Providing Application QoS through Intelligent Sensor Management

Mark Perillo, Wendi B. Heinzelman
Department of Electrical and Computer Engineering
University of Rochester
Rochester, NY 14627
{*perillo, wheinzel*}*@ece.rochester.edu*

***Abstract*— Wireless sensor networks are uniquely characterized by tight energy and bandwidth constraints. These networks should be designed to provide enough data to their application so that a reliable description of the environment can be derived, while operating as energy-efficiently as possible and at the same time meeting bandwidth constraints. These goals are typically contradicting and must be balanced at the point where the application is best satisfied. In this paper, we address the problem of maximizing lifetime for a wireless sensor network while meeting a minimum level of reliability. This maximization is achieved by jointly scheduling active sensor sets and finding paths for data routing. Simulation results show that network lifetime can be significantly increased through such methods.**

I. Introduction

Interest in the use of wireless sensor networks has blossomed over the last several years due to technological advances enabling smaller devices and the realization of the potential benefit of such networks in many applications. In some situations, sensor networks may consist of sensors with overlapping coverage areas providing redundant information, providing an application with a level of reliability that is more than necessary. Rather than provide this unnecessary reliability, it may be desirable to reduce power consumption and conserve energy in these sensors to lengthen the lifetime of the network or minimize the rate at which the sensors must be replenished with energy. This energy conservation can be accomplished through a number of ways. For example, sensors' reporting rate or data resolution can be adjusted, or the sensors can be turned off completely for an extended period of time. Balancing the application reliability with this goal of energy-efficiency essentially provides a type of application quality of service (QoS). To efficiently provide this QoS to the application, interaction with lower levels of the sensor network's protocol stack is required. Recently, efforts have been made to develop middleware providing this interaction while simplifying software development efforts [1] [2]. In this work, we discuss the advantages of efficient sensor management when used in such a middleware system.

In this work, we show how the use of two strategies - turning off redundant sensors and energy-efficient routing - can be used to extend network lifetime for a given required level of data reliability. Recent research has focused on methods of in-network data aggregation to reduce the amount of communication in dense wireless sensor networks. In this case, low-level fusion is typically performed on data from neighboring sensors before being sent to a data sink. As an alternative to this approach, redundant sensors can be turned completely off for periods of time to save energy. Of course, there is a tradeoff between power consumption and data reliability when choosing which approach to use. We consider the latter approach in this work but realize the benefits of the former. Our work also takes careful consideration at the routing layer. Many power/energy-aware routing algorithms have been proposed over the last several years. Most of these algorithms are designed to work without knowledge of future traffic patterns in the network, which turns out to be quite critical. We show in this paper that joint optimization of sensor scheduling and data routing can extend the lifetime of a network considerably compared to approaches using unintelligent scheduling, even when used with power-aware routing algorithms.

The rest of this paper is organized as follows. Section II gives the problem formalization. Section III provides simulation results. Section IV provides context for where this intelligent scheduling/routing can be used. Section V addresses related work. Section VI concludes the paper.

II. Multihop Sensor Network Management Problem

If an application is able to perform at an acceptable level using data from a number of different sensor sets, we would like to schedule the sets so as to maximize the sum of the time that all sensor sets are used. Acknowledging the impact that route selection will have on network lifetime, we would like to determine route selection in conjunction with the sensor schedule. In general, the routes should be chosen so that nodes that are more critical for use as sensors are routed around as often as possible. On the other hand, when determining the length of time for which a sensor set should be used, it is important to consider that the affected sensors are not only those that are active in the set, but also those being used in the chosen path(s) to the data sink. Obviously, it is wise to tightly couple the scheduling of sensor sets with selection of routes. In this section, we formalize this problem and model it as a generalized maximum flow problem with additional constraints. For a description of the generalized maximum flow problem, the reader is referred to [3].

We assume that for the majority of the network lifetime, the sensors act in a vigilant state, looking for a potential phenomenon in the environment being monitored. In this case, the state of the application remains constant over time. In applications such as object tracking where higher reliability is required in the vicinity of the object and nearby sensors become more critical, the application state changes frequently and the problem becomes much more difficult to model. In this work, we assume the simpler model of a constant state application but discuss the implications that multi-state applications would have on our model in Section II-C.

A. Problem Formulation

In previous work, we have shown how to maximize network lifetime via optimal scheduling in single hop wireless sensor networks [4]. Here, we extend the model to account for multihop networks. We consider a multihop network consisting of N_S multi-mode sensors and refer to the complete set of sensors as $S = \{S_j, j \in \{1 \dots N_S\}\}$. In general, we will assume that all sensors in the network are capable of operating in $N_{m,j}$ active modes and additionally in sleep mode, where the sensor's power consumption is negligible. An example of a sensor that is capable of operating in multiple active modes is a video camera that can send data at variable resolution or an ECG system that can work with different numbers of leads.

In order to achieve the application's required QoS, it may be possible to use a number of the sensors by themselves or in combination. A sensor set is determined to be *feasible* if i) the total bandwidth necessary to support the traffic of the set in any region is below the capacity of the network and the data is guaranteed to be schedulable and ii) the set provides the necessary reliability to the application. We will refer to the set of feasible sensor sets as $F = \{F_i, i \in \{1 \dots N_F\}\}$. In order to describe the makeup of each feasible sensor set F_i, we use a variable a_{ijk}, which is equal to one if sensor S_j is being used in mode k during feasible sensor set F_i and equal to zero otherwise. We must also define $P_{s,jk}$, which represents the power consumption of sensor S_j when used in mode k. This power consumption includes both the sensing power and the transmission power (typically the dominant term).

Now we will define a representation of the available routes from each node with variable $r_{j_1 j_2 l}$, which is equal to one if sensor S_{j_1} is included on sensor S_{j_2}'s lth distinct path to the data sink (routes are arbitrarily numbered from 1 to N_{P,j_2}, which represents the number of distinct paths from sensor S_{j_2} to the data sink). $P_{r,j_1 j_2 k}$ represents the power consumption of sensor S_{j_1} when routing sensor S_{j_2}'s data during mode k and includes receiving, processing, and transmit power.

We wish to develop a schedule that determines which sensor combinations should be used to monitor the environment and for how long and how the data from these sensors should be routed to the application. Let T_i represent the length of time that feasible sensor set F_i is being used in the schedule. T_i provides all the information that needs to be known about the schedule since the order of usage is of no consequence as long as startup costs of using the sets are negligible.

We also need to determine f_{jli}, the fraction of time that path l is used to route sensor S_j's data during the time that feasible sensor set F_i is used.

Sensor set lifetime is limited by the sensor battery levels E_j, which introduces our first constraint.

$$\sum_{i=1}^{N_F} \sum_{j_2=1}^{N_S} \sum_{k=1}^{N_{m,j_2}} \sum_{l=1}^{N_{P,j_2}} r_{j_1 j_2 l} P_{r,j_1 j_2 k} f_{j_2 l i} a_{i j_2 k} T_i + \sum_{i=1}^{N_F} \sum_{k=1}^{N_{m,j_1}} a_{i j_1 k} P_{s,j_1 k} T_i \le E_{j_1} \qquad \forall j_1 \tag{1}$$

This constraint says that the total amount of time any node can route other nodes' data and the total amount of time any node can be an active sensor are limited by that node's initial energy.

Another obvious constraint in this problem is that a sensor cannot realistically operate in multiple modes within a single sensor set, which means that

$$\sum_{k=1}^{N_{m,j}} a_{ijk} \le 1 \qquad \forall (i,j) \tag{2}$$

If a sensor is not in direct communication range with the data sink, its data needs to be forwarded to the data sink for the entire duration of each of its sensor sets' scheduled time. This is true for each feasible sensor set in which the sensor is used. This gives the constraint

$$\sum_{l=1}^{N_{P,j}} f_{jli} = \begin{cases} 1 & \text{data sink not in } S_j\text{'s Tx range} \\ 0 & \text{otherwise} \end{cases} \qquad \forall (i,j) \tag{3}$$

The objective of the problem is to maximize

$$T_{total} = \sum_{i=1}^{N_F} T_i \tag{4}$$

B. Interpretation of Joint Scheduling and Routing Problem as a Generalized Maximum Flow Problem

The joint scheduling and routing problem that was formalized in the previous section can be modeled as a generalized maximum flow graph problem. Consider an energy bank represented by source node s in Figure 1. In our model, we will initially represent energy consumption as the flow x_{ij} along arc (i, j). Nodes that represent sensors that are available for use by the application can be seen in the second column in Figure 1. Since the energy bank supplies the sensor nodes with their energy, arcs (s, S_j) must be drawn from the source node s to each of the sensor nodes S_j. Each sensor can only be supplied with the energy that is contained locally, and so there is a capacity u_{sS_j} on each arc (s, S_j), such that

$$0 \le x_{sS_j} \le u_{sS_j} \tag{5}$$

and

$$u_{sS_j} = E_j \tag{6}$$

Additionally, there are nodes in the graph representing the feasible sensor sets F_i. An arc (S_j, F_i) is drawn on the graph iff sensor S_j is included in feasible sensor set F_i. Once flow arrives at one of the nodes in F, it is more appropriate to consider the flow as time rather than energy. The arc multipliers to accomplish this will be described later.

If we wish to map the multihop case to the graph problem, we must represent data forwarding on the graph. In order to do this, we introduce two additional sets of nodes to the graph. A node in the first set R_{ji} (fourth column of nodes in Figure 1) represents the collective task of routing sensor S_j's data during the operation of feasible sensor set F_i. Since the routing contributes to the sensor set, there are arcs drawn from these nodes to the sensor set during which they forward data. The number of these arcs is equal to the number of active sensors in F_i that are not within direct communication range of the data sink. For example, in Figure 1, sensor set F_1 consists of sensors S_2 and S_3, but there is only one such arc, for S_2, since S_3 is within transmission range of the data sink.

A node P_{jil} in the second additional set of nodes (third column of nodes in Figure 1) represents the lth distinct path used in forwarding S_j's data during F_i. In the case of multipath routing, several of these path nodes exist for each node described in the previous paragraph. The number of these nodes P_{jil} having an arc drawn to R_{ji} is equal to the number of distinct paths from sensor S_j to the data sink. Arcs are also drawn from each sensor on the path to the path node.

A destination node d that represents the application is shown in the last column in Figure 1. Arcs (F_i, d) are drawn from each feasible sensor set to the destination node. As in all generalized maximum flow problems, the objective is to maximize the total flow into the destination node.

Finally, we must define the multiplier values for each arc. On arcs (S_j, F_i) and arcs (S_j, P_{jil}), we wish to convert flow from units of energy to units of time. The multiplier on (S_j, F_i) is the inverse of the power consumption of S_j when used in F_i.

$$\gamma_{S_j F_i} = \frac{1}{\sum_{k=1}^{N_{m,j}} a_{ijk} P_{s,jk}} \tag{7}$$

The multiplier on (S_{j_1}, $P_{j_2 il}$) is the inverse of the power consumption of sensor S_{j_1} when routing sensor S_{j_2}'s data during F_i.

$$\gamma_{S_{j_1} P_{j_2 il}} = \frac{1}{\sum_{k=1}^{N_{m,j_2}} a_{ij_2 k} P_{r,j_1 j_2 k}} \tag{8}$$

Multipliers also need to be included on the arcs connecting the feasible sensor sets and the destination node. These need to be included because the time that set F_i is used is equal to the time that each of its active sensors is used, not the sum as the problem currently stands. The multiplier $\gamma_{F_i d}$ on each arc (F_i, d) must be equal to the inverse of the sum of the number of active sensors in the set and the number of active sensors requiring data routing.

$$\gamma_{F_i d} = \frac{1}{\sum_{j=1}^{N_S} \sum_{k=1}^{N_{m,j}} a_{ijk}(1 + \sum_{l=1}^{N_{P,j}} f_{jli})} \tag{9}$$

The multiplier $\gamma_{P_{jil} R_{ji}}$ on each arc (P_{jil}, R_{ji}) must be equal to the inverse of the number of intermediate nodes on the lth path from sensor S_j to the data sink.

$$\gamma_{P_{j_1 il} R_{j_1 i}} = \frac{1}{\sum_{j_2=1}^{N_S} r_{j_2 j_1 l}} \tag{10}$$

Additional constraints are that the flows on each individual arc entering each feasible sensor set (from both nodes representing collective routing tasks as well as those representing sensors) must be equal. This is because of the fact that all sensors must be active for the entire duration of the sensor set's scheduled time in order for the set to be considered feasible from the application's perspective.

$$\gamma_{S_j F_i} x_{S_j F_i} = \gamma_{F_i d} x_{F_i d} \quad \forall (i, j) \tag{11}$$

$$\gamma_{R_{ji} F_i} x_{R_{ji} F_i} = \gamma_{F_i d} x_{F_i d} \quad \forall (i, j) \tag{12}$$

Also, the time that each node used in a path contributes to that path must be equal.

$$\gamma_{S_{j_1} P_{j_2 il}} x_{S_{j_1} P_{j_2 il}} = \gamma_{P_{j_2 il} R_{j_2 i}} x_{P_{j_2 il} R_{j_2 i}} \quad \forall (i, j_1, j_2, l) \tag{13}$$

Since we are attempting to maximize application lifetime, the objective of this graph problem is

$$\text{Maximize} \qquad \sum_{i=1}^{N_F} \gamma_{F_i d} x_{F_i d} \tag{14}$$

Because of the constraints in Equations 11, 12, and 13, none of the algorithms that are commonly used to solve generalized maximum flow problems in polynomial time can be used for the sensor scheduling problem. Instead, we use a simple linear programming approach. It has been shown that linear programs are solvable in polynomial time [5].

C. Extension to Multi-State Applications

The previously described model applies specifically to applications where the set of sensor sets providing reliable data to the application remains fixed over time. In many cases, current environmental conditions may cause these sets to change over time. An example of this is a personal health monitoring system where abnormal conditions detected at the base station node require more intense monitoring of certain vital signs. Another extreme case of this is an object tracking application. For these types of applications, more sensors may be required to be active in the current vicinity of the object being tracked. These applications can have a huge number of possible states. To incorporate this into the graph problem,

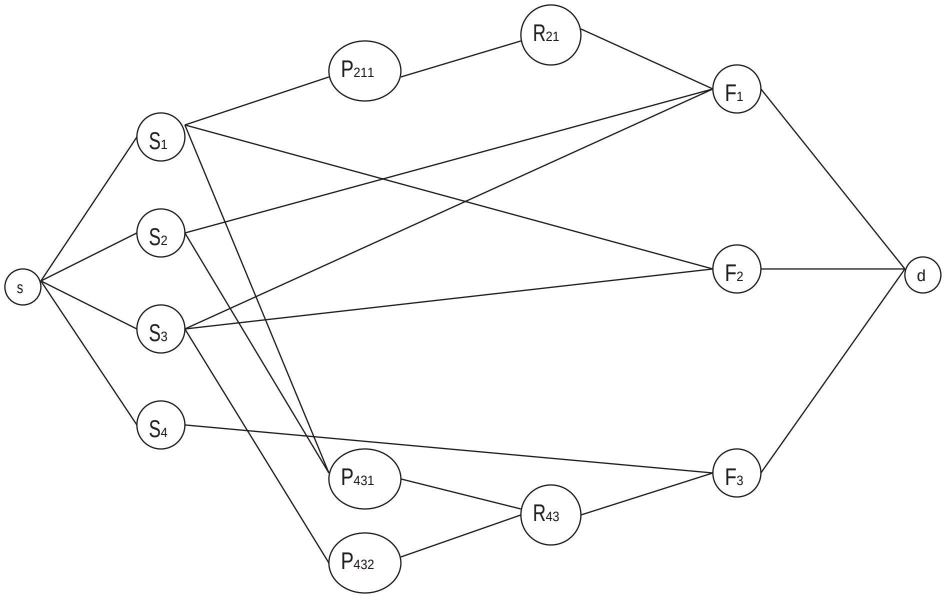


Fig. 1. Interpretation of the scheduling problem in a multihop network as a generalized maximum flow problem.

one would simply need to list all feasible sets for all possible states in the column where feasible sets are currently listed. The arcs from the feasible sets to the destination node would be replaced by arcs to a new set of nodes representing the possible application states. Arcs would be drawn from nodes representing feasible sensor sets to each state in which the set provides the acceptable application QoS. Finally, arcs would be drawn from these state nodes to the destination node. Further constraints would need to be placed on the problem so that the flows on the arcs from the state nodes to the destination are proportional to the fraction of time that the system is expected to be in the respective states. Depending on the application, building a probabilistic model to find these fractions may be simple or very complicated. Also, for such applications, it may be necessary to rerun the optimization program periodically to to reflect deviations from the expected behavior of the application state over time.

III. Simulations and Analysis

We expected that the potential relative lifetime improvement achieved through schedule/routing optimization rather than unintelligent sensor management methods would be quite large, especially in multi-hop networks, where the choice of sensor sets affect a larger number of sensors, namely all those in the selected path(s) to the data sink. In order to observe the size of the relative improvement that could be achieved and the effects of various factors on this improvement, we ran simulations of applications that require a minimum sensor coverage area over an environment where a potential phenomenon is expected to occur. We contrasted the lifetime of the optimal scheduling/routing with those under identical situations when sensor sets are chosen randomly from the set of feasible sensor sets. In these randomly chosen sets, we used shortest path routing and shortest cost routing, where the cost of routing a packet through node j was $\frac{1}{1-E_j}$ and E_j represents the current normalized energy capacity of sensor S_j. Our results show that depending on the situation, relative lifetime improvement can vary from nothing to more than a factor of 4.

It should be noted that our optimization uses global information. In larger networks, this is not realistic as the cost of gathering this information and the complexity of solving the optimization problem grows quickly. With larger networks, an alternative approach might be to perform several smaller optimization problems instead. Also, in all simulations, we did not consider the overhead of setting up and tearing down routes, although we acknowledge that in real scenarios, this could in fact impact average power consumption and network lifetime.

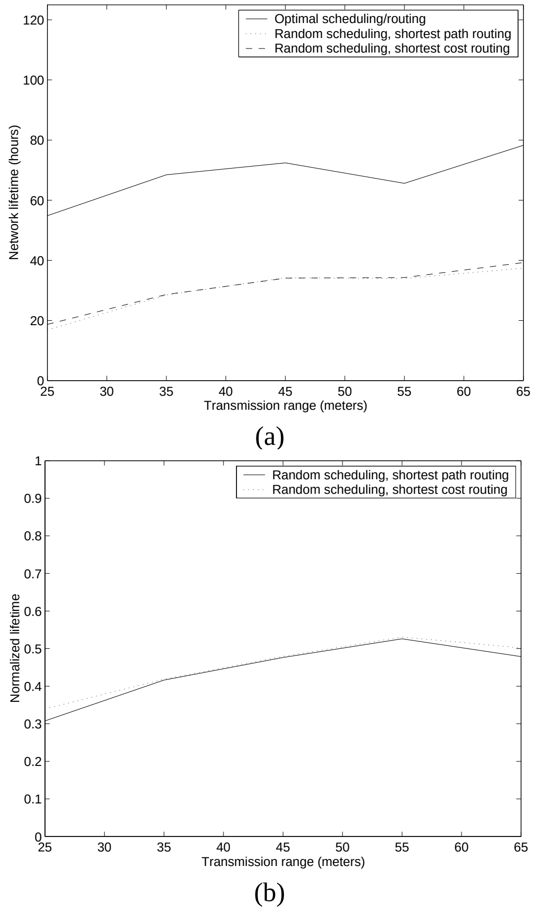
We placed a number of sensors, each capable of monitoring the environment within a constant sensing range, at random locations in a rectangular environment. The feasible sensor sets were found by determining which combinations of sensors would allow 100% of a predetermined portion of the area to be monitored. The choice of this method for specifying feasible sets is not necessarily important; any method that provides the possible sets of sensors that provide acceptable application QoS would be able to map to our framework. Because the complexity of finding every possible feasible set increases exponentially with the number of sensors used, we found 50 feasible sensor sets that represent a subset of F for each simulation. The way that these feasible sensor sets were chosen was as follows. Starting with a randomly chosen sensor, additional sensors were iteratively added to the set until the 100% coverage threshold was met. Additional sensors were chosen randomly with probability proportional to the added coverage that they would potentially provide.

In all simulations, we set initial sensor energy to 1 J for each node. The energy to transmit a packet was 15 μJ and the energy to receive a packet was 10 μJ. The energy to forward a packet was equal to the sum of the energy to transmit a packet and the energy to receive a packet (sensing and processing power was ignored, although this could easily be incorporated into the model). Each sensor was capable of operating in a single mode in which it sent packets at a rate of 1 packet/second. Assuming that we were working well under the point where network congestion becomes an issue, all sets were considered feasible from the network perspective. Each data point is averaged over 10 scenarios. For each scenario, we found the optimal scheduling/routing and ran 10 trials using the random scheduling with each routing method. All simulations for a single scenario used identical sensor locations and feasible sensor sets.

In the first simulations, we kept the number of nodes constant at 100 and varied the nodes' transmission range to observe the effect of path length on the potential relative improvement in application lifetime through optimization. The size of the environment was fixed at 100 m x 100 m and the sensing range was fixed at 25 m. The sink to which all data was routed was located in the center of the environment. Plots of application lifetime using the optimal scheduling and routing and using the randomly chosen sets with shortest path and shortest cost routing are shown in Figure 2a. A plot of the lifetime achieved through the random approaches normalized to the optimal solution's lifetime is shown Figure 2b. The size of the benefit of scheduling/routing optimization seems to decrease with transmission range until the network becomes more and more of a single-hop network where all sensors can communicate directly with the data sink. After this point, the size of the benefit should remain relatively constant, as our result verify. We have also plotted the relative lifetime improvement against the average length of the shortest path from the sensors in the network to the data sink in Figure 3. As expected, it can be seen that as routing becomes more and more critical, random set selection with shortest path and with shortest cost routing performs poorly.

Next, we varied the number of sensors while keeping the transmission radius fixed at 25 m and the size of the environment fixed at 100 m x 100 m to observe the effect of sensor node density. A plot of relative lifetime improvement using the optimal scheduling/routing and using randomly chosen sets with shortest path routing and with shortest cost routing is shown in Figure 4a. As more energy is distributed throughout the network, network lifetime is extended for all scheduling and routing methods. The relative improvement achieved through optimization is shown in Figure 4b. In previous work [4], we have shown that the size of the relative improvement seemed to be influenced very little by sensor node density for the single hop case. For multihop networks, sensor node density seems to have a small effect on the size of relative improvement as well.

Finally, we kept the node density constant at 0.01 node/m^2 and varied the size of the environment to observe the effect of the size of the problem. If the network were consistently single hop and sensors were distributed very uniformly throughout the environment (e.g., in grid-like fashion), we would expect the absolute lifetime to stay relatively constant. Since sensor location is random, however, the possibility of a lightly covered area in the environment, limiting network lifetime, increases. Also, the average power consumption in the network should increase as the sensor data needs to be forwarded along more hops on average. These two factors should cause network lifetime to decrease as network size increases. Figure 5a verifies this intuition and shows that network lifetime decreases with an increase in the size of the environment using all three approaches. Figure 5b shows that as the problem becomes large and there are more free variables to be solved, the benefit that can be achieved through optimization of the scheduling and routing becomes higher.



(a)

(b)

Fig. 2. Absolute lifetime (a) and lifetime normalized to optimal solution (b) of the sensor application, as a function of transmission range.

A summary of our results is given in Table I.

IV. MOTIVATION

We have begun development of a middleware for sensor networks whose purpose is to optimally balance the tradeoff between application performance, power consumption, and network lifetime. MiLAN (Middleware Linking Applications and Networks), an ongoing project at University of Rochester's

Routing method	Transmission range	Number of nodes	Field Size	Normalized lifetime
Optimal	(25m - 65m)	100	100 m x 100 m	1
Shortest path routing	(25m - 65m)	100	100 m x 100 m	(0.31±0.02 - 0.48±0.05)
Shortest cost routing	(25m - 65m)	100	100 m x 100 m	(0.34±0.03 - 0.50±0.04)
Optimal	25m	(50 - 150)	100 m x 100 m	1
Shortest path routing	25m	(50 - 150)	100 m x 100 m	(0.34±0.11 - 0.30±0.04)
Shortest cost routing	25m	(50 - 150)	100 m x 100 m	(0.41±0.11 - 0.34±0.04)
Optimal	25m	100	(60 m x 60 m - 100 m x 100 m)	1
Shortest path routing	25m	100	(60 m x 60 m - 100 m x 100 m)	(0.47±0.06 - 0.24±0.05)
Shortest cost routing	25m	100	(60 m x 60 m - 100 m x 100 m)	(0.51±0.09 - 0.29±0.04)

TABLE I

SUMMARY OF RESULTS.

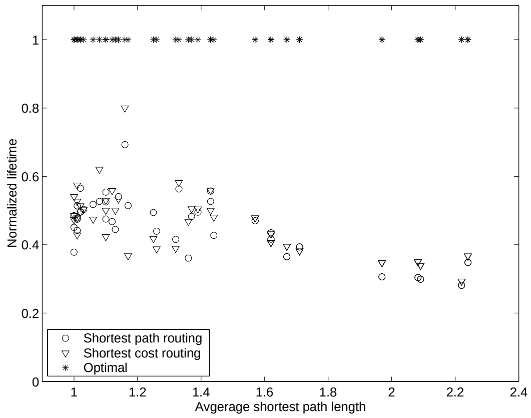
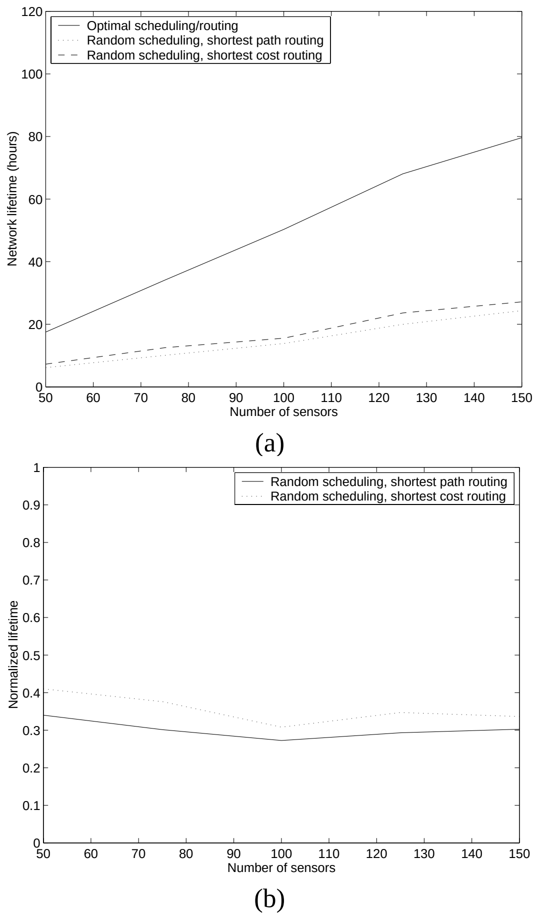


Fig. 3. Lifetime of the sensor application using optimal scheduling/routing and using random feasible sets with shortest path and shortest cost routing normalized to the optimal schedule, as a function of average shortest path length.



(a)

(b)

Fig. 4. Absolute lifetime (a) and lifetime normalized to optimal solution (b) of the sensor application, as a function of the number of sensor nodes.

Center for Future Health [6], will hide all network configuration information and decision-making from the application, simplifying program development efforts [1]. The main idea behind MiLAN is that sensor networks often consist of sensors that provide redundant information and are capable of providing more information than what the underlying application requires. MiLAN should decide the states of each sensor and be able to dynamically modify these states in response to changes in network topology and application state. Initially, we are developing MiLAN as a centralized middleware, where a central base station running the application makes all decisions regarding network configuration. Eventually, we plan to develop a distributed version of MiLAN, where sensors can either receive commands from local administrators instructing them about the ideal state that they should be in at a given time or actually make decisions by themselves based on local information.

There are many ways that an energy-performance tradeoff can be balanced. In some detection problems, the reporting rate of sensors can be varied so that a low delay between the time of event occurrence and the time that the application is notified can be given up for lower power consumption. Also, sensors can be turned off completely if their data is not important to the application. In that case, one might question why these additional sensors were deployed in the first place. The additional sensor deployment is justified because the state of the application or environment being monitored can change over time, and some sensors' data might be more critical during one state than in others. Also, it may be wise to deploy additional sensors for redundancy in the sense that if one sensor dies or malfunctions, it can be easily replaced by another.

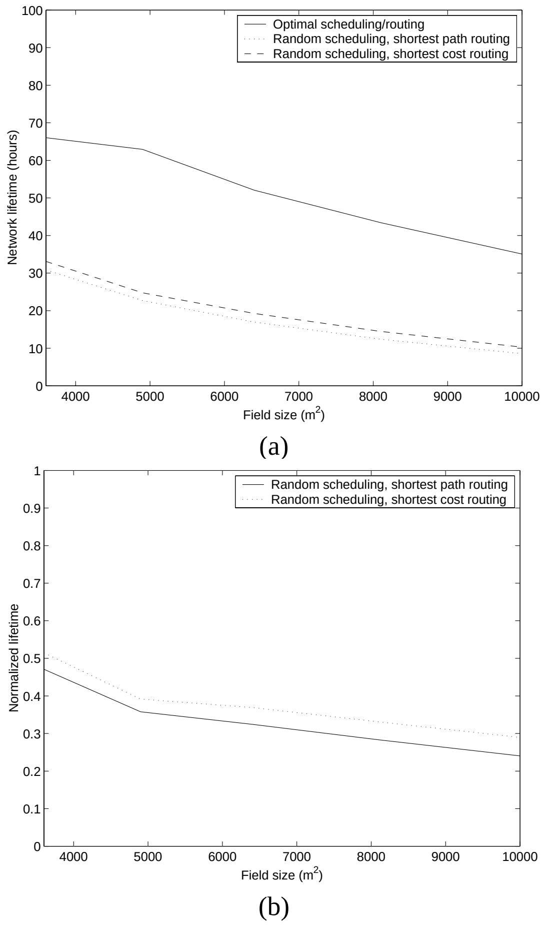


Fig. 5. Absolute lifetime (a) and lifetime normalized to optimal solution (b) of the sensor application, as a function of the field size.

Sensors can have diversity or similarity either spatially or in the type of information that they provide to the application. Activating sensors that are similar in either sense may be unnecessary and even undesirable when power consumption is a critical issue, as is often the case. For example, consider a heart-monitoring application that we are developing as the testbed for MiLAN. One type of sensor that we are using in our network is a blood pressure monitor. The application may want to monitor blood pressure with a number of sensors located at different places to see if there are blood pressure abnormalities anywhere on the body. The correlation of the data being sent from these sensors depends on the spatial diversity of the sensors. Using a sensor whose coverage area has large overlap with another sensor that is already being used is not wise since the marginal information that it provides to the application is minimal. To illustrate diversity (or similarity) in the type of information that a sensor provides, consider the use of an ECG system in conjunction with a simple direct heart rate monitor. While the heart rate monitor provides useful information in the absence of the ECG system, it does not provide much additional information when the ECG is already being used, and its energy could be conserved by turning the sensor off with no penalty in application performance. More importantly, the bandwidth consumed by such a sensor could prevent a sensor with more diverse information, such as a pulse oxymeter or gait analyzer, from joining the network if network congestion is an issue.

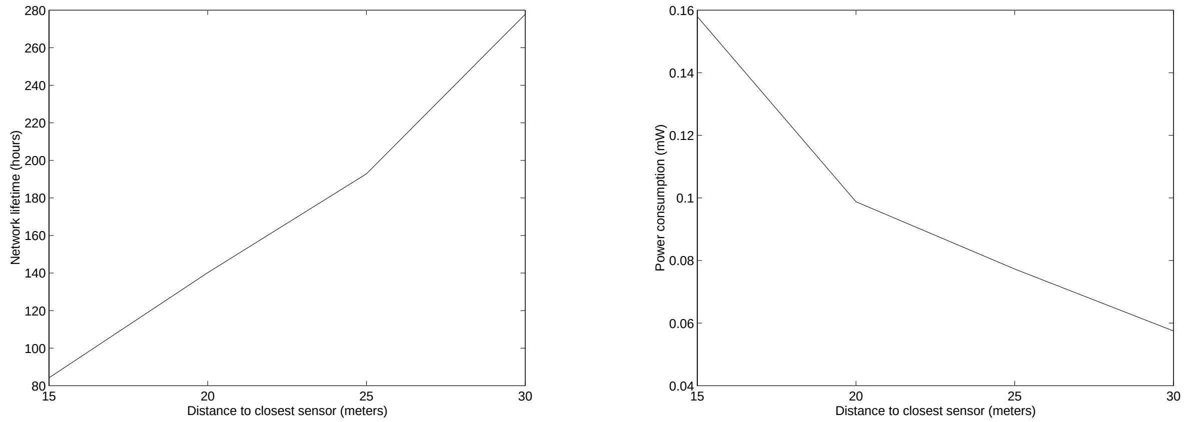
In this paper, we have considered sensor management for the spatially diverse sensor management problem. Specifically, for a given level of application reliability (in our simulations, we consider reliability to be the fraction of area covered by the sensor network), we show how to optimally schedule sensors to extend the system lifetime to its maximum. In addition to the blood pressure monitor that we have mentioned, it is obvious that this strategy will have benefits in many other types of sensor networks, such as smoke detection in forests and chemical detection in disaster situations.

In some network architectures, the entire set of spatially diverse sensors can be fused and interpreted as a single "virtual sensor." This "virtual sensor" would be capable of operating in multiple modes, varying from large coverage/high confidence and short lifetime/large power consumption to little coverage/low confidence and long lifetime/little power consumption, depending on what level of reliability is met. For an application where reliability is determined by the minimum distance from any point in the field to the nearest sensor, an example of this tradeoff curve is shown in Figures 6a and 6b. Before determining the sensor set scheduling and routing, the application could look at one of these curves to determine its available options. If the application absolutely needs data to have a certain level of reliability, it would set a threshold on the horizontal axis and determine the lower bound on power consumption or the upper bound on network lifetime. If on the other hand, the application sets a hard constraint on power consumption or network lifetime, it would set a threshold on one of the vertical axes, determining the reliability that should be obtained before running the optimization.

Running these optimization programs may in fact be too computationally intense for some networks, especially in those where the base station itself is an energy-constrained node. One option that can be used in situations with spatially determined feasible sets is to divide the problem into several local optimizations. This could simplify and distribute the optimization computation significantly. If these optimizations, even when scaled down through distribution, are still too costly, this optimization method will at least be a tool to let us know how close to the optimum alternative sensor scheduling and routing algorithms actually work.

V. Related Work

A great deal of research has focused on developing energy/power-aware routing algorithms to be used to extend the lifetime of wireless networks. Singh *et al.* argued for the use of new power-oriented metrics to be used when determining paths in routing protocols [7]. In Li's $max - minxP_{min}$ method, paths are chosen by maximizing the minimum residual energy of all nodes on the chosen path while bounding the



(a) (b)

Fig. 6. Reliability/lifetime/power consumption tradeoff for sensor application.

total power consumption [8]. Chang *et al.* found local algorithms that provide routing solutions allowing network lifetime to be extended near or equal to the optimal off-line calculated solution [9]. Many others have written about ways to achieve energy-efficiency in routing protocols [10] [11] [12]. These methods all have the same fundamental limitation of not having knowledge of future traffic patterns (namely, in our case, the predetermined sensor set schedule). In this paper, we have shown the advantage that this knowledge can provide when determining routes.

Many researchers have considered sensor management from a variety of perspectives [13] [14] [15] [16]. Limited energy resources are typically considered the primary design constraint for wireless sensor networks. Several sensor management schemes have focused on providing energy-efficiency in wireless sensor networks. STEM, proposed by Schurgers *et al.* [17], and SPAN, proposed by Chen *et al.* [18], allow smart sensors to be turned off whenever they are not being used as a traffic source or in a vital role in packet forwarding. Other cluster-based protocols such as LEACH, developed by Heinzelman *et al.* [19], take advantage of the redundant nature of data from densely populated wireless sensor networks by aggregating sensor data before forwarding to the data sinks, greatly reducing the communication involved. While most of the research in this area has attempted to minimize power consumption by reducing communication, not much has been done to modify sensor modes (and in turn reduce the amount of sensed data in the network) when the amount of data being collected is more than necessary. Tilak *et al.* have shown the advantages of modifying the sensor mode (reporting rate) so that an efficient point on the reliability curve is met [13]. In this work, the limiting factor being considered was network congestion rather than sensor energy.

The work of Bhardwaj *et al.* is the most similar to that described in this paper, although we were unaware of their work at the time of our work [20]. Bhardwaj solves a similar sensor management optimization problem from a slightly different perspective. Our work considers the problem from a high-level perspective, and is flexible in terms of finding feasible sensor sets for a variety of sensor network applications. The model described in this paper was intended for applications which spend the most of their time in a "vigilant" state, with alarm states occurring infrequently and for short periods of time. Meanwhile, the work of Bhardwaj *et al.* is more focused on efficiently sending event-triggered sensor data to the data sink, and the model is aimed more toward tracking-type applications.

VI. Conclusions and Future Work

Intelligent sensor management is one way to provide Quality of Service to an application. We have formalized a sensor management problem and shown through simulations that for applications using spatially diverse sensors, improvement larger than a factor of 4 can be obtained when sensor scheduling and routing are solved optimally and jointly when compared with more random methods. Simulation results show that improvement increases rapidly with the volume of the data routing but is less dependent of the network density.

In our model, average power consumption of a node to route data is proportional to the routing load. In real networks, this power cost may actually depend on the state of the node. For example, in an IEEE 802.11 network, the marginal power cost to route data for a node already being used as an active sensor in the current set is significantly less than for one that is otherwise idle. For the active sensor, the energy cost is approximately proportional to the difference between idle power and the average of the receive power and transmit power. Meanwhile, a sensor that is not being used in the current sensor set would otherwise have its radio turned off, and the marginal cost is proportional to the sum of the receive power and transmit power. Similarly, if a node is being used to route data in multiple paths, its energy cost per forwarded packet is significantly less. When modeling such types of networks, the problem can no longer be modeled as a generalized maximum flow problem and becomes NP-hard, requiring a heuristical approach to solve.

While not representative of all networks, we believe that our model does represent some typical networks that are likely to be used in sensor networks. Many TDMA-based networks can be modeled as having power dissipation that is nearly proportional with the routing load, as we have modeled. Such networks include Bluetooth networks when an intelligent scheduling method is used. Still, even in these TDMA-based networks, traffic schedules may need to be set up to allow for this efficiency, and this overhead is not accounted for in our model.

In all simulations, we did not consider the overhead of setting up and tearing down routes. We acknowledge that in real situations this could in fact impact average power consumption and network lifetime and our model would need to account for these additional factors before running the optimization program. It should also be noted that these simulations were quite simplified and that quirks of specific routing protocols and other factors could cause discrepancies in performance measurements. For example, in our simulations, we considered only routes in which each successive hop moves toward the base station to be valid. In fact, there are additional valid indirect routes that are not accounted for in our simulations. Our approach also requires global information about the neighborhoods of each node. For small networks, this may not be a problem, but requiring this information to be propagated back to the base station would not scale well for larger networks.

The motivation for this work was a larger project in which we are developing a middleware system for use in wireless sensor networks. We will incorporate the ideas and strategies presented in this paper during the development of this middleware. We are also developing a distributed algorithm in which sensors use information from neighboring nodes, gathered through eavesdropping, to determine what their current state should be. We hope that this algorithm will achieve lifetime results that are near the upper bounds that can be found by the methods described in this paper.

ACKNOWLEDGEMENTS

The authors would like to thank the peer reviewers for their helpful comments. Mark Perillo is supported by the Office of Naval Research through a NDSEG fellowship.

REFERENCES

[1] A. Murphy and W. Heinzelman, “Milan: Middleware linking applications and networks,” TR-795, University of Rochester, Computer Science, November 2002.

[2] K. Nahrstedt and D. Xu, “Qos-aware middleware for ubiquitous and heterogeneous environments,” *IEEE Communications Magazine*, vol. 1, no. 4, October 2001.

[3] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, Inc., 1993.

[4] M. Perillo and W. Heinzelman, “Optimal sensor management under energy and reliability constraints,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2003.

[5] L. Khachiyan, “A polynomial algorithm in linear programming (in russian),” *Soviet Mathematics Doklady*, vol. 20, 1979.

[6] “University of rochester center for future health,” http://www.futurehealth.rochester.edu, 2002.

[7] S. Singh, M. Wu, and S. Raghavendra, “Power aware routing in mobile ad hoc networks,” in *Proceedings of the Fourth Annual International Conference on Mobile Computing and Networking*, 1998.

[8] Q. Li, J. Aslam, and D. Rus, “Online power aware routing in wireless ad hoc networks,” in *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking*, 2001.

[9] J. Chang and L. Tassiulas, “Energy conserving routing in wireless ad hoc networks,” in *Proceedings of INFOCOM*, 2000.

[10] D. Braginsky and D. Estrin, “Rumor routing algorithm for sensor networks,” in *Proceedings of the International Conference on Distributed Computing Systems (ICDCS-22)*, 2001.

[11] C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed diffusion: A scalable and robust communication paradigm for sensor networks,” in *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking*, 2000.

[12] Y. Xu, J. Heidemann, and D. Estrin, “Geography-informed energy conservation for ad hoc routing,” in *Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking*, 2001.

[13] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, “Infrastructure tradeoffs for sensor networks,” in *Proceedings of the ACM International Workshop on Wireless Sensor Networks and Applications Workshop*, 2002.

[14] M. Chu, H. Haussecker, and F. Zhao, “Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks,” *International Journal of High Performance Computing Applications*, 2002.

[15] J. Liu, P. Cheung, L. Guibas, and F. Zhao, “A dual-space approach to tracking and sensor management in wireless sensor networks,” in *Proceedings of the ACM International Workshop on Wireless Sensor Networks and Applications Workshop*, 2002.

[16] S. Meguerdichian, S. Slijepcevic, V. Karayan, and M. Potkonjak, “Localized algorithms in wireless ad-hoc networks: Location discovery and sensor exposure,” in *Proceedings of MobiHOC*, 2001.

[17] C. Schurgers, V. Tsiatsis, S. Ganeriwal, and M. Srivastava, “Optimizing sensor networks in the energy-latency-density design space,” *IEEE Transactions on Mobile Computing*, vol. 1, no. 1, January 2002.

[18] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, “Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks,” in *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networkin*, 2000.

[19] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, October 2002.

[20] M. Bhardwaj and A. Chandrakasan, “Bounding the lifetime of sensor networks via optimal role assignments,” in *Proceedings of INFOCOM*, 2002.