Center for Research in Wireless Mobility and Networking (CReWMaN)

Pervasively Secure Infrastructures (PSI) Grant No. IIS-0326505

NSF Second Annual Report (June 2004 - May 2005)
http://crewman.uta.edu/psi

Abstract

The second year of the PSI project was a successful cohesion of work from five overlapping domains: Wireless Sensor Networks, Computer and Network Security, Pervasive Computing, Machine Learning and Databases. We have significantly expanded our research in each of these areas so as to position ourselves in building an integrated security system that can observe and react to threats posed by the ever changing environment in an adaptive manner.

At the end of the second project year, we are shifting our focus from research to prototyping; we are in the first phase of development of a high quality testbed consisting of wireless sensors, RFID tags, mobile devices backed by a high speed core network that will help us better understand the real world challenges of deploying a Pervasively Secure Infrastructure.

1 Wireless Sensor Networks (WSN) and Network Security.

Faculty Involved : Kalyan Basu, Sajal Das, Frank Lewis.
Student Members : Afrand Agah, Amitabha Ghosh, Habib Ammari, Hyun Jung Choe, Jun-Won Ho, Pradip De, Prasanna Ballal, Sumantra Kundu, Wei Zhang, Wook Choi.
WebPage: http://www.crewman.uta.edu/psi, http://arri.uta.edu/acs/

In the realm of WSN and Network Security, during the second year, we researched energy efficient data gathering, secure data dissemination and routing strategies among partipating sensor nodes. The research has also helped us to understand and develop representative heuristics and algorithms for detecting rogue nodes and intrusion in WSN. Because data collected from the sensor nodes is voluminous, it needs to intelligently forwarded over the wired network. We have successfully developed and patented an algorithm that is capable of structured data processing in the traditional infrastructure network. We highlight key features of our research that exemplify our progress in the above areas.

Energy and Application-Aware Data Gathering in Wireless Sensor Networks

A wireless sensor network is an application-specific information gathering platform where sensors are required to sense their vicinity (sensing coverage) continuously, consuming highly limited resources such as energy which may not often be replenishable. Thus, an important issue in sensor networks is to design energy-aware algorithms and protocols that optimize energy consumption with a goal to extend the network lifetime while meeting the user requirements such as coverage and data reporting latency.

The sensitivity to these requirements varies depending on the *type of applications*, implying that the designed algorithms and protocols must also be application-aware. In this part of the research, we have developed a *coverage-adaptive energy and application-aware data gathering strategy*, called CA-DGS, for wireless sensor networks [14][15][18]. We have also investigated the Poisson sampling technique to improve the spatial regularity of selecting k sensors and have proposed an enhanced random sensor selection scheme, called constrained random sensor selection (CROSS) [13].

The CA-DGS is based on a trade-off between sensing coverage and data reporting latency. The basic idea is to select in each round k data reporters (sensors) which are sufficient for the desired sensing coverage (DSC) specified by the users/applications. For selecting k reporters in a round, we make use of three efficient coverage-adaptive random sensor selection (CANSEE) schemes [16][17]. These reporters form a data gathering tree and are scheduled to remain active for that round only. This process incurs some delay but saves energy. We have derived a probabilistic bound on k and also measured the probability for almost surely having k data reporters in each round.

Based on the initial probabilistic analysis, we have observed that the CROSS scheme improves the connectivity of the selected sensors and reduces the variance of the sensor covered area in each round. Simulation results have demonstrated that CA-DGS results in a significant conservation of energy with a small trade-off in terms of data reporting latency. In particular, the higher the network density, the higher is the energy conservation without any additional computation cost.

Sensor and Decision Networks for Machine Diagnosis and Prognosis

For machinery condition monitoring, we have developed an energy-efficient method for data gathering in a wireless sensor network . A generalized battery consumption model was developed that is based on matrices for computation of the awake/sleep times of nodes in the WSN. This equation allows fast computation of new times when nodes are added to or removed from the network. A supervisory controller was developed that combines scheduling and contention. A finite state machine runs on each node to keep track of changes in state needed for monitoring, battery conservation, and data transmission.

Supervisory Control of Mobile Wireless Sensor Networks

We have applied a matrix-based discrete event controller (US Patent) to wireless sensor networks for supervision in task sequencing and resource assignment. The needs of WSN are not the same as other discrete event systems, and the controller was modified to allow for multiple missions, mission priority interrupts, and fast assignment of sensors without blocking phenomena. The WSN consists of some fixed unattended ground sensors, and some mobile sentry nodes that can vary their location to enhance the capabilities of the WSN in terms of repairing damage, compensating for faults, providing additional sensor information, and responding to detected events. The discrete event controller assigns the mobile nodes to tasks to carry out programmed missions, and has alarm capabilities when certain events are detected. Nodes can coordinate to improve sensory information about events. Routing decisions are required for decision-making in multiple assigned missions.

Adaptive Sampling in WSN

We have developed an adaptive sampling algorithm in mobile WSN for environmental monitoring. A dynamical model is created that describes both the robot mobility and the uncertainty due to data gathering. This allows combining navigation and measurement so that the mobile sensors are directed to make measurements at those locations that provide the greatest decrease in uncertainty. A Bayesian method based on error covariances allows prediction of the best next-sampling locations.

Modeling Broadcast Protocols

Disseminating a piece of information from any node to all the nodes of a sensor network or to a large group of nodes is an important requirement for wireless sensor networks. We are curently working on a mathematical model based on the underlying concepts of Epidemic Theory for estimating the spread of information dissemination in a large scale sensor network; since it is necessary from a source node's perspective to estimate the extent of the propagation.

We are modeling the propagation of data dissemination as an active infection spread in a susceptible population. This model will provides us with a time varying estimate of the fraction of the whole network that the information has ultimately spread to.

Distributed Greedy Algorithm for Connected Sensor Cover in Dense Sensor Networks

Achieving optimal battery usage and prolonged network lifetime are two of the most fundamental issues in wireless sensor networks. By exploiting node and data redundancy in dense networks, and by scheduling nodes efficiently, minimum battery drainage is possible. In this research, we focussed on the problem of Minimum Connected Sensor Cover (MCSC), an NP-hard problem, and develop a distributed greedy algorithm to generate sub-optimal connected sensor covers for homogeneous dense static sensor networks. Our greedy algorithm is based on the notions of maximal independent sets on random geometric graphs, and on the structure of Voronoi diagram. We developed complexity analysis and bounds on the cardinalities of maximal independent sets (MIS) for our problem scenario, and derive an analytical expression for the size of the sub-optimal minimum connected sensor cover.

Characterizing Least Detectable Paths with Coverage in Heterogeneous Sensor Networks

The basic Voronoi diagram based approach does not fit well when one considers a realistic sensing model and the heterogeneity of nodes. In this research, we describe a new model, the Apollonius diagram, that fits the more generic sensing model, that of a circular disk, and can take into account nodes with different sensing radii. We have defined the notion of accumulated total area coverage of an object traversing the sensing field over a certain time interval, and then related it to the existence of vulnerable paths in the network and its detection capability. We also studied the variation of the confidence probability for non existence of a least detectable path with the number of nodes in the network for different sensing radii using Matlab.

Secure Auction based Routing (SAR)

A malicious node can misrepresent its identity in the network and issue route error messages to misdirect the path or drop incoming packets. To circumvent and counter such behaviour, we have developed a routing protocol based on an auction framework. This will empower us to isolate malicious nodes and provide secure routing in wireless sensor networks.

In the proposed protocol, sensors willing to route data compete against each other and ultimately gain reputation based on their actions. The competition is based on auction theory. The amount of a bid that each node offers is equal to its utility value; and the price that a winner of a bid pays is a reduction of its original energy power. We have shown that a node's truthful bidding remains a dominant strategy and so to have a secure routing protocol, we need to isolate malicious nodes who do not bid truthfully. The scheme is called Secure Auction based Routing (SAR). It incorporates the total bid of each route in data packets. A performance analysis of SAR has also been carried out and compared with other well known protocols like INSENS and CONFIDANT protocols. Based on simulation and synthetic data, experimental results show that the proposed auction based framework significantly increases the chances of success in defense strategy for wireless sensor networks.

Secure Data Aggregation in Wireless Sensor Networks

Data aggregation (in-network processing) is one of the most common used approaches to reduce the redundant messages and save energy. The objective of this research is to secure aggregation against active attacks in sensor networks, especially when some of the sensor nodes are compromised.

In this work, we are developing a reputation based trust model framework so as to detect and filter false data caused by either faulty sensor nodes or compromised nodes, thus, provide protection against false data injection attacks. Specifically, techniques based on information theory are being developed as as to quantify the reputation of a node. Preliminary simulation results show that our framework can effectively detect the compromised nodes based on a node's reputations matrix.

Decentralized Detection of Sensor Replicates in Wireless Sensor Networks

For cost-effectiveness, sensor nodes are usually not equipped with tamper-resistant hardware. When these unshielded sensor nodes are deployed in hostile environments, the adversary can easily capture and generate malicious clones of some of the nodes. The adversary may also launch various kinds of attacks to paralyze normal operations after putting such cloned nodes at pre-destined places. To protect sensor networks from attacks mounted by replicated nodes, we are researching on an efficient decentralized detection mechanism.

The basic idea of our detection mechanism is to discover, among randomly chosen nodes, duplicated instances of identity claims announced by the nodes. Prior to the deployment of the nodes, N unique identity codes are generated and hash image of each identity code is installed on each of the N nodes, where N is the number of deployed nodes in the network. This N unique identity codes generation can be achieved by generating N unique random numbers with ANSI X9.17 Pseudorandom number generator (PRNG). For each identity code, N sub-identity codes are generated by threshold secret sharing scheme in such a way that an identity code can be reconstructed with any two sub-identity codes. N distinct sub-identity codes that map to N distinct identity codes are installed in each node.

Initial simulation results indicate that such an approach can indeed thwart mounted attacks by replicated sensor nodes in WSN.

Publications

[1] A. Agah, S. K. Das and K. Basu, "Preventing DoS attack in Sensor and Actor Networks: A Game Theoretic Approach", IEEE International Conference on Communications (ICC), Seoul, Korea, May 2005.

[2] A. Agah, S. K. Das and K. Basu, "A game theory based approach for security in sensor networks", IEEE International Performance Computing and Communications Conference (IPCC), 2004.

[3] S. K. Das, A. Agah and M. Kumar,"Security Issues in Pervasive Computing," *Wireless Information Highways* (Ed: D. Katsaros), Chapter XV, pp. 421-440, 2005.

[4] S. K. Das and A. Agah, "Security in Wireless Mobile and Sensor Networks," *Wireless Communications Systems and Networks* (Ed: M. Guizani), Chapter 18, pp. 531-557, Kluwer Academic Pub, 2004.

[5] A. Ghosh and S. K. Das, "A Distributed Greedy Algorithm for Connected Sensor Cover in Dense Sensor Networks," *Proceeeings of IEEE/ACM International Conference on Distributed Computing in Sensor Systems* (DCOSS), Marina del Ray, CA, June 30 - July 1, 2005.

[6] J. Mireles, G. Bravo, F.L. Lewis, R. Campos, A. Ramirez, On the development of MatrixC, a discrete event controller supervisor: implementation of a fuzzy logic dispatching approach, Proc. IFAC Symp. Systems Structure and Control, Oaxaca, Mexico, Dec. 2004.

[7] M. Sun, F.L. Lewis, S.S. Ge, Platoon-Stable Adaptive Controller Design, Proc. IEEE CDC, pp. 5481-5486, Bahamas, Dec. 2004.

[8] S. Chaudhuri, R. K. Ghosh and S. K. Das, "Towards Optimal Sensor Placement with Hypercube Cutting Planes" *Proceedings of IEEE Wireless Communications and Networking Conference* (WCNC), New Orleans, Mar 2005 (Invited Paper).

[9] S. R. Kundu, S. K. Das and K.Basu, "Finite State Markov Model for Effective Bandwdth Calculation in Wireless Packet Networks", 3rd Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt),Riva del Garda, Trentino, Italy 2005.

[10] C. Schuba, S. R. Kundu and J. Goldschimdt, "Method and Apparatus for Fast Scalable Matching of Structured Data Stream", Patent application filed in the United States, Jan., 2005.

[11] V. Giordano, F.L. Lewis, J. Mireles, B. Turchiano, Coordination control policy for mobile sensor networks with shared heterogeneous resources, Proc. Int. Conf. Control & Automation, Budapest, June, 2005, to appear.

[12] V. Giordano, F.L. Lewis, B. Turchaino, P. Ballal, V. Yeshala, Matrix computational framework for discrete event control of wireless sensor networks with some mobile agents, Proc. Mediterranean Conf. Control & Automation, Limassol, Cyprus, June 2005, to appear.

[13] W. Choi, S. K. Das, and H. J. Choe,“Constrained Random Sensor Selection for Application-specific Data Gathering in Wireless Sensor Networks”, In Proceedings of IEEE International Workshop on Embedded Networked Sensors (EmNetS-II), 2004. (To Appear)

[14] W. Choi, ”A Novel Framework for Energy and Application-Aware Data Gathering in Wireless Sensor Networks”, PhD Thesis, CSE Dept, UTA, May 2005.

[15] W. Choi and S. K. Das, ”A Novel Framework for Energy-Conserving Data Gathering in Wireless Sensor Networks”, In Proceedings of IEEE conference on computer communications (INFOCOM), 2005.

[16] W. Choi and S. K. Das,”Coverage-adaptive Random Sensor Selection with Latency Control for Application-specific Data Gathering in Wireless Sensor Networks”, In Proceedings of IEEE International Conference on Control and Automation (ICCA), 2005. (Invited Paper)

[17] W. Choi and S. K. Das, “Trade-off Between Coverage and Data Reporting Latency for Energy-Conserving Data Gathering in Wireless Sensor Networks,” *Proceedings of the First IEEE Interational Conference on Mobile Ad hoc and Sensor Systems* (MASS), Ft. Lauderdale, Florida, pp. 503-512, Oct 2004.

[18] W. Choi and S. K. Das, and K. Basu, “Angle-based Dynamic Path Construction for Route Load Balancing in Wireless Sensor Networks,” *Proceedings of IEEE Wireless Communications and Networking Conference*, Atlanta, GA, Mar 2004.

2 Pervasive Computing and Communications

Faculty Involved : Behrooz Shirazi, Mohan Kumar, Sajal Das
Student Members : Hitha Alex, Huaping Shen, Nirmalya Roy, Pradip De, Preetam Ghosh, Zhijun Wang.
WebPage: http://www.cse.uta.edu/pico@cse/

In Pervasive Computing and Communication, we have successfully:

- proposed and developed a MidFusion middleware architecture to facilitate information fusion in sensor network applications.
- Proposed Bayesian network based scheme that efficiently meets the challenges of heterogeneity of sensors and information fusion.
- Developed a learning based method, LEVeL, to integrate ontologically mapped information exchange between mobile agents.
- In the process of developing a prototype on sensor nodes to evaluate the feasibility of the proposed algorithms.

Information acquired by large number of heterogeneous sensors needs to be integrated in a proactive, intelligent, and situation-aware manner to predict the occurrence of events (including security) in the PSI framework. In this part of the project, we are investigating the applicability of sensors by deploying collaborating software agents that meet the needs of dynamic applications.

Two major challenges for proactive and real time collaboration among agents are (1) heterogeneity of sensors, information representation and granularity and (2) fusion of uncertain, redundant, complementary and time sensitive information from various sensors.

We are investigating the coupling of sensors and associated agents for real time information fusion and decision making in distributed and dynamic applications. The agents cooperate in real-time to make intelligent and informed decisions using Bayesian Network reasoning. We have proposed and demonstrated a learning based approach called LEVeL to effectively measure the confidence in cooperating agent observations.

We have also developed MidFusion, an adaptive middleware architecture to facilitate information fusion in sensor network applications. MidFusion discovers and selects the best set of sensors or sensor agents on behalf of applications (transparently), depending on the quality of service (QoS) guarantees and the cost of information acquisition. A sensor selection algorithm (SSA) for selecting the best set of sensors is being designed, and its implementation on a real sensor network is being prototyped [1][2][3]. We plan to complete the following in the next phase of the project:

- Incorporation of multi-agent based collaborative and distributed information fusion and decision making..
- Development of collaborating agent community framework that allows heterogeneous agents to interoperate, collaborate and perform decision making about application goals.
- Improvization of the MidFusion mechanism to address issues of efficient computation and scalability.
- Development of suboptimal sensor selection algorithms for real-time applications by using context information and middleware techniques.

Enhancing Availability of Grid Computational Services to Ubiquitous Computing Applications

The Grid is an integrated infrastructure that can play the dual roles of a coordinated resource consumer as well as a donator in distributed computing environments. In a mobile grid environment, the Grid acts as a resource hungry consumer whereas in a ubiquitous computing environment, it has the inherent potential to provide services to applications. The enormous growth in the use of mobile and embedded devices in ubiquitous computing environment and their interaction with human beings, produces a huge amount of data that need to be processed efficiently anytime anywhere. However, such devices often have limited resources in terms of CPU, storage, battery power and communication bandwidth. Thus, there is a need to transfer ubiquitous computing application services to more powerful computational resources.

In this research [6], we have investigated the use of the Grid as a candidate for provisioning computational services to applications in ubiquitous computing environments. We have developed a competitive model that describes the possible interaction between the competing resources in the Grid Infrastructure as service providers and ubiquitous applications as subscribers. The competition takes place in terms of quality of service (QoS) and cost offered by different Grid Service Providers. The ubiquitous users' resource demands depend not only on the QoS parameters (e.g., response time and loss probability i.e., loss of ubiquitous users experienced by a grid service provider in our context) offered by that service provider but also upon those of its competitors. We develop a stochastic equilibrium model for QoS assurance. Based upon Simulation results show that our proposed framework maximizes the provider's expected revenue at the optimal point of different QoS parameters.

Publications

[1] H. Alex, M. Kumar, B. Shirazi, Collaborating Agent Communities for Information Fusion and Decision Making, International Conference on Knowledge Integration and Multi Agent Systems, Boston, USA, April 2005.

[2] H. Alex, M. Kumar, B. Shirazi, MidFusion: Middleware for Information Fusion in Sensor Network Applications, International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Melbourne, Australia, December 2004.

[3] H. Alex, M. Kumar and B. Shirazi, MidFusion: An Adaptive Middleware for Information Fusion in Sensor Network Applications, Information Fusion Journal, Special Issue on Information Fusion in Distributed Sensor Networks. (under final review)

[4] H. Shen, M. Kumar, S. K. Das, and Z. Wang, "Energy-Efficient Data Caching and Prefetching for Mobile Devices Based on Utility," *ACM/Springer Journal on Mobile Networks and Applications* (Special Issue on Mobile Services, Guest Eds: Q. H. Mahmoud and U. Varshney), Vol. 10, pp. 475-486, 2005.

[5] J. Cao, X. Feng, J. Lu, H. Chan, and S. K. Das, "Reliable Message Delivery for Mobile Agents: Push or Pull?", *IEEE Transactions on Systems, Man and Cybernetics, Part A*, Vol. 34, No. 5, pp. 577-587, Sept 2004.

[6] N. Roy, S. K. Das, K. Basu, M. Kumar, "Enhancing Availability of Grid Computational Services to Ubiquitous Computing Applications," *Proceedings of 19th IEEE International Parallel and Distributed Processing Symposium* (IPDPS), Denver, Colorado, Apr 2005.

[7] N. Banerjee, A. Acharya and S. K. Das, "SIP-based Mobility Architecture for Next Generation Wireless Networks," *Proceedings of 3rd IEEE Symposium on Pervasive Computing and Communications* (PerCom), Kauai Island, Hawaii, Mar 2005.

[8] P. De, K. Basu and S. K. Das, "An RFID Based Framework for Object Distribution, Tracking and Recalls in Pervasive Transaction Environments," *Proceedings of the First IEEE Interational Conference on Mobile Ad hoc and Sensor Systems* (MASS), Ft. Lauderdale, Florida, pp. 349-358, Oct 2004.

[9] P. De, K. Basu and S. K. Das, "An Ubiquitous Architectural Framework and Protocol for Object Tracking using RFID Tags," *Proceedings of MobiQuitous Networking Conference*, Boston, pp. 174-182, Aug 2004.

[10] P. Ghosh, N. Roy, S. K. Das, and K. Basu, "A Game Theory Based Pricing Strategy for Job Allocation in Mobile Grids," *Proceedings of International Parallel and Distributed Processing Symposium*, Santa Fe, New Mexico, Apr 2004.

[11] P. Ghosh, N. Roy, S. K. Das, and K. Basu, "A Pricing Strategy for Job Allocation in Mobile Grids Using a Non-cooperative Bargaining Theory Framework," *Journal of Parallel and Distributed Computing* (Special Issue on Design and Performance of Networks for Super-, Cluster-, and Grid-Computing, Guest Eds: A. Zomaya, M. Ould-Khaoua and H. Sarbazi-Azad), to appear, 2005.

[12] S. K. Das, M. Kumar and B. Shirazi, "Pervasive Community Computing (PCC) Paradigm with Application to Healthcare," *Proceedings of the 1st International Conference on Ubiquitous Computing*, Seoul, South Korea, pp. 3-8, Oct 2003.

[13] Z. Wang, S. K. Das, H. Che, and M. Kumar, "A Scalable Asynchronous Cache Consistency Scheme (SACCS) for Mobile Environments," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 15, No. 11, Nov. 2004.

3 Mobile AdHoc Networks

Faculty Involved : Mukhesh Singhal, Raphael Finkel, Sajal Das
Student Members : Abhishek Roy, Habib Amari, Wook Choi
WebPage:

Ad hoc networks, which do not rely on any infrastructure such as access points or base station, can be deployed rapidly and inexpensively even in situations with geographical or time constraints. In ad hoc networks, each node acts both as a router and as a host. The topology of an ad hoc network may change dynamically, which makes it difficult to design an efficient routing protocol. Nowadays, more and more wireless devices are being used, which can form large ad hoc networks. It is important to design scalable routing protocol for ad hoc networks. In this project, we developed an Anchor Based Routing Protocol (ABRP), a scalable routing protocol for ad hoc networks. It is a hybrid routing protocol, which combines table-based routing strategy with geographic routing strategy. We simulated different size of networks from 200 nodes to 1600 nodes. Simulation results show that ABRP is efficient and scales well to large networks. ABRP combines the advantages of multi-path routing strategy and geographic routing strategy—efficiency and scalability, and avoids the burden—GPS support.

In mobile ad hoc networks (MANETs), nodes depend on each other for routing and forwarding packets. Cooperation among nodes is a key issue in such an environment. However, some of the nodes in MANETs may behave selfishly and may not forward packets to save battery and other resources. Since nodes in MANETs communicate on a peer-to-peer basis, without any central authority (which can monitor selfish behavior of nodes), a centralized solution to stimulate cooperation is not suitable. We developed a fully distributed solution for nodes' cooperation in MANETS. The solution is light-weight in that neighbor monitoring is on-demand, and nodes work in promiscuous mode only part time to save battery. Besides, we provide fairness to nodes in the network by considering their battery status. We incorporated our light-weight solution for selfish node problems considering battery status in AODV. To evaluate the performance of our scheme, we conducted simulation using Glomosim. Simulation results show that the proposed scheme improves data forwarding capability substantially in presence of selfish nodes, identifies the selfish nodes accurately and has low overhead.

Ad-hoc on-demand routing protocols like AODV establish and maintain routes on-demand. However, the paths established by on-demand protocols, which are optimal during route establishment phase, become sub-optimal over time due to node mobility. In this project, we developed a Path Compression Algorithm (PCA) that optimizes the established routes when feasible without incuring extra overhead. By utilizing promiscuous mode of operation, nodes 'hear' hop counts embedded in IP-optional header of data packets to find shorter paths. PCA avoids unnecessary aggressive and ephemeral route updates, which improves the protocol performance. Routing path optimality is important as optimal paths reduce the packet drop ratio, end-to-end delay, and reduces energy dissipation of end-to-end data transmission. We conducted simulations using GloMoSim to evaluate the performance of our algorithm and compared it with AODV and a related protocol in the literature - SHORT. Our simulation results show that PCA achieves higher delivery rate, lower control overhead, lower end-to-end delay and lower average hop count compared to both AODV and SHORT.

Bluetooth is an emerging wireless networking technology for short range applications. The basic unit in a Bluetooth network is piconet, which can contain up to 8 active nodes. Two or more piconets may connect with each other to form the so-called scatternet. Scatternet is indispensable for the future of Bluetooth technology. Security protocols for Bluetooth piconets, such as key management and authentication have been defined by the Bluetooth Specification. However, little work has been done on security of Bluetooth scatternets. In this project, we investigated secure key establishment in Bluetooth Piconets. We developed a link key establishment protocol for Bluetooth scatternets by extending the piconet key agreement protocol to scatternets. The proposed protocol has the same security level as that of the piconet key establishment protocol and is compatible with the Bluetooth Specification.

Education activities included involving graduate students in the research project and developing curricula that incorporated the most recent research results. Specifically the project funded four Ph.D. students as Research Assistants. These students actively participated in the project and developed and evaluated the techniques [6][7][8].

Publications

[1] A. Roy and S. K. Das, "QM^2RP: A QoS-Based Mobile Multicast Routing Protocol using Multi-Objective Genetic Algorithm", *ACM/Kluwer Wireless Networks Journal* (Special Issue on Networking 2002 Selected Papers, Guest Eds: M. Conti and E. Gregori), Vol. 10, No. 3, pp. 271-286, May 2004.

[2] Huaizhi Li and M. Singhal, "A Scalable Routing Protocol for Ad Hoc Networks", to appear in the Proc. of IEEE 61st Semiannual Vehicular Technology Conference, (Mobile Networks), May 30 - June 1, 2005, Stockholm, Sweden.

[3] Huaizhi Li and M. Singhal, "A key establishment protocol for bluetooth scatternets", 3rd International Workshop on Mobile Distributed Computing, (In conjunction with the 25th Int'l Conf. on Distributed Computing Systems) Columbus, OH, June 2005.

[4] Venkata C. Giruka and Mukesh Singhal, "Optimized and Stable path On-demand Routing Protocols for Mobile Ad-hoc Networks". Submitted to IEEE Transactions on Mobile Computing.

[5] Venkata C. Giruka, Mukesh Singhal and Siva Prasad Yarravarapu "A Path Compression Algorithm for On-demand Routing protocols for Ad-Hoc networks", IEEE International Conference on Mobile Ad-hoc and Sensor Systems, Fort Lauderdale, Florida. October 2004.

[6] W. Choi, S. K. Das, J. Cao and A. K. Datta, "Randomized Dynamic Route Maintenance for Adaptive Routing in Multihop Mobile Ad hoc Networks," *Journal of Parallel and Distributed Computing* (Special Issue on Sensor, Ad hoc Wireless and Peer-to-Peer Networks, Eds. J. Wu and M. Cardei), Vol. 65, No. 2, pp. 107-123, Feb. 2005.

[7] W. Choi and S. K. Das, "Design and Analysis of a Proxy-Based Indirect Routing Scheme in Ad hoc Wireless Networks," *ACM Mobile Networking and Applications Journal* (Special Issue on Mobile

Ad-hoc Networks, Guest Eds: A. Campbell, M. Conti and S. Giordano), Vol. 8, No. 5, pp. 499-515, Oct 2003.

[8] W. Choi and S. K. Das, “Nomadic Control Packet-Based Dynamic Route Maintenance Scheme for Adaptive Routing in Mobile Ad Hoc Networks,” *Proceedings of Annual IEEE Conference on Local Computer Networks* (LCN), Bonn, Germany, pp. 140-149, Oct 2003.

[9] Yongwei Wang and Mukesh Singhal, “A Light-weight Solution for Selfish Nodes Problem Considering Battery Status in Wireless Ad-Hoc Networks”, to appearin IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob2005), August 22-24, 2005, Montreal, Canada.

4 Machine Learning, Data Mining and Computational Intelligence

Faculty Involved : Diane Cook, Lawrence Holder
Student Members : A. Rakhshan, C. Noble, I. Jonyer, J. Coble, J. Kukluk, J. Potts, M. Mukherjee, R. Rathi
WebPage: A.I Lab http://ailab.uta.edu/

The ability to learn concepts from relational data has become a crucial challenge in many security-related domains. For example, the U.S. House and Senate Intelligence Committees’ report on their inquiry into the activities of the intelligence community before and after the September 11, 2001 terrorist attacks revealed the necessity for ”connecting the dots”, that is, focusing on the relationships between entities in the data, rather than merely on an entity’s attributes. The ability to discover relationship-driven patterns can impact our ability to prevent future attacks and ensure national security. A natural representation for relational information is a graph, and the ability to discover previously-unknown patterns in such information could lead to a significant improvement in our ability to identify potential threats.

This research has investigated approaches to improve the scalability and effectiveness of the Subdue [1][2][3][4][5][6][7][8][9][10][11][12] graph-based relational learning system. We have also extended Subdue to perform not only unsupervised discovery of patterns from graph data, but also supervised learning from graph-based examples, hierarchical clustering,and graph grammar learning from graph data.

Unsupervised Substructure Discovery

Subdue accepts as input directed or undirected graphs with labeled vertices (nodes) and edges (links). As an unsupervised discovery algorithm, Subdue searches for a substructure, or subgraph of the input graph, that best compresses the input graph. The search terminates upon reaching a limit on the number of substructures extended (this defaults to half the number of edges in the graph), or upon exhaustion of the search space. Once the search terminates and Subdue returns the list of best substructures, the graph can be compressed using the best substructure. The compression procedure replaces all instances of the substructure in the input graph by single vertices, which represent the substructure definition. Incoming and outgoing edges to and from the replaced instances will point to or originate from the new vertex that represents the instance. The Subdue algorithm can be invoked again on this compressed graph. This procedure can be repeated multiple times, and is referred to as an iteration.

To allow slight variations between instances of a discovered pattern, Subdue applies an inexact graph match between the substructure definition and potential instances. The discovery algorithm can be biased by incorporating prior knowledge in the form of predefined substructures or preference weights on desirable (undesirable) edge and vertex labels. Subdue’s run time is polynomial in the size of the input graph. Substructure discovery using Subdue has yielded expert-evaluated significant results in domains including predictive toxicology, network intrusion detection, earthquake analysis, web structure mining, and protein data analysis.

Subdue can also be used to learn concepts that distinguish examples of different classes. New examples that contain the discovered substructures are classified as positive examples, otherwise they are classified as negative examples.

Issues Related to Learning from Counter-Terrorism Data

As part of the U.S. Air Force program on Evidence Assessment, Grouping, Linking and Evaluation (EAGLE), a domain has been built to simulate the evidence available about terrorist groups and their plans prior to their execution. The simulator was developed based on feedback from intelligence analysts and several studies with respect to appropriate signal-to-noise ratios. The data we use represents the activities of terrorist organizations as they attempt to exploit vulnerable targets and consists of a number of concepts, including threat and non-threat actors, threat and non-threat groups, targets, exploitation modes (vulnerability modes are exploited by threat groups, productivity modes are exploited by threat and non-threat groups), capabilities, resources, communications, visits to targets, and transfer of resources between actors, groups and targets.

The domain follows a general plan of starting a group, recruiting members with needed capabilities, acquiring needed resources, visiting a target, and then exploiting the target. These events involve various forms of communication and transfer of assets. The EAGLE simulator generates various threat and non-threat groups, and then executes various vulnerability and productivity exploitations. The simulator generates evidence related to all these events, and this evidence is passed through filters varying the degree of observability and noise in the final evidence. This final evidence is the data from which we are to learn.

We have already achieved some success in identifying patterns and classifying threat groups in the EAGLE data. However, researching a complex domain such as counter-terrorism has raised a number of issues that need to be addressed, some of which cannot be effectively addressed using current techniques.

The first challenge is handling structural data that is received in incremental blocks. Instead of reprocessing the accumulated graph after each increment, we provide summary statistics for each increment and process the new block independently. The result in quality of discovered results is not affected because we explicitly look for pattern instances that grow across increment boundaries. However, the run time of Subdue is significantly improved over the batch processing version found in the original Subdue algorithm.

Second, we have explored a new representation for supervised learning that allows instances of all classes to be embedded in one connected graph. The examples themselves are of varied size and can overlap. We have shown with NASA data and with EAGLE data that distinguishing concepts can be learned with this representation and that the accuracy of the result is superior to the case where examples are extracted from the single graph.

Finally, we are currently investigating methods of scaling the entire algorithm using stochastic approaches. While the first step is to explore stochastic approximation techniques for graph isomorphism and subgraph isomorphism, we are also considering applying stochastic approaches to the task of discovering an entire set of interesting patterns in the original structural data.

Publications

[1] A. Rakhshan, L. B. Holder and D. J. Cook, Structural Web Search Engine, *International Journal on Artificial Intelligence Tools*, 13(1), pages 27-44, 2004.

[2] I. Jonyer, L. B. Holder and D. J. Cook, MDL-Based Context-Free Graph Grammar Induction and Applications, *International Journal on Artificial Intelligence Tools*, 13(1), pages 45-64, 2004.

[3] I. Jonyer, L. B. Holder, and D. J. Cook, Attribute-Value Selection Based on the Minimum Description Length, *Proceedings of the International Conference on Artificial Intelligence*, 2004.

[4] J. Coble, D. J. Cook, R. Rathi, and L. B. Holder, Structure Discovery from Sequential Data, *Proceedings of the Florida Artificial Intelligence Research Symposium*, 2004.

[5] J. Coble, D. J. Cook, R. Rathi, and L. B. Holder, Structure Discovery from Sequential Data, *International Journal of Artificial Intelligence Techniques*, 14(1-2), 2005.

[6] J. Kukluk, L. B. Holder, and D. J. Cook, Algorithm and Experiments in Testing Planar Graphs for Isomorphism, *Journal of Graph Algorithms and Applications* (To Appear).

[7] J. Potts, D. J. Cook, and L. B. Holder, Learning from Examples in a Single Graph, *Proceedings of the Florida Artificial Intelligence Research Symposium*, 2005.

[8] J. Potts, L. B. Holder, D. J. Cook, and J. Coble, Learning Concepts from Intelligence Data Embedded in a Supervised Graph, *Proceedings of the International Conference on Intelligence Analysis*, 2005.

[9] J. Coble and D. J. Cook, Structure Discovery in Sequentially Connected Data, *Proceedings of the Florida Artificial Intelligence Research Symposium*, 2005.

[10] L.B. Holder, D. Cook, J. Coble and M. Mukherjee, Graph-based Relational Learning with Application to Security, *Fundamenta Informaticae Special Issue on Mining Graphs, Trees and Sequences* (To Appear).

[11] L. B. Holder and D. J. Cook, Graph-based Data Mining, in J. Wang (ed.), *Encyclopedia of Data Warehousing and Mining*, Idea Group Publishing, 2005.

[12] R. Rathi and D. J. Cook, A Serial Partitioning Approach to Scaling Graph-Based Knowledge Discovery, *Proceedings of the Florida Artificial Intelligence Research Symposium*, 2005.

5 InfoFilter: A Content-Based Information Filtering System

Faculty Involved : Sharma Chakravarthy
Student Members : *WebPage*: http://itlab.uta.edu/
Information filtering includes monitoring text streams to detect patterns that are more complex than those handled by search engines. Text stream monitoring and pattern detection have far reaching applications such as tracking information flow among terrorist outfits, web parental control, and business intelligence. Pattern characterization requirements of the applications entail an expressive language for specifying patterns than what is currently provided by Information Retrieval Query Languages (IRQLs) and current information filtering systems. In essence, pattern specification and its detection play a major role in information filtering.

User Specification

In InfoFilter, users can specify simple and composite patterns using the Pattern Specification Language, PSL. It supports the following operators and options: frequency, synonyms, sequence, Boolean operators, structural, wild card, and proximity. Furthermore, any arbitrary complex pattern can be composed using the above operators. Some of the operators in PSL have some similarities with event specification languages (e.g., [1]) used for the specification of events. Even though some of the operators are similar, semantics of pattern operators are different as it includes the notion of proximity, which is crucial in information filtering. In addition, PSL supports pattern operators such as regular expressions, frequency and synonyms. On the other hand, PSL allows composition of all the operators specified above.

Pattern Types

In PSL, a pattern $\boldsymbol{P}$ is formally represented as P_i^j, where i is the pattern identifier and j is the instance of the pattern identifier. A pattern $\boldsymbol{P}$ is a function that maps from the offset interval domain onto the boolean values, "True/False" corresponding to the occurrence or non-occurrence of the pattern. O_s is the start offset, and O_e is the end offset of the pattern, where offset is the position of the word relative to the beginning of the text stream. For example, for a phrase "user pattern", if "user" occurs at 50 and "pattern" occurs at 51, then O_s is 50 and O_e is 51.

According to the semantics of PSL, patterns are classified as:

Simple patterns: Simple patterns are the basic building blocks and can be either *System-defined* (i.e., built into the system), or *User-defined*. System-defined patterns are pre-defined and they correspond to structural elements present in the text streams, such as the beginning of a sentence, a paragraph, or a document/stream. For example, two system-defined patterns *BeginPara* and *EndPara* are used to define the beginning and end of a paragraph. On the other hand, possible user-defined patterns include a single word or any of its synonyms (e.g., *"filter"*), multi-word or phrases (e.g., *"information filtering system"*), or simple regular expressions (e.g., *"filter*"*).

Complex Patterns: Complex patterns are composed of simple patterns, complex patterns, pattern operators and options. PSL provides a comprehensive set of pattern operators and they are: Boolean (OR, NOT, NEAR), sequence (FOLLOWED BY), structure (WITHIN), frequency (FREQUENCY), proximity (NEAR/N, FOLLOWED BY/N) and the option synonym (SYN).

Operators for Pattern Specification

Semantics of PSL operators and options are explained below (formal definitions for these operators and options are not provided in this report due to lack of space). Please refer to [1] for details.

OR: Disjunction of two simple or complex patterns P_1 and P_2, denoted by (P_1 *OR* P_2), occurs when either P_1 or P_2 occurs. For example, *"information" OR "filtering"* will be detected whenever either one of the keywords occurs. Since simultaneous occurrences of the same patterns are not possible in a stream (essentially a sequence), exclusive OR semantics is used.

NOT: Non-occurrence of the simple or composite pattern P_2 in the range formed by the end offset of P_1 and the start offset of P_3, where P_1 and P_3 can also be simple or composite patterns, is denoted by *(NOT [/F](P_2)(P_1, P_3))*. "*F*" indicates the minimum number of occurrences and its default value is 1. For example, *NOT ("filtering")("information", "retrieval")* will be detected whenever *"information"* is followed by *"retrieval"* without the word *"filtering"* occurring at least once in between them.

NEAR: Conjunction of two simple or composite patterns P_1 and P_2, denoted by *(P_1 NEAR [/D] P_2)*, occurs when both P_1 and P_2 occur, irrespective of their order of occurrence. "*D*" is the maximum distance allowed between the two patterns P_1 and P_2. Default value of "*D*" is the scope of the operator (which can be the entire document), and it refers to the *AND* operator of the Boolean model. The minimum value of D is 1. For example, *"information" NEAR/10 "filtering"* will be detected whenever both these words co-occur within a distance of 10.

FOLLOWED BY: Sequence of two simple or composite patterns P_1 and P_2, denoted by *(P_1 FOLLOWED BY [/D] P_2)*, occurs when the occurrence of P_1 is followed by the occurrence of P_2. The end offset of P_1 is less than the start offset of P_2; that is, the occurrence interval of P_1 and P_2 should not overlap. "*D*" is the maximum distance allowed between the two patterns P_1 and P_2. If "*D*" is not specified, the distance is bounded by the scope of the operator (can be the entire document). If the value of "*D*" is 1 (minimum value), this indicates that the patterns P_1 and P_2 form a phrase. For example, *"information" FOLLOWED BY /10 "filtering"* will be detected whenever the word "*information*" precedes "*filtering*" within a distance of 10 words.

WITHIN: Occurrence of a simple or composite pattern P_2 in the range formed by the end offset of the pattern P_1 and the start offset of P_3, denoted by *(P_2 WITHIN (P_1, P_3))*. The pattern is detected each time pattern P_2 occurs in the range defined by patterns P_1 and P_3. For example, *"information filtering" WITHIN (BeginPara, EndPara)* will be detected whenever the phrase *"information filtering"* occurs within a paragraph. When an expression is specified without a system-defined pattern, the default structure (e.g., a document, a web page) is used as the default. This operator is crucial while expressing the scope of the stream being processed.

FREQUENCY: Multiple occurrences of a simple or composite pattern that exceeds or equal to *F*, denoted by *(FREQUENCY /[F] (P))*. A pattern *P* is detected each time *P* occurs at least F times, where "*F*" is the minimum number of occurrences specified by the user. The first *F* occurrences are considered for this detection as the stream/document is processed sequentially. The default value of *F* is 1, which is the minimum value. Multiple occurrences are assumed to be disjoint as instances for a pattern detection are used only once. The same instance is not used for detecting multiple patterns. For example, *FREQUENCY/10 ("information filtering")* will be detected whenever the phrase *"information filtering*" occurs at least 10 times. Frequency can be applied to any pattern expression; the composition of NEAR and FREQUENCY allows the specification of very expressive patterns.

SYN: This is an option and is specified along with a single-word pattern (currently), denoted by *(P[SYN])*, to indicate multiple single-word patterns that have the same meaning, in a succinct manner. In Psnoop, specifying a single-word pattern with SYN option is equivalent to specifying *N* simple patterns that carry the same meaning (synonyms) as the original pattern. For example, if you specify the word *"bomb"[SYN]* is equivalent to specifying *"bomb" OR "explosives device" OR "weaponry" OR "arms" OR "implements of war" OR "weapons system" OR "munition"* . If any of these words or phrases appears in the text, the pattern *"bomb"[SYN]* is detected. This option adds simplicity and flexibility to the specification of single-word patterns. The same is true for composite patterns with embedded synonym specification, e.g. *"missile"[SYN] NEAR "Iraq"*.

Using the above operators, users can specify complex and meaningful patterns. For instance, a complex pattern *("bomb" occurring prior to "ground zero" for at least two times, with a single occurrence of "automotive" or its synonyms)*, can be specified using the PSL as

P_1 = ("bomb" FOLLOWED BY "ground zero")

P_2 = FREQUENCY /2 (P_1)

P_3 = P_2 NEAR "automotive" [SYN]

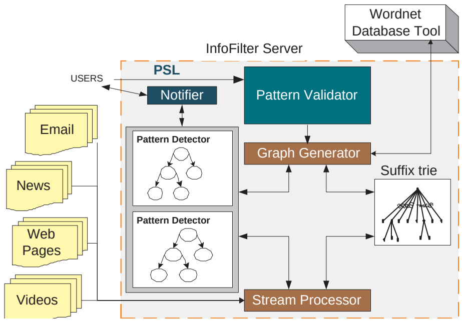

Figure 1: Architecture of InfoFilter

6 Implementation

InfoFilter analyzes text streams based on the content and structural information, and notifies users when their patterns of interest are detected. The system is based on the client/server architecture, wherein a group of clients register with a server to specify patterns of interest. Figure 1 shows the various modules of the InfoFilter server: Pattern Validator, Graph Generator, Stream Processor, Pattern Detector, and other external components. The patterns can be associated with different types of text streams (e.g., documents, web pages, and video captions). User patterns are handled by certain modules and incoming streams are handled by other modules.

The InfoFilter server has been implemented as a stand-alone Java application. The server consists of Pattern Validator, Graph Generator, Stream Processor, Pattern Detector, Notifier, and Suffix trie. The pattern validator, a JavaCC parser, accepts user patterns in the form of a linear text using PSL BNF (see appendix), such as *("Information" FOLLOWED BY "Filtering")*. It parses, validates, and tokenizes the patterns based on the syntax of the PSL language. Then, it places the tokens into a stack in postfix order and passes it to the graph generator.

The graph generator reads the postfix notation and constructs the PDG. It interacts with the WordNet database tool to extract synonyms of words. It uses Java WordNet Library (JWNL)[6] to provide this interaction. The extracted words, synonyms, phrases, and regular expressions are stored in a shared suffix trie. Suffix tries are constructed over all the suffixes of the text. They are characterized as space efficient and have less search time for small text [2][3][5]. Pattern detector is a Java library that provides the APIs necessary to construct the Pattern Detection Graphs (PDGs) that detects complex patterns.

Stream processor is used to handle various incoming text streams and is implemented using several parsers. For handling unstructured text, a parser has been implemented using the BreakIterator Class in Java. The BreakIterator Class can detect text boundaries such as word boundaries, sentence boundaries, and so forth.The Java Mail API has been used to parse emails. The parser should extract the textual information present in the emails, which is then tokenized using the BreakIterator. All the tokens generated by these parsers are then matched against the extracted keywords, phrases, and regular expressions stored in the suffix tries. If there is a match, the stream processor notifies the pattern detector, which alerts the notifier. The notifier is implemented using Java Mail APIs [?] to send messages to users subscribing to the system.

Conclusions and Future Work

In this report, we have presented the InfoFilter system, a content-based system for filtering text streams. InfoFilter has been developed with an intent to support expressive user patterns using PSL and to provide filtering on streams and notification. We have implemented the InfoFilter server consisting of Pattern Validator, Pattern Detector, Graph Generator, and Stream Processor. PSL, proposed in this report, with its expressiveness and well-defined semantics (see [1] for details), overcomes the limitations of the current information systems used for specifying and detecting user patterns. It provides a complete set of pattern operators and options such as *frequency*, *synonyms*, *followed by*, *Boolean operators*, *structural*, *wild card*, and *proximity*.

We are currently extending the InfoFilter in a number of ways. PSL is being extended to support complex regular expressions, and synonyms for phrases. Higher level specification of patterns that can be converted into PSL is another direction. InfoFilter can be linked with the web monitoring systems to filter web contents in a selective manner. The system can also be extended to search for patterns in an entire web and for the generation of web ontology based on the patterns detected in web pages in a web server. Patterns can be detected over other types of streams such as XML can be supported by modifying the stream processor and pattern validator modules of the InfoFilter architecture. Handling of partial pattern matching can be explored further. In the long term, this work can be extended to incorporate adaptive filtering, and discovering patterns that can be of an interest to the user.

Publications

[1] L. Elkhalifa, "InfoFilter: Complex Pattern Specifications and Detection over Text Streams", Master's thesis, Department of Computer Science and Engineering, The University of Texas at Arlington, 2004. Online at http://itlab.uta.edu/ITLABWEB/Students/sharma.

[2] M. Nelson, "Fast String Searching with Suffix Trees", Dr. Dobb's Journal, August 1996.

[3] R. Baeza-Yates and B. Ribeiro-Neto, "Modern Information Retrieval", New York: ACM Press/Addison-Wesley, 1999.

[4] S. Chakravarthy, V. Krishnaprasad, E. Anwar and S. K. Kim, "Composite Events for Active Databases: Semantics, Contexts and Detection", 1994, pp 606-617.

[5] S. Wu and U. Manber, "Fast Text Searching Allowing Errors", October 1992.

[6] "JWNL (Java Wordnet Library)". Online at http://sourceforge.net/projects/jwordnet

[7] "Sun MicroSystems: JavaMail API Specifications v1.3.1", 2003.

7 Mobile Database

Faculty Involved : Ali Hurson
Student Members : A. Tangpong, B. Yang, J. Ploskonka, J. Ghaznavi, J. Sustersic, P. Yan, X. Gao, Y. Jiao.
WebPage:

Research In MAMDAS

After witnessing the success of many mobile agent applications, we have proposed and implemented a new infrastructure called MAMDAS Mobile Agents within the framework of Mobile Data Access System. MAMDAS combines the merits of SSM (Summary Schemas Model) and mobile agents in building a distributed large-scale information access systems. It aims to achieve higher performance, while providing special support for mobile users. Our experimental results have shown that MAMDAS is 6 times faster than the client-server based SSM prototype, because of the reduced network traffic.

Adaptive Power Management for Mobile Agent-Based Information Retrieval

In this project, we proposed and simulated an adaptive application-driven power management (AADPM) protocol for mobile agent-based applications within the IEEE 802.11b infrastructure WLAN environment. Our goal was to minimize energy while achieving low round trip time (RTT) delay. The protocol was evaluated using the network simulator NS2. We also draw horizontal comparisons among a variety of power management methods reported in the literature. Experimental results showed that, compared to the power save mode supported by 802.11b, AADPM reduces the network interface card energy consumption by 65

Security in MAMDAS

Security can be viewed from different angles. One way to define a secure system is that a system can guarantee information confidentiality, authentication, integrity, and non-repudiation. Another way of looking at security is to protect data and services of a system against security threats. As noted in the literature, there are four types of security threats: interception, interruption, modification, and fabrication. Security can also be viewed based on the underlying application domain. In the application domain that we are interested in, mobile agent-based applications, three key security issues can be identified: protect the agent against malicious hosts, protect the host against malicious agents, and protect the network communication. Albert Einstein once said, Problems cannot be solved at the same level of awareness that created them. Simply trying to identify existing security threats and devising a countermeasure for each one of them is not the way to build a secure system, because remedies of current problems may not protect the system against emerging new threats in the future. Moreover, we cannot build a secure system by just piling security solution pieces together, because security strategies often need to collaborate in order to protect the system. In addition, resources may be wasted due to the overlapping of security strategies. A better approach to tackle this problem is to separate security policy from security mechanisms. Security policies precisely describe which actions the entities in a system are allowed to take and which ones are prohibited; security mechanisms are means by which a security policy can be enforced. Conventional notions of security mechanisms include authentication, access control (authorization), and auditing. We believe that they are not sufficient for user protection, in that database access control is designed to protect data from unauthorized use and therefore, only protects the system. A mounting concern is the potential privacy loss on the users side due to the exposure of their queries. This concern has stimulated a new branch of security mechanism research Private Information Retrieval (PIR), and we proposed to extend the conventional notions of security mechanisms to include PIR. Following the same design philosophy of mechanism-policy separation, we believe that the security architecture of a mobile agent-based application should include two parts: security mechanisms provided by the underlying agent platform and security policies suitable for that particular application. In this project, we proposed a security architecture for MAMDAS and discuss how it can be realized using existing technology. Our goal is to allow flexible policies to be enforced by mechanisms selected from a large repertoire.

Automated Integration of Multimedia Semanteme Using Summary Schemas

Content-based multimedia access has mostly focused on feature-based representation rather than semantic description of the multimedia data. However, human perception of multimedia objects is often based on semantic contents. Given that there is a need for automated extraction algorithms of semantic contents, this work intends to develop a method that allows content-based representation, organization, and retrieval of the multimedia data in a heterogeneous distributed database platform. The novelty of this framework comes from (1) its mathematical capability to represent semantic contents, (2) the hierarchical organization and classification of multimedia data objects according to their semantic contents, and (3) the ease of nearest-neighbor searching through synonym links. Using the summary schemas model, as our heterogeneous database infrastructure, the proposed scheme is intended to be simulated and compared against several other proposals that have advanced in the literature. Achievements:

We developed a scheme for semantic-based image content representation in distributed heterogeneous database environment. The summary-schemas model (SSM) is used as the underlying platform. With the ability of summarizing the content information and guiding the data distribution, the SSM provides a quality-guaranteed and time-efficient image accessing strategy. The advantages of representing and indexing images on the platform of SSM are as follows: (1) The SSM has mathematical capability of representing semantic content precisely and concisely; (2) Content-based queries are efficiently resolved through hierarchical organization and classification of image objects according to their contents; and (3) Imprecise text descriptions can be included in queries, which further enhances the searching capability of image databases. Conceptual framework of the proposed model is defined. The proposed scheme has been simulated, and the results are analyzed. Mobile ad hoc networks have gained more and more research attentions by provisions of wireless communications without location limitations and pre-built fixed infrastructure. Because of the absence of any static support structure, ad hoc networks are prone to several limitations such as bandwidth, connectivity, and power. Multimedia retrieval is a challenging task in wireless ad hoc networks because of the multiple limitations. We showed that data content distribution can be employed to facilitate content-based multimedia retrieval in ad hoc networks. Motivated by the data organization methodology of the Summary Schemas Model (SSM), we proposed a logic-based content summary framework that is able to represent semantic contents of multimedia data using concise logic terms. Furthermore, a virtual infrastructure was built to cluster mobile nodes according to semantic contents. The proposed framework has been simulated, and the simulation results are analyzed based on various performance metrics.

Publications

[1] A. R. Hurson and Y. Jiao,, Data Broadcasting in a Mobile Environment, Wireless Information Highway, 2005, chapter IV, pp. 96-154.

[2] A. R. Hurson and B. Yang, Multimedia Content Representation Technologies, Encyclopedia of Multimedia Technology and Networking, 2005.

[3] A. R. Hurson and Y. Jiao, Database System Architecture A Walk Through Time: From Centralized Platform to Mobile Computing. In Proc of the 5th IEEE International Symposium and School on Advance Distributed Systems (ISSADS05), Guadalajara, Mexico, 2005.

[4] B. Sun, A.R. Hurson, and J. Hannan, Energy-Efficient Scheduling Algorithms of Object Retrieval on Indexed Parallel Broadcast Channels, International Conference on Parallel Processing, 2004, pp. 440-447.

[5] B. Yang and A.R. Hurson, Adaptive Content-Aware Multimedia Clustering in Wireless Ad Hoc Networks, International Conference on Advances in Mobile Multimedia, 2004, pp. 35-44.

[6] B. Yang and A.R. Hurson, An Extendible Semantic-Based Content Representation Method for Distributed Image Databases, IEEE Sixth International Symposium on Multimedia Software Engineering, 2004, pp. 222-226.

[7] B. Yang and A.R. Hurson, Ad Hoc Image Retrieval Using Hierarchical Semantic-Based Index, IEEE AINA2005 International Conference, 2005, pp. 629-634 (best paper).

[8] B. Yang and A.R. Hurson, A Content-Aware Multimedia Accessing Model in Ad Hoc Networks, 11th International Conference on Parallel and Distributed Systems, 2005.

[9] B. Yang and A.R. Hurson, Semantic-Based Clustering of Multimedia Data in Ad Hoc Networks, IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, 2005.

[10] J. Sustersic and A.R. Hurson, Quality of Service (QoS) in Internet Cache Coherence, International Journal of High Performance Computing and Networking, 2005.

[11] X. Gao and A.R. Hurson, Location Dependent Query Proxy, Twentieth Annual ACM Symposium on Applied Computing, 2005, pp. 1120-1124.

[12] Y. Jiao and A.R. Hurson, Adaptive Power Management for Mobile Agent-Based Information Retrieval, IEEE AINA2005 International Conference, 2005, pp. 675-680.

[13] Y. Jiao and A.R. Hurson, Application-Driven Dynamic Wireless Network Power Management, invited paper, International Conference on Advances in Computer Science and Technology, 2004, pp. 291-296.

[14] Y. Jiao and A.R. Hurson, Performance Analysis of Mobile Agents in Mobile Database Access Systems A Quantitative Case Study, Journal of Interconnection Networks, Volume 5, Number 3, 2004, pp. 351-372.

8 PSI TestBed at CReWMaN

At the end of the second year of PSI project, we are creating a high quality testbed that would complement and enhance research in PSI. Figure 2 is an illustration of the various components of the testbed architecture that we are in the process of building at CReWMaN lab.

To build skills for developing application specific algorithms and protocols for sensor networks using Tiny-OS and nesC under Berkeley mote platform, a small experiment on Intrusion Detection and Data Aggregation using Multi-hop Routing is being carried out on the prototype architecture.The modus-operandi is as follows:
Vibration and/or acoustic sensors are deployed near the entrance of a large room and only acoustic sensors inside a surveillance area. When an object tries to enter the secure area, the vibration sensors detect the presence of the object and activates the acoustic sensors that are inside. Once the acoustic sensors are activated they start sensing any acoustic signals generated by the intruder object and routes their data via multi-hop to the aggregator/base station. The object might carry an RFID tag. The aggregator aggregates the values it received from all the nodes and checks for the presence of a valid RFID tag. If it does not find a valid tag or a blacl-listed tag is detected, the system raises an alarm.
This experiment marks the first step twards prototype development. We are taking an incremental iterative approach where experiments across different research domains are going to be merged for the final PSI prototype.

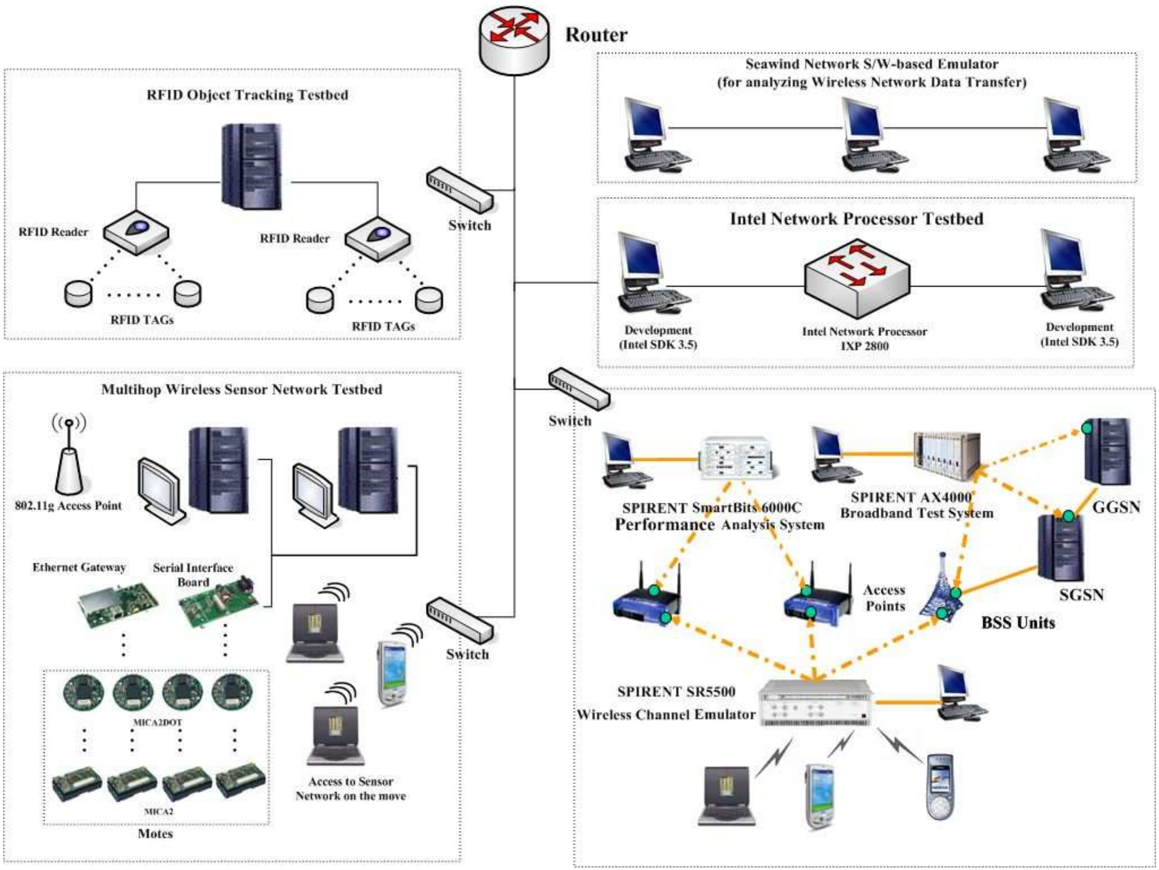

Figure 2: PSI TestBed at CReWMaN (currently under development)

9 Direct Impact of PSI

Project Coordination and Interaction

The PSI project has fostered excellent working relation and collaboration among the PIs and Co-PIs from UTA, UKY and PSU. There is intense exchange of ideas between the various personnel so as to develop an understanding of the mechanism and architecture on how to transform useful research into meaningful working prototype.
At UTA, there are regular monthly face-to-face meetings with students and faculty working in various project groups and subgroups. At the recently concluded second annual workshop on PSI project at UTA, all PIs and their students got together and made technical presentations on research findings, shared ideas, and chalked out a plan that would lead to further collaborative research and enhance synergy.

Project Outcome - To Date

The PSI project has resulted in three PhD theses from UTA and one from PSU. At the same time, six Masters theses were awarded from UTA and one from PSU. To date, there are around fifteen PhD theses being pursued at UTA and four each at UKY and PSU. Two undergrad students have trained through REU supplements at UTA. Two new graduate level courses on network security have been introduced at UTA and UKY with the idea of disseminating and encouraging research among students. The research findings were presented at conferences, workshops and also published as articles and book chapters.

Project Leverage

Two new projects emerged out of this ITR project. One is called SafetyNet project for border security including perimeter control, airport/harbor security. Another is on bulding high assurance video sensor networks in highly uncertain dynamic environments. They were submitted to NSF and Federal Earmark funding.
An NSF MRI proposal got funded in 2004 that help procure equipment needed to develop a prototype of the reseach proposed in this ITR project. This ITR project has also resulted in a spin-off in Bioinformatics and System Biology. This has further led to collaboration with UT Southwestern Medical center in Dallas.