Parametric Probabilistic Sensor Network Routing[*]

Christopher L. Barrett
barrett@lanl.gov

Stephan J. Eidenbenz
eidenben@lanl.gov

Lukas Kroc
kroc@lanl.gov

Madhav Marathe
marathe@lanl.gov

James P. Smith
jpsmith@lanl.gov

Los Alamos National Laboratory
Basic and Applied Simulation Science (CCS-5)
MS M997, P.O. Box 1663
Los Alamos, NM 87544, USA

ABSTRACT

Motivated by realistic sensor network scenarios that have mis-informed nodes and variable network topologies, we propose a fundamentally different approach to routing that combines the best features of limited-flooding and information-sensitive path-finding protocols into a reliable, low-power method that can make delivery guarantees independent of parameter values or information noise levels. We introduce *Parametric Probabilistic Sensor Network Routing Protocols*, a family of light-weight and robust multi-path routing protocols for sensor networks in which an intermediate sensor decides to forward a message with a probability that depends on various parameters, such as the distance of the sensor to the destination, the distance of the source sensor to the destination, or the number of hops a packet has already traveled. We propose two protocol variants of this family and compare the new methods to other probabilistic and deterministic protocols, namely constant-probability gossiping, uncontrolled flooding, random wandering, shortest path routing (and a variation), and a load-spreading shortest-path protocol inspired by [Servetto, Barrenechea, 2002]. We consider sensor networks where a sensor's knowledge of the local or global information is uncertain (parametrically noised) due to sensor mobility, and investigate the trade-off between robustness of the protocol as measured by quality of service (in particular, successful delivery rate and delivery lag) and use of resources (total network load). Our results show that the multi-path protocols are less sensitive to misinformation, and suggest that in the presence of noisy data, a limited flooding strategy will actually perform better and use fewer resources than an attempted single-path routing strategy, with the *Parametric Probabilistic Sensor Network Routing Protocols* outperforming other protocols. Our results also suggest that protocols using network information perform better than protocols that do not, even in the presence of strong noise.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols—*Routing Protocols*

General Terms

Algorithms, Measurement, Performance

Keywords

Sensor Networks, Probabilistic Routing, Simulation

1. INTRODUCTION

As the deployment of large-scale sensor networks is imminent, the need for efficient communication protocols that govern the interaction of the sensors with each other continues to be of utmost importance. A sensor network consists of a large number of mobile devices that measure an environmental variable such as temperature, pressure, radioactivity, or geographic position. The sensors are equipped with transceivers that enable them to transmit and receive data to and from other sensors. In our model, most sensors only have a very small transmission range, while a single sensor is equipped with a powerful transmitter device (along with a more powerful battery) that enables it to forward data to a fixed base station. The need for communication between a regular sensor, called source, and the powerful sensor, called destination can arise at any time possibly triggered by unexpected changes in the environment variable or by a timer; we assume that a communication session consists of sending a single packet of data to the powerful sensor who then forwards it to the base station; this assumption is reasonable as the data of an environmental variable only consist of a few bytes. If a source sends out its data packet, chances are that the packet will not reach the destination in a single hop due to the typically very limited transmission range of the source; in order to obtain a functioning system, other sensors that receive the packet transmitted by the source need to forward the data packet themselves until it reaches the destination. This model of communication results in a multi-hop path from source to destination with intermediate sensors forwarding the packet.

The problem of how to route a data packet has been the subject of intense study in wire-line and wireless settings. The unique requirements of a sensor network setting (e.g., limited battery power, high network node density, and frequent network topology changes) have also led to proposed routing protocols specifically tuned for sensor networks such as SMECN, Flooding, Gossiping, SPIN, SAR, and LEACH, see [1] for an overview. Traditionally, if a source and a destination is given, most routing protocols—particularly in ad hoc networks—first compute a routing path and then route data packets along the computed path. In our setting, this approach seems to be somewhat overkill as we only need to send a single packet to the destination. In fact, a straight-forward flooding approach (i.e., every sensor who receives the packet retransmits it to all its neighbors) would solve our routing problem, but at the cost of involving all sensors in every transmission and thus unnecessarily depleting their batteries. A wandering approach alleviates this problem: in the WANDERER protocol, a sensor that receives a packet forwards the protocol to only one of its neighbors, which it chooses according to some given probability distribution. Such an approach, however, may severely increase path lengths. A third approach, PURE GOSSIP tries to combine the advantages of FLOODING and WANDERER by deciding at each node whether to forward the packet to all neighboring nodes with a given probability. A well-known result from percolation theory (see [6] for details) says such an approach will result in either almost all nodes receiving the packet or almost no nodes receiving the packet.

Our Contribution: In this paper, we propose a family of routing protocols called Parametric Probabilistic Sensor Network Routing protocols, which substantially improve the performance of controlled flooding methods like gossiping by making the probability of retransmission a function of several parameters. Parametric Probabilistic Sensor Network Routing protocols are completely described by the retransmission probability function: each node in the network—upon receiving a packet—retransmits the packet to its neighbors with a certain probability according to the probability function. While this basic principle is straight-forward, we must be careful not to overload the probability function: the retransmission probability could depend on parameters as diverse as the numbers of copies of the packet already received by a certain node or the distance to the destination. In order to keep the routing protocol as light-weight as possible, we try to keep the retransmission probability as simple as possible with respect to the complexity it takes to compute or estimate the parameters that the function depends on. We propose two variations of Parametric Probabilistic Sensor Network Routing: The retransmission probability function of our first variation DESTINATION ATTRACTOR depends on the distance counted in hops from the source node to the destination node and the hop-distance from the node currently holding a packet to the destination; in our second variation DIRECTED TRANSMISSION, the probability depends on the same two factors and additionally on the number of hops that the packet has already traveled.

We measure the performance of DESTINATION ATTRACTOR and DIRECTED TRANSMISSION against each other as well as against a number of other proposed sensor network routing protocols through simulation in a realistic setting. The protocols we consider are FLOODING, PURE GOSSIP, and WANDERER as described above. We also compare against a SHORTEST PATH protocol that always reaches the destination in a minimum number of hops and a variation called SHORT PATH. The SHORTEST PATH protocol serves as a theoretical optimum that no real-life protocol can ever match: the large number of sensors and their level of mobility make the cost of updating routing tables prohibitive. It is included to investigate the degree to which our Parametric Probabilistic Sensor Network

Routing protocols achieve this optimum performance. Finally, we compare against a protocol we call SHORTEST PATH COUNTING, in which a packet is sent to one neighbor who is closer to the destination and sits on a shortest path to the destination; the probability of choosing a specific neighbor depends on the number of shortest paths on which the neighbor sits. SHORTEST PATH COUNTING is a generalization of an earlier protocol [13] that works on general graphs (rather than only on grids as in [13]). Most of the methods we use[1] rely on global information that changes as the sensors change positions. As exact and continuously updated computation of this information would incur prohibitively high overhead in these protocols, estimator methods are required. Rather than implementing a particular estimator method, we model sensor mobility through inaccurate data: the parameter values that the sensors use are "noised".

In our simulations, we investigate the trade-off between quality of service measures (in particular the fraction of times that at least one packet reaches the destination and the average lag before a packet reaches the destination), load incurred in the network (measured as the number of times that a sensor emits a packet summed over all time steps and nodes) for different levels of noise. The key finding of our simulations is that between the two protocols from the Parametric Probabilistic Sensor Network Routing family, DIRECTED TRANSMISSION outperforms DESTINATION ATTRACTOR on all levels of noise achieving better quality of service results while incurring less load in the network. Both Parametric Probabilistic Sensor Network Routing protocols in turn clearly outperform PURE GOSSIP on all noise levels despite the fact that PURE GOSSIP is not sensitive to noise at all as it does not rely on global information. These three multi-path protocols incur more load in the network than the single-path protocols, but only for low levels of noise; for high noise levels, single path algorithms such as SHORT PATH can break leading to excessive lag times and even load, while delivering only a small fraction. Among the single-path protocols, SHORT PATH outperforms SHORTEST PATH COUNTING, while WANDERER and SHORTEST PATH come in last. Our results thus suggest that multi-path algorithms are more robust to mis-information, and can actually have a higher delivery rate, lower lag *and* lower load under the same conditions.

Related Work: Routing in wireless and ad hoc networks has been the subject of active research over the last two decades, see [14] for the state of the art until 1995 and more recent survey papers [12, 3, 11]. The tutorial by N. Vaidya at MobiHoc 2002 also provides an excellent upto date survey of recent protocols. As argued in [9], one possible reason for researchers proposing a large number of routing protocols is that efficient routing aims to optimize a variety of different measures, including, path lengths, load balancing, robustness, etc. Here we concern ourselves with the following design constraints that are particular to sensor networks: (i) low energy requirements, (ii) simple and local routing methods, (iii) robustness under node/edge failures. Probabilistic routing is a simple method that designed correctly satisfies all the above constraints.

Probabilistic routing and gossiping in particular have been applied in various setting, such as spreading distributed database updates [4] and multicast techniques (see [10] for a recent result and an overview). While earlier work on gossip usually assumes a complete graph as network and nodes then randomly choose to send a message to one or several of its neighbors, we make no assumptions on the structure of the underlying network. Here, we are concerned with wireless communication and thus assume omnidi-

---

[1]Namely, DESTINATION ATTRACTOR, DIRECTED TRANSMISSION, SHORTEST PATH, SHORT PATH and SHORTEST PATH COUNTING.

rectional transceivers that either forward a message to all of their neighbors within range or to none of them. This model has been introduced by Haas et al. [6], where the authors propose to use gossiping as a means to reduce flooding in the route discovery phase of ad hoc routing protocols such as AODV. Most previous work on gossiping assumes a constant probability for forwarding messages and tries to exploit the bimodal behavior of gossiping. Non-constant probability gossiping has been developed by Kempe et al. [7], in which the authors use a gossip-based approach to spread alarm messages, where the probability of forwarding to a particular neighbor (again assuming a complete and infinite graph) depends on the geographic distance to that neighbor. Multiple path routing has been applied in [5] in order to make a network resilient against node failures.

The SHORTEST PATH COUNTING protocol is inspired by the recent work of Servetto, Barrenechea [13]. In that paper, the authors considered the problem of routing in sensor networks, when the sensors are placed on a uniform 2D-grid and sensors can fail with certain probability. This model is closely related to the classical work on percolation theory, with the important distinction that we are now interested in the length of paths between a given source and destination. The authors leave open the design of probabilistic local methods for unstructured networks. Our work is a step in this direction. In our work, we model network irregularities such as node failures or mobility through noise.

Organization: The remainder of this paper is organized as follows. Section 2 presents the Parametric Probabilistic Sensor Network Routing protocols. We briefly present the other protocols that we consider in Section 3, and present a method for estimating global information in Section 4. Section 5 presents the experimental design; Section 6 contains a discussion of our simulation results. Concluding remarks and directions for future research can be found in Section 7.

2. PARAMETRIC PROBABILISTIC SENSOR NETWORK ROUTING PROTOCOLS

Parametric Probabilistic Sensor Network Routing protocols apply a limited-flooding strategy. The key element of our protocol is the idea that the retransmission probability for a packet at a sensor node is a function of various parameters rather than a constant as used in previous gossiping approaches such as in [6]. We will show that this allows us to further limit flooding (and thus further reduce the load of the network) as compared to constant-probability gossip solutions.

Let S denote the source sensor (the origin of data), D denote the final destination sensor (the one we want to send the data to); in time step i, all sensors who received a packet in the previous time step $i-1$ decide whether to retransmit the packet and, if affirmative, do so; let R_i denote a sensor holding a packet at time step i, and let R_{i-1} denote the sensor that sent the packet to R_i (if R_i received the packet form more than one sensor, R_{i-1} is defined to be any one of them).

The same sensor can hold a packet on different timesteps, and the same packet can be on multiple sensors at the same time. Source S emits its packet in time step 0, thus $S = R_0$. We denote the hop-distance between two sensors, say S and D, by $d(S, D)$.

The retransmission probability of a protocol from the Parametric Probabilistic Sensor Network Routing family could depend on various parameters, such as $d(S, D)$, $d(S, R_i)$, $d(S, R_{i-1})$, $d(R_i, D)$, $d(R_{i-1}, D)$, the number of neighbors of R_i or R_{i-1}, or the number of received copies of the same packet. Using subsets of these parameters, we choose two forms that yield qualitatively different

protocols that we call DESTINATION ATTRACTOR and DIRECTED TRANSMISSION.

2.1 Destination Attractor

The DESTINATION ATTRACTOR evolves from a very simple and sensible concept: If the packet is getting closer to the destination, then its retransmission probability is increased; if it is getting further away from the destination, then the retransmission probability is reduced. For a packet at the ith node R_i in its path from the source, this probability P_{R_i} can be represented in a recursive manner as

$$P_{R_i} = \begin{cases} (1 + k) \, P_{R_{i-1}} & \text{closer to destination} \\ (1 - k) \, P_{R_{i-1}} & \text{further from destination} \\ P_{R_{i-1}} & \text{same or indeterminate} \end{cases} \quad (1)$$

where k is the relative amount by which the probability should be increased or decreased. The three cases of Equation 1 can be combined if we use the change in the distance to destination $d(R_{i-1}, D) - d(R_i, D)$ in the last time step as an additional factor. In the case of accurate information, $d(R_{i-1}, D) - d(R_i, D)$ always has a value of ± 1 or zero, as in Equation 1. Thus,

$$\begin{aligned} P_{R_i} &= \{1 + k \, [d(R_{i-1}, D) - d(R_i, D)]\} \, P_{R_{i-1}} \\ &\approx \exp \{k \, [d(R_{i-1}, D) - d(R_i, D)]\} \, P_{R_{i-1}} \quad (2) \end{aligned}$$

The second line of Equation 2 is derived from the similarity to the low order terms of a Taylor series expansion of the exponential function $(\exp(x) \sim 1 + x)$. We arrive at a non-recursive form that is more amenable to analysis by evaluating the recursive terms back to the source $S = R_0$ and by using the fact that the sum of a contracting sequence is the difference of its end-points:

$$\begin{aligned} P_{R_i} &= \exp \left\{ k \sum_{j=1}^{i} [d(R_{j-1}, D) - d(R_j, D)] \right\} \\ &= \exp \{k \, [d(S, D) - d(R_i, D)]\} \quad (3) \end{aligned}$$

Obviously, $P_{R_i} > 1$ is possible (in which case the probability is one) if the packet moves closer to the destination over several time steps; this is intended to allow a packet to build up "positive energy", which can only be lost over a series of time steps moving farther away from the destination.

Thus, the retransmission probability for DESTINATION ATTRACTOR depends on the hop-distance of the source to the destination as well as the distance of the node currently holding a packet to the destination. With an infinitely dense planar graph, the local values of the retransmission probability function are illustrated in Figure 1 (left panel), where the destination is located in the center of the butte and the source node is located at any place on the rim of the butte. This explains the name DESTINATION ATTRACTOR. Parameter k models the exponential decay constant whose increase will cause the contour to more closely resemble a pedestal with vertical sides; decreasing k will spread the tails of the contour. In both cases, the probability within the circle of distance $d(S, D)$ around the destination will always be unity. However, this is all done in the presence of perfect information. In the presence of noisy data, the contour in Figure 1 (left panel) would acquire mesas extending out from the main pedestal in the areas of missing information. It would also naturally change in the presence of mis-information. Parameter k can be used to help alleviate problems caused by mis-information: decreasing k makes DESTINATION ATTRACTOR behave more like flooding. Thus, for a particular sensor network setting (for example sensor with excessive mobility in a tornado)
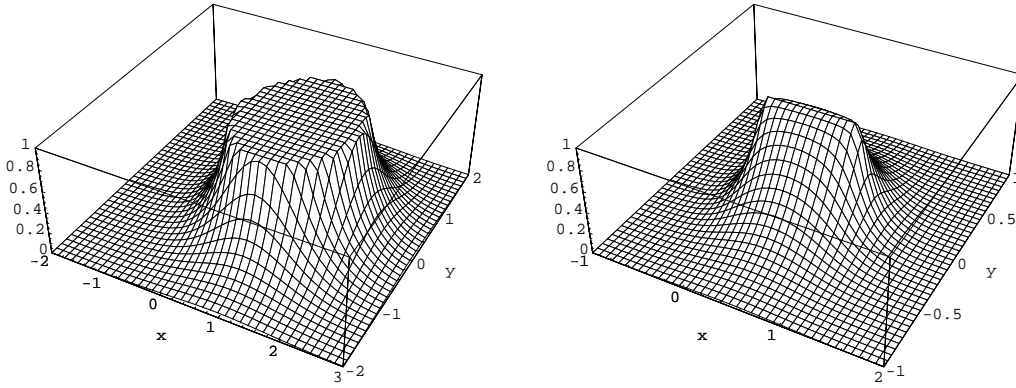
Figure 1: Retransmission Probability contours for DESTINATION ATTRACTOR **and** DIRECTED TRANSMISSION**, where the source is located at coordinates (0,0) and the destination at ((1,0).**

we can tune parameter k so as to still guarantee a required QoS-measure.

2.2 Directed Transmission

The retransmission probability function that we use for the DIRECTED TRANSMISSION protocol is based on the idea that the nodes that lie on a shortest path from the source to the destination should forward packets with a very high probability and the farther away a node is from the shortest path, the smaller its retransmission probability should be. Let R_i' denote any node that holds a packet at time i. The node minimizing $\min_{R_i'} d(R_i', D)$ is the node closest to the destination. Thus, the retransmission probability P_{R_i} of node R_i holding a packet at time i is defined as:

$$P_{R_i} = \exp\left\{ k \left[\min_{R_i'} d(R_i', D) - d(R_i, D) \right] \right\}$$

We have chosen an exponential form in order to remain amenable to analysis. Since we assume that the network topology does not change dramatically during a transmission (and if it does, we model it through noise), there is at least one packet that travels along a shortest path from source to destination; moreover, we know that such a packet will take $d(S, D)$ time steps to reach destination as it advances one hop in each time step. Thus:

$$\min_{R_i'} d(R_i', D) = d(S, D) - i$$

This allows us to write the retransmission probability in a simpler form:

$$P_{R_i} = \exp\left\{ k \left[d(S, D) - d(R_i, D) - i \right] \right\}$$

Even though their motivation is very different, DIRECTED TRANSMISSION and DESTINATION ATTRACTOR use the same two parameters (i.e., $d(S, D)$ and $d(R_i, D)$); in addition, the DIRECTED TRANSMISSION protocol uses the number of time steps i that the packet has traveled so far.

With an infinitely dense planar graph, the local values of the retransmission probability function for DIRECTED TRANSMISSION is illustrated in Figure 1 (right panel), where the destination is located on one end of the ridge and the source node is located at the other end of the ridge. Since the protocol aims to forward packets in the direction of the destination, we call it DIRECTED TRANSMISSION. Decreasing k will cause the shape to have less steep walls. Parameter k can be tuned by a system designer to achieve a certain QoS-level.

3. OTHER PROTOCOLS AND OVERVIEW

This section describes the other routing protocols used in our simulations.

Wanderer: The simplest routing protocol that we use is called WANDERER, which is simply a random walk on the network of possible radio connections starting from the source. When transmitting a packet, the sender chooses only one of its neighbors to be the receiver. All neighbors receive the packet (since we assumed omnidirectional radios), but only the neighbor specified as the receiver will retransmit. WANDERER does not use global information, but each node must maintain a list of its neighbors in order to choose one randomly. As a node may receive the same packet multiple times, the node keeps a record of how many times it has forwarded a particular packet. It may make sense to only forward the same packet a certain number of times; this number, called duplicate-forwarding parameter, can be set by a system designer.

Short and Shortest Path: In the presence of perfect, static global information, nothing will out-perform a SHORTEST PATH algorithm, which keeps only one copy of the packet alive at any given time and chooses the next node in the path uniformly among those that are closest to the destination. Since this method breaks down in the presence of noise, we introduce a method we call SHORT PATH: instead of choosing the next node exclusively from those that are closest to the destination, a sensor node R_i chooses the next sensor node uniformly at random among all the sensor nodes who claim to be closer to the destination than the present node R_i. For these methods, a system designer can set the duplicate-forwarding parameter, which makes sense for SHORTEST PATH only if we have noise.[2]

Shortest Path Counting: The SHORTEST PATH COUNTING protocol is an extension of SHORT PATH: sensor R_i currently holding the packet chooses to forward to one of its neighbors, where the probability of forwarding to neighbor R_{i+1} (with liberal use of notation) is a function of the number of shortest paths from the source or to the destination on which R_{i+1} lies. This was inspired by the work of Servetto and Barranechea [13]. Our modification is in the edge weighting, which is generalized to arbitrary graphs. The probability $P(R_i, R_{i+1})$ that sensor R_i chooses to forward to

---

[2]Setting the duplicate-forwarding paramenter prevents a packet from looping in the network infinitely long, which may happen in the presence of noise; if a node does no longer forwards a packet because of the parameter, the destination is not reached but on hte positive side the network load incurred is limited.

sensor R_{i+1} is

$$P(R_i, R_{i+1}) = \begin{cases} \frac{x_{i+1}}{\sum x_j} & \text{if } d(R_i, D) \leq d(R_i, S) \\ \frac{1/y_{i+1}}{\sum 1/y_j} & \text{if } d(R_i, D) > d(R_i, S) \end{cases} \quad (4)$$

where x_i is the number of shortest paths to the destination from sensor R_{i+1} and y_i is the number of shortest paths from the source to sensor R_{i+1}, and the sum is over all neighbors of sensor R_i. The design goals from [13] are different from ours: [13] designs a protocol that spreads the load across all equivalent paths (over many packet transmissions) on a grid topology; we want a protocol to be robust to arbitrary network topology and misinformation. In the absence of noise, this protocol performs as well as the shortest-path with the added feature of load-spreading-on-average, but having only one packet on the network at any time makes it inherently sensitive to misinformation. SHORTEST PATH COUNTING has again the duplicate-forwarding parameter, which can be set by a system designer.

Flooding: FLOODING is a protocol that requires no network knowledge and has no parameters: every node that receives a packet retransmits it once to all of its neighbors. Subsequent copies will be dropped, and the result is that every node on the network in the same component as the source will transmit the packet one time.

Gossiping: A simple parametric extension of FLOODING is PURE GOSSIP, which assigns a retransmission probability to a packet at the source, and every sensor to receive the packet will choose to retransmit with that probability. It tends to have a percolation behavior, in that for a given retransmission probability it is most likely that either very few devices on the network will receive the packet, or most devices will receive the packet. The retransmission probability p is a parameter that can be set by a system designer.

All protocols that we consider can be classified as to whether they use network information and whether they are single-path or multi-path protocols. Table 1 gives an overview.

4. ESTIMATING GLOBAL INFORMATION

The retransmission probability functions for DESTINATION ATTRACTOR and DIRECTED TRANSMISSION depend on knowing the hop-distance $d(R_i, D)$ from the node R_i to the destination D, distance $d(S, D)$ from source to destination, and time step i. The time step information is very simple to compute: we add a field into the header of the packet that contains the number of hops that it has already taken; this value is initialized to zero by the source and is incremented by every node that retransmits the packet. Computing hop-distances is harder: as the nodes are mobile, the distance values may change over time, but we assume that these changes will not be very substantial during the delivery of a single packet to the destination. Nevertheless, we cannot assume perfect knowledge as computing this distance takes several time steps such that the computed information might be outdated once it reaches the nodes. No matter what method we choose to estimate these distances, we may have wrong values. In our simulations, we model this issue by introducing noise such that a node will be likely to either underestimate or overestimate its current distance to the destination. Having global information and thus the tedious task to find estimates adds algorithmic complexity to our protocol, which is the price we pay for obtaining improved performance. Several methods could be developed to estimate $d(R_i, D)$. We outline only one of them that we believe to be one of the most light-weight methods possible. Other estimators may yield better results; in fact, finding improved estimator methods is a promising direction for future research.

A light-weight method of estimating the distance to the destination requires each sensor to include its current estimate of its distance to the destination in each packet that it retransmits; the destination "acknowledges" each packet it receives by sending out a packet containing the value zero in the distance-to-destination field (this acknowledgement packet is only received by the destination's direct neighbors; they do not forward the acknowledgement). Whenever a sensor R receives a packet from one of its neighbors R', sensor R checks whether its current distance estimate for $d(R, D)$ is more than one unit larger than the distance estimate $d(R', D)$ of its neighbor R'; if this is the case, sensor R updates its estimate by setting $d(R, D) = d(R', D) + 1$. In order to allow a sensor to increase its distance estimate, we assign each information a time-to-live T after which the sensor will assume it to be invalid and reset its estimate to the minimum of the most recently obtained: thus, the distance estimate of a sensor is one unit larger than the minimum of the estimates of its neighbors that the sensor received in most recent T time steps. Using this method, correct distance information propagates from the destination to the other nodes, where the wave front advances by one hop for each packet that the destination acknowledges: a sensor at (correct) distance d from the destination will be able to update its estimate to the correct value after retransmitting (or sending) d packets, if the network topology does not change during this time. Changes in the network topology get propagated equally slowly, however, if network traffic is not too infrequent, the update speed is still sufficient. Initially, the nodes set their distance estimate to some constant.

Rather than implementing this particular estimator method, we have opted to model the mobility and the inevitably resulting imperfectness of the distance information with noise in our simulations. The SHORTEST PATH COUNTING protocol uses number of shortest paths as parameter, which would require a different estimator but can be modeled through noise as well.

5. EXPERIMENTAL DESIGN

In order to assess the performance of the routing protocols described above, we analyze the results of simulations that we conducted under various levels of mis-information (noise) and through the entire reasonable parameter range[3].

We use the following setup for our experiments: 5000 sensors are spread uniformly at random in a square field, where all sensors have equal range of transmission, inducing a graph with average node degree of approximately 6.7. This average connectivity is chosen so that there is a huge connected component (and thus a high probability of having a path between source and destination), while there is still an interesting non-uniform structure in the network [8]. For each simulation run, we place the destination in the middle of the sensor field, and we randomly choose a source sensor among all remaining sensors. We generate one packet at the source that needs to get to the destination (using the tested protocol) with no competing traffic on the network. In the simulations, we ignore the specifics of all layers in the protocol stack other than the network layer, where the routing takes place, except for the medium access control (MAC) sublayer for which we assume a TDMA model. All multiple-packet protocols (i.e., FLOODING, PURE GOSSIP, DESTINATION ATTRACTOR, DIRECTED TRANSMISSION) use a duplicate-dropping functionality in our simulations:

---

[3]The tunable parameter k values for DESTINATION ATTRACTOR and DIRECTED TRANSMISSION are: 0.001, 0.01, 0.022, 0.046, 0.1, 0.22, 0.46, 1.0, 10.0, 100.0, for PURE GOSSIP the probabilities of retransmission are: 20%, 25%, 30%, 35%, 40%, 42.5%, 45%, 47.5%, 50%, 55%, 60%, 70%, 80%, 90% and for the duplicate-forwarding parameter for the single-path methods we used values of 0 (no retransmission of the same packet),3,10,30,100,300, and infinity (no duplicate-dropping)

Table 1: Protocol overview

	uses network information	multi-path
WANDERER	○	○
SHORTEST PATH	●	○
SHORT PATH	●	○
FLOODING	○	●
PURE GOSSIP	○	●
DESTINATION ATTRACTOR	●	●
DIRECTED TRANSMISSION	●	●

if a packet is received a second time, it is discarded unconditionally.[4] For the single-path methods (i.e., SHORT PATH, SHORTEST PATH, SHORTEST PATH COUNTING, and WANDERER), the value of the duplicate-fowarding parameter determines this behavior.

Each network and source-destination pair are used for only one simulation run, and 1000 runs are made for each scenario to acquire statistical results. A scenario consists of a protocol, a parameter value, and a noise level. We let each simulation run as long as there are packets on the network, or up to 5000 timesteps (this only happens with single-path algorithms). It turns out to be necessary to impose this limit because the packet may travel in loops in the network when the information is inaccurate.

An important aspect of our study is the dependence of the performance of a protocol on the accuracy of information it has available (for algorithms that use information about the network). We therefore test the protocols for several levels of "noisiness" of the information. The network information (distance to destination, etc.) is computed off-line in our simulations. The model used to add noise is such that each value representing a piece of information is replaced by an (integer) number randomly chosen from a uniform distribution with mean equal to the original (accurate) value and width of the interval equal to certain percentage of the original value. We use the width of the interval as parameter to model different noise-levels. The noise-levels we choose for each protocol are 0%, 3%, 10%, 30%, 100%, and 300% (values are bounded at 0, so that 300% noise is actually +300%/-100%). The noise is persistent in the sense that the information available to a sensor does not change with time. This is equivalent to the sensors not updating their information about the network in the course of one source-to-destination transmission. Higher noise-levels could also infer higher mobility levels: the faster nodes move, the faster and the more severely distance estimators become outdated, for all sensible (that is, reasonably light-weight) estimator methods.

We asses the performance of a protocol with a given parameter and a specified noise level by looking at the following measures:

Load: number of times any sensor transmits a packet (sum across all timesteps and all sensors)

Lag: number of timesteps it takes to deliver the packet to the destination

Fraction delivered: fraction of runs where the destination receives the data (at least one packet copy reaches the destination)

6. COMPARATIVE SIMULATION RESULTS

In order to illustrate how the multi-path protocols DIRECTED TRANSMISSION, DESTINATION ATTRACTOR, and PURE GOSSIP work, we show snapshots of a run of these protocols on the same

network in Figures 2, 3 and 4, respectively. The network is picked from the ones we used for the simulations and no noise is present in the information available to the sensors. The source is marked by S and destination by D. Time progresses from left to right, timesteps 10, 20 and 30 are shown. Black nodes are sensors that have a packet at that time step, grey nodes are sensors without packets. In the chosen run, DIRECTED TRANSMISSION imposes less load to the network than DESTINATION ATTRACTOR, which in turn is lighter than PURE GOSSIP. As we shall see below, this ranking is actually representative for the three multi-path protocols.

Protocol Comparison: We now compare the algorithms' performance with respect to lag, fraction delivered, and load. For this comparison, we fix the value of noise to 30% (we will discuss the influence of noise on the protocols later). Figure 5 shows how the algorithms compare, where the x-axis depicts the measure fraction delivered and the y-axis stands for the average load (left panel) and average lag divided by the length of the shortest path (right panel), where averages are taken over the 1000 runs.

For all tunable methods, the tunable parameter determines the measure fraction delivered and these protocols are thus shown as a line going through points which correspond to the values of the parameter used in the simulations. The parameter values increase from left to right for PURE GOSSIP and decrease from left to right for DIRECTED TRANSMISSION and DESTINATION ATTRACTOR. For the single-path methods, which all have the duplicate forwarding parameter, the values increase from 0 (for no duplicate-fowarding) to infinity (for forwarding the same packet as often as it is received) from left to right. In order to keep Fig. 5 readable, we only show the value for infinity for WANDERER and SHORTEST PATH.

The optimal point in Figure 5 (left panel) is the lower right corner, with very low load and 100% delivery. The upper left corner represents the worst performance in both delivery and load. For the single-path methods, not surprisingly, WANDERER is in the poor performance region (it usually does not reach the destination, and wanders for the 5000 timesteps until the simulation is stopped[5]) and FLOODING is in the upper right corner (delivery is guaranteed because all sensors receive the packet, but that imposes maximum load). The poor performance of SHORTEST PATH is somewhat surprising but indicates that this method is highly sensitive to noise in information (with perfect information, it is at the optimal point). For small values of the duplicate-forwarding parameter, SHORTEST PATH drops down to the origin very quickly (not shown in the figure).

SHORTEST PATH COUNTING performs adequately for this noise level and actually beats PURE GOSSIP; interestingly, the curve for SHORTEST PATH COUNTING seems to end where the curve for

---

[4]This requires that the sensors keep a list of packets already forwarded; in reality, such a list can be of limited constant length.

[5]If the duplicate-forwarding parameter is set to a small constant rather than infinity, the load decreases (not shown in the figure).
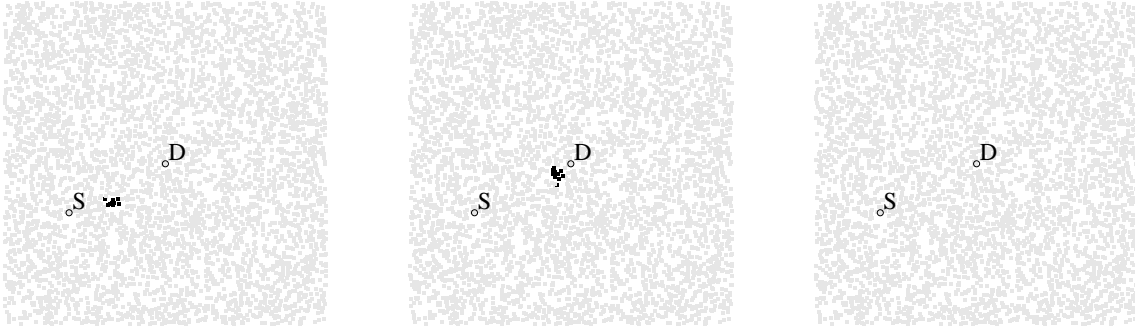
Figure 2: Sample snapshot of DIRECTED TRANSMISSION**: This method does not leave any packets on the network by timestep 20. It issues a cloud of packets and moves it towards the destination, keeping its spread roughly constant.**
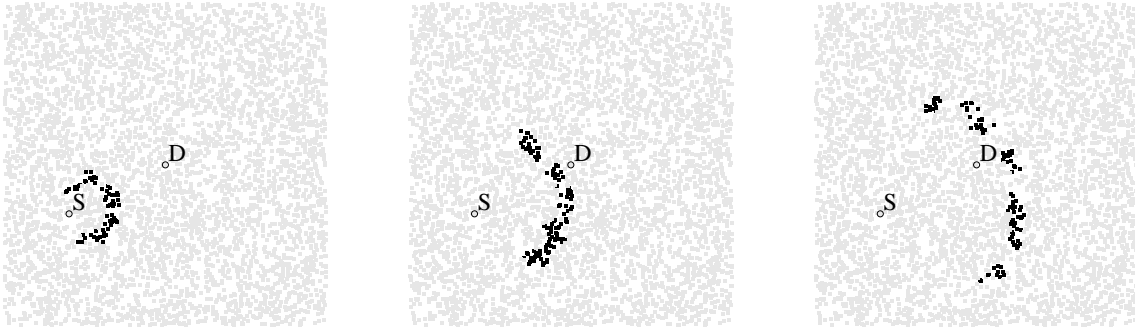

Figure 3: Sample snapshot of DESTINATION ATTRACTOR**: Limited flooding is performed within a circle of radius roughly equal to the original distance of S and D (no flooding is done in the opposite direction to D). The wavefront passes through D and slowly disappears as it gets further from D.**

DESTINATION ATTRACTOR starts. SHORT PATH performs very well on this noise level; however, the highest fraction delivered achievable is at roughly 80 %.

The fact that all lines lead from the upper right corner to the bottom left corner represents the trade-off between average delivery and average load as the value of the parameter is changed. If we restrict ourselves to the high-delivery part of the graph (75% delivery and higher), we can conclude that with 30% noise, DIRECTED TRANSMISSION is better than DESTINATION ATTRACTOR which is in turn better than PURE GOSSIP in the sense that for any given parameter to a method, there exists a value of a parameter for the other method such that it outperforms the first in *both* load and delivery. In the same sense, it can also be concluded that for this level of noise, all single-path methods perform worse than DIRECTED TRANSMISSION. For the low-delivery part of the graph, DIRECTED TRANSMISSION and SHORT PATH perform almost equally well.

The right panel of Figure 5 allows us to extend the same observations about the relative performance ordering of DIRECTED TRANSMISSION, DESTINATION ATTRACTOR and PURE GOSSIP to the lag measurements. Again, DIRECTED TRANSMISSION and DESTINATION ATTRACTOR outperform PURE GOSSIP. The single-path protocols are clearly inferior even if the duplicate-forwarding parameter is set to infinity as shown in the figure.

Dependence on Noise Level: No algorithm in our comparison in Figure 5 reaches the optimal point, which is due to the noise in information available to the sensors. As we will see in the next plots, perfect information allows some methods to reach this point, but as

the noise level increases, it is not possible to remain there. Methods that do not use any information about the network (WANDERER, FLOODING and PURE GOSSIP) are totally insensitive to noise, and their positions in the plot do not change with noise. Thus, when studying how the algorithms perform for different levels of noise levels, we can restrict ourselves to methods that actually make use of information: DIRECTED TRANSMISSION, DESTINATION ATTRACTOR, SHORTEST PATH COUNTING, SHORTEST PATH and SHORT PATH. Figure 6 illustrates this dependence: the x-axis again stands for fraction delivered and the y-axis stands for the average load, but the results are shown for different noise levels for DIRECTED TRANSMISSION (left panel), DESTINATION ATTRACTOR (middle panel), and SHORT PATH (right panel). The single-path methods SHORTEST PATH COUNTING, SHORTEST PATH and SHORT PATH all perform similarly with respect to changing noise: they move from the lower right to the upper left corner as noise increases. The speed of this transition (i.e., the sensitivity to noise) is different: SHORTEST PATH is most sensitive to noise, followed by SHORTEST PATH COUNTING, followed by SHORT PATH, which is the only method shown in the right panel of Figure 6.

Figure 6 shows that as the noise level increases, all methods shift in the direction from the optimal point towards the worst region. Both Parametric Probabilistic Sensor Network Routing methods move closer to the graph of PURE GOSSIP (which remains stationary), but retain their relative performance ordering. The speed at which particular protocols move away from the optimal point depends critically on how much and how efficiently they use the information available. The more efficiently information is used,
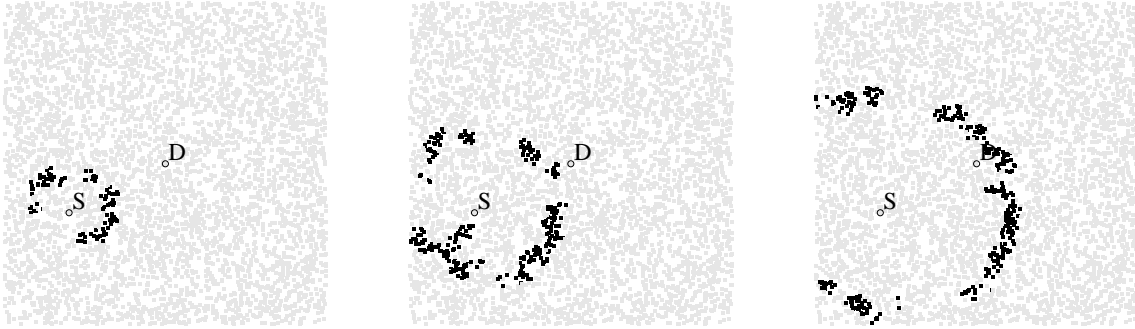
128

Figure 4: Sample snapshot of PURE GOSSIP**: Its wavefront propagates in all directions, eventually reaching all the sensors. It uses no information about the network to control the spread, although the extent of spreading is controlled by a parameter.**
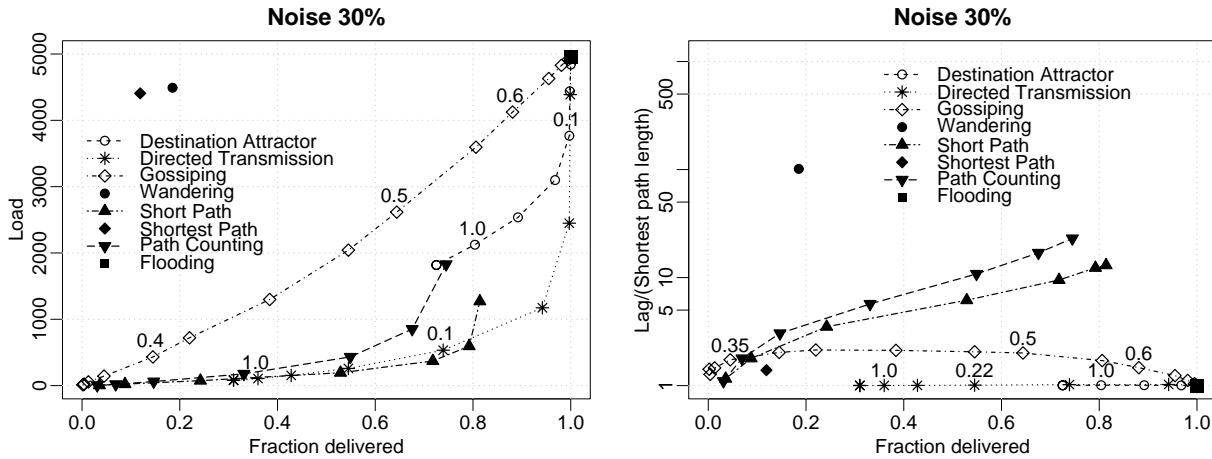

Figure 5: Load and Lag vs. Fraction Delivered is plotted for each algorithm (and various values of its parameter, if applicable) for noisiness of 30%. Note the logarithmic scale on the Lag axis. The Lag is divided by length of the shortest path from source to destination for each run to make the runs comparable.

the better the algorithm performs when the information is accurate, but the more rapidly it deteriorates when the information becomes inaccurate. This effect is most pronounced when comparing single-path to multiple-path methods. Single-path methods only rely on a single packet traveling in the network, which can easily get lost or take a long time to reach the destination if the routing decisions are based on unreliable information. In this way, they can impose a larger average load than multi-path methods, which illustrates the trade-off between sensitivity to noise and the other measures, fraction delivered and load.

Tunable Parameter Dependence: As a final study, we investigate in more detail how the tunable parameter determines the QoS measure fraction delivered. Figure 7 shows this depence for DIRECTED TRANSMISSION, DESTINATION ATTRACTOR, and PURE GOSSIP, where the y-axis is the concrete parameter value. The dependence is *monotonic* in all three cases, making it easy for a system designer to search for the parameter value resulting in the desired QoS level. The curves for all three methods have a flat region around a certain parameter value. This shows that close to that value, the probability of delivery is quite sensitive to the value of the parameter, which corresponds to the behavior of PURE GOSSIP as reported in [6]: for these critical values, either almost all of the sensors receive the packet or hardly any of the sensors receive the packet. As can be seen from the figure, this effect is less pronounced for DI-

RECTED TRANSMISSION and DESTINATION ATTRACTOR than it is for PURE GOSSIP.

Figure 7 also shows that for any given level of noise, DIRECTED TRANSMISSION and DESTINATION ATTRACTOR guarantee non-zero fraction delivered *no matter what the parameter value is*. This is very important, especially in situations when it is difficult to set the correct value of the parameter. Such a claim cannot be made for the case of PURE GOSSIP. We believe that the reason the curve for 300% crosses the other curves in the case of DIRECTED TRANS-MISSION is that this level of noise causes the average distance to be overestimated (as explained in Section 5) and thus the method generates more packets (resulting in better fraction of packets delivered, but also higher load). A similar tendency can be observed for DESTINATION ATTRACTOR, but not as strong.

In summary, our results show that DIRECTED TRANSMISSION and DESTINATION ATTRACTOR outperform all other protocols for a given quality of service level and network load incurred if a certain noise level is exceeded (30%). In fact, if we look at three characteristics of a protocol, namely load, fraction delivered, and sensitivity to noise, we see that to optimize any two is simple: FLOODING optimizes fraction delivered and is insensitive to noise, SHORTEST PATH optimizes fraction delivered and minimizes load, time-limited WANDERER achieves low load and is insensitive to
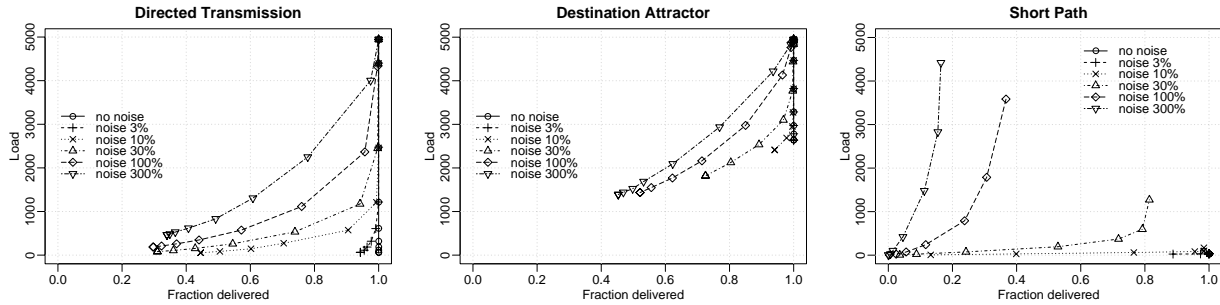
Figure 6: Load vs. Fraction delivered is plotted for noise levels of 0%, 3%, 10%, 30%, 100% and 300%, for three different algorithms.
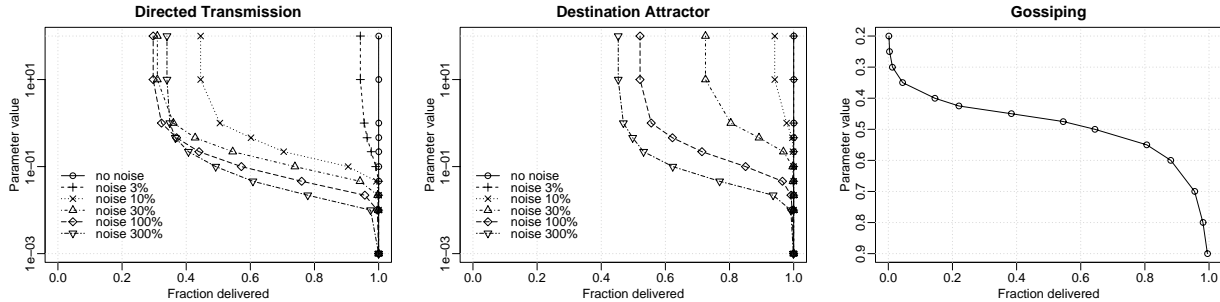

Figure 7: Parameter value vs. Fraction delivered is plotted for noise levels of 0%, 3%, 10%, 30%, 100% and 300%, for three different algorithms. Note the logarithmic scale on the parameter axis for DIRECTED TRANSMISSION **and** DESTINATION ATTRACTOR**.**

noise. Optimizing all three goals is a balancing act, and we suspect that no algorithm fully achieving all three goals simultaneously can exist. A great advantage of the tunable methods is that they allow a system designer to set the parameter so that the resulting performance meets the need of the application. As long as such a need lies on the curve plotted in the graph, the trade-off between performance in load and fraction delivered can be controlled. Among the tunable methods, DIRECTED TRANSMISSION achieves the best balance for these three goals.

As a final note, we conducted similar studies on denser networks with uniform distributions of nodes (average degree about 14), and on non-uniform spatial distributions of sensor devices. The parameter values required to achieve a certain QoS-level (Figure 7) remain almost constant for DESTINATION ATTRACTOR and DIRECTED TRANSMISSION; however, they change dramatically for PURE GOSSIP, which suggests that the Parametric Probabilistic Sensor Network Routing protocols are more stable in this sense as well. Figure 8 illustrates a single instance of a realistic (non-uniform) urban sensor network where each vehicle with a driver and all pedestrians on the streets carry a sensor device [2], each with a range of 100 meters. As with the denser uniform networks, plots for this scenario are omitted due to lack of space, but can be obtained from the authors. The results are qualitatively similar to those presented here. The principal difference lies in the ability of the freeway to guide packets through the network, partly because most nodes are along the freeways, resulting in an improved performance with all protocols for the same average node degree and average path length as in the uniform case.

7. CONCLUSION

We have proposed the routing protocol family Parametric Probabilistic Sensor Network Routing, whose members are character-

ized by a description of their retransmission probability function that a sensor uses to decide whether it forwards a received packet. Our simulations have shown that the two Parametric Probabilistic Sensor Network Routing protocols DESTINATION ATTRACTOR and DIRECTED TRANSMISSION can guarantee any quality of service level (measured as fraction delivered) even in the presence of highly noised network information. Our results also suggest that multi-path methods outperform single-path methods in the presence of noise. Our results serve as a proof-of-concept for the basic idea of Parametric Probabilistic Sensor Network Routing: making the retransmission probability depend on network information parameters is a promising step towards more robust routing protocols in sensor networks.

There is no reason to believe that DIRECTED TRANSMISSION is the best protocol from the Parametric Probabilistic Sensor Network Routing family. In fact, various points remain open for future research in the area of parametric sensor network routing protocols. Designing a protocol variant that outperforms DIRECTED TRANSMISSION is the most prominent challenge. Changing model assumptions to make the scenarios even more realistic is another research direction that has several aspects. Different network topologies, for instance, can result from the same node positions using different radio ranges, or by changing the density of nodes with the same power. Preliminary analysis, including that on the realistic urban structured distributions mentioned in the previous section, indicates that the results are similar, but we have not yet proven any general results independent of network topology. Also, there are many different noise models that can be tested, especially absolute bounds instead of relative bounds, or an information propagation model, in which noise on individual nodes is correlated to noise on its neighboring nodes. Although mobility can be considered to be a type of mis-information, this should be investigated separately. While we assume that these results apply to arbitrary
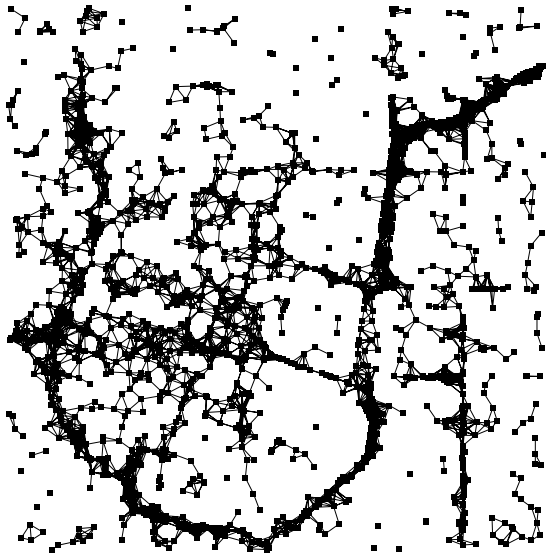
Figure 8: Realistic urban sensor network: 1750 mobile devices in 8.5 km^2 of downtown Portland.

geometric random graphs, there are some peculiarities to grids that may warrant special attention.

We are currently building a full-fledged simulator for large-scale networks; we plan to implement Parametric Probabilistic Sensor Network Routing protols in this tool, which will allow us to take into account interactions among different protocol stack layers as well as provide a more realistic framework for comparing methods to estimate global information.

8. REFERENCES

[1] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, Erdal Cayirci; *A Survey on Sensor Networks*; IEEE Communications Magazine, August 2002.

[2] C.L. Barrett, M. Drozda, M.V. Marathe, S.S. Ravi, and J.P. Smith. *A Mobility and Traffic Generation Framework for Modeling and Simulating Ad Hoc Communication Networks*; in press, Scientific Programming 2003.

[3] J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva. Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. *Proc. 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking,* ACM, Dallas, TX, October 1998.

[4] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, D. Terry; *Epidemic algorithms for replicated database maintenance*; Proceedings ACM Symposium on Principles of Distributed Computing (PODC), pp. 1–12, 1987.

[5] Deepak Ganesan, Ramesh Govindan, Scott Shenker, Deborah Estrin; *Highly Resilient, Energy Efficient Multipath Routing in Wireless Sensor Networks*; Mobile Computing and Communications Review, Vol. 1, No. 2.,

[6] Zygmunt Haas, Joseph Y. Halpern, Li Li; *Gossip-Based Ad Hoc Routing*; Proceedings of INFOCOM 2002, 2002.

[7] David Kempe, Jon Kleinberg, Alan Demers; *Spatial Gossip and Resource Location Protocols*; Proceedings of 33rd ACM Symposium on the Theory of Computing (STOC), 2001.

[8] L. Kleinrock and J. Sylvester; *Optimum Transmission Radii for Packet Radio Networks or Why Six is a Magic Number*; Proc. IEEE National Telecommunications Conference, Birmingham, AL, pp. 4.3.1-4.3.5, Dec. 1978.

[9] F. Kuhn, R. Wattenhofer and A. Zollinger, Worst Case Optimal and Average Case Efficient Goemetric Ad-hoc Routing *Proc. 4th ACM International Symposium on Mobile Networking and Computing,* pp. 267-278, 2003.

[10] Jun Luo, Patrick Th. Eugster, Jean-Pierre Hubaux; *Route Driven Gossip: Probabilistic Reliable Multicast in Ad Hoc Networks*; Proceedings of INFOCOM 2003, 2003.

[11] M. Mauve, J. Widmer and H. Hartenstein, A Survey on Position Based Routing in Mobile Ad-hoc Networks, *IEEE Network*, Nov. 2001.

[12] S. Ramanathan and M. Steenstrup, A survey of Routing Techniques for Mobile Communication Networks. *Mobile Networks and Applications*, Vol. 1-2, 1996, pp. 89-104.

[13] Sergio D. Servetto, Guillermo Barranechea; *Constrained Random Walks on Random Graphs: Routing Algorithms for Large Scale Wireless Sensor Networks*; Proceedings of 1st ACM International Workshop and Wireless Sensor Networks and Applications (WSNA 2002), 2002.

[14] *Routing in Communication Networks,* M. Steenstrup, Editor, Prentice Hall, New Jersey, 1995