Localized Probabilistic Routing for Data Gathering in Wireless Ad hoc Networks

Eun-Sook Sung, Miodrag Potkonjak
Computer Science Department
University of California, Los Angeles, CA 90095
{eunsook, miodrag}@cs.ucla.edu

Abstract— **Longevity is a critical requirement for wireless sensor networks to application domains like environmental monitoring. Deterministic communication protocols that assume that packet loss implies congestion suffer due to the non-negligible probability of a wireless link being lossy even at very short distances. In this paper, we propose a localized probabilistic routing scheme aimed at extending network lifetime while coping with lossy links. The proposed protocol is informed by our study of the centralized algorithm that optimally computes paths. We obtain the most relevant properties from the centralized optimal routing algorithm and present their statistical models. The proposed routing scheme, based on the statistical models, is a promising step towards an optimal solution.**

Keywords: decentralized probabilistic routing, wireless ad-hoc networks

I. INTRODUCTION

Ad hoc networks have been an active research area due to a tremendous number of potential applications. Wireless sensor networks are an important type of ad hoc networks in which a large number of inexpensive nodes equipped with a variety of sensors are deployed in an area of interest to monitor environments. Such networks permit remote regions to be monitored; however, they are fundamentally limited by the size of the batteries deployed. Thus, energy-efficient operation is a design goal for wireless sensor networks. One of the important criteria in designing a communication protocol for ad hoc networks is scalability, which allows for good performance as the network size grows. Scalability can be achieved by protocols that are decentralized because forwarding decisions can be made locally, avoiding the cost of collecting and distributing global knowledge of networks.

We aim to develop a decentralized and energy-efficient protocol for use in periodic data collection scenarios with a single data collection point, the sink node. To this end, we first study a centralized optimal routing protocol in terms of maximizing network lifetime, because we believe that a better understanding of centralized optimal routing schemes can inform the design of localized algorithms. There can be two approaches to solve the optimal routing problem of maximizing network lifetime. The first approach casts the problem in the light of graph theory as a min cut/max flow problem, which provides insight into the maximum number of times each link in the network can be used. However, the min cut/max flow algorithm assumes that the links that connect nodes perfect. Clearly, this assumption is inconsistent with the recent findings that wireless links are lossy and dynamic, rather than perfect. A reduction to the min cut/max flow problem with imperfect links is ambitious and intractable.

The other approach we can take is to solve the problem by using linear programming techniques. We formulate the maximization of network lifetime as a linear programming problem, and then solve this problem using a free software package, Lpsolve [11]. Although the variables for the incoming and outgoing flux are not integers, the solutions to the variables give us reasonable bounds. After we examine the behavior of this centralized optimal routing algorithm in medium sized networks, we demonstrate how we can apply those lessons learned to a completely localized routing protocol. To do this, we build statistical models for the behavior of the centralized optimal routing and use them for our localized routing protocol.

A common property of routing trees used for periodic data gathering is that nodes near the root will inevitably consume more energy and deplete their energy supply sooner than other nodes. When a node near the sink fails, it is important to quickly detect the resulting broken links and compensate for them, which is not possible in many application scenarios. Meanwhile, provided that all the packets can successfully reach the data sink, we can give nodes the opportunity to choose their next hops probabilistically each time their packets have to be sent. The motivation behind this probabilistic approach is that choosing a different next hop each time can help balancing the data collection load. However, to save energy and to find a desirable path, neighbors that are connected by high quality links and that can guide fast to the sink should be shown a preference. There are several ways to determine the priorities of one-hop neighbors, using different probabilities, such as [3].

Our contributions are three-fold: Firstly, we propose a linear programming (LP) data gathering formulation to obtain the optimal execution solution while maximizing operational network lifetime. We then analyze the optimal traffic patterns and build statistical models for efficient design and operation of a wireless distributed system. Based on those models, we introduce a probabilistic routing protocol suited for lossy wireless ad-hoc networks. Our protocol is fully decentralized. Lastly, we show that the properties identified and statistical models developed can be applied to other communication tasks as well.

978-0-7695-3649-1/09 $25.00 © 2009 IEEE
DOI 10.1109/CNSR.2009.62

II. RELATED WORK

A probabilistic routing called PGR is proposed in [6]. In PGR, the next hop neighbor is chosen among a set of neighbors that are within a small angle of a line that directly connects the node to the sink. A choice is made among those candidates by considering their remaining energy as well as the link quality. This approach is similar to a minimum energy routing protocol like the Dijkstra's algorithm, in the sense that PGR prefers the most stable links requiring the fewest number of retransmissions. In particular, the authors exclude the neighbors whose angle is great from the next-hop candidates so that a packet will never be sent in the backward direction. However, due to the dynamics of wireless links, a neighbor that is further from the sink than the current hop may be a more desirable candidate than a node that is geographically closer to the destination. In [3], an energy-aware probabilistic routing protocol for low power ad hoc networks is presented. The operation of the protocol mainly relies on the information from the setup phase in which a node initiates a route request (by flooding) and a routing table is built up by finding all paths and their energy cost. When a node is ready to send its packet to the destination, among multiple routes, a route is probabilistically selected, based on the amount of energy required for that path and the remaining energy at each node so that the route with energy rich nodes is more likely chosen. A disadvantage of this method is the overhead of the setup phase. Also, to keep all the paths alive, flooding is continuously required in the maintenance phase. In [2], a parametric probabilistic routing protocol is initially introduced. The original idea of probabilistic routing is to limit the flooding by using a concept of a probability in forwarding. The authors examine to take a few parameters into account, such as destination attractor and directed transmission. Their proposed retransmission probability, used for flooding, simply increases as the packet is getting closer to the destination. They inform that protocols using network information perform better than protocols that do not, although there is strong noise. This approach is similar to ours in that our scheme takes advantage of network information. Probabilistic flooding (PF) routing is introduced in [7]. In PF, a node that has received a packet independently decides whether the node relays the packet or just discards it. The probability of forwarding a packet in PF protocol is set to some constant. Lower this probability restricts the amount of flooding, but also reduces the number of paths that connect the source to destination. Nonetheless, our approach is different from them. Our routing scheme probabilistically selects a next hop neighbor based upon our novel statistical models.

III. SYSTEM MODELS AND ASSUMPTIONS

We consider a single-tiered wireless ad hoc network consisting of many sensing nodes deployed and one data sink node whose location is away from all the other nodes. Each node is equipped with a low power radio and acts as both a data source and data forwarder. Also, we assume that the network is static after deployment.

Network lifetime: The operational lifetime of a network is defined to be the maximum time duration during which all nodes in the network are operational, i.e. the time until the first node's battery is depleted. We use this definition because a network monitoring application may be impaired as nodes fail.

Link model: We use the cumulative distribution function (CDF) of low-power wireless links developed in [9]. Since wireless links are dynamic and asymmetric, a simplified link propagation model or unit-disk model cannot be used. Furthermore, lossy wireless link properties can greatly affect the performance of routing algorithms and thus should be considered in designing communication algorithms. In our model, forward and backward links are estimated separately, and reception rates are not fixed according to geometric distances, but are estimated based on the CDF by using the inverse CDF method. We believe that our statistical link model is realistic enough to be used for ad hoc wireless network simulations.

Energy model: The components that consume power on a node include the radio, the microprocessor, and the attached sensors. We assume that the energy spent in a node mainly depends on the communication between nodes. We estimate energy consumption using the measurements proposed in [5]. In this formulation, the power spent in transmission is referred to as tx and set to 1.9W, the power spent in reception is referred to as rx and set to 1.5W, and the power spent in idle time is referred to as idle and set to 0.75W. We have implemented a simple energy model in which every time a packet is transmitted, the total energy of the node decreases by the value tx multiplied by the transmission time [8]. The same formula is applied to determine the decrease in battery upon reception (using rx), or when idle (using idle.)

Communication model: In the communication in this work, when a node sends a packet, it expects the acknowledgement from the receiver in order to ensure that the packet is correctly received. In addition, the waiting/idle period is a few times longer than the sum of the time required for the receiver to receive the packet and to transmit an acknowledgement back [8]. After the waiting period interval, if the sender does not receive an acknowledgement from the receiver, the sender retransmits the packet.

Traffic model: Each node is assumed to generate data at a (low) constant rate. For the basic data-gathering task, each node transmits one packet to the data sink node in a round robin fashion (one transmitter at a time.)

Assumptions: A set of nodes is deployed at random over a two dimensional plane. Each node knows its coordinates within the plane, equipped with a GPS device or a localization protocol such as the signal-strength based localization scheme. In addition, each node knows the

location of the destination node, using a location database or a location service scheme.

IV. PROPOSED PROBABILITY MODELS

In this section we first implement an LP solution of a centralized optimal way that maximizes the lifetime of a network in which all packets are destined to a single data-gathering node. We then examine several beneficial characteristics of the centralized optimal solution. After highlighting those beneficial aspects, we present a few statistical models for those properties that are the most relevant for an efficient routing protocol. We then attempt to follow/imitate the optimal centralized behaviors in our localized routing protocol.

A. Problem formulation as an LP solution

Our goal is, firstly, to solve the optimal routing problem in a centralized manner, then, to abstract a set of important properties that we observe in the optimal routing algorithm, and to model the relationship between the important properties. An LP formulation whose purpose is to maximize network lifetime provides the routing solution in a centralized optimal way. We propose an Integer Linear Programming formulation that generates the optimal solutions for small networks. We then convert this formulation to an LP formulation in order to address medium sized networks. Although, due to the conversion to LP, this formulation does not provide the exact optimal solution, it can still provide a reasonable lower bound for the estimated network lifetime. We finally collect datasets that include information about how a node selects its next hop neighbor in networks of varying densities.

Problem formulation: To maximize network lifetime we attempt to minimize the maximum energy consumption value among all nodes within the network. The constraints of this problem can be written informally as follows. First, the packets should be routed to the data sink node. Second, an incoming packet must leave a node unless the node is the destination for the packet. Third, each node spends its energy when transmitting and receiving a packet, as well as being idle.

We refer to the data sink as N_0, and assume that each node in the network of size n is numbered as N_1 to N_n. The primary variables of the formulation are the number of packets received by node N_j from node N_i and is referred to as flux, x_{ij}. Other notations are as follows: C_i is the required energy at node N_i, $OutMsg_i$ is the total number packets transmitted from N_i, and $InMsg_i$ is the total number packets received by N_i. The objective of this optimization problem is to minimize the energy requirement of the critical node, which is the most energy-depleted node in the network.

Objective Function: Minimize Cmax
subject to

$$\text{Cmax} >= C_i, \text{ for } i \in \{1,\ldots, n\} \quad (4.1)$$

$$OutMsg_i = 1 + InMsg_i, \text{ for } i \in \{1,\ldots, n\} \quad (4.2)$$

$$InMsg_0 = n \quad (4.3)$$

$$\sum_i x_{ij} = InMsg_j, \text{ for } j \in \{1,\ldots, n\} \quad (4.4)$$

$$\sum_j x_{ij} = OutMsg_i, \text{ for } i \in \{1,\ldots, n\} \quad (4.5)$$

$$C_i = \sum_j x_{ij} \times \text{energy consumption during the communication between } N_i \text{ and } N_j, \text{ for } i \in \{1,\ldots, n\} \quad (4.6)$$

The constraint shown in (4.1) indicates that the Cmax is the required energy amount at the critical node such that the objective function is to minimize this Cmax. The constraints shown in (4.2), (4.4) and (4.5) denote the sum of flux at each node conserved, which is for the second requirement. The constraint shown in (4.3) indicates, for the first requirement, the requirement of the flow of each packet; the expected net flux to the destination (data sink node) should be equal to the number of nodes in the network since the network traffic is data gathering and each node sends a packet only to the destination. The energy consumed C_i includes the energy spent in transmission and reception of packets as well as idle time at node N_i. The constraint shown in (4.6), for the third requirement, denotes the energy consumption as a function of our energy model and the number of packet transmission. Because the optimized variables, x_{ij}, and C_i, are allowed to be any non-negative real number, the problem defined above is a linear programming problem, which can be solved by the simplex method or any other polynomial complexity methods. Now, our problem is reduced to strategically finding routes that minimize Cmax.

B. Analysis of datasets from LP solutions

We briefly discuss the dataset used in the analysis. We outline only a few characteristics that we believe to be most intuitive and relevant, although further exploration of better characteristics is needed and is a focus of our future work.

Let S denote a message sender, and D denote the final destination node. For simplicity, S is assumed to be located at the origin, and D is located at some positive x coordinate. As S changes (as the message propagates through the network) the coordinate system can be changed to maintain these conditions. The neighbors within the one-hop neighborhood of S, along with their associated link qualities, and location information are collected. From the location data a *reduction in distance to the destination*, difference between distance from S to D and distance from one-hop neighbor to D, is also calculated. The intuition concerning the chosen data is that the node that lies on a great reduction and that is connected by high quality link should often be used for a forwarder with high probability, while the node that brings in a smaller reduction should have a smaller probability of being used. Based on this intuition, we examine the following three characteristics.

1) Relative coordinates: We calculate the relative coordinate of a next hop with respect to a message sender and assign the coordinate the number of packets transmitted.

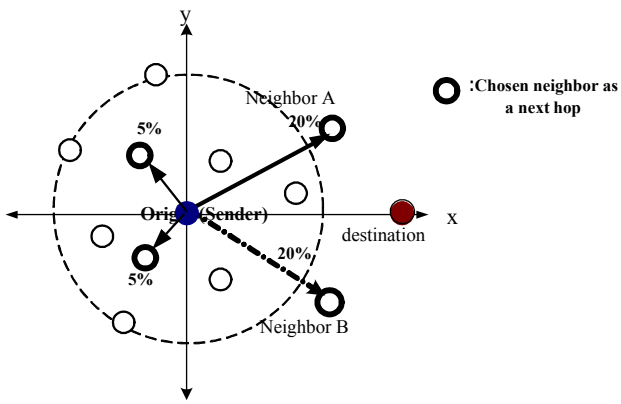

Figure 1. An illustration of choosing the next hop along the forwarding path

This information can be used to inform the decision of choosing a next-hop. As previously mentioned, the coordinate system of all the nodes within one-hop is translated and rotated so that the sending node S lies at the origin. Using the total number of messages transmitted from the sending node and the number of messages sent to the node at a specified coordinate, we can calculate what percent of messages node S is forwarding to a node at a specified coordinate. Thus, the relative distance property provides an idea of where to send and how frequently. Importantly, this property indicates which regions are "popular" areas and which are not. Intuitively, the model of relative distance exhibits a symmetric shape along the x-axis and higher likelihood as the x coordinate grows. This symmetric shape arises from the fact that the goal of the centralized LP solution is to minimize the maximal energy consumption, which attempts to balance the forwarding load. Depicted in Fig. 1, if the sender forwards 20 percents of its load to node A, then node B becomes a likely candidate for its next-hop to forward the remaining 20%.

2) Forward Link quality and reverse link quality: We examine the reception rate of the LP solution data. Note that link quality is asymmetric. Thus we also collected reverse direction reception rates, as well as forward reception rates.

3) Distance to each possible next hop, along with the distance from each possible next hop to the destination: Note that the distance from a source to the destination can be divided into two components: first, the distance from a source to a neighbor, and second, the distance from the neighbor to the destination. Although it may seem advantageous to choose a neighbor that is closer to the destination, this heuristic may not perform well. It might require a large amount of energy to take the packet to the next-hop neighbor. Rather than use this heuristic, we investigate a reduction in distance to the destination defined before.

In building models, we investigate the relationship between 2) and 3). In the model, the y-axis corresponds to the link quality between a sending node to the next hop, and the x-axis denotes a reduction to the destination.

C. Statistical non-parametric regression

When we build our statistical model for packet forwarding as a function of distance and link quality, we are faced with two major difficulties: (i) we do not have data for each possible distance and link quality, and (ii) for a given distance and given link quality we do not have enough collected samples. For these reasons, we take an approach of non-parametric regression. A non-parametric regression is a technique of data modeling in which the predictor is constructed from the data that help explaining the response variable of interest according to the available data. Kriging [10] is one such technique from the class of linear estimator, and it has been used to estimate spatially varying phenomenon with a limited number of samples. The estimators from Kriging are optimal in the sense that the variances of estimators are minimized. For our purposes, we assume second order stationary, which means that the correlation between two random variables at different locations only depends on the distance between the locations and not on the locations themselves. The algorithm for Kriging is to add sampling locations one at a time, choosing the best possible sampling location in an iterative fashion. The method derived from the algorithm is approximately a space filling design.

The model construction using Kriging is based on the estimation on the correlation between the values at different locations. In our model, the semivariogram is taken from a family of exponential functions, and thus the further two locations are apart from each other, the less correlation between corresponding values. The distance in which the value is predicted to the closest sample determines the estimation variance because to minimize the covariance the distance to a sample must be minimized. We construct our statistical models using Kriging in which the data collected is a partial realization of random variables of a distance and link quality. Applying the Kriging method to resolve the earlier difficulties mentioned, we achieve the smoothness property from the distance and link quality model. We used an open source package, R project [12], for our figures. In summary, using this non-parametric regression method, if two neighbors of a sending node have similar distance and link quality, the probabilities of choosing them as a next hop are similar.

D. Probability model construction

We present the results of our data analysis. For the sake of brevity and clarity, we focus on a few specific properties: the relative coordinate of a node selected as the next hop neighbor and the relationship between distance and link quality. The first task is to observe the outbound traffic to the next hop neighbor chosen by a sending node.

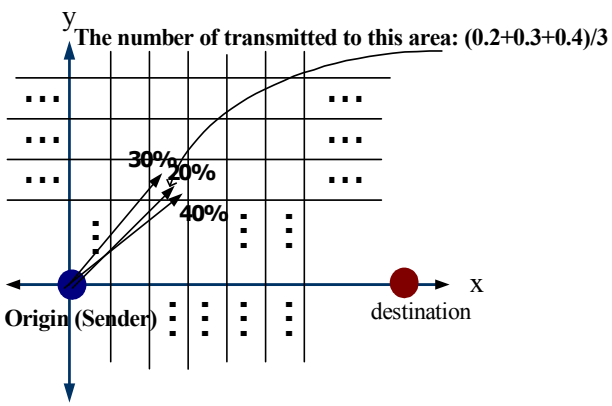

Figure 2. Normalized percentage of packet transmissions

Outbound traffic model: Relative coordinates

In our analysis, we decompose a network embedded in a plane into a number of small areas of an equal size, like a grid, and then average the collected datasets (*how many times packets are sent to the specified coordinate value*) over a same sized area, as shown in Fig. 2. The details of the way in which we analyze datasets are as follows. When we examine the routing table from the LP solution, remember that the sender is at the origin, and the destination lies along the positive part of the x-axis. We calculate the relative coordinate values from the geographical location data and collect the associated numbers of packet transmission sent to the coordinates. In other words, for the coordinate of a next hop, we have a frequency value, which denotes how many times a message sender has transmitted packets to the node at the coordinate point. This frequency is then weighted by dividing the rate of the number of neighbors. Specifically, when calculating the final frequency value, we give a certain weight to its next-hop, and the weight corresponds to the number of possible one-hop neighbor options of the message sender for its next-hop. The motivation behind this approach is that if there are many possible next hop options and among them, one is selected, then the selected neighbor (by the LP solution) should be a very desirable neighbor as a next-hop. We thus attempt to give the next-hop a preference, a high weight. In the same manner, if there are only few possible next-hop options and one is selected, the selected neighbor is given a small weight. The frequency value is, finally, divided by the total number of outbound messages from the sender. This final value provides each coordinate the probability of being selected as a next hop. Next, we use a window for all points within a similarity range of x and y coordinates. The probability values within a similar range are averaged, as depicted in Fig.2. In summary, the normalized relative distance provides an idea of where and how frequently a message is sent by a sender at the origin.

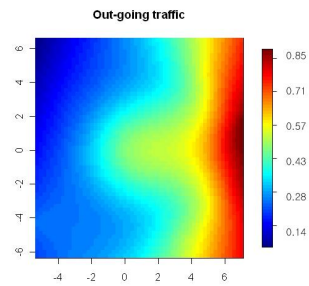

Figure 3. Frequency model for outbound (flowing-out)

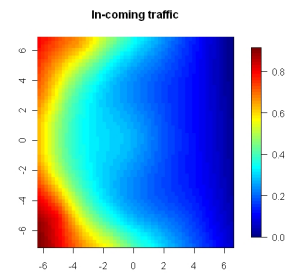

Figure 4. Frequency model for inbound (flowing-in)

In Fig.3, we present our first proposed model, based on relative out-bound traffic analysis, which illustrates how likely a node at a particular point may be selected as a next hop. The (x, y) coordinate corresponds to a relative coordinate value with respect to the message sender, and the color value denotes the normalized frequency of the area corresponding to (x, y), and the frequency is calculated by the number of times a message is sent to the node from the sending node, divided by the total number messages transmitted by the sending node, and then divided by the rate of the number of one-hop neighbors of the sending node. The frequency values are finally averaged over a same sized area.

The calculation can be shown as ($\sum_{area} \frac{x_{ij}}{OutMsg_i}$ × number of one-hop neighbors of N_i) divided by the number of samples in the area, where x_{ij} is the number of packets sent from N_i to N_j, and $OutMsg_i$ is the total number packets transmitted from N_i .

In Fig.3 and Fig. 4, a sender is located at the origin (0,0). In Fig.3, we observe that the most popular region as a next hop is the 1st and 4th quadrant, close to the x-axis and far away from the origin. In other words, when a node selects its next-hop, the nodes that are distant from the origin but close to x-axis are primary candidates. We also observe that the model shows somewhat symmetric behavior along the x-axis, which agrees with our intuition.

Using the above, we now have a probability model by which a node is assigned a likelihood of being chosen as a next hop, by calculating the node's relative x-axis and y-

axis distances with respect to the sender. Likewise, the inbound flux is analyzed, and its model is presented in Fig.4. The figure shows that those nodes that are far to the left of the origin frequently transmit their packets in the optimal LP routing protocol.

Next, we examine the probability model for the distance between a message sender and its next hop versus the link quality that connects them.

Model for Distance vs. Link quality

Here, we consider the geographical distance between a source (S) and the destination (D), denoted by (S,D), as well as the distance between the next-hop (NH) and the destination, denoted by (NH, D). A reduction in distance to the destination is calculated by the difference between (S,D) and (NH, D). From here, unless stated, a distance denotes the reduction in distance to the destination.

This distance information can be used to determine whether the candidate neighbor is close to the final destination or not. Link quality denotes the reception rate between a source and a one-hop neighbor node. We construct a model whose inputs are the distance (*reduction to the destination*) between S and a candidate next hop neighbor, and a measured reception rate between the two. The output is the probability of being chosen as the next-hop. Using a method similar to our outbound traffic analysis, we analyze our link quality datasets, which contains distance and measured link quality. To normalize our estimates, we use the same windowing method for all points that are about the same distance apart. For each distance we create a model for reception rate and the distance between the sender and receiver. Data samples where both the link quality and the intervening distance are similar are averaged. The model built on these averaged values indicates how likely a node is to be selected as the next-hop given the distance and the link quality between the sender and the candidate next hop.

In Fig 5, we present the model of the relation between distance and link quality. The x-axis denotes the distance, which is a reduction in distance to the destination, and the y-axis shows the link quality between a sender and the next hop. The color value (z-axis) denotes a normalized frequency of the corresponding area containing (x, y). The frequency is calculated by the number of times of being sent to the node at the corresponding coordinate value, divided by the number of total transmitted messages by the message sender. The frequency values within a same sized area are then averaged. The calculation can be shown as

$$\left(\sum_{area} \frac{x_{ij}}{OutMsg_i}\right)$$ divided by the number of samples in the area,

where x_{ij} is the number of packets sent from N_i to N_j, and $OutMsg_i$ is the total number packets transmitted from N_i. Above, we expected the model would exhibit higher likelihood as both x and y values are larger. Also note that distance is *negative* when the sender decides to choose a next-hop that is geographically further from the sink.

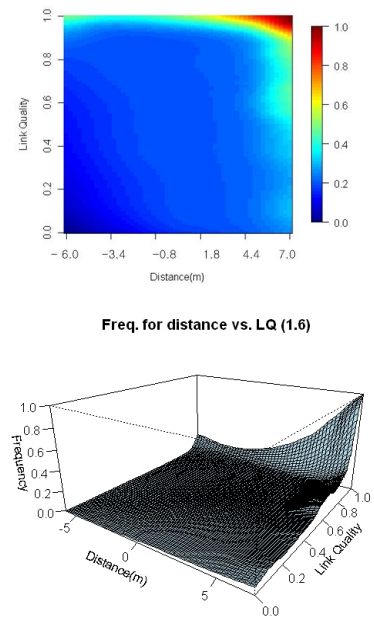

Figure 5. Frequency model for Link quality vs. distance in high-density networks

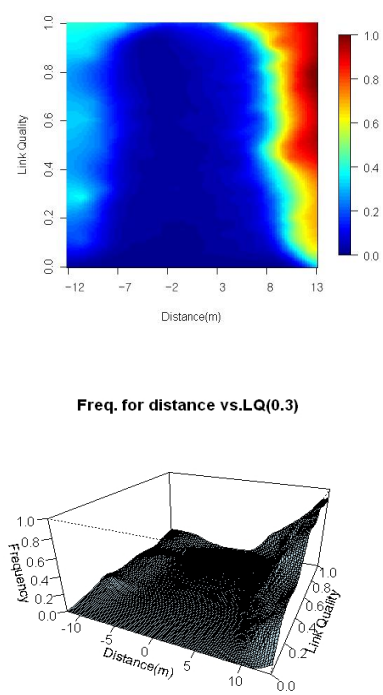

Figure 6. Frequency model for Link quality vs. distance in low-density networks

We observe, from Fig. 5, that the nodes that are located at a relatively great distance from the sender and that have good link quality are preferred most of time as a next hop. The next preferred neighbors are the nodes with good link quality but a medium distance from the sender. Unlike [6], the neighbors that are further from the sink node than the current hop are selected with even low probability. In particular, we see that nodes always choose their one-hop neighbors that are connected with good link quality

In Fig 6, we present the probability model for distance versus link quality for the case of low-density (in the case of 0.3 node per square meter) networks. The nodes in a low-density network generally prefer those neighbors with a great distance and good link quality as their next one-hop neighbors, just like the nodes in high-density (more than one node per square meter) networks do. However, it is more difficult to find a one-hop neighbor that is at a great distance and that is connected with good link quality. For this reason, it selects its one-hop neighbor that is far way and that is connected with medium quality link. Furthermore, in networks with low node density, a neighbor that is further from the sink than the current sender and that is connected with medium link quality is quite often selected, unlike when the network density is high. Also, importantly, we observe that nodes generally prefer neighbors that are a great distance away, which allows for quick advancement towards the destination.

V. LOCALIZED ROUTING PROTOCOL

In this section, we propose a simple but useful routing scheme to maximize network lifetime, which is one of the crucial design objectives in developing routing protocols for wireless ad hoc sensor networks. In our protocol, to route packets for the data sink node, the sensing nodes use only local information. Here, we leverage those lessons learned from our study of the centralized optimal routing protocol. In our proposed routing algorithm, a node assigns probabilities to the set of candidate nodes based upon characteristics measured by the sender. In this way, we achieve a fully decentralized algorithm that closely approximates the centralized solution.

Neighbor discovery and maintenance

After deployment, each node collects information on what nodes are within one hop, and what link quality connects them. This is done by having nodes send “hello” messages, which contain the sender's location and energy information, in addition to allowing the sender and receiver to measure the link quality that connects them. After this probing stage, a node knows the set of neighbors that it can communicate with and build its neighbor table. This neighbor table is refreshed periodically to ensure that *bad* links are excluded, which would be handled by a neighborhood service or link layer facility, and the badness can be defined by the application. Each entry in the neighbor table contains the neighbor id, the neighbor location information, the corresponding link quality, and the energy information of the neighbor. When node S wants to transmit a message to the data sink node D, it looks in its neighbor table.

Although probabilistic routing solves the cycle problem, to ensure that a packet does not follow a large cycle we use the same heuristic adopted in [8]. A node first compares its energy level locally to its neighbors. If its energy level is low, the node transmits the packet to the neighbor with the best link quality to minimize its energy expenditure. In the case of a routing loop, the participants will gradually have their batteries drained, which will then cause them to use another path, breaking the routing loop. If the energy of node S is not low among its neighbors, then S proceeds to assign probabilities to each of the candidate neighbors (as described above), referencing the model we have constructed in the previous section.

Using our proposed model of link quality versus distance, a neighbor with a good link quality and greater reduction is selected as a next hop with high probability. In this way, our localized scheme attempts to maximize the advancement toward the destination while minimizing the required energy consumption at each step with high probability.

VI. EXPERIMENTATION

We simulated our proposed localized protocol using our own simulator. The results reported here are obtained by taking more than 50 simulation runs. As a point of comparison, we also simulated a minimum energy routing scheme that uses the Dijkstra’s algorithm, which finds the global minimum energy path from a source to the data sink node. In our experiments, the nodes are placed at random in a square area.

We compare these two protocols using the following two metrics:

Lifetime of a network: The operational lifetime is defined as the time it takes until the first node in the network is completely drained of energy.

Path length: The path length is defined as the number of hops a packet takes to reach its destination.

We expect that the Dijkstra’s minimum energy path routing will have a shorter network lifetime on average because it always selects the path that will minimize the total amount of energy spent, without regard for nodes along the path that have low energy. On the other hand, our proposed protocol gives nodes the opportunity to choose their next hops probabilistically each time their packets have to be sent. As a result, they choose a different next hop each time, which prevents a certain set of good links from being overly used and helps prolonging the operational lifetime. We also expect that the minimum energy path algorithm produces much shorter paths than our localized algorithm, because our algorithm does not try to optimize for path length. In fact, the path length of our protocol shows, on average, about more than three times longer than the minimum energy routing protocol.

In summary, because our algorithm selects the next hop probabilistically, it does not explicitly try to minimize the energy consumption on each path. However, a neighbor

with a good link quality and greater reduction to the destination has a higher probability of getting selected as a next hop. As a result, the proposed protocol implicitly minimizes the required energy amount locally at every hop.

Simulation results: In low-density (below than 0.4) networks, our protocol does not perform better than the minimum energy routing protocol; it achieves only about 70~80% of network lifetime of the Dijkstra's minimum energy routing protocol. One possible reason for this is that there are other beneficial properties that we could incorporate into our algorithm. Investigating more sophisticated properties is left for future work. For now, we present results from dense (higher than 1) deployments. Fig.7 demonstrates the ratio of network lifetime of the proposed localized routing protocol and the Dijkstra's minimum energy routing protocol. Based on the simulation results, our protocol exhibits consistent throughput over a range of network sizes (deployed number of nodes.) This suggests that our protocol scales as the network size increases. Despite being a localized algorithm, our median network lifetime is more than 30% better than our point of comparison that uses global information.

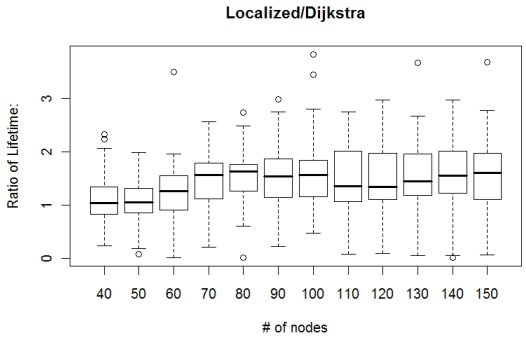

Figure 7. Ratio of lifetime (proposed localized : Dijkstra) vs. node numbers deployed

VII. CONCLUSION

We have shown that a better understanding of centralized optimal protocols can guide us towards great advances in designing localized communication protocols. Our study aims at discovery and modeling of beneficial characteristics for the purpose of enhancing the performance of routing operation. To this end, we first formulated the problem of maximizing network lifetime as a linear programming problem. The solution is used to identify properties of the optimal routing protocol that we would like our localized algorithm to approximate. We then presented an energy-aware routing protocol. Our protocol uses location information of one-hop neighbor nodes, along with link quality information, to make routing decisions. Instead of deterministically choosing the next hop, a node is assigned a probability of being used. Our approach ensures that "good" nodes along the minimum energy path are not drained of their energy too quickly, which in turn increases the lifetime of a network. In addition, our algorithm attempts to locally minimize energy consumption by selecting good links with high probability. Selecting nodes with good links contributes to energy saving, and increases the overall lifetime of the network. The other advantage of our algorithm is that it is loop-free without requiring that state be kept per route. Therefore, it reduces the amount of routing information that needs to be stored at each node. Above all, the proposed localized routing scheme is conceptually simple and can be easily used. Although many schemes have been proposed to achieve the energy savings at the link layer, by integrating link quality information into the routing layer our proposed scheme contributes to saving energy. Our simulation results show that the lifetime of a network increases significantly (by about 30%) compared with the minimum energy path routing protocol. Our protocol, based on our study of centralized routing behavior, is a promising step towards an optimal energy-efficient solution that is also localized.

REFERENCES

[1] M. Zorzi and R. R Rao. "Geographic Random Forwarding(GeRaF) for Ad Hoc and Sensor Networks: Energy and latency Performance". IEEE WCNC 2003

[2] C. L. Barrett, S. J. Eidenbenz and L. Kroc. "Parametric Probabilistic Sensor Network Routing". WSNA 2003

[3] R. C. Shah and J. M. Rabaey. "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks". IEEE WCNC 2002

[4] Z. Haas, J. Y. Halpern and L. Li. "Gossip-Based Ad Hoc Routing". Proceedings of INFOCOM 2002.

[5] Oliver Kasten. Energy consumption Model. ETH-Zurich, Swiss Federal Institute of Technology

[6] T. Roosta, M. Menzo and S. Sastry. "Probabilistic Geographic Routing in Ad Hoc and Sensor Networks," in IWWAN, May 2005.

[7] Y. Sasson, D. Cavin, and A. Schiper. "Probabilistic broadcast for flooding in wireless mobile ad hoc networks". IEEE WCNC, 2003

[8] E. Sung and M. Potkonjak, "Energy balancing routing schemes for low-power wireless hoc networks", CNSR 2009

[9] A. Cerpa, J. L. Wong, L. Kuang, M. Potkonjak and D. Estrin "Statistical Model of Lossy Links in Wireless Sensor Networks". IPSN 20005

[10] Kriging, http://en.wikipedia.org/wiki/Kriging

[11] Lpsolve, http://lpsolve.sourceforge.net/5.5/

[12] R:www.r-project.org