

An IT Infrastructure for Responding to the Unexpected

Magda El Zarki, PhD

Ramesh Rao, PhD

Sharad Mehrotra, PhD

Nalini Venkatasubramanian, PhD

Proposal ID: 0403433

University of California, Irvine

University of California, San Diego

June 1st, 2005

Table of Contents

Table of Contents 2

AN IT INFRASTRUCTURE FOR RESPONDING TO THE UNEXPECTED 4

 Executive Summary..... 4

 Infrastructure 5

 Responsphere Management 7

 Faculty and Researchers 8

 Responsphere Research Thrusts 10

 Networks and Distributed Systems..... 10

 Privacy Preserving Media Spaces..... 10

 Activities and Findings..... 10

 Products..... 13

 Contributions..... 14

 Adaptive Data Collection in Dynamic Distributed Systems..... 14

 Activities and Findings..... 14

 Products..... 15

 Contributions..... 16

 Rapid and Customized Information Dissemination..... 20

 Activities and Findings..... 20

 Products..... 20

 Contributions..... 21

 Situation Awareness and Monitoring 23

 Activities and Findings..... 23

 Products..... 27

 Contributions..... 29

 Efficient Data Dissemination in Client-Server Environments..... 29

 Activities and Findings..... 29

 Products..... 30

 Contributions..... 30

 Power Line Sensing Nets: An Alternative to Wireless Networks for Civil Infrastructure Monitoring 30

 Activities and Findings..... 30

 Products..... 31

 Contributions..... 31

 Wireless Infrastructure/GLQ 32

 Activities and Findings..... 32

 Products..... 32

 Contributions..... 33

 Networking in the Extreme..... 36

 Activities and Findings..... 36

 Contributions..... 36

 Products..... 41

 CVC: Collaborative Visualization Center 42

 Activities and Findings..... 42

 Products..... 44

 Contributions..... 44

 CyberShuttle / CyberAmbulance..... 44

 Activities and Findings..... 44

 Products..... 49

 Contributions..... 49

 A Transportable Outdoor, Tiled Display Wall for Emergency Response 50

 Activities and Findings..... 50

 Products..... 51

 Contributions..... 51

 “Manpacks” and the Always Best Located Platform..... 51

 Activities and Findings..... 51

- Products..... 52
- Contributions..... 55
- Cellular Network Testbed..... 59
 - Activities and Findings..... 59
 - Products..... 60
 - Contributions..... 60
- Data Management and Analysis..... 62
 - Automated Reasoning and Artificial Intelligence..... 62
 - Activities and Findings..... 62
 - Products..... 63
 - Contributions..... 63
 - Supporting Keyword Search on GIS data..... 64
 - Activities and Findings..... 64
 - Products..... 64
 - Contributions..... 64
 - Supporting Approximate Similarities Queries with Quality Guarantees in P2P Systems..... 64
 - Activities and Findings..... 64
 - Products..... 65
 - Contributions..... 65
 - Statistical Data Analysis of Mining of Spatio-Temporal data for Crisis Assessment and Response..... 65
 - Activities and Findings..... 65
 - Products..... 66
 - Contributions..... 67
- Social & Organizational Science..... 68
 - RESCUE Organizational Issues..... 68
 - Activities and Findings..... 68
 - Products..... 69
 - Contributions..... 70
 - Social Science Research utilizing Responsphere Equipment..... 70
 - Activities and Findings..... 70
 - Products..... 74
 - Contributions..... 74
- Transportation and Engineering..... 76
 - CAMAS Testbed..... 76
 - Activities and Findings..... 76
 - Products..... 78
 - Contributions..... 78
 - Research to Support Transportation Testbed and Damage Assessment..... 79
 - Activities and Findings..... 79
 - Products..... 80
 - Contributions..... 81
- Additional Responsphere Papers and Publications..... 84
- Courses..... 86
- Equipment..... 87

AN IT INFRASTRUCTURE FOR RESPONDING TO THE UNEXPECTED

Executive Summary

The University of California, Irvine (UCI) and the University of California, San Diego (UCSD) received NSF Institutional Infrastructure Award 0403433 under NSF Program 2885 CISE Research Infrastructure. This award is a five year continuing grant and the following report is the Year One Annual Report.

The NSF funds from year one (\$628, 025) were split between UCI and UCSD with half going to each institution. The funds were used to begin creation of the campus-level research information technology infrastructure known as Responsphere at the UCI campus as well as beginning the creation of mobile command infrastructure at UCSD.

The results from year one include 88 research papers published in fulfillment of our academic mission. A number of drills were conducted either in the Responsphere infrastructure or equipped with Responsphere equipment in fulfillment of our community outreach mission. Additionally, we have made many contacts with the First Responder community and have opened our infrastructure to their input and advice. Finally, as part of our education mission, we have used the infrastructure equipment to teach or facilitate a number of graduate and undergraduate courses at UCI including: ICS 105 (Projects in Human Computer Interaction), ICS 123 (Software Architecture and Distributed Systems), ICS 280 (System Support for Sensor Networks), and SOC 289 (Issues in Crisis Response). At UCSD, there are two series of engineering design project courses – ECE 191 (undergraduate) and ECE 291 (graduate) that are required courses for students. During the 2004-2005 academic year, 9 undergraduate and 3 graduate level projects were facilitated using infrastructure from Responsphere. In addition, Dr Leslie Lenert gave a guest lecture as part of a symposium series titled “Public Policy and Biological Threats,” sponsored by the UCSD Institute on Global Conflict and Cooperation (Graduate School of International Relations and Pacific Studies).

From an accountancy perspective, at UCI we have approximately \$6K left in the 2004 NSF account, \$9K in the UCI cost-share account, and \$17K left in the CalIT2 cost-share account at the time of writing this report. We have taken care to create partnerships with vendors (IBM, Canon) and service providers (NACS) in order to achieve academic pricing in the creation of the infrastructures.

By the end of Year 1, UCSD plans to use the majority of the Year 1 funds allocated to them. UCSD has purchased wireless infrastructure equipment to support the Gaslamp Quarter Testbed and wireless mesh networking activities; and GPS equipment in support of location aware tracking systems (Manpack activities). In support of the mobile command and control vehicle, UCSD has continued development of Cybershuttle and also plans to purchase an outdoor tiled display wall to use as a display for the vehicle, and for other visualization purposes; UCSD has budgeted \$90,000 in year 1 for this purpose. The portability of the proposed display will enable various project teams, which participate in drills with first responders, to bring the display along in the mobile command vehicle and use it at a mobile command post during drills. The display wall will also be used as a collaborative visualization tool in the Calit2 building for research activities on other projects (e.g., ResCUE, WIISARD, OptiPuter).

Spending plans for next year at UCI include: continuation/extension of the pervasive infrastructure, increasing storage, adding computation, and beginning to build the visualization cluster. Additionally, we will host a number of drills and evacuations in the Responsphere infrastructure.

Spending plans for next year at UCSD include: construction and deployment of a multipanel visualization display (through the end of year 1) and corresponding computation equipment (year 2); continued development and enhancement of the mobile command and control vehicle; purchase and implementation of 3D scanning equipment; continuation of Manpack and Always Best Located (ABL platform development; and development and deployment of Gaslamp Quarter infrastructure. In addition, we plan to support various drills associated with the ResCUE and WIISARD projects.

Infrastructure

Responsphere is the hardware and software infrastructure for the Responding to Crisis and Unexpected Events (ResCUE) NSF-funded project. The vision for Responsphere is to instrument selected buildings and an approximate one third section of the UCI campus (see map below) with a number of sensing modalities. In addition to these sensing technologies, the researchers have instrumented this space with pervasive IEEE 802.11a/b/g Wi-Fi and IEEE 802.3 to selected sensors. They have termed this instrumented space the “UCI Smart-Space.”

The sensing modalities within the Smart-Space include audio, video, powerline networking, motion detectors, RFID, and *people counting* (ingress and egress) technologies. The video technology consists of a number of fixed Linksys WVC54G cameras (streaming audio as well as video) and several Canon VB-C50 tilt/pan/zoom cameras. These sensors communicate with an 8-processor (3Ghz) IBM e445 server. Data from the sensors is stored on an attached IBM EXP 400 with a 4TB RAID5EE storage array. This data is utilized to provide emergency response plan calibration, perform information technology research, as well as feeding into our Evacuation and Drill Simulator (DrillSim).

This budget cycle (2004-2005), the Cal-IT2 building (building 325 on map below) has been fully-instrumented. The spending plan for next budget cycle (2005-2006) is to extend the infrastructure through the rest of the 300-series buildings (see map below) and possibly the Smart-Corridor (see map below), if funding allows. We anticipate having the entire UCI Smart-Space instrumented by budget year three.

UCSD is developing fixed infrastructures at the UCSD campus and at the downtown Gaslamp Quarter in addition to a mobile infrastructure to be used for drill activities as well as joint research activities in conjunction with UCI’s Smart Space. In late September 2004, UCSD formally opened the Collaborative Visualization Center (CVC), which serves as a command and control prototyping facility as well as a collaborative research space for various other projects.

Development of the Mobile Command and Control Vehicle has continued throughout year 1 with the activities associated with Cybershuttle and CyberAmbulance. The portable visual display system planned for design and construction during the end of year 1 will act as the display platform for the mobile command vehicle.

A wireless mesh gateway with multiple relay nodes and wireless clients has been built at UCSD, where mesh networking research is being conducted in preparation for

deployment of the infrastructure in San Diego's Gaslamp Quarter. The network consists of seven Tropos wireless network access points and three Motorola Canopy backhubs.

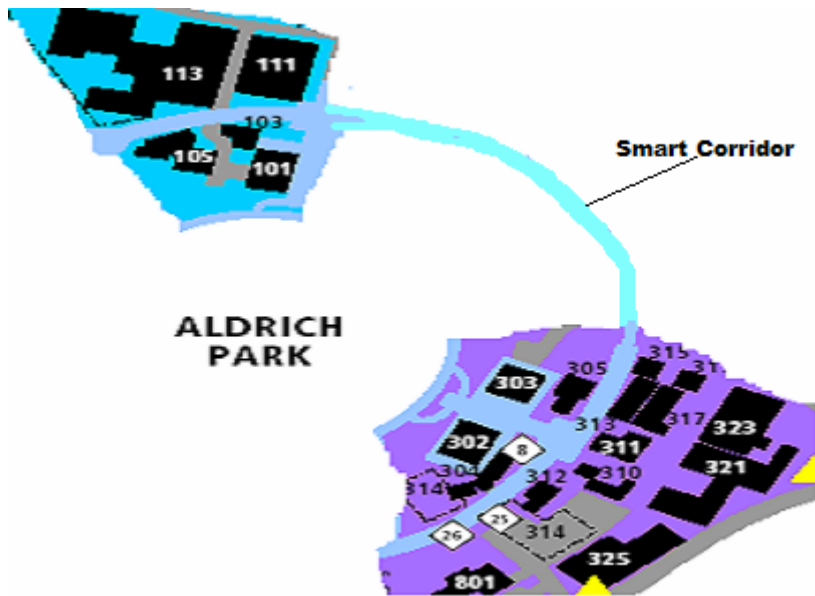
The CalRadio and Always Best Located (ABL) have been prototyped; GPS development kits and localization equipment have been purchased.

The original proposed budget requested a larger amount of funds in the first years of the grant; the intention was to purchase equipment in conjunction with the Calit2 San Diego Division's move into a newly built building. Although it was expected that we would move into our new building in January 2005, construction delays have pushed back the move-in date until approximately the end of September 2005. As a result, Responsphere has delayed in purchasing major pieces of equipment. By the end of Year 1, we expect to have spent the majority of the funds allocated to us.

In addition, we plan to purchase an outdoor tiled display wall as described in the visualization section of the report, and have budgeted \$95,000 in year 1 for this purpose. The proposed budget had requested funds for building a command and control prototyping facility; or, to supplement equipment purchased for a Collaborative Visualization Center (CVC) that was launched at UCSD Jacobs School of Engineering in September 2004. A decision was made to use Responsphere funds to purchase collaborative visualization tools in the Calit2 building for research activities on ResCUE and other projects (e.g., OptIPuter and WIISARD); and the portability of the proposed display will enable teams on various projects who participate in drills with first responders to bring the display to be used at a mobile command post during drills. As such, we request the full amount of \$202,605 for year 2 (September 2005-2006).

In fulfillment of the outreach mission of the Responsphere project, one of the goals of the researchers at the project is to open this infrastructure to the first responder community, the larger academic community including K-12, and the solutions provider community. The researchers' desire is to provide an infrastructure that can test emergency response technology and provide metrics such as evacuation time, casualty information, and behavioral models. These metrics provided by this test-bed can be utilized to provide a quantitative assessment of information technology effectiveness.

One of the ways that the Responsphere project has opened the infrastructure to the disaster response community is through the creation of a Web portal. On the www.responsphere.org website there is a portal for the community. This portal provides access to data sets, computational resources and storage resources for disaster response researchers, contingent upon their complying with our IRB-approved access protocols. At the time of writing this report, IRB is currently reviewing our protocol application.



Responsphere Management

The researchers from the Responsphere and ResCUE projects have deemed it necessary to hire a professional management staff to oversee the operations of both projects. The management staff consists of a Technology Manager, Project Manager, and Program Manager at UCI. At UCSD, the management staff consists of a Project Manager and Project Support Coordinator.

As the Responsphere infrastructure is a complex infrastructure, it was necessary to hire managers with technical as well and managerial skills. In addition to having MBAs, the managers have a diverse technical skill set including: Network Management, Technology Management, VLSI design, and cellular communications. This skill set is crucial to the design, specification, purchasing, deployment, and management of the Responsphere infrastructure.

Part of the executive-level decision making involved with accessing the open infrastructure of Responsphere (discussed in the Infrastructure portion of this report) is the specification of access protocols. Responsphere management has decided on a 3-tiered approach to accessing the services provided to the first responder community as well as the disaster response and recovery researchers.

Tier 1 access to Responsphere involves a read-only access to the data sets as well as limited access to the drills, software and hardware components. To request Tier 1 access, the protocol is to submit the request, via www.responsphere.org, and await approval from the Responsphere staff as well as the IRB in the case of federally funded research. Typically, this access is for industry affiliates and government partners under the supervision of Responsphere management.

Tier 2 access to Responsphere is reserved for staff and researchers specifically assigned to the ResCUE and Responsphere grant. This access, covered by the affiliated Institution's IRB, is more general in that hardware, software, as well as storage capacity can be utilized for research. This level of access typically will have read/write access to the data sets, participation or instantiation of drills, and configuration rights to most equipment. The protocol to obtain Tier 2 access begins with a written request on behalf

of the requestor. Next, approval must be granted by the Responsphere team and, if applicable, by the responsible IRB.

Tier 3 access to Responsphere is reserved for Responsphere technical management and support. This is typically “root” or “administrator” access on the hardware. Drill designers could have Tier 3 access in some cases. The Tier 3 access protocol requires that all Tier 3 personnel be UCI or UCSD employees and cleared through the local IRB.

Faculty and Researchers

[Naveen Ashish](#), Visiting Assistant Project Scientist, UCI

Sheldon Brown, Professor of Visual Arts; Director, Center for Research in Computing and the Arts (CRCA), UCSD

[Carter Butts](#), Assistant Professor of Sociology and the Institute for Mathematical Behavioral Sciences, UCI

Howard Chung, ImageCat, Inc.

Ganapathy Chockalingam, Principal Development Engineer, Calit2 UCSD Division

Remy Cross, Graduate Student, UCI

[Mahesh Datt](#), Graduate Student, UCI

Greg Dawe, Principal Development Engineer, Calit2 UCSD Division

Rina Dechter, Professor, UCI

Tom DeFanti, Research Scientist-Electronic Visualization, Calit2 UCSD Division

[Mayur Deshpande](#), Graduate Student, UCI

Raheleh Dilmaghani, Graduate Student, UCSD

[Ronald Eguchi](#), President and CEO, ImageCat, Inc.

[Magda El Zarki](#), Professor of Computer Science, UCI, PI Responsphere

Vibhav Gogate, Graduate Student, UCI

[Qi Han](#), Graduate Student, UCI

[Ramaswamy Hariharan](#), Graduate Student, UCI

Greg Hidley, Chief Infrastructure Officer, Calit2 UCSD Division

Bill Hodgkiss, Associate Director, Calit2 UCSD Division

[Bijit Hore](#), Graduate Student, UCI

David Hutches, Director of Engineering Computing, UCSD

John Hutchins, Graduate Student, UCI

[Charles Huyck](#), Senior Vice President, ImageCat, Inc.

Babak Jafarian, Senior Development Engineer, Calit2 UCSD Division

[Ramesh Jain](#), Bren Professor of Information and Computer Science, UCI

Aaron Jow, Graduate Student, UCSD

[Dmitri Kalashnikov](#), Post-Doctoral Researcher, UCI

Donald Kimball, Principal Development Engineer, Calit2 UCSD Division

[Iosif Lazaridis](#), Graduate Student, UCI

Leslie Lenert, Associate Director for Medical Informatics, Calit2 UCSD Division; Professor of Medicine, UCSD

[Chen Li](#), Assistant Professor of Information and Computer Science, UCI

[Yiming Ma](#), Graduate Student, UCI

BS Manoj, Postdoctoral Researcher, Calit2 UCSD Division

[Gloria Mark](#), Associate Professor of Information and Computer Science, UCI

[Daniel Massaguer](#), Graduate Student, UCI

[Sharad Mehrotra](#), RESCUE Project Director, Professor of Information and Computer Science, UCI

[Amnon Meyers](#), Programmer/Analyst, UCI

John Miller, Senior Development Engineer, Calit2 UCSD Division

Rajesh Mishra, Ericsson Researcher, Calit2 UCSD Division

[Shivajit Mohapatra](#), Graduate Student, UCI

Javier Rodriguez Molina, Undergraduate Student, UCSD

Diep Ngoc Nguyen, Undergraduate Student, UCSD

Douglas Palmer, Principal Development Engineer, Calit2 UCSD Division

Sangho Park, Postdoctoral Researcher, UCSD

Stephen Pasco, Senior Development Engineer, Calit2 UCSD Division

Miruna Petrescu-Prahova, Graduate Student, UCI

Ramesh Rao, PI, Director of Calit2 UCSD Division, Qualcomm Professor of Electrical and Computer Engineering

[Vinayak Ram](#), Graduate Student, UCI

[Will Recker](#), Professor of Civil and Environmental Engineering, Advanced Power and Energy Program, UCI

Nitesh Saxena, Graduate Student, UCI

[Dawit Seid](#), Graduate Student, UCI

[Masanobu Shinozuka](#), Chair and Distinguished Professor of Civil and Environmental Engineering, UCI

[Michal Shmueli-Scheuer](#), Graduate Student, UCI

[Padhraic Smyth](#), Professor of Information and Computer Science, UCI

Jeanette Sutton, Natural Hazards Research and Applications Information Center, University of Colorado at Boulder

[Nalini Venkatasubramanian](#), Associate Professor of Information and Computer Science, UCI

Kathleen Tierney, Professor of Sociology, University of Colorado at Boulder

Bob Welty, Director of Homeland Security Projects, San Diego State University (SDSU) Research Foundation

[Jehan Wickramasuriya](#), Graduate Student, UCI

[Xingbo Yu](#), Graduate Student, UCI

Responsphere Research Thrusts

The following research and research papers were facilitated by the Responsphere Infrastructure, or utilized the Responsphere equipment:

Networks and Distributed Systems

Privacy Preserving Media Spaces

Activities and Findings.

Emerging sensing, embedded computing, and networking technologies have created opportunities to blend computation with the physical world and its activities. Utilizing the National Science Foundation Responsphere grant, we have created a test-bed that allows us to create a blended world at UC, Irvine. The resulting pervasive environments offer numerous opportunities including customization (e.g. personalized advertising), automation (e.g. inventory tracking and control) and access control (e.g. biometric authentication, triggered surveillance). Our interest, in this project is on human-centric pervasive environments in which human subjects are embedded in the pervasive space. This naturally leads to concerns of privacy. Many types of privacy challenges can be identified (a) privacy of individuals from the environment, (b) privacy of individuals from other users (e.g. eavesdropping) and (c) privacy of interactions among individuals (communication etc.). Our focus, at present, is on privacy concerns stemming from interactions between individuals and the environment.

Below we list our contributions to the architecture for privacy preserving pervasive spaces, techniques to query over encrypted data (an important component of our privacy preserving technique), methods to quantify privacy, as well as, our efforts on building a privacy preserving video surveillance system.

Privacy-Preserving Triggers for Pervasive Spaces: Providing pervasive services often necessitates gathering information about individuals that may be considered sensitive. In this work, we consider a trigger-based pervasive environment in which end-user services are built using triggers over events detected through sensors. Privacy concerns in such a pervasive space are analogous to disclosure and misuse of event logs. We use cryptographic techniques to prevent loss of privacy, but that raises a fundamental challenge of evaluating triggers over the encrypted data representation. We address this challenge for a class of triggers that are powerful enough to realize multiple pervasive functionalities. Specifically, we consider a limited class of historical triggers which we refer to as *counting triggers*. The (test) condition in such triggers is of the form: the Sum of X_i is equal to Δ , where X_i is the transaction attribute and Δ is the threshold. Such triggers can be used for both discrete threshold detection (e.g., a person should not enter

the third floor of a building more than five times) and cumulative threshold detection (e.g., an employee's daily travel costs should not exceed more than \$100 per day total). Using secret-sharing techniques from applied cryptography, we devise protocols to test such conditions in a way that the user data is not accessible or viewable up until the time at which the condition (or set of conditions) are met. Our approach is especially useful in the context of implementing access control policies of the pervasive space. Using our approach, it is guaranteed that the adversary (i.e., people with access to the servers and logs of the pervasive space) does not know any additional information about individuals except what it can decipher from the knowledge of trigger execution. For instance, if a trigger is used to implement an access control policy, the adversary will learn nothing other than the fact that the policy has been violated.

While the nature of triggers for which our technique is designed is limited to counting triggers, we show that such triggers are powerful enough to support a variety of functionalities in the surveillance application.

A Privacy-Preserving Index for Range Queries: In this paper we address privacy threats when relational data is stored over untrusted servers which must provide efficient access to multiple clients. Specifically, we analyze the data partitioning (bucketization) technique and algorithmically develop this technique to build privacy-preserving indices on sensitive attributes of a relational table. Such indices enable an untrusted server to evaluate obfuscated range queries with minimal information leakage. We analyze the worst-case scenario of inference attacks that can potentially lead to breach of privacy (e.g., estimating the value of a data element within a small error margin) and identify statistical measures of data privacy in the context of these attacks. We also investigate precise privacy guarantees of data partitioning which form the basic building blocks of our index. We then develop a model for the fundamental privacy-utility tradeoff and design a novel algorithm for achieving the desired balance between privacy and utility (accuracy of range query evaluation) of the index .

Querying Encrypted XML documents: This work proposes techniques to query encrypted XML documents. Such a problem predominantly occurs in environments when XML data is stored over untrusted servers which must support efficient access to clients. Encryption offers a natural solution to preserve the privacy of the client's data. The challenge now is to execute queries over the encrypted data, without decrypting them at the server side. The first contribution is the development of the security mechanisms on the XML documents, which help the client to encrypt portions or totality of the XML documents. Techniques to run SPJ(Selection-projection-join) over encrypted XML documents are analyzed. A strategy, where indices/ancillary information is maintained along with the encrypted XML documents is exploited, which helps in pruning the search space during query processing.

A Systematic Search Method for Optimal K-anonymity. Fueled by the advancements in network and storage technologies, organizations are collecting huge amounts of information pertaining to individuals. Great demands are placed on the organizations to share the personal information collected. Improper exposure of information raises serious privacy concerns and could potentially have serious consequences. The notion of K-anonymity was introduced to protect privacy concerns. A K-anonymized dataset has the property, where any tuple in the dataset is indistinguishable from k-1 other tuples. Protecting the privacy of the individuals is not the only goal, the data needs to be

information rich as much as possible for its subsequent use. Recent approaches for k-anonymizing the dataset were based on attribute generalization where attribute values were replaced by coarser representations and suppression where entire tuples are deleted. These approaches tend to form clusters of tuples and all the tuples are represented by their cluster level values. In this paper we will show that attribute generalization approaches have significant information loss and we introduce the tuple generalization approach. In tuple generalization approach, tuples are generalized instead of the attributes. Tuple generalization is a data centric approach which clusters similar tuples in a fashion that releases more information than that was possible due to the attribute generalization. Tuple generalization approaches explore a bigger solution space thereby increasing the chances of finding better solutions. We conduct experiments to validate our claim that tuple generalization approaches indeed do produce better solutions than previous methods proposed.

Privacy Protecting Data Collection In Media Spaces: Around the world as both crime and technology become more prevalent, officials find themselves relying increasingly on video surveillance as a cure-all in the name of public safety. Used properly, video cameras help expose wrongdoing but typically come at the cost of privacy to those not involved in any maleficent activity. What if we could design intelligent systems that are more selective in what video they capture, and focus on anomalous events while protecting the privacy of authorized personnel?

This work proposes a novel way of combining sensor technology with traditional video surveillance in building a privacy protecting framework that exploits the strengths of these modalities and complements their individual limitations. Our fully functional system utilizes off the shelf sensor hardware (i.e. RFID, motion detection) for localization, and combines this with a XML-based policy framework for access control to determine violations within the space. This information is fused with video surveillance streams in order to make decisions about how to display the individuals being surveilled. To achieve this, we have implemented several video masking techniques that correspond to varying user privacy levels. These results were achievable in real-time at acceptable frame rates, while meeting our requirements for privacy preservation.

Privacy-protecting Video Surveillance. Forms of surveillance are very quickly becoming an integral part of crime control policy, crisis management, social control theory and community consciousness. In turn, it has been used as a simple and effective solution to many of these problems. However, privacy-related concerns have been expressed over the development and deployment of this technology. Used properly, video cameras help expose wrongdoing but typically come at the cost of privacy to those not involved in any maleficent activity. As video surveillance becomes more accessible and pervasive, systems are needed that can provide automated techniques for surveillance which preserve the privacy of subjects as well as aid in detecting anomalous behavior by filtering out 'normal' behavior. Here we describe the design and implementation of our privacy protecting video surveillance Responsphere infrastructure that fuses additional sensor information (i.e. RFID, motion detection) with video streams in order to make decisions about how and when to display the individuals being surveilled. In this paper, we describe the overall architecture of our fully implemented, real-time system but concentrate on the video surveillance subsystem. In particular, enhancements to the

algorithms for object detection, tracking and masking are provided, as well as preliminary work on handling the fusion of multiple surveillance streams.

Dynamic Access Control for Ubiquitous Environments. Current ubiquitous computing environments provide many kinds of information. This information may be accessed by different users under varying conditions depending on various contexts (e.g. location). These ubiquitous computing environments also impose new requirements on security. The ability for users to access their information in a secure and transparent manner, while adapting to changing contexts of the spaces they operate in is highly desirable in these environments. This paper presents a domain-based approach to access control in distributed environments with mobile, distributed objects and nodes. We utilize a slightly different notion of an object's ``view'', by linking its context to the state information available to it for access control purposes. In this work, we tackle the problem of hiding sensitive information in insecure environments by providing objects in the system a view of their state information, and subsequently manage this view. Combining access control requirements and multilevel security with mobile and contextual requirements of active objects allow us to re-evaluate security considerations for mobile objects. We present a middleware-based architecture for providing access control in such an environment and view-sensitive mechanisms for protection of resources while both objects and hosts are mobile. We also examine some issues with delegation of rights in these environments. Performance issues are discussed in supporting these solutions, as well as an initial prototype implementation and accompanying results.

Products.

1. J. Wickramasuriya, M. Datt, S. Mehrotra and N. Venkatasubramanian
Privacy Protecting Data Collection In Media Spaces.
ACM International Conference on Multimedia (ACM Multimedia 2004) New York, NY,
Oct. 2004
2. J. Wickramasuriya and N. Venkatasubramanian
Dynamic Access Control for Ubiquitous Environments.
International Symposium on Distributed Objects and Applications (DOA 2004)
Agia Napa, Cyprus, Oct. 2004
3. J. Wickramasuriya, M. Alhazzazi, M. Datt, S. Mehrotra and N. Venkatasubramanian
Privacy-protecting Video Surveillance.
SPIE International Symposium on Electronic Imaging (Real-Time Imaging IX)
San Jose, CA, Jan. 2005
4. Bijit Hore, Sharad Mehrotra, Gene Tsudik.
A Privacy-Preserving Index for Range Queries (VLDB 2004)
5. Jehan Wickramasuriya, Mahesh Datt, Bijit Hore, Stanislaw Jarecki, Sharad Mehrotra,
Nalini Venkatasubramanian.
Privacy-Preserving Triggers for Pervasive Spaces.
Technical Report, 2005 (Submitted for publication)

6. B. Hore, R. Jamalla, and S. Mehrotra.
A Systematic Search Method for Optimal K-anonymity
Technical Report, 2005.

7. R. Jammalla, S. Mehrotra.
Querying Encrypted XML documents.
Technical Report 2005.

Contributions.

Quantifying Privacy & Information Disclosure: we have explored measures to quantify privacy and confidentiality of data in the setting of pervasive environments. Specifically, measures of anonymity as well information theoretic measures to quantify privacy have been developed. Techniques to efficiently transform data via perturbation/blurring/bucketization to maintain a balance between privacy and data utility have been developed.

Privacy Preserving Architecture for pervasive spaces: We developed an approach to realizing privacy preserving pervasive spaces in the context of triggered pervasive architectures. In such an environment, pervasive functionalities are built as triggers over logs of events captured via the sensor infrastructure in the pervasive space. The problem of privacy preservation constitutes methods for privacy preserving event detection, privacy preserving condition checking over event logs, and action execution. Techniques to support a large class of triggers without violating privacy of individuals immersed in the pervasive space have been developed. An integral aspect of our approach is the creation of techniques to evaluate queries over encrypted data representation.

Prototype system Implementation: We have developed a prototype system for privacy preserving data collection in pervasive environments at Responsphere. Using the prototype, a fully functional video surveillance system has been developed. The system is operational and use in the instrumented Smart-Space of the Calit2 building at UC, Irvine.

Adaptive Data Collection in Dynamic Distributed Systems

Activities and Findings.

The primary goal of our work in the past year within this project is to further investigate system support for dynamic data collection in highly heterogeneous environments and under extreme situations (presence of faults, timeliness guarantees etc.). In the kinds of dynamic data collection and processing systems that we are dealing with, there are numerous overlapping and conflicting concerns, such as the availability of energy in wireless devices, the availability of bandwidth, the availability of the physical infrastructure during crises, competition between different kinds of applications for resources, and the tension between the need for simple and robust “best-effort” solutions vs. more complex and semantically complete ones. This represents a huge design space, and it is a challenge to account for all these concerns while retaining the workability and simplicity of the proposed solutions.

In particular, we have developed techniques for the collection and processing of data from heterogeneous dynamic data sources, such as sensors and cell phones, and the

querying of such data stored imprecisely in databases. We have developed techniques and protocols for supporting a *wide range of queries* in sensor networks including aggregate queries general SQL queries, in-network generated queries and continuous queries in an energy efficient manner.

Furthermore, in a crisis, information is needed rapidly and over failing infrastructures. We also describe in this report our work in the areas of *fault tolerant and real-time data collection*. We have developed an initial prototype software for collection, processing and visualization of localization information from cell phones in a framework called CELLO.

While the current development effort focuses on obtaining location-type data, we believe the techniques developed can be generalized to other dynamic information as well.

Products.

1. Iosif Lazaridis and Sharad Mehrotra. *Approximate Selection Queries over Imprecise Data*, International Conference on Data Engineering (ICDE 2004), Boston, March 2004
2. Iosif Lazaridis, Qi Han, Xingbo Yu, Sharad Mehrotra, Nalini Venkatasubramanian, Dmitri V. Kalashnikov, Weiwen Yang. *QUASAR: Quality-Aware Sensing Architecture*, SIGMOD Record 33(1): 26-31, March 2004.
3. Qi Han and Nalini Venkatasubramanian, Information Collection Services for QoS-aware Mobile Applications, IEEE Transactions on Mobile Computing, 2005, To Appear.
4. Qi Han, Iosif Lazaridis, Sharad Mehrotra, Nalini Venkatasubramanian *Sensor Data Collection with Expected Reliability Guarantees*, The First International Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS), (in conjunction with PerCom), Kauai Island, Hawaii, March 8, 2005
5. Xingbo Yu, Ambarish De, Sharad Mehrotra, Nalini Venkatasubramanian, Nikil Dutt, Energy-Efficient Adaptive Aggregate Monitoring in Wireless Sensor Networks, submitted for publication.
6. Xingbo Yu, Sharad Mehrotra, Nalini Venkatasubramanian, SURCH: Distributed Query Processing over Wireless Sensor Networks, submitted for publication.
7. Qi Han, Sharad Mehrotra and Nalini Venkatasubramanian, Energy Efficient Data Collection in Distributed Sensor Environments, The 24th IEEE International Conference on Distributed Computing Systems (ICDCS), March 2004
8. Qi Han, Nalini Venkatasubramanian, Real-time Collection of Dynamic Data with Quality Guarantees, IEEE Tran. Parallel and Distributed Systems, in revision
9. Xingbo Yu, Koushik Niyogi, Sharad Mehrotra, and Nalini Venkatasubramanian.

Adaptive Target Tracking in Sensor Networks
 The 2004 Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDS'04), San Diego, January 2004.

Contributions.

Basic Architecture for Dynamic Data Collection: Sensor devices are promising to revolutionize our interaction with the physical world by allowing continuous monitoring and reaction to natural and artificial processes at an unprecedented level of spatial and temporal resolution. As sensors become smaller, cheaper and more configurable, systems incorporating large numbers of them become feasible. Besides the technological aspects of sensor design, a critical factor enabling future sensor-driven applications will be the availability of an integrated infrastructure taking care of the onus of data management. Ideally, accessing sensor data should be no difficult or burdensome than using simple SQL. In a SIGMOD RECORD 2004 paper we investigate some of the issues that such an infrastructure must address. Unlike conventional distributed database systems, a sensor data architecture must handle extremely high data generation rates from a large number of small autonomous components. And, unlike the emerging paradigm of data streams, it is infeasible to think that all this data can be streamed into the query processing site, due to severe bandwidth and energy constraints of battery-operated wireless sensors. Thus, sensing data architectures must become quality-aware, regulating the quality of data at all levels of the distributed system, and supporting user applications' quality requirements in the most efficient manner possible.

We evaluated the feasibility of the basic adaptive data collection premise using a target tracking application[YNMV 2004]. In this application, the goal was to determine the track of a moving target (at a determined level of approximation) through in a field of instrumented sensor (acoustic sensors, in this case). We developed a quality aware information collection protocol in sensor networks for such tracking applications. The protocol explores trade-off between energy and application quality to significantly reduce energy consumption in the sensor network thereby enhancing the lifetimes of sensor nodes. Simulation results over elementary movement patterns in a tracking application scenario strengthen the merits of our adaptive information collection framework.

Query processing over imprecise data: Utilizing the NSF-funded Responsphere infrastructure, we investigated strategies to support query processing over imprecise data. Over the past year, we have made significant progress in the quality-aware processing of a wide class of queries described below:

Selection Queries: We examined the problem of evaluating selection queries over imprecisely represented objects. General SQL statements must sometimes be posed over approximate data, due to performance considerations necessitating imprecise data caching at the query processing site, or due to inherent data imprecision. Imprecise representation may arise due to a number of factors. For instance, the represented object may be much smaller in size than the precise data (e.g., compressed versions of time series), or imprecise replicas of fast-changing objects across the network may need to be maintained (e.g., interval approximations for time-varying sensor readings). In general, it may be impossible to determine whether an imprecise object meets the selection predicate.

To make results over such data meaningful, their accuracy needs to be quantified, and an approximate result needs to be returned with sufficient accuracy. Retrieving the precise objects themselves (at additional cost) can be used to increase the quality of the reported answer. In our paper we allow queries to specify their own answer quality requirements. We show how the query evaluation system may do the minimal amount of work to meet these requirements. Our work presents two important contributions: First, by considering queries with set-based answers rather than the approximate aggregate queries over numerical data examined in the literature; second, by aiming to minimize the combined cost of both data processing and probe operations in a single framework. Thus, we establish that the answer accuracy/performance tradeoff can be realized in a more general setting than previously seen.

Continuous and Aggregation Queries: Wireless sensor networks are characterized by their capability of generating, processing, and communicating data. One important application of such networks is aggregate monitoring in which a continuous aggregate query is posed and processed within networks. Sensor devices are battery operated, limited power (energy budget) is a challenging factor in deploying wireless sensor networks. Supporting precise strategies for data aggregation can be very expensive in terms of energy. In this paper, we consider approximate aggregation in which a certain amount of error can be tolerated by users. We develop an approach to processing continuous aggregate queries in sensor networks which fully exploit the error tolerance to achieve energy efficiency. Specifically, we schedule transmission and listening operations for each sensor node to achieve collision-free communication for data aggregation. In order to minimize power consumption, the error bounds at sensor nodes are dynamically adjusted. The adjustment period is also adapted to cater to dynamic data behavior. We also allow the adjustment to be performed both locally on sections of a network and globally on the entire network. Our simulation results indicate that, with small error tolerance, about {20~60%} energy saving can be achieved from using the proposed adaptive techniques.

In-network queries: In this NSF report, we present SURCH, an novel decentralized algorithm for efficient processing of queries generated in networks. Given the ad hoc nature of these queries, there is no pre-established infrastructures (e.g., a routing tree) within sensor networks to support them. SURCH fully exploits the distributed nature of wireless sensor networks. With small amount of messages, the algorithm searches the query region specified by a query. Partial results of the query are aggregated and identified before they are delivered to a destination node for final processing. We show that in addition to its power of handling ad hoc queries generated in-network, SURCH is extremely efficient for an important type of queries that solicit data from only small portion of relevant sensor nodes. It also possesses nice features which allow it to be power-aware and failure-resilient. Performance results are presented to evaluate various policies available in SURCH. It can be shown that with a normal network setup, the algorithm saves around 60% messages as compared with other aggregation techniques.

Energy Efficiency, Fault Tolerance, Timeliness in Sensor Networks: Supporting the wide range of queries in general purpose sensor networks is a challenge. As discussed earlier, energy efficient query processing is a significant challenge in sensor networks. The challenge is even more acute when there are failures in the communication/sensing infrastructure or when there are real-time constraints on the data collection process. This

section addresses key architectural challenges.

A key issue in designing data collection systems for sensor networks is energy efficiency. Sensors are typically deployed to gather data about the physical world and its artifacts for a variety of purposes that range from environment monitoring, control, to data analysis. Since sensors are resource constrained, often sensor data is collected into a sensor database that resides at (more powerful) servers. A natural tradeoff exists between the sensor resources (bandwidth, energy) consumed and the quality of data collected at the server. Blindly transmitting sensor updates at a fixed periodicity to the server results in a suboptimal solution due to the differences in stability of sensor values and due to the varying application needs that impose different quality requirements across sensors. This paper proposes adaptive data collection mechanisms for sensor environments that adjusts to these variations while at the same time optimizing the energy consumption of sensors. Modern sensors try to be power aware, shutting down components (e.g., radio) when they are not needed in order to conserve energy. We consider a series of sensor models that progressively expose increasing number of power saving states. For each of the sensor models considered, we develop quality-aware data collection mechanisms that enable quality requirements of the queries to be satisfied while minimizing the resource (energy) consumption. Our experimental results show significant energy savings compared to the naïve approach to data collection.

The second issue is that of fault tolerance. We plan to investigate fault tolerance issues in sensor networks producing information about the world. The goal is to capture information at a high level of quality from distributed wireless sensors. But, since these sensors communicate over unreliable wireless channels and often fail, the “quality” of the captured information is uncertain. We will develop techniques to gauge this quality. Due to the fragility of small sensors, their finite energy supply and the loss of packets in the wireless channel, reports sent from sensors may not reach their destination nodes. In an IEEE PerSenS 2004 paper we consider the problem of sensor data collection in the presence of faults in sensor networks. We develop a data collection protocol, which provides expected reliability guarantees while minimizing resource consumption by adaptively adjusting the number of retransmissions based on current network fault conditions. Due to the fragility of small sensors, their finite energy supply and the loss of packets in the wireless channel, reports from sensors may not reach the sink node. In this paper we consider the problem of sensor data collection in the presence of faults in sensor networks. We develop a data collection protocol, which provides expected reliability guarantees while minimizing resource consumption by adaptively adjusting the number of retransmissions based on current network fault conditions. The basic idea of the protocol is to use re-transmission to achieve user required reliability. At the end of each round, the sensor sends the server the information about the number of data items sent in this round, and the number of messages sent in this round. Based on these information, the server estimates current network situation, as well as the actual reliability achieved in this round. The sensor derives the optimal re-transmission times for data items generated in the next round based on the feedback from the server. The third issue we address is that of supporting timeliness in the data collection process. In this work, we aim to provide a real-time context information collection service that delivers the right context information to the right user at the right time. The complexity of providing the real-time context information service arises from (i) dynamically

changing status of information sources; (ii) diverse user requirements in terms of data accuracy and service latency; and (iii) constantly changing system conditions. Building upon our prior work, we develop techniques to address the tradeoffs between timeliness, accuracy and cost for real-time information collection in distributed real-time environments.

In this work (currently in revision with a journal), we develop the problem of quality-aware information collection work in the presence of deadlines using the notions of Quality-of-Service (QoS) to express timeliness, Quality-of-Data (QoD) to express accuracy and collection cost. We propose a middleware framework for the real-time information collection that incorporates a family of algorithms that accommodate the diverse characteristics of information sources and varying requirements from information consumers. A key component of this framework is an information mediator to coordinate and facilitate communication between information sources and consumers. We develop distributed mediator architectures to support QoS and QoD. We are currently evaluating standard sensor-server architectures as well as distributed and centralized mediator architectures to study the QoS/QoD/Cost balance. We are also in the process of integrating this work in the context of AutoSEC, a prototype service composition framework being developed at UCI.

Cell phone as a sensor: Cell phones have become widespread and localization technologies, allowing their precise pinpointing in space are becoming available. During crises, a cell phone can be used to give location-specific information to users and to give information about the spatial distribution of individuals to responders. We will investigate ways to collect location data and to predict motion patterns of users despite failures in infrastructure that are likely to take place. During the summer of 2004, we developed a prototype system for gathering location information from cell phones in real time. This system used BREW (www.qualcomm.com/brew)-based clients running on cell phones which sampled the unit's location periodically, using assisted GPS technology. This data is then forwarded via data transmission (IP) to a Java-based server with a relational database back-end where it is stored. Additionally, simple motion prediction techniques are applied on the location stream to anticipate the future location of each object, and there is a mechanism for adapting the frequency at which samples are acquired. We are currently considering how to flesh out the basic architecture with adaptive data collection protocols, whereby the frequency and accuracy of the incoming data stream is modified depending on (i) the "predictability" of the user's trajectory, (ii) the interest of applications in his trajectory, and (iii) the correlation of users' location with other spatially distributed "events," such as first responder teams, or the distribution of hazards in the crisis region. This will be part of a larger reflective cellular architecture which will be resilient to disasters and fluctuations of resource availability, and be able to function in the most mission-effective manner.

Efficient resource provisioning that allows for cost-effective enforcement of application QoS relies on accurate system state information. However, maintaining accurate information about available system resources is complex and expensive especially in mobile environments where system conditions are highly dynamic. Resource provisioning mechanisms for such dynamic environments must therefore be able to tolerate imprecision in system state while ensuring adequate QoS to the end-user. In this work, we address the information collection problem for QoS-based services in

mobile environments. Specifically, we propose a family of information collection policies that vary in the granularity at which system state information is represented and maintained. We empirically evaluate the impact of these policies on the performance of diverse resource provisioning strategies. We generally observe that resource provisioning benefits significantly from the customized information collection mechanisms that take advantage of user mobility information. Furthermore, our performance results indicate that effective utilization of coarse-grained user mobility information renders better system performance than using fine-grained user mobility information. Using results from our empirical studies, we derive a set of rules that supports seamless integration of information collection and resource provisioning mechanisms for mobile environments.

Rapid and Customized Information Dissemination

Activities and Findings.

The primary goal of our work in the past year within this project is to further investigate specific issues in data dissemination to both first responders and the public. During this time, we introduced and explored a new form of dissemination that arises in mission critical applications such as crisis response, called flash dissemination. Flash dissemination is the rapid dissemination of critical information to a large number of recipients in a very short period of time. A key characteristic of flash dissemination is its unpredictability (e.g. natural hazards). A flash dissemination system is often idle, but when invoked, it must harness all possible resources to minimize the dissemination time. In addition, heterogeneity of receivers, communication networks poses challenges in determining efficient policies for rapid dissemination of information to a large population of receivers. Our goal is to explore this problem from both theoretical and pragmatic perspectives and develop optimized protocols and algorithms to support flash dissemination. We also aimed to incorporate our solutions in a prototype system.

In the next budget cycle, we also plan to understand the problem of customized dissemination more formally. Risk communication has been a topic of extensive studied in the disaster science community – it deals with models for communication to the public, the social implications of disaster warnings and insights into what causes problems such as over-reaction or under-reaction to messages in a disaster. Using this rich body of work from the social and disaster science communities, we develop IT solutions that will incorporate the above perspectives while ensuring that the information to end-users is customized appropriately and delivered reliably.

We then plan to focus on evacuation as a crisis related activity and developed techniques for :

- (a) customized dissemination to users in pervasive environments given information on user reactivity to different (and multiple) forms of communication and
- (b) customized delivery of navigation information to users with disabilities.

Products.

1. Mayur Deshpande, Nalini Venkatasubramanian. The Different Dimensions of Dynamicity. Fourth International Conference on Peer-to-Peer Systems. 2004.
2. Mayur Deshpande, Nalini Venkatasubramanian, Sharad Mehrotra. Scalable,

Flash Dissemination of Information in a Heterogeneous Network. Submitted for publication, 2005.

3. Alessandro Ghighi. Dissemination of Evacuation Information in Crisis. RESCUE Tech Report, Also M.S. Thesis, U of Bologna, (thesis work was done while the author was an exchange student at ITR-RESCUE)
4. Dissemination white paper
5. SAFE white paper.

Contributions.

Flash Dissemination in Heterogeneous Networks:

Any solution to the flash-dissemination problem must address the unpredictable nature of information dissemination in a crisis. Since the dissemination events (e.g. major disasters) are unpredictable and maybe rare - deployed solutions for flash-dissemination may be idle for a majority of the time. However, upon invocation of the flash dissemination, large amounts of resources must be (almost) instantaneously available to deliver the information in the shortest possible time. This sets up a paradoxical requirement from the system. During periods of non-use the *idling-cost* (e.g dedicated servers, background network traffic) of the system must be minimum (ideally zero) while during times of need, maximum server and network resources must be available.

In addition, flash dissemination may need to reach a very large number of recipients leading to *scalability* issues. For example, populations affected by an earthquake must know within minutes, about protective actions that must be taken to deal with aftershocks or secondary hazards (hazardous materials release) after the main shock. Given the potentially large number of recipients that connect through a variety of channels and devices, heterogeneity is a key aspect of a solution. In the earthquake example, customized information on care, shelter and first-aid must be delivered to tens of thousands of people with varying network capabilities (Dialup, DSL, Cable, cellular, T1, etc.). Flash dissemination in the presence of various forms of heterogeneity is a significant challenge. The heterogeneity is manifested in the data (varying sizes, modalities) and in the underlying systems (sources, recipients and network channels). Furthermore, varying network topologies induce heterogeneous connectivity conditions - organizational structures dictate overlay topologies and geographic distances between nodes in the physical layer influence end-to-end transfer latencies.

Designing low-idling overhead protocols with good scaling properties that can also accommodate heterogeneity is a challenge, since the basic problem of disseminating data in a heterogeneous network is itself NP-Hard. We investigated a peer based approach based on the simple maxim of transferring the dissemination load to information receivers.

Beginning with theoretical foundations from broadcast-theory, random graphs and gossip-theory, we develop protocols for flash dissemination that work under a variety of situations. Of particular interest are two protocols proposed that we proposed – DIM-RANK (a centralized greedy heuristic based approach) and CREW (Concurrent

Random Expanding Walkers) - an extremely fast, decentralized, and fault-tolerant protocol that incurs minimal (to zero) idling overhead. Further, we implemented and tested these protocols using a CORBA-based Peer to peer middleware platform called RapID. RapID is a middleware platform that we are developing for plug-and-play operation of dissemination protocols. Performance results obtained by systematically evaluating Rank-DIM and CREW with a variety of heterogeneity profiles illustrate the superiority of the proposed protocols.

The Rapid Middleware Platform:

RapID is a P2P system that provides for scalable distribution of data. A network of machines form a RapID overlay network. A sender machine can then send data/content to all the other machines in a fast, scalable fashion. RapID is agnostic of the content being disseminated. It can be used to distribute any file. The file to be disseminated is input at a command line. The file is broken into chunks and 'swarmed' into the overlay. On the receiving end, the chunks are collected in the 'right-order' (using a 'sliding window') and sent to STDOUT. The output can be piped into another program or redirected into a file (to be stored) as per the users wish. We have prototyped a family of flash dissemination protocols in Rapid including DIM, DIM-RANK, Distributed DIM and CREW.

RapID is built upon a CORBA-like middleware called ICE (www.zeroc.com). This allows us to leverage the many advantages of a CORBA based middleware platform. Given the Interfaces of the objects, the software can be developed or ported in many languages (C++/Java) and many operating systems. Lower-level network programming details are cleanly abstracted and hence prototyping is much faster and easier. This is invaluable to quickly test new modifications and ideas, and if they work well, to transition them seamlessly to release software. Abstraction of RPC calls into method invocations further reduces development time.

Further, the flexibility of CORBA allows us to use the same software base for both release versions and emulated scalability testing. For example, a Peer-Object can run be in many ways: as a single network entity that binds to a port on a machine (release version) or many Peer Objects that are all listening on the same port (testing version); all without changing any of the code within the Peer's implementation. This is possible because of network transparency and object call semantics in CORBA. Not only does this liberate us from maintaining two separate code bases but also gives greater confidence that the release system will perform as expected.

Navigation Support for Users with Disabilities:

Evacuation is an important activity in crisis response. While evacuation of general populations from affected regions is a challenging issue from an IT perspective, this proposal focuses on ensuring that IT solutions for evacuation should also consider the needs of special populations in order that the evacuation is successful. Additionally crisis scenarios can cause temporary impairments for users, and hence evacuation procedures cannot be generalized for normal populations. Technologies to aid in evacuation of users with disabilities and those impaired due to the crisis are important in crisis response. The goal of the proposed system SAFE- Systems-based Accessibility For Evacuation is to support evacuation of users with disabilities from a region that has been affected by a disaster or get first responders to the users in need.

Using several positioning and navigation technologies the system would enable individuals with sensory impairments to find their way out of a building to a designated area with available emergency personnel. For individuals physically unable to exit (due to an existing or incident acquired impairment), the proposed system would accurately report their precise location to first responders and rescue workers who could then more effectively organize rescue efforts. The system design is composed of both the hardware sensing technologies (GPS, dead reckoning, wireless signal strength/beaconing) as well as the software/middleware to provide multi-modal instructions to users in an emergency and to enable authorities to access the information in real-time when needed. The instrumented environment is expected to consist of sensors and other hardware form the localization module, which will help identifying user location and user requirements. The software design will include the design of the overall framework, a plan generation agent that will generate the evacuation plan based on the localization and user requirements. In addition the software system will also have a customization and dissemination module to send accessible information to the end users.

Over the last year, we have developed a software architecture and key techniques that will be incorporated in the software modules within SAFE. The key modules of this system are:

- a localization module for user localization,
- evacuation plan generation using inputs from users, localization information and spatial map of the information
- information customization that generates accessible information to the user using an ability based model from the cognitive science community.
- Dissemination that decides which information to broadcast to which set of users and selects the appropriate dissemination policy based on the constraints in the communication channels like bandwidth, error rate, delays etc. Based on these constraints the information disseminator can also reduce the quality of information based on accuracy, timeliness and accessibility requirements.

Situation Awareness and Monitoring

Activities and Findings.

Situation awareness and monitoring tools are important to enable real-time and off-line crisis event analysis. Such tools should allow supporting almost-real-time tracking of the actions and decisions of people involved in a large-scale event. Capturing of events as they occur enables real-time querying, prediction of subsequent events, predicting emerging disasters, etc. Such tools should facilitate the capture of event data despite the chaos during a disaster which can then be used for post-disaster analysis and reconstruction of events. Such an effort can serve various purposes including identifying factors that caused or exacerbated the situation, detecting and analyzing failures of established procedures/protocols, avoiding repetition of mistakes, better preparedness and planning for future disasters, and so on.

The primary goal of our project is to develop technologies and research solutions to support storing/representing and subsequent *analysis* of crisis events. To achieve this overall goal, during the current year we focused on the following problems:

Event DBMS.

Many contemporary scientific and business applications, and specifically situation awareness applications, operate with collections of events, where a basic event is an occurrence of a certain type in space and time. Applications that deal with events might require the support of complex domain-specific operations and involved querying of events, yet many manipulations with events are common across domains. Events can be very diverse from application to application. However they all have common attributes such as type, time, and location, whereas their diversity is reflected by the presence of auxiliary attributes. We argue that many of such applications can benefit from an Event Database Management System (EDBMS). Such a system should provide a convenient and efficient way to represent, store, and query large collections of events. EDBMSs provide a natural mechanism for representing and reasoning about evolving situations. We envision EDBMS to serve as the core of Situation Awareness tools that we are building and will also benefit larger research community. Making EDBMS a reality requires addressing multiple challenges, some of which we address during the current year and which are listed below.

Event Disambiguation.

Thus far we primarily focused on creating Situation Awareness by analyzing human-created reports in free text. The information in such reports is not well-structured, compared for example with sensory data. This creates the problem of extracting events, entities and relationships from reports. The raw datasets from which this information is extracted are often inherently ambiguous: the descriptions of entities and relationships provided in this dataset are often not sufficient to uniquely identify/disambiguate them. Similar problem with datasets are fairly generic and frequently found in other areas of research. For example, if a particular report refers to a person “J. Smith”, the challenge might be to decide which “J. Smith” it refers to. Similarly, a city named “Springfield” may refer to one of many different possibilities (there are 26 different cities by this name in continental USA). Such problems with data are studied by the research area known as *information quality and data cleaning*. Thus our goal was to look into methods to support event disambiguation.

Similarity Retrieval and Analysis.

Most of the event disambiguation methods, in turn, rely on similarity retrieval/analysis methods at the intermediate stages of their computations. These methods typically have wider applicability than merely disambiguation. Our goal was to look into such methods and in particular into the methods that analyze relationships to measure similarity and methods that include relevance feedback (from humans) in the loop. When humans describe events in free text, the information about properties of the events, such as space and time, are often uncertain. For example, a person might specify the location of an event as being ‘near’ a major landmark, without specifying the exact point-location. To disambiguate events we should be able to compare two events based on their properties, and consequently we must be able to

measure similarity between attributes of the events, including (uncertain) spatial and temporal event attributes.

Graph Analysis Methods.

Events, in our model form a web (or a graph) based on various relationships. The graph view of events, referred to as EventWeb, provides a powerful mechanism to query and analyze events in the context of disaster analytics and situation understanding. Our goal was to study general purpose graph languages and graph algebra that would allow creating EventWeb, manipulations with EventWeb, and provide support for analytical queries on top of it.

Major Findings for Situation Awareness and Monitoring Research:

Spatial queries over (imprecise) event descriptions.

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. To address this issue, we propose an approach to model and represent (potentially uncertain) event locations described by human reporters in the form of free text. We analyze several types of spatial queries of interest in SA applications. We design techniques to store and index uncertain spatial information, to support the efficient processing of queries. Our extensive experimental evaluation over synthetic and real datasets demonstrates the efficiency of our approach.

Exploiting relationships for domain-independent event disambiguation.

We address the problem known as “reference disambiguation”, which is a subproblem of the event disambiguation problem. The overall task of event disambiguation is to address the uncertainty that arises after information extraction, to populate EventWeb. Specifically, we consider a situation where entities in the database are referred to using descriptions (e.g., a set of instantiated attributes). The objective of reference disambiguation is to identify the unique entity to which each description corresponds. The key difference between the approach we propose (called RED) and the traditional techniques, is that RED analyzes not only object features but also inter-object relationships to improve the disambiguation quality. Our extensive experiments over two real datasets and also over synthetic datasets show that analysis of relationships significantly improves quality of the result.

Learning importance of relationships for reference disambiguation.

The RED framework analyzes relationships for event disambiguation. It defines the notion of “connection strength” which is a similarity measure. To analyze relationships our approach utilizes a model that measures the degree of interconnectivity of two entities to each other via chains of relationships that exist between them. Such a model is a key component of the proposed technique. We concentrate on a method for *learning* such a model by learning the importance of different relationships directly from the data itself. This allows our reference disambiguation algorithm to automatically adapt to different datasets being cleaned.

Capturing uncertainty via probabilistic ARGs.

Many data mining algorithms view datasets as attributed relation graphs (ARGs) in which nodes correspond to entities and edges to relationships. However, frequently

real-world datasets contain uncertainty and can be represented only as probabilistic ARGs (pARGs), i.e. as ARGs where each edge has associated with it a non-zero probability of that edge to exist - probabilistic edge. We introduce a novel concept of probabilistic ARGs and then analyze some of its interesting properties.

Exploiting relationships for object consolidation.

We address another event disambiguation challenge, commonly known as “object consolidation”, using the RED framework. This important challenge arises because objects in datasets are frequently represented via descriptions (a set of instantiated attributes), which alone might not always uniquely identify the object. The goal of object consolidation is to correctly consolidate (i.e., to group/determine) all the representations of the same object, for each object in the dataset. In contrast to traditional domain-independent object consolidation techniques, our approach analyzes not only object features, but also additional semantic information: inter-objects relationships. The approach views datasets as attributed relational graphs (ARGs) of object representations (nodes), connected via relationships (edges). The approach then applies graph partitioning techniques to accurately cluster object representations. Our empirical study over real datasets shows that analyzing relationships significantly improves the quality of the result.

Efficient Relationship Pattern Mining using Multi-relational Iceberg-cubes

Multi-relational data mining (MRDM) is concerned with data that contains heterogeneous and semantically rich relationships among various entity types. We introduce multi-relational iceberg-cubes (MRI-Cubes) as a scalable approach to efficiently compute data cubes (aggregations) over multiple database relations and, in particular, as mechanisms to compute frequent multi-relational patterns (“itemsets”). We also present a summary of performance results of our algorithm.

GAAL: A Language and Algebra for Complex Analytical Queries over Large Attributed Graph Data.

Attributed graphs are widely applicable data models that represent both complex linkage as well as attribute data of objects. We define a query language for attributed graphs that builds on a three-tiered recursive attributed graph query algebra. Our algebra, in addition to recursively extending basic relational operators, provides a well-defined semantics for graph grouping and aggregation that perform structural transformation and aggregation in a principled manner. We show that a wide range of complex analytical queries can be expressed in a succinct and precise manner by composing the powerful set of primitives constituting our algebra. Finally, we discuss implementation and optimization issues.

PSAP: Personalized Situation Awareness Portal for Large Scale Disasters.

A large scale disaster can cause wide spread damage, its effects can last a long period of time. Effective response to such disasters relies on obtaining information from multiple, heterogeneous data sources to help assess the geo-graphic extent and relative severity of the disaster. One of the key problems faced by analysts trying to scope out and estimate the impact of a disaster is that of information overload. In addition to well established sources of information (e.g. USGS), that continuously stream data, potentially useful public information sources such as Internet news reports and blogs swell exponentially at the time of a disaster, resulting in enormous redundancy and clutter. PSAP (Personalized Situation Awareness Portal) is a system

that helps analysts to quickly and effectively gather task-specific, personalized information over the Internet. PSAP uses a standard three tier server/client architecture where the server incrementally gathers data from a variety of sources (satellite imagery, international news reports, Internet blogs) and performs extraction of spatial and temporal information on the assimilated data. These materials are further organized in automatically generated topics and stored in the server database. The personalized portal interface gives analysts a flexible way to view the personalized information stored on PSAP server based on the user specific information needs. We focus on the design and implementation issues of PSAP, and evaluate our system with the real-life analytical tasks using information gathered about the most recent Tsunami disaster in the south-east Asia since its onset.

I-Skyline: A Completeness Guaranteed Interactive SQL Similarity Query Framework.

A query model of a similarity query combines a set of similarity predicates using an operator tree. Return set of a similarity query consists of a ranked list of objects that best matches the query model. The query model is constructed by a user, but model parameters or model itself can be inaccurately specified in the beginning of the search. The question of similarity query completeness is whether a system can guarantee that all the relevant tuples under the ideal query model be retrieved. We address this problem by first defining the concept of similarity query completeness; we then propose a novel interactive framework called interactive skyline (I-Skyline), which only shows an optimal set of tuples to the user and always guarantees to retrieve all the relevant tuples. However, the size of I-Skyline can be large. To reduce the I-Skyline size and still be able to guarantee certain level of completeness, we introduce a progressive I-Skyline extension that explores user knowledge on the query model. Experimental analysis confirms that I-Skyline is an effective and efficient framework for guaranteeing similarity query completeness.

A Framework for Refining Structured Similarity Queries Using Learning Techniques.

In numerous applications that deal with similarity search, a user may not have an exact idea of his information need and/or may not be able to construct a query that exactly captures his notion of similarity. A promising approach to mitigate this problem is to enable the user to submit a rough approximation of the desired query and use the feedback on the relevance of the retrieved objects to refine the query. We explore such a refinement strategy for a general class of SQL similarity queries. Our approach casts the refinement problem as that of learning concepts using examples. This is achieved by viewing the tuples on which a user provides feedback as a labeled training set for a learner. Under this setup, SQL query refinement consists of two learning tasks, namely learning the structure of the SQL query and learning the relative importance of the query components. We develop appropriate machine learning approaches suitable for these two learning tasks. The primary contribution of our work is a general refinement framework that decides when each learner is invoked in order to quickly learn the user query. Experimental analyses over many real-world datasets and queries show that our strategy outperforms the existing approaches significantly in terms of retrieval accuracy and query simplicity.

Products.

1. Dmitri V. Kalashnikov, Yiming Ma, Ramaswamy Hariharan, and Sharad Mehrotra
Spatial queries over (imprecise) event descriptions
Submitted to VLDB 2005
2. Dmitri V. Kalashnikov, Sharad Mehrotra, and Zhaoqi Chen.
Exploiting relationships for domain-independent data cleaning.
In SIAM International Conference on Data Mining (SIAM SDM), Newport Beach, CA, USA, April 21-23 2005
3. Dmitri V. Kalashnikov and Sharad Mehrotra.
Exploiting relationships for domain-independent data cleaning.
Submitted to ACM TODS journal (ext. ver. of SDM-05) available as TR-RESCUE-04-21, Dec 2004.
4. Dmitri V. Kalashnikov and Sharad Mehrotra.
Learning importance of relationships for reference disambiguation.
TR-RESCUE-04-23, Dec 8, 2004.
5. Dmitri V. Kalashnikov and Sharad Mehrotra.
Computing connection strength in probabilistic graphs.
TR-RESCUE-04-22, Dec 8, 2004.
6. Zhaoqi Chen, Dmitri V. Kalashnikov and Sharad Mehrotra.
Exploiting relationships for object consolidation.
Submitted to ACM SIGMOD IQIS Workshop, available as TR-RESCUE-05-01, Jan 7, 2005.
7. D. Y. Seid and S. Mehrotra.
Efficient Relationship Pattern Mining using Multi-relational Iceberg-cubes.
Fourth IEEE International Conference on Data Mining (ICDM'04), pp. 515-518, 2004.
8. D. Y. Seid and S. Mehrotra.
GAAL: A Language and Algebra for Complex Analytical Queries over Large Attributed Graph Data.
Submitted to 31st Int. Conf. on Very Large Data Bases (VLDB'05), 2005.
9. Y. Ma, R. Hariharan, A. Meyers, Y. Wu, M. Mio, C. Chemudugunta, S. Mehrotra
N. Venkatasubramanian, P. Smyth
PSAP: Personalized Situation Awareness Portal for Large Scale Disasters
Submitted to VLDB 2005 (demo publication)
10. Yiming Ma, Qi Zhong, Sharad Mehrotra
I-Skyline: A Completeness Guaranteed Interactive SQL Similarity Query Framework
Tech report, 2004
11. Yiming Ma, Dawit Seid, Qi Zhong, Sharad Mehrotra

A Framework for Refining Structured Similarity Queries Using Learning Techniques
 In Proc of CIKM Conf, 2004 (poster publication)

Contributions.

Progress has been achieved as specified in the goal for the current year. Both, novel tools/technologies as well as new research solutions have been proposed. The most significant progress has been achieved in the following areas:

1. **High-level design of the Event DBMS.** Event DBMS is rather a long term project, which requires progress in several research directions before it can be fully implemented. The ultimate goal of the project is to enable situation representation, awareness, and analytical querying of events that constitute a large crisis event. We have designed a high-level representation of Event DBMS. Specifically, we propose the EventWeb view of the Event DBMS.
2. **Event disambiguation.** We have achieved progress in designing event disambiguation methods. Specifically we proposed novel Relationship-based Event Disambiguation (RED) framework that is currently capable of addressing two important disambiguation challenges, known as “reference disambiguation” and “object consolidation” in the literature. While traditional methods, at the core, employ feature-based similarity to address the problem, RED proposed to enhance the core of these techniques by also analyzing relationships that exist between entities in the EventWeb.
3. **Similarity retrieval and analysis.** We have designed several techniques for similarity retrieval via relevance feedback as well as methods for measuring the similarity of uncertain spatial information of events. The latter is necessary for event disambiguation and enables querying of (uncertain) spatial event information.
4. **Graph language and algebra.** Given EventWeb is a graph, we have successfully researched and implemented a novel graph language and algebra that allows creation, manipulation and complex analytical queries on top of graphs in general, and specifically on the EventWeb.

Efficient Data Dissemination in Client-Server Environments

Activities and Findings.

We studied the following problem related to data dissemination. Consider a client-server environment that has dynamically changing data on the server, and queries on the client need to be answered using the latest server data. Since the queries and data updates have different “hot spots” and “cold spots,” we want to combine the PUSH paradigm and the PULL paradigm in order to reduce the communication cost. This problem is important in many client-server applications, such as traffic report systems, mediation systems, etc. In the PUSH regions, the server notifies the client about updates, and in the PULL regions the client sends queries to the server. We call the problem of finding the minimum possible communication cost by partitioning the data space into PUSH/PULL regions for a given workload the problem of data gerrymandering. We present solutions under different communication cost models

for a frequently encountered scenario: with range queries and point updates. Specifically, we give a provably optimal-cost dynamic programming algorithm for gerrymandering on a single range query attribute. We propose a family of heuristics for gerrymandering on multiple range query attributes. We also handle the dynamic case in which the workload evolves over time. We validate our methods through extensive experiments on real and synthetic data sets.

In addition to working on the gerrymandering problem, we are also looking at other related variations, including: (1) the server is “passive,” i.e., it does not do any PUSH; (2) the client allows answers using stale data. We study how to answer client queries using server data to minimize the communication cost while maximizing certain goodness function.

Products.

A paper was submitted to a conference. We are trying to finish a workshop paper soon.

Contributions.

Achieved: Systematic study on different algorithms for push/pull-based data dissemination. Developed a family of gerrymandering algorithms. Conducted extensive experiments. Submitted a research paper.

We have developed a family of algorithms for different cases to do data gerrymandering. We have conducted extensive experiments to compare these algorithms. We have written a research paper submitted to a conference.

We are currently investigating research problems related to this setting.

Power Line Sensing Nets: An Alternative to Wireless Networks for Civil Infrastructure Monitoring

Activities and Findings.

We studied the potential use of power lines as a communication medium for sensor networks based in civil infrastructure environments.

The main drawback of the existing sensing technologies is the fact that the sensors are invariably battery powered and thus have a limited lifetime. Further, wireless technologies have a limited range of communication. These two drawbacks are the main motivation for seeking an alternative sensing technology for specific sensing environments. Civil Infrastructures typically are connected to the power grid and have abundant power lines that run through the structure.

For sensing in such environments, it seems obvious to use the power grid to power the sensors but in addition, one can also use the power lines to build a secure and reliable communications network. In other words, using power lines, it is possible to develop sensors that draw power as well as transmit data using the same wire. Since batteries are eliminated in this case (only used as backup when power loss occurs) the life of the sensors is limited only by the event of a mechanical or electrical failure in the sensor. Further, it has been shown that data can be easily propagated over the power lines over long distances.

We propose to use the LonWorks Technology by Echelon for Powerline Communication. Our intent is to prototype and build integrated sensing nodes which comprise of the Neuron Chip, the power line transceiver and the power line coupler. These sensing nodes are arranged in a de-centralized peer-to-peer architecture and are capable of making intelligent decisions. Each sensing node is addressable by a unique 48 bit Neuron ID or a logical ID. The sensors use the power line for sending data as well as a source of power supply.

We have set up a small prototype testbed in our labs which demonstrate the use of power lines as a medium for communication of data amongst the sensors. Further, we have also set up a remote server which allows any user to log in to the server from the Internet or a PSTN network and view the data that has been logged by the sensors. We are looking at setting up a live test across a bridge on the I-5 near the La Jolla Parkway, UCSD. This experiment would help us understand the effectiveness of using power lines in lieu of wireless as a base medium for sensing technologies.

We are also investigating the use of RFID's over the power lines for monitoring and tracking purposes. In this scenario, the system would comprise of passive RFID tags. The RFID readers would be interfaced to the power line in a manner similar to the sensors and would transmit data either to the central database or use the data to trigger other sensors via the power lines. A sample scenario that we plan to develop is the use of RFID tags to detect people entering any civil infrastructure. Every person who is allowed to enter the building is given a RFID tag. Thus, when the RFID readers detect an invalid tag or something out of place, they immediately send messages to the video cameras, security and possibly the alarm through power lines. The video camera in turn takes a snapshot and sends it to the security personnel for further analysis.

We believe that as a communication medium, power lines provide us with a cost-effective solution for scenarios such as infrastructure monitoring/sensing since power lines are a part of the existing infrastructure and logically will be used to power sensors. Although still in its infancy, this field holds a lot of potential in particular with data communication speeds in the range of 1.5 Mbps. Efficient and cost effective scenarios for sensor deployment using PLC are now possible, replacing wireless approaches for infrastructure monitoring.

Products.

We plan to submit an initial paper on this to the USN (International workshop on RFID and Ubiquitous Sensor Networks) 2005 conference in June. We are also looking to incorporate some results and statistical data on the results of using power line communications for sensors in our paper.

Contributions.

A thorough study on Powerline Communications and the existing technologies prevalent in the market was done. Based on that, the possibility of using power lines as an alternative to wireless sensing technologies was looked into.

We have set up an initial testbed for prototyping the nodes and the communication messages amongst the sensors. We are currently planning to conduct experiments on communications amongst various types of sensors using the power lines to determine the limitations, if any.

A Paper is due for submission on June 15th to the USN 2005 Conference.

Wireless Infrastructure/GLQ

Activities and Findings.

The GLQ is a testbed and an example of a living laboratory. It is a hybrid wireless mesh network connected to the Internet over multiple long-haul point-to-point wireless links. It offers not only crucial data on the pattern of user traffic over a wireless mesh network but also a wideband Internet access infrastructure to public and law enforcement agencies. The infrastructure would enable monitoring of the busy, crowded, multi-block business area in downtown San Diego by law enforcement agencies for timely responses in emergencies. Many of the current research products that are being developed as part of RESCUE, such as vision-based and location-based situational awareness, and protocols for efficient and reliable mesh networking, can be studied on this testbed. The importance of this testbed is underlined by the fact that wireless mesh networks play a very crucial role in Ground Zero communications.

To test and validate the GLQ testbed plan before full deployment, UCSD has designed and built a wireless mesh network in the RESCUE Laboratory at UCSD. This laboratory deployment is helping the research team run experimental studies before the actual deployment in the Gas Lamp Quarter. In the first stage of this laboratory deployment, several experiments were conducted to confirm the testbed's functionality. Wireless client nodes were also configured to communicate through this gateway. In the second stage of this deployment, we will set up multi-hop relaying, routing and communication to the Internet through the gateway. Currently, we are in the process of building the laboratory testbed for the GLQ wireless mesh network. The actual GLQ testbed will be outfitted early in Year 3. At that time, the prototype of the quality-aware data collection framework will be implemented and tested within the GLQ testbed.

Products.

Current Progress. In order to test and validate the GLQ testbed plan before full deployment, UCSD has designed and built a wireless mesh network within the RESCUE Laboratory in EBU I at UCSD. This laboratory deployment is helping the research team run experimental studies before the actual deployment in the Gas Lamp Quarter. This laboratory deployment has involved four stages of implementation and evaluation; the first two stages have been completed so far. The four stages are:

1. Setting up the wireless mesh network gateway which acts as a bridge for both wired and wireless parts of the mesh network and conduct studies on bridging, routing, name resolution and Internet access across wireless and wired networks.
2. Setting up a wireless mesh gateway with multiple relay nodes and wireless clients, and conduct studies on relaying, detection, association, and disassociation. A linear string topology is considered for this stage.
3. Setting up planned mesh topologies with as many relay nodes and clients as possible and conduct extensive studies on the testbed. The studies include path selection, load balancing, hand-off, seamless roaming, video delivery, quality of service measurements, and will include video traffic. For video delivery performance and associated studies, we have decided to engage Ortiva Wireless, Inc. in order to evaluate their video conditioning product, Ostreamer, within the GLQ testbed. A multistage evaluation plan is under preparation. These studies will help in understanding key issues in efficiently delivering multimedia streams to first responders. Quantitative and qualitative studies involving networking and multimedia parameters will be carried out in this task.
4. Completing the development of the laboratory GLQ testbed using the long-haul Canopy wireless link, and finalizing the setup before deployment in the San Diego GLQ testbed. In this task, we plan on using multiple gateways connected to the Canopy client transceivers. The Canopy clients will then communicate with a Canopy server that will have a wired network connection.

In the first stage of this laboratory deployment, several experiments were conducted to confirm testbed's functionality and wireless client nodes were also configured to communicate through this gateway. In the second stage of this deployment, we will setup multi-hop relaying, routing, and communication to the Internet through the gateway. Currently, we are in the process of proceeding with the stages (3) and (4) to build the laboratory testbed for the GLQ wireless mesh network.

Contributions.

Gas Lamp Quarter (GLQ) Testbed (UCSD: R. Rao, B. Jafarian, BS Manoj, R. Dilmaghani)

The Gas Lamp Quarter (GLQ) testbed consists of a rapidly deployable mobile networking, computing, and geo-localization infrastructure in the context of incident-level response to spatially-localized disasters such as the World Trade Center attack. The testbed focuses on situations where the crisis site either does not have an existing infrastructure, or alternatively, the infrastructure is severely damaged. This testbed focuses on supporting basic services essential to the first responders that can be brought over to crisis sites for rapid deployment. Such services include communication among the first responders, accurate geo-localization both inside and outside of buildings, in urban as well as rural areas, computation infrastructure, incidence level command center, and technology to support information flow from/to crisis sites to/from regional emergency centers.

This testbed will be deployed in the Gas Lamp Quarter district in downtown San Diego. The testbed will provide a seamless Wi-Fi (802.11b) connectivity for first responders in this area. GLQ is currently divided into three zones, where each zone has its central post in direct line of site to the top of the NBC building. The transmitter on top of NBC building provides the broadband access to these three lampposts via a 5.2 – 5.7 GHz backhaul. By using Tropos Networks 5110 outdoor units, the coverage of these three zones will be expanded and we are able to provide the support for standard 802.11b users.

5110 units will find each other and mesh together using a standard ISM band which will also act as an 802.11b access point for end users. By connecting these three locations to the network, the testbed will cover a large area. Also by having three entry points, the reliability of the system will be enhanced. The three lampposts with backhaul connectivity (5th and Market, 5th and E, 4th and G) will have a Motorola Canopy receiver and a Tropos 5110 unit installed. The others will just have a Tropos 5110 unit installation. The 802.11b cells communicate with each other wirelessly through a mesh routing algorithm implemented within the access points. The control protocol is part of the Tropos Sphere operating and management tool. The following discussion highlights lessons learned during design and lab trials prior to network deployment.

Network Design. For this testbed, there is a need for a reliable and controlled wireless network that covers the entire GLQ area in downtown San Diego. Although there are several wireless ISPs in this area, most of have patchy coverage and are not reliable.

RESCUE is in a good position to design and deploy its own network to cover this area. This will provide an opportunity for researchers to have a more reliable network and to maintain control of the system. In deploying this system, the following issues are being addressed.

Frequency Band. The system can deploy in an unlicensed band and use standard off-the-shelf equipment. This will reduce the cost of deployment significantly. Our approach is to have a *hybrid network* design. While using the standard 802.11b for reaching the end-users, 5.2 GHz and 5.7 GHz will be used for backhauls to increase the capacity of the system. The other option is using Ensemble (a San Diego based company)/XO partnership. Recently, Ensemble and XO announced that they are able to provide equipment and spectrum to support any wireless deployment in the downtown San Diego area. This will certainly increase the reliability of the system since that system will operate in a licensed band.

Site Acquisition. Sentre Partners, one of the leading real estate companies in San Diego, has already committed itself to several projects in order to promote the city of San Diego. They own three of the tallest buildings in the downtown area. By using their rooftops, it will be easy to have a large footprint and cover the downtown area.

Backhauls. The cost of the backhaul is the most critical factor affecting long-term deployment. One possible solution for meeting this need is to involve some of the local telephone companies, e.g., Verizon or SBC, in providing the backhaul. Another solution

is to use the under-utilized bandwidth in the buildings to connect our base stations to the network. Sentre Partners, the owner of NBC building, is committed to provide this access and enough bandwidth for the project; therefore, this is the solution we have chosen.

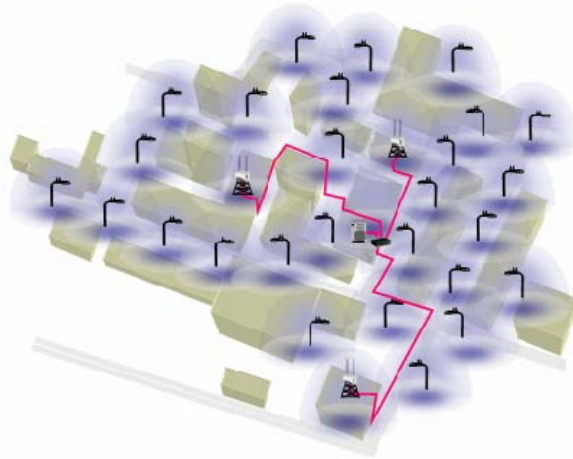


Figure 4. A Metro-Scale Cellular Wi-Fi Deployment

Throughput Analysis in Large Networks. Maximizing throughput in large networks requires minimal network bandwidth for protocol traffic control and optimal data paths for users in the face of highly variable RF conditions. Wireless link bandwidth is a finite resource and any traffic for control signals will reduce the capacity for user traffic. Traditional mesh nodes maintain routes between all nodes in the network, using either link state or distance vector protocols. As a result, the routing tables and information exchanged between nodes grows proportionally to the size of network. After the network reaches a certain size, the routing overhead will exceed the data traffic. In the present GLQ testbed architecture, the nodes and their routing mechanism, maintains constant routing overhead as network elements grow in number.

Wireless links are prone to multi-path fading and interference. These effects are dynamic, asymmetric and vary over time, particularly in a mobile environment. Ultimately, these effects manifest themselves as 802.11 packet reception errors, making them the major source of throughput loss on the wireless data link. In sharp contrast to wired networks where link-status is binary, throughput measured across one or more wireless hops can fluctuate anywhere between 0 and 100% of its theoretical maximum due to packet errors. Recent research has shown that routing algorithms that minimize hop-count or rely solely on RF signal strength to make routing decisions will fail to converge on a useful network topology, and offer poor throughput. Since these routing decisions are uncorrelated with throughput, they achieve far less throughput over time. This would include the vast

majority of wired routing protocols as well as the wireless routing algorithms employed by interconnected-based client mesh networks.

In contrast, the Predictive Wireless Routing Protocol (PWRP), used in GLQ deployment, is sensitive to these variations in throughput, i.e., by taking bi-directional measurement samples multiple times a second across wireless links. Based on a history of these measurements, predictive algorithms dynamically tune the selection of the multi-hop paths from the available paths in the mesh network. By estimating the throughput of each alternative path using advanced multi-hop metrics, PWRP ensures that it consistently selects paths in the top few percentiles of all available paths. On average, this achieves more than twice the throughput of competing routing approaches which are, in effect, choosing their paths at random with respect to throughput. In fact, PRWP consistently ensures a stable and high level throughput for Wi-Fi clients.

Network Layer Resiliency. In the initial development of this testbed for RESCUE, resiliency to unexpected failure was the main assumption in design and deployment. Cellular networks are prone to service interruption due to loss of network components (BSS and MSC). Wi-Fi networks, on the other hand, are vulnerable to severe interferences. PWRP incorporates network layer resiliency and self-healing features that enable network deployment with the desired reliability. PWRP is fully distributed and eliminates all single points of failure allowing for geographic distribution and rapid restoration. The GLQ testbed quickly detects backhaul failures and degradation and re-route the traffic through other nodes to other available backhaul link. The routing protocol typically relies on a small number of “hello” packets to detect the state of a link. It is important to discriminate between temporary wireless fades and an actual loss of link. During all these process the application and all active session must be maintained without interruption.

Networking in the Extreme

Activities and Findings.

Ground Zero communication is a kind of networking-in-the-extreme which offers a very challenging task in providing adequate communication experience in the event of disasters. The major aspect of networking in the extreme is the chaotic presence of multiple networks or absence of any communication networking infrastructure. As part of Responsphere project, we have developed solutions for developing reliable infrastructures for networking in the extreme. These solutions include (i) Robust Multi-Access Schemes and (ii) efficient wireless mesh networking solutions. Multi-Access schemes focus on using multiple wireless access networks for achieving the following: (i) high end-to-end throughput, (ii) glitch free communication experience, (iii) connectivity in the presence of chaotic heterogeneous wireless networks, (iv) aggregate bandwidth over multiple interfaces for multi-homed hosts, and (v) to utilize ad hoc wireless networks or wireless mesh networks.

Contributions.

We have developed several robust multi-access schemes for Networking in the Extreme. A multi-access scheme enables a wireless host to communicate with multiple access networks in such a way to choose either single best possible network or multiple networks simultaneously. We have developed the **Always Best Connected (ABC)** multi-access scheme that chooses the single best possible network in such a way that a seamless communication experience is provided. In the presence of multiple communication networks, utilizing multiple networks simultaneously provides a better option in obtaining a high end-to-end bandwidth. The critical issues in using multiple communication networks are (i) aggregation of bandwidth obtained through each of the network and (ii) packet scheduling through multiple paths. We developed the Session Layer Bandwidth Aggregation (SEBAG) scheme for efficiently using multiple end-to-end transport layer connections to achieve bandwidth aggregation.

Always Best Connected (UCSD/ R. Rao, B. S. Manoj, and R. Mishra)

The Always Best Connected (ABC) concept refers to an environment where several different types of access networks and different devices are available to a user for communication. The user can choose at any time the access network and device that best suits his or her needs depending on the applications that he or she is currently running, and change whenever something better becomes available. In recent years, the following functionalities of ABC have been built and tested in our labs: (i) Access Discovery and Selection, (ii) Mobility Management, (iii) Bandwidth Aggregation in over wireless mesh networks, (iv) Measurements, and (v) Profile-Based Access Control.

Access Discovery performs the task of determining what network accesses are currently available for the client and whether there is connectivity to a point in his home network. Access Selection allows the user to select different access networks depending on his predefined preference or based on the dynamic preference based on his profile stored in the network. In order to achieve its purpose, the access discovery and selection module continuously supervises all the interfaces on the client. It checks if the interface is up and running and has connectivity to the Home Agent. For connectivity check, the client sends probes (TCP/UDP requests) to the Home Agent and verifies the response. Access selection is done using the GUI where a user could choose the interface he wants all the applications on his device to use.

Mobility Management manages the session continuity for various services. The current infrastructure of ABC supports both Session Layer Mobility solutions as well as Mobile IP with NAT traversal support. Session layer mobility solution currently supports only TCP traffic and hence is applicable for applications like web browser, real media player, etc. The session layer mobility solution captures the initial connect request from an application in the client and redirect the request to a local proxy in the client. This local proxy connects to network based proxy (server), which in its turn issues the final (proxy) connect request to the actual destination. Hence, these redirection procedures results in three TCP connections out of the original single request: one from the client to the local client proxy, one from the local client proxy to the server proxy and one from the server proxy to the web server. When a user switches to another interface (another access

network) on the terminal, the only TCP connection that breaks is the one from the client proxy to the server proxy. However, the application and the application server (web server) still maintain the end-to-end state (IP address and port numbers) and remains in a connected state. Mobile-IP on the other hand is a network layer mobility solution where a virtual interface (the endpoint of the Mobile-IP tunnel) is created to hide the actual interface. Since the application always binds to Mobile-IP tunnel interface, any change in the actual interface does not affect the existing session. The current ABC framework supports Mobile-IP with NAT traversal. It works on a Linux platform and the tunnel interface has been implemented as a Linux driver running in kernel mode.

Bandwidth Aggregation handles the use of multiple interfaces simultaneously. Extending the concept of ABC we can now utilize all the accesses that a user has, primarily for load balancing (thereby increasing the network utilization) as well as providing redundancy. This could also be used when an ABC node is acting as a router in the network, as it would have multiple access networks it could do IP load balancing across multiple interfaces. The prototype for the ABC access aggregation in its current form runs on Linux. The setup contains multiple wireless mesh nodes and the assumption for this version of aggregation is that all the wireless mesh network nodes are gateways to external access networks (can route traffic to Internet) as well. We use AODV (Ad-Hoc On demand Distance Vector protocol) to discover all the gateway nodes within the wireless mesh network. The ABC client node is part of the AODV subnet as well and performs session level access aggregation either for the traffic originating from this machine or for a traffic that is forwarded through this machine. The current load balancing mechanism is round robin, *i.e.*, the sessions are equally distributed across multiple outgoing gateways in the wireless mesh network. Performance measurements are carried out for evaluating the effectiveness of ABC system and for obtaining the throughput, delay and other statistics for the ABC client.

Profile Server helps the ABC client to store and manage its various preferences. Some of the preferences can include the network preferences, application characteristics, and device characteristics. Earlier, the ABC service relied on user control for the selection of network access type or, alternatively, a fixed selection policy residing in the client. However, in order for the ABC service to reflect dynamics in, *e.g.*, the access network performance, traffic pricing, and client status, the ABC client policies is made adaptive. Remotely controlled ABC service policies are now a part of the network management system that also hosts the ABC service. The profile server implements some policies for dynamic control of network access. There are three main software subsystems that implement the profile server. These are (i) the client-side profile module, (ii) the profile server, and (iii) the network profile admin module. The client-side profile module resides on the ABC client and is responsible for updating the client related profile information (*e.g.* network statistics, application usage, GPS information) to the server. The profile server stores the client profile information in a database (My-SQL), and provides a communication interface to the administrators. The Network Admin application provides a user interface to view the currently active profile clients, as well as the ability to control these clients. The communications between these subsystems are done using a CORBA (Common Object Request Broker Architecture) like protocol.

One of the challenges in transitioning between these access technologies is the dramatic difference in their capacities from an application's perspective. For example, in moving from a WLAN (IEEE 802.11b) to cellular access (such as CDMA2000, 1xRTT, or GPRS), the capacity decreases, causing the application to have to adapt – and quickly – so as not to adversely affect the user. This becomes a particularly serious problem if the application is multimedia-based, with a heavy reliance on voice and/or video.

With the addition of all these features, ABC has become a versatile testbed for current and future projects. Seamless interoperability among various network infrastructures to maintain constant – and the most powerful – connection to the Internet for mobile users using both licensed (e.g., cellular networks) and unlicensed (e.g., IEEE 802.11b) spectrum. This work is making it possible for user devices to seamlessly roam among varying network access technologies, including IEEE 802.11b, CDMA2000, GPRS, and Ethernet, with controlled adaptation of applications' performance to the network capabilities at hand.

Session Layer Bandwidth Aggregation (SEBAG) Scheme (R. Rao, B. S. Manoj, and R. Mishra)

The Session layer Bandwidth Aggregation (SEBAG) scheme is an always-all-connected communication access paradigm which provides end-to-end multiple paths over multiple communication interfaces. Here, the bandwidth aggregation module is a thin layer that operates logically between the application layer and the transport layer and hence we considered it a session layer solution. At the client node, the session layer module is called Client-side SEBAG Aggregator Module (CSAM) and its equivalent at the server's end is Server-side SEBAG Aggregator Module (SSAM). De-linking the bandwidth aggregation process from application layer is important as it simplifies and optimizes the operation of bandwidth aggregation mechanisms. The primary responsibility of SEBAG aggregator modules (SAMs) is to manage packet striping and aggregation based on the availability of network resources over multiple end-to-end paths. For example, when a mobile node moves into the coverage area of a new network, the CSAM identifies it and communicates over to the SSAM and initiates another transport layer connection to utilize the new access network. The SSAM would now start including the new transport connection in the packet scheduling process at the server side. Similarly when mobile hosts move out of the coverage of a particular network, it can remove the transport level connection that was setup through that interface.

The cross layer design approach used at the client side is facilitated by a Cross Layer Interaction Module (CLIM) which interacts with session and MAC layers. The major advantage of this approach is that we can dynamically update the properties of the end-to-end transport session depending upon the changes in the network access system. The number of TCP connections serving a single transport session is limited by the number of access networks present. Choice of access networks includes IEEE 802.3, Bluetooth, IEEE 802.11b/g/a, 3rd generation cellular networks, CDMA data networks, and satellite links. The number of multipath transport connections can be decided based on policies

such as cost of access, bandwidth of access networks, and power consumed by the interfaces.

In order to achieve high throughput for an end-to-end data transfer session, SEBAG utilizes an efficient traffic aggregation mechanism. For example, consider the end-to-end transport layer connection where each connection (for example TCP) is identified by IP addresses and TCP ports of both the sender and the receiver. When an application layer at the client side needs to open a connection, it passes the request to the Client-side SAM. Instead of setting up a single TCP connection, the CSAM sets up multiple TCP connections with the server end-point. All these connections are maintained as part of a session layer connection. Now each of these connections can in fact operate as an independent TCP connection. Each TCP connection established between CSAM and SSAM transfers data and in-band signaling packets for SEBAG. For example, the first packet sent over a new transport layer connection contains the connection identifier. Application layer protocols need not be aware that it uses multiple transport layer connections. Therefore, SEBAG is transparent to application layer protocols and transport layer protocols. This is achieved by handling all the interface primitives between application and transport layers.

The CSAM communicates with a Cross Layer Interaction Module (CLIM) which in turn interacts with the Network layer as well as the MAC layer. CLIM understands the number of communication interfaces and the capability, status of connectivity, and the raw data rate of each of them. By periodically updating the network access capability of the node, the CLIM aids the CSAM to change the number of transport connections dynamically. The major actions carried out by the link manager are monitoring and conveying the status of each of the available communication interfaces and executing the commands from the access manager. Every interface at the MH has one link manager. Upon the instructions from the access manager, a link manager could initiate several actions such as turning up or down, enables IP connectivity to the responding node (either server or client), and monitoring the availability of link.

SEBAG uses a new scheme called Expected Earliest Delivery Path First (EEDPF) scheduling. In this case, we estimate the end-to-end bandwidth and assign a packet to the connection which is expected to deliver the packet at the earliest. When a packet is to be scheduled, the SEBAG module finds the connection which could deliver it to the destination at the earliest. The SAM modules obtain the end-to-end throughput as well as the average end-to-end latency of each connection to determine this. The expected delivery time is estimated as the sum of the end-to-end latency and the packet transmission time. In addition to the data rate at the interface, this scheme uses the end-to-end data rate as the primary input to the scheduler. In our experiments, we noticed that the high data rate interface received a proportionally higher share of packets.

Wireless Mesh Networking Research (R. Rao, B. S. Manoj, and R. Mishra)

In order to provide communication services to disaster sites, especially in the event of complete destruction of existing communication infrastructure, it is essential to depend on ad hoc deployment of wireless networks. Wireless mesh networks provide a quickly

deployable network infrastructure with high flexibility that can adapt to fast changing communication requirements at Ground Zero. Such a network can be deployed by placing wireless mesh relay nodes that operate using IEEE 802.11 technology. In order to have a reliable and efficient wireless mesh network for Ground Zero communications, we need to develop new solutions that provide a suitable infrastructure for response activities. Existing solutions for mesh networking include the formation of a Wireless Distribution System (WDS) with Spanning Tree Protocol (STP) for a loop free communication paths between communication end-points. Such a system; though has an advantage of simplicity also faces several issues when deployed for emergency situations. The major issues are (i) the requirement for multiple redundant paths between communication end-points, (ii) formation of a single point of failure in the network by the root node of the spanning tree, (iii) inefficient use of multiple gateway nodes by STP, (iv) low capacity utilization due to bandwidth bottlenecks, and (v) in efficiency due to the choice of lowest-address-node as the root of the SPT. We developed a robust wireless mesh networking solution which uses dual SPTs over multiple interfaces to mitigate the issues faced by existing single SPT based wireless mesh networks. In our solution, every wireless mesh network relay node uses two IEEE 802.11b interfaces to provide better performance. We formed two separate SPTs, one over each interface. The gateway nodes are chosen to be the roots of each of the SPTs. This system has many significant advantages, some of which are the following: (i) simplicity and transparency provided by the WDS and SPT based mesh networking, (ii) high throughput capacity, (iii) removal of single point of failure in the network, (iv) provisioning of multiple redundant paths, and (v) aggregation of bandwidth obtained through multiple gateway nodes.

Products.

Always Best Connected Software Ramesh Rao, Rajesh Mishra, B.S. Manoj (UCSD)

Seamless interoperability among various network infrastructures to maintain constant – and the most powerful – connection to the Internet for mobile users using both licensed (e.g., cellular networks) and unlicensed (e.g., IEEE 802.11b) spectrum. This work is making it possible for user devices to seamlessly roam among varying network access technologies, including IEEE 802.11b, CDMA2000, GPRS, and Ethernet, with controlled adaptation of applications' performance to the network capabilities at hand.

One of the challenges in transitioning between these access technologies is the dramatic difference in their capacities from an application's perspective. For example, in moving from a WLAN (IEEE 802.11b) to cellular access (such as CDMA2000 1xRTT or GPRS), the capacity decreases, causing the application to have to adapt – and quickly – so as not to adversely affect the user. This becomes a particularly serious problem if the application is multimedia-based, with a heavy reliance on voice and/or video.

1. P.Nuggehalli, V.Srinivasan, R. R. Rao, “Energy Efficient Transmission Scheduling for Delay Constrained Wireless Networks,” accepted for publication in *IEEE Transactions on Wireless Communications*

2. K. Wang, C. F. Chiasserini, J. G. Proakis, R. R. Rao, "Joint Scheduling and Power Control Supporting Multicasting in Wireless Ad Hoc Networks", submitted to Ad Hoc Networks Journal.
3. K. Chebrolu, R. Rao "Bandwidth Aggregation for Real-Time Applications in Heterogeneous Wireless Networks", IEEE Transactions on Wireless Communications, Accepted.
4. K. Chebrolu, B. Raman, R. Rao, "A Network Layer Approach to Enable TCP over Multiple Interfaces", to appear in ACM/Kluwer Journal of Wireless networks (WINET).
5. P. Zhou, X. Wang, R. Rao, "Throughput Capacity of Wireless Mesh Networks," submitted to IEEE Wireless Communications, April, 2005.
6. V. Srinivasan, P. Nuggehalli, C.-F. Chiasserini, and R. R. Rao, "An Analytical Approach to the Study of Cooperation in Wireless Ad Hoc Networks," in *IEEE Transactions on Wireless Communications*, Vol. 4, No. 2, Mar. 2005, pp.--.
7. M. Liu, R. Rao, L. Milstein, "On the Existence of Fair Wireless Scheduling Protocols," submitted to *IEEE Transactions on Wireless Communications*, 2004.
8. K. Wang, C.F. Chiasserini, J. Proakis, R. Rao, "A Distributed Fair Scheduling Framework for Wireless Ad Hoc Networks," submitted to IEEE Transactions on Wireless Communications, Proceedings of Globecom 2004, Dallas, TX, November 2004.
9. K. Wang, C. F. Chiasserini, J. G. Proakis, R. R. Rao, "Distributed Fair Scheduling and Power Control in Wireless Ad Hoc Networks," Proc. of Globecom 2004, Dallas, TX, November 2004.
10. J. Fang, R. Rao, "Optimal flow control for end-to-end delay and power constrained wireless multihop networks," Proceedings of IEEE Military Communication Conference (Milcom), Monterey, CA., October, 2004.
11. BS Manoj (UCSD) is organizing a workshop titled "Next Generation Wireless Networks: Design for Reliability, Availability, and Emergency Response," at the IEEE HiPC 2005 conference (www.hipc.org). This year's focus theme is on the areas relevant to the ITR project RESCUE. Ramesh Rao will give a keynote talk in this workshop. Information on the workshop can be found at <http://wongen.calit2.net>

CVC: Collaborative Visualization Center

Activities and Findings.

UCSD Campus, EBU I rm 6504

Leveraging prior investments, the Collaborative Visualization Center (CVC) was created to serve as a command and control prototype facility. The CVC is a research and presentation facility for artistic, scientific and engineering visualization. The room is a multidisciplinary haven designed to facilitate research, presentations, meetings, and collaboration in the same location. The CVC is envisioned as a dynamic, collaborative research environment. In addition to its functionality as a meeting room and videoconferencing facility, it provides advanced visualization equipment capable of expanding the scope of ongoing research and creating opportunities for new research activities based on the collaborative, interdisciplinary nature of such a space. Far from static, the CVC will continue to evolve as needs and technology demand. The CVC was formally launched on September 24, 2004.

Training courses on the CVC have been taught by Helena Bristow to the UCSD community to expand the scope of use of the facility. It has been tested for drills and other purposes as a command and control facility by ResponSphere and synergistic projects including RESCUE, WIISARD, OptiPuter, and BioNet.

What's in the CVC?

The room is currently outfitted with the following equipment:

- *Barco Galaxy Projector*
The room is dominated by a large Stewart screen and Barco Galaxy projector. The Stewart screen acts as one of the four walls of the CVC and measures approximately 8 feet high by 20 feet wide. The Barco projects a high-resolution (1280x1024) image to a size of 8 feet high by 10 feet wide. This facility is unique because the Barco projects from *behind* the screen, so persons in the CVC can walk right up to the screen without projecting a shadow onto the image. This is driven by a powerful Windows PC with two state-of-the-art graphics boards, but can also be connected to any laptop via a standard VGA cable.
- *Quadrant Projectors*
Quadrant projectors are able to display separate images to the four corners of the dominating screen, allowing for visualization of multiple applications simultaneously, and further allowing the facility to function as a test command center.
- *PC & Graphics Drivers*
The projector is driven by a powerful Windows PC with two state-of-the-art graphics boards, an ATI FireGL X1-256 and a TeraRecon Volume Pro 1000. The ATI card accelerates the display of 3D lines and polygons. It also can run the new OpenGL shader capability, which UCSD is Beta testing for ATI. OpenGL shaders can be used very effectively in visualization research and graphic rendering applications. The TeraRecon card accelerates the display of volume data. Acceleration programs target features of the TeraRecon board to assist specific visualization needs, maximizing use of the Barco projector's active stereo feature in which OpenGL quad framebuffers are

used to project both left and right eye views. The result is a very dramatic and information-rich 3D interactive display.

- *Sound System*

The facility is outfitted with a 5.1 channel surround sound system, which is routed through a stereo and can be accessed by the DVD player, the PC, or even a laptop.

- *Videoconferencing*

The facility is equipped with a 50 inch plasma display which is connected to a Sony VTC1 video-teleconferencing system with both a wireless Shure microphone and an omnidirectional VTC microphone.

- *Smart Board*

There is a second plasma display equipped with a SMART Board Interactive Overlay for Flat Panel Displays. This device transforms the plasma screen into an interactive touch-screen display, allowing the user to control computer applications, navigate websites, and write notes in “electronic ink”. Users can then save these notes into one of many different formats, print them on the bluetooth printer, and/or send them to meeting participants or to those who missed the meeting.

- *Bluetooth Printer*

The Bluetooth Printer can print documents from any bluetooth-enabled device in the room, including the Smart Board and the main PC, or any bluetooth-enabled laptop.

Products.

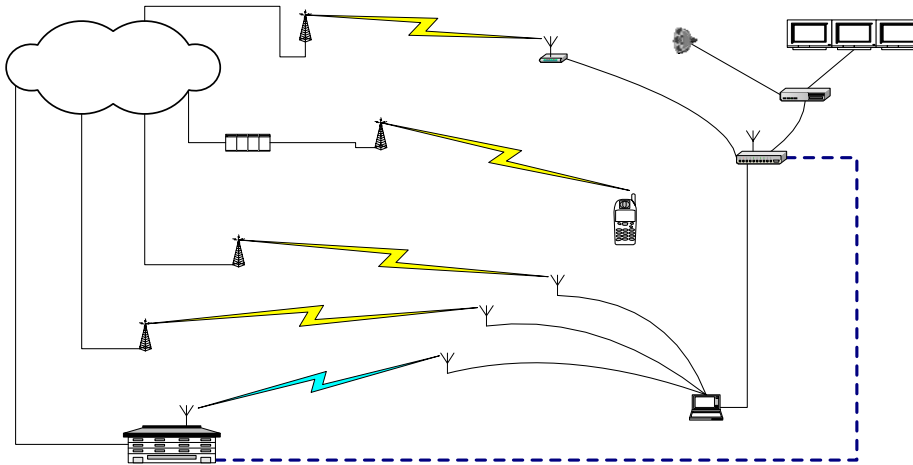
None.

Contributions.

None.

CyberShuttle / CyberAmbulance

Activities and Findings.



The UCSD CyberShuttle and CyberAmbulance projects consist of initial experiments in the deployment of mobile, hybrid wireless platforms. The intent of these experiments was to determine the feasibility and practicality of using high speed Internet services as the uplink technology between computing and other devices residing in a mobile environment, and the commodity Internet. To date, these experiments have occurred in three phases: CyberShuttle 1.0, CyberAmbulance, and CyberShuttle 2.0.

CyberShuttle 1.0

The first phase of the mobile, hybrid wireless platform experiments involved a very simple design consisting of a Qualcomm prototype HDR (1xEV-DO) modem and a local wireless access point, both installed on several UCSD campus shuttle busses. The HDR modem was initially connected to a private IS-95 (TIA-EIA-95) network operated by Qualcomm, and with base stations at the Qualcomm corporate headquarters in San Diego and at the main Engineering building at the University of California, San Diego. The HDR modem was based on the release 0 standard, and provided a maximum forward (downlink) bandwidth of 2.4Mbps and a maximum reverse (uplink) bandwidth of 153Kbps. A network path from the HDR modem to the commodity Internet was provided via a wired backend infrastructure co-managed by Qualcomm and UCSD. The wireless (802.11b) access point in the shuttle provided the typical connection services available on such a device (DHCP, NAT, etc.) and used the HDR modem as the uplink (WAN) via a wired Ethernet connection (see Figure A below).

Various wired and wireless devices could be deployed and tested on the mobile platform without special configuration, owing to the transparency of the uplink technology. This enabled a variety of experiments to be moved successfully from the laboratory environment to the shuttle bus, generally without modification of connection topologies. Along with the deployment and testing of commodity applications (Web, email, downloading, instant messaging, etc.), a variety of specialized experiments were deployed in this environment, including those requiring bidirectional, interactive behavior. In general, while not a significant research experiment *per se*, CyberShuttle 1.0 provided a flexible platform for

testing both common and novel applications in a mobile environment. One of these applications led to the second phase of the experiment, the CyberAmbulance.

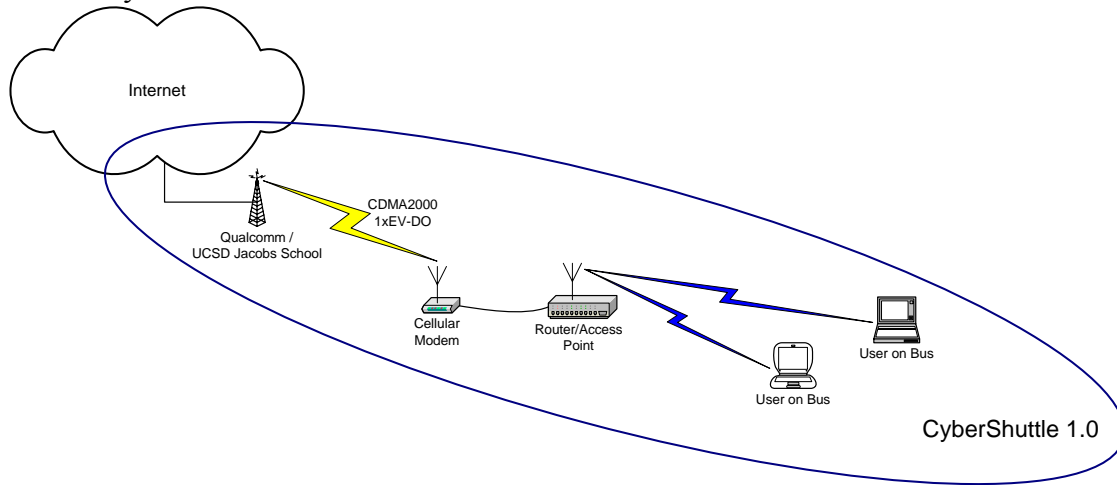


Figure A

CyberAmbulance

The CyberAmbulance project built upon the basic shuttle infrastructure introduced in CyberShuttle 1.0 to create a platform for testing how the introduction of a high-speed uplink, plus additional technologies, could be used to enhance the effectiveness of emergency medical professionals in transit from field locations to regional medical centers. A UCSD campus shuttle was configured to function as a mock ambulance setting (see Figure B below) via the addition of an enhanced physical (e.g., patient stretcher, monitoring equipment, etc.) and telecommunications infrastructure. Complementing the existing infrastructure from CyberShuttle 1.0, the CyberAmbulance configuration included a VPN client and server to validate the need to secure patient data, as well as a laptop configured with a camera and headset, and connected into the wireless infrastructure on the shuttle. The laptop and associated devices were used to transmit and receive live audio, video, and telemetry between the shuttle and emergency department personnel at the UCSD Medical Center. Professional actors were recruited to simulate a variety of patient conditions in real time, while the shuttle was in motion.

The key point of the experiment was to investigate what benefits could be gained by the introduction of a high-speed data connection, versus the lower-speed radio-based connections commonly in use. In general, the availability of the improved connection was viewed positively by both mobile and remote emergency medical professionals, particularly in the context of being able to transmit larger data sets, such as still and motion video images of patient condition and behavior. However, the IP-based audio connection was not a substantial improvement over

existing telecommunications technologies, and in fact was seen as less reliable and intelligible. Frequently, the audio portion of the connection was abandoned in favor of more traditional means of voice communication. Also, while it was determined that the introduction of a VPN circuit between the shuttle and the medical center did not result in an unacceptable degradation of network performance, the reliability of the VPN connection was often not robust in the face of general cellular connectivity losses, as the shuttle traveled in and out of range of the Qualcomm and UCSD base stations. While IP-based technologies would generally have good failure recovery during such network partitions, those overlaid on top of a VPN connection did not fare as well because of the inability of the VPN hardware to reestablish a circuit upon loss of the underlying IP connection. Nevertheless, because the CyberAmbulance experiment provided useful data to emergency medical services researchers, it validated the general CyberShuttle concept as a useful platform for experimental deployment of mobile technologies. For this reason, CyberShuttle 2.0 was designed and deployed as a logical next step.

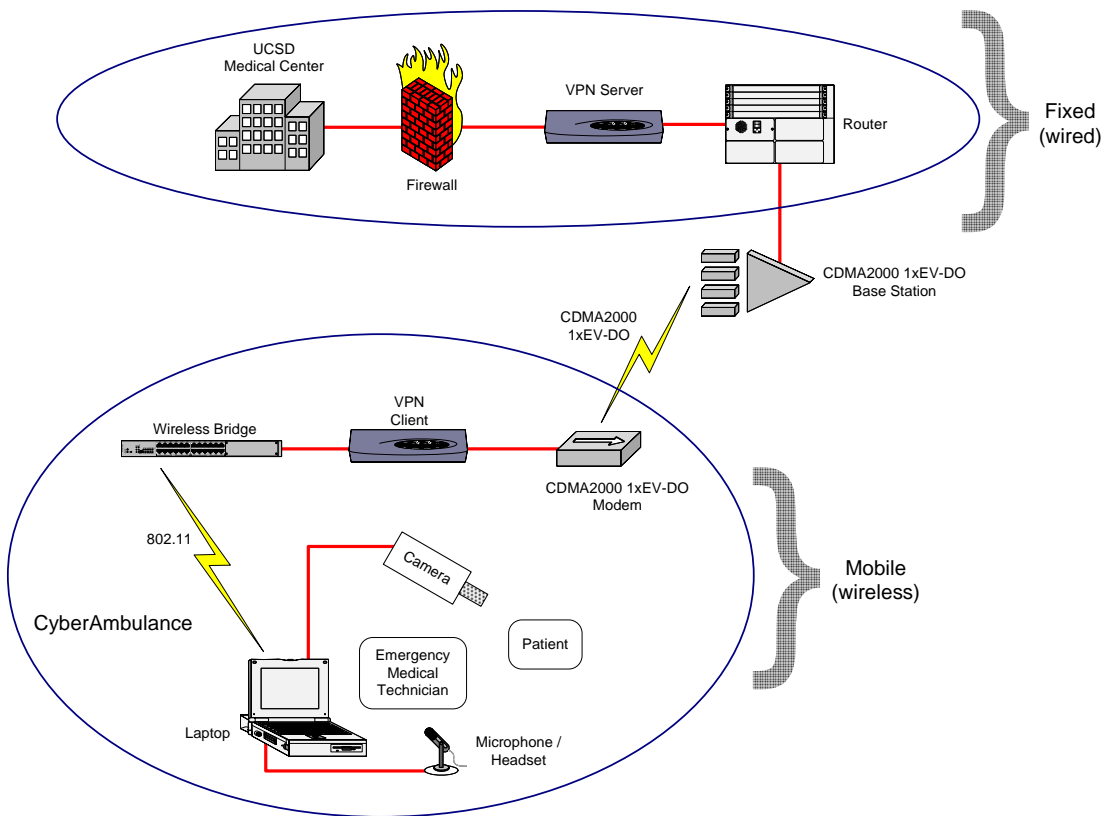


Figure B

CyberShuttle 2.0

While version 1.0 of the CyberShuttle provided a basic platform for deploying mobile computing experiments, such experiments were generally not

“permanent”, in the sense that they actively existed only during the course of the experiment itself. The next phase in the development of the CyberShuttle introduced additional technologies to enable the creation of a more robust mobile, experimental platform (see Figure C below). Augmenting the infrastructure from CyberShuttle 1.0 was the inclusion of on-board computing and A/V hardware. Using these systems enabled the deployment of more complex production and experimental services, including ones which were “always on”. In addition, the availability of a permanent computing platform permitted the introduction of shuttle-resident services that could not only be used by riders, but could also collect data from the shuttle environment and transmit them to a remote site independent of the on-line interaction of researchers.

The enhanced experimental platform in CyberShuttle 2.0 has proven to be very useful as a deployment mechanism for a wide variety of experiments substantially more complex than those supported by version 1.0. In some of these, researchers and end users interact with services deployed on the shuttle-resident computing platform via voice and touch-tone commands issued from cellular telephones (see Figure C). CyberShuttle 2.0 has also been successfully used as a command post in projects such as RESCUE (www.itr-rescue.org), WIISARD (www.wiisard.org), and Responsphere (www.responsphere.org). The shuttle’s use in a diverse collection of research activities speaks to its value as a flexible platform for the deployment of production and experimental mobile services.

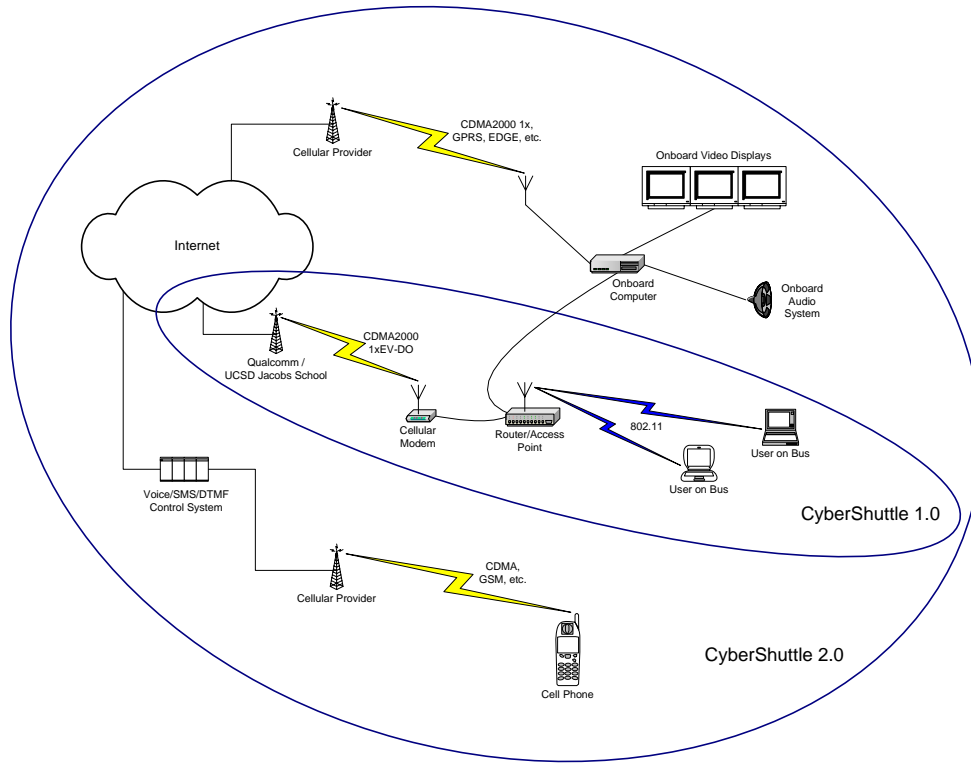


Figure C

Products.

The CyberShuttle and CyberAmbulance projects are not conceived as research activities in and of themselves, except insofar as they have provided and continue provide a platform for subsequent research. As such, while there have been informal publications of the specifications of these systems, work products have been and are more appropriately discussed in the context of the research activities supported by the platforms. Similarly, while the projects are not necessarily significant contributions to the body of scientific knowledge, they have substantially aided in the design and implementation of other projects which have made such contributions (as described briefly here, and more substantially elsewhere in this document). Some of the publications resulting from these activities are:

Chan TC, Killeen J, Griswold W, Lenert L. "Information Technology and Emergency Medical Care during Disasters." *ACAD Emergency Medicine*, Nov 2004. Vol 11 No 11.

Contributions.

The CyberShuttle and CyberAmbulance projects are not conceived as research activities in and of themselves, except insofar as they have provided and continue provide a platform for subsequent research. As such, while there have been informal publications of the specifications of these systems, work products have been and are more appropriately discussed in the context of the research activities supported by the platforms. Similarly, while the projects are not necessarily significant contributions to the body of scientific knowledge, they have substantially aided in the design and implementation of other projects which have made such contributions (as described briefly here, and more substantially elsewhere in this document).

A Transportable Outdoor, Tiled Display Wall for Emergency Response

Activities and Findings.

The Calit2 Visualization Group plans the construction and integration of visualization tools (hardware and software) for Responsphere during Year 1 as part of a partnership between two large ITRs at UCSD: RESCUE and OptIPuter. We propose the construction of a portable, rugged, brilliant display wall and cluster system for in field emergency response deployment.

The system would include a 5x3 grid of approximately 21.3 inch ruggedized transfective displays providing a >25 megapixel viewing surface, a multi-node visualization cluster to drive it, and ruggedized transport cases for both. Applications of tiled display walls include the visualization of large remote sensing, volume rendering imagery, mapping, seismic interpretation, museum exhibits and other applications that require a large collaborative screen area. This system would provide advanced emergency response visualization capabilities when coupled with a collection of existing and emerging software tools under development in the OptIPuter project.

These tools include JuxtaView, LambdaRAM, Vol-a-Tile and SAGE. JuxtaView is a visualization tool that displays ultra-high resolution 2D images, such as USGS aerial and satellite maps. JuxtaView invokes LambdaRAM, a middleware layer offering a remote memory interface. LambdaRAM takes advantage of multiple-gigabit network connection by pre-fetching portions of very large datasets before an application is likely to need them (similar to how RAM caches work in computers today) and storing them in the local cluster's memory for display. Aggressive pre-fetching and large bandwidth utilization can overcome the wide-area network latency that hinders interactive applications such as JuxtaView.

Vol-a-Tile is a volume rendering tool for large-scale, time-series scientific datasets for scalable resolution displays. These large scale datasets can be dynamically processed and retrieved from remote data stores over photonic networks. Vol-a-Tile focuses on very large volumetric data and time-dependent simulation results such as seismic data and ultra-high resolution microscopy.

SAGE (Scalable Adaptive Graphics Environment) will support the unification of a collection of display tiles into a virtual super display and allow for the integration and control of multiple graphic and video formats. SAGE also allows for the laptop/tablet level control of large display walls to include pen based annotations

This system participates in the following Responsphere Project goals:

- Constructing a command and control prototyping environment
- Building a vehicular based mobile command and control platform

Resources required include display and cluster hardware.

Products.

1. Larry Smarr, Joe Ford, Phil Papadopoulos, Shaya Fainman, Thomas DeFanti, Maxine Brown, Jason Leigh, "[The OptIPuter, Quartzite, and Starlight Projects: A Campus to Global-Scale Testbed for Optical Technologies Enabling LambdaGrid Computing](#)," (invited paper), Optical Fiber Communication Conference & Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC) 2005, Anaheim, California, March 6-11, 2005, CD ROM.

Contributions.

The portability of such a display screen, plus its ability to be used outdoors, will be beneficial to the RESCUE project, as well as other projects and applications. For example, the display will be a part of the mobile command and control vehicle, and can be set up at the mobile command post at the various drills that are part of both the RESCUE and the WIISARD projects; the software tools under development by the OptIPuter project would enable advanced emergency response capabilities.

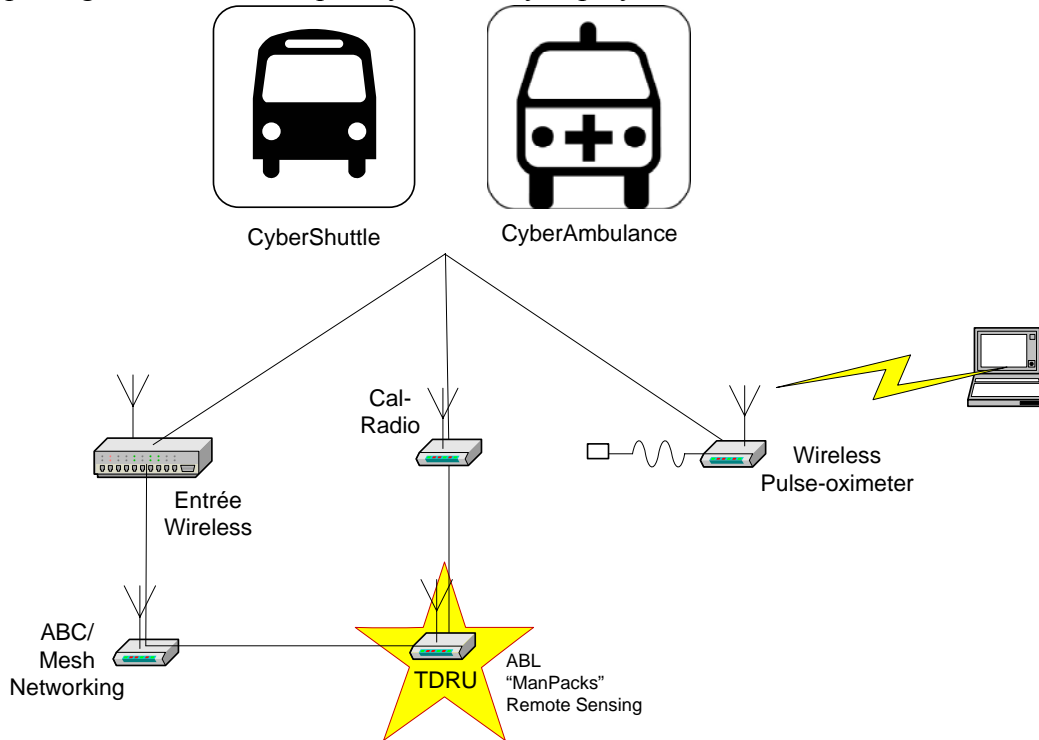
In addition, the display can be used in other visualization demonstrations. Specifically, the upcoming iGRID conference at UC San Diego in September 2005 will have several GIS and collaborative visualization demonstrations that would benefit from this display.

“Manpacks” and the Always Best Located Platform

Activities and Findings.

The CyberShuttle (RESCUE) and CyberAmbulance (WIISARD) were developed as prototypes for different kinds of mobile command/control and response vehicles. Alone or together, neither was sufficient to meet the needs of a disaster scenario requiring medical or non-medical response. As a result, a series of artifacts began to emerge with varying capabilities. Out of the original CyberShuttle came the need to expand its WiFi “bubble” beyond the bus itself. A local company, Entrée Wireless, productized technology which was developed by Calit2 researchers, allowing a 1XEVD0 data signal to be processed and shared over 802.11b. This solution was further improved upon by Calit2 researchers in the context of Mesh Networking research for RESCUE and applications for WIISARD – getting these “breadcrumbs” to dynamically reconfigure the network and intelligently store and re-send data in case of loss of connectivity. Other platforms emerged as solutions to different kinds of problems: the Wireless Pulse-oximeter transmits patient pulseox information in real-time over a wireless network and can be deployed in disaster-site triage scenarios, and Cal-Radio is an emerging platform

for fully configurable network research, programmable down to the MAC layer thus allowing greater flexibility in applications development and implementation. Yet new needs emerge with each iteration: we plan to develop a Cal-Radio/Mesh hybrid with GPS and other location-based technology, on which the Always Best Located platform and various other sensors can be deployed, allowing for more complete situational awareness in packages which can be quickly and easily deployed at the site of a disaster.



Products.

Peer to Peer 511 System

The goal develop a fully automated system that will build on the concept of "humans as sensors" to collect and relay disaster related information to the general public and to the first responders. Though government agencies and the private sector have some of the basic data needed for effective disaster prevention and management, the means to effectively disseminate the data in an intelligent manner (i.e, delivery of relevant and timely information to the right target population) is lacking. Typically the data is disseminated in a broadcast mode, which could create a mass panic type of situation. Also, in many situations, there is significant lag in the collection of crisis related data by the government agencies. This lag can be eliminated by empowering the general public to report relevant information.

We will use San Diego as a test bed to develop, deploy and test the above mentioned system that will empower the general public (in particular the commuters) of the county to act as human sensors and relay information about incidents ranging from wild fires, mudslides and other major accidents to the general public and to the 911 control center. The system can be accessed simply by making a phone call and will be based on speech

recognition. We have learnt from past experience, that the general public will not adopt such a system if you inject a new phone number during the time of a disaster (such as the San Diego wild fires). The system should be available on a regular basis, disseminating information that is valuable to public on an every day basis.

We will address these problems by using a traffic notification system that has been operational for the past two years and used by thousands of San Diego commuters every day as the basis for prototype. The system currently provides personalized real time traffic information to the commuters via cell phones (<http://traffic.calit2.net>). We will modify this system so that commuters can report incidents 24x7, including the time, location, severity and the urgency of the event. We will analyze the data for validity and populate the events in a GIS database. Other commuters calling in, will hear these events if they happen to fall in their commute segment. Also based on the severity of the incident, we can notify all or part of the users via voice calls and text messages in a parallel and scaleable manner. We will create a hierarchical voice user interface that will accommodate for the severity of the incidents being reported. Examples of scenarios are the following. In the simplest case, a commuter might see a major accident that has closed several lanes of a highway. He can report this incident via the system and other users who are calling in for traffic information will hear this event if it happens to fall in their commute segment. An example of a more severe case would be the San Diego wild fires spreading to I-15 resulting in a shutdown of the freeway. If one reports such an event, due to the severity of the event, the system will trigger an alert all the users, to avoid that region of the freeway.

Route Preview Software with Audio Cues for First Responders

The goal of this research was to investigate how to present audio cues to a first-responder using a GIS database and a location-based application so that the first-responder may follow a pre-defined path. We also investigated how audio cues can allow a first-responder to become oriented in an unfamiliar environment. Assuming a first-responder relies on his own senses for collision avoidance and personal safety, we created a location-based application that will enable a first-responder to either maintain orientation or follow a pre-defined route.

Our software application includes several features; it allows marking favorite waypoints and paths, and supports presenting waypoints via an audio description in a meaningful way. It marks waypoints as visited or unvisited, displays a list of neighboring waypoints to the user's current waypoint, identifies the last waypoint visited using the software's "back" feature, and allows a user to traverse a series of waypoints from the current position back to the starting position. The software also presents neighboring waypoints at an intersection in clockwise order, starting with the northern-most waypoint and listing the remaining neighboring waypoints in the order of East, South and West. The user interface (UI) models the UI of a cell phone application, enabling a user to select information from a menu by entering numerical digits 0 through 9, and get information as with a standard phone menu.

The software can guide a rescue worker through any defined path. From the beginning of this path, the software will mark waypoints at an intersection that will lead him to the end of this path. This feature will allow a rescue team to examine a hazardous situation and

guide a first responder along a pre-defined path. The software allows the user to create a new path so that a first responder can mark a route to a desired waypoint.

Our research found that presenting neighboring waypoints in clockwise order efficiently oriented the user with audio cues.

Software that assists a first-responder in both maintaining orientation and following a pre-defined route currently includes several features. The software allows marking favorite waypoints and paths. It also supports presenting waypoints via an audio description in a meaningful way. For example, it marks waypoints as visited or unvisited. It displays a list of neighboring waypoints to the user's current waypoint and identifies the last waypoint visited using the software's "back" feature, it allows the user to traverse a series of waypoints from the user's current position back to the starting position. The software also presents neighboring waypoints at an intersection in clockwise order. It first lists the northern-most waypoint and proceeds listing the remaining neighboring waypoints in the order of East, South and West. The software's user interface (UI) will model the UI of a cell phone application. The user selects information from a menu by entering numerical digits 0 through 9 and the UI conveys information to the user as with a standard phone menu.

The software will guide a rescue worker through any defined path. Starting out at the beginning of this path, the software will mark waypoints at an intersection that will lead him to the end of this path. With this feature, a rescue team will be able to examine a hazardous situation and guide a first-responder along a pre-defined path. In addition, the software allows the user to create a new path so that a first-responder can mark a route to a desired way point.

Zig-Zag: Sense-of-Touch Guidance Device

The Zig-Zag sense of touch hand-held prototype is an RF transmitter/receiver analog device. It is composed of a 3-channel FM 2-stick radio transmitter which sends the signal to an analog module box. The receiver standing inside the analog module box feeds the signal to the RC Switch (35 RAM). From here, the input received gets modulated in order to output the necessary square wave signal. The signal leaves the analog module through a DB9 connection to reach the hand-held box consisting of a HS 322 servo and a vibrating motor. This last box will be the output of our system. The servo will point to the direction indicated by the transmitter and the vibrating motor will go off signifying that the final target location has been reached or the necessity of an immediate stop.

Existing technology, such as GPS, shows the position of an object in a map. It also shows the direction in which this object is pointing with respect to the North. Map quest, or car GPS systems are excellent examples of the possibilities. Both of these examples rely on the sense of vision. In emergency situations in which vision is not possible these systems would be inefficient. The Zig-Zag sense of touch device gives a promising solution for these emergency situations.

John Miller's (UCSD) research in route preview software with audio cues for first responders also has an application for the blind community, as the blind can follow audio cues in the same way as first responders. He presented this technology to the Research

and Development committee for the National Federation of the Blind (NFB) 02/01/05. He will also present this technology to the blind community at large at the national convention for NFB in Louisville, KY, July 2005 (expected attendance of 2000). He plans to bring 6 prototype devices to the conference and will deploy them as a field trial there.

Courses:

- ECE 191, Winter 2005. Route Preview software with Audio Cues for first responders. (John Miller, UCSD) The research team worked to develop a suite of hardware and software location-based solutions to help first responders. The teams built two devices: 1) sense-of-touch evaluation module; 2) basic stamp computer control, which uses digital input for the sense-of-touch guidance system. These devices will eventually incorporate the Route-Preview software to help a first responder to orient himself during disaster situations.
- ECE 191, Spring 2005. Warehouse Assistant for First Responders (John Miller, UCSD). Follow-on project to ECE191, winter 2005. This project investigates how software, a barcode reader, and a web cam can help efficient and accurate inventory collection in, for example, a disaster scenario in which a non-medical person has been given a list of medical supplies from a medical warehouse, where they are not familiar with the layout. The software will help the first responder collect needed supplies efficiently without assistance of warehouse personnel. Details of the project and a free download of the application can be found at <http://rescue.calit2.net/warehouse>.
- ECE 291, Spring 2005. ZIGZAG Sense-of-Touch Guiding System with Computer Control and Remote Video (John Miller, UCSD). Follow-on project to ECE191, winter 2005. The team's responsibility was to design a computer interface to connect a prototype device to a digitally-controlled sense of touch servo system. The prototype device was designed to assist a blind person to navigate a path, or direct first responders whose senses of vision and hearing are saturated with real-time information or are reduced in a smoky or loud environment to use sense-of-touch to follow a route in a disaster situation. This opens the possibilities of GPS navigation with a sense-of-touch user interface. An initial emphasis of the design will be on computer control of a servo and a computer collecting GPS information. Additional work will complement the student's interest, be it hardware design, hardware integration, or software design. A related article showcasing the ECE 191 projects from Winter 2005 can be found at the following link: <http://www.calit2.net/articles/article.php?id=365>

As a follow on to the ZIGZAG project, and under the direction of RESCUE researcher Dr John Miller, UCSD undergraduate Javier Rodriguez Molina was hired to build 10 prototype ZIGZAG devices for field testing. In addition, undergraduate Diep Nguyen spent the academic year 2004-05 working on developing the Route-preview software for a sense of touch guidance system

Contributions.

Humans as Sensors: Notification System via Voice and Text

<http://traffic.calit2.net>

After being petitioned by the DOT, FCC has designated 511 (like 911) as the number for traveler information such as traffic, transit info and weather. We are developing a next-generation traveler information system that will serve as a powerful tool for first responders and commuters to relay, share and disseminate all types of critical information

The goal is to develop a fully automated system to collect and relay freeway-related information to the general public and to the first-responders. Though government agencies and the private sector have some of the basic data needed for effective traffic management, the means to effectively disseminate the data in an intelligent manner (i.e., delivery of relevant and timely information to the right target population) is lacking. Typically the data is disseminated in a broadcast mode. Also, in many situations, there is significant lag in the collection of crisis-related data by the government agencies. This lag can be eliminated by empowering the general public to report relevant information. Also, the current technique for obtaining traffic flow information by burying loop inductors at discrete points on the freeways is expensive and does not provide uniform coverage. Our approach is to tap into GPS-equipped cell phones to obtain position and velocity information. This approach provides uniform coverage.

The most compelling aspect of such a system is that information is disseminated in a targeted manner to people, with minimal delay. Currently, people call 911 if they see a severe accident and that information never cascades to the commuters other than through a vague traffic report on the radio with a long delay. Also, we can detect abnormalities based on the volume of calls received in any hour. If the volume of the calls spike, we know something must be wrong on the freeways. Indirectly the commuters are acting as sensors by calling in. We can also determine the location of the problem, by the highway they are requesting information for. One must also take into account the validity and truthfulness of the information the commuters are reporting since it will be easy for a user to spam the system. We will adopt a rating system which let only users who are regular users of the system to report incidents initially. Others will not have sufficient privileges. Given that traffic is the number one problem in San Diego according to a recent poll, if we can get 10%-20% of the population to adopt the system, this will serve as a powerful tool for the general public to relay, share and disseminate all types of critical information.

Cellular Platform for Location Tracking

Mobile phones have evolved into smart computing devices that provide instantaneous information anywhere, anytime. We have built a system to monitor and track the location of UCSD campus shuttle buses, based on the latest technology developed from Qualcomm, to allow GPS fixing from mobile phones. The system combines location information with multimedia data sharing among mobiles and servers. It includes a real-time map view of buses indicating their location, speed and identity (e.g., campus loop, east parking, etc.) on both web and mobile interfaces; and a management system

including mobile phone assignment to individual buses, messaging to one or more buses, and enabling or disabling the mobile-tracking feature (this feature is only available to bus dispatcher).

We have developed a set of specifications/requirements and built a client- and server-based system, where the client resides in the mobile phone and the server resides in a Java-based web server. The deployed system will demonstrate the availability of the position of a bus and multimedia data in real-time. The complete system includes the following components and related software: a mobile phone (with Assisted GPS, Camera, BREW 2.x); software at the mobile to perform GPS fixing and networking for data exchange with server (the GPS data includes the latitude, longitude, horizontal speed, heading and altitude); and a Java-based server for web-based presentation of GPS position, bus speed and bus identity. The system will be integrated with a GIS engine to display a UCSD campus map. Our continued work includes completing system integration, testing the system against the specifications/requirements, and conducting a field trial by installing the system onto campus buses at UCSD and testing.

Courses:

- ECE 191, Winter 2005: UCSD Campus Bus Monitoring System (John Zhu, UCSD) Two students participated in the project. The system is built with latest Assisted GPS (AGPS) technology. Based on AGPS, a mobile phone is used to track the position and speed of campus bus. A real-time map is provided to view location, speed, and identity of buses. The system combines the location information and multimedia data sharing among mobile phones and servers. The objective of this project is to build a client and server based system, where the client resides in the mobile phone and the server resides in Java based web server. The system will be deployed in a field setting, demonstrating the availability of position and multimedia data in real-time.

WIISARD Program

Description: WIISARD stands for: Wireless Internet Information System for Medical Response in Disasters. It entails the use of sophisticated wireless technology to coordinate and enhance care of mass casualties in a terrorist attack or natural disaster. The project brings together broad-based participation from academia, industry, the military, and emergency responders from the City and County of San Diego. WIISARD will provide emergency personnel and disaster command centers with medical data to track and monitor the condition of hundreds to thousands of victims on a moment-to-moment basis at the disaster site.

In summary, WIISARD is an integrated application that is bringing cutting edge wireless Internet technologies to the field treatment station. It has components that enhance the situational awareness of first responders, facilitate recording of medical data, aid in the monitoring of severely ill patients, and facilitate communication of data to hospitals. The WIISARD system will undergo evaluation throughout the 3-year contract, beginning with controlled studies of individual components and



culminating with a randomized trial conducted during a simulated WMD attack. This work is sported by contract N01-LM-3-3511 from the National Library of Medicine. The WIISARD team is grateful to DPAC Technologies and Ubicom, Inc. for donating equipment and software to aid the effort and to the San Diego Metropolitan Medical Response System.

WIISARD is producing a several unique components for possible manufacture and distribution: patient triage tags, wireless blood pulse oximeter, manpacks, and a software support package.

1. Doug Palmer, “An Intelligent 802.11 Triage Tag For Medical Response to Disasters”, AMIA 2005 Spring Conference.
2. Doug Palmer, “An Ontology of Geo-Reasoning to Aid in Medical Response to Attacks with Weapons of Mass Destruction”, AMIA 2005 Spring Conference.
3. Doug Palmer “MASCAL: RFID Tracking of Patients, Staff and Equipment to Enhance
4. Hospital Response to Mass Casualty Events”, , AMIA 2005 Spring Conference.
5. Doug Palmer, “802.11 Wireless Infrastructure To Enhance Medical Response to Disasters”, AMIA 2005 Spring Conference.

Responsphere infrastructure funding is aiding the WIISARD program effort by enhancing capabilities in these areas: Mobile Command Post, Tracking and Data Relay Units (TDRUs), mesh networking software/firmware, and terrain situation mapping of objects and geographic information system.

Courses/Academic Contribution: Seven ECE191 undergraduate projects and two ECE191 graduate projects were sponsored by WIISARD. These yielded significant training in real-world projects and were of benefit to the program overall. One Master’s thesis student is supported for work on the patient tag.

CalRadio Platform

Description: CalRadio is a wireless transceiver research program that has the goal of developing general broad application radio/networking research and development test platforms. These platforms are for broad access by the wireless community and are made available to the public on an open basis for Research and Development. A single integrated test platform gives a new dimension to radio design in the future: the capability of publishing standards in software/firmware and hardware. This has the utility of greatly speeding the design, implementation and adoption of new standards. CalRadio 1, which is already in limited production, supports an 802.11b soft-MAC that enables extensive research in mesh networking and sensor networks.

A new mezzanine board for CalRadio 1 is under development that



will allow software defined WiFi, BlueTooth, Zigbee and UWB development and is due in late summer 2005. It will also have the capability of ABL (Always Best Located) via triangulation with Time-of-Flight determination.

ABL is essential to many Calit2 programs in Homeland Security. The key to any information system to aid first-responders is to track people and assets. GPS technology is convenient and very powerful for some applications but it is ineffective in downtown areas and indoors. Calit2 intends to utilize TDRUs to relay information and Time-of-Flight data for all “tagged” clients in the disaster area to enable accurate positioning for the required asset tracking and management. The mezzanine board under development for CalRadio will utilize RFMagic chipsets for the PHY to enable a wide range of synthesis and receive waveforms and precision time synchronization to achieve these goals.

The CalRadio platform is now available to research institutions on an open Gnu style license. See <http://calradio.calit2.net/>.

Presentations/publications: CalRadio was unveiled at the 2005 IPSN/SPOTS symposium at UCLA in May and took top honors.

ResponSphere infrastructure support is enabling the field deployment of up to 25 CalRadio mesh network access points in one of the most extensive permanent field test facilities in the world. This will serve as a permanent test facility for researchers from around the world. Presently, deploying a mesh network for test and evaluation is a lengthy and time consuming process. This field setup will also allow ABL research to progress in a real-world environment.

Courses/Academic Contribution: Two ECE291 graduate students have participated in CalRadio development and one PhD student and one MSEE student are conducting thesis work on the platform.

Cellular Network Testbed

Activities and Findings.

System Overview

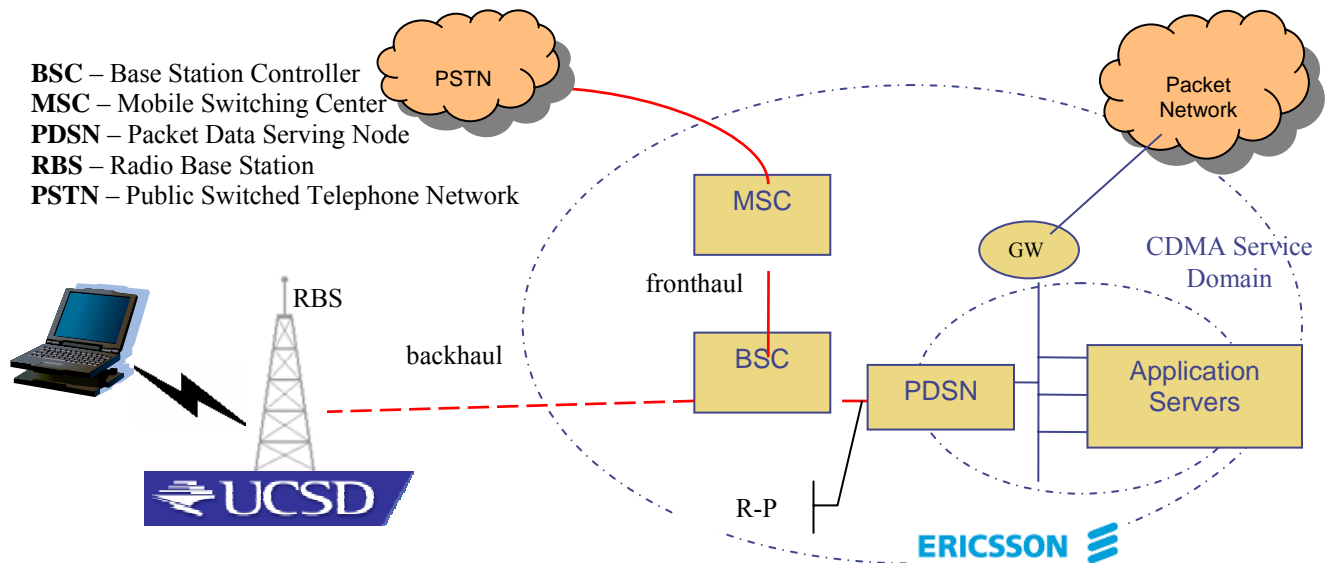
The facilities at Calit2 pertinent to this effort include access to an experimental CDMA 2000 1XRTT network through a base station that is housed at the UCSD campus. The base station was a donation from Ericsson and is being used by projects such as RESCUE, Adaptive Systems, and WhyNet to conduct research on network behavior.

A meeting was held between personnel from interested projects at UCSD, UCI, and UCLA on April 6, 2005 to discuss potential uses for the infrastructure, the tools, and potential research collaborations. A server has been setup at UCSD so that remote access to these tools can be granted to ResponSphere collaborators at UCI and WhyNet collaborators at UCLA. We plan to use TEMS Investigator & Deskcat to correlate RSSI(Receive signal strength indication) and BER; and collect RLP statistics and power

control data to understand the underlying relation to upper layers. We are also planning to collect long-term data logs (such as running a laptop on a shuttle) using TEMS products and make them available offline so that any researcher can work on it to understand the cross-layer interactions.

Products.

The Radio Base Stations (RBS) at UCSD is connected to a Base Station Controller (BSC) at Ericsson over a micro-link. Note that this is a commercial system but used only for experimental purposes, which allows us to control network activity and create outages, etc without disturbing actual customers.



This experimental network will be used to better understand the wireless channel and voice/packet data traffic characteristics

Contributions.

Data Access Points

In this network data collection for post-processing can be done at BSC, RBS, and mobile host.

BSC

Following parameters can be accessed from BSC reported during an interval

- Maximum number of simultaneous calls
- Number of active users on the fronthaul span
- For each sector (a RBS typically has three sectors)
 - Forward dedicated channel power utilization
 - Total transmit power

- Air interface Voice Equivalence (AVE) of current voice and data calls
- Number of active users
- Percentage of link utilization shown separately for forward and reverse links for each span at backhaul and fronthaul
- R-P link packet data user data capacity utilization
- Total transmit power
- Active voice users

Following real-time/ parameters per user may also be accessible

- Assigned bandwidth (supplemental channel rate)
- RLP (Link layer protocol used over the air for packet data) queue size
- Number of new sent and re-sent bytes per frame

RBS

Following data can be extracted from RBS during the reporting interval

For each sector

Duration of overall fundamental and supplemental channel usage

Number of supplemental channel setup requests

Average call load with respect to power (AVE)

Average number of fundamental channels

A histogram of the forward dedicated channel power utilization per transmit antenna

Mobile

An air interface tool, TEMS, is used to collect data at the mobile. Using TEMS, multiple access terminals can be measured simultaneously and measurements can be coupled with a GPS receiver to provide location information.

Mobile can provide the following information

Mobile Information (IMSI, ESN, MOB_P_REV, Slot cycle index)

System Information (System/Network ID, Frequency, Pilot PN, Walsh code etc.)

RLP statistics and power control information

Handoff information and number of RBSs in range and their pilot strengths

Measured data rate, observed error rate and frame error rate

TEMS can also collect air interface statistics from any commercially available CDMA2000 network such as Verizon and Sprint.

For post-processing TEMS Desccat can be used. Using this tool, different vendors' drive test data files can also be imported and analyzed. Both the TEMS and TEMS Desccat tools were jointly purchased by ResponSphere and WhyNet, as both projects stand to gain from the infrastructure as well as the collaboration.

Data Management and Analysis

Automated Reasoning and Artificial Intelligence

Activities and Findings.

Activities include:

- Develop new modeling frameworks that can effectively represent both discrete and continuous variables in presence of deterministic and probabilistic information on these variables.
- Develop algorithms for Learning our new modeling frameworks from data.
- Develop anytime algorithms for answering various queries on the learned model.
- Use our frameworks and algorithms to model a range of problems in transportation literature like finding Origin/Destination matrices for a town/region and determining behavior of individuals in different situations.

Findings Include:

Domain-independent modeling frameworks:

Our findings from the stand-point of domain-independent reasoning are two new modeling frameworks which we call Hybrid Mixed Networks and Hybrid Dynamic Mixed Networks.

Hybrid Bayesian Networks are the most commonly used modeling framework for modeling real-world phenomena that contains probabilistic information on discrete and continuous variables. However, when deterministic information is present, its representation in Hybrid Bayesian Networks can be computationally inefficient. We remedy this problem by introducing a new modeling framework called Hybrid Mixed Networks that extend the Hybrid Bayesian Network framework to efficiently represent discrete deterministic information.

Most real-world phenomena like tracking an object/individual also require the ability to represent complex stochastic processes. Hybrid Dynamic Bayesian Networks (HDBNs) were recently proposed for modeling such phenomena. In essence, these are factored representation of Markov processes that allow discrete and continuous variables. Since they are designed to express uncertain information they represent deterministic constraints as probabilistic entities which may have negative computational consequences. We remedy this problem by introducing a new modeling framework called Hybrid Dynamic Mixed Network that extends Hybrid Dynamic Bayesian Networks to handle discrete deterministic information in the form of constraints.

A drawback of our frameworks is that they are not able to represent deterministic information on continuous variables and a combination of discrete and continuous variables. We seek to overcome this drawback by extending our modeling frameworks to handle these variants.

Approximate Algorithms for Inference in our frameworks:

Once a probabilistic model is constructed, a typical task is to “query” the model which is commonly referred to as the “inference problem” in literature. We have developed various inference algorithms for performing inference in Hybrid Mixed Networks and Hybrid Dynamic Mixed Networks which we describe below.

Since the general inference problem is NP-hard in Hybrid Mixed Networks, we have to resort to approximate algorithms. The two commonly used approximate algorithms for inference in probabilistic frameworks are Generalized Belief Propagation and Rao-Blackwellised Importance Sampling. We seek to extend these algorithms to Hybrid Mixed Networks. Extending Generalized Belief Propagation to Hybrid Mixed Networks is straight-forward, however a straight-forward extension of importance sampling results in poor performance. To remedy this problem, we present a new sampling algorithm that uses Generalized Belief Propagation as a pre-processing step. Our empirical results indicate that our new algorithm performs better than our straight-forward extensions of Generalized Belief Propagation and importance sampling in terms of the approximation error. Our algorithms also allow trading of time with accuracy in a systematic way.

We also adapt and extend the algorithms discussed above to Hybrid Dynamic Mixed Networks which model sequential stochastic processes. We empirically evaluate the extensions of our algorithms to sequential computation on a real-world problem of modeling individual's car travel activity. Our evaluation shows how accuracy and time can be traded using our algorithms in an effective manner.

In future, we seek to extend our algorithms to handle variants of the modeling framework discussed in the previous subsection.

Modeling Car travel activity of individuals

We have developed a probabilistic model that given a time and the current GPS reading of an individual (if available) can predict the destination and the route to destination of the individual. We have used the Hybrid Dynamic Mixed Network framework for modeling this application. Some important features of our model are:

- Infer High level goals or the locations where the person spends significant amount of time.
- Infer Origin-Destination matrix for the individual. This matrix contains the number of times a user moves between given locations.
- Answer queries like:
 - Where the person will be after 10 minutes.
 - Is the person heading home or picking up his kids from a day-care center.

In future, we seek to extend this model to compute aggregate Origin-Destination models for a region/town and model behavior of individual to situations on the road like a traffic jam/accident.

Products.

1. Vibhav Gogate, Rina Dechter, James Marca and Craig Rindt (2005): Modeling Transportation Routines using Hybrid Dynamic Mixed Networks. Submitted to Uncertainty in Artificial Intelligence (UAI), 2005.
2. Vibhav Gogate and Rina Dechter (2005): Approximate Inference Algorithms for Hybrid Bayesian Networks in Presence of Constraints. Submitted to Uncertainty in Artificial Intelligence (UAI), 2005.

Contributions.

Our contributions to the body of scientific knowledge include:

Domain-independent modeling frameworks: We have developed a new framework that can be used to represent deterministic and probabilistic information on discrete and continuous variables under some restrictions.

Approximate algorithms for Inference in our frameworks: We have developed new approximate algorithms that extend and integrate well-known algorithmic principles like Generalized Belief Propagation and Rao-Blackwellised sampling for reasoning about our restricted framework. We evaluated our algorithms experimentally and found that the approximation achieved by our algorithms was better than the state-of-the-art algorithms.

Modeling Activity of individuals: We have developed a probabilistic model for reasoning about car-travel activity of an individual. This model takes as input the current time and the GPS reading of the person (if available) and predicts the destination and route to the destination of the person. Our probabilistic model is able to achieve better prediction accuracy than the state-of-the-art models

Supporting Keyword Search on GIS data

Activities and Findings.

In crisis situations there are various kinds of GIS information stored at different places, such as map information about pipes, gas stations, hospitals, etc. Since first responders need fast access to important, relevant information, it becomes important to provide an easy-to-use interface to support keyword search on GIS data. The goal of this project is to provide such a system, which crawls and/or archives GIS data files from different places, and builds index structures. It provides a user-friendly interface, using which a user can type in a few keywords, and the system can return all the relevant objects (in a ranked order) stored in the files. For instance, if a user types in “Irvine school”, the system will return schools in Irvine stored in the GIS files, the corresponding GIS files, possibly displayed on a map. This feature is similar to online services such as Google Local, with more emphasis on information stored in GIS data files.

Products.

We mainly focus on GIS data stored in structured relations on the Responsphere database server. We have identified a few challenges. No publications as of yet.

Contributions.

We are building a system prototype on keyword search on GIS data. Additionally, we are developing corresponding techniques. I believe these techniques could be applicable to other domains as well.

Supporting Approximate Similarities Queries with Quality Guarantees in P2P Systems

Activities and Findings.

Data sharing in crisis situations can be supported using a peer-to-peer (P2P) architecture, in which each data owner publishes its data and shares it with other owners. We study

how to support similarity queries in such a system. Similarity queries ask for the most relevant objects in a P2P network, where the relevance is based on a predefined similarity function; the user is interested in obtaining objects with the highest relevance. Retrieving all objects and computing the exact answer over a large-scale network is impractical. We propose a novel approximate answering framework which computes an answer by visiting only a subset of network peers. Users are presented with progressively refined answers consisting of the best objects seen so far, together with continuously improving quality guarantees providing feedback about the progress of the search. We develop statistical techniques to determine quality guarantees in this framework. We propose mechanisms to incorporate quality estimators into the search process. Our work makes it possible to implement similarity search as a new method of accessing data from a P2P network, and shows how this can be achieved efficiently.

Major findings:

- Similarity queries are very important in P2P systems.
- A similarity query can be answered approximately with certain quality guarantees by accessing a small number of peers.

Products.

One paper was submitted to a conference.

Contributions.

We developed a framework to support progressive query answering in P2P systems, and develop techniques to estimate qualities of answers based on objects retrieved in a search process. We have conducted experiments to evaluate the techniques. A paper was submitted to a conference.

Statistical Data Analysis of Mining of Spatio-Temporal data for Crisis Assessment and Response

Activities and Findings.

- Inferring “Who is Where” from Responsphere Sensor Data
 - Develop an archive of time-series data (over a time-period of at least 1 month) consisting of historical data from the UCI campus that relates to human activity over time on campus, including:
 - Real-time sensor data from the CALIT2 building, e.g., people-counts as a function of time based on detectors and video cameras located near entrances and exits.
 - Internet and intranet network traffic information from NACS, indicating usage activity over time (e.g., number of page requests from computers within UCI, every 5 minutes).
 - Traffic data (loop sensor from ITS) from freeways and surface streets that are close to UCI.
 - Context information such as number of full-time employees at UCI in a given week, academic calendar information such as academic holidays, class schedules/enrollment/locations, etc.

- Information on daily and weekly parking patterns, electricity usage, etc., if available.
 - Conduct preliminary exploratory analysis: data verification and interpretation, data visualization, detection of time-patterns (e.g., daily usage patterns), detection of outliers, characterization of noise, identification of obvious correlations across time series, etc.
 - Develop initial probabilistic model, estimation algorithms, and prediction methodology that can infer a probability distribution over how many people are (or were) on the UCI campus (or in a specific area or building on campus) at any given time (in the past, currently, or in the future). The framework will (a) fit a statistical model to historical time-series data, taking into account known systematic effects such as calendar and time of day effects and (b) be able to use this model to infer a distribution over the “hidden” unobserved time-series, namely the true number of people.
- Topic Extraction from Text
 - Apply topic extraction algorithms to collections of news reports, Web pages, blogs, etc., related to disasters, to assist in development of crisis-reponse Web portals.
- Video Data Analysis:
 - Finish and document work on developing general purpose stochastic models for modeling trajectories of human paths in areas such as plazas and similar spaces.

Products.

1. S. Parise and P. Smyth, Learning stochastic path-planning models from video images, Technical Report UCI-ICS 04-12, Department of Computer Science, UC Irvine, June 2004.
2. S. Gaffney and P. Smyth, Joint probabilistic curve-clustering and alignment, in Neural Information Processing Conference (NIPS 2004), Vancouver, MIT Press, in press, 2005.
3. M. Steyvers, P. Smyth, M. Rosen-Zvi, and T. Griffiths, Probabilistic author-topic models for information discovery, in Proceedings of the Tenth ACM International Conference on
4. Knowledge Discovery and Data Mining, ACM Press, August 2004.
5. M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, The author-topic model for authors and documents, in Proceedings of the 20th International Conference on Uncertainty in AI, Morgan
6. Kaufmann, 2004.
7. M. Rosen-Zvi, T. Griffiths, P. Smyth, and M. Steyvers, Learning author topic models from text corpora, Submitted to the Journal of Machine Learning Research, February 2005.

Contributions.

- Inferring “Who is Where” from ResponSphere Sensor Data
 - We have begun developing an archive of time-series data from the UCI campus that relates to human activity over time on campus:
 - Commercial "people-counter" devices and software (based on IR technology for door entrances and exits) have been ordered and will be tested within the next month or 2. If they work as planned a small number will be installed to collect data (in an anonymous manner) at the main CALIT2 entrances.
 - We have obtained preliminary internet and intranet network traffic data from ICS support and NACS, such as sizes of outgoing email buffer queues every 5 minutes, indicating email activity on campus.
 - Electricity usage data for the campus, consisting of hourly time series data, has been collected.
 - Current and historical class schedule and class size information has been collected from relevant UCI Web sites.
 - We have had detailed discussions with UCI Parking and Transportation about traffic and parking patterns over time at UCI and obtained various data sets related to parking.
 - We have conducted preliminary exploratory analysis of the various time-series data sets, including data verification and interpretation, data visualization, and detection of time-dependent patterns (e.g., daily usage patterns).
 - We have developed an initial probabilistic model (a Bayesian network) that integrates noisy measurements across multiple spatial scales and that can connect measurements across different times. We are currently working on how to parametrize this model and testing it on simple scenarios.
- Topic Extraction from Text
 - Topic extraction algorithms were successfully applied to news reports obtained from the Web related to the December 2004 tsunami disaster, and the resulting topic models were used as part of the framework for a prototype "tsunami browser".
- Video Data Analysis:
 - We developed a general framework, including statistical models and learning algorithms, which can automatically extract and learn models of trajectories of individual pedestrian movement from a fixed outdoor video camera. This work was written up in a technical report by student Sridevi Parise. The work has currently been discontinued since the student has left the RESCUE project.

Social & Organizational Science

RESCUE Organizational Issues

Activities and Findings.

Our activities for this year included applying for and receiving institutional approval to complete our research with human subjects, completing the planning process and developing the data collecting strategy, and initiating interviews with key personnel in the City of Los Angeles.

Our major findings for this year include:

- Barriers to the adoption of advanced technology are largely based upon issues of cost and culture.
- Information Technology for communication between field responders and incident command or emergency operations center is largely based upon voice communications; mainly hand held radios as well as a limited number of cell phones. Satellite phones and landlines are used in the EOC. Very few responders (mainly those at the command level) have access to handheld text-messaging devices (palm-pilots, PDAs, or blackberries) for communication in emergencies.
- Three radio systems are used for departmental communications in L.A. One system is held by LAPD, one by LAFD, and one for all the other departments. Because these systems are each built on different frequencies, they are unable to communicate with each without some form of interoperability technology.
- Visual communications in the form of streaming video are available to Incident Command Posts and the EOC via three means: LAPD helicopters, LAFD helicopters, and private industry mass-media. In addition, the Department of Transportation controls a series of video cameras (ATSAC) at key traffic-congested locations across the city. These cameras have been utilized for some EOC activations to monitor affected areas. LAPD has installed surveillance cameras in one community (for crime reduction purposes) and have plans to do so in others.
- Police and Fire vehicles are equipped with Mobile Data Terminals allowing them to communicate with dispatch centers. GPS systems are not yet available in these vehicles. They do not currently receive GIS imagery or streaming video due to bandwidth limitations. MDTs are useful for report-writing and resource tracking in the field, and displaying data on specific building structures.
- Data collection in the field is carried out by personnel from various departments and relayed through their Departmental Operations Centers to Incident Command or the Emergency Operations Center using low-tech means for communications. A common format is hand-written reports prepared in the field and then entered into personal computers back in the office.
- The Information Management System in the City of L.A. is self-hosted and is currently being changed over to a new system that is more intuitive for users, adaptable to different situations, and integrates more easily with GIS applications. The IMS for L.A. differs from the systems used by the County and the State posing some problems of interoperability and/or data sharing via computer systems.
- Information security for most city departments is managed through the Information Technology Agency. All departments are linked through the secure intranet; the

Police Department has an additional layer of security within the local intranet. The three proprietary departments of the city, Department of Water and Power, Harbor, and Airport, maintain their own security systems outside of the city's intranet. Access to all systems is protected through ID and passwords.

- The City of Los Angeles is actively involved in the development of interoperable radio systems within two area-wide response networks: the Urban Area Security Initiative (UASI), and the L.A. Regional Tactical Communications System (LARTCS).
- L.A. has three additional data hubs that have been established as data repositories. It also has three fully built-out alternate Emergency Operations Centers distributed through the city.
- Geographic Information Systems (GIS) data are maintained through the Information Technology Agency, the Bureau of Engineering in Public Works, and the Department of Building and Safety. Limited access to GIS data is available via on the each departmental website. There is no centralized data base for this mapping data or a common cataloging system. In a disaster activation, mapping experts are mobilized from ITA and Engineering.
- The most important forms of communication and decision making remains face-to-face communications followed by voice communications and mapping applications.

Products.

1. Tierney, K. J. 2005. "Interorganizational and Intergovernmental Coordination: Issues and Challenges." Presentation at "Crossings" Workshop on Cross-Border Collaboration, Wayne State University, Detroit, MI, March 15.
2. Tierney, K. J. 2005. "Social Science, Disasters, and Homeland Security." Invited presentation at the Office of Science and Technology Policy, Washington, DC, Feb. 8.
3. Sutton, J. 2004. "Determined Promise: Information Technology and Organizational Dynamics." Invited presentation at the RESCUE Annual Investigator's Meeting, Irvine, CA, Nov. 14.
4. Sutton, J. 2004. "The RESCUE Project: IT and Communications in Emergency Management." Invited presentation at the Institute of Behavioral Science, University of Colorado, Boulder, CO, Nov. 1.
5. Tierney, K. J. 2004. "What Goes Right When Things Go Wrong: Resilience in the World Trade Center Disaster." Invited lecture, Department of Sociology, University of California, Davis. Speaker series on "When Things Break Down: Social Systems in Crisis," Oct. 15, 2004
6. Interviewees from the City of Los Angeles, as well as RESCUE collaborators, will be invited to attend the Hazards Workshop in Boulder, Colorado held annually by the Natural Hazards Research and Applications Information Center.

Contributions.

We have completed interviews with personnel from six departments in the City of Los Angeles whose General Managers are representatives on the Emergency Operations Board. These departments include: Emergency Preparedness Department, Information Technology Agency, Los Angeles Police Department, Los Angeles Fire Department, Recreation and Parks, and the Board of Public Works – Bureau of Engineers. The total number of interviewees is 37. Additional interviews are being scheduled with personnel from the Department of Transportation, Department of Water and Power, General Services, Airport, and Harbor.

Interview questions focus on two areas: 1) identifying current technologies supported by the department for information/data gathering, information analysis, information sharing (with other agencies) and information dissemination to the public; and 2) security concerns and protocols for data sharing between agencies.

We attended two disaster drills; Determined Promise 2004/Operation Golden Guardian (August 5-6, 2004), and Operation Talavera (December 9, 2004). At these disaster drills, researchers made a written record of observations and conversations with responders or evaluators and took digital pictures of various technologies being used by command personnel and field responders. We also observed one training exercise with the LAPD (February 15, 2005), and attended an Emergency Operations Board meeting (March 21, 2005).

Current research papers in progress include “Barriers to the Adoption of Technology” and “Dynamic Evolving Virtual Organizations: DEVOS in Emergency Management.”

We have begun collaborations with UCI researchers to analyze data on crisis-related emergent multi-organizational networks (EMONs). We met in April, 2005 for a two-day workshop to discuss strategies for data cleaning, organizing, and analysis. An initial examination of our data set has been completed and we have identified a series of papers based upon this data set that will be developed during phase 2 of this research.

We have also provided research-based guidance to RESCUE researchers on social and behavioral aspects of information dissemination and provided basic domain knowledge on emergency preparedness, response, and recovery.

Social Science Research utilizing Responsphere Equipment

Activities and Findings.

Analysis of Responder Emergency Phase Communication and Interaction Networks (w/Miruna Petrescu-Prahova and Remy Cross): The emergency phase of an unfolding disaster is characterized by the need for immediate response to reduce losses and to mitigate secondary hazards (i.e., hazards induced by the initial event). Such response activities must be conducted quickly, within a high-uncertainty environment; thus, coordination demands are generally high. In this project, we study the networks of communication and work-based interaction among responders during crisis situations, with an emphasis on identifying implications for communication system design.

Thrust areas: Information Analysis, Information Sharing, Information Dissemination

Sub-thrust areas: Event Extraction from Text, Standard and Optimal Structural Forms in DEVOs, Behavioral Models for Social Phenomena

Approach: Extraction of networks from text, social network analysis

Objectives:

Extract network data sets from transcripts and other sources

Conduct initial analyses

Analysis of Response Organization Networks in the Crisis Context (w/Kathleen Tierney, Miruna Petrescu-Prahova, and Remy Cross): Crisis situations generate demand for coordination both within and across organizational borders; such coordination, it has been argued, is sufficiently extensive to allow the entire complex of crisis response organizations to be characterized as a single “Dynamically Evolving Virtual Organization” (or DEVO). In this project, we examine organizational networks associated with a major crisis situation, applying new methods to analyze structural features associated with communication and coordination.

Thrust areas: Information Analysis, Information Sharing, Information Dissemination

Sub-thrust areas: Event Extraction from Text, Standard and Optimal Structural Forms in DEVOs, Behavioral Models for Social Phenomena

Approach: Extraction of networks, social network analysis

Objectives:

Extract network data sets from archival and other sources

Conduct initial analyses

Bayesian Analysis of Informant Reports (w/Remy Cross): Effective decision making for crisis response depends upon the rapid integration of limited information from possibly unreliable human sources. In this project, a Bayesian modeling framework is being developed for inference from informant reports which allows simultaneous inference for informant accuracy and an underlying event history.

Thrust areas: Information Collection

Sub-thrust areas: Inferential Modeling of Unreliable Sources

Approach: Bayesian data analysis

Objectives:

Obtain and code a data set with crisis-related informant reports

Emergency Response Drills (w/Linda Bogue, Remy Cross, and Sharad Mehrotra): In order to both gather information on emergency response activities and test RESCUE research, we are currently collaborating with campus Environmental Health and Safety personnel in the development of emergency response drills informed by RESCUE research.

Thrust areas: Information Collection, Information Sharing

Sub-thrust areas: Video Analysis, Localization, Registration, Multi-modal Event Assimilation

Approach: Planning/consultation

Objectives:

Development of initial exercise plan

Plan implementation

Development of outline for future exercises

Network Analysis of the Incident Command System (w/Miruna Petrescu-Prahova and Remy Cross): The Incident Command System (ICS) is the primary organizational framework for conducting emergency response activities at the local, state, and federal levels. Effective IT interventions for crisis response organizations must take the properties of the ICS into account; this includes identifying structural features which may create unanticipated coordination demands and/or opportunities for information sharing. In this project, we conduct a structural analysis of the Incident Command System, using the Metamatrix/PCANS framework.

Thrust areas: Information Sharing

Sub-thrust areas: Standard and Optimal Structural Forms in DEVOs

Approach: Extraction of networks from planning documents, social network analysis

Objectives:

Identification of relevant documentation

Identification of Metamatrix vertex sets

Construction of adjacency matrices

Vertex Disambiguation within Communication Networks (w/Miruna Petrescu-Prahova): In contrast to the usual problems of error and informant accuracy arising from network data elicited via observation or self-report, network data gleaned from archival sources (particularly recordings or transcripts of interpersonal communications) raises the problem of vertex disambiguation: determination of the identity of individual senders and receivers. In this project, we develop stochastic models for network inference on communication networks with ambiguous vertex identities.

Thrust areas: Information Analysis, Information Collection

Sub-thrust areas: Stochastic Models, Inferential Modeling of Unreliable Sources

Approach: Simulation, analysis of test data

Objectives:

Development of initial model

Implementation of model

Application to test data

The major findings from this research over the past year include the following:

- In examining the radio communication networks of responders at the WTC disaster, we have found that emergency phase communications were dominated by a relatively small group of “hubs” working with a larger number of “pendants” (i.e., responders with a single communication partner). Very little clustering was observed, a property which strongly differentiates these networks from most communication structures observed in non-emergency settings. Still more surprising was the finding that communication structures among specialist responders (e.g., police, security, etc.) did not differ significantly from those of non-specialist responders (e.g., maintenance or airport personnel) on most structural properties. These findings would seem to suggest that emergency radio communication in WTC-like settings is not a function of training or organizational context, and that similar considerations should enter into the design of communication systems for both types of entities.
- While communication structure at the WTC showed surprising homogeneity across responder type, substantial differences were observed across individual responders. While most responders communicated a small number of times with very few

partners (often only one), some had many partners (in some cases covering 30-50% of the whole network). Those with many partners tended to act as coordinators, bridging many responders who did not otherwise have contact with one another. Such strong heterogeneity between positions suggests that effective communication systems for first responders must be able to cope with substantial interpersonal differences in usage. Even more significantly, subsequent analysis of coordinators at the WTC found approximately 85% to be “emergent,” in the sense of lacking an official coordinative role in their respective organizations. Thus, not only must effective responder communication systems deal with heterogeneous usage patterns, but system designers cannot assume that intensive users will be identifiable *ex ante*. The lack of significant differences between specialist and non-specialist responders in this regard indicates that this concern is no less relevant for established crisis response organizations than for organizations not conventionally involved in crisis response.

- Although analyses of the WTC police response are only preliminary, two major findings have already emerged. First – and in contrast with some observers' reports – the network of interaction among Port Authority police at the WTC site appears to be quite well-connected. While individual reports suggest that the response fragmented into small, isolated groups, the interactions of these groups over the course of the response appears to have led to a fairly well-integrated whole; because of the scale of this integration, it appears not to have been easily understood by individuals in the field. Despite this overall connectivity, police reports strongly suggest that effective localization technology might have aided the WTC response, particularly during the period following the collapse of the first tower (when many units lost visual contact due to the resulting dust clouds). Consonant with that observation, the extensive effort devoted to finding and tracking personnel seen in the WTC radio communications would seem to indicate that localization could permit faster evacuation during events of this kind.

In addition to the above findings, several major achievements of the past year include the following:

- A number of data sets have been created for use by RESCUE researchers. These include the WTC radio communications and police interaction data sets, data on interactions among response organizations, and weblog data relating to the Boxing Day Tsunami. In addition to yielding basic science insights in their own right, these data sets will facilitate “ground truthing” of technology development for many projects within RESCUE.
- RESCUE researchers have successfully conducted IT-enabled observation of an emergency response exercise. In addition to providing an opportunity to test RESCUE technology, this activity has provided experience which will be extremely valuable in supporting more extensive data collection. This activity also demonstrates successful cooperation with a community partner organization, generating (through opportunities for feedback and self-study) value-added for the wider community.
- A model has been constructed for network inference from data (such as radio communications) in which vertex identity is ambiguous. Ultimately, this model should reduce dependence on human coders to analyze transcripts, greatly increasing

the amount of responder communications data which can be processed by disaster researchers. In addition, this work is expected to serve as a point of contact for various groups within RESCUE (including those lead by Mehrotra and Smyth), all of whom are working on disambiguation problems of one sort or another.

Products.

1. Butts, C.T. 2005. "Permutation Models for Relational Data." IMBS Technical Report MBS 05-02, University of California, Irvine.
2. Butts, C.T. 2005. "Exact Bounds for Degree Centralization." *Social Networks*, forthcoming.
3. Butts, C.T. and Petrescu-Prahova, M. 2005. "Emergent Coordination in the World Trade Center Disaster." IMBS Technical Report MBS 05-03, University of California, Irvine.
4. Butts, C.T. and Petrescu-Prahova, M. 2005. "Radio Communication Networks in the World Trade Center Disaster." IMBS Technical Report MBS 05-04, University of California, Irvine.

Invited talks:

1. Butts, Carter T. "Building Inferentially Tractable Models for Complex Social Systems: a Generalized Location Framework." (8/2005). ASA Section on Mathematical Sociology Invited Paper Session, "Mathematical Sociology Today: Current State and Prospects." ASA Meeting, Philadelphia, PA.
2. Butts, Carter T. "Beyond QAP: Parametric Permutation Models for Relational Data." (10/2004). Quantitative Methods for Social Science Colloquium, University of California at Santa Barbara. Santa Barbara, California.

Conference presentations:

1. Butts, Carter T. and Petrescu-Prahova, Miruna. "Radio Communication Networks in the World Trade Center Disaster." (8/2005). ASA Meeting, Philadelphia, PA.
2. Butts, Carter T.; Petrescu-Prahova, Miruna; and Cross, Remy. "Emergency Phase Networks During the World Trade Center Disaster." (6/2005). Third Joint US-Japan Conference on Mathematical Sociology, Sapporo, Japan.
3. Butts, Carter T.; Petrescu-Prahova, Miruna; and Cross, Remy. "Responder Communication Networks During the World Trade Center Disaster." (2/2005). 25th Sunbelt Network Conference (INSNA), Redondo Beach, CA.

Contributions.

Analysis of Responder Emergency Phase Communication and Interaction Networks:

We have utilized data from the World Trade Center disaster (obtained from the Port Authority of New York and New Jersey) to construct two large network data sets for emergency phase responder interaction. The first of these data sets involves 17 networks (ranging in size from approximately 50-250 responders) describing interpersonal radio communications among various responder groups (including both specialist (e.g., security, police) and non-specialist (e.g., WTC maintenance workers) responders). The second data set consists of approximately 160 networks of reported communication and interaction, based on police reports filed by the Port Authority Police Department. Initial analyses of this data have already yielded several surprising findings regarding the structure of responder communication at the WTC.

Objectives:

Extract network data sets from transcripts and other sources
(accomplished)

Conduct initial analyses (accomplished)

Analysis of Response Organization Networks in the Crisis Context: Over the past year, we have obtained data on interorganizational response networks from six crisis events (originally collected by Drabek et al. (1981)) and have begun an initial analysis of this data. We have also worked with Kathleen Tierney vis a vis the coding of network data obtained from archival sources at the World Trade Center disaster. Although some work on this project was delayed due to administrative difficulties in obtaining the (legally sensitive) WTC data, the team has nevertheless met the primary objectives of gaining access to an initial data set and initiating analysis; due to the delay, more emphasis was placed on the analysis of responder communication networks (see above). Further progress was also made on techniques for analyzing DEVO structure, as reflected in the publications and technical reports (see above *Products*).

Objectives:

Extract network data sets from archival and other sources (accomplished)

Conduct initial analyses (accomplished)

Bayesian Analysis of Informant Reports: Progress centered on seeking a data source to test and validate the intertemporal informant report model. In particular, a systematic data collection framework was created for monitoring news items posted to English-language weblogs, based on a total sample of approximately 1800 blogs monitored four times per day. This framework was in place at the time of the Boxing Day Tsunami, and as such was able to collect initial reports of the disaster (as well as baseline data from the pre-disaster period). Monitoring of the sample was continued for the next several months, yielding a rich data set with extensive intertemporal information. Next year, this data set will be used to calibrate and validate the intertemporal informant reporting model for crisis-related events.

Objectives:

Obtain and code a data set with crisis-related informant reports
(accomplished)

Emergency Response Drills: Over the past year, regular contact was established between UCI's Environmental Health and Safety department (via Emergency Management Coordinator Linda Bogue) and RESCUE researchers. As a result of this relationship, we were able to participate in a number of EH&S training activities, and

were able to conduct an IT-enabled observation of a radiological hazard drill. This activity allowed the RESCUE team to test several technologies, including in situ video surveillance during crisis events, cell-phone based localization technology, streaming of real-time video from crisis events via wireless devices, and coordination of on-site observers with response operations. Experience from drill participation this year has led to a plan of action for further RESCUE involvement with exercises in the coming years, ultimately including the incorporation of Responsphere solutions in the drill itself.

Objectives:

- Development of initial exercise plan (accomplished)
- Plan implementation (accomplished)
- Development of outline for future exercises (accomplished)

Network Analysis of the Incident Command System: Practitioner documentation on the ICS was obtained from FEMA training manuals and other sources. This documentation was then hand-coded to obtain a census of standard vertex classes (positions, tasks, resources, etc.) for a typical ICS structure. Some difficulty was encountered during this process, due to the presence of considerable disagreement among primary sources. For this reason, it was decided to proceed with a relatively basic list of organizational elements for an initial analysis. Based on this list, adjacency matrices were constructed based on practitioner accounts (e.g., of task assignment, authority/reporting relations, task/resource dependence, etc.). Further validation of these relationships by domain experts will be conducted, prior to analysis of the system as a whole.

Objectives:

- Identification of relevant documentation (accomplished)
- Identification of Metamatrix vertex sets (accomplished)
- Construction of adjacency matrices (accomplished)

Vertex Disambiguation within Communication Networks: Transcripts of responder radio communications at the WTC were utilized as the test case for this analysis. Based on heuristics developed by a human coding of the radio transcripts, a discrete exponential family model was developed for maximum likelihood estimation of a latent multigraph with an unknown vertex set. This model utilizes contextual features of the transcript (e.g., order of address, embedded name or gender information, etc.) to place a probability distribution on the set of potential underlying structures; the maximum probability multigraph is then obtained via simulated annealing. This model has been applied to the WTC test data, where results appear to be roughly comparable to human coding. Further refinement of the model is expected to enhance performance vis a vis the human baseline.

Objectives:

- Development of initial model (accomplished)
- Implementation of model (accomplished)
- Application to test data (accomplished)

Transportation and Engineering

CAMAS Testbed

Activities and Findings.

Responding to natural or man-made disasters in a timely and effective manner can reduce deaths and injuries, contain or prevent secondary disasters, and reduce the resulting economic losses and social disruption. Organized crisis response activities include measures undertaken to protect life and property immediately before (for disasters where there is at least some warning period), during, and immediately after disaster impact. The objective of the RESCUE (Responding to Crisis and Unexpected Events) team at UCI is to radically transform the ability of organizations to gather, manage, analyze and disseminate information when responding to man-made and natural catastrophes. This project focuses on the design and development of a multi-agent crisis simulator for crisis activity monitoring. The objective is to be able to play out a variety of crisis response activities (evacuation, medical triaging, firefighting, reconnaissance) for multiple purposes – IT solution integration, training, testing, decision analysis, impact analysis etc.

Rationale of Research: The effectiveness of response management in a crisis scenario depends heavily on situational information (e.g., state of the civil, transportation and information infrastructures) and information about available resources. While crisis response is composed of different activities, the response can also be modeled as an information flow process comprising of information collection, information analysis, sharing and dissemination. The simulator will play out specific response activities can include evacuation, traffic management, triage and provision of medical services, damage assessment etc, that are usually under the control of an on-site, incident commander who in turn reports to a central Emergency Operations Center (EOC). The proposed simulator will model these different activities at both the macro and micro level and model the information flow between different entities. Models, technologies, decisions and solutions will be injected at interface points between these activities in the simulator or at specific junctures in the information flow to study the effectiveness of solutions/decisions in the response process.

Such an activity-based simulator consists of (a) different models that drive the activity, (b) entities that drive the simulation and (c) information that flows between different components. Models represent the spatial information, the scenario (location of entities, movement etc) as it is being played out, the crisis and its effect as it occurs, and the activity of different agents in the system as they take decisions. The entities of the simulator include agents, which might represent (or be) civilians, response personnel etc, communication devices, sensing technology and resources (like equipment, building etc). Agents have access to information based on the sensors available to them. Using this information agents make decisions, and take actions. Human models drive the decision making process of agents. In addition, we model the concept of an agent's role. Based on the role, the activities of the agent differ, and the access privilege to information varies. The activity model models the log of activities captured in the simulator (this could be done on a per-agent basis or at a more coarse-grained level).

To add further realism into the activity, the simulation will be integrated into a real-world instrumented framework (the UCI Responsphere Infrastructure) that can capture physical reality during and activity as it occurs. Responsphere is an IT infrastructure test-bed that incorporates a multidisciplinary approach to emergency response drawing from academia, government, and private enterprise. The instrumented space will be used to conduct and monitor emergency drills within UCI. Sensed

information of the physical world will be fed into the simulator to update the simulation and hence calibrate the activities at different stages. Such integration will also allow human to assume specific roles in the multi-agent simulator (e.g. a floor warden in an evacuation process) and capture decisions made by humans (citizens, response personnel) involved in the response process. The integration will therefore extend the scope of the simulation framework to capture the virtual and physical worlds and merge them into an augmented reality framework.

Products.

1. Alessandro Ghigi, Customized Dissemination in the context of emergencies, MS Thesis, University of California, Irvine, 2005
2. Jonathan Cristoforetti, "Multimodal Systems in the Management of Emergency Situations", MS Thesis, University of California, Irvine, 2005
3. Vidhya Balasubramanian, Daniel Massaguer, Qi Zhong , Sharad Mehrotra, Nalini Venkatasubramanian, The CAMAS Testbed, UCI Technical Report, 2005

Contributions.

Much of this research is a result of the refinement of the vision/goals/objectives set forth in the original proposal and the research tasks associated with the area were not explicitly called out in the proposal or agreement we had with NSF. So most of the task here are new and we have made significant progress along numerous fronts as listed below:

1. Design and Architecture of CAMAS: We are developing a multi-agent discrete event simulator that simulates crisis response. In this simulator response event is simulated with the information exchange highlighted. In order to play the response event we have designed agents which take the roles of common people, and response personnel. Human models are used to model the agents and some aspects of their decision taking abilities. In addition the agent behavior can be calibrated based on their human counterparts. Models of spatial temporal data are being studied in order to model the geography, and model the events as they occur.
2. Instrumented Framework for CAMAS: The CAMAS testbed consists of a pervasive infrastructure with various sensing (including video, audio), communication, computation and storage capabilities within the Responsphere infrastructure. While the testbed is designed to support crisis related activities including simulations and drills, during normal use time (when data is not being captured for a crisis exercise), a variety of other applications and users will be supported through the same Responsphere hardware/software infrastructure. An example of such an application is monitoring of equipment and people-related activities within the CallIT2 space. These applications provide the right framework and context for evaluating the research being developed in this project.
3. Modeling Spatial Information: There is a lot of spatial information that needs to be processed in this simulator, such as the geographical information about a given region, the location of agents etc. Geographical

space includes both indoor and outdoor spaces. Representation of spatial information is in databases and there is an interface available for agents to access this spatial information in order to make decisions. We are designing navigation algorithms for the evacuation of people. For the purpose of navigation different representations are superimposed on the basic database based on the requirements of agents and the algorithm. The navigation includes a hierarchical path planning, and obstacle avoidance.

4. Prototype system Implementation: We have developed a prototype v 0.1 of the CAMAS testbed that simulates evacuation of people inside a building (DrillSim).
5. Integration of Dissemination and Collection techniques in the testbed: Specific solutions for information dissemination and information collection have been developed and implemented in order to test on our framework. The information collection is a multimodal collection algorithm that collects localization information. The dissemination solution targets dissemination of information in a multi-agent context where agents have access different communication devices. Efforts are on to integrate these solutions into the CAMAS testbed.

Research to Support Transportation Testbed and Damage Assessment

Activities and Findings.

Significant progress has been made in creating a centralized web-based loss estimation and transportation modeling platform that will be used to test and evaluate the efficacy of Information Technology (IT) solutions in reducing the impacts of natural and manmade hazards on transportation systems. The simulation platform (INLET for Internet-based Loss Estimation Tool) incorporates a Geographic Information System (GIS) component, unlimited internet accessibility, a risk estimation engine, and a sophisticated transportation simulation engine. When combined, these elements provide a robust online calculation capability that can estimate damage and losses to buildings and critical transportation infrastructure in real time during manmade or natural disasters. A pre-Beta version of this internet web-based program has been developed using Active Server Pages (ASP) and Java Script to dynamically generate web pages, Manifold IMS as a spatial engine, and Access as the underlying database. This development was performed primarily on the Responsphere computational infrastructure. The basic components of this system have been tested and validated in the calculation of building losses and casualties for a series of historic earthquake events in southern California.

Preliminary system models have been created that are based on two major components; namely disaster simulation and transportation simulation. These two components interact with the information acquisition unit which is where the disaster event is detected and where disaster information is distributed. Figure 1 illustrates how these major elements interact with each other in the preliminary system model. The Information block represents where the IT solutions will emerge in this study. The Transportation Simulation block represents where detailed modeling takes place and where transportation system performance is assessed based on data and information collected on the extent and severity of the disaster. In the Disaster Simulation block, the impact of the

disaster in terms of economic losses and other impacts (such as casualty levels) is calculated. In this scheme, the results of the disaster simulation will also feed directly into the transportation simulation engine to identify damage to key transportation components and to assess probable impacts (such as traffic delays or disruption) to the transportation system.

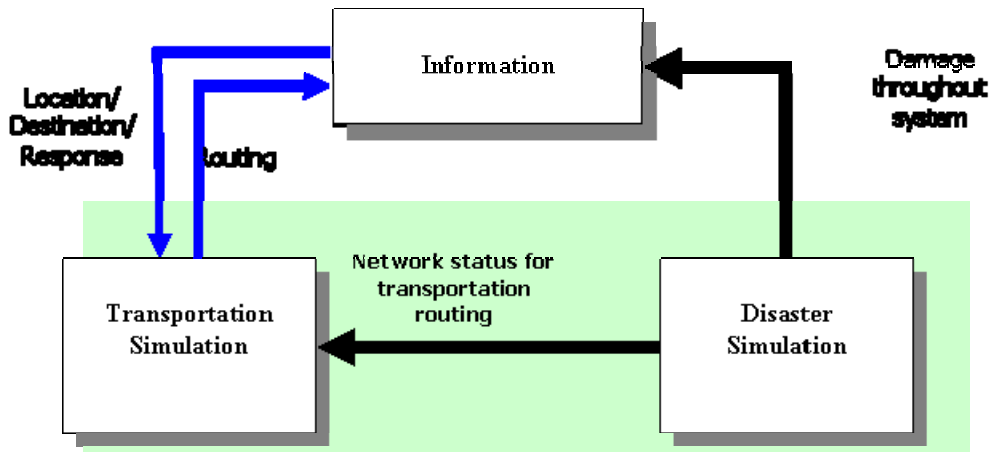


Figure 1. Web-based System for Loss Estimation and Transportation Modeling Platform. In addition to developing the computational platform for the Transportation Testbed, the research team was able to test and validate the application of the VIEWS (Visualizing the Impacts of Earthquakes with Satellite) after the December 26, 2004 Indian Ocean earthquake and tsunami. VIEWS was initially developed with funding from the Multidisciplinary Center for Earthquake Engineering Research (MCEER); it's application during the Indian Ocean earthquake and tsunami was a RESCUE research activity where data and information collected by other RESCUE investigators (Mehrotra, Venkatasubramanian) was integrated with moderate- and high-resolution satellite imagery served from within the VIEWS system.

Products.

1. Chung, Hung C., Huyck, Charles K., Cho, Sungbin, Mio, Michael Z., Eguchi, Ronald T., Shinozuka, Masanobu, and Sharad Mehrotra, "A Centralized Web-based Loss Estimation and Transportation Modeling Platform for Disaster Response," *Proceedings of the 9th International Conference on Structural Safety and Reliability*, Rome, Italy, June 19-23, 2005.
2. Sarabandi, Pooya, Adams, Beverley J., Kiremidjian, Anne S., and Ronald T. Eguchi, "Methodological Approach for Extracting Building Inventory Data from Stereo Satellite Images," *Proceedings of the 2nd International Workshop on Remote Sensing for Post-Disaster Response*, Newport Beach, CA., October 7-8, 2004.
3. Huyck, Charles K., Adams, Beverley J., and Luca Gusella, "Damage Detection using Neighborhood Edge Dissimilarity in Very High-Resolution Optical Data,"

Proceedings of the 2nd International Workshop on Remote Sensing for Post-Disaster Response, Newport Beach, CA., October 7-8, 2004.

4. Mansouri, Babak, Shinozuka, M., Huyck, Charles, K., and Bijan Houshmand, "SAR Complex Data Analysis for Change Detection for the Bam Earthquake using Envisat Satellite ASAR Data," *Proceedings of the 2nd International Workshop on Remote Sensing for Post-Disaster Response,* Newport Beach, CA., October 7-8, 2004.
5. Chung, Hung-Chi, Adams, Beverley J., Huyck, Charles K., Ghosh, Shubharoop, and Ronald T. Eguchi, "Remote Sensing for Building Inventory Update and Improved Loss Estimation in HAZUS-99," *Proceedings of the 2nd International Workshop on Remote Sensing for Post-Disaster Response,* Newport Beach, CA., October 7-8, 2004.
6. Eguchi, Ronald T., Huyck, Charles K., and Beverley J. Adams, "An Urban Damage Scale based on Satellite and Airborne Imagery," *Proceedings of the 1st International Conference on Urban Disaster Reduction,* Kobe, Japan, January 18-20, 2005.
7. Huyck, Charles K., Adams, Beverley J., Cho, Sungbin, and Ronald T. Eguchi, "Towards Rapid City-wide Damage Mapping using Neighborhood Edge Dissimilarities in Very High Resolution Optical Satellite Imagery – Application to the December 26, 2003 Bam, Iran Earthquake," *Earthquake Spectra,* In press.
8. Mansouri, Babak, Shinozuka, Masanobu, Huyck, Charles, and Bijan Houshmand, "Earthquake-induced Change Detection in BAM by Complex Analysis using Envisat ASAR Data," *Earthquake Spectra,* In press.
9. Chung, H., Enomoto, T., and M. Shinozuka, "Real-time Visualization of Structural Response with Wireless MEMS Sensors," *Proceedings of the 9th International Symposium on NDE for Health Monitoring and Diagnostics,* San Diego, CA., March 14-18, 2004.

Contributions.

1. Loss Estimation Modeling Platform

In the years following the 1994 Northridge earthquake, many researchers in the earthquake community focused on the development of loss estimation tools, such as HAZUS (HAZards-US). Advances in GIS technology, including those associated with desktop GIS programs and robust application development languages enabled rapid scenario generation that previously required intensive data processing work (in many cases, manual activities.) In some cases, these programs were linked with real-time earthquake information systems, such as the Caltech USGS Broadcast-of-Earthquakes (CUBE) system, or ShakeMaps (USGS) to generate earthquake losses in real-time. The Federal Emergency Management Agency (FEMA) funded the

development of a high-profile and widely-distributed loss estimation program called HAZUS (HAZards-US). Because of its flexibility in allowing users to import into the program custom datasets, the use of HAZUS in real events (e.g., 2001 Nisqually, Washington earthquake) sometimes resulted in conflicting results, especially when different users were involved in modeling the same event.

This issue (i.e., conflicting results) highlighted the need for a centralized system where data and model updates could cascade to multiple users. Currently, the research team is developing INLET as a simulation platform to test the efficacy of information technology solutions for crisis response. INLET could potentially become an online real-time loss estimation system that would be available to the all emergency management personnel.

The current state of GIS technology that is available in Internet Map Servers (IMS) complicates high-level online GIS analysis. IMS systems are used almost exclusively to serve data. Although these data may be dynamic, there are no geo-processing or analytical capabilities that are available to the programmer or user. Integrating higher-order GIS capabilities into an IMS environment required that the research team perform a thorough review of the technology. In this development process, several prototypes were produced: a) an initial prototype developed with JAVA, the University of Minnesota's MapServer software, and a POSTGRE SQL database with a PostGIS spatial component, b) an ARCGIS/ARCSDE implementation with DB2 and Spatial Extender, and 3) a Manifold implementation with a Microsoft Access database engine. After a thorough analysis of the benefits and costs of each prototype, the research team determined that the Microsoft programming environment was the most suitable alternative for rapid product development and testing, acknowledging that the eventual tool would have to be an open simulation tool. In early 2005, the Manifold prototype was expanded to a pre-Beta Internet web-based program using Active Server Pages (ASP) and Java Script, Manifold IMS, and Microsoft Access.

The development of INLET has (in most cases) utilized publicly-available damage and loss models. In many cases, these models have been simplified in order to adapt them to an online environment. Summarized below are changes that were made to these existing models:

Where possible, the program utilized databases and SQL for all calculations. SQL is a native database language, and queries written in SQL are optimized by the database. Custom functions in loss estimation programs generally use many individual calls to an underlying database, requiring data type conversions and causing slow read/write times. Using SQL queries, the entire calculation is stored in the database, hence, bypassing any need for type conversions and providing the fastest possible disk i/o.

The online program uses simplified damage functions to estimate loss, i.e., no "demand-capacity" curves are used. Our validation tests show there is a very good match between results from these simplified models and the original damage functions. Also, processing time has been drastically reduced.

Significant efficiencies result when INLET streamlines building stock databases, i.e., reduces the data into a small set of tables. Many building inventory values, such as

total square footage per census tract for each building category are pre-calculated. This reduces the number of calculations that INLET needs to perform.

Frequently accessed, intermediate calculations are saved to temporary tables, which avoids “joins” with two or more tables.

A user can customize a scenario by placing an earthquake (magnitude and location) on a map and INLET will associate that event with a particular fault. INLET will then calculate the ground motion patterns from that event and import that information into the loss estimation module.

Additionally, INLET will have the capability of estimating damage to single-site, such as highway bridges or cellular towers.

2. Transportation Modeling

In this research, modeling traffic patterns and drivers’ behavior has required addressing a number of factors, including assessing driver’s aggression and awareness levels. Modeling the collective behavior of drivers in panicked states has yet to be studied in detail (Santo, 2004) and is a major focus of this research. A prototype simulation tool has been developed and is illustrated in the simple test network described below. Figure 2(a) shows the simplex form of a transportation network with seven (7) zone centroids (red dots) connected by links. An unexpected event is simulated by the closing of a link. Furthermore, to exacerbate the situation a toxic plume is assumed to migrate towards the center of the grid, thus identifying a need to evacuate populations in the center. Figure 2(b) presents the results of a proof-of-concept experiment where travel times are calculated based on having no information on the event, half the drivers having information on the event, and all drivers having information on the event. The results clearly show that having some or complete information on the event can increase the number of drivers who escape (e.g., evacuating within 40 minutes) by 30% and 60%, respectively.

In the remaining part of this year, we will explore various alternatives for modeling the behavior of driver’s under disaster conditions. In particular, we will investigate how information technologies and intelligent messaging help to improve decision making and overall response.

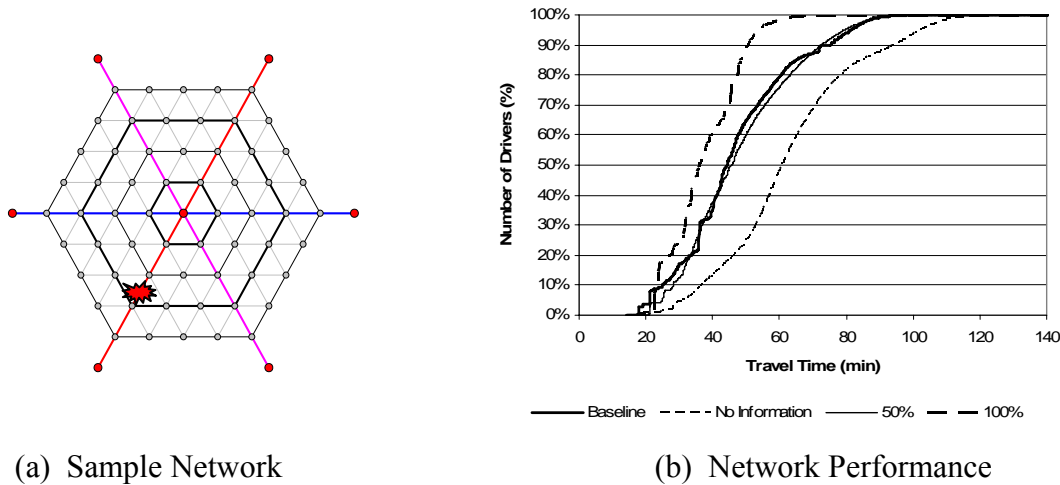


Figure 2. Sample Application of Transportation Simulation Tool

Additional Responsphere Papers and Publications

The following list of papers and publications represent a list of additional 2004-02005 research efforts utilizing the Responsphere research infrastructure:

1. Han, Qi, and Nalini Venkatasubramanian, Information Collection Services for QoS-based Resource Provisioning for Mobile Applications. *IEEE Transactions on Mobile Computing*, in revision.
2. Han, Qi, and Nalini Venkatasubramanian, Real-time Collection of Dynamic Data with Quality Guarantees. Technical Report TR-RESCUE-04-22. 2004. Submitted for publication.
3. Hore, Bijit, Sharad Mehrotra, and Gene Tsudik. Privacy-Preserving Index for Range Queries. Technical Report TR-RESCUE-04-18
4. Wickramasuriya, J., M. Datt, S. Mehrotra, and N. Venkatasubramanian. Privacy Protecting Data Collection for Media Spaces. Technical Report TR-RESCUE-04-21. Submitted for Publication 2004.
5. Yu, Xingbo, Sharad Mehrotra, Nalini Venkatasubramanian, and Weiwen Yang. Approximate Monitoring in Wireless Sensor Networks. Technical Report TR-RESCUE-04-16. Paper under submission.
6. Adams, B.J., C.K. Huyck, R.T. Eguchi, F. Yamazaki, M. Estrada, and C. Herring. QuickBird Imagery of Algerian Earthquake used to Study Benefits of High-

Resolution Satellite Imagery for Earthquake Damage Detection and Recovery Efforts (in press), by *Earth Imaging Journal*.

7. Butts, C.T. Exact Bounds for Degree Centralization. Institute for Mathematical Behavioral Sciences Technical Report MBS 04-09, University of California, Irvine.2004.
8. Kalashnikov, D. and S. Mehrotra. An algorithm for entity disambiguation. University of California, Irvine, TR-RESCUE-04-10.
9. Kalashnikov, Dmitri V., and Sharad Mehrotra. ReIDC: a novel framework for data cleaning. RESCUE-TR-03-04.
10. McCall J. and M. M. Trivedi, "Pose Invariant Affect Analysis using Thin-Plate Splines" *Proceedings of International Conference on Pattern Recognition 2004*, to appear.
11. Ma, Yiming, Sharad Mehrotra, Dawit Yimam Seid, Qi Zhong. Interactive Filtering of Data Streams by Refining Similarity Queries. Technical Report TR-RESCUE-04-15.
12. Ma, Yiming, Qi Zhong, Sharad Mehrotra, Dawit Yimam Seid. A Framework for Refining Similarity Queries Using Learning Techniques. Submitted for publication, 2004.
13. Hore, Bijit, Hakan Hacigumus, Bala Iyer, Sharad Mehrotra. Indexing Text Data under Space Constraints. Technical Report TR-RESCUE-04-23. 2004. Submitted for review.
14. Iyer, Bala, Sharad Mehrotra, Einar Mykletun, Gene Tsudik, and Yonghua Wu. A Framework for Efficient Storage Security in RDBMS. E. Bertino et. al. (Eds.) *EDBT 2004, LNCS 2992*, pages 147-164. Springer-Verlag: Berlin Heidelberg, 2004.
15. Wickramasuriya, J., N. Venkatasubramanian. Middleware-based Access Control for Ubiquitous Environments. Technical Report TR-RESCUE-04-24. 2004. Submitted for publication.
16. Yang, Xiaochun, and Chen Li. Secure XML Publishing without Information Leakage in the Presence of Data Inference. *VLDB 2004*.
17. Butts, C.T. and M. Petrescu-Prahova. "Radio Communication Networks in the World Trade Center Disaster." IMBS Technical Report MBS 05-04, University of California, Irvine, 2005.

18. Deshpande, M., Venkatasubramanian, N. and S. Mehrotra. "Scalable, Flash Dissemination of Information in a Heterogeneous Network". Submitted for publication, 2005.
19. Deshpande, M., Venkatasubramanian, N. and S. Mehrotra. "I.T. Support for Information Dissemination During Crisis". UCI Technical Report, 2005

Courses

The following undergraduate and graduate courses are facilitated by the Responsphere Infrastructure, use Responsphere equipment for research purposes, or are taught using Responsphere equipment:

ICS 105 (Projects in Human Computer Interaction),

ICS 123 (Software Architecture and Distributed Systems),

ICS 280 (System Support for Sensor Networks)

SOC 289 (Issues in Crisis Response).

ECE 191 (Fall 2004) (Wireless Sensor Network)

ECE 191 (Fall 2004) (PulseOx Component Design)

ECE 191 (Winter 2005) (Man Down Detection Device)

ECE 191 (Winter 2005) (ZIGZAG Sense-of-Touch Guiding System with
Computer Control)

ECE 191 (Winter 2005) (WiFi PulseOx Component Design)

ECE 191 (Winter 2005) (Wireless Sensor Network)

ECE 191 (Spring 2005) (Wireless Sensor Networks)

ECE 191 (Spring 2005) (Warehouse Assistant for First Responders)

ECE 191 (Spring 2005) (WiFi Pulse-Oximeter Breadboard and Integration)

ECE 291 (Spring 2005) (Cal-RADIO WiFi Research Project)

ECE 291 (Spring 2005) (ZIGZAG Sense-of-Touch Guiding System with

Computer Control and Remote Video)

ECE 291 (Spring 2005) (Software Based MIMO Wireless Communication

Testbed)

Dr Leslie Lenert gave a guest lecture as part of a symposium series titled “Public Policy and Biological Threats,” sponsored by the UCSD Institute on Global Conflict and Cooperation (Graduate School of International Relations and Pacific Studies).

In addition to project-based courses, we are exploring creating new engagements with existing curricula and programs, such as:

- We will explore opportunities for engagement with the Occupational Safety and Health Administration (OSHA) Training Institute at UCSD Extension. This institute offers OSHA training for the private sector and federal agencies, and has had over 15,000 safety and health professionals in attendance since 1992.
- Through our collaborations with Robert Welty, Director for Homeland Security Projects for the San Diego State University Foundation and Associate Director of the SDSU Visualization Center, we will explore participating in the new Master of Science degree specialization program in Emergency Preparedness and Response. Graduate which their School of Public Health has designed to produce public health leaders specifically trained to protect communities and respond to the unique health threats posed by disease outbreaks, acts of terrorism or massive natural disasters.
- We will explore guest lecturing for symposium courses on homeland security offered by the graduate school of International Relations and Pacific Studies (IRPS) at UCSD, and the possibility of working with IRPS to incorporate Responsphere-related topics into these symposium courses.

Equipment

This page summarizes the types of equipment we obtained for the project. The most significant purchases were an IBM e445 Xseries 8-processor computer, multi-modal sensing devices (cameras, people-counters, RFID, PDAs, and audio sensors), IBM 4 TB disk array, visualization projectors, and a number of PCs or laptops for graduate students. Our strategic partnership between Responsphere and IBM allowed us to purchase the e445 server and the RAID, originally priced at \$110,279, for \$76,204. Additionally, a partnership with CDWG (retailers for Canon and Linksys) resulted in significant cost savings for cameras pricing. In all cases, education pricing and discounts were pursued during the purchasing cycle.

UCI Equipment		
Date	Equipment	Usage
8/1//2004	Laptop	Grad student programming
9/12/2004	Activewave Inc.	RFID equipment
9/17/2004	Text Analysis Int.	Information extraction software

10/11/2004	DNS Registration	DNS for: www.responsphere.org
11/1/2004	Sony Handi-Cam	Portable camcorder for filming drills
1/15/2005	HP iPAQ 6315	sensing and communication devices for first responders
2/1/2005	Linksys WVC54G	Ehernet + WiFi audio/video cameras
2/1/2005	POE Adapters	Power Over Ethernet injectors
2/1/2005	Echelon	Powerline Networking topology
2/1/2005	Canon VB-C50i	Tilt/pan/zoom cameras
3/9/2005	HP Navigation system	GPS system
3/29/2005	Laptop	Sharad
4/1/2005	IBM e445 server	Main compute/processing system
4/1/2005	IBM EXP 400	4TB RAID storage for Responsphere
4/1/2005	Laptop	Presentation
4/1/2005	APC	UPS for server and RAID
4/1/2005	Electricians	Wiring/cabling for Smart-Space (Responsphere)
4/21/2005	Dell	12 Grad Student PCs
5/1/2005	Dell	5 Grad Student PCs
5/1/2005	RetailReserach	People Counters
5/1/2005	Electricians	Wiring/cabling for Smart-Space (Responsphere)
5/10/2995	Microsoft	Software licensing
6/1/2005	Canon	Projectors: Visualization System

UCSD Equipment

Date	Equipment	Usage
9/28/2004	Networks Plus Technology Point Research Corporation	Projectors
9/29/2004	CSI Wireless	GPS receiver evaluation kit and development board
12/8/2004	QVidia Technologies	GPS localization equipment
8/23/2004	Officetronics	High-definition video transmission system
9/24/2004	Ericsson Netqual	Wireless whiteboard capture systems
11/15/2005	Tropos Networks	TEMS network analysis kit (cost split with WhyNet)
2/14/2005	PDM Net, Inc.	Antennas for GLQ
5/18/2005	Ericsson Netqual	Wireless nodes, transmitters/receivers for GLQ
4/6/2005	UCSD Bookstore	TEMS deskcat network analysis software (cost split with WhyNet)
4/12/2005	UCSD Bookstore	Storage media
4/7/2005	Synchrotech, Inc	Digital Camera and accessories
4/22/2005	Anixter, Inc	PCMCIA card bus adapter for CyberShuttle
5/16/2005	Fry's Electronics	VGA cable for CVC
		Handheld Radios