Introduction to Advances in Geosensor Networks

Silvia Nittel[1], Alexandros Labrinidis[2], and Anthony Stefanidis[3]

[1] Department of Spatial Information Science and Engineering
University of Maine Orono, ME 04469
nittel@spatial.maine.edu
[2] Department of Computer Science
University of Pittsburgh
Pittsburgh, PA 15260
labrinid@cs.pitt.edu
[3] Department of Geography and Geoinformation Sciences
George Mason University
Fairfax, VA 22030
astefani@gmu.edu

Advances in microsensor technology as well as the development of miniaturized computing platforms enable us to scatter numerous untethered sensing devices in hard to reach terrains, and continuously collect geospatial information in never before seen spatial and temporal scales. These geosensor network technologies are revolutionizing the way that geospatial information is collected, analyzed and integrated, with the geospatial content of the information being of fundamental importance. Analysis and event detection in a geosensor network may be performed in real-time by sensor nodes, or off-line in several distributed, in-situ or centralized base stations.

A large variety of novel applications for geosensor networks have already emerged. For example, real-time event detection of toxic gas plumes in open public spaces is crucial for public safety, while monitoring the progression of an oil spill is environmentally valuable. At the same time, more sophisticated applications are becoming feasible for the first time, for example, people using mobile phones equipped with microsensors and short-range communication to collect, exchange, and analyze local information with each other, whether it be about the freshness of produce in a supermarket, or the presence of influenza viruses in the air, or the presence of explosives at an airport. Geosensor networks may find applications in diverse fields, such as environmental monitoring (e.g., habitat observation and preservation, ocean and coastal monitoring), precision agriculture and fisheries, ad-hoc mobile computing for transportation, or surveillance and battlefield situations.

A geosensor network can loosely be defined as a sensor network that monitors phenomena in geographic space. Geographic space can range in scale from the confined environment of a room to the highly complex dynamics of an ecosystem region. Nodes in the network are static or mobile, or attached to mobile objects (e.g., on buses) or used by humans (e.g., cell phones). For example, cameras and GPS sensors on-board static or mobile small-form, mobile or stationary platforms have the ability to provide continuous streams of geospatially-rich information. Today, these types of geosensor networks can range in scale from a few cameras monitoring traffic to thousands of nodes monitoring an entire ecosystem.

S. Nittel, A. Labrinidis, and A. Stefanidis (Eds.): GSN 2006, LNCS 4540, pp. 1–6, 2008.

Over the last 10 years, research efforts have taken place to develop the basic hardware infrastructure for small-scale sensor network systems consisting of large numbers of small, battery-driven sensor nodes that collaborate unattended and are self-organizing on tasks, and communicate via short-range radio frequency with neighboring nodes. Operating system software for these devices is available in the open source domain today, and the basic strategies of deploying sensor networks have a solid foundation.

Many challenges for geosensor networks still exist. The main challenge of sensor networks is to program the sensor nodes as a single, task-oriented computational infrastructure, which is able to produce globally meaningful information from raw local data obtained by individual sensor nodes. Another challenge is to integrate the sensor network platforms with existing, large-scale sensors such are remote sensing instrument or large, stationary ocean buoys, and process the information in real-time using a data streaming paradigm. Overall, the current research problems are centered at data management in general. Data management challenges for geosensor network applications can be divided into the areas of basic data querying, managing, and collection, which is today researched in the database community, and the more advanced, higher-level data modeling and formal data representation issues as well as data integration strategies, which are novel topics in spatial information science.

The spatial aspects of the overall technology are important on several (abstraction) levels of a geosensor network, as the concepts of space, location, topology, and spatio-temporal events are modeled on various abstraction levels. For example, the hardware and communication layers handle the physical space of sensor deployment, and communication topologies. The database layer generates execution plans for spatio-temporal queries that relate to sensor node locations and groups of sensors. Data representation and modeling deals with the relation between collected raw sensor data as fields and phenomena in geographic space.

The papers collected in this volume represent key research areas that are fundamental in order to realize the full potential of the emerging geosensor network paradigm. They cover the spectrum from low-level energy consumption issues at the individual sensor level to the high-level abstraction of events and ontologies or models to recognize and monitor phenomena using geosensor networks. We tried to cluster them across two separate research areas, namely *data acquisition and processing*, and *data analysis* and *integration*. This separation is to a certain extent arbitrary, as most papers address challenges that permeate across different research areas. Additionally, to better illustrate the complexity of transferring research into practice, we have also included three papers representing the diversity of geosensor network applications.

Section 1: Data Acquisition and Processing

The first section of this book is dedicated to papers that deal with *data acquisition and processing* challenges in geosensor networks, mostly on the sensor nodes themselves, or even within the sensor network. Since sensor networks depend on in-network processing to preserve energy and deal with the restricted communication bandwidth between sensor nodes, the challenge exists to come up with intelligent data collection strategies. Another challenge is the actual deployment strategy for sensor nodes, since they are likely scattered in the area of interest, but must be able to achieve network connectivity for all nodes. This also applies for mobile sensor networks, i.e. networks in which one, or more nodes are mobile.

Regarding *sensor deployment*, the paper of Ferentinos, Trigoni, and Nittel is addressing mobile sensor network data acquisition in a turbulent environment, where sensor nodes move involuntarily. The motivating application is ocean current tracking using a network of sensors that are drifting on the sea surface.

Energy management remains a crucial issue in sensor network deployment, with in-network communication being the major energy-consuming activity. Addressing this issue, the paper of Kulik, Tanin, and Umer is presenting a novel algorithm for queries in sub-network structures. Such queries need to reference only parts of a sensor network, thus offering opportunities to optimize data collection paths while minimizing energy consumption. Motivated by the same goal, namely the implementation of energy-efficient query processing in sensor networks, Deligiannakis and Kotidis present a comparative analysis of several data reduction techniques, ranging in scope from the simple monitoring of a node's variance to the identification of spatiotemporal correlations among nodes.

Processing the tsunami of data generated by sensor networks presents a host of well-known challenges that overwhelm classic database management systems, with their store-then-query processing paradigm. Data stream processing systems (DSPS) have been introduced to *handle* substantial volumes of data in a variety of applications, ranging from financial data to environmental monitoring. Typical queries in a DSPS are registered ahead of time; these queries are continuous, constantly being evaluated against a never-ending stream of incoming data to generate output streams. Accordingly, they are highly suitable for the processing requirements of sensor network applications. The paper of Tatbul, Ahmad, Cetintemel, Hwang, Xing, and Zdonik addresses the scalability and high availability aspects of a distributed stream processing system through the Borealis prototype. In particular, the authors discuss how to dynamically modify data and query properties of Borealis without disrupting the system's runtime operation, and how to adapt Borealis to tolerate changes and failures.

With the *current* trend of moving from spatial to spatio-temporal analysis, *mobility* is increasingly being seen as a first-class citizen in sensor networks. Addressing this issue, the paper by Agouris, Gunopulos, Kalogeraki, and Stefanidis introduces a spatio-temporal framework to support object and sensor mobility, inspired by active surveillance applications using optical sensors (video and still cameras). The key challenge they address is how to track moving objects using a small network of sensors that are also mobile (e.g., cameras on-board unmanned aerial vehicles). In the process, the authors touch upon issues related to the modeling of spatiotemporal information (e.g., the movement of a car), the development of similarity metrics to compare spatiotemporal activities (e.g., the movement of two different cars), and the management of a network to optimally track moving objects (e.g., repositioning sensors in order to follow certain activities). Continuing in the same context, the paper by Bakalov and Tsotras presents a novel indexing scheme for streaming spatio-temporal data, and efficient algorithms for evaluating spatio-temporal trajectory join queries, which used to identify a set of objects with similar behavior over a query-specified time interval. This supports queries about previous states of the spatio-temporal stream, provides approximations for object trajectories, and supports incremental query evaluation.

Section 2: Data Analysis and Integration

The second section of this book is dedicated to papers that deal with *data analysis and integration* challenges in geosensor networks, focusing especially on issues like 3D visual analysis, geosensor webs and standardization/interoperability, and higher-level semantic modeling.

Picking up the thread that ended the first section of this book, with camera-based surveillance, we have the paper of Akdere, Cetintemel, Crispell, Jannotti, Mao, and Taubin on visual sensor networks for 3D sensing. It presents the concept of a "virtual view" as a novel query mechanism to mediate access to distributed image data in a video sensor network. The ultimate objective is to establish efficient and effective networks of smart cameras that can process video data in real time, extracting features and 3D geometry in a collaborative manner.

Today, many traditional, larger-scale sensor field stations are already in place. With ubiquitous wireless communication networks, data can be retrieved via satellite links in real-time. Thus, the overall data collection paradigm in sensor data management is changing to a real-time scenario, and ultimately sensor networks will be integrated with larger-scale field stations and remote sensing imagery. Calling this scenario the *Geosensor Web*, we envision that access to sensor data will be as uniform and easy as access to data on the World Wide Web today. However, several problems exist that have to be addressed first before enabling such a vision. The article by Agrawal, Ferhatosmanoglu, Niu, Bedford, and Li focuses on the first challenge for sensor data integration, i.e. a framework comprising of real-time data streams from live sensors, a *stream-based middleware* for on-the-fly sensor data integration and analysis, linking it with ontologies and stored domain knowledge. Their driving application area is coastal forecasting and change analysis.

Viewing practical interoperable sensor data integration, industry-supported standards for access and sensor data presentation protocols are key. The paper by Botts, Percivall, Reed, and Davidson presents the Open Geospatial Consortium's (OGC) standard with regard to the standardized architecture for Sensor Web Enablement. Today, this standard is used for interoperable access to remote sensing instruments; however, it will become a highly important, enabling mechanism for the overall Geosensor Web. Many open research problems can still be found in the area of real-time sensor data integration such as novel meta data, access rights, copyright/privacy issues, uniform scale representation, scalability, and others.

Looking at the flood of collected and integrated real-time sensor data, it becomes clear that the cognitive aspects of users must be addressed and that higher-level, semantically rich data representation models and query languages are necessary. Users need to be able to express higher-level events such as "Track the toxic cloud, and report any topological changes" easily. This type of data representation model is also necessary to integrate sensor data with available domain knowledge and/or historic data.

Another mechanism for high-level and interoperable data representation of low-level sensor data are ontologies. Hornsby and King investigate the supporting role of ontologies for geosensor network data. In particular, they explore methods to link ontologies with real-time geosensor networks in order to augment the collected data with generalization or specialization relations from an ontology. For example, a geosensor network can observe moving cars on a freeway, tracking their location via a car identifier. The car identifier links the car to a classification scheme stored in a database; here, vehicles can be classified as military support trucks or medical support vehicles. Hornsby and King have implemented a mechanism to associate data values from the geosensor database with classes in the ontology, and support generalization or specialization queries based on the ontology.

Section 3: Applications

Today, the applications for geosensor networks are appearing in different domains ranging from habitat monitoring, watershed management, environmental pollution monitoring, deep sea explorations to monitoring food safety in South Africa and precision agriculture for large vineyards in Southern Australia. One can observe that a mindset change in the application areas is taking place. Scientists as well as practitioners are aware of technology advancements, which provide means for real-time availability of observational data and allowing existing sensor platforms to be networked. Networking sensor platforms of different types and scale provide an increased capability to correlate spatio-temporal information covering an entire region of interest. This paradigm shift from post-event, estimation based, historic data analysis to real-time, sensor-rich event detection and monitoring is fundamental in environmental applications. Adding small-scale geosensor network technology will change the awareness of the potential data scale over the next decade.

The paper by Pettigrew, Roessler, Neville, and Deese presents the GoMOOS project in the Gulf of Maine. Currently, the long-running coastal observing project consists of several, large-scale buoys distributed in the Gulf of Maine. Each buoy has several surface and underwater sensors attached, and uploads collected information via satellite link to a central computer. This information is used to model sea surface currents using a neural network approach since the available information currently is point based, thus, the currents in large areas need to be coarsely estimated. This project also relates to the paper by Trigoni, Ferentinos and Nittel in the second section discussing the deployment of a network of untethered ocean drifters to investigate the ocean currents directly, and potential collection information about nutrients and algae.

The next paper, by Terhorst, Moodley, Simonis, Frost, McFerren, Roos and van den Bergh, describes the deployment of another *environmental monitoring application*, targeted at detecting vegetation fires over Africa. Vegetation fires that burn over high-voltage electricity transmission lines can create line faults, which can disrupt regional electricity supply. Improving detection response time (ideally, making it real-time) can help mitigate the impact of such wild fires. Terhorst et al present the architecture of their Advanced Fire Information System and illustrate how a combination of two separate satellite systems with different characteristics can lead to improvement in detection accuracy.

Another interesting application area of geosensor networks is intelligent transportation systems. Location-based techniques have been used for the last decade to acquire information about the surroundings of the current location of a car or a pedestrian. Assuming the existence of sensors on devices as well as the ability for short-range, ad-hoc communication of nodes in close spatial proximity, a new range of application becomes possible. For example, several nodes collecting sensor information about their environment can collaborate and exchange, aggregate and forward this information. Often, sensed information is mostly relevant in the immediate neighborhood, and becomes less relevant with increasing distance. Wu, Winter and Guan address an approach to ride sharing using ad-hoc geosensor networks. Transportation clients, e.g. pedestrians, needing a ride from location l_1 to location l_2, request offers from transportation providers nearby. Transportation or ride providers can be private automobiles, taxis or public transportation. Based on the offers for complete or partial trips and accounting for the presence of competing clients, a client node has to strategize on accepting offers to cover the route at hand.

Outlook and Open Issues

The papers collected in this volume clearly demonstrate the interdisciplinary nature of geosensor networks, and their rapidly emerging potential to revolutionize the way in which we observe the physical world. There is an increased realization of the need for accurate and continuous monitoring of our environment. This environment is characterized by its inherent complexity (e.g., with the evolution of an ecosystem and the corresponding climate dynamics affecting and affected by human activities) thus mandating the realization of the potential offered by geosensor network applications.

If we attempt to identify relevant topics that have not yet emerged to the prominence that they deserve in our community, we can start with the issue of *sensor data* privacy. With the requirements to design ultra-light wireless communication protocols for small-form devices, there is limited room left for advanced encryption schemes. A related issue is the need for *authentication* of sensed data. If sensor networks are deployed in security-sensitive areas, built-in mechanisms need to be available to provide for such data authentication. A third open issue is *data quality*. Mechanisms need to assure that defective or incorrectly calibrated sensors are excluded from the computation, and that calibration is established individually as well as collectively before deployment and also continuously later on. Today, many research efforts in sensor networks are conducted under assumptions derived from the constraints of current hardware platforms such as the Berkeley motes. Many of these assumptions such as using radio broadcasting as communication modality or restricted battery life might not be valid anymore in a few years, and these assumptions might change completely. Lastly, one observation we can make based on the current state-of-the-art is that existing approaches to geosensor networks are rather passive: sensors observe events, and just report/record information. As geosensor networks become more mature and more pervasive, we expect the level of interaction between humans, geosensor networks, and the environment to increase dramatically.

Given the many different aspects and challenges of this interdisciplinary field, a single volume cannot possibly exhaust the emerging research agenda, but we believe that the papers collected here offer a valuable snapshot of current research, as it was reflected at the GSN'06 conference. We hope that this volume will serve as a reference point for new scientists venturing in this area, and hope that we will have them participate in future GSN conferences.