

Using Mobile Relays to Prolong the Lifetime of Wireless Sensor Networks

Wei Wang, Vikram Srinivasan, Kee-Chaing Chua

Department of Electrical and Computer Engineering
National University of Singapore
Singapore, 117576

{g0402587,elevs,eleckc}@nus.edu.sg

ABSTRACT

In this paper we investigate the benefits of a heterogeneous architecture for wireless sensor networks composed of a few resource rich mobile nodes and a large number of simple static nodes. These mobile nodes can either act as mobile relays or mobile sinks. To investigate the performance of these two options and the trade-offs associated with these two options, we first consider a finite network. We then compute the lifetime for different routing algorithms for three cases (i) when the network is all static (ii) when there is one mobile sink and (iii) when there is one mobile relay. We find that using the mobile node as a sink results in the maximum improvement in lifetime. We contend however that in hostile terrains, it might not always be possible for the sink to be mobile. We then investigate the performance of a large dense network with one mobile relay and show that the improvement in network lifetime over an all static network is upper bounded by a factor of four. Also, the proof implies that the mobile relay needs to stay only within a two hop radius of the sink. We then construct a joint mobility and routing algorithm which comes close to the upper bound. However this algorithm requires all the nodes in the network to be aware of the location of the mobile node. We then proposed an alternative algorithm, which achieves the same performance, but requires only a limited number of nodes in the network to be aware of the location of the mobile. We finally compare the performance of the mobile relay and mobile sink and show that for a densely deployed sensor field of radius R hops, we require $O(R)$ mobile relays to achieve the same performance as the mobile sink.

Categories and Subject Descriptors

C.2.1 [Network Architecture and Design]: Wireless communication

General Terms

Algorithms, Design, Performance

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobiCom '05, August 28–September 2, 2005, Cologne, Germany.
Copyright 2005 ACM 1-59593-020-5/05/0008 ...\$5.00.

Keywords

Sensor networks, Network lifetime, Mobile relay.

1. INTRODUCTION

Wireless sensor networks are expected to be deployed in inaccessible and hostile environments such as dense jungles (for habitat monitoring applications), battlefields (for enemy troop movement monitoring) etc. Since these environments are not amenable to careful deployment of sensor nodes, it is expected that a large number of cheap, simple sensor devices will be randomly scattered over the region of interest. These devices are then expected to self-organize to form a multi-hop network and relay critical information back to a "sink" which acts as a gateway to a backbone network. Thus in addition to the sensing task, every node also bears the burden of relaying information from other sensors in the network to the sink.

Since these devices are expected to be extremely cheap, they are also severely restricted in the resources that they possess. Each node is battery powered and has limited processing and memory capabilities. For example the Berkeley mote is powered by two alkaline AA batteries and has a 4MHz processor with 128kB of instruction memory and 4kB of RAM. Therefore it is critical that these resources be used judiciously in order to maximize the benefit from the network before it dies. Although there is a concerted effort from the device research community at designing low power hardware and efficient energy sources, the network research community has also realized that inefficient algorithms at the various networking layers can result in nodes dying prematurely. There are several proposals at the MAC ([1, 2]) and Network layers ([3, 4]). However most of these proposals are based on the assumption that the entire network is composed of static nodes. In this paper we assume that in addition to a large sensor field composed of simple low cost static nodes, we have a few extremely resource rich (in terms of processing, memory and energy) mobile nodes. We then pose the following questions: (i) Should a mobile node act as a sink or a relay? (ii) What is the lifetime of the network for each of these alternatives? (iii) What are the trade-offs associated with each of these alternatives?

The motivation for investigating such a heterogeneous network is the following. Suppose you are a network designer who is asked to randomly deploy a sensor network over some given area and provide a guarantee in terms of lifetime, connectivity and coverage. You are only allowed to use extremely simple, cheap and fault prone sensor devices which

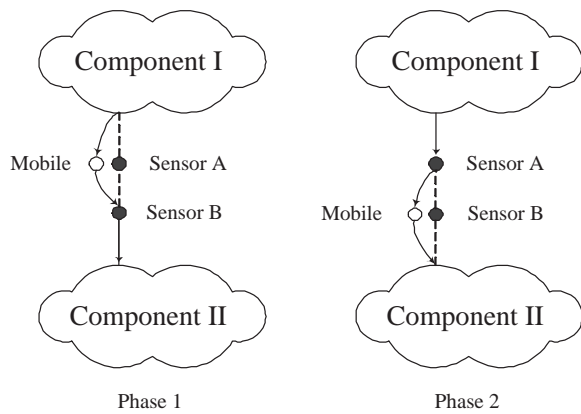


Figure 1: Using mobile relay to extend the lifetime for the bottleneck nodes

cost 15 cents a device. In this scenario, the only way to provide the required guarantees is to over provision and deploy a very large number of static devices (say a million). Therefore even if each sensor device by itself is cheap (15 cents), having to deploy a million devices will cost \$150,000. An alternative is to have a few expensive (say 10 devices at \$100 each) and resource rich mobile devices and many simple static nodes. In this scenario, it is intuitive that to provide the same guarantee of lifetime, connectivity and coverage, one would require far fewer static devices (say 100000), resulting in a total cost of about \$16,000. The motivation of this paper is to investigate what performance improvement can accrue from mobile devices and the trade-offs associated with such an architecture.

Figure 1 shows a sample network, where the use of mobile nodes as relays can significantly improve the lifetime of the network. Suppose that the network is composed of two components which are connected via two sensor nodes A and B. Assume that these are the critical bottleneck nodes in the network and the lifetime of these two nodes is T , while other sensors have lifetime much longer than T . If we have one mobile node with the same transmission range and reception range as the sensor nodes, then the network lifetime can be at least doubled. A simple algorithm for this would be for the mobile node to shuttle between node A and node B and inherit the responsibilities of the node with which it is co-located (including sensing and relaying). It is clear that with an appropriate shuttling schedule, the network lifetime can be doubled to $2T$. We assume here that the energy resource at the mobile node is far greater than that of any of the sensing nodes.

In this paper, we first consider the case of finite networks, and pose the problem of maximizing lifetime as a linear programming problem and derive the optimal schedule for the mobile nodes. The system model used here for mobile relays is similar to the one for mobile sinks in [5, 6]. We then compare the performance of the mobile relay with minimal hop routing, energy-conserving routing and the mobile sink approach proposed in [7]. We see that the mobile sink approach always out performs the mobile relay approach. The intuitive reason for this is the following. When the sink is static, the nodes around the sink become bottleneck nodes since they relay traffic for all the other nodes in the net-

work. However, by making the sink mobile, we distribute the bottleneck nodes all around the network. We contend that it is not always feasible to have a mobile sink, since it is expected to act as a gateway to a backbone network. In hostile and inaccessible environments, it might not be possible to maintain continuous connectivity with the backbone network when the sink is mobile.

We then consider a large densely deployed sensor network and show that an upper bound on lifetime with one mobile relay is 4 times that of the static network. More interestingly, this upper bound computation shows that the mobile relay will never have to venture farther than a two hop distance from the sink. We then construct a joint mobility and routing algorithm which improves the lifetime of the network by almost a factor of 4. The disadvantage of this routing algorithm is that all the nodes in the network will have to be aware of the current location of the mobile relay. We then construct another routing algorithm in which only nodes within a certain distance of the sink need to be aware of the location of the mobile relay. We show that we can achieve the same lifetime improvement with this algorithm as the earlier algorithm. We then show how these algorithms can be extended to the case when there are m mobile nodes. We finally compare the lifetime of the network in the mobile sink case with that of the mobile relay case. We show that for a large dense network deployed in a circular region of radius R , we need $O(R)$ mobile relays to achieve the same lifetime as that of the mobile sink.

The rest of the paper is organized as follows: Section 2 summarizes related work. Section 3 compares the mobile relay approach with other energy conserving methods for finite networks. Section 4 investigates the performance of a large dense network with a few mobile relays and gives a joint mobility and routing algorithm. Section 5 analyzes the performance of the mobile sink approach and mobile relay approach. Section 6 gives the simulation results on finite networks. Finally, section 7 concludes the paper.

2. RELATED WORK

Existing literature utilizes mobile nodes as mobile sinks to save energy. Shah et al.[8] proposed to use randomly moving "Data Mules" to help gather the sensing data. Mobile sinks with predictable and controllable movement patterns are studied in [9, 10, 11]. In these approaches, the static sensors only send out their data when the sink is moving close enough to them. The disadvantage of such proposals is that there will be considerable delay in acquiring sensed data, since a node need to wait for the sink to approach it. In order to minimize the delay, several methods of delivering the sensing data through multi-hop communication to the mobile sink are proposed [5, 6, 7]. The mobile sink can either 'jump' between several predefined positions or patrol on an continuous route. In the first case, the problem can be posed as a linear programming problem where mobile sinks can find the optimal time schedule to stay at these predefined points [5, 6]. Another method is introduced in [7], where the optimal route is obtained through a geographic traffic load model. In this approach, as the mobile sink goes around the network, sensors will continuously track the position of the sink and send their packets to the sink via multi-hop communication. The disadvantages of the mobile sink paradigm are three fold, (i) the mobile sink need to roam around the periphery of the network to maximize the network lifetime

[7], (ii) all nodes in the network will have to be aware of the current location of the mobile sink in order to route information to it and (iii) in most scenarios, the sink is a gateway to a backbone network, over which human operators can monitor the status of the sensor field. In such scenarios, it will be difficult to engineer a system whereby a mobile sink is connected at all times to the backbone network.

Our approach is different in that we consider a static sink and mobile nodes act as relays to improve the lifetime of the network. Moreover, we will show that in a large and dense network, the region over which the mobile nodes will have to patrol in order to maximize lifetime is limited to a two hop radius around the sink.

Another large category of energy conserving methods is to use flow control algorithms to find the optimal energy conserving routes [3, 12]. The energy conserving routing and the mobile sink approach share the same idea of distributing the traffic load evenly in the network so that the lifetime of the network is maximized. Unlike these methods, our approach tries to "add" more energy to heavily loaded nodes. What is interesting is that in a large and dense network, even the addition of one mobile node can result in a lifetime improvement of up to four times over that of the all static network.

Other solutions for energy saving have also been intensely studied, including data aggregation and topology control methods. Data aggregation and clustering methods such as [13, 14, 15] aggregate the sensing data to decrease traffic volume and thereby prolong network lifetime. Topology control methods such as [16, 17] use controllable transmission range to achieve the most energy efficient network topology. In our work we do not address the issues of data aggregation or topology control. However, these ideas can be useful complements to our proposal of using mobile relays. As we will describe later, depending on the position of the mobile relay, traffic is intentionally routed via a few specific network nodes. This could facilitate the data aggregation process.

3. COMPARISON OF ENERGY CONSERVING METHODS

In this section, we will compare some of the different energy conserving protocols proposed in the literature with the mobile relay approach. For the sake of these illustrations, we assume that the mobile node has complete knowledge of the topology of the network. We will start by describing two different ways in which a mobile relay can be used to prolong lifetime¹.

i). Static Routing

In this approach, the mobile relay decides a schedule, which is given by a sequence $\{(l_1, u_1), (l_2, u_2), \dots, (l_N, u_N)\}$, where N is the number of nodes in the network, l_i is the location of node i in the network and u_i is the duration of time the mobile relay spends at that location. When the mobile relay is at the location l_i , it inherits the responsibilities of node i which is at l_i and allows node i to rest for a duration u_i . Therefore, for the duration u_i , the routes in the network do not change, other than to route all traffic passing through node i via the mobile relay. Since the static node i can stay in sleep mode for the duration u_i , its lifetime will

¹There are many other possibilities, we illustrate the simplest alternatives here

be extended u_i time units. In practice, one mobile relay can only stay in one place at a time, so the sum of the duration cannot exceed the lifetime of the network. It is clear that the optimal schedule is to sort the original lifetime of the static nodes in increasing order, then use the mobile relay to help the static nodes in this order. This process is similar to the water-filling argument used in information theory [18]. For the m mobile relay case, we have:

$$m \times T_s = \sum_{t_i < T_s} (T_s - t_i) \quad (1)$$

where T_s is the lifetime (defined as the time until the first static node dies as in [3]) after we add m mobile nodes, and t_i is the lifetime of node i before we add mobile nodes. The relay durations of m mobiles are water filled to those static nodes whose lifetime is shorter than the optimal lifetime T_s .

ii). Dynamic Routing

In this approach, the routes from a node to a sink will depend on the position of the mobile relays. In other words, the sensors need to know where the mobile relays are and redirect their packets according to the current position of the mobile relays. In general, the more packets passing through the mobile relays, the larger the performance gain from mobile relays.

For the network with one mobile relay and N static nodes, the optimal routes can be calculated through linear optimization in a way similar to [6]. To mitigate the computational burden, we suppose that the mobile relay only stays at positions where there is a static sensor. If the mobile relay will stay at the position of node k for a time period t_k until some node in the network dies, then the linear optimization problem can be formulated as follows:

$$\text{Maximize } \sum t_k \quad (2)$$

$$\text{s.t. } \sum_j x_{ij}^k - \sum_j x_{ji}^k = s_i \times t_k \quad \forall i, k \quad (3)$$

$$x_{ij}^k \geq 0 \quad \forall (i, j), \forall k \quad (4)$$

$$\sum_{\forall k \neq i} \sum_j x_{ij}^k \times e \leq E \quad \forall i \neq 0 \quad (5)$$

where x_{ij}^k is the total traffic flow from node i to node j during time period t_k . E and e represent the initial energy of the sensor nodes and the energy for relaying one packet respectively. s_i is the packet generation rate for node i . Assume the data generation rate is uniform and $s_i = 1$. Then the data generation rate for the sink, denoted as s_0 , is set to $-N$. Constraints (3) and (4) give the network flow constraint for node i during the period t_k , the total out-flow minus the total in-flow during the period will be equal to the packets node i generated in this period. Constraint (5) is the energy constraint: the sum of energy over all the periods during which node i relays traffic should not exceed the initial energy E . The energy used in period $k = i$ is not counted, since we assume that in period t_i the mobile is on node i and takes over the task of node i , including sensing and relaying. Note that as opposed to the static routing scheme, where the routes to the sink do not depend on the position of the mobile relay, in the dynamic routing scheme, the routes depend on the position of the relay.

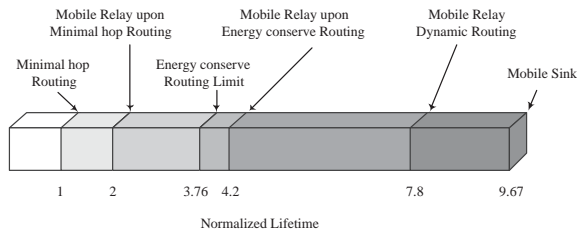


Figure 2: Comparing network lifetime for different approaches

Figure 2 compares the two mobile relay approaches with other methods. We average over 100 realizations of a random network with 100 nodes and one sink in a 5×5 square area. In this figure, the lifetime of non-energy-aware shortest hop routing networks is normalized to 1.

As the figure shows, when we use one mobile relay over shortest path routing, we can get almost 100% improvement by using static routing algorithms. If the network lifetime has been optimized by energy conserving routing [3], putting in one mobile relay with static routing can improve the lifetime by about 11%. This shows that the static routing is only useful in networks with unbalanced loads.

The dynamic routing algorithm performs better; as figure 2 shows, the improvement is almost 100% over the energy conserving routing. In the dynamic routing case, the mobile can attract more traffic to pass through it and the traffic through bottleneck nodes can be decreased. For the remainder of this paper, we will mostly discuss the dynamic routing case. Figure 2 also shows that the mobile sink outperforms the mobile relay approach. The intuitive explanation for this is the following. Since all traffic has to be routed to the sink, the nodes around the sink are bottleneck links and become congested very easily. If the sink is mobile, these bottleneck nodes change with time and are uniformly distributed over the entire network. However, the mobile sink will have to roam at the periphery of the network. All nodes in the network will have to be aware of the position and location of the sink and it might not always be possible for the sink to relay information over the backbone. Comparing with mobile sink approach, the mobile relay approach is much more robust. The network can keep functioning by routing over static sensors alone when the mobile relay is absence (e.g. for recharging). As we will see in the next section, in a dense network, the mobile relay only needs to move within a two hop radius around the sink.

4. MOBILE RELAYS IN DENSE NETWORKS

4.1 Assumptions and Network Model

In this section we will describe our network models and basic assumptions. We assume that the sensors are densely deployed according to a Poisson point process with density λ in a circular area of radius $R \gg 1$. By dense we mean that in each hop the packet can travel as far as the transmission range in any direction. We assume that there are N sensor nodes in the network with one sink n_0 at the center of the circular area.

We assume a data logging application, where the sensors are required to send their sensing data at certain rate. Fur-

thermore, for the sake of simplicity, we assume that the data generation rate for all the sensors is one. We assume that the transmission range of all the sensors is equal to 1 and the sensors do not change their transmission powers. We define ρ as the average number of neighbors for the sink. Since the transmission range is 1, the average number of nodes in the transmission range of the sink will be $\rho = \pi\lambda$.

We assume that all sensor nodes have the same initial energy E and the energy of the sink is unlimited. For the energy consumption models, observing that there are many sleep scheduling methods to put idle sensors into sleep [1, 2], we suppose the network has adopted optimal sleep scheduling protocols. Thus, the energy consumption in idle state can be ignored and we only consider the energy used in sensing, receiving and transmission. We assume that sensors will consume e_s unit of energy for sensing data in each time unit, e_r and e_t for receiving and transmitting one packet of data, respectively. To further simplify our energy model, we assume that the transmission energy dominates the total energy consumption, so that the difference between e_s and e_r can be ignored. In our model, the total energy consumed by sensors in sending out one packet is a constant e , which is the sum of the transmission energy and the receiving (or sensing) energy. Thus, if the average number of packets flowing out from the sensor k per time unit is f_k , the lifetime of this sensor would be:

$$T_k = \frac{E}{f_k e} \quad (6)$$

In order to extend the lifetime of the network, we need to maximize the lifetime of the node who dies first. This is equivalent to minimizing $\max_{1 \leq k \leq N} \{f_k\}$.

We assume mobile relays have the same sensing ability and transmission range as the static sensor nodes and they have rechargeable batteries thus there are no energy limits on mobile relays.

To facilitate our discussion, we divide static sensors to different sets according to their distance to the sink. The set P_i contains all the nodes which can reach the sink with minimal hop count i . For example, the set of all the immediate neighbors of the sink will be P_1 . In a dense network, the sensor node n will be in the set P_k iff $k-1 < d(n, n_0) \leq k$, where $d(n, n_0)$ is the Euclidean distance between node n and the sink n_0 . Thus, the nodes in P_k will be in the k^{th} annulus around the sink as showed in figure 3. We denote the nodes outside the transmission range of the sink as $\overline{P_1}$. The set of all the nodes which can reach the sink within j hops is denoted as $Q_j = \bigcup_{k \leq j} P_k$. We use n_k to represent a node in P_k and $n_{\leq j}$ to represent a node in Q_j .

4.2 Upper Bounds on Lifetime

We will first give the upper bound on the lifetime of the static network.

In a static network, it is clear that the nodes in P_1 will be the bottleneck nodes of the network, since all traffic in the network will have to pass through some node in P_1 . Theorem 1 gives the lifetime upper bound for the static network.

THEOREM 1. *The lifetime of a dense static network is upper bounded by $\frac{E}{R^2 e}$ time units.*

PROOF. When the network density λ is large enough, there would be $\rho = \pi\lambda$ nodes in the transmission range of the sink a.s. by law of large numbers. Then the total initial energy stored in P_1 would be ρE .

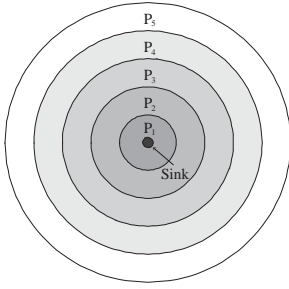


Figure 3: Dividing nodes to different subsets in the circular network

Suppose the lifetime of the network is $\tilde{T} > \frac{E}{R^2e}$. In each time unit, there would be $N = \rho R^2$ packets generated by sensors in the network, and all of them should be delivered to the sink. Since the sink can only receive data from the nodes in P_1 , both the packet generated by nodes in P_1 and in $\overline{P_1}$ must pass through nodes in P_1 at least once. Thus, the total packets passing through nodes in P_1 per time unit should satisfy:

$$\sum_{n_1 \in P_1} f_{n_1} \geq N \quad (7)$$

The total number of packets delivered by nodes in P_1 in time \tilde{T} will be

$$\begin{aligned} D &= \tilde{T} \times \sum_{n_1 \in P_1} f_{n_1} \\ &\geq \tilde{T} \times N \\ &> \frac{E}{R^2e} N \\ &= \frac{\rho E}{e} \end{aligned} \quad (8)$$

The total energy used by nodes in P_1 would be $D \times e > \rho E$. This contradicts our assumption that the total initial energy stored in P_1 is ρE . So the lifetime of the static network must be less than or equal to $\frac{E}{R^2e}$. \square

Note that the proof in Theorem 1 is the best possible lifetime over all possible routing algorithms². We now prove that the lifetime upper bound can be improved by a factor of four with one mobile relay.

THEOREM 2. *With one mobile relay, the lifetime of a dense network is upper bounded by $4\frac{E}{R^2e}$ time units.*

PROOF. The amount of traffic passing through nodes in Q_i is at least the sum of the traffic generated in $\overline{Q_i}$, which is $N - \rho i^2$ per time unit. Since the mobile relay has a transmission range of one and it can only be at one place at a time, $\overline{Q_i}$'s traffic should be relayed for at least $i - 1$ hops by

²Since we have a symmetric network, shortest path routing is equivalent to energy conserving routing

static nodes in Q_i . Since the number of nodes in Q_i is ρi^2 , we can bound the lifetime of the network by

$$\begin{aligned} T^1 &\leq \frac{\rho i^2 E}{(N - \rho i^2) \times (i - 1)e} \\ &= \frac{i^2 E}{(R^2 - i^2)(i - 1)e} \end{aligned}$$

Therefore

$$T^1 \leq \min_i \frac{\rho i^2 E}{(N - \rho i^2) \times (i - 1)e} \quad (9)$$

When $i \geq 2$, as i increases, the right hand side of the inequality (9) will monotonically increase and when $i = 1$, the right hand side is infinity. Therefore the least upper bound on lifetime is when $i = 2$. Therefore $T^1 \leq \frac{4E}{(R^2 - 4)e}$.

By investigating the problem closer, we note that in equation 9, we have only considered the traffic coming into Q_2 . By now taking into account the traffic generated by Q_2 , which also need to pass through nodes in Q_2 at least once, we can further tighten the bound to $4\frac{E}{R^2e}$. \square

What Theorem 2 shows is that the mobile relay needs to stay only within a *two hop radius* of the sink in order to maximize the lifetime.

4.3 Joint Mobility and Routing Algorithm

We assume a dense static network with one mobile relay and construct a joint mobility and routing algorithm whose lifetime is close to the upper bound derived in Theorem 2.

A broad outline of the algorithm is as follows. From Theorem 2, we know that the mobile relay needs to only stay within a two hop radius in order to maximize the lifetime. Therefore the mobility pattern of the mobile relay is as follows: Starting from the sink, the mobile relay traverses a path which forms a set of concentric circles, centered around the sink with increasing radii, until it reaches the periphery of Q_2 . It stays on each point on this path for a certain duration and relays traffic to the sink.

The routing algorithm is roughly as follows. Assume that the mobile is located at position M , then all traffic in $\overline{Q_2}$ is first aggregated to points on the line OM , where O is the position of the sink. This traffic is then directed hop by hop along the line OM until it reaches the sink. We call this routing algorithm ARA (Aggregation Routing Algorithm) for the rest of this paper.

THEOREM 3. *There exists a routing scheme which can extend the network lifetime to at least $4\frac{E}{R^2e} - \frac{16E}{R^3e}$ with one mobile node, when the network radius $R > 16\pi + 4$.³*

PROOF. First, we will use the mobile relay to build 4ρ relay paths in Q_2 . Each path will only contain one static node $n_{\leq 2}$ in Q_2 and each static node in Q_2 will only be used once.

Consider three arbitrary small areas a_1 , a_2 and a_3 illustrated in figure 4 (a), where $0 < r \leq 1$ is the distance from the area a_1 to the sink. The distance between the three areas is 1. By our definition, the nodes in area a_i will be in the set P_i . As $\Delta\theta$ and Δr are small, any node in area a_1 can directly talk to the sink and the nodes in a_2 . Also, nodes in a_2 and a_3 can communicate with each other. If we put

³Here the bound for R is very loose, we conjecture that this lifetime can be achieved with $R = 20$, since less strict methods yields $R = 20$.

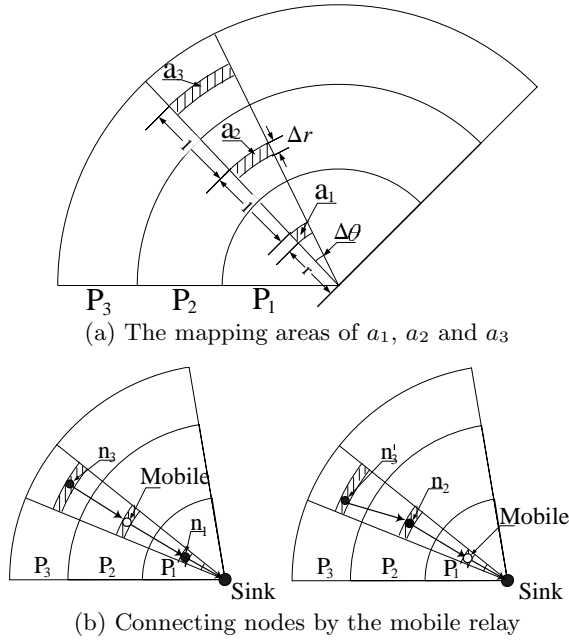


Figure 4: Aggregation node selection and the relay path construction

the mobile in area a_2 , we can connect a static node n_3 in a_3 with a static node n_1 in a_1 , thus the packet received by n_3 can flow to the sink with the help of the mobile and n_1 . Also, when the mobile is in the area of a_1 , it can connect the sink with a node n_2 in area a_2 which can draw data from another node n'_3 in area a_3 , as figure 4 (b) shows. Since the network density is $\lambda = \frac{\rho}{\pi}$, the number of nodes in area a_1 would be $x_1 = \frac{\rho}{\pi} r \Delta r \Delta \theta$. The number of nodes in area a_2, a_3 would be $x_2 = \frac{\rho}{\pi} (r+1) \Delta r \Delta \theta$ and $x_3 = \frac{\rho}{\pi} (r+2) \Delta r \Delta \theta$. For $0 < r \leq 1$, we have:

$$\begin{aligned} x_1 + x_2 &= \frac{\rho}{\pi} (2r+1) \Delta r \Delta \theta \\ &\leq \frac{\rho}{\pi} (r+2) \Delta r \Delta \theta = x_3 \end{aligned} \quad (10)$$

which means that the number of nodes in a_3 is always bigger than the sum of that in area a_1 and a_2 . Therefore, for any node in a_1 or a_2 , we can associate a unique node in a_3 to it. Varying r and θ , we can cover all the nodes in Q_2 , and build an injective mapping $f: Q_2 \rightarrow P_3$. This implies that for each node $n_{\leq 2}^r$ in Q_2 , we can associate a unique node, $n_3^a = f(n_{\leq 2}^r) \in P_3$ such that n_3^a can communicate to the sink by relaying via the mobile node and $n_{\leq 2}^r$. We call the node $n_{\leq 2}^r$ the relay node and the node n_3^a the aggregation node (in the latter discussion we will drop the subscript and just denote them as n^r and n^a). We define the range of mapping f as the aggregation set, denoted by G . What this means is that if the mobile is at a distance r_m , $0 < r_m \leq 1$ from the sink, then n^r is in Q_2 one hop from the mobile and at a distance $1 + r_m$ from the sink. A unique aggregation point n^a which is at a distance 1 from n^r and $2 + r_m$ from the sink is chosen. Similarly, n^r and n^a can be defined appropriately when $1 < r_m \leq 2$. Therefore, depending on the position of the mobile, n^r and n^a can be defined. If the mobile covers all the positions in Q_2 , then every node in Q_2 will be chosen

as n^r once⁴. If, depending on the position of the mobile, all traffic in Q_2 is routed via the associated aggregation node n^a , then we will have 4ρ unique paths from P_3 to the sink.

Suppose that we can aggregate the packets generated by nodes in Q_2 to the aggregation node n^a (later we will show this is possible), then there will be at most N packets passing through any of the 4ρ paths in each time unit. When a node n^a is aggregating all traffic from Q_2 , the rate of energy consumption for n^r and n^a will be very high. Suppose, we reserve $E' = e \times 4 \frac{E}{R^2 e} = 4 \frac{E}{R^2}$ in these nodes to send out the data that they sense, then they will have $E - E'$ units of energy for relaying traffic from other nodes in the network. Therefore, the mobile relay will stand in one location for a duration $\frac{E - E'}{Ne}$ time units before moving to the next location.

Since there are 4ρ routes in total, the total lifetime for nodes in Q_2 will at least be:

$$4\rho \times \left(\frac{E - E'}{Ne} \right) = 4 \frac{E}{R^2 e} - \frac{16E}{R^4 e} \quad (11)$$

We will now construct a mobility algorithm and a routing algorithm for the nodes in the network and show that none of the nodes in the network will deplete their energy prior to $4 \frac{E}{R^2 e} - \frac{16E}{R^4 e}$.

To get the lifetime of $4 \frac{E}{R^2 e} - \frac{16E}{R^4 e}$, we need to aggregate the traffic generated outside Q_2 to the node n^a . Figure 5 shows the mobility algorithm for the mobile relay. Note that the mobile relay remains within Q_2 . Figure 6 describes the Aggregation Routing Algorithm (ARA) in detail. We outline below the salient points of this algorithm.

1. Nodes in Q_3 :

The data generated in Q_3 will be delivered as follows: Nodes in P_1 directly send their data to the sink in one hop. Nodes in P_2 will send their data to nodes in $P_3 \setminus G$, which are the set of spare nodes which is not used as aggregation points in P_3 . Nodes in P_3 will also send their data to nodes in $P_3 \setminus G$. The task for nodes in $P_3 \setminus G$ is to redirect all the data they receive to the current aggregation node n^a using nodes in $P_3 \setminus G$ as relays.

2. Nodes in Q_3 :

- (a) The nodes in Q_3 will first relay the packets they generate to the line OM . A node in annulus P_k , $k > 3$ which is at a distance $k-1 \leq l < k$ sends the packets it sensed to a point on the line OM which is also at a distance l from the sink. It does so by relaying its traffic only via nodes which lie on the circle of radius l as shown in figure 7.
- (b) Then we use the points on OM to deliver the traffic to the aggregation point n^a . For packets generated in P_k , $k > 3$, after they reach the line OM , they will first be sent to the node on OM which is at a distance $k-1+r$ from the sink. Then the packets are sent hop by hop, passing through nodes on line OM with distance $i-1+r$, $4 \leq i < k$ from the sink, until they reach the aggregation point.

⁴Since we assume a Poisson point process, the probability that two nodes fall at the same location can be ignored.

Mobility algorithm for the mobile relay

Parameters:

(r_m, θ) : the coordinate for the mobile's position in a polar coordinate system where the sink is at the origin.

Method – Mobility management:

Set $r_m = 0, \theta = 2\pi$;

while ($r_m < 2$)

if $\theta < 2\pi$

 Set $\theta = \theta + \Delta\theta$;

 Move to the new position (r_m, θ) ;

else

 Set $r_m = r_m + \Delta r, \theta = 0$;

 Move to the new position (r_m, θ) ;

endif

if $r_m < 1$

while There exists an unselected node whose coordinate is $(r_m + 1, \theta)$

 Pick up one unselected node n whose coordinate is $(r_m + 1, \theta)$;

 Set $n^r = n, n^a = f(n)$;

 Broadcast the information about n^r and n^a to all nodes;

 Relay packets from n^r to the sink for $\frac{E-E'}{N_e}$ time units;

 Mark n as selected;

endwhile

else

while There exists an unselected node n whose coordinate is $(r_m - 1, \theta)$

 Pick up one unselected node n whose coordinate is $(r_m - 1, \theta)$;

 Set $n^r = n, n^a = f(n)$;

 Broadcast the information about n^r and n^a to all nodes;

 Relay packets from n^a to n^r for $\frac{E-E'}{N_e}$ time units;

 Mark n as selected;

endwhile

endif

endwhile

Figure 5: Mobility algorithm for the mobile relay

- (c) After aggregating the traffic at n^a , we will use one mobile and one node n^r in Q_2 to build a path from n^a to the sink.

As there are 4ρ candidates for n^a , the routing table will change as the aggregation point changes. From the symmetry of the mobility and routing algorithms described in figures 5 and 6, it is clear that the traffic load for the nodes which lie on a circle with center at the the sink will be equal. Although some nodes may be heavily burdened for a short duration, the total traffic load for nodes on a circle centered at the sink will be equal over the lifetime of the network. In the rest of this paper, we will frequently refer to nodes within a ring with width of Δr . We denote a ring of $[i-1+r, i-1+r+\Delta r]$ as $Ring_{i,r}$ in the following discussion.

Aggregation Routing Algorithm running on a static node

$n \in P_k$

Parameters:

n^a : the current aggregation node

n^r : the current relay node

r : the distance between n^a and the sink is $r + 2$

Method – ARA:

switch (k : the index of P_k where $n \in P_k$)

case 1:

if $n = n^r$

 Relay the received packet to the sink;

else

 Send the sensed data to the sink;

endif

case 2:

if $n = n^r$

 Relay the received packet to the mobile;

else

 Find a neighbor in $P_3 \setminus G$ and send sensed data to it;

endif

case 3:

if $n = n^a$

 Relay the received packet to the mobile or n^r ;

else if $n \in G$

 Find a neighbor in $P_3 \setminus G$ and send sensed data to it;

else

 Relay the received packet towards n^a using a neighbor in $P_3 \setminus G$;

endif

case 4, ..., R:

if $d(n, n_0) = k - 1 + r$ and n is on the line OM

 Find a neighbor in P_{k-1} whose distance to the sink is $k - 2 + r$ and send the packet to it;

else if n is on the line OM

 Find a neighbor on OM whose distance to the the sink is $k - 1 + r$ and send the packet to it;

else

 Find a neighbor whose is closest to line OM and has the same distance to the sink, send the packet to it;

endif

Figure 6: The Aggregation Routing Algorithm

We will investigate the energy consumption of all nodes in the network under this joint mobility and routing scheme, and show that the network lifetime will be at least $4\frac{E}{R^2e} - \frac{16E}{R^4e}$.

1) Lifetime for nodes in Q_2 :

A node in P_1 either relays traffic for the entire network or relays only its own traffic directly to the sink. For each node in P_1 , we have reserved $E' = \frac{4E}{R^2}$ units of energy for transmitting its own traffic. As we discussed in the mobility algorithm, the lifetime of the nodes in P_1 is at least $\frac{4E}{R^2} - \frac{16E}{R^4e}$. Similarly, for nodes in P_2 , since they transmit the

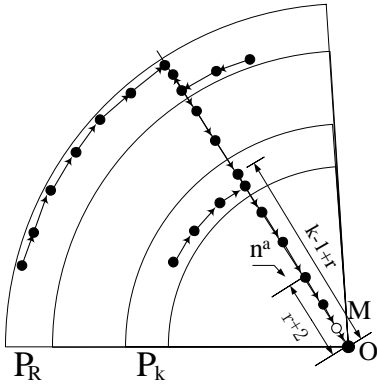


Figure 7: Packet aggregating routes with ARA

information they sense to a node in $P_3 \setminus G$ and we have reserved $\frac{4E}{R^2}$ units of energy for this, they can also live for at least $\frac{4E}{R^2} - \frac{16E}{R^4e}$ units of time.

2) Lifetime for nodes in P_3 :

Like the nodes in Q_2 , the nodes in G will be only left with enough energy to send out their own data after they have acted as aggregation points.

Therefore nodes in $P_3 \setminus G$ need to carry the task of relaying the data generated by nodes in P_3 and P_2 . There will be ρ nodes in $P_3 \setminus G$, and they must relay 8ρ packets to the aggregation node in each time unit, 3ρ for nodes in P_2 and 5ρ for nodes in P_3 . They will first deliver the packets for at most $3\pi + 1$ hops to get them to reach the line OM . Then they need one more hop to send them to the aggregation node. So, a packet from P_2 or P_3 will be relayed for no more than $3\pi + 2$ hops by nodes in $P_3 \setminus G$.

There will be $2\rho(1-r)\Delta r$ nodes in $P_3 \setminus G$ in $Ring_{3,r}$ and they will be symmetrically distributed around the ring. From the symmetry of the mobility algorithm and consequently the symmetry of the location of aggregation points, the traffic load in this ring can be evenly distributed. On average each node in $P_3 \setminus G$ will deliver packets for 8 nodes in P_2 and P_3 including itself. Here the energy consumption for delivering one packet to the aggregation point is counted in the energy budget of the first node relaying it in $P_3 \setminus G$, since the energy consumption will be evenly distributed among nodes in $Ring_{3,r}$. In order to distribute the load evenly among different rings, we need to map $16\rho(1-r)\Delta r$ nodes in P_2 and P_3 to nodes in $P_3 \setminus G$ in $Ring_{3,r}$ and deliver the packets generated by them only through nodes in this ring. Such a mapping can be built as follows: First, map P_2 and P_3 nodes to the $P_3 \setminus G$ nodes in the outermost ring of $Ring_{3,r=1}$, then decrease r and map inner rings in sequence, until we reach the innermost ring of $Ring_{3,r=0}$. When mapping nodes to a particular ring $Ring_{3,r}$, all nodes in $\{P_3 \setminus G\} \cap \{[r+2+\Delta r, 3]\}$ would have already been mapped to 8 nodes in P_2 and P_3 . Therefore, when we are at $Ring_{3,r}$, total number of nodes in P_2 and P_3 which have already been mapped will be $8 \int_{x=r}^1 2\rho(1-x)dx = \rho(8-16r+8r^2)$. The nodes in P_2 and P_3 which can communicate with a node in $Ring_{3,r}$ are in the area of $[r+1, 3]$, which has $\rho(8-2r-r^2)$ nodes in total. Since $\rho(8-2r-r^2)$ is always bigger than

$\rho(8-16r+8r^2)$ for $0 < r \leq 1$, this implies that there are unmapped nodes in P_2 and P_3 which can be mapped to $Ring_{3,r}$. This holds true for every r as we go from the outermost ring to the innermost ring. Therefore, we can build a mapping from the 8ρ nodes in P_2 and P_3 to the ρ nodes in $P_3 \setminus G$, with each node in $P_3 \setminus G$ exactly being mapped to 8 unique nodes in P_2 and P_3 . Since each node in $P_3 \setminus G$ will have to relay for 8 nodes and each packet is routed for at most $(3\pi + 2)$ hops. The lifetime for any node in $P_3 \setminus G$ is at least $\frac{E}{8(3\pi+2)e}$. When $R > 20$, we can guarantee the lifetime of them will be larger than $\frac{4E}{R^2e}$.

3) Lifetime for nodes in P_k with $k \geq 4$

The nodes in P_k , $k \geq 4$ will have to relay traffic for information generated in P_k and for information generated in Q_k . First consider the packets generated in P_k : For nodes in $Ring_{k,r}$, the packets generated in this ring will be relayed to the line OM by nodes in this ring. Each packet will travel at most π in angle before it can reach the line OM . It also needs to be relayed for one more hop to reach some node on line OM with exactly $k-3$ distance to the aggregation point. Deducting the first hop reserved for nodes to send out its own data, the maximal hops a packet will travel in $Ring_{k,r}$ will be $\pi(k-1+r) + 1$. Since the mobility and routing algorithm is symmetric, the traffic will be evenly distributed among nodes in this ring. As there are $2\rho(k-1+r)\Delta r$ nodes in $Ring_{k,r}$, in each time unit the ring will generate $2\rho(k-1+r)\Delta r$ packets. During the lifetime of the network, which is $\frac{4(E-E')}{R^2e}$, the total energy used in delivering this part of traffic will be upper bounded by:

$$\begin{aligned} E_1(k, r) &\leq 2\rho(k-1+r)\Delta r \times \frac{4(E-E')}{R^2e} \\ &\quad \times e \times (\pi(k-1+r) + 1) \\ &= \frac{8\rho(k-1+r)(\pi(k-1+r) + 1)}{R^2} \\ &\quad \times (E-E')\Delta r \end{aligned} \quad (12)$$

The next part is the packet generated by nodes in Q_{k-1} . They will be relayed for one hop to the nodes in P_{k-1} . Observe that the nodes in $Ring_{k,r}$ only will be involved in relaying this part of traffic when the distance between the current aggregation point n^a and the sink is $2+r$. As we have mentioned, there will be $2\rho(2r+1)\Delta r$ aggregation points whose distance to the sink is $r+2$. Each of them will be used for at most $\frac{E-E'}{\rho R^2e}$ time units. Then, the nodes in $Ring_{k,r}$ will need to route traffic from Q_{k-1} for at most $2(2r+1)\frac{E-E'}{R^2e}\Delta r$ time units. Since the aggregation points in circle $2+r$ are sequentially chosen in a clockwise direction, the traffic load on every node in the ring $Ring_{k,r}$ as a relay to ring $Ring_{k-1,r}$ will be equal. Therefore, we need to calculate the total energy used in this ring to ensure that no node in the ring will use up its energy prematurely. There will be at most N packets from Q_{k-1} passing through $Ring_{k,r}$ per time unit. So the total energy consumption during the lifetime will be upper bounded by:

$$\begin{aligned} E_2(k, r) &\leq 2(2r+1)\frac{E-E'}{R^2e}\Delta r \times N \times e \\ &= 2\rho(2r+1)(E-E')\Delta r \end{aligned} \quad (13)$$

Since there are $2\rho(k-1+r)\Delta r$ nodes in $Ring_{k,r}$, the total

energy that can be used for relaying will be $2\rho(k-1+r)(E-E')\Delta r$. So the total residual energy for nodes in $Ring_{k,r}$ will be lower bounded by:

$$\begin{aligned}
E_{re}(k,r) &\geq 2\rho(k-1+r)(E-E')\Delta r - [E_1(k,r) + E_2(k,r)] \\
&= 2\rho(k-r-2)(E-E')\Delta r - E_1(k,r) \\
&\geq 2\rho(k-3)(E-E')\Delta r - \frac{8\rho k(\pi k+1)(E-E')\Delta r}{R^2} \\
&\geq 2\rho(k-3)(E-E')\Delta r - \frac{8\rho(\pi k+1)(E-E')\Delta r}{R} \\
&= \frac{2\rho((R-4\pi)k-3R-4)(E-E')\Delta r}{R} \\
&\geq \frac{2\rho((R-4\pi)\times 4-3R-4)(E-E')\Delta r}{R} \\
&= \frac{2\rho(R-16\pi-4)(E-E')\Delta r}{R} \tag{14}
\end{aligned}$$

When R is bigger than $16\pi + 4$, the total residual energy will be greater than 0. Since the traffic will be distributed evenly among nodes in $Ring_{k,r}$, the residual energy will also be evenly distributed among them and none of them will die before $4\frac{E}{R^2e} - \frac{16E}{R^4e}$. The bound holds for all k and r , so no node in $\bigcup_{k>4} P_k$ will die prematurely when $R > 16\pi + 4$.

We have shown that every node in the network will have lifetime more than $4\frac{E}{R^2e} - \frac{16E}{R^4e}$, thus the joint mobility and routing algorithm that we have constructed can achieve a lifetime of at least $4\frac{E}{R^2e} - \frac{16E}{R^4e}$ with one mobile. \square

Theorem 3 shows that we can construct a routing algorithm which can achieve the lifetime of $4\frac{E}{R^2e} - \frac{16E}{R^4e}$ with one mobile. Since R^4 decays much faster than R^2 , as the network radius R becomes large, the lifetime will approach 4 times that of the static network with one mobile relay.

4.4 Routing With Few Nodes Aware of Mobile Location

In the previous discussion, we constructed the ARA which can achieve the lifetime of $4\frac{E}{R^2e} - \frac{16E}{R^4e}$ when R is large enough. In ARA, every node in the network needs to know the position of the mobile and appropriately route traffic. However this implies large overheads in disseminating knowledge of the location of the mobile node to all the nodes in the network. We could of course argue that since the mobility algorithm is deterministic, it will only involve a one time dissemination of information to the nodes in the network. However, in a distributed sensor network, synchronization is extremely difficult and over time nodes in the network may make incorrect assumptions about the location of the mobile node. In this section, we show that we can construct a routing algorithm, whereby only a limited number of nodes in the network need to know the location of the mobile relay. More interestingly, we show that with this routing algorithm, we can still achieve a lifetime bound of $4\frac{E}{R^2e} - \frac{16E}{R^4e}$. We call this routing algorithm as ARALN (Aggregation Routing Algorithm with Limited Nodes).

ARALN is described in detail in figure 9. We outline the ideas of the algorithm below.

1. Nodes which are outside the circle with radius s do not need to know the position of the mobile and they can use shortest path routing algorithm to send their packets towards the sink.

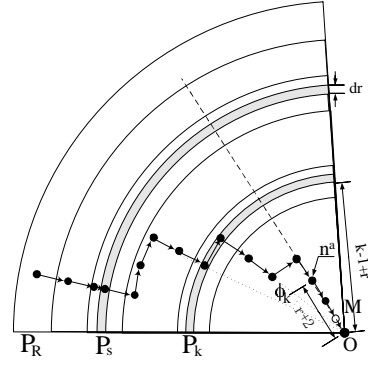


Figure 8: Packet aggregating routes with ARALN

2. Once the information from $\overline{Q_s}$ reaches P_s , it is relayed in one hop to the aggregation ring - $Ring_{s,r}$ in P_s , where the distance from the aggregation point n^a to the sink is $2+r$. Once it reaches a node in this aggregation ring, it will be delivered by a series of aggregation rings - $Ring_{i,r}, 4 \leq i \leq s-1$ until it reaches n^a . In each aggregation ring, it will be relayed around an angle ϕ_i within $Ring_{i,r}$ before it is relayed to the next aggregation ring - $Ring_{i-1,r}$. When this traffic reaches the line OM , it is then routed hop by hop along OM as before. This is shown in figure 8.
3. Packets generated by nodes in Q_s are routed as in the ARA described in figure 6.

THEOREM 4. *With Aggregation Routing Algorithm with Limited Nodes, the network lifetime is lower bounded by $4\frac{E}{R^2e} - \frac{16E}{R^4e}$, for $s = 22, R > 84$.⁵*

PROOF. The proof of the lifetime for this algorithm is outlined in Appendix A. \square

4.5 Network with m Mobile Relays

In this section, we will extend our discussion to a network with m mobile relays. When we have m mobile nodes in the network, they will stay within Q_{2m} and get nearly $4m$ times lifetime when R is large.

THEOREM 5. *The lifetime of a uniform dense network with m mobile relays is upper bounded by $4m\frac{E}{(R^2-4m^2)e}$ time units.*

PROOF. Consider the traffic load in Q_i with $i \geq m$. The traffic generated in $\overline{Q_i}$ would be $N - \rho i^2$ per time unit. This traffic will be relayed at least $i - m$ times by static nodes in Q_i . The number of nodes in Q_i is ρi^2 . Constraining the m mobile nodes to remain within Q_i and using similar arguments in the proof of Theorem 2, we can bound the network lifetime by:

$$\begin{aligned}
T^m &\leq \frac{\rho i^2 E}{(N - \rho i^2) \times (i - m)e} \\
&= \frac{i^2 E}{(R^2 - i^2)(i - m)e} \tag{15}
\end{aligned}$$

⁵Here the bounds for both R and s are quite loose, we suspect we can achieve the lifetime with s near 10 and R near 40.

ARALN running on a static node $n \in P_k$

Parameters:

- n^a : the current aggregation node
- n^r : the current relay node
- r : the distance between n^a and the sink is $r + 2$

Method – ARALN:

switch (k :the index of P_k where $n \in P_k$)

case 1, 2, 3:

Call method ARA;

case 4, . . . , $s - 1$:

if $d(n, n_0) = k - 1 + r$
if the packet is generated in $\overline{Q_{s-1}}$ and it has traveled ϕ_k in P_k
Find a neighbor in P_{k-1} whose distance to the sink is $k - 2 + r$ and send the packet to it;
elseif the packet has reached line OM
Find a neighbor in P_{k-1} whose distance to the sink is $k - 2 + r$ and send the packet to it;
else
Find a neighbor who is closest to line OM and whose distance to the sink is $k - 1 + r$, send the packet to it;
endif
else if n is on the line OM
Find a neighbor on OM whose distance to the sink is $k - 1 + r$ and send the packet to it;
else
Find a neighbor whose is closest to line OM and has the same distance to the sink, send the packet to it;
endif

case s :

if $d(n, n_0) = k - 1 + r$
Find a neighbor in P_{k-1} whose distance to the sink is $k - 2 + r$ and send the packet to it;
else
Find a neighbor whose distance to the sink is $k - 1 + r$, send the packet to it;
endif

case $s + 1, \dots, R$:

Find a neighbor who is closest to the sink, send the packet to it;

Figure 9: The Aggregation Routing Algorithm with Limited Nodes

For $i < m$, such bound will be infinity. For $i \geq m$, the function $\frac{i^2}{i-m}$ takes the smallest value at $i = 2m$. So, set $i = 2m$ we get the least upper bound, which is given by $4m \frac{E}{(R^2 - 4m^2)e}$ time units. \square

Notice that, here the bound is looser than the one we derived in Theorem 2.

THEOREM 6. *There exists a routing scheme which can extend the network lifetime to $4m \frac{E}{R^2e} - \frac{32\pi m^3 E}{R^4e}$ with m mobile nodes, when R is large enough.*

For the sake of brevity, we just state this result without pro-

viding the details in route construction. The improvement factor for m mobiles is $4m - \frac{32\pi m^3}{R^2}$; when R is large enough, we can get an improvement factor of $4m$ for m mobile relays.

5. MOBILE SINK VERSUS MOBILE RELAYS

5.1 The Performance of the Mobile Sink Approach

As shown in figure 2, the mobile sink approach in [7] can do better than the mobile relay approach. Here we will show that when the network is dense and large, the mobile sink approach can give an improvement factor of $O(R)$ on network lifetime.

We suppose a mobile sink is moving around the perimeter of the network, as shown in figure 10(a). We will construct a routing scheme which can achieve network lifetime of at least $\frac{E}{(2+\pi)Re}$.

In each ring around the center O of the network with radius $r + i - 1$ and width Δr , we try to forward all the packets generated in this ring to the nodes positioned on the line OS which is the line between the center O and the sink n_0 . Then the packets are forwarded toward the sink on the line OS with one hop progress distance of 1, except the last hop to reach the sink.

The traffic load for nodes in $Ring_{i,r}$ will be evenly distributed among all the nodes in it, since our scheme is symmetric and the movement of the sink is also symmetric.

The traffic load of the ring $Ring_{i,r}$ will be composed of two parts:

First, they should forward for those nodes in rings of $Ring_{j,r}$, $1 \leq j \leq i$. Since there are $2\rho(r + j - 1)\Delta r$ nodes in $Ring_{j,r}$, there would be $\sum_{j=1}^i 2\rho(r + j - 1)\Delta r = \rho i(i - 1 + 2r)\Delta r$ packets passing through the nodes in $Ring_{i,r}$ per time unit.

Second, the packets generated in $Ring_{i,r}$ will need to travel for at most $\pi(r + i - 1) + 1$ hops to reach the line OS .

Then the average traffic load for the $2\rho(r + i - 1)\Delta r$ nodes in $Ring_{i,r}$ per time unit would be:

$$\begin{aligned}
Load(i, r)' &= \frac{\rho i(i - 1 + 2r)\Delta r}{2\rho(r + i - 1)\Delta r} + \pi(r + i - 1) + 1 \\
&= \frac{i(i - 1 + 2r)}{2(r + i - 1)} + \pi(r + i - 1) + 1 \\
&< \frac{i}{2} \left(1 + \frac{r}{r + i - 1}\right) + \pi R + 1 \\
&\leq i + 1 + \pi R \\
&< (2 + \pi)R
\end{aligned} \tag{16}$$

So the network lifetime would at least be $\frac{E}{(2+\pi)Re}$, which is $O(R)$ times better than the all static network.

5.2 The Performance of the Mobile Relay Approach

Here, we will show that to achieve a lifetime improvement of $O(R)$ over an all static network, we require $O(R)$ mobile relays. This can be seen from Theorem 5, with m mobile relays, the lifetime improvement factor is upper bounded by $4m$. This implies that we need at least $O(R)$ mobile relays

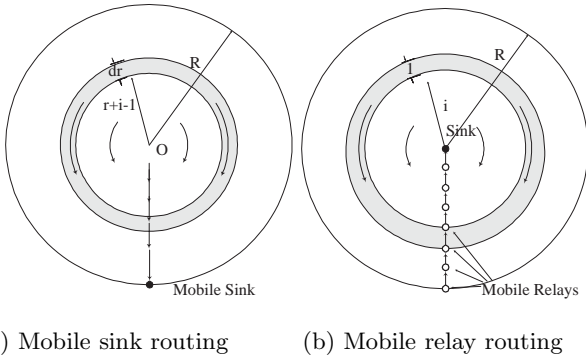


Figure 10: Comparing the mobile sink and the mobile relay approach

to get the same lifetime improvement as that of a mobile sink approach.

Furthermore, with $O(R)$ mobile relays, we can actually achieve the improvement factor of $O(R)$. Suppose that we have R mobile relays lining up from the network center to the periphery of the network with distance of 1 between each other, as shown in figure 10(b). The mobile relays will circle around the network while keeping this form to evenly distribute the network load. There will be one mobile relay in each area P_i . The static node in P_i will send their data to the mobile relay in P_i in a similar way as the mobile sink case. After the traffic is gathered by the mobile relays, they will be delivered hop by hop by mobile relays to the sink staying at the center of the network. We can easily see that the network load for the R mobile relay case is smaller than the mobile sink case which we have just investigated by comparing the routing scheme in different approaches. Thus, with R mobile nodes, we can achieve $O(R)$ improvements as the mobile sink solution.

6. EXPERIMENT RESULTS

As the analysis in Section 4 is based on dense and large network assumption, in this section we will show through simulation how mobile relay will work in a finite network, for example, network with 200 nodes and moderate node density.

Our simulation is based on the simplified energy model stated in Section 4 without considering the MAC or physical layer. The sensors are randomly deployed on regions with different size and shape. For each instance of deployment, we calculate the lifetime for the static network through the linear optimization algorithm as described in [3]. The lifetime for network with one mobile is obtained through the dynamic routing scheme described in Section 3. The result is averaged over 50 instances for each set of network size and density.

6.1 Network Lifetime

Figure 11 shows the lifetime for static networks and networks with one mobile in a circular region. The lifetime is normalized by setting the maximal node lifetime of $\frac{E}{e}$ to 1. For the static network, the optimal lifetimes are quite close to the upper bound of $\frac{E}{R^2 e}$. Also the lifetime curves for the all static network with different network densities almost overlap with each other. This shows that the lifetime of a static network is independent of the network density

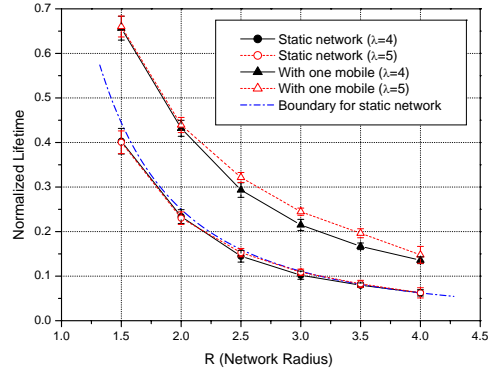


Figure 11: Network lifetime for nodes randomly deployed on a circular region, for $\lambda = 4$ and 5 (confidence interval 95%)

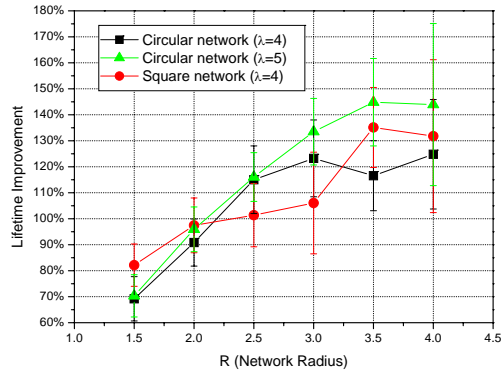
as shown in Theorem 1. For networks with one mobile relay, when the network density increases, the network lifetime also increases. The reason for this is that we assumed very high density in our theoretical derivation, as the network density increases, the lifetime improvement ratio will increase and approach a value of 4.

6.2 Lifetime improvement with One Mobile Relay

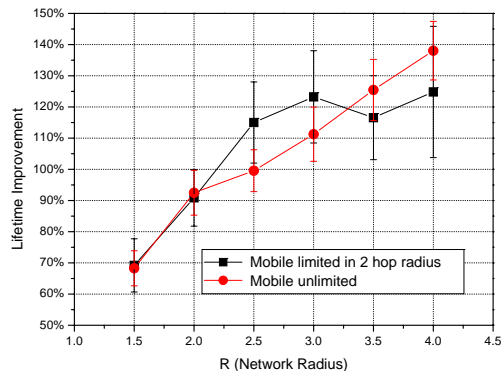
Figure 12(a) shows the lifetime improvement for networks with one mobile going around in a two hop radius of the sink. Although we assumed dense and large networks in Section 4 to derive the lifetime bounds, the improvement for network with moderate size and density is still considerable. For example, in a network deployed on a circular region with radius 4 and density $\lambda = 4$, which has about 200 nodes, we can improve its lifetime by 130% with only one mobile relay. Figure 12 also shows that the improvement ratio increases as the network size and density increase. This coincides with the intuition that as the network becomes larger and denser, the lifetime improvement will approach the bound of 300%. Due to the complexity of the linear optimization problem, we only conduct the experiments on networks with $R \leq 4$, which is much smaller than network radius requirement of 84 in section 4. Nevertheless, the results still suggest that the improvement ratio is a nondecreasing function of the network size and density.

Another observation is that the mobile relay approach can also work in non circular network regions. Although we only investigated the case for circular network regions in Section 4, our theory can be easily extended to non circular network regions. For a static network with N static nodes, the lifetime will be bounded by $\frac{\rho E}{N e}$. When a mobile node is introduced in the network, it can work in a similar way as in the circular network. Figure 12(a) shows the lifetime improvement for network deployed on a square area is similar to the one on circular regions.

Our discussion in section 4 shows that the mobile relay only need to go around in the two hop range of the sink to fully extend the network lifetime. Figure 12(b) compares the lifetime improvement between optimizations with lim-



(a) Comparing between different network settings



(b) Comparing between different mobile moving range (nodes randomly deployed on a circular region with $\lambda = 4$)

Figure 12: Average lifetime improvement for networks with one mobile over the static network (confidence interval 95%). For the square network, the sink is putted at the center of the square and the network radius R is defined as half the side length of the square.

ited mobile moving range and unlimited moving range. The results shows that the lifetime improvement are quite close, which means the mobile also seldom go out of two hop range in a finite network.

7. DISCUSSION AND CONCLUSION

To show the advantage of mobility, we will compare our approach with two approaches which rely on using additional static nodes to improve performance. In the first approach, we add more simple static devices, while in the second, we add more resource rich static nodes,

i) Simple Static Nodes

One way to increase the network lifetime is to redeploy more static nodes in the area near the sink. These additional static nodes serve as reservoirs of energy. Normally these nodes are in the sleep state. Only when nodes near the sink use up their energy do these additional nodes wake up to take over the relay tasks. To achieve a lifetime improvement of 4 by this approach, we need to increase the density of

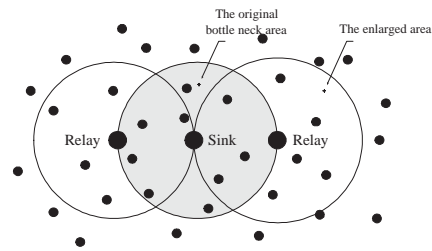


Figure 13: Adding static relays to mitigate the traffic load of bottle neck nodes

nodes within two hop radius of the sink. It is easy to see that at least 4ρ additional static nodes are required to achieve the same performance as the network with one mobile relay.

ii) Resource Rich Static Nodes

Assume that instead of mobile nodes, we add resource rich static nodes as relays. These static relay nodes have far greater energy but the same communication range as the static nodes. Given a limited number of resource rich static nodes, the most obvious choice of location for these devices will be near the sink such that they can directly communicate with the sink. Therefore, the amount of traffic flowing through the original bottleneck (or minimum lifetime) nodes will be reduced. The new set of bottleneck nodes, will be the set of neighbors of the sink and the set of neighbors of the resource rich static nodes. For example, consider figure 13. Assume we add two static relays which can directly communicate with the sink. Some of the relieving burden of the bottleneck nodes around the sink is mitigated. In other words, in this new network, the set of bottleneck nodes which determine the lifetime of the network will be the nodes which are one hop neighbors of the sink and the two relays. Each of these new bottleneck nodes will carry far less traffic than the original network, thereby improving the lifetime of the network. Therefore, one can view, the addition of these static devices as increasing the set of bottleneck nodes in the network, but reducing the network load on each of them. Since the static relay must be connected to the sink or other relays, adding one static relay can wield its influence over an area of at most 0.6090π , as can be seen in figure 13. Thus, one static resource rich relay can increase the network lifetime by at most 60% over the original network, while one mobile relay can increase 300% as we have shown previously.

In conclusion, we have investigated the possibility of using a heterogeneous network composed of many simple static nodes and a few mobile nodes. We have investigated two possibilities, (i) using the mobile node as a relay and (ii) using the mobile node as a sink. Although, it is clear from our analysis that using a mobile sink is always beneficial in terms of the lifetime of the network, but there are application scenarios in which a mobile sink is not feasible. We show that even with one node as a mobile relay, we can get a lifetime improvement of up to four times over the static network. Another interesting property of this relay approach is that we only need to change the routing algorithm for a relatively small area to use the mobile relay. As shown in our proof, we only need the nodes within 22 hops of the sink need to be aware of the location of the mobile relay. Although 22 seems like a large number, it must be noted that

this is a very loose bound that we have derived. We suspect that this bound on s can be improved substantially to the order of around 10 hops. This is motivated by equation 18 in appendix A. Furthermore, the mobile need not travel all around the network. It never needs to venture further than two hops from the sink.

Our analysis makes some simplifying assumptions about the density and the topology of the network. We further have assumed a data logging application, with the sensors incapable of power control. In such a scenario it is reasonable to assume that the power consumed in transmitting a packet over one hop is constant, if we assume some underlying MAC layer mechanism. Moreover, if most of the nodes are static and the environment is also mostly static, it is reasonable to assume that the wireless channel is time invariant and therefore the energy consumption per packet transmission is constant.

8. REFERENCES

- [1] R. Zheng, J. C. Hou, and L. Sha, "Asynchronous wakeup for ad hoc networks," in *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, Jun 2003.
- [2] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 21st IEEE INFOCOM*, Jun 2002.
- [3] J. H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proceedings of the 19th IEEE INFOCOM*, Mar 2000.
- [4] N. Sadagopan and B. Krishnamachari, "Maximizing data extraction in energy-limited sensor networks," in *Proceedings of the 23rd IEEE INFOCOM*, Mar 2004.
- [5] S. Gandham, M. Dawande, R. Prakash, and S. Venkatesan, "Energy-efficient schemes for wireless sensor networks with multiple mobile base stations," in *Proceedings of IEEE GLOBECOM*, Dec 2003.
- [6] Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli, "Exploiting sink mobility for maximizing sensor networks lifetime," in *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS)*, 2005.
- [7] J. Luo and J. P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proceedings of the 24th IEEE INFOCOM*, Mar 2005.
- [8] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks," in *Proceedings of the IEEE Workshop on Sensor Network Protocols and Applications (SNPA)*, 2003.
- [9] A. Chakrabarti, A. Sabharwal, and B. Aazhang, "Using Predictable Observer Mobility for Power Efficient Design of Sensor Networks," in *The second International Workshop on Information Processing in Sensor Networks (IPSN)*, Palo Alto, CA, April 2003.
- [10] A. Kansal, A. Somasundara, D. Jea, M. Srivastava, and D. Estrin, "Intelligent fluid infrastructure for embedded networks," in *Proceedings of the 2nd international conference on Mobile systems, applications, and services (MobiSYS)*, 2004.
- [11] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*. ACM Press, 2004, pp. 187–198.
- [12] A. Shankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," in *Proceedings of the 23rd IEEE INFOCOM*, Mar 2004.
- [13] J. Chou, D. Petrovic, and K. Ramchandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *Proceedings of the 22nd IEEE INFOCOM*, Mar 2003.
- [14] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS)*, 2000.
- [15] O. Younis and S. Fahmy, "Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach," in *Proceedings of the 23rd IEEE INFOCOM*, Mar 2004.
- [16] N. Li, J. Hou, and J. Sha, "Design and analysis of an MST-based topology control algorithm," in *Proceedings of the 22nd IEEE INFOCOM*, 2003.
- [17] J. Pan, Y. Hou, L. Cai, Y. Shi, and S. Shen, "Topology control for wireless sensor networks," in *Proceedings of the 9th annual international conference on Mobile computing and networking (MobiCom)*, 2003.
- [18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc, 1991.

APPENDIX

A. NETWORK LIFETIME FOR ARALN

Theorem 4: With Aggregation Routing Algorithm with Limited Nodes, the network lifetime is lower bounded by $4\frac{E}{R^2e} - \frac{16E}{R^4e}$, for $s = 22$, $R > 84$.

PROOF. Here we will show that all the nodes can stay alive for at least $4\frac{E}{R^2e} - \frac{16E}{R^4e}$ time units under ARALN described in section 4. We will also show that s is a constant and only nodes in a limited area have to know the position of mobile relay.

1). Lifetime for nodes in Q_3

Since the ARALN in Q_3 is same as the ARA, the lifetime for nodes in Q_3 can be derived in the same way as in ARA.

2). Lifetime for nodes in P_k with $4 \leq k < s$

As in the proof of Theorem 3, the total energy consumption for relaying packets generated in the aggregation ring will be $\frac{8\rho(\pi(k-1+r)+1)(k-1+r)(E-E')}{R^2}\Delta r$. For $k \leq \frac{R}{4}$, this part is further bounded by $2\rho(E-E')\Delta r$. After the packets get to the line OM , these packets will be treated as the packets generated in $\overline{Q_s}$ and we will have to take into account the energy consumption for relaying these packets.

For the data generated in $\overline{Q_k}$, the nodes in the aggregation ring of $Ring_{k,r}$ will use their residual energy for two tasks: First, they need to forward the packets one hop in the normal direction to make them reach the nodes in the next ring $Ring_{k-1,r}$; Second, they should relay the packet in the tangential direction for a certain angle ϕ_k to deliver the traffic to the line OM .

As shown in the proof of Theorem 3, the nodes in a par-

ticular aggregation ring will only relay packets for at most $2(2r+1)\frac{E-E'}{R^2e}\Delta r$ time units. Since nodes in $\overline{Q_k}$ can not generate more than N packets per time unit, the total number of packets to be relayed by the aggregation ring in the network lifetime will not exceed $2\rho(2r+1)\frac{E-E'}{e}\Delta r$ packets.

There will be $2\rho(k-1+r)\Delta r$ nodes in the aggregation ring in P_k , so the total energy can be used in relaying for others would be $2\rho(k-1+r)(E-E')\Delta r$. Delivering $2\rho(2r+1)\frac{E-E'}{e}\Delta r$ packets to the next ring will consume $2\rho(2r+1)(E-E')\Delta r$ energy. For $k \leq \frac{R}{4}$, the nodes in the aggregation ring can spend at least $2\rho(k-1+r)(E-E')\Delta r - 2\rho(2r+1)(E-E')\Delta r - 2\rho(E-E')\Delta r = 2\rho(k-r-3)(E-E')\Delta r$ energy to relay the packets in the tangential direction towards line OM . The angle ϕ_k , by which every packet generated by $\overline{Q_k}$ can travel in P_k , would be:

$$\begin{aligned}\phi_k &= \min_r \frac{2\rho(k-r-3)(E-E')\Delta r}{(k-1+r) \times 2\rho(2r+1)\frac{E-E'}{e}\Delta r \times e} \\ &= \min_r \frac{k-r-3}{(k-1+r)(2r+1)} \\ &= \frac{k-4}{3k}\end{aligned}\quad (17)$$

Each packet will at most need to be relayed for an angle of π to reach the line OM . Considering that we relay the packets in discrete steps, some packets may over shoot. For example, when the packets are routed around the circle with radius 3 (the inner circle for P_4), one hop will carry the packet for an angle of 0.335. Since the over shoot will never exceed the largest step, which is 0.335, the total angle a packet need to travel will not exceed $\pi + 0.335$. In ARALN, each packet will be relayed by ϕ_k in the aggregation ring in P_k before it reaches the line OM . For $s = 22$, the total angle packets can travel is $\sum_{k=4}^{21} \phi_k = 3.584$, which is larger than $\pi + 0.335$. So we only need to use nodes in $\bigcup_{k=4}^{22} P_k$ to aggregate the traffic. As we assumed that $k \leq \frac{R}{4}$ for $4 \leq k \leq 21$, then R should be bigger than 84 in this case.

3). Lifetime for nodes in P_k with $k = s$

For the nodes in P_s , first, they need to send their data to nodes in $Ring_{s,r}$, then the nodes in this aggregation ring will relay the packet to nodes in $Ring_{s-1,r}$. Since the nodes in $P_k, k > s$ will not know the position of the mobile, they will deliver their packets via the shortest path. The traffic passing through $Ring_{k,r}, k \geq s$ will be $\sum_{l=k}^R 2\rho(l-1+r)\Delta r = \rho(R-k+1)(R+k+2r-2)\Delta r$. So the nodes in $Ring_{s,r}$ should send out $\rho(R-s+1)(R+s+2r-2)\Delta r$ packets per time unit. In addition, they need to deliver $2\rho(2r+1)\frac{E}{e}\Delta r$ packets to the next ring in total. The traffic load for this task is $\frac{2\rho(2r+1)\frac{E}{e}\Delta r}{4\frac{E}{R^2e}}$ per time unit. Then the average load per time unit for one node in $Ring_{s,r}$ will be:

$$\begin{aligned}Load(s,r) &= \frac{\rho(R-s+1)(R+s+2r-2)\Delta r + \frac{2\rho(2r+1)\frac{E}{e}\Delta r}{4\frac{E}{R^2e}}}{2\rho(s-1+r)\Delta r} \\ &= \frac{2(R-s+1)(R+s+2r-2) + R^2(2r+1)}{4(s-1+r)} \\ &\leq \frac{2(R-s+1)(R+s) + 3R^2}{4(s-1)} \\ &< \frac{7R^2}{4(s-1)}.\end{aligned}\quad (18)$$

As we have shown that s is much bigger than 8, the load for one node in the aggregation ring is well below $\frac{R^2}{4}$ and the lifetime of these nodes will exceed $4\frac{E-E'}{R^2e}$.

4). Lifetime for nodes in P_k with $k > s$

For the nodes in P_k with $k > s$, they only need to relay traffic passing through them for one hop towards the sink. It is easy to see that the load for these nodes is smaller than nodes in P_s . So the lifetime of nodes in P_k with $k > s$ will also exceed $4\frac{E-E'}{R^2e}$. \square