Sensor: The Atomic Computing Particle

Vijay Kumar
Guest Editor
University of Missouri-Kansas City

Abstract: We visualize the world as a fully connected information space where each object communicates with all other objects without any temporal and geographical constraints. We can model this fully connected space using fine granularity processing which can be implemented using sensors technology. We regard sensors as atomic computing particles which can deployed to geographical locations for capturing and processing data of their surrounding. This report introduces a number of excellent research articles which present unique problems and their success in finding efficient solutions for them. It also peeks intothe future of ever changing information processing discipline.

1. Introduction

We visualize the space we live in as a *fully connected information space*. Each object, which could be mobile or immobile, of this space is fully and continuously connected to all other objects. This connectivity allows each object to communicate with all other objects without any temporal and spatial constraints. Figure 1 illustrates the information space we envision.

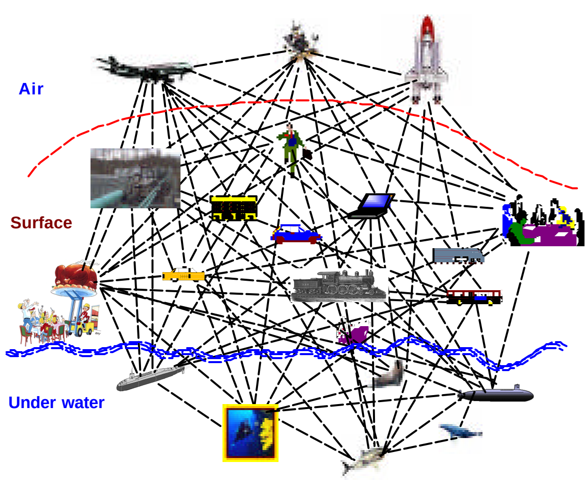

Figure 1. A Fully Connected Information Space.

The unlimited pervasiveness of the information space, although links every objects, creates some serious and complex problems related to the capture of information from geographical locations not easily reachable by humans such as ocean bed, enemy territories, deep space, and so on. The medical field also has similar difficulty in collecting data of internal live organs such as liver, heart, etc. The pervasiveness aspect of this space cannot be satisfactorily implemented and managed by conventional approaches such as distributed systems and the state of the art mobile system because of their highly coarse granularity of processing. Researchers have come to realize that a finer granularity processing discipline and a unified approach is necessary to manage the information space we desire.

2. Atomic Computing

The concept of atomic computing defines fine processing granularity which allows us to embed processing units at geographical locations to create a fully connected network of countably infinite nodes. With this framework it is possible to visualize an instance of fully connected information space. Every node of this network has data processing capability (data capture, data validation, and data dispatch) and functional autonomously without going out of sync with other nodes of the network. We envision that each node can be reprogrammed locally or remotely to append necessary capability anytime and the network topology can be made adaptable to the environment where it is deployed [1]. We strongly believe that the current state of sensor technology can provide us our fully connected information space. We try to establish this claim in subsequent sections.

3. Sensor Technology

A sensor is a programmable, low-cost, low-power, multi-functional device and usually has a much shorter working life span [1]. Essentially it is an *atomic processing unit* which is highly suitable for working as a node of distributed peer-to-peer or client server architecture. We refer to this as a *sensor node* or *S-node* which can be programmed to capture and validate the data of its environment, to perform some operation and dispatch the result to other nodes or to a designated server. In addition to this, it can also receive data from other S-nodes and can be queried by any peer or client of the network. With these unique properties it makes sense to embed S-nodes to any objects of this information space for monitoring, controlling, and logging their activities. For example, programmed sensors may be embedded at various points in cars of a family, in the house, in the office, in children's school bags, in parents' briefcases, etc. Each S-node will capture data and communicate with other S-nodes, which will help the parents to be aware of and be fully connected with everyday activities of their children. At the time of need any family members can be reached instantly. On a large scale we can visualize sensor deployment at various places (buildings, malls, factories, high rise building, banks, etc.) of a city for continuous monitoring of events to manage security [4, 10, 12]. Similarly to protect water supply, gas pipelines, and so on, programmed sensors can be deployed at strategic locations. We observe that in medical field two types of sensor (a) immersive and (b) non-immersive were developed and deployed to study functions of external and internal human organs for diagnostic purpose. This kind of use of S-nodes, of course raises a number of complex ethical and moral issues but we do not address them here.

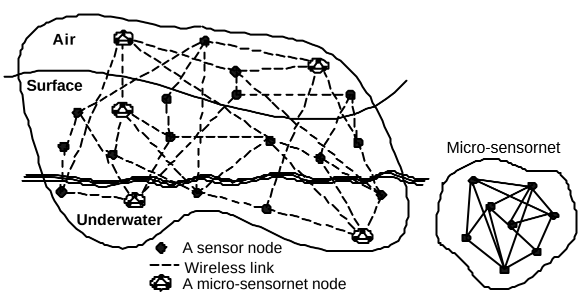

Figure 7. An ESS with Micro-sensornet.

We define the concept of *Embedded Sensor Space (ESS)*, which is a countably infinite set of uniquely programmed network of S-nodes. Thus, $ESS = \{s_1, s_2, ..., s_\infty\}$ where s_i $(i = 1, 2, ..., \infty)$ are programmed S-nodes and s_i and s_j $(i \neq j)$ are fully connected and have direct communication facility. A S-node of ESS captures data of its environment and dispatches it to other S-nodes through routers [2, 3, 6, 7, 8, 9, 11].

We also introduce the concept of *micro-sensor* which is network of a small number of specialized fully connected S-nodes of ESS. One of the nodes in a micro-sensor net is responsible for coordinating the activities of other sensors in the set. This gives rise to the problem of *leader election* problem, which we do not discuss in this report. To the best of our knowledge we are not aware of any earlier work on our ESS with micro-sensor net.

Figure 7 illustrates a reference architecture of our ESS. The ESS pervades air (captures the space), surface and underwater world. The purpose of micro-sensor net is to modularize the entire reference structure for developing security schemes to handle Denial of Services (DoS) attack, authentication, certification, and so on. For example, if a S-node of an ESS wants to communicate with multiple S-nodes of a micro-sensor net, then instead of getting certification from these S-nodes, it can get certification only from the leader of the micro-sensor net.

We observe that ESS has a universal scope for managing information processing, communication, and information dissemination and because of this it has received tremendous attention from media, researchers, industry and common people. It is, therefore, not surprising that academic community is promoting this technology through conferences and workshops, industry is deploying it to solve problems in the area of security, communication, monitoring, etc., and the common people are enjoying the fancy gadgets for managing their every day activities.

We are investigating the use of ESS for a number of activities. Figure 8 illustrates one of our research projects which deal with the development of ESS to monitor gas emission at various landfills and send information to the DBMS for further processing and dissemination [5]. The ESS we envision is very large and highly data intensive since sensors will be capturing data continuously. The emission of gases and their volume are not predictable so these sensors will have to be active continuously and there must also have a fail-safe scheme, which would ensure that any sensor failure or malfunctioning is promptly propagated to the servers for immediate action to minimize the damage. This implies that they will be sending different types of data with different constraints associated with them. Some data will be temporal in nature with limited validity. These must be processed in real-time and decision, if any, must be propagated to target sensors for changing or altering their functionality. The diversity of data category, their real-time characteristics, propagation of results to target sensors present a number of complex data management problems.

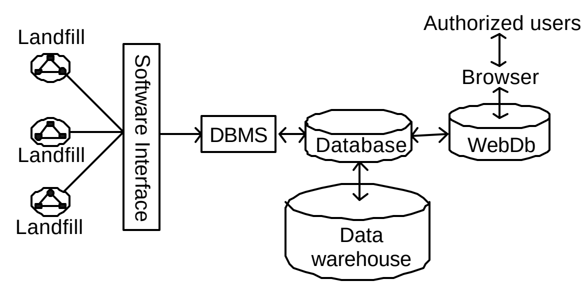

Figure 8. Landfill monitoring through sensornet

The ESS, in addition to unique problems of its own, will inherit all existing problems in the areas of data processing, data communication, network routing, security, query processing, database recovery and transaction management, etc. Unfortunately, existing solutions to these problems will not work satisfactorily with any amount of tweaking. It needs innovative solutions with much higher adaptability and scope. They should be able to work with millions of S-nodes and micro-sensor nets, they must be highly secured, they must be energy efficient, they must not fail or if they fail, their recovery must not affect system performance and they must offer a high degree of QoS.

4. Sensor Data Mining

S-nodes of an ESS have its own semantics, that is, each S-node may have special functionality and its own way of dealing with data. This characteristic of S-node creates a *sensorbase* similar to database where sensor mining may become inevitable to mine a set of sensors with special characteristics. For example, we may have a search request *Find sensors and their geographical locations which capture temperature data*. We believe that there are similarities in sensorbase and database but there are significant differences too. We are investigating if data mining approaches can be applied in sensorbase mining.

5. An Introduction to the Special Issues

We opted to complement the efforts of academic community for promoting the basic and applied research in sensor technology through two special issues of SIGMOD record. These issues are devoted to research in sensor system infrastructure, security, and data processing (data capture, validation and consistency preservation). It aims to unify multi-discipline highly complementary sensor technology research and development through the unique link, which is clearly visible when seen but invisible when not seen. The research scope of sensor technology is vast consisting of a large number of problems. It is not possible to include all of them and one has to be selective to be in sync with the interest and priority of research direction. We identified the following areas (a) Sensor network security, (b) Stream data processing and information management, (c) Sensor mobility and communication, (d) Sensor data capture and validation, (e) Sensor deployment, and (f) Sensor system management for SIGMOD record. We received 25 high quality papers in most of these areas out of which we selected 17 for publication and decided to disseminate these papers in two consecutive issues. We selected 10 papers related to application domain to appear in December 2003 issue and the remaining 7 will appear in March 2004 issue.

6. An Overview of Articles

The December 2003 issue papers effectively presented the state of sensor technology and its application in solving some of the important real life problems which could not be achieved easily through conventional systems. An excellent introduction to sensor technology capability was given by Joseph Hellerstein and his group. They illustrated that among all sensing technology, sensor has a special place because of its pervasiveness and low-cost features. They started from massive sensing equipments such as satellite and gracefully converged to tiny and cheap devices, which I have called as Atomic Computing Particles (ACP), touching many other devices on the way. It is true that every discipline needs database support and this paper highlighted what sensor technology has to include for managing its data and what database technology has to do to provide this service. One of the challenges offered by sensor technology to database technology was related to stream data processing and the database community is frantically trying to develop efficient solutions. In addition to this the paper gave a lot of research and development quality stuff to the academic community to ponder over.

Murali Mani presented his inspiring thoughts on how to deal with the data captured by sensors where they are deployed. He argued that sensors capture raw data in a continuous fashion probably with some real-time constraints and it must be stored in a database as *this* sensor's data. It seems that the database is then sensor specific and application processing is also sensor specific. A database must have query facility, therefore, in this paper Mani presented his approach for querying sensor specific database with a query language and his brand of query processor.

Philippe Bonnet and his group integrated Bluetooth technology to develop a sensor network. They argued about the usefulness of these atomic computing particles (sensors) when they are mobile in a small geographical domain. In their paper they demonstrated two things (a) how Bluetooth impacts the functionality of mobile sensor network and (b) how the application software interacts with such a mini wireless sensor network. The conducted some experiment with Bluetooth-Based Sensor Network and observed very useful information related to energy consumption and network topology.

The paper by Raynold Cheng and Sunil Prabhakar presented their observations on uncertainty inherent in random movement of sensor while capturing data. They looked at the difficult question of maintaining accuracy in query processing for randomly changing data values and proposed the idea of *Probabilistic Threshold Query* to keep the inaccuracy to minimum in dealing with data related to highly mobile objects.

Their future work is to investigate how other types of queries can be dealt with under their approach.

Anastassia Ailamaki and the group presented their work on the application of sensor technology to protect drinking water and keep it safe for general consumption. EPA (Environment Protection Agency) is quite serious about threats aimed to pollute drinking water. They proposed to enhance the capability of sensor technology by eliminating or minimizing their inherent limitations. They looked into the complex problem of optimal sensor deployment at strategic places for providing security (making threats ineffective) and monitoring the pollution level of drinking water and keeping it safe to drink. Optimal data distribution problem is still lingering around. Hopefully we have a decent solution through sensor approach.

Alan Demers and the group introduced their Cougar Project which is investigating energy-efficient data dissemination and query processing schemes. Under this project the group is investigating a number of related issues such as data transmission scheduling, view selection for query processing, power consumption in data processing activities, synchronization among sensor nodes of sensor network, and so on. Their future work is to related to fault-tolerance of their system.

Energy consumption has been a serious problem in atomic computing. Rajgopal Kannan and his team argued that the problem becomes more complex because sensor networks are typically unattended since they are deployed in hazardous, hostile and remote locations not easily reachable by humans. This kind of setup requires a high degree of intelligence for saving energy. The team presented a MAC protocol for sensor network and showed its high degree of energy saving properties through simulation.

Hüseyin Özgür Tan and Ìbrahim Körpeoglu presented their work on power efficient data gathering and aggregation in wireless sensor network. They assumed a wireless network of sensors where sensors gather data about their environment and dispatch it to a base station for further processing. Their objective was to maximize the system lifetime by energy conservation per communication round. They proposed two new algorithms and through simulation result they indicated that their schemes are near optimal.

Mario Gerla and Kaixin showed that through mobile *swarm* they developed high quality multimedia streaming can be supported in large sensor network. The motivation was that sensors nodes with their limited capability are unable to support multimedia streams, which are necessary for acquiring detail information about a geographical location. They also outlined their network and routing schemes for supporting their swarms.

Themistoklis Palpanas and the group attacked the problem of data collection of a disaster, natural or manmade, for its timely control to minimize damage. The control of such events, such as chemical spill, gas pipeline burst, poisonous gas leak, and so on, requires that necessary information must be collected and processed in under hard-real time constraints. This makes data collection and dissemination very difficult. They presented their scheme for a highly distributed infrastructure.

The following seven papers mostly related to system and architecture issues will be published in March 2004 issue of SIGMOD record.

The paper of Malik Tubaishat and the group presented a hierarchical model for an ad-hoc sensor network. They emphasized on minimizing energy consumption in group communication and security. They designed a secured routing protocol called *Secure Routing protocol for Sensor Network (SRPSN)* which is supposed to guarantee the correctness of query result.

Yongge Wang presented work on security. It was argued that in some situation it was impractical for a sensor node to contact the base station for establishing a secured communication channel. The algorithm presented in this paper eliminated the need of base station for secured communication between any two mobile sensor nodes.

The challenge of communication among unmanned mobile vehicles was discussed in the paper of Mario Gerla and Yunjung. They utilized the idea of *team multicast* for establishing communication among mobile vehicles which created a scenario of ad-hoc network. The team studied the performance and reported the benefits of their scheme compared to traditional multicast approach.

Iosif Lazaridis and the team addressed the issue of quality (data and result) in sensor network. They reported their ongoing work on QUASAR (Quality Aware Sensing Architecture) and discussed the future of sensor-based data architecture.

The paper of Nam Hun Park and Won Suk Lee proposed a statistical grid-based data clustering scheme for sensor data. They observe that the data stream is unbounded and could be of enormous volume consequently many algorithms compromise the correctness for minimizing processing time. Their algorithm seems to overcome some of the problems which they illustrate by experimental analysis.

Takahiro Hara, Norishige and Shojiro Nishio presented the extension of their work on data replication in ad hoc sensor network. In this extended work they considered data correlation for improving their replication algorithm. They proposed three replica allocation approaches and illustrated their merits through simulation experiment.

Amol Despande reported some study on a highly query processing operator called Eddy for data stream in sensor network environment. In the paper the implementation of eddies architecture in the PostgresSQL is reported with initial experimental study to study its overheads.

A sensor can send data in any form, i.e., pulses, packets, sound, etc. There are a large number of data capture schemes are available [15, 18, 19]. We are in the process of developing, which will take into consideration energy conservation aspect. This stream data must be converted before it can be stored and processed. Since sensors in our network could be location specific we propose to program the L/L information in the sensor, which will be appended to each dispatch from that sensor. We agree that this will increase the cost but the benefit will outweigh the cost. This conversion can be easily done by a simple mapping function, since the type of the dispatch from a sensor will be known in advance. A simple conversion table, residing in the interface, will be satisfactory for the conversion.

S-nodes in a ESS are insecure repositories and routers of data. There are many applications where sensors are deployed in hazardous environments in which they are subject to failure or destruction under attack by a malicious adversary. For example, consider seismic sensor networks in earthquake or rubble zones or sensors in military battlegrounds under enemy threat. Wireless sensor networks are also extremely vulnerable to data loss under denial of service (DoS) attacks. Nodes use wireless communication because the network's large scale, ad-hoc deployment and limited energy resources makes wired or satellite communication impractical. Jamming a transmitting nodes frequency makes its data unavailable. Thus, any model for ensuring effective query reporting and collaborative mining in sensor networks, while incorporating the constraints of energy efficiency and distributed decision-making, should simultaneously take sensor failure and security considerations into account. This will require the development of specific algorithms to ensure that the tasks of (a) data storage and content in distributed repositories (which could be special sensor nodes within the network) and (b) data retrieval are not affected by the inhospitable environment.

7. Conclusions and Future Directions

This report with the help and accepted papers looked into the world of sensor technology and how it is affecting the world of information management. These papers discussed problems unique to sensor technology and presented possible solutions to some of the problems. It appears that the future of information management will be taken over by finer granularity processing paradigm and at present sensor technology is the only approach in site for implementing such processing discipline. S-nodes will pervade every object, live and dead, mobile and immobile, and visible and invisible and control information in and out of the object. Sensors have been extensively embedded in many objects such as watches, televisions, cell phones, toys, cars, etc., but the concept of sensornet was not perceived. It is here now and becoming an integral part of our life. The future will bring more powerful and smaller size sensors, which will sit in every geographical points of this world and quietly do the job they are designed for.

Acknowledgements

Thanks to Rajgopal Kannan, S.S. Iyengar and Dharma Agrawal for their valuable help in composing this article.

References

1. Estrin, R. Govindan and J. Heidemann, "Embedding the Internet", ACM Comm., Vol. 43, No. 4, May 2000.
2. Alberto Cerpa and Deborah Estrin, "Ascent: Adaptive Self-Configuring sensor Network Topologies," *Technical Report* UCLA/CSD-TR-01-0009" Computer Science Department, UCLA.
3. T. Clouqueur, P. Ramanathan, K. K. Saluja, and K-C. Wang, "Value-fusion versus decision-fusion for fault-tolerance in collaborative target detection in sensor networks," *Proceedings of the Fourth International Conference on Information Fusion*, August 2001.
4. H. Deng, W. Li, and Dharma P. Agrawal, "Routing Security in Ad Hoc Networks," *IEEE Communications Magazine*, Special Topics on Security in Telecommunication Networks, Vol. 40, No. 10, October 2002.
5. EPA publication, "Nonhazardous waste landfill publications," *CR-ROM*, EPA530-C-00-002.
6. Soheil Ghiasi, Ankur Srivastava, Xiaojian Yang, and Majid Sarrafzadeh, "Optimal Energy Aware

Clustering in Sensor Network," *Sensors*, Vol. 2, pp. 258-269 July 2002.

7. W. B. Heinzelman and A.P. Chandrakasan and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications*, Vol. 1, No. 4, pp. 660-670, October 2002.

8. Chalermek Intanagonwiwat and Ramesh Govindan and Deborah Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM)*, pp. 56-67, August 2000

9. Stephanie Lindsey, Cauligi Raghvendra, and Krishna Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics," *IEEE Trans. on Parrallel and Distributed Sys*. 13, No. 9, September 2002.

10. Perrig, P. Szewczyk et at., SPINS: Security Protocols for Sensor Networks, *Wireless Networks* 8, 521-53, 2002.

11. R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I - fundamentals," *Proceedings of the IEEE*, Vol. 85, No. 1, pp. 54-63, January 1997.

12. D. Wood and J. A. Stankovic, "Denial of Service in Sensor Networks," *IEEE Computer*, 35(10), Oct. 02.