e.g., resolution of a PDA used by a consumer). The pub/sub system besides being responsible for matching and routing the published content, also becomes responsible for converting the content into the suitable format for each user. Content conversion is achieved through a set of content adaptation operators (e.g., image transcoder, document translator, etc.). We study algorithms for placement of such operators in the pub/sub broker overlay so as to minimize the communication and computation consumption.

In implementing CCD, we extend the functionality of the brokers in the distributed pub/sub system beyond normal pub/sub processing to also support customization of the published content by executing appropriate content adaptation operators. We model the content customization possibilities by using a directed graph called Content Adaptation Graph (CAG) whose nodes are the content formats that can be supported by the system. The edges of a CAG represent content adaptation operators that convert content from the format represented by the source node to the format represented by the end node of the directed edge. The brokers are

connected through a DHT-based overlay network and content dissemination occurs via a dissemination tree constructed using the DHT-based routing schema.

Using the proposed CAG and dissemination tree we model the customized content dissemination process by calculating a dissemination plan which is a dissemination tree where nodes are annotated with customization operators and links are annotated with content formats. Then, we assign a cost for a dissemination plan based on the computations performed in the tree nodes and communications performed between tree nodes by transmitting the content in the annotated formats. Based on this modeling we define the customized content dissemination problem for a published content as the task of finding the dissemination plan with the minimum cost. We show that this is an NP-hard problem by reducing the directed Steiner Tree problem to this problem.

We propose two classes of solutions for the problem. The first class of solution is for the scenarios where the number of available formats in CAG is small. In this case, we propose an optimal plan computation algorithms based on dynamic programming. The second class of solution is for the general case where the available formats in the system are very diverse. This solution is a heuristic similar to query optimization techniques in DBMS systems. It uses a multilayer graph data structure that is created by combining CAG and the dissemination tree. The heuristic starts with an initial plan and tries to refine the plan to another plan with smaller cost in an iterative fashion. The optimization process has several parameters including the criteria to select the next plan in each iteration and number of iterations that affect the quality of the resulted dissemination plan. Our experimental evaluations show that the proposed heuristic provides dissemination plans very close to the optimal dissemination plan. The proposed architecture also mimics fault-tolerant and load balancing capabilities that are provided by DHT-based pub/sub framework. This work has been accepted for publication at ACM Middleware 2009.

In addition to supporting the basic functionality of customized content dissemination, we also addressed issues of reliability and scalability in Cluster-based Publish/Subscribe Systems. In our earlier work in Year 4/5, we developed a cluster based distributed publish-subscribe infrastructure that addressed the dual needs of customization and robustness. Building on the cluster pub-sub system, this year, we developed static and dynamic load balancing techniques to support scalable customization. Static load balancing uses the multi-cluster architecture for the broker overlay network to ensure uniform subscription distribution knowledge in the event space. Dynamic load balancing, on the other hand, is achieved through exploiting multiple inter-cluster and intra-cluster connections to dynamically distribute publication and subscription forwarding load among brokers during run time. Our experimental results show that the proposed load balancing techniques effectively prevent overloaded brokers without having significant effect on content dissemination efficiency. This work was published at the IEEE SAINT 2009 conference.

Products and Contributions

Websites for crisis workshops, software development efforts accessible from www.itr-rescue.org

FaReCast

A Fast, Reliable Application-Layer Multicast Protocol for Flash Dissemination in wired networks.

RADcast

A Reliable Application Data broadcast protocol.

FACE

An algorithm for locating the gateway of a wireless network. FACE is an approximate algorithm that works in a time- and cost-efficient manner without compromising on the optimality of solutions

PassItOn

A fully distributed opportunistic messaging application on off-the-shelf mobile handheld devices.

MICS

MICS (Multidimensional Indexing for Content Space) is a model for efficient representation and processing of large numbers of subscriptions and publications.

CCD

An efficient Customized Content Dissemination framework that is a content-based publish/subscribe framework that delivers matching content to subscribers in their desired format.

Project 5: Privacy Implications of Technology

Project Summary

Privacy concerns associated with the infusion of technology into real-world processes 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 develop 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 society. Our objective in the project has been to understand privacy concerns in adopting technology from the social and cultural perspective, and to design socio-technological solutions to alleviate such concerns. We have focused on applications that are key to effective crisis management. For example, applications for situational awareness might involve personnel and resource tracking, data sharing between multiple individuals across several levels of hierarchy and authority, information integration across databases belonging to different organizations. While many of these applications have to integrate and work with existing systems and procedures across a variety of organizations, another ongoing effort is to build a “sentient” space from the ground up where privacy concerns are addressed right from the inception, trying to adhere to the principle of “minimal data collection.”

Activities and Findings

In year 6 we continued our work in the two main areas:

1. creating privacy-aware observation systems (pervasive spaces). Observation systems, where sensors are used to create awareness, are of paramount importance to crisis response and also one of the primary directions of research in RESCUE. Our goal, in this project, has been to explore implications of deploying these systems to privacy and to identify/develop privacy technology that could address the privacy needs.
2. Secure data management and sharing systems in the outsourcing scenario. Data outsourcing is an approach to a emerging model to get quality managed data services without the need for incurring the high cost of building an IT infrastructure which could be

cost-prohibitive for smaller agencies. Through outsourcing, organizations can achieve not only better management of their resources but participate effectively in data sharing and mobile access to data. Indeed, concepts such as Ebox explored in the SAMI project are a natural realization of data outsourcing model as applied to emergency response.

In year 6, we continued our investigation of privacy technologies along both the above directions which is summarized below.

Privacy in Observation Systems: In year 6, we made a significant departure from our prior work (until year 5) by considering weaker trust models compared to fully untrusted model we had adopted in prior work. In the untrusted model, the observation system itself was entirely untrusted and data was encrypted at the sensors. Approaches/techniques were developed to do event detection obliviously such that the adversary cannot infer any privacy properties of subjects in the observation space. While our prior work had developed such an approach, restrictions posed on the nature of events it could detect restricted its impact. Given the limitations of untrusted model, as well as the abundance of applications where weaker models might suffice, we considered a weaker trust model wherein the observation system in itself is trusted though individuals/subjects immersed in the observation space may have privacy policies about who can (and under what conditions) access information that may lead to inferences about attributes considered sensitive by the subjects. The weaker trust model allowed us to design a functionally richer system where the environment is modeled as a data streaming source on which users can pose real-time, continuous queries. Continuous queries are the long running and need to be evaluated against the dynamically changing state of the space. Here, the idea is to let users pose generic queries about the environment or other individuals and also allow them to specify their own policies for disclosing information pertaining to them. The goal is to release as much information as possible without violating the privacy preferences of individuals. Specifically, we built a middleware for pervasive spaces that plays the tradeoff between these two conflicting requirements - maximizing the utility of the information exposed and ensuring users' privacy preferences are met. In order to resolve these two (often) conflicting objectives, we propose a framework grounded in utility theory where users dynamically control the level of disclosure about their information. We provide appropriate definitions of privacy and utility for RegionMonitor type of applications that are intended for supporting collaborative work. Existing definitions of privacy and anonymity do not apply too well in this context. We propose a distributed solution that, given a user's background knowledge, maximizes the utility of the information being disclosed to information recipients while meeting the privacy requirements of users. We implement our solution in the context of a real pervasive space middleware and provide experiments that demonstrate its behavior.

Secure data outsourcing and sharing systems: In years 3, 4 and 5 we developed two new products DataGuard & DataVault which allowed an individual to securely store his/her data on a remote server and access it from a diverse set of clients possibly located in unsafe environments. Data confidentiality and user authentication were two important challenges that were addressed in these software. In the past year, we focused on a slightly different setup – we looked at existing web applications where users stored and accessed personal data on servers remotely and asked ourselves, how can we make this secure? In other words, we wanted to explore approaches which would prevent the server from learning the content in user's data, but nonetheless would allow the user to avail most of the functionalities of the application. We specifically looked at the Google calendar application which allows users to store events and their descriptions. Search is allowed using date range and keywords on “what”, “where” fields as well as free text search on the “description” field. We created a middleware that could adequately “obfuscate” the data representation on the server so that server-side entities are not able to learn these values. Specifically, we wanted to obfuscate the

representation of dates of events and the keywords present in the text fields while still enabling search on these attributes. We implement the order-preserving encryption techniques proposed in EUROCRYPT-09 for enabling search on the date field and implement the hash-and-encrypt strategy proposed in CRYPTO-07 and DBSEC-07 for encrypting keywords. We realized that the techniques proposed in the above mentioned papers can be significantly strengthened in practice by incorporating a few heuristics. We use ideas of bucketization similar to what was used in our previous work to enhance the security of range queries on the date field. For keyword queries, we enhance the security of frequent words (to prevent frequency-based attacks) by using multiple hash functions for some chosen words. While such a technique increases the overhead of keyword matching and retrieval, the number of hash functions can be used as a tuning parameter for trading off security and performance.

Products and Contributions

PrivGoogleCalendar

PrivGoogleCalendar is a personal privacy management middleware targeted at the Google Calendar to empower individual users to control the privacy of their information. The model used to develop PrivGoogleCalendar is that the service provider is untrusted and the users cannot modify the Google calendar application. PrivGoogleCalendar sits between a Google calendar client (Google calendar web page, embedded Google calendar or other Google calendar clients) and the Google calendar server. It is browser independent and does not require any modification to the Google calendar storage model, data access protocol, and interface. It is developed based on proxy technologies, which enable it to intercept and secure the communication between the Google clients and the server. The current implementation supports most of the Google calendar functionality including searching and sharing. PrivGoogleCalendar was designed with usability as an important factor with the goal of making it easy to use for the average Internet users.

RegionMonitor

RegionMonitor is a SATware application that allows users to pose queries pertaining to other individuals and resources in the pervasive space. User needs and privacy requirements are modeled in a utility framework where users dynamically specify via policies their information and privacy needs as the positive and negative utilities associated with the release of certain information. SATware attempts to maximize the utility of the information being released while preserving privacy with a polynomial optimization algorithm based on distributed simulated annealing. This algorithm uses a probabilistic rule-based system to compute potential privacy violations.

Project 6: MetaSIM

Project Summary

MetaSIM is a web-based collection of simulation tools developed to test the efficacy of new and emerging information technologies within the context of natural and manmade disasters, where the level of effectiveness can be determined for each technology developed. MetaSIM currently incorporates three simulators: 1) Crisis simulator InLET; 2) Transportation simulator, and 3) Simulator for agent based modeling (Drillsim).

Activities and Findings

In Year 6, the core earthquake loss estimation component, InLET, has been decoupled from the MetaSIM modeling environment, and ported to Microsoft Virtual Earth. InLET was deployed at the Great Southern California ShakeOut Exercise for two agencies: i. City of Inglewood, ii. CalEMA (California Emergency Management Agency). InLET was used to assess preliminary damage and generate situational awareness for the responding agencies and the local residents for the simulated earthquake measuring 7.8 on the Richter scale. A custom version of InLET was developed with the City's building and lifeline inventory for the ShakeOut scenario. Demonstrations were made to several high profile public and private agencies, as well as NGOs. The resolution of base data was expanded to the building parcel level for the City of Inglewood, demonstrating that the platform could produce and map building level results.

Additionally, the InLET platform was presented to several City of Los Angeles staff and is currently being considered for integration into the Emergency Operations Center. Participants included- Nekpen Aimiwu, Department Emergency Coordinator, City of Los Angeles, Planning Dept; Arif Alikhan, Deputy Mayor, City of Los Angeles; Todd Chamberlain, Captain, LAPD Special Operations Bureau; Scott Kroeber, Commander, LAPD Special Operations Bureau; Olivia Mitchell, Deputy Director, City of Los Angeles Community Development; Andy Neiman, Lieutenant, LAPD Special Operations Bureau; Richard Roupoli, Deputy Chief, LAPD Special Operations Bureau; and Tony Varela, Assistant Chief, LAFD Homeland Security. These high level end users were able to walk away with a clear understanding of how the results of research could practically be integrated into their organizational structures, as well as provided critical feedback for refining technology for end users.

Products and Contributions

The primary artifact of the Transportation Testbed is MetaSIM. METASIM is a web-based collection of simulation tools developed to test the efficacy of new and emerging information technologies within the context of natural and manmade disasters, where the level of effectiveness can be determined for each technology developed. METASIM incorporates a crisis simulator, a transportation simulator, and a simulator for agent based modeling (Drillsim). METASIM is envisioned as a comprehensive modeling platform for plug-and-play simulation tools for emergency managers and first responders to support response, recovery and mitigation activities.

A preliminary website has been developed in HTML and stored in the backend database to produce web pages on-the-fly through Java script. The web pages call the various simulators and allow users to define parameters for the various simulations. The parameters are saved in user specified scenarios and the simulations are run through the interface. After each run the results are stored in the database and the website calls and displays intermediate and final results.

A description of the individual simulators and components integrated into the METASIM framework is provided below:

Crisis Simulator/ InLET

The Crisis Simulator currently simulates an earthquake event and estimates damage and casualties for Los Angeles and Orange counties. The crisis simulator integrates the earthquake loss estimation components of InLET; the Internet based Loss Estimation Tool.

DrillSim

DrillSim is an agent-based activity simulator that models human behavior at the individual or micro level. DrillSim tests IT solutions by modeling situation awareness and provides it to the agent to react accordingly. For example, an early warning system might be used to modify the timing of agent evacuation. Micro-level activity modeling provides the ability to mimic agent behavior in crisis, as well as interactions between people during crisis, thereby providing a more robust framework for integrating responses to information and technology.

Transportation Simulator

The transportation simulator consists of an integrated model of simplified quasi-dynamic traffic assignments, and a destination choice model. Information that becomes available through IT solutions is simulated through parameters, such as subscription to routing support information via cell phone or email, information arrival time and update frequency, system credibility and acceptance, to reduce uncertainties associated with decision making when evacuating a congested network.

GIS Applet for Visualization

A GIS applet has been developed for the crisis simulator for visualization of the different geographic data layers and the simulation results.

2.2 Other RESCUE-Related Research

Carboxyhemoglobin (SpCO) Levels in Firefighters. This research study investigated the exposure levels of firefighters to Carbon Monoxide (i.e., sampling SpCO levels) at the incident site. Firefighters were equipped with a number of personal and environmental sensing devices. These devices utilized zibgee and 802.11 RF technologies to report sensed levels in real time. We obtained over 7,000 data points throughout a 4 hour Live Burn exercise. This data is being analyzed and correlated in order to calibrate our systems and understand how environmental factors impact SpCO readings.

2.3 RESCUE Artifacts

CalMesh Networking System: CalMesh is an affordable mesh networking solution enabling Internet access and team communication where the infrastructure has been compromised or damaged. It is a quickly self-organizing WiFi mesh network of small, lightweight and easily reconfigurable nodes.

At the request of the San Diego Science festival organizers, we set up a CalMesh ad-hoc wireless network covering most of the exhibit areas along the Prado (around 40+ booths), many of which would otherwise have not had WiFi connectivity or only spotty access to the Internet, much less high-speed web access. The coverage area was along the eastern part of the Prado from the Lilly Pond to the Fountain. This was the largest deployment of CalMesh ever.

The CalMesh network was linked to the Internet via the High Performance Wireless Research and Education Network (HPWREN) and a commercial provider, Sky River, both had access points on the Natural History Museum's roof. An access bonding solution from Mushroom Network Inc. aggregated the two channels together. Combined, they provided 45 megabits per second of bandwidth.

The deployment also provided an excellent opportunity to collect data, to further our research on communications in cases of emergency. Many experiments were conducted and measurements taken on the network and surrounding environment.

Peer-to-Peer Adaptive Information Collection System: This artifact is a fully automated peer-to-peer system (<http://traffic.calit2.net>) in San Diego, Los Angeles, and the Bay Area (in northern California) that collects and relays highway incident information to the general public and to first responders. Our automated peer to peer traffic system (<http://traffic.calit2.net>) has been further disseminated with an iPhone app: commuters in California equipped with the Apple iPhone can now get personalized traffic information via the "California Traffic Report," the first iPhone application from Calit2 at UCSD. In the first ten days since the app became available through Apple's App Store on Feb. 7, roughly 2,650 people have downloaded the application, and downloads continue to run at a clip of roughly 250 per day. The California Traffic Report made it into the first page of "Top Free" apps in the Travel section of the App Store. (ref: <http://www.calit2.net/newsroom/release.php?id=1471>).

Rich Feeds System/Optiportable: Rich Feeds (<http://rescue.calit2.net/>) is a system that demonstrates how unconventional data feeds and emergent data feeds can be captured,

preserved, integrated, and exposed either in real time or after the fact. Rich Feeds promotes situational awareness during a disaster by integrating and displaying these feeds. Rich Feeds was used to display a variety of live and real-time data (with and without terrain overlay); clicking on the indicators (tacks) revealed the data and/or detail. The data was also archived for future analysis. The following technologies were integrated:

- maps of node locations
- radio frequency (RF) spectrum sampling points
- GPS location- based tracking technology in vehicles showing the location of the Calit2 vehicles at the scene.

Optiportable was deployed, using both CGLX and XDMX to show off the system's capabilities, in the booth in a 5x3 configuration (5 30"-screens across, 3 down). Various webpages, both static and dynamic (including live CogNet monitoring data graphs) were displayed, as well as the "Backpack Cam" video feed, on the XDMX side. We were also able to pull up several extremely high resolution images on the CGLX side (spread across several screens) and zoom in as far as the image resolution would allow. (You could see hikers in photos of Half-Dome in Yosemite that were too small in normal resolution.) Early in the day, there were some serious issues which were resolved by the team using an advanced configuration of the nearby MESH node using NAT, and some quick reconfiguration of Optiportable's routing. Optiportable is a portable visualization system consisting of fifteen 30-inch displays. It runs Rocks 5.1 x86_64, Viz 5.0 as the core system and the latest Hiper [CGLX] software.

Disaster Portal: The Disaster Portal (www.disasterportal.org) is an easily customizable web portal and set of component applications which can be used by first-responders to provide the public with real-time access to information related to disasters and emergency situations in their community. Current features include a situation overview with interactive maps, announcements and press notifications, emergency shelter status, and tools for family reunification and donation management. The Disaster Portal dramatically improves communication between first-responders/government agencies and the public, allowing for rapid dissemination of information to a wide audience.

The development of the Disaster Portal is based on two primary considerations. While we aim to provide practical applications and services of immediate utility to citizens and emergency managers, we also strive to significantly leverage many relevant pieces of IT research within RESCUE. The advanced technologies that are currently incorporated into the Disaster Portal include components for customizable alerting, family reunification, scalable load handling, unusual event detection and internet information monitoring.

Recent development on the Disaster Portal software has focused on documentation and packaging for additional deployments by other city or county governments. Support of the original pilot deployment for the City of Ontario, California has been transitioned to city IT resources, and a new deployment is being made by Champaign, IL. The team is in discussions with the County of San Diego for a possible large scale deployment to that region.

SAW: The PISA project's SAW authentication technology eliminates the need for users to remember a password. SAW accomplishes this by leveraging popular messaging services, including email, text messages, pagers, and instant messaging. This technology was implemented as a deployment option within the Disaster Portal application.

TrustBuilder2: To address the systems issues associated with the adoption of trust negotiation as a flexible authorization approach for virtual organizations, we have designed, implemented,

evaluated, and released the TrustBuilder2 framework for trust negotiation. TrustBuilder2 is a flexible framework designed to allow researchers to quickly prototype and experiment with various approaches to the trust negotiation process. In TrustBuilder2, the primary components of a trust negotiation system are represented using abstract interfaces, any of which can be overridden or extended by users of the system. TrustBuilder2 is also agnostic with respect to the formats of credentials and policies used during the negotiation process; support for new credential formats and policy languages can easily be added by extending the appropriate classes in TrustBuilder2's abstract type hierarchy. The flexibility of TrustBuilder2, along with its support for the interposition of user-defined plug-ins at communication points between system components, not only enables users to rapidly implement support for new features, but also provides a framework within which the trade-offs between various system configurations can be quantitatively analyzed. The full source code of the TrustBuilder2 framework is available under a BSD-style open source license at <http://dais.cs.uiuc.edu/tn>, and has been downloaded more than 200 times, and is currently being used by several research laboratories.

Within the TrustBuilder2 framework, we have investigated the design of efficient policy compliance checkers. Given a policy p and a set C of credentials, a compliance checker is responsible for determining one or more minimal subsets of C that satisfy p . A compliance checker capable of finding all such satisfying sets of credentials would afford its users significant strategic and privacy-preservation benefits. Such a compliance checker would give users a complete view of the next-step state space for the negotiation, thereby allowing them to choose a locally-optimal satisfying set to disclose (e.g., the least sensitive satisfying set). Because many trust management policy languages are based on declarative logics (e.g., Datalog), traditional top-down theorem-proving approaches have long been used to solve the compliance checker problem. These types of theorem provers find at most one proof for a particular conclusion; as a result, compliance checkers built using this approach find at most one satisfying set of credentials at a time. Unfortunately, results in the research literature show that the time overheads of using these types of compliance checkers to find all satisfying sets of credentials grow exponentially in the size of the union of the satisfying sets discovered. This approach is not only asymptotically undesirable, but also impractical, as it can take seconds to find only a small number of satisfying sets. In an effort to increase the efficiency of trust negotiation systems, as well as enable participants to make more informed decisions, we developed the Clouseau compliance checker.

Clouseau: Clouseau takes a non-traditional approach to the type of theorem proving used in policy compliance checking. Specifically, Clouseau compiles trust negotiation authorization policies into patterns, translates credentials into abstracted objects, and then leverages efficient pattern matching algorithms developed by the artificial intelligence community to determine all satisfying sets of credentials for a particular policy. The running time of this algorithm scales as $O(NA)$, where N is the number of satisfying sets and A is the average size of each satisfying set; this is much more efficient than existing approaches. In more practical terms, our implementation of Clouseau - which runs inside of the TrustBuilder2 framework - can find hundreds of satisfying sets per second and can find all satisfying sets of credentials for practical policies in about the time required for one hard disk access. However, the model theories used to prove the correctness of more traditional compliance checkers do not map directly onto this pattern matching approach in the general case. To account for this, we have proven the correctness of Clouseau under a broad class of circumstances. Specifically, we have described compilation procedures for translating trust management policies written in the RT0, RT1, and WS-SecurityPolicy policy languages into a format suitable for analysis by Clouseau. We then proved that when operating on these compiled policies, Clouseau finds all satisfying sets for the

original policies, finds no extraneous satisfying sets, and does all of this much faster than existing approaches.

SATware system: SATware (<http://ics.uci.edu/~projects/SATware>) is a multimodal sensor data stream querying, analysis, and transformation middleware that aims at realizing a sensor-based observation system that was originally designed as a RESCUE artifact to serve as a testbed for research on privacy and on situational awareness. As part of RESCUE, we have instrumented about a 1/3 of the campus including the Calit2 building as well as the newly constructed Bren Hall with a variety of sensors including cameras, audio sensors, motes with temperature sensors, mobile devices with GPS technologies, radios, people counters, etc. This sensing infrastructure (that we refer to as Responsphere) is used to implement and monitor a variety of campus drills and experiments that test and validate RESCUE technologies.

In contrast with previous work on distributed stream systems and distributed sensor databases, in SATware our architectural objective is to hide the raw data being sensed and instead provide a semantically meaningful view of a geographic region and its entities to the applications. To this end, we modeled sentient spaces with two main layers of abstraction: the physical layer that consists of sensors, actuators, networks, and computing and storage nodes, and the semantic layer which captures the semantics of the real-world being sensed. Such a semantic level is modeled using data modeling techniques such as ER diagrams. Sensor data is translated through the concept of “virtual sensors” into meaningful events that are modeled as changes in attributes associated with the entities and relationships.

Based on the above semantic approach, we designed and implemented SATware which now serves as a middleware architecture connecting over 300 diverse types of sensors (including over 200 cameras) spread all over UCI. In addition, SATware runs on mobile platforms and provides interfaces to associate new sensors, new operators, as well as a powerful declarative query language based on SQL for writing applications. SATware supports a policy language for encoding privacy policies which are specified at the semantic level and translated to the sensor level. Furthermore, SATware supports SATmonitor and SATscheduler components that monitor and automatically schedule resources to optimize quality of data capture. In addition, techniques have been incorporated to exploit semantics to dynamically calibrate/recalibrate sensors.

Progress in SATware has witnessed an advance of SATware from being an artifact for demonstration of privacy & situational awareness research to being an important derivative research goal for RESCUE in itself. There are currently two PhD students working on SATware and the SATware team has applied for additional extramural funds to sustain the ongoing research in the area. In particular, SATware has transitioned into becoming the software “glue” for the recently funded DHS FEMA project at UCI.

Crisis Alert: an Artifact for Customized Dissemination: The Crisis Alert system artifact has been built during the past year with the goal of integrating research directly into information dissemination technology and to respond to the issues identified in the warning literature regarding over-response and under-response in crisis situations. In fact, Crisis Alert has the ability to send emergency notifications that are customized to the needs of each recipient. Crisis Alert also contains rich information such as maps of the area, location of the open shelters closer to the recipient's location, current state of hospitals and their address and contact information and they can be automatically created by the system according to a set of rules defined during the risk-knowledge phase of deployment of a warning system.

To reach a greater part of the population and to overcome the partial failure of communication infrastructure, Crisis Alert delivers emergency notifications through different modalities. In addition, Crisis Alert takes advantage of the emergency response plan of each organization, integrating social networks in the emergency dissemination process.

The main goal of this year for the Crisis Alert system has been to test it in real scenarios and incorporate the findings due to these tests. The Ontario drill has provided useful inputs to improve the usability of the system, regarding the case where no policies have been identified for the current emergency. In order to face this situation, the policy language has been enriched with the concept of “protective action,” allowing the emergency personnel to specify policies that can be applied to different and unpredictable events. The drill has also highlighted the need for defining the concept of group of events. In fact, when a major disaster strikes, it usually generates a set of emergencies that are related to the major one but that can be of different nature and require different countermeasures. In this case, the population response could be improved if complete information is provided by the authorities through a single notification or update.

We are also trying to validate the Crisis Alert system through a series of pilot studies that involve schools and educational institutions. There are multiple purposes for these studies: to deploy the software prototype that we already have into a test scenario, to gather feedback from both the emergency personnel and the alert recipients involved, and to compare the information learned from this feedback to information obtained from actual drills. In this study, we would like to take advantage of the infrastructure that has already been put into place by Fonevia, which is used on a day-to-day base for disseminating information from schools to parents during non-disaster times. In fact, if people are already familiar with the dissemination system, it is more likely to obtain a prompt reaction when a warning is issued. Furthermore, we will supplement our technology testing with a simulation framework that will help us understand the alert dissemination in the whole community. Given knowledge of the geographies, policies and protocols – we can conduct a what-if analysis of the speed at which information can be spread in the community given different technology and usage scenarios.

Finally, in the next months we are planning a workshop with representatives from Southern California schools and school districts to collect information about existing warning systems, processes and procedures for emergency warnings and alerts to schools. This information will help us design the next generation of technologies and processes to help educational institutions better prepare for disasters and effectively respond in real-time to emergencies.

MetaSim: METASIM, both a project and an artifact, is a web-based collection of simulation tools developed to test the efficacy of new and emerging information technologies within the context of natural and manmade disasters, where the level of effectiveness can be determined for each technology developed. METASIM currently incorporates three simulators: 1). Crisis simulator InLET, 2) Transportation simulator, and 3) simulator for agent based modeling (Drillsim). For a detailed discussion of this artifact, see Section 2.1.

Responsphere: Responsphere is an IT infrastructure set of testbeds that incorporates a multidisciplinary approach to emergency response drawing from academia, government, and private enterprise. The IT infrastructure is used to test the efficacy of RESCUE technologies and extract meaningful metrics regarding those technologies. We view the testbeds as proving grounds for disruptive technology. During Year 6, the focus was to maintain the testbed. Several RESCUE artifacts (e.g., Disaster Portal and MetaSim) are hosted within Responsphere

and our priority is to keep these artifacts running and available to first responders and the communities. Additionally, several first response activities such as drills and technology deployments were either conducted within Responsphere or facilitated by Responsphere.

Data Repository: RESCUE has created a publicly available, with proper IRB clearance, repository of disaster response related data sets (<http://rescue-ibm.calit2.uci.edu/datasets>). These data sets are an extensive collection of information including (but not limited to): GIS maps, 911 calls, news broadcasts, mote telemetry data, and data gathered from emergency response exercises. We believe this repository to be the largest collection of disaster response related data in existence.

3 BROADER IMPACT

3.1 Community Outreach

Outreach to First Responders. Through the creation of the Center for Emergency Response Technoloiges (CERT) at UCI, we have significantly increased our community and first responder outreach program. CERT is the continuity vehicle for the Rescue project beyond the 4+ years of NSF funding.

We have performed several joint technology testing exercises with the first responders, including several Hazmat drills and one "Live Burn" exercise. The Hazmat drills were conducted on campus with the cooperation of the police, fire, and EH&S departments. The Live Burn exercise was a technology deployment exercise conducted at the Anaheim California Fire Training Facility. We deployed a number of Rescue technologies at the exercise and collected a large amount of data as well as practitioner feedback.

Outreach activities at UCSD included demonstrating our infrastructure and research technologies for industry groups, domestic and international governmental delegations, and conferences that took place at Calit2. Collaboration with UCSD Campus Police and UCSD Emergency Management has continued to evolve. At the conclusion of the Responsphere and RESCUE projects at UCSD, we will begin to engage with our campus partners as part of the WIISARD SAGE project. The most significant outreach activity was our participation in the 2009 San Diego Science Festival Expo, attended by over 50,000 people.

Researchers from Rescue and the NIH/NLM WIISARD projects participated in the grand finale of the month-long San Diego Science Festival: Expo Day at Balboa Park. More than 50,000 people attended the event held on Saturday, April, 4, 2009, which featured 200+ exhibition booths. Organizers called it "the largest one-day science gathering ever in the United States."

More than three dozen researchers (PIs, faculty, staff, postdocs, graduate and undergraduate students) were on-hand at Balboa Park to run demonstrations, provide information to visitors, and manage the wireless network and the many experiments.

Multiple technologies (devices, software and systems) were deployed and demonstrated, including several versions of Gizmo, Rescue's family of wireless mobile platforms designed to transport cameras, other sensors, and wireless access points to and around disaster sites in order to get communications going again in an emergency. Gizmo put a smile of the face of dozens and dozens of kids, big and small. Children as young as 5 years old waited patiently in

line to "test-drive" one. The mobile touch screen kiosk based on the Gizmo technology also made its public debut in the Calit2 booth.

At the request of organizers, we implemented the largest ever deployment of CalMesh. It provided bandwidth to dozens of exhibitors (who otherwise would not have enjoyed high-speed Web access because Balboa Park wasn't equipped for it); some of the access points were also open for the 50,000 Expo attendees to use. Photo Gallery is available here: http://projects.calit2.net/gallery/main.php?g2_itemId=3436.

Industry Interactions. We continue to actively cultivate industrial affiliates as well as government partners as part of our outreach mission. We have transitioned our Rescue Industrial Affiliates over to CERT and continue to work with these organizations on technology development and testing.

At UCSD, we have continued to work with Anritsu in gathering and understanding wireless spectrum data using their Electromagnetic Interference Measurement Software for portable spectrum analyzers, and with Nokia Siemens Networks (NSN) on FEMTO Cell Interface Tools and Mobiles In the last year we also collaborated with SkyRiver Communications, Mushroom Networks and the High Performance Wireless Research and Education Network (HPWREN) on the wireless network deployment for the San Diego Science Festival.

Broader Community: The RESCUE Disaster Portal and the Peer-to-Peer traffic system are examples of successful outreach to the broader user community. Citizens are empowered (through RESCUE Technology) to not only make day-to-day decisions regarding their commutes but make potential life saving decisions by virtue of the information disseminated to them.

Several high-level meetings have been conducted with State and Federal officials in emergency management to discuss opportunities for deployment of RESCUE technologies. For example, RESCUE team members have demonstrated the capabilities of MetaSIM to the GIS and Earthquake/Tsunami program managers at the California State Office of Emergency Services and the Mitigation Directorate of DHS/FEMA.

3.2 Education Outreach

Graduate and Undergraduate Education. RESCUE continues to have an impact on course curriculum throughout all the universities involved in the project. Specific courses and class projects have been designed to have a direct tie to the research being done at RESCUE. Details of the ongoing RESCUE seminar series (now a part of the CERT seminar series) can be found at <http://cert.ics.uci.edu>. Throughout the year, the RESCUE project has encouraged undergraduate students to be a part of ongoing research through individual study courses, honors courses, the NSF-funded California Alliance for Minority Program (CAMP), Calit2's SURF-IT program, and undergraduate research appointments.

At UCI, we also participated in the Calit2-sponsored SURF-IT program for undergraduate research. The program is a 10-week summer program and RESCUE hosted <STUDENT>. Sentence on Student work. In addition to the SURF-IT program we taught the following Rescue related courses at UCI: ICS 192, ICS 214A, ICS 214B, ICS 215, ICS 203A, ICS 278, ICS 199, ICS 290, ICS 280, ICS 299.

The following UCSD undergraduate (1xx) and graduate (2xx) courses have been based on the RESCUE research area and/or utilized Responsphere infrastructure; in some cases, project-based courses have either contributed to infrastructure improvements or built new components for the infrastructure: ECE 191 (6 projects), MAE 156B (1 project), ENG 100, CSE 294 and CSE 218. In addition, RESCUE Robust Networking and Information Collection research area leader BS Manoj continues to teach ECE 158B (Advanced Data Networks, which covers challenges in communications during disasters).

K-12 Education. RESCUE continues to reach out to the K-12 community by sponsoring high school interns and participating in campus events for high school students. During the 2008-2009 academic year, Rescue hosted several high school senior interns from the Preuss School, a charter school under the San Diego Unified School District whose mission is to provide an intensive college preparatory curriculum to low-income student populations and to improve educational practices in grades 6-12.

At UCI's CERT Center, we recently hosted an Information Dissemination workshop with representatives from Southern California schools and school districts to collect information about existing warning systems, processes and procedures for emergency warnings and alerts to schools. This information will help us design the next generation of technologies and processes to help educational institutions better prepare for disasters and effectively respond in real-time to emergencies.

Internships and Student Exchange Programs. Students participated in the creation and evaluation of several RESCUE technologies, several students from the senior group design courses in electrical or mechanical engineering went on to intern on various subprojects.

Academic Community: Rescue researchers and technologists from UCSD campus gave a number of keynote addresses and invited talks. These addresses provide the Rescue team the opportunity to engage a number of stakeholders (Government, industry, Academia, and First Responders) within the emergency response domain. Researchers at Rescue have hosted or participated in a number of events, sessions, and workshops throughout Year 6. For detailed information regarding these events, please reference the "latest news" section under www.cert.ics.uci.edu

4 FUTURE PLANS

Future directions for research in policy and information sharing for disaster management include ultra-lightweight mechanisms for gracefully adapting to dynamic changes in virtual organization membership and connectivity; automated support for analysis of access logs and audit trails generated while a system was operating with lowered authorization barriers due to emergency conditions; and user-friendly policy management tools that support visualization, exploration, and analysis of large sets of policies and proposed policy changes.

Various options are being evaluated for future research directions deploying the core components of MetaSIM individually as well as a complete system. We seek to leverage the capabilities of METASIM and are working to transform it into a tool that can be used for multiple purposes. Some possible applications include: (a) a loss estimation and decision support tool for public agencies, (b) the integration into Web Portals such as the City of Ontario's portal www.disasterportal.org, (c) the creation of a training tool for first responders, (d) the creation of a modeling tool for online services, and (e) the creation of an online disaster game for the gaming industry.

Continuity of both the Rescue project and its accompanying infrastructure (Responsphere) is being pursued through the CERT Center at UCI. The Rescue Community Advisory Board is transitioning to CERT as is the first responder advisors, staff and affiliated faculty. Funding opportunities for continued operations of CERT are being pursued through NSF, NIST, DHS, as well as other avenues (e.g., Raytheon Corporation).

In addition RESCUE has acquired supplemental funding to further develop and refine its work on Entity Resolution focusing towards web search technologies. This supplemental award is of direct importance to artifacts such as family reunification. Progress on this work will be included in the final report for RESCUE.

APPENDIX A: PUBLICATIONS IN YEAR 6

1. Bhorkar,Abhijeet; BS,Manoj. Selection Diversity based MAC Protocol for MIMO Ad hoc Wireless Networks, IEEE Globecom 2008,2008-11
2. Jafarpour,Hojjat; Hore,Bijit; Mehrotra,Sharad; Venkatasubramanian,Nalini. Subscription Subsumption Evaluation for Content-based Publish/Subscribe Systems, ACM/IFIP/USENIX Middleware 2008,2008-12
3. van der Horst,Timothy W.; Seamons,Kent E.. pwdArmor: Protecting Conventional Password-based Authentications, 24th Annual Computer Security Applications Conference (ACSAC 2008),2008-12
4. Vaisenberg,Ronen; Mehrotra,Sharad; Ramanan,Deva. Exploiting Semantics For Scheduling Data Collection From Sensors On Real-Time To Maximize Event Detection, 2009 MMCN Conference on Multimedia and Computer Networks,2009
5. Dilmaghani,Raheleh B.; Rao,Ramesh R.. A Systematic Approach to Improve Communication for Emergency Response, 2009 Hawaii International Conference on System Sciences,2009-01
6. Kalashnikov,Dmitri V.; Chen,Zhaoqi; Mehrotra,Sharad; Nuray-Turan,Rabia. Web People Search via Connection Analysis, The IEEE Transactions on Knowledge and Data Engineering,2009-01-01
7. Behm,Alexander; Ji,Shengyue; Li,Chen; Lu,Jiaheng. Space-Constrained Gram-Based Indexing for Efficient Approximate String Search, ICDE 2009,2009-03-29
8. Kalashnikov,Dmitri V.; Chen,Zhaoqi; Nuray-Turan,Rabia; Mehrotra,Sharad; Zhang,Zheng. WEST: Modern Technologies for Web People Search., In Proc. of IEEE International Conference on Data Engineering (IEEE ICDE),2009-04
9. Desai,Chaitanya; Kalashnikov,Dmitri V.; Mehrotra,Sharad; Venkatasubramanian,Nalini. Using Semantics for Speech Annotation of Images., In Proc. of IEEE International Conference on Data Engineering (IEEE ICDE),,2009-04
10. Xing,Bo; Mehrotra,Sharad; Venkatasubramanian,Nalini. RADcast: Enabling Reliability Guarantees for Content Dissemination in Ad Hoc Networks, The 28th Conference on Computer Communications (IEEE INFOCOM 2009),2009-04
11. Xing,Bo; Seada,Karim; Venkatasubramanian,Nalini. An Experimental Study on Wi-Fi Ad-Hoc Mode for Mobile Device-to-Device Video Delivery, IEEE INFOCOM 2009 Workshop on Mobile Video Delivery (MoViD 2009),2009-04-24
12. Nuray-Turan,Rabia; Chen,Zhaoqi; Kalashnikov,Dmitri V.; Mehrotra,Sharad. Exploiting Web querying for Web People Search in WePS2., In 2nd Web People Search Evaluation Workshop (WePS 2009), 18th WWW Conference,,2009-05
13. Zhang,Wenyi; Rao,Bhaskar. Two Microphone Based Direction of Arrival Estimation for Multiple Speech Sources using Spectral Properties of Speech, none,2009-05
14. Zhang,Wenyi; Rao,Bhaskar. Combining Independent Component Analysis with Geometric Information and its Application to Speech Processing, ICASSP,2009-05
15. Lee,Adam; Winslett,Marianne; Perano,Kenneth. TrustBuilder2: A Reconfigurable Framework for Trust Negotiation, Proceedings of the International Conference on Trust Management (IFIPTM 2009),2009-06

16. Chen,Zhaoqi; Kalashnikov,Dmitri V.; Mehrotra,Sharad. Exploiting Context Analysis for Combining Multiple Entity Resolution Systems., In Proc. of ACM SIGMOD Int'l Conf. on Management of Data (ACM SIGMOD),2009-06-29
17. Jafarpour,Hojjat; Mehrotra,Sharad; Venkatasubramanian,Nalini. Dynamic Load Balancing for Cluster-based Publish/Subscribe System, IEEE SAINT 2009,2009-07
18. Jafarpour,Hojjat; Mehrotra,Sharad; Venkatasubramanian,Nalini; Montanari,Mirko. MICS: An Efficient Content Space Representation Model for Publish/Subscribe Systems, ACM DEBS 2009,2009-07
19. Hojjat Jafarpour, Bijit Hore, Sharad Mehrotra and Nalini Venkatasubramanian."CCD: Efficient Customized Content Dissemination in Distributed Publish/Subscribe", ACM/IFIP/USENIX Middleware 2009, Urbana-Champaign, IL,USA, November 2009.
20. Hojjat Jafarpour, Sharad Mehrotra, Nalini Venkatasubramanian and Mirko Montanari, "MICS: An Efficient Content Space Representation Model for Publish/Subscribe Systems", ACM International Conference on Distributed Event-Based Systems (DEBS) 2009, Nashville, TN, USA, July 2009.
21. Hojjat Jafarpour, Sharad Mehrotra and Nalini Venkatasubramanian."Dynamic Load Balancing for Cluster-based Publish/Subscribe System", IEEE International Symposium on Applications and the Internet (SAINT) 2009, Seattle, WA, USA, July 2009.
22. Hojjat Jafarpour, Bijit Hore, Sharad Mehrotra and Nalini Venkatasubramanian. "Subscription Subsumption Evaluation for Content-based Publish/Subscribe Systems", ACM/IFIP/USENIX Middleware 2008, Leuven, Belgium, December 2008.
23. Bo Xing, Sharad Mehrotra and Nalini Venkatasubramanian. RADcast: Enabling Reliability Guarantees for Content Dissemination in Ad Hoc Networks. The 28th Conference on Computer Communications (IEEE INFOCOM 2009). April 19 - 25, 2009. Rio de Janeiro, Brazil.
24. Bo Xing, Karim Seada and Nalini Venkatasubramanian. An Experimental Study on Wi-Fi Ad-Hoc Mode for Mobile Device-to-Device Video Delivery. IEEE INFOCOM 2009 Workshop on Mobile Video Delivery (MoViD 2009). April 24, 2009. Rio de Janeiro, Brazil.
25. Bo Xing, Karim Seada and Nalini Venkatasubramanian. Proximiter: Enabling Mobile Proximity-Based Content Sharing on Portable Devices (Demo Paper). The 7th IEEE International Conference on Pervasive Computing and Communications (PerCom 2009). March 9 - 13, 2009. Galveston, Texas USA.
26. Bo Xing, Karim Seada, Peter Boda and Nalini Venkatasubramanian. PassItOn: An Opportunistic Messaging Prototype on Mobile Devices (Demo Paper). The 6th Annual IEEE Consumer Communications & Networking Conference (CCNC 2009). January 10 - 12, 2009. Las Vegas, Nevada USA.
27. Bo Xing, Sharad Mehrotra and Nalini Venkatasubramanian. RADcast: Enabling Reliability Guarantees for Content Dissemination in Ad Hoc Networks. Technical Report. August 2010.

28. Bo Xing, Sharad Mehrotra and Nalini Venkatasubramanian. Disruption-Tolerant Spatial Dissemination. Technical Report. August 2009.
29. Bo Xing, Sharad Mehrotra and Nalini Venkatasubramanian. Gateway Designation for Timely Communications in Emergency Response Mesh Networks. Technical Report. September 2009.
30. Adams, Beverley J. and McMillan, Anneley (2008), "Multi-Sensor Image Fusion Techniques for Robust Neighborhood-Scale Urban Damage Assessment," Remote Sensing for Resilient Multi-Hazard Disaster Response, Vol. III, Technical Report MCEER-08-0022, Buffalo, N.Y., November 2008
31. Adams, Beverley J. and Ronald T. Eguchi (2008), "Introduction to Damage Assessment Methodologies," Remote Sensing for Resilient Multi-Hazard Disaster Response, Vol. I, Technical Report MCEER-08-0020, Buffalo, N.Y., November 2008
32. Eguchi, Ronald T. and Hope A. Seligson, (2008), "Loss Estimation Models and Metrics," Chapter 6 in Risk Assessment, Modeling and Decision Support, Editors: A. Bostrom, S. French and S. Gottlieb.
33. Eguchi, R.T., Huyck, C.K., Ghosh, S. and Adams, B.J., (2008), "The Application of Remote Sensing Technologies for Disaster Management," Proceedings of The 14th World Conference on Earthquake Engineering, Oct 12-17th, Beijing China.
34. Friedland, C., Levitan, M., and Adams, B.J., (2008), "Sustainability of Remote Sensing for "Per-Building" Assessment of Residential Buildings Subjected to Hurricane Storm Surge," Proceedings of the 6th International Workshop on Remote Sensing for Disaster Management Applications, September 11-12, 2008, Pavia, Italy.
35. Gusella, Luca, Huyck, Charles K., Adams, Beverley J., "Counting the Number of Collapsed Buildings using an Object-Oriented Analysis: Case Study of the 2003 Bam Earthquake," Remote Sensing for Resilient Multi-Hazard Disaster Response, Vol. II, Technical Report MCEER-08-0021, Buffalo, N.Y., November 2008
36. McMillan, Anneley, Morley, J.G., Adams, Beverley J., and Chesworth, Simon, (2008), "A Study of Multi-Temporal and Multi-Resolution SAR Imagery for Post-Katrina Flood Monitoring in New Orleans," Remote Sensing for Resilient Multi-Hazard Disaster Response, Vol. IV, Buffalo, N.Y., November 2008
37. McMillan, Anneley, Adams, Beverley J., Reynolds, Amber, Brown, Tanya, Liang, Daan, and Womble, J. Arn, (2008), "Advanced Technology for Rapid Tornado Damage Assessment Following the 'Super Tuesday' Tornado Outbreak of February 2008," Special Report MCEER-08-SP01, Buffalo, N.Y., April 18, 2008.
38. McMillan, Anneley, Adams, Beverley J., Ghosh, Shubharoop, and Huyck, Charles K., (2008), "Collection and Preliminary Analysis of Aerial and In-Field Building Damage Information in the Aftermath of the 2007 California Wildfires," Special Report MCEER-08-SP03, Buffalo, N.Y., April 30, 2008.
39. Reynolds, A., Krupar, R., Brown, T., McMillan, A., Liang, D., and Womble, J.A. (2008), "February 2008 Super Tuesday Tornado Outbreak Damage Assessment," Proceedings of

the 6th International Workshop on Remote Sensing for Disaster Management Applications, September 11-12, 2008, Pavia, Italy.

40. Womble, J.A., Mehta, K., and Adams, B.J. (2008), "Integration of Remote Sensing Imagery and VIEWS™ Field Data for Post-Hurricane Charley Building Damage Assessment," Remote Sensing for Resilient Multi-Hazard Disaster Response Series, Vol. V, Technical Report MCEER-08-0024, Buffalo, N.Y., November 17, 2008.
41. Sarabandi, Pooya, Kiremidjian, Anne S., Eguchi, Ronald T., and Adams, Beverley J. (2008), Building Inventory Compilation for Disaster Management: Application of Remote Sensing and Statistical Modeling, Technical Report MCEER-08-0025, Buffalo, N.Y., November 20, 2008.
42. Svekla, Walter and Verrucci, Enrica, (2008), "Hurricane Gustav Reconnaissance: Damage Survey using the VIEWS™ System," Special Report MCEER-08-SP08, October 2008.
43. Zhaoqi Stella Chen, Dmitri V. Kalashnikov, and Sharad Mehrotra. Exploiting Context Analysis for Combining Multiple Entity Resolution Systems. In Proc. of ACM SIGMOD Int'l Conf. on Management of Data (ACM SIGMOD), June 29-July 2, 2009.
44. Dmitri V. Kalashnikov, Zhaoqi Chen, Rabia Nuray-Turan, and Sharad Mehrotra. Web people search via connection analysis. In IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), 20(11), November 2008
45. Rabia Nuray-Turan, Zhaoqi Chen, Dmitri V. Kalashnikov, and Sharad Mehrotra. Exploiting Web querying for Web People Search in WePS2. In 2nd Web People Search Evaluation Workshop (WePS 2009), 18th WWW Conference, April, 2009.
46. Dmitri V. Kalashnikov, Zhaoqi Chen, Rabia Nuray-Turan, Sharad Mehrotra, and Zheng Zhang. WEST: Modern Technologies for Web People Search. In Proc. of IEEE International Conference on Data Engineering (IEEE ICDE), demo publication, March 29 - April 4, 2009.
47. Chaitanya Desai, Dmitri V. Kalashnikov, Sharad Mehrotra, and Nalini Venkatasubramanian. Using Semantics for Speech Annotation of Images. In Proc. of IEEE International Conference on Data Engineering (IEEE ICDE), short publication, March 29 - April 4, 2009.
48. Rabia Nuray-Turan, Dmitri V. Kalashnikov, and Sharad Mehrotra. High Quality Retrieval of Probabilistically Tagged Content. Submitted to a Conference, 2009.
49. Sharad Mehrotra, Dmitri V. Kalashnikov, Nalini Venkatasubramanian, and Jie Xu. A Semantics-Based Approach for Speech Annotation of Images. Submitted to a journal, 2009.
50. J. Hutchins, A. Ihler, and P. Smyth . Probabilistic analysis of a large-scale urban traffic data set. Second International Workshop on Knowledge Discovery from Sensor Data (ACM SIGKDD Conference, KDD-08), August 2008.
51. Yiming Ma, Dmitri V. Kalashnikov, and Sharad Mehrotra. Towards Managing Uncertain Spatial Information for Situational Awareness Applications. In IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), 20(10), October 2008.

52. Naveen Ashish and Sharad Mehrotra. XAR: An Integrated Framework for Semantic Extraction and Annotation Cases on Semantic Interoperability for Information Systems Integration: Practices and Applications. Yannis Kalfoglou (Editor), IGI Global, June 2009
53. Naveen Ashish, Sharad Mehrotra and Pouria Pirzadeh. Incorporating Integrity Constraints in Uncertain Databases. CoRR:arXiv:0907:1632
54. Naveen Ashish, Jay Lickfett, Sharad Mehrotra and Nalini Venkatasubramanian. The Software EBox: Information Integration for Situational Awareness. IEEE Intelligence and Security Informatics (ISI), June 2009, Dallas.
55. Naven Ashish, Sharad Mehrota and Pouria Pirzadeh. XAR: An Integrated Framework for Free Text Information Extraction. IEEE CSIE Conference, April 2009, Los Angeles.
56. Stefano Bonetti. A framework for progressive localization query evaluation. UC Irvine, Master's thesis, March 2009
57. Chiara Chiappini. Visualization of query results in multimedia databases. UC Irvine, Master's thesis, March 2009.