Extending the Lifetime of Wireless Sensor Networks through Mobile Relays

Wei Wang, *Student Member, IEEE,* Vikram Srinivasan, *Member, IEEE,* and Kee-Chaing Chua, *Member, IEEE*

***Abstract*— We investigate the benefits of a heterogeneous architecture for wireless sensor networks composed of a few resource rich mobile relay nodes and a large number of simple static nodes. The mobile relays have more energy than the static sensors. They can dynamically move around the network and help relieve sensors that are heavily burdened by high network traffic, thus extending the latter's lifetime. We first study the performance of a large dense network with one mobile relay and show that network lifetime improves over that of a purely static network by up to a factor of four. Also, 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 can yield a network lifetime close to the upper bound. The advantage of this algorithm is that it only requires a limited number of nodes in the network to be aware of the location of the mobile relay. Our simulation results show that one mobile relay can at least double the network lifetime in a randomly deployed WSN. By comparing the mobile relay approach with various static energy provisioning methods, we demonstrate the importance of node mobility for resource provisioning in a WSN.**

***Index Terms*— Sensor networks, Network lifetime, Mobile relay.**

I. Introduction

Wireless sensor networks (WSNs) 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 also severely restricted in the resources that they possess. Each is powered by batteries and has limited processing and memory capabilities. For example, the Berkeley mote is powered by two alkaline AA batteries and has only a 4MHz processor with 128kB of instruction memory and 4kB of RAM.

One of the great challenges for WSN designers is to use such resource constrained sensors to guarantee certain network requirements, such as network lifetime, sensing coverage and end-to-end delay. One possible solution is to deploy a dense homogenous network, i.e., cheap sensors are scattered densely to increase the amount of resources deployed per unit area. For example, we can deploy sensors several times denser than required, then design a scheduling scheme to make them work in batches, so that the total network lifetime can be extended [1]. However, dense deployment brings many problems, such as difficulties in network management and severe MAC contentions. Another possible solution, which we consider in this paper, is to deploy a heterogenous network, having a few resource rich (in terms of processing, memory and energy) mobile nodes in addition to a large number of simple low cost static nodes. Unlike the "Data Mule" solution [2], where mobile nodes are used for carrying data packets, we use mobile nodes to dynamically distribute network resources, such as, energy, computational power, sensing and communication abilities. Mobility gives us a more efficient way to meet the network resource requirements, e.g., mobile nodes can move to the areas where resources are most needed, such as areas where node density is low due to the randomness in deployment or areas where more resources are required for increased sensing activities. Therefore, the resources carried by the mobile nodes can appear at the right place and time to be used efficiently. In the latter discussions, we will demonstrate that adding a few resource rich mobile nodes can provide same performance as increasing the network density by several times. Thus, the heterogenous network approach is more effective in terms of hardware cost than the dense deployment approach.

In this paper, we are motivated to investigate what performance improvement can accrue from mobile devices and the trade-offs associated with a heterogenous network architecture. We focus on using mobile nodes which have more energy than static sensors to extend the network lifetime. Static sensors only have limited energy from nonrechargeable batteries. Once the battery runs out, the sensor will die. Therefore it is critical that this energy 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 ([3], [4]) and Network layers ([5], [6]), however, most of these proposals are based on the assumption that the entire network is composed of static nodes.

We use the mobile nodes as relays to help energy limited static sensors in packet relaying, which is one of the most energy consuming tasks. The mobile relays have the same communication range and sensing ability as the static sensors. Fig. 1 shows an example of how a mobile relay may work. Suppose that the network is composed of two components which are connected via two sensor nodes A and B. Since all traffic flows between these two components pass through these two sensors, their battery will drain quickly. Suppose

Authors are with Department of Electrical and Computer Engineering, National University of Singapore, Email: {wang.wei, elevs, eleckc}@nus.edu.sg.

A shorter version of this work appeared in ACM MobiCom 2005 [24].

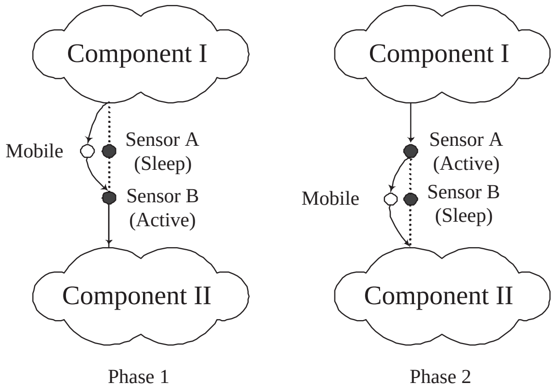


Fig. 1. One mobile relay can take over the tasks of *multiple* bottleneck nodes (sensor A and B) during different network time periods.

the lifetime of these two nodes is T. Even if other sensors have lifetime much longer than T, the network will get partitioned when these two sensors die at time T. If we have one mobile relay helping them, 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 static nodes. This example shows that a single mobile relay can greatly increase the network lifetime when there are bottlenecks in the network. In this paper, we focus on a common sensor network structure with a single sink at the center, where the network bottleneck is around the sink.

We first consider a large densely deployed WSN 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 advantage of this routing algorithm is that only nodes within a certain distance of the sink need to be aware of the location of the mobile relay. This algorithm can also be extended to the case when there are m mobile nodes and provide improvements close to $4m$ times.

The performance of a mobile relay is further studied in the case of finite and random networks. We pose the problem of maximizing lifetime as a linear programming problem and derive the optimal schedule for the mobile node. The system model used here for mobile relay is similar to the one for mobile sink in [7], [8]. The performance of the mobile relay is compared with minimal hop routing, energy-conserving routing and the mobile sink approach proposed in [9]. We show that using a mobile relay is better than most of the static energy provisioning methods. However, the mobile sink approach always out performs the mobile relay approach. Actually, 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 a mobile sink. 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 network. 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.

The main contributions of this paper are as follows:

1) We proposed a new way for resource redistribution in wireless sensor networks, which uses resource rich mobile nodes to help simple static senors. We demonstrate the usefulness of this approach by showing that in the ideal case, one energy rich mobile node can improve the lifetime of a large and dense network by 4 times.

2) We derived the network lifetime improvement upper bound for single mobile relay and multiple mobile relays. We then construct a joint mobility and routing algorithm to show that this bound is asymptotically achievable in large and dense networks.

3) We study the performance of mobile relay in random and finite network by formulating it as a linear programming problem. We compare the performance of mobile relay to various other static and mobile approaches and show the advantages of mobile relay approach.

The rest of the paper is organized as follows: Section II summarizes related work. Section III investigates the performance of a large dense network with a few mobile relays and gives a joint mobility and routing algorithm. Section IV gives the simulation results on finite random networks. Section V compares the performance of the mobile relay approach with other approaches. Finally, section VI concludes the paper.

II. Related Work

Existing literature utilizes mobile nodes as mobile sinks to save energy. Shah *et al.* [2] propose to use randomly moving "Data Mules" for data gathering. Mobile sinks with predictable and controllable movement patterns are studied in [10], [11], [12]. In these approaches, the static sensors only send out their data when the sink is close enough to them. The disadvantage of such proposals is that there will be considerable delay in packet delivery, since a node needs to wait for the sink to approach it. In order to minimize the delay, several methods of transmitting the sensed data through multi-hop communication to the mobile sink are proposed in [7], [8], [9], [13]. 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 a mobile sink can find the optimal time schedule to stay at these predefined points [7], [8]. Another method is introduced in [9], [14], 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. In most networks, 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. As we will show later, the mobile relay approach is simpler and more robust, since the network can keep operating even when the mobile relay leaves the network for recharging. Also, the mobile relay only needs to move within a small area, while the mobile sink solution requires the mobile to roam around the periphery of the network to maximize the network lifetime [9].

Another large category of energy conserving methods is to use flow control algorithms to find the optimal energy conserving routes [5], [15]. The energy conserving routing and the mobile sink approach share the same idea of distributing the traffic load evenly around the network so that the lifetime of the network is maximized. Energy provisioning in static sensor networks is studied in [16], where a total amount of energy is added in relay nodes deployed at selected positions. Such static relay nodes can heal the topology defects in randomly deployed networks, so the network lifetime can be improved greatly when the network is sparse. However, as the network density grows beyond a certain threshold, the improvement gets saturated since most of the topology defects have been mitigated. Compared to the static relay approach, the mobile relay approach can provide considerable improvements on dense networks as well as healing the topology defects.

Other solutions for energy saving have also been intensely studied, including data aggregation and topology control methods. Data aggregation and clustering methods such as [17], [18], [19] aggregate the sensed data to decrease traffic volume and thereby prolong network lifetime. Topology control methods such as [20], [21] 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.

III. Mobile Relays in Dense Networks

A. Assumptions and Network Model

We assume that there are N sensors uniformly distributed in a circular area of radius R, which is much larger than the communication range of sensors. There is only one sink n_0 at the center of the circular area. We assume the network density $\lambda = N/\pi R^2$ is large, so that in each hop the packet can travel as far as the transmission range in any direction and the number of sensors in area A is λA almost surely.

We consider a data logging application, where the sensors are required to send their sensed data at a fixed rate. Furthermore, for simplicity, the data generation rates for all the sensors are the same, normalized to one packet per unit time. The transmission range of all sensors is equal to 1 and the sensors do not change their transmission powers. Let ρ be 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 $\rho = \pi\lambda$.

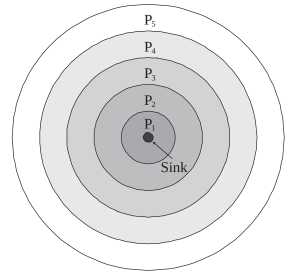


Fig. 2. Dividing nodes to different subsets in the circular network

We assume that all static sensors have the same initial energy E and the energy of the sink is unlimited. For energy conservation, the network adopts certain sleep scheduling protocols [3], [4]. Thus, the energy consumption in idle state can be ignored and we only consider the energy used in sensing, receiving and transmitting. Suppose that sensors consume e_s unit of energy for sensing data in each time unit, e_r and e_t for receiving and transmitting one packet, respectively. To further simplify our energy model, we assume that the difference between e_s and e_r is small compared to $e_r + e_t$. Thus, the total energy consumed by a sensor in transmitting one packet is a constant e, which is the sum of the transmitting energy and the receiving (or sensing) energy. If the average number of packets flowing out from sensor k per time unit is f_k, the lifetime of this sensor will be:

$$T_k = \frac{E}{ef_k} \tag{1}$$

In this paper, the lifetime of the whole network is defined as the time that the first node dies as in [5]. Since energy conserving routing is used, the network gets partitioned when the first node dies [22].

We assume that the sensor network contains a small number of mobile relays, which can move around to improve the network performance. Mobile relays have the same sensing ability and transmission range as the static sensors but they have rechargeable batteries and thus have no energy limits. To facilitate our discussion, we divide the static sensors into 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, sensor n will be in the set P_k iff $d(n, n_0) \in (k-1, k]$, 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 shown in Fig. 2. 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 $Q_j = \bigcup_{k \le j} P_k$. We use n_k to represent a node in P_k and $n_{\le j}$ to represent a node in Q_j.

B. Upper Bounds on Lifetime

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

Proof: Consider a network with $N = \pi\lambda R^2$ sensors uniformly and independently deployed in the circular area with radius R. Each sensor will have equal probability of $\frac{Communication\, area\, of\, the\, sink}{Total\, network\, area} = \frac{\pi}{\pi R^2} = \frac{1}{R^2}$ to be deployed within the transmission range of the sink. The number of one hop neighbors of the sink, denoted as $|P_1|$, will be a Binomial random variable with parameters N and $\frac{1}{R^2}$ [23]. By the Law of Large Numbers [23], for any $\epsilon > 0$ the probability that $|P_1| \in [(1-\epsilon)\frac{N}{R^2}, (1+\epsilon)\frac{N}{R^2}]$ tends to 1 as $\lambda \to \infty$. In other words, the number of sensors in P_1 is nearly $\frac{N}{R^2} = \pi\lambda = \rho$ for high density networks. Then the total initial energy stored in P_1 is nearly ρE. Suppose the lifetime of the network is $\widetilde{T} > \frac{E}{R^2 e}$. As we assumed, each sensor generates one packet per time unit. Thus, there are $N = \rho R^2$ packets generated by sensors in the network and delivered to the sink in each time unit. Since the sink can only receive data from the nodes in P_1, packets generated by nodes in P_1 and in $\overline{P_1}$ must pass through nodes in P_1 at least once. Thus, the total number of packets passing through nodes in P_1 per time unit satisfies:

$$\sum_{n_1 \in P_1} f_{n_1} \geq N \tag{2}$$

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

$$D = \widetilde{T} \times \sum_{n_1 \in P_1} f_{n_1} \geq \widetilde{T} \times N > \frac{\rho E}{e} \tag{3}$$

The total energy used by nodes in P_1 is $D \times e > \rho E$. 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 $\frac{E}{R^2 e}$. ■

Theorem 1 can be easily extended to a network with arbitrary topology. For a network with area A, the lifetime is bounded by $\frac{\pi E}{Ae}$. Note that the result in Theorem 1 is the best possible lifetime over all routing algorithms. It shows that for a static network with only one sink, the network lifetime decreases as the network size increases. However, if we introduce one mobile relay into the network, the bottleneck will no longer be the neighbors of the sink, since the mobile relay can act as a "bridge" connecting nodes in P_2 to the sink in this case.

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

Proof: The amount of traffic passing through nodes in Q_i is at least the total 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 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

$$T^1 \quad \leq \quad \frac{\rho i^2 E}{(N - \rho i^2) \times (i-1)e} = \frac{i^2 E}{(R^2 - i^2)(i-1)e} \tag{4}$$

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

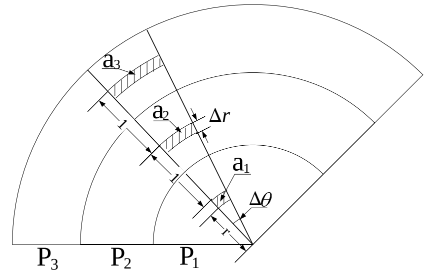


Fig. 3. The mapping areas of a_1, a_2 and a_3

C. Joint Mobility and Routing Algorithm

We now 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 can be 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 at each point on this path for a certain duration and relays traffic to the sink.

More specifically, when the mobile is at position M, 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 $4\frac{E}{R^2 e} - \frac{16E}{R^4 e}$ with one mobile node, when the network radius $R > 16\pi + 4$. [1]

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 Fig. 3, 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 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 Fig. 4 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

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

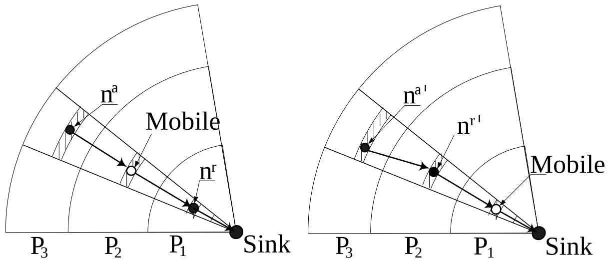


Fig. 4. Aggregation node selection and the relay path construction

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:

$$x_1 + x_2 = \frac{\rho}{\pi}(2r+1)\Delta r\Delta\theta \leq \frac{\rho}{\pi}(r+2)\Delta r\Delta\theta = x_3 \quad (5)$$

which means that the number of nodes in a_3 is always bigger than the sum of that in areas 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 $n^r_{\leq 2}$ in Q_2, we can associate a unique node, $n^a_3 = f(n^r_{\leq 2}) \in P_3$ such that n^a_3 can communicate to the sink by relaying via the mobile node and $n^r_{\leq 2}$. We call the node $n^r_{\leq 2}$ the relay node and the node n^a_3 the aggregation node (in later 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 once. If, depending on the position of the mobile, all traffic in $\overline{Q_2}$ 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 $\overline{Q_2}$ 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 $\overline{Q_2}$, 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 be at least:

$$4\rho \times (\frac{E - E'}{Ne}) = 4\frac{E}{R^2 e} - \frac{16E}{R^4 e} \quad (6)$$

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 its energy prior to $4\frac{E}{R^2e} - \frac{16E}{R^4e}$.

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, θ = 2π;
while (r_m < 2)
    if θ < 2π
        Set θ = θ + Δθ;
        Move to the new position (r_m, θ);
    else
        Set r_m = r_m + Δr, θ = 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, θ)
            Pick up one unselected node n whose coordinate is
              (r_m + 1, θ) and n' with coordinate (r_m + 2, θ);
            Set n^r = n, n^a = n';
            Broadcast the information about n^r and n^a to the nodes
              in the aggregation area;
            Relay packets from n^r to the sink for E/(ρR^2 e) − 4 E/(ρR^4 e)
              time units;
            Mark n and n' as selected;
        endwhile
    else
        while There exists an unselected node n whose coordinate is
              (r_m − 1, θ)
            Pick up one unselected node n whose coordinate is
              (r_m − 1, θ) and n' with coordinate (r_m + 1, θ);
            Set n^r = n, n^a = n';
            Broadcast the information about n^r and n^a to the nodes
              in the aggregation area;
            Relay packets from n^a to n^r for E/(ρR^2 e) − 4 E/(ρR^4 e) time units;
            Mark n and n' as selected;
        endwhile
    endif
endwhile
```

Fig. 5. Mobility algorithm for the mobile relay

To get the lifetime of $4\frac{E}{R^2e} - \frac{16E}{R^4e}$, we need to aggregate the traffic generated outside Q_2 to the node n^a. Fig. 5 shows the mobility algorithm for the mobile relay. Note that the mobile relay remains within Q_2. Fig. 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 $\overline{Q_3}$:

a) The nodes in $\overline{Q_3}$ 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 senses 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

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 ∈ 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 \ 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 ∈ G
     Find a neighbor in P_3 \ G and send sensed data to it;
  else
     Relay the received packet towards n^a using a neighbor in
     P_3 \ 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
```

Fig. 6. The Aggregation Routing Algorithm

the circle of radius l as shown in Fig. 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 $i-1+r$, $4 \le i < k$ from the sink, until they reach the aggregation point.

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 Figs. 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 overall 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 focus on traffic load for 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 sequel.

We now investigate the energy consumption of all nodes in the network under this joint mobility and routing scheme, and

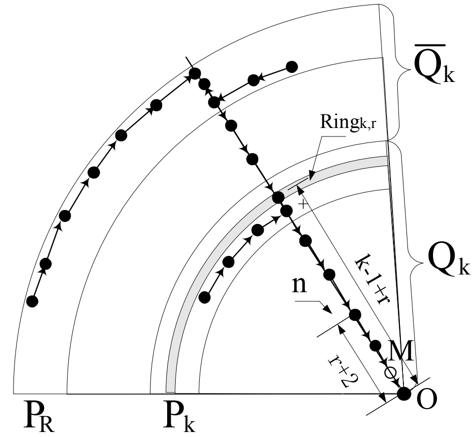


Fig. 7. Packet aggregating routes with ARA

show that the network lifetime will be at least $4\frac{E}{R^2 e} - \frac{16E}{R^4 e}$.

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 $\frac{4E}{R^2} - \frac{16E}{R^4 e}$. Similarly, for nodes in P_2, since they transmit the 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 $\frac{4E}{R^2} - \frac{16E}{R^4 e}$ 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 $Ring_{3,r}$, the 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 $0 < r \leq 1$, 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 their lifetime will be larger than $\frac{4E}{R^2 e}$.

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 $\overline{Q_k}$. 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 $\frac{4(E-E')}{R^2 e}$, the total energy used in delivering this part of traffic will be upper bounded by:

$$\begin{aligned} E_1(k,r) &\leq 2\rho(k-1+r)\Delta r \times \frac{4(E-E')}{R^2 e} \times \\ &\qquad e(\pi(k-1+r)+1) \\ &= \frac{8\rho(k-1+r)(\pi(k-1+r)+1)(E-E')\Delta r}{R^2} \end{aligned} \quad (7)$$

The next part is the packet generated by nodes in $\overline{Q_{k-1}}$. 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 $\frac{E-E'}{\rho R^2 e}$ time units. Then, the nodes in $Ring_{k,r}$ will need to route traffic from $\overline{Q_{k-1}}$ for at most $2(2r+1)\frac{E-E'}{R^2 e}\Delta r$ 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 $\overline{Q_{k-1}}$ passing through $Ring_{k,r}$ per time unit. So the total energy consumption during the lifetime will be upper bounded by:

$$\begin{aligned} E_2(k,r) &\leq 2(2r+1)\frac{E-E'}{R^2 e}\Delta r \times N \times e \\ &= 2\rho(2r+1)(E-E')\Delta r \end{aligned} \quad (8)$$

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:

$$\begin{aligned} E_{re}(k,r) &\geq 2\rho(k-1+r)(E-E')\Delta r - [E_1(k,r) + E_2(k,r)] \\ &\geq 2\rho(k-3)(E-E')\Delta r - \frac{8\rho(\pi k+1)(E-E')\Delta r}{R} \\ &\geq \frac{2\rho(R-16\pi-4)(E-E')\Delta r}{R} \end{aligned} \quad (9)$$

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 $4\frac{E}{R^2 e} - \frac{16E}{R^4 e}$. The bound holds for all k and r, so no node in $\bigcup_{k\geq 4} P_k$ will die prematurely when $R > 16\pi + 4$. ∎

Theorem 3 shows that we can construct a routing algorithm which can achieve the lifetime of $4\frac{E}{R^2 e} - \frac{16E}{R^4 e}$ 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.

D. Routing With Few Nodes Aware of Mobile Location

In ARA, every node in the network needs to know the position of the mobile and appropriately route traffic. This implies large overheads in disseminating information on the location of the mobile relay to all the nodes in the network. We now construct a routing algorithm in which only a limited number of nodes in the network need to know the location of the mobile relay. We show that with this routing algorithm we can still achieve a lifetime bound of $4\frac{E}{R^2 e} - \frac{16E}{R^4 e}$. We call this routing algorithm ARALN (Aggregation Routing Algorithm with Limited Nodes).

ARALN is described in detail in Fig. 8. We outline the ideas of the algorithm below.

1) Nodes which are outside a 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.

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 Fig. 9.

```
ARALN running on a static node n ∈ P_k
Parameters:
    n^a: the current aggregation node
    n^r: the current static relay node
    r: the distance between n^a and the sink is r + 2
    OM: the straight line connecting the sink and the mobile
Method – ARALN:
switch (k:the index of P_k where n ∈ 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;
    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;
  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;
  case s + 1, ..., R:
    Find a neighbor who is closest to the sink, send the packet to it;
```

Fig. 8. The Aggregation Routing Algorithm with Limited Nodes

3) Packets generated by nodes in Q_s are routed as in ARA described in Fig. 6.

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

Detailed proof of Theorem 4 can be found in [24].

E. Network with m Mobile Relays

We next 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 uniformly distributed 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:

$$T^m \leq \frac{\rho i^2 E}{(N - \rho i^2) \times (i - m)e} = \frac{i^2 E}{(R^2 - i^2)(i - m)e} \quad (10)$$

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

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 $4m\frac{E}{R^2 e} - \frac{32\pi m^3 E}{R^4 e}$ with m mobile nodes, when R is large enough.

The proof of Theorem 6 follows the proof of Theorem 4. However, a different mobility model is used with m mobile relays. The m mobile relays need to stay in a line and form a path to the sink with m sensors in Q_{2m} to maximize the network lifetime. There are $(2k-1)\rho$ sensors in P_k and $4m^2\rho$ sensors in Q_{2m}. Each path will use m static sensors and m mobiles, so there will be $4m$ such paths. Also each path can survive close to $\frac{E}{R^2 e}$, so the total network lifetime will be close to $4m\frac{E}{R^2 e}$. The component $\frac{32\pi m^3 E}{R^4 e}$ accounts for the need to reserve energy to deliver the packets generated in Q_{2m}.

IV. Mobile Relay in Finite Randomly Deployed Networks

In this Section, we consider the mobile relay in a randomly deployed network of moderate size and relax the assumption of a dense and large network in Section III.

A. System Model

For a random network with moderate size, such as 100 sensors randomly deployed in a 5×5 hops area, the randomness of sensor distribution may generate topology defects, i.e., voids or low sensor density areas. Such topology defects prevent the construction of perfect symmetric routing as described in Section III. However, our experiments show that a mobile relay can still improve the network lifetime by more than two times for networks with such topology defects.

[2]Here 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.

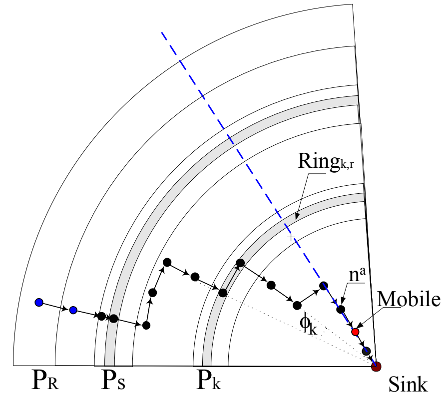


Fig. 9. Packet aggregating routes with ARALN

We first construct an optimization problem for the routing and mobility algorithm in a random finite network with only one mobile relay. The network topology is abstracted as a *Random Geometric Graph* with edge (i, j) between any pair of vertex i and j of distance smaller than 1 from each other. We maximize the overall network lifetime under the energy constraints of static sensors. Similar to the assumptions used in [8] for mobile sinks, we assume that the mobile relay will stay at positions where there is a static sensor. When the mobile relay is at the position of sensor k, it will take over the task of sensor k and sensor k will sleep for that time period. The mobile relay will shift between sensors and try to help as many sensors as possible during the network lifetime. The mobile will always stay at the position of some static sensor during the network lifetime, since this will always give a longer lifetime than removing the mobile node. Thus, maximizing the sum of periods for which the mobile stays at each location will give the optimal network lifetime. Here, we ignore the time used for moving the mobile nodes, which will be further discussed in Sec.V.

We formulate the linear optimization problem for maximizing the network lifetime as follows:

$$Maximize \quad \sum t_k \tag{11}$$

$$s.t. \quad \sum_j x_{ij}^k - \sum_j x_{ji}^k \quad = \quad s_i \times t_k \quad \forall i, k \tag{12}$$

$$x_{ij}^k \quad \geq \quad 0 \quad \forall (i,j), \forall k \tag{13}$$

$$\sum_{\forall k \neq i} \quad \sum_j x_{ij}^k \times e \quad \leq \quad E \quad \forall i \neq 0 \tag{14}$$

where t_k is the length of time that the mobile relay will stay at sensor k and x_{ij}^k is the total traffic flow from sensor i to sensor j during that time period, E and e represent the initial energy of the sensor nodes and the energy for relaying one packet, respectively, and s_i is the packet generation rate for node i. We assume that the data generation rate is uniform, i.e., $s_i = 1$. Then, the data generation rate for the sink, denoted as s_0, is set to $-N$. Constraints (12) and (13) give the network flow constraint for node i during the period t_k, which restricts the difference between the total out-flow and the total in-flow during the period to be equal to the packets node i generates in the period. Constraint (14) is the energy constraint, namely, the sum of energy over all the periods should not exceed the initial energy E. The energy used in period t_i is not counted, since during t_i the mobile is at the position of sensor i and sensor i is sleeping. On average there are $O(\rho N)$ edges in the static network and the mobile can appear in N different places. So, there are $O(\rho N^2)$ variables and $O(N^2)$ constraints in the linear programming problem, which is solvable in polynomial time.

B. Simulation Study

The experiment is based on the simplified energy model stated in Section III without considering the MAC or physical layer. The sensors are randomly deployed on fields with different size and shapes. For each network instance, we calculate the lifetime of the static network by the linear optimization algorithm described in [5], which gives the optimal lifetime for the static network. The lifetime of the mobile relay solution on the same network instance is calculated through the optimization problem as in (11)-(14). The lifetime improvement is averaged over 100 network instances.

1) Traffic Distribution

We first illustrate the traffic distribution for a randomly deployed network obtained by solving the linear programming problem. The network contains 100 randomly deployed sensors on 5×5 square area (equivalent to $\lambda = 4$) with one sink at the center (marked as a square). There are topology defects in this network instance as shown in Fig. 10(a), e.g., there is no sensor in the hole up-left to the sink. Fig. 10(a) describes the optimal traffic load of the static network, where the numbers on sensors indicate the number of packets they relay per unit time. Since the traffic is optimized by energy conserving routing, the neighbors of the sink share the same load of 13 packets per unit time. However, the load for these sensors is much higher than sensors in the periphery. Thus, these immediate neighbors of the sink will die first, then the sink will be disconnected from other parts of the network.

Figs. 10(b) and 10(c) show the traffic flow with one mobile relay in the network. The mobile relay stays at different places (represented by triangles) during different time periods. The traffic flow in each time period will be adjusted to maximize the overall network lifetime. Although we do not explicitly use the routing schemes developed in Section III, the routes obtained through lifetime optimization still follow the ideas used in Section III:

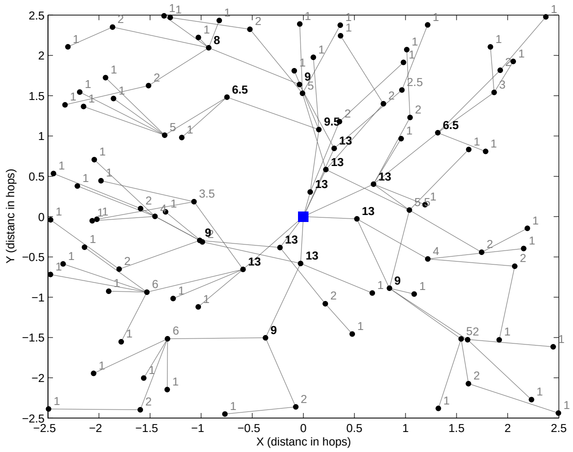
a) We do not explicitly oblige the mobile to stay around the sink in the optimization, however, the mobile relay still stays within two hops range of the sink. The experiments with and without such movement constraints provide similar network lifetime [24]. This shows the mobile relay actually only needs to stay in a two-hop distance around the sink to maximize the network lifetime.

b) Sensors will redirect their packets towards the mobile relay so that the mobile relay can carry more traffic. Also, the aggregation process will mostly use the sensors on the edge of the network. We can see from Figs. 10(b) and 10(c), almost 70% of the packets are flowing through the mobile relay.

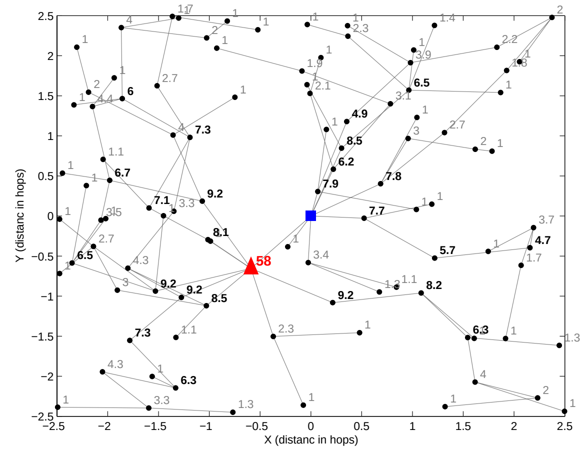
c) Sensors will work as the static relay node alternately. There exist some time periods that a particular sensor may be highly loaded, for example, some sensors need to carry 15 packets per unit time in Fig. 10(c). However, after they perform their relay task, they seldom need to relay for others during the rest of their lifetime. Such alternation actually reduces the overall traffic load for sensors. As shown in Fig. 10(d), most sensors share the average traffic load of 4.4 packets over the total network lifetime after we put in the mobile relay. This maximal traffic load is much smaller than the maximal load in the static network case, so the network can survive for a much longer time. Also, only a few sensors have traffic load smaller than 4.4, which means all sensors will use up their energy almost at the same time, thus there is little waste in residual energy.

2) Lifetime Improvements

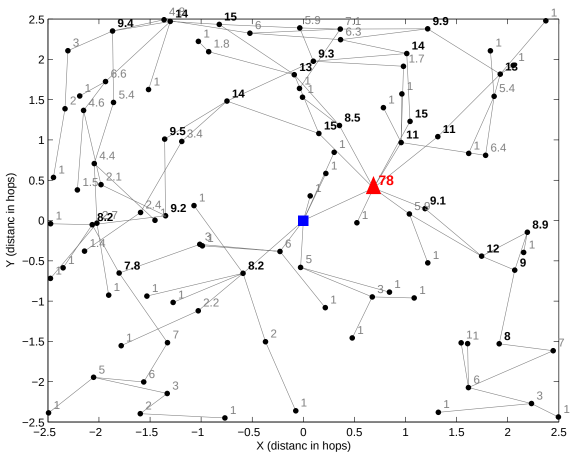
Fig. 11 gives the lifetime for static circular networks and the lifetime with one mobile relay for different network sizes. The



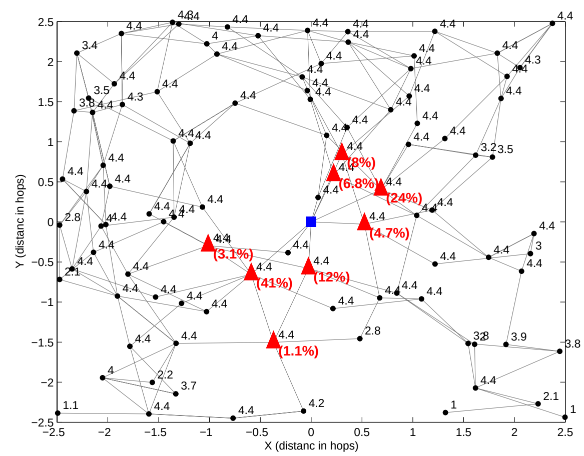
(a) Traffic distribution of static network



(b) Traffic distribution when mobile relay stays at position I (41% network lifetime)



(c) Traffic distribution when mobile relay stays at position II (24% network lifetime)



(d) The traffic flow averaged over the network lifetime with one mobile relay. Triangles are positions where the mobile relay will stay and the numbers below it are the length of stay

Fig. 10. Traffic flow in network with 100 sensors in 5×5 square area. The sink is marked as the square in the center of the network. Mobile relay is marked as triangle. In this example, we only show two positions and the corresponding traffic flow, there are actually 8 positions at which the mobile relay will stay during the network lifetime.

lifetime shown in the figure is normalized by the maximum node lifetime $\frac{E}{e}$. The experimental results of the static network fit the analysis in Theorem 1 quite well.

The lifetime improvement due to the mobile relay is more than two times for networks with more than 100 nodes, irrespective of the network shape, as shown in Fig. 12. The improvement ratio also increases with the network size and density. This agrees with the analysis in Section III. The density and size of the network are two major factors which limit the performance of the mobile relay in the random network. A low network density increases the chance of topology defects around the sink. Such defects will limit the number of routes that can be built across the two hop area around the sink. A limited network size means that only part of the traffic can be redirected and carried by the mobile relay, thus limiting its efficiency. As Fig. IV-B shows, all sensors in the network participate in the traffic aggregation, however, there is still not enough space for all packets to be aggregated to the mobile relay. As the network density and size grow, the effects of topology defects and limited space are mitigated so that the lifetime improvement is increased.

3) Route Dilation

Another important performance metric is the average path length, which affects many other metrics including the average delay. As our scheme requires the traffic to be redirected towards the mobile relay, packets will take a longer route compared to shortest path routing. However, the route dilation, defined as the path length of our scheme divided by the length of the shortest path, is bounded by $\pi+1$, since the packet will at most go around an angle of π and then directly go towards the sink. Fig. 13 compares the route dilation of the mobile relay solution with other schemes. The results show that energy conserving routing does not incur much route dilation and its average hop count is quite close to the shortest path. The mobile relay scheme uses almost twice as many hops to deliver the packet. In our experiments, route dilation converges to 2.6 as the network size grows. So the route dilation and the consequent longer packet delay is bounded for the mobile

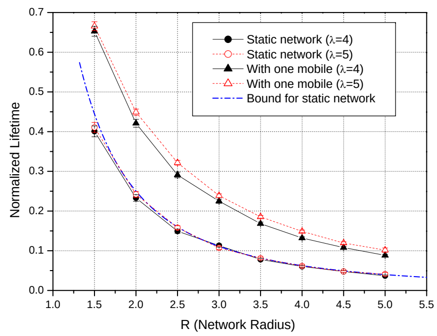


Fig. 11. Network lifetime for nodes randomly deployed on a circular region, for λ = 4 and 5 (confidence interval 95%)

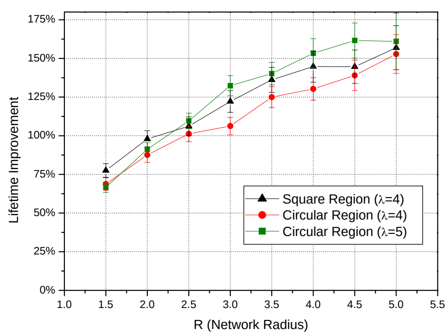


Fig. 12. Average lifetime improvement for networks with one mobile over the static network (confidence interval 95%). For the square network, the sink is put at the center of the square and the network radius R is defined as half the side length of the square.

relay scheme. Note that the delay of the mobile relay scheme is much smaller than "data mule" solutions where delay depends on the moving speed of the mobile [2].

V. Performance Comparison of Mobile Relay to Other Related Approaches

Fig. 14 compares the network lifetime for different approaches. In this figure, the lifetime of non-energy-aware minimal hop routing networks is normalized to 1. Energy conserving routing improves the network lifetime nearly 4 times over minimal hop routing. This scheme gives the upper bound for a homogenous static network. By adding only one mobile relay, we can double the lifetime of energy conserving routing. However, when we use static energy provisioning schemes, e.g., adding 4 times more energy to randomly selected 25% of the static sensors, the network lifetime can only be extended by 40%. The discussion below will show that using a mobile relay is better than most static energy provisioning approaches. Another method to extend lifetime is to use the mobile node as the sink. This gives a lifetime improvement that is even better than the mobile relay. However, the mobile sink approach has certain constraints as we discuss in the sequel.

A. Static Approach

1) Increasing the density of static nodes

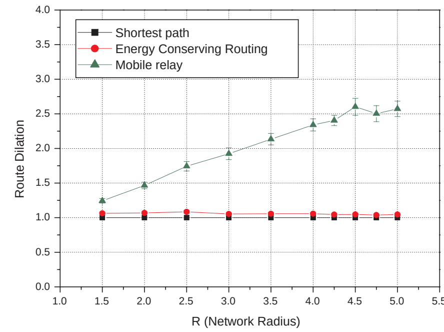


Fig. 13. Comparing the route dilation for different approaches. (Network on square area with $\lambda = 4$, confidence interval 95%)

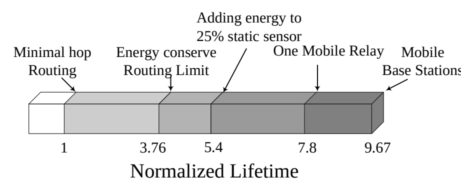


Fig. 14. Comparing network lifetime for different approaches

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. If the nearby sensors die, these additional sensors will 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 nodes within the two hop radius of the sink. It is easy to see that we need to deploy at least 4ρ additional static nodes around the sink to achieve the same performance as one mobile relay, where ρ is the number of sensors around the sink in the original network.

2) Increasing the energy carried by static nodes

We can also upgrade the battery of static sensors to extend their lifetime. Assume that we have some static sensors equipped with better batteries, which provide four times more energy than normal batteries. In the deployment phase, those more powerful static sensors are mixed up with normal sensors and randomly deployed in the field. Therefore, the network lifetime will be extended since some of the static sensors gain more initial energy. Fig. 15 compares the lifetime for such hybrid static network with networks having one mobile relay. The lifetime improvement of the mobile relay solution is close to the static network where 50% of the static sensors have four times more energy. However, in the mobile relay solution we only need to add one mobile relay in the network, which may be much cheaper than the static solutions where we must upgrade the batteries of 50% of the sensors to gain similar performance.

3) Multiple static sinks

Instead of mobile nodes, we can also add multiple static sinks to improve the network lifetime. By adding more static sinks, we can scale down the size of the sensor network. As

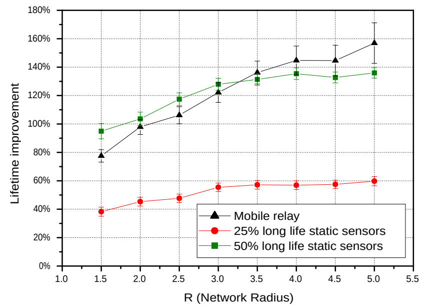


Fig. 15. Comparing the lifetime improvement of adding one mobile relay with solutions that add more energy to static sensors. (Network on square area with $\lambda = 4$, confidence interval 95%.)

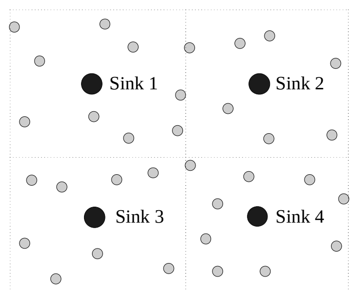


Fig. 16. Adding static sinks to improve the network lifetime

shown in Fig. 16, we can split the original large network into four smaller networks by adding three more static sinks. Sensors only need to send their data to the closest sink, thus the traffic is largely reduced. Sensors around the static sinks will still be the bottleneck. Since each small network only contains $1/4$ of the sensors of the original network, the traffic flow through the bottleneck sensors will be reduced to $1/4$. The newly added sinks still have ρ neighbors as in the original network, so the lifetime of the new network will be 4 times of the original due to the reduction in traffic. As we have shown in Section III, adding one mobile relay can achieve 4 times improvement in the ideal case, which is similar to adding 3 more static sinks.

By investigating different static approaches, we see that the mobile relay is better than most static approaches in the sense of lifetime improvement. However, static solutions also have certain advantages since their routing schemes can be very simple and the routing overhead is low.

B. Mobile Sink Approach

The mobile sink approach uses the mobile node as the sink. As the mobile node goes around the network, the bottleneck around the sink is distributed over the whole network. As shown in Fig. 14, the mobile sink approach in [9] can provide longer lifetime than a mobile relay. When the sink is moving around the network periphery, the mobile sink approach can give an improvement factor of $O(R)$ on network lifetime, while we need $O(R)$ mobile relays to achieve the same improvement [24].

However, there are certain tradeoffs between the mobile relay and mobile sink solution. First, the mobile relay solution is more flexible than the mobile sink. The sink is the gateway to the outside network, and certain applications may not permit the sink to be mobile. Also, mobile sink may create more routing overheads. For a realtime data logging application, the mobile sink should always be connected to the sensors. So, each single movement of the sink must be traced and the data gathering path should be changed accordingly. For the mobile relay approach, the network is functional without the mobile relay. Thus, the mobile relay can detach from the network, move, then reattach at a new position. Such flexibility in movement generates much less overhead than the mobile sink. The mobile relay also do not need to move very fast. When mobile relay is moving and detached from the network, the static nodes can still relay the packets and only part of static sensors will deplete their energy at a faster rate. Such faster energy depletion will not severely impair node lifetime, since the time spent for mobile to move around in two hop distance is ignorable compared to the network lifetime, which is generally from several months to years. Second, the mobile relay can heal the topology defects in the network. Due to the randomness in deployment, there may be topology defects in the network, e.g., the network is partitioned in the beginning or the communication to certain parts of the network relies on few sensors, as shown in Fig. 1. The mobile relay can improve the network lifetime greatly in this case, while the mobile sink cannot. Finally, the mobile relay requires only a small moving range, and only sensors in a limited area need to know where the mobile relay is. For the mobile sink approach, the lifetime improvement depends on the moving range of the sink. If the sink only moves in a small area, the lifetime improvement will be similar to mobile relay approach.

VI. Conclusion

We have investigated the possibility of using a heterogeneous network composed of many simple static nodes and a few mobile nodes. 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 in the ideal case. Another interesting property of this mobile relay approach is that we only need to change the routing algorithm for a relatively small area to use the mobile relay. Furthermore, the mobile relay need not travel all around the network. It never needs to venture farther than two hops from the sink. We also compared the mobile relay approach with other possible solutions. We see that mobility is actually a great advantage since the mobile relay is more efficient than most static energy provisioning methods. We also investigate other ways to use mobile nodes, such as mobile sink approach. Although it is clear from our analysis that using a mobile sink is always beneficial in terms of the lifetime of the network, there are certain tradeoffs to make the sink mobile.

In this paper, we make some simplifying assumptions, e.g., the network is running a data logging application and sensors are incapable of power control. However, in a network which is event based, using mobile relay may be even more beneficial.

Since the traffic is not uniformly distributed in such a network, we can move the mobile relay in the directions where traffic is high. In this case we may not need to redirect the traffic as in the data logging application, so that the overhead caused by mobile relay will be reduced. Our scheme can also work together with power control or data aggregation/compression methods. Although the traffic can be reduced by data compression, the bottleneck described in this paper still exists in such network since the information generated per unit area is still fixed, and our model of uniform packet generation rate can be applied. The network with power control can also fit in the optimization problem introduced in Section IV by setting the link cost according to the energy consumption over the particular link.

REFERENCES

[1] H. Liu, P. Wan, C. Yi, X. Jia, S. Makki, and P. Niki, "Maximal lifetime scheduling in sensor surveillance networks," in *Proc. IEEE INFOCOM*, Mar. 2005, pp. 2482-2491.

[2] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor networks," in *Proc. IEEE SNPA*, May 2003, pp. 30-41.

[3] R. Zheng, J. C. Hou, and L. Sha, "Asynchronous wakeup for ad hoc networks," in *Proc. ACM MobiHoc*, Jun 2003, pp. 35-45.

[4] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. IEEE INFOCOM*, Jun 2002, pp. 1567-1576.

[5] J. H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM* , Mar 2000, pp. 22-31.

[6] N. Sadagopan and B. Krishnamachari, "Maximizing data extraction in energy-limited sensor networks," in *Proc. IEEE INFOCOM*, Mar 2004, pp. 1717-1727.

[7] S. Gandham, M. Dawande, R. Prakash, and S. Venkatesan, "Energy-efficient schemes for wireless sensor networks with multiple mobile base stations," in *Proc. IEEE GLOBECOM*, Dec 2003, pp. 377-381.

[8] Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli, "Exploiting sink mobility for maximizing sensor networks lifetime," in *Proc. HICSS*, Jan 2005.

[9] J.Luo and J. P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proc. IEEE INFOCOM*, Mar 2005, pp. 1735-1746.

[10] A. Chakrabarti, A. Sabharwal, and B. Aazhang, "Using predictable observer mobility for power efficient design of sensor networks," in *Proc. IPSN*, Apr 2003, pp. 129-145.

[11] A. Kansal, A. Somasundara, D. Jea, M. Srivastava, and D. Estrin, "Intelligent fluid infrastructure for embedded networks," in *Proc. ACM MobiSys*, Jun 2004, pp. 111-124.

[12] W. Zhao, M. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *Proc. ACM MobiHoc*, May 2004, pp.187-198.

[13] I. Papadimitriou and L. Georgiadis, "Maximum lifetime routing to mobile sink in wireless sensor networks," in *Proc. IEEE SoftCOM*, 2005.

[14] J. Luo, J. Panchard, M. Piorkowski, M. Grossglauser, and J.-P. Hubaux, "Mobiroute: Routing towards a mobile sink for improving lifetime in sensor networks," in *Proc. DCOSS*, 2006, pp. 480-497.

[15] A. Shankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, Mar 2004, pp. 1089-1097.

[16] Y. T. Hou, Y. Shi, H. D. Sherali, and S. F. Midkiff, "Prolonging sensor network lifetime with energy provisioning and relay node placement," in *Proc. IEEE SECON*, Sep 2005, pp. 295-304.

[17] J. Chou, D. Petrovic, and K. Ramchandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *Proc. IEEE INFOCOM*, Mar 2003, pp. 1054-1062.

[18] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. HICSS*, Jan 2000.

[19] O. Younis and S. Fahmy, "Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach," in *Proc. IEEE INFOCOM*, Mar 2004, pp. 629-640.

[20] N. Li, J. Hou, and J. Sha, "Design and analysis of an MST based topology control algorithm," in *Proc. IEEE INFOCOM*, Mar 2003, pp. 1702-1712.

[21] J. Pan, Y. Hou, L. Cai, Y. Shi, and S. Shen, "Topology control for wireless sensor networks," in *Proc. ACM MobiCom*, Sep 2003, pp. 286-299.

[22] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. ACM MobiCom*, 1998, pp. 181-190.

[23] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. 4th Ed. McGraw Hill, 2002.

[24] W. Wang, V. Srinivasan, and K. C. Chua, "Using mobile relays to prolong the lifetime of wireless sensor networks," in *Proc. ACM MobiCom*, Aug 2005, pp. 270-283.

Wei Wang (S'07) obtained his B.Eng and M.Sc in Electronics Science and Engineering from Nanjing University, China, in 1997 and 2000 respectively. He is currently a Ph.D student in National University of Singapore. His research interests are in the area of wireless sensor networks.

Vikram Srinivasan (M'98) obtained his Bachelor of Science in Physics from the University of Chennai in 1994 and an M.E in Electrical Communications Engineering from the Indian Institute of Science, Bangalore in 1998. He obtained his PhD in Electrical and Computer Engineering from the University of California at San Diego in 2003. He is currently an Assistant Professor in the Electrical and Computer Engineering Department at the National University of Singapore. His research interests are broadly in the area of wireless networks.

Kee-Chaing Chua (M'87) received a PhD degree in Electrical Engineering from the University of Auckland, New Zealand in 1990. He joined the National University of Singapore (NUS) as a Lecturer in 1990 and is now a Professor in the Department of Electrical & Computer Engineering. He served as the Faculty of Engineering's Vice Dean for Research from June 2003 to March 2006. From 1995 to 2000, he was seconded to the Center for Wireless Communications (now part of the Institute for Infocomm Research), a national telecommunication R&D centre funded by the Singapore Agency for Science, Technology and Research as its Deputy Director. From 2001 to 2003, he was on leave of absence from NUS to work at Siemens Singapore where he was the founding head of the Mobile Core R&D Department funded by Siemens' ICM Group. Since March 2006, he has been seconded to the National Research Foundation as a Director.

Dr. Chua has carried out research in various areas of communication networks and has published more than 180 papers in these areas in international refereed journals and conferences. His current research interests are in wireless networks (in particular wireless sensor networks) and optical burst switched networks. He has also been an active member of the Institute of Electrical & Electronics Engineers (IEEE), Inc., and is a recipient of an IEEE 3rd Millennium medal.