# Using Geospatial Information in Sensor Networks*

John Heidemann
*USC/Information Sciences Institute*

September 17, 2001

## Abstract

This paper describes several ways sensor networks can benefit from geospatial information and identifies two research directions. First, better models of localization error, logical location, and communications costs are required to understand the interactions between spatial information and control and communications algorithms in sensor networks. Second, wider use of spatial information in densely deployed sensor networks will move sensor networking applications from simple tracking to object counting and area monitoring, and can enable the use of data mining techniques to sensor networks for "spatial sensor mining".

## 1 Introduction

Recent developments in inexpensive, short-range wireless communication and sensors have built on inexpensive computers to enable *sensor networks*: collections of small devices spatially distributed around an environment. Sensor networks are being applied in areas including environmental monitoring, conditional based maintenance, surveillance, computer augmented or smart spaces, and inventory tracking; we expect many other applications to appear as the technology becomes more widely available.

Unlike traditional sensor systems, sensor networks depend on *dense sensor deployment* and *physical co-location* with their targets to accomplish their goals. *Dense deployment* implies the use of 100s or 1000s of sensor nodes in small areas and is enabled by low-cost devices and short-range wireless communication. With *physical co-location*, sensors are tightly coupled with their environment: they may be attached to packages being tracked, or deployed a few meters apart to cover an intersection or field. Physical co-location simplifies signal processing problems [15]. Dense deployment allows use of redundancy [20], can reduce communication costs [15], and provides sufficient nodes to allow physical co-location.

*This work was supported by DARPA under grant DABT63-99-1-0011 as part of the SCAADS project.

The author can be contacted by e-mail at *johnh@isi.edu*.

*Spatial location* is central to sensor network operation. The purpose of sensor networks is often to answer spatial queries such as "what is moving down the road and how fast is it?" or "how many animals are in the northwest field?". Sensor networks also make use of spatial information to simplify their problems. Sensor deployment requires localization to determine the quality of coverage [14, 6] and to constrain communications to geographic areas [13, 11, 22]. Collaborative signal processing techniques such as beamforming [21] and information-based approaches (for example, [4, 10]) to combine the results of multiple sensors to provide a whole greater than the sum of the parts. At an operational level, spatial information can be used to conserve energy by load balancing (for example, [8, 20]) and to control network utilization [7].

Although sensor networks today use spatial information as part of current applications and operations, additional research is required to systematize these gains and enable new applications. We suggest two directions of research: first, *better models of spatial information* in sensor networks are needed. We must understand the error patterns of current localization systems and how location interacts with sensing and communication. Second, we suggest that *spatial sensor mining* is a promising direction of research. Long-term data collection from densely deployed, spatially distributed sensors can enable fine-grained trend analysis if we can cope with data collection and processing constraints. We describe these problems in more detail below.

## 2 Modeling Spatial Information in Sensor Networks

Much recent research has focused on how to determine node location in a range of circumstances. In some cases GPS may be used, but other solutions are required for indoor use or for cases where the power or size requirements GPS are excessive (examples include [18, 19, 1, 5, 16]). To build on this promising research we must next understand the error patterns of current localization systems and how location interacts with sensing and communication.

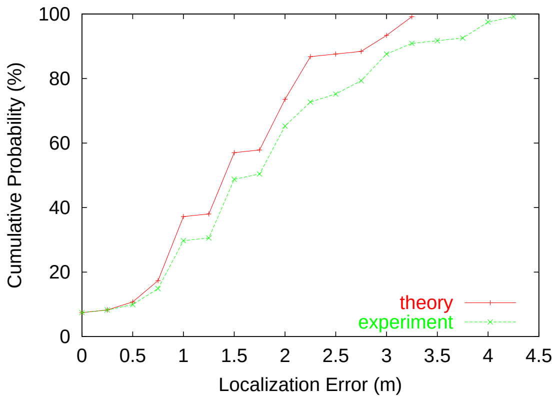


Figure 1: Cumulative distribution of localization error, both experimental and theoretically predicted, in a simple proximity-based localization system (Figure 5 from [5]).

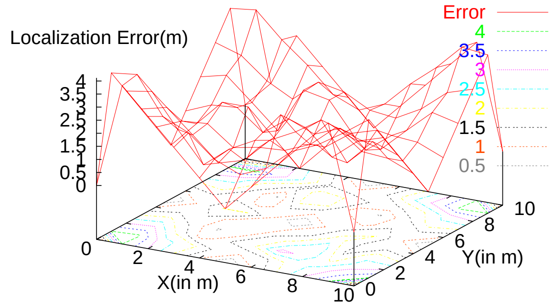


Figure 2: Localization error as a function of position in a simple proximity-based localization system (Figure 4 from [5]).

Existing localization systems have quantified localization error. Figure 1 shows the probability of a given level of localization error in a simple proximity-based localization system, both from experiment and as predicted by a simple spherical radio propagation model. Similar studies have been done for most other localization systems (for example, GPS [9]). Although statistical studies of localization error exist, simple statistical models are not sufficient to understand the impact of localization error on applications that depend on spatial information. Is the error independent across all nodes, or do two nodes in close proximity exhibit positively or negatively correlated error? For GPS, we know that closely placed nodes tend to exhibit correlated absolute error (hence the ability to do differential GPS), and so their relative positions are quite good even though absolute positions my be off by 10–100m [9]. Other localization systems exhibit very different patterns of error. For example, error in the system described earlier is strongly dependent on node location (as shown in Figure 2), so nearby nodes may observe very different levels of absolute position error and hence large errors in their relative positions. Good error models for a range of localization systems is necessary to understand the effects of error on algorithms that build on spatial information.

Although a localization system may provide the physical location of a node, a better understanding is needed of node *logical* locations: not just where the user or sensor is, but is affected or observed. For example, condsider the "museum guide" context-aware application. In addition to where the user is located, it needs to know where he or she faces to describe the correct artwork. In a sensor network, node location may be a point, but seismic or acoustic sensor coverage may be a much larger disk or cone depending on the type of sensor and the details of its physical placement. Algorithm correctness can be guaranteed by assuming maximum sensor range and correcting after-the-fact, but with large sensor ranges and dense deployment this approach may be very costly. We need better models of sensor coverage, and better techniques for propagating this information through applications.

Finally, there is considerable latitude to optimize communication based on spatial information. If you ask for information to the west, there is no need to propagate that query to the east (examples using this principle include ad hoc routing [13], directed diffusion [11, 22] and collaborative signal processing [10]). However, one must consider communications overheads when evaluating these costs. For example, geographic-assisted routing must be robust to "holes" in the topology that produce local routing minima, and the cost of adjusting and readjusting a geographically scoped query must be weighed against the overhead of an excessively large query or the penalty of an inaccurate answer. Better models of protocol overheads are needed to understand the benefits of protocols and algorithms that use spatial information to optimize communication.

# 3 Spatial Sensor Mining

In addition to a better understanding of spatial information, dense and long-term sensor deployment allows new applications. Challenges here include moving from sensor network tracking of one or a small number of targets to *monitoring an area with many objects* (distributed in space) and *spatial sensor mining*: drawing conclusions from information gathered over time.

Current uses of sensor networks focus on *object tracking*: using a group of sensor nodes to identify an object such as a vehicle. More sophisticated versions of this problem consider tracking multiple objects moving together, or objects that meet and disperse. Better integration of sensing and location allow examination of properties of areas as a whole and groups of objects rather than individual objects. Rather than focusing many sensors on a single target, the challenge is how to focus many sensors on many targets. Algorithms include determining object population (counting) and density, object flows or movement trends, with applications to the environment (wildlife tracking), society (tracking crowds of people in buildings), and the military. Early developments here includes sensor tomography (sensor network self-monitoring) [23], and distributed database techniques for sensor networks [2, 3].

More challenging still is the goal of monitoring sensor fields over time. If collaborative signal processing uses multiple sensors to track a target or area at a point in time, *spatial sensor mining* is the use of sensors to track an area over a long duration, combining data from both the spatial and temporal domains.

We illustrate spatial sensor mining with two examples. First, consider a rapidly deployed field of visual or infrared sensors that will operate for about a week. The first day of readings can accumulate data about how the sensor net changes as a function of time-of-day. Although the broad arc of variation is dominated by the time, individual sensors will behave differently given placement (for example, while all will experience daylight at the basically the same time, some may be in the shadow of a hill or tree for part of the day). This baseline can then be used on subsequent days to judge whether changes are appropriate or abnormal.

Second, consider a wired sensor network deployed for longer terms (months or years), perhaps monitoring utility systems such as water or power [17], or a fixed region. Such a system is unlike traditional sensor networks in that some elements may be wired and powered, but it shares the goals of dense sensor deployment, spatially distributed nature, and physical co-location with the sensed targets. It also benefits from sensor network approaches such as data diffusion that allow easy deployment and reconfiguration and data-centric operation. A primary constraint in such systems is communications bandwidth: clearly sensors placed on every meter of pipe cannot report flow rates every few seconds to a central site. Instead, techniques for distributed data processing and mining are needed. Approaches such as filters to allow in-network-processing [12] and sensor network tomography [23] offer promise to minimize communications by aggregating information. These approaches must be combined with capabilities for "drill-down" on unusual phenomena, and with appropriate data mining and analysis techniques.

Spatial sensor mining in some ways is similar to condition based maintenance: the general approach is to use long-term monitoring to detect problems. The approaches differ, however, in terms of sensor density and amount of pre-planning. Condition based maintenance may employ one or a few sensors per monitored object, and sensors are deployed with hand-configured code to watch that specific object. Spatial sensor mining instead considers sensors densely deployed in the environment, but perhaps not attached to specific equipment or pre-programmed to detect specific conditions. Instead, sensors will collaborate to monitor objects, and automated techniques from data mining may be used to detect long-term trends and anomalies.

# 4 Conclusions

Sensor networks depend on spatial information; their physically distributed nature provides advantages in energy efficiency and signal processing compared to centralized systems. Current sensor networks too often depend on ad hoc or non-existing models of localization, logical location, and communications costs. Better models are required in each of these areas to achieve best operation. Moreover, better integration between spatial and sensor information is necessary for sensor networks to move from simply tracking to counting and monitoring areas, and approaches such as spatial sensor mining suggest a role for ad hoc sensor networks in long-term data analysis and problem detection.

Although there is great industry interest in computing in the "post-PC" world, most of industry remains focused on user-centric computing platforms such as personal digital assistants or network computers. Continued research and academic leadership in sensor networks is necessary to explore the much different domain of spatially and physically distributed computers that interact first with the environment and only secondarily (and in the aggregate) with human users.

# Acknowlegements

This work builds upon technical discussions and input from many people: Deborah Estrin and Ramesh Govindan provided input on many areas of this work, Nirupama Bulusu and Lewis Girod on localization, Philippe Bonnet and Johannes Gehrke about database and sensor network interation (although any misstatements about data mining are mine), and Masanobu Shinozuka suggested and discussed applications of sensor networks to utility systems.

## References


[1] Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *Proceedings of the IEEE Infocom*, page xxx, Tel Aviv, Israel, March 2000. IEEE.

[2] Philippe Bonnet, Johannes Gehrke, Tobias Mayr, and Praveen Seshadri. Query processing in a device database system. Technical Report TR99-1775, Cornell University, October 1999.

[3] Philippe Bonnet, Johannes Gehrke, and Praveen Seshadri. Querying the physical world. *IEEE Personal Communications Magazine*, 7(5):10–15, October 2000.

[4] Richard R. Brooks and S. Sidtharama Iyengar. Robust distributed computing and sensing algorithm. *IEEE Computer*, 29(6):53–60, June 1996.

[5] Nirupama Bulusu, John Heidemann, and Deborah Estrin. GPS-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5):28–34, October 2000.

[6] Nirupama Bulusu, John Heidemann, and Deborah Estrin. Adaptive beacon placement. In *Proceedings of the 21st International Conference on Distributed Computing Systems*, pages 489–498, Phoenix, AZ, April 2001.

[7] Alberto Cerpa and Deborah Estrin. Ascent: Adaptive self-configuring sensor network topologies. Technical Report CSD-TR 01-0009, University of California, Los Angeles, Computer Science Department, May 2001.

[8] Jae-Hwan Chang and Leandros Tassiulas. Energy conserving routing in wireless ad-hoc networking. In *Proceedings of the IEEE Infocom*, pages 22–31, Tel Aviv, Israel, March 2000. IEEE.

[9] Dan Charrois. Study of the accuracy of averaged non-differential GPS measurements. Web page `http://www.syz.com/gps/gpsaveraging.html`, February 1999.

[10] Maurice Chu, Horst Haussecker, and Feng Zhao. Scalable information-dirven sensor querying and routing for ad hoc hetereogeneous sensor networks. *International Journal of High Performance Computing Applications*, page (to appear), 2002.

[11] Daniel A. Coffin, Daniel J. Van Hook, Stephen M. McGarry, and Stephen R. Kolek. Declarative ad-hoc sensor networking. In *Proceedings of the SPIE Integrated Command Environments Conference*, San Diego, California, USA, July 2000. SPIE. (part of SPIE International Symposium on Optical Science and Technology).

[12] John Heidemann, Fabio Silva, Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, and Deepak Ganesan. Building efficient wireless sensor networks with low-level naming. In *Proceedings of the Symposium on Operating Systems Principles*, page (to appear), Lake Louise, Banff, Canada, October 2001. ACM.

[13] Brad Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing for wireless networks. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, pages 243–254, Boston, Mass., USA, August 2000. ACM.

[14] Seapahn Meguerdichian, Farinaz Koushanfar, Miodrag Potkonjak, and Mani B. Srivastava. Coverage problems in wireless ad-hoc sensor networks. In *Proceedings of the IEEE Infocom*, page (to appear), Anchorage, Alaska, USA, April 2001. IEEE.

[15] Gregory J. Pottie and William J. Kaiser. Embedding the internet: wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58, May 2000.

[16] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6thACM/IEEE International Conference on Mobile Computing and Networking*, pages 32–43, Boston, MA, USA, August 2000. ACM.

[17] Masanobu Shinozuka. Detection of damage location and extent of water supply systems. In *Proceedings of the 1stWorkshop on Mitigation of Earthquake Disaster by Advanced Technologies*, pages 55–61, Los Angeles, California, USA, March 2000.

[18] Roy Want, Andy Hopper, Veronica Falcao, and Jonathon Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.

[19] Andy Ward, Alan Jones, and Andy Hopper. A new location technique for the active office. *IEEE Personal Communications Magazine*, 4(5):8–15, October 1997.

[20] Ya Xu, John Heidemann, and Deborah Estrin. Geography-informed energy conservation for ad hoc routing. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking*, pages 70–84, Rome, Italy, July 2001. USC/Information Sciences Institute, ACM.

[21] Kung Yao, Ralph E. Hudson, Chris W. Reed, Daching Chen, and Flavio Lorenzelli. Blind beamforming on a randomly distributed sensor array system. *IEEE Journal of Selected Areas in Communication*, 16(8):1555–1567, October 1998.

[22] Yan Yu, Ramesh Govindan, and Deborah Estrin. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical Report TR-01-0023, University of California, Los Angeles, Computer Science Department, 2001.

[23] Yonggang Zhao, Ramesh Govindan, and Deborah Estrin. Residual energy scans for monitoring wireless sensor networks. Technical Report 01-745, University of Southern California Computer Science Department, May 2001.