Geometric Approaches to Ad Hoc and Sensor Networks:

Report of the 2006 NSF Workshop

Subhash Suri
Department of Computer Science
University of California
Santa Barbara, CA 93106

Leonidas J. Guibas
Computer Science Department
Stanford University
Stanford, CA 94305

Alon Efrat
Department of Computer Science
University of Arizona
Tuscon, AZ 85721

Anish Arora (Ohio State U.), **Sándor P. Fekete** (U. Braunschweig, Germany), **Jie Gao** (SUNY Stonybrook), **Robert Ghrist** (U. Illinois), **Carlos Guestrin** (CMU), **Thomas C. Henderson** (U. Utah), **Brad Karp** (U. College London), **Thomas J. Peters** (U. Connecticut), **Sylvia Ratnasamy** (Intel Research, Berkeley), **Andrea Richa** (Arizona State U.), **Gaurav Sukhatme** (USC), **Andreas Terzis** (Johns Hopkins U.), **Roger Wattenhofer** (ETH Zurich), **Feng Zhao** (Microsoft)

Abstract

Embedded networked sensing devices are becoming ubiquitous across many activities that are important to our economy and life, from manufacturing and industrial sensing, to agriculture and environmental monitoring, to hospital operations and patient observation, to battlefield awareness and other military applications. In each such deployment modest to large numbers of simple devices that possess sensing and processing capabilities are networked together to form a sensor network. The fact that nodes in these networks are embedded in the physical world and their sensed data is highly correlated with their physical locations imparts a uniquely geometric character to these systems. The geometry and topology of both the sensor field layout, as well as that of the signal landscapes studied, greatly affects issues such as routing, data aggregation and information brokerage, outlier detection and other statistical processing, and so on.

This report summarizes the key findings about the opportunities presented by the exploitation of geometric methods in ad hoc and sensor networks, based on a two-day NSF-sponsored workshop held at the University of California at Santa Barbara during June 12–13, 2006.

1 Introduction

Enabled by recent advances in micro-electronics and fabrication, a new generation of integrated embedded devices, called *smart sensors*, has emerged that seems capable of realizing the long-cherished vision of *sensory omnipresence* or *ubiquitous awareness*. Through collaboration and *ad hoc* wireless networking, a collection of such devices can provide real-time, fine-grained sensing, monitoring, and actuation across large geographical areas. Because of their small form factors and ability to operate in an untethered mode, these *sensor networks* can achieve an unprecedented level of universality: they can be deployed almost anywhere (even air dropped), are able to organize themselves into a network through self-localization and *ad hoc* wireless communication, and function unattended for long durations. Building on these ideas, a number of exciting research prototypes have already been proposed and implemented as proofs of concept during the last few years, with varying goals of habitat monitoring, wildlife dynamics, aquatic observations, surveillance, structural monitoring, as well as global research initiatives like the International Polar Year (IPY). Given the scientific and engineering boldness of this vision and the enormous potential benefit to the society at large, it is not surprising at all that sensor networks have attracted strong interest from both academia and the industry.

Within academia, sensor networks have elicited interest from an unusually broad spectrum: from device designers to domain scientists, from computer architecture to operating systems, from programming languages to database systems, from signal processing to information theory, from physical layer medium to wireless networking protocols, from algorithms to computational geometry, from graph theory to topology, and so on. In industry, sensor networks have drawn the interest and investment from tech giants such as Intel and Microsoft to many startups such as Crossbow, Ember, MillenialNet, and Dust Inc, among others. This level and breadth of activity underscores both the *intellectual* richness and complexity of these systems as well as their *commercial* potential.

Today, with nearly a decade of research and development behind it, the field of sensor networks is at a formative stage where the scope and complexity of many basic issues is well-established and early research prototypes have either validated some of the design principles or shown their limitations. These early efforts have also highlighted the need for *inter-disciplinary research* because many of the fundamental issues in sensor networks span across multiple areas that have traditionally had only limited collaboration and interaction.

With this backdrop and motivation, a workshop on “Geometric Approaches to Ad Hoc and Sensor Networks” was organized at University of California, Santa Barbara, during June 12–13, 2006, under the auspices of the National Science Foundation. The workshop’s goal was to provide a forum for some of the leading experts in areas most relevant to “algorithmic and geometric foundations” of sensor networks, to discuss their research, help identify important research challenges, and formulate a set of recommendations that can help advance the field and accelerate the adoption and deployment of sensor networks. The following report is an “executive summary” of what the workshop participants believe are the most significant (algorithmic) research challenges in this endeavor. We conclude with a set of recommendations that, with the help of NSF initiatives, could provide a significant boost to the future scale, scope and adoption of sensor networks.

2 Sensor Networks and the Role of Geometry

A key fact distinguishing sensor networks from other networked systems is that sensor nodes are deeply attached to the physical environment in which they function — *they are embedded systems*. As a result, geometry plays a fundamental and crucial role in all aspects of the sensor network, including their design and operation. In particular, the physical layout of the network deeply affects issues such as routing and information discovery; communication depends on node proximity, and node proximity in turn determines the correlation between sensed values and affects what information is necessary to transmit; the global structure of signal landscapes determines whether local greedy methods get stuck locally or can reach the desired global optimum; and tracking mobile phenomena requires migrating processes in the network and robust end-to-end connections between them. Unlike more traditional networks such as the Internet or the phone network, communication in sensor networks is dictated less by the desires of the end nodes and more by the geography of the sensor field and the associated signal landscapes, as well as the overall network task.

At the same time, the geometry of ad hoc and sensor networks is not as explicit as the geometry traditionally studied in computational geometry and related disciplines. Node locations are not always available, proximity does not always imply connectivity, and wireless link variability creates volatile connectivity graphs. There is a sense in the community that there is geometry in sensor networks “in the large” — though perhaps not at the scale of an individual node or two.

Motivated by these observations, the discussion during the workshop focused largely on techniques of a geometric or topological character that are particularly relevant to sensor networks. The following sections discuss the six areas of technical problems whose solution, we feel, can significantly advance the state of sensor networks.

2.1 Structure Discovery and Self-Organization

A sensor network is a *self-organizing* network: it must discover its own network architecture and dynamically adapt to addition or loss of network elements. Unlike most conventional computing systems, however, even the basic initialization process can be quite challenging for such a network, due to the lightweight nature of its components and the variability present in node placement and network links. Yet, a proper understanding of this architecture is essential for an efficient implementation of all the higher level abstractions. The sensor networks is also programmable to a large extent, and the system must adapt its structure to best deal with the data it is likely to sense and the functions it has to perform. Among the key challenges identified under this topic are the problems of *network self-localization* and *sensor field morphology*.

Self-localization is important in sensor networks because manual calibration is not scalable, and GPS hardware at every node is not practical due to cost, power consumption, or large form factor. In self-localization, nodes use inexpensive ranging devices to estimate distances (or angles) among neighbors, and then deduce global coordinates from this partial distance matrix. This is the classical problem of *graph embedding*, with a long history in graph theory, rigidity theory, distance geometry and topology. Despite the fundamental nature of the localization, the problem still lacks a satisfactory solution — all current localization methods fail to localize parts of the network, even for relatively small network sizes and well-distributed nodes. A closer collaboration between geometers, mathematicians, and networking researchers is needed to accelerate progress on this key problem. When confronted with a difficult problem, engineers are especially good at “altering” the problem. In sensor localization, this can occur through the use of additional machinery, such as beacons, mobile nodes, or through incremental localization.

The goal of morphology is to understand the global layout of the sensor field, including the identification of boundaries, both external and those of network holes, the detection of narrow passages and other communication bottlenecks, etc. Morphological understanding is key to performing loadbalanced routing and many other essential tasks of the network: a hole can simply indicate the presence of some physical obstruction where nodes could not be placed (e.g., a pond) or it could signal a region where something bad is going on, such as an area where nodes have depleted their resources or were destroyed. Topological methods (in the sense of the so-named part of mathematics) can play a fundamental role in this context, as they are both global and robust — we are now seeing a number of topologists taking more interest in sensor networks.

A global understanding of the network topography is also important for addressing security concerns, including the detection of wormholes, malicious impostor nodes who lie about their location, etc.

2.2 Naming and Routing

The next fundamental step beyond structure discovery is to provide a mechanism for sensor nodes to "identify" themselves and build a routing structure to communicate among themselves. Telephone numbers, IP addresses, postal addresses, etc. are examples of "naming" schemes that help identify, locate and establish communication with entities within a network. The unique characteristics of sensor networks, however, require rethinking the traditional role of (and, indeed, even the need for) addressing in these networks. Because sensor networks aim at easy network deployment and initialization, they are less amenable to carefully configured addresses like the IP or telephone networks. In addition, unlike IP and telephone networks that are used primarily for pairwise communication, sensor networks are largely envisioned as data-centric systems that emphasize collection, aggregation, brokerage and dissemination of information. The consensus among the workshop participants was that significant new research is needed to achieve lightweight and scalable routing and naming for sensor networks. The current state of the art can only support fairly primitive uses of sensor networks. In particular, there are two leading approaches to scalable routing in sensor networks, location-based and virtual-coordinates based, and they both have significant problems to overcome.

The location-based routing, at first glance, offers a very elegant solution — the coordinates of a node, obtained either using GPS or other localization techniques, serve as its name, and the messages between two nodes are routed using the *geographic* paths (greedy forwarding, face routing, etc). While elegant in conception, geographic routing has been hampered in its adoption due to at least two major concerns. First, obtaining real coordinates is both expensive and unreliable due to the cost, power consumption, and limitations of GPS in mobile or indoor environments. Second, and more unexpectedly, empirical studies have repeatedly shown that physical proximity and robust wireless connectivity are not always congruent, violating a key assumption implicit in geographic routing. Current research aims to develop novel routing algorithms that can tolerate incongruencies between physical distance and connectivity.

The idea behind virtual coordinates is to construct a virtual system of coordinates based entirely on the (measured) connectivity among the nodes. Since no localization is needed, virtual coordinates offer an inexpensive naming solution for routing and a network-wide frame of reference. In spite of extensive work, however, current methods for constructing virtual coordinates are neither efficient nor robust. In particular, the current schemes work well only for dense and uniform networks where communication graph distances are good estimators of the underlying (but unknown) Euclidean distances. More approaches need to be explored, including collections of local embeddings, coupled with alternative, high-level routing mechanisms.

Geographic routing with either real or virtual coordinates is only known to work in 2-D, or by enforcing limited 2-D connectivity even when the network deployment is 3-D. More abstract routing methods based on landmark coordinates have been proposed to make routing dimension-independent.

In order to realize the vision of highly scalable and sophisticated sensor networks, we will need scalable routing support, and geometry-based approaches appear to be the most promising in this regard. The challenge is to find geometric solutions that are simple and yet robust to (i) location inaccuracies, (ii) the challenges of radio propagation, (iii) network deployment dimension (2D or 3D), and (iv) temporal link volatility and node failures. Node mobility, when present, adds its own complexities in terms of variable node neighborhoods and intermittent connectivity. In addition, as sensor networks grow in scope, coverage, complexity and diversity, we will also need to broaden our view and handle networks in which heterogeneous sensors are deployed within a larger interconnected world — mobile phones and cameras in an urban environment, sensors in a home, vehicular networks and so forth. In such a setting, a closer coupling with the wired Internet and its users and applications would require addressing schemes that focus on interoperability with, and management within, the public Internet. The requirements of, and options for, addressing in such highly-networked environments add another interesting and challenging dimension to the role of naming in sensor networks that is as yet not well understood.

2.3 Information Aggregation and Dissemination

The next algorithmic challenge is to deal with the problems of data representation and management. If sensor networks are to become enablers of pervasive awareness, they must provide “visibility” into a (remote) physical environment through measurements and observations, as well as distributed intelligence for various computational or actuation tasks. These tasks are made especially challenging due to the scale of sensor networks, both in time and space, and their limited resources (energy, bandwidth). The concerns of resource limitations have led to much interesting research in distributed signal compression, in-network processing, etc. for economizing and managing network resources while executing their tasks. In this quest, a dominant viewpoint for sensor networks has been as a data collection system whose primary goal is to deliver data to a central base station. As sensor networks grow larger in size, however, and especially as we envision systems consisting of interconnected sensor networks spanning large regions or even the entire country, such centralized data collection is neither desirable nor feasible. We must therefore explore scalable and collaborative mechanisms for the network in which the nodes can decide what to sense, and where to store past observations, so as to best serve the current network users — users possibly embedded in the same physical space where the network is operating. Applications may require information to be delivered with stringent latency constraints, while at the same time conserving resources in order to maintain network coverage, connectivity, and longevity for future use. This becomes especially important for “closing the loop” around sensor networks — when the goal is to use sensor data to act in real-time on the physical world. Additionally, while most of the past work has focused on numeric data, much of sensor data is likely to be “spatial” in nature, a signal landscape, so to speak, which is best modeled and reasoned about using geometric methods.

Several research challenges were identified by the workshop participants, including methods for information dissemination, aggregation, summarization, and brokerage. In particular, flooding and blind gossiping are the de facto methods for information dissemination or discovery in unstructured networks, but they are clearly ill-suited for resource-constrained sensor networks. Techniques that exploit the geometric embedding of a sensor networks can potentially lead to much more efficient methods. Second, most sensor network tasks do not require *raw data* — instead, users are typically interested in useful statistics about data, *anomalies* in data, or important changes in trends or patterns.

Thus, transporting raw, unprocessed data wastes not just the resources of the source sensor but also drains the bandwidth and the energy of many other nodes along data's multi-hop route. But most current schemes for in-aggregation data suffer from serious shortcomings: they assume reliable tree-based routing, or are limited to fairly simple statistical aggregation (such as min, max). We need generic and lightweight methods for summarizing spatial data sensed by the network, and efficient distributed schemes for recognizing changes in trends. Finally, a key *information brokerage* problem is to efficiently match producers (sources) of information with consumers (queries), especially when actionable information needs to be delivered to users embedded and operating in the same space as the network.

We expect that the key to solving many of these problems is to heavily exploit the fact that a sensor network is embedded in physical space and employ techniques motivated by geometric, topological, or physical analogies. As a physical example, we are all familiar with following a sound gradient to arrive at a sound source, such as a water fall — and such methods have been used in sensor networks for access to non-physical quantities such as information. Indeed, these gradient methods for locating information have nice robustness properties. But imagine now there are a myriad of sounds (information) sources in the environment — can we follow the chirping of particular bird in the cacophony of a rain forest? A variety of information coding techniques can be brought to bear so as to allow us to selectively follow only the gradients of information that we are truly interested in. Such coding techniques also add some measure of information security to the system. As another geometric example, "road-systems" can be established in a network that naturally create rendez-vous points for information producers and information consumers, following the natural morphological features of the network.

2.4 Sensor Placement and Coverage

A sensor network, ultimately, is only as good as the quality of its sensed data. While the *capabilities* of a sensor are a problem of engineering, and the *communication* of the sensor signal is a problem of information and network protocols, the problems of sensor *placement and coverage* are fundamental problems in the geometric design of the network. These problems, which range from determining optimal placement and allocation of sensor nodes to adapting a given distribution of nodes to the desired task, are uniquely geometric in nature. Further, one cannot decouple sensing or communication issues from the geometry of the embedded network.

For example, in order to provide a concrete context for many of the challenges involved in sensor placement and coverage, let us consider a few motivating (and very real) applications of sensor networks. (1) Biologists investigating aquatic ecosystems are interested in the study of biomass, including algal blooms. A proposed deployment will include a large number of sensors, both static (buoys) and mobile (boats), along with robotic sensors that traverse a cable and descend to different depths in the lake. In such a deployment, the positions of the buoys, paths of the boats, and the motion of networked infomechanical systems (NIMS) must be jointly optimized. (2) Placement of sensors in water distribution network poses at least two fundamental challenges: the evaluation of the quality of any given sensor placement, and the optimization over a combinatorial number of such placements. The quality of a sensor placement is based on various performance criteria, such as average time to detection, population affected by an intrusion, consumption of contaminated water and the fraction of detected scenarios. (3) Military sensor deployments for surveillance and search are immensely complicated because of the complex nature of acoustic signals as well as the need to often coordinate the motion of UAVs with the placement of acoustic sensors.

Thus, the non-uniformity of physical space, presence of obstacles, non-isotropic nature of sensing, multi-modal sensor fusion, and constraints on the possible sensor locations present some very challenging yet fundamental problems. There is great need for efficient algorithms to find good placements, with provable quality. The issue of dealing with noisy sensors is largely unexplored territory. Another significant research direction is to understand the power of *incremental* placement, where each additional sensor placement optimally helps the algorithm *discover* and *learn* the properties and the structure of the sensing environment, as well as provide guidance for future placements. Low-cost sensors are prone to failures, and so we also need to find placements that are also robust to adversarial attacks. Note also the tension between the need, on the one hand, to spread out the sensors to attain better coverage, and on the other, to provide a sufficient local sensor density to ensure reliable connectivity. The usefulness of submodular functions in such optimization problems has been recently demonstrated.

2.5 Sensor Networks and Mobility

An important aspect of our vision of sensor networks is their ability to *actuate* and act upon on their physical environment — and mobility is an important component in this regard. Indeed, as technology matures, we expect to see many systems that involve cooperation between static and mobile nodes. Naturally, geometric methods are fundamental in dealing with mobility and the workshop participants outlined several research problems that need to be addressed in order for sensor networks to realize the full potential of mobility. These challenges can be articulated at three levels, depending on the level of interaction between the robots and the network of (static) sensors.

In the simplest setting, we essentially have a mobile sensor board: a single robot that carries all the sensors onboard. A key challenge in this scenario is to plan the path of this robot in such a way that the collected data can be successfully interpolated during the motion. In many situations, the path of the robot may need to be updated dynamically in response to sensed data, for instance, to obtain additional samples in certain other areas. In other settings, the environment may be partially or fully unknown, requiring online mapping. In general, the problem may require simultaneous localization, mapping, and exploration subject to communication constraints.

At the next level of interaction, multiple robots with onboard sensors coexist and cooperate with an embedded set of static sensors. The communication and interaction among these devices can take multiple forms. The robots may act as relay stations for the static nodes, the robots may act as *data mules* for the static nodes, or the static nodes may act as “sign posts” for the robots, providing navigation and path planning advice. The robots could also carry energy replenishments for the static nodes, and, finally, the robots may even pick up, drop off or reposition static nodes for a physical reconfiguration of the network. This rich set of possibilities suggests many geometric questions concerning (optimal) path planning for the robots in ways that enable them to “visit” the statics sensor nodes.

Finally, one can envision a swarm of inexpensive robots, with limited localization and computational capabilities, performing sensing, data collection and actuation tasks. An algorithmic theory of large scale collaborative exploration, geometric reasoning, motion coordination, data collection is only in very early stages of development, and needs significant progress before it can provide a solid foundation for mobility in sensor networks.

2.6 Wireless Network Models

While it is tempting to assume that the physical layer of the communication in sensor networks, namely the wireless medium, is well understood, the reality suggests otherwise. Modeling the complexity and spatio-temporal variability of a wireless channel remains a major challenge, and even today many of the fundamental questions about the network's connectivity, transport capacity, power levels, etc., remain largely unanswered.

Historically, a convenient and theoretically tractable model has been the *unit disk* model, which describes a radio device's transmission (or reception) range as a disk of radius R, called its transmission radius. Each and every receiver inside the transmission range can "hear" the message, while no one outside the range can. While simple for analysis, many experiments have shown that the true nature of the transmission range is far from this ideal: the real range can be highly non-uniform, extending far beyond R in some places and significantly less than R in others; it may even have holes; and also shows significant temporal variation.

Thus, even determining the network connectivity, namely, which nodes can communicate, is a challenge in wireless networks, and it is amplified in wireless sensor networks due to the low energy level of nodes and the scarce bandwidth in the system. Nodes in sensor networks must also rely on multi-hop communication, further stressing the system. Some recent work has gone beyond the idealized unit-disk model, and considered variants such as the *quasi disk* model, but even those models fall well short of capturing the true nature of radio communication and significant work is needed to develop a satisfactory theory of network connectivity in wireless networks.

A second important feature of wireless transmission, which has only recently begun to receive attention, is the phenomenon of *interference*. Unlike wired networks, communication in wireless networks is broadcast based: a node's transmission interferes with the *receiving* ability of all other nodes in its neighborhood. Several different proposals for modeling this effect have been made: *unit disk with distance interference* model, which assumes that a pair can communicate as long as the receiver is not disturbed by a third nearby node; the *protocol model*, which requires that the receiver be outside a guard (interference) zone of every other transmitting node; and the *physical or signal-to-interference plus noise* model, which requires that the intended signal level at the receiver should be above a certain threshold relative to ambient noise and the interference caused by other simultaneous transmissions. Due to interference, the *throughput* capacity of the network depends critically on the choice and the schedules of transmissions, and is really a problem about *geometric packing*. This viewpoint has been advanced in several recent papers, but the problem is only partially understood and, more significantly, lacks clean and usable *algorithmic* solutions.

3 Opportunities, Impact, and Recommendations

Geometric approaches, through concepts and techniques, offer a number of opportunities in sensor networks to address problems at structural, functional and application levels. These opportunities are new in the sense that they have not been previously investigated in the context of Internet-scale networking. They are also expected to play an important role in the ongoing debate of how to integrate sensor networks with Internet's next generation, as addressed by NSF's FIND and GENI projects.

Indeed, on the networking side, sensor networks have already generated a lot of research addressing issues of low-power, irregular topologies, wireless link volatility, etc. But perhaps the most unique characteristic of sensor networks is that they create an environment where data transport cannot be separated from data content — and both are intimately tied to geometry. Indeed, the geometry of signal landscapes becomes deeply intertwined with the geometry of data acquisition and transport. Thus NSF may want to consider joint solicitations between CNS and other CISE divisions, like CCF or IIS, blending networking with geometric techniques in data management, data mining, probabilistic reasoning, etc., all in the context of embedded networked sensor systems.

It is also important to address the problem of "closing the loop" — taking sensor networks beyond just passive information collection instruments and using them to act on the world in real time (many surveillance/security applications have this character), incorporating sensor mobility in the process. This can be both active mobility, by a small number of robotic platforms, as well as passive mobility, by nodes passively advected by moving physical objects. This creates many new challenges for the network communication layer, such as intermittent connectivity, latency guarantees on information delivery, QoS type issues, security, and so on — again many of a geometric character. Several issues in planning and evaluating sensor coverage also arise. Some of these topics have of course been addressed by the networking community in the context of other networks, but the collaborative aspect of sensor networks presents many new twists, and provides new opportunities for collaboration between sensor networks, control theory, and geometry/topology researchers.

Conversely, there are also tremendous opportunities for novel research in computational geometry. These opportunities go well beyond simply applying known geometric concepts to a new application domain. There is a clear need to develop a more qualitative understanding of the kinds of spaces to which geometric machinery can be applied. Moreover, the lightweight and distributed nature of the sensor platform, coupled with the constraints of energy, reliability, robustness, and security, necessitate a fresh approach to geometric algorithms. Moreover, as research on sensor networks discovers the true geometry of electromagnetic and other physical phenomena in diverse settings, we can expect new probabilistic, statistical, and model-based techniques to be devised for geometric solutions. Therefore, we feel strongly that sensornet-directed research in computational geometry, algebraic topology, statistics, and optimization theory will be particularly beneficial to resolving many of the key challenges outlined in this report, and pushing the envelope of scalability, robustness and scope of sensor networks.

One possible approach to seed progress along these opportunities is through large-scale sensor network "community" testbeds. These testbeds would enable high-fidelity validation of geometric solutions as well as facilitate apples-to-apples comparison of geometry-inspired sensor network solutions with more traditional internet-style graph solutions. A large-scale testbed would also lend insight into how different algorithmic components in the sensor network (e.g., initialization, localization, structure discovery, routing, data storage, information brokerage, aggregation) interact with each other.

4 Workshop Web Site

The workshop web site is hosted at `http://www.cs.ucsb.edu/~suri/Workshop06/workshop.html`, where the original workshop schedule, position papers, talk slides, and additional participant write-ups can be found.