

# Mobile Relay Configuration in Data-Intensive Wireless Sensor Networks

Mr. Anil S Band  
CSE (M.Tech 1<sup>ST</sup>)  
JIT Nagpur

MS. Suwarna Hajare  
CSE(M.TECH)  
JIT Nagpur

**Abstract**— Wireless Sensor Networks (WSNs) are increasingly used in data-intensive applications such as microclimate monitoring, precision agriculture, and audio/video surveillance. A key challenge faced by data-intensive WSNs is to transmit all the data generated within an application's lifetime to the base station despite the fact that sensor nodes have limited power supplies. We propose using low-cost disposable mobile relays to reduce the energy consumption of data-intensive WSNs. Our approach differs from previous work in two main aspects. First, it does not require complex motion planning of mobile nodes, so it can be implemented on a number of low-cost mobile sensor platforms. Second, we integrate the energy consumption due to both mobility and wireless transmissions into a holistic optimization framework. Our framework consists of three main algorithms. The first algorithm computes an optimal routing tree assuming no nodes can move. The second algorithm improves the topology of the routing tree by greedily adding new nodes exploiting mobility of the newly added nodes. The third algorithm improves the routing tree by relocating its nodes without changing its topology. This iterative algorithm converges on the optimal position for each node given the constraint that the routing tree topology does not change. We present efficient distributed implementations for each algorithm that require only limited, localized synchronization. Because we do not necessarily compute an optimal topology, our final routing tree is not necessarily optimal. However, our simulation results show that our algorithms significantly outperform the best existing solutions.

## I. INTRODUCTION

SNS have been deployed in a variety of data-intensive applications including microclimate and habitat [1], precision agriculture, and audio/video surveillance [2]. A moderate-size WSN can gather up to 1 Gb/year from a biological habitat [3]. Due to the limited storage capacity of sensor nodes, most data must be transmitted to the base station for archiving and analysis. However, sensor nodes must operate on limited power supplies such as batteries or small solar panels. Therefore, a key challenge faced by data-intensive WSNs is to minimize the energy consumption of sensor nodes so that all the data generated within the lifetime of the application can be

Several different approaches have been proposed to significantly reduce the energy cost of WSNs by using the mobility of nodes. A robotic unit may move around the 25 minutes in full motion. network and

collect data from static nodes through one-hop or multihop transmissions [4], [5], [6], [7], [8]. The mobile node may serve as the base station or a "data mule" that transports data between static nodes and the base station [9], [10], [11]. Mobile nodes may also be used as relays [12] that forward data from source nodes to the base station. Several movement strategies for mobile relays have been

Although the effectiveness of mobility in energy conservation is demonstrated by previous studies, the

. The authors are with the Department of Computer Science, Michigan State

following key issues have not been collectively addressed. First, the movement cost of mobile nodes is not accounted for in the total network energy consumption. Instead, mobile nodes are often assumed to have replenishable energy supplies [7] which is not always feasible due to the constraints of the physical environment. Second, complex motion planning of mobile nodes is often assumed in existing solutions which introduces significant design complexity and manufacturing costs. In [7], [8], [14], [15], mobile nodes need to repeatedly compute optimal motion paths and change their location, their orientation and/or speed of movement. Such capabilities are usually not supported by existing low-cost mobile sensor platforms. For instance, Robomote [16] nodes are designed using 8-bit CPUs and small batteries that only last for about

In this paper, we use low-cost disposable mobile relays to reduce the total energy consumption of data-intensive WSNs. Different from mobile base station or data mules, mobile relays do not transport data; instead, they move to different locations and then remain stationary to forward data along the paths from the sources to the base station. Thus, the communication delays can be significantly reduced compared with using mobile sinks or data mules. Moreover, each mobile node performs a single relocation unlike other approaches which require repeated relocations.

Our approach is motivated by the current state of mobile sensor platform technology. On the one hand, numerous low-cost mobile sensor prototypes such as Robomote [16], [18] are now available. Their manufacturing cost is comparable to that of typical static sensor platforms. As a result, they can be massively deployed in a network and used in a disposable manner. Our approach takes advantage of this capability by

assuming that we have a large number of mobile relay nodes. On the other hand, due to low manufacturing cost, existing mobile sensor platforms are typically powered by batteries and only capable of limited mobility. Consistent with this constraint, our approach only requires one-shot relocation to designated positions after deployment. Compared with our approach, existing mobility approaches typically assume a small number of powerful mobile nodes, which does not exploit the availability of many low-cost mobile nodes.

We make the following contributions in this paper:

1. The rest of the paper is organized as follows: Section 2 reviews related work. In Section 3, we formally define the problem of optimal mobile relay configuration. In Section 4, we present our centralized optimization framework, an optimal solution for the base case with a single mobile relay, an optimal algorithm for constructing a routing tree given no mobility of nodes, and a greedy algorithm for improving the routing tree by adding new nodes. In Section 5, we present an optimal relocation algorithm given a fixed routing topology. In Section 6, we discuss the efficiency and optimality of our framework. In Section 7, we propose efficient distributed implementations for our algorithms. Section 8 describes our simulation results and Section 9 concludes this paper.

We formulate the problem of Optimal Mobile Relay Configuration (OMRC) in data-intensive WSNs. Our objective of energy conservation is holistic in that the total energy consumed by both mobility of relays and wireless transmissions is minimized, which is in contrast to existing mobility approaches that only minimize the transmission energy consumption. The tradeoff in energy consumption between mobility and transmission is exploited by configuring the positions of mobile relays. We study the effect of the initial configuration on the final result. We compare different initial tree building strategies and propose an optimal tree construction strategy for static nodes with no mobility. We develop two algorithms that iteratively refine the configuration of mobile relays. The first improves the tree topology by adding new nodes. It is not guaranteed to find the optimal topology. The second improves the routing tree by relocating nodes without changing the tree topology. It converges to the optimal node positions for the given topology. Our algorithms have efficient distributed implementations that require only limited, localized synchronization. We conduct extensive simulations based on realistic energy models obtained from existing mobile and static sensor platforms. Our results show that our algorithms can reduce energy consumption by up to 45 percent compared to the best existing solutions.

base station moves around the network and collects data from the nodes. In some work, all nodes are always performing multiple hop transmissions to the base station, and the goal is to rotate which nodes are close to the base station in order to balance the transmission load [4], [5], [6]. In other work, nodes only transmit to the base station

when it is close to them (or a neighbor). The goal is to compute a mobility path to collect data from visited nodes before those nodes suffer buffer overflows [7], [8], [14], [15]. In [8], [19], [20], several rendezvous-based data collection algorithms are proposed, where the mobile base station only visits a selected set of nodes referred to as rendezvous points within a deadline and the rendezvous points buffer the data from sources. These approaches incur high latencies due to the low to moderate speed, e.g., 0.1-1 m/s [14], [16], of mobile base stations. Data mules are similar to the second form of mobile base stations [9], [10], [11]. They pick up data from the sensors and transport it to the sink. In [21], the data mule visits all the sources to collect data, transports data over some distance, and then transmits it to the static base station through the network. The goal is to find a movement path that minimizes both communication and mobility energy consumption. Similar to mobile base stations, data mules introduce large delays since sensors have to wait for a mule to pass by before starting their transmission. In the third approach, the network consists of mobile relay nodes along with static base station and data sources. Relay nodes do not transport data; instead, they move to different locations to decrease the transmission costs. We use the mobile relay approach in this work. Goldenberg et al. [13] showed that an iterative mobility algorithm where each relay node moves to the midpoint of its neighbors converges on the optimal solution for a single routing path. However, they do not account for the cost of moving the relay nodes. In [22], mobile nodes decide to move only when moving is beneficial, but the only position considered is the midpoint of neighbors. Unlike mobile base stations and data mules, our OMRC problem considers the energy consumption of both mobility and transmission. Our approach also relocates each mobile relay only once immediately after deployment. Unlike previous mobile relay schemes [13] and [22], we consider all possible locations as possible target locations for a mobile node instead of just the midpoint of its neighbors. Mobility has been extensively studied in sensor network and robotics applications which consider only mobility costs but not communication costs. For example, in [23], the authors propose approximation algorithms to minimize maximum and total movement of the mobile nodes such that the network becomes connected. In [24], the authors propose an optimal algorithm to bridge the gap between two static nodes by moving nearby mobile nodes along the line connecting the static points while also minimizing the total/maximum distance moved. In [25], [26], the authors

## RELATED WORK

We review three different approaches, mobile base stations, data mules, and mobile relays, that use mobility to reduce energy consumption in wireless sensor networks. A mobile propose algorithms to find motion paths for robots to explore the area and perform a certain task while taking into consideration the energy available at each robot. These problems ignore communication costs which add an

Fig. 1. Reduction in energy consumption due to mobile relay. As the data chunk size increases, the optimal position converges to the midpoint of  $s_1 s_3$ .

increased complexity to OMRC, and consequently their results are not applicable.

Our OMRC problem is somewhat similar to a number of graph theory problems such as the Steiner tree problem [27], [28], [29] and the facility location problem [30], [31]. However, because the OMRC cost function is fundamentally different from the cost function for these other received signal strength of problems, existing solutions to these problems cannot be applied directly and do not provide good solutions to a  $\frac{1}{4}$  OMRC. For example, there is no obvious way to include

#### PROBLEM DEFINITION Energy Consumption Models

$10^{-7} J = \text{bit}$  and  $b = \frac{1}{4} \times 2$

Nodes consume energy during communication, computation, and movement, but communication and mobility parameter  $k$  is roughly  $10^{10}$  energy consumption are the major cause of battery drainage. Radios consume considerable energy even in an idle listening state, but the idle listening time of radios can be significantly reduced by a number of sleep scheduling protocols [32]. In this work, we focus on reducing the total energy consumption due to transmissions and mobility. Such a holistic objective of energy conservation is motivated 3.2 by the fact that mobile relays act the same as static

For mobility, we consider wheeled sensor nodes with differential drives such as Khepera [17], Robomote [16], and FIRA [18]. This type of node usually has two wheels, each controlled by independent engines. We adopt the distance proportional energy consumption model which is appro

[33]. The energy  $E_{\text{mob}} = \frac{1}{2} k d$ : The value of the parameter  $k$  depends on the speed of the node. In general, there is an optimal speed at which  $k$  is  $x_1 x_3$ , which is suggested in lowest. In [33], the authors discuss in detail the variation of the energy consumption with respect to the speed of the mote. When the node is running at optimal speed,  $k = \frac{1}{4} \times 2$  [33].

To model the energy consumed through transmissions, models:  $k = \frac{1}{4} \times 2$ ;  $a = \frac{1}{4} \times 0.6$  we analyze the empirical results obtained by two radios CC2420 [34] and CC1000 [35] that are widely used on solution is to move  $s_2$  to  $x_i$

where  $m$  is the number of bits transmitted and  $a$  and  $b$  are constants depending on the environment. We now discuss the instantiation of the above model for both CC2420 and CC1000 radio platforms. In an outdoor environment, for

80 dbm (which corresponds to a packet reception ratio higher than 95 percent), we obtain

$10^{-7} J = \text{bit}$  and  $b = \frac{1}{4} \times 4 \times 10^{-10} J = \text{bit}$  from the measurements on CC2420 in [36]. This model is

consistent with the theoretical analysis discussed in [37]. We also consider the energy needed by CC1000 to output the same levels. We get lower consumption parameters:  $a = \frac{1}{4} \times 0.3$

$10^{-10} J = \text{bit}$ . We will see in Section 5 that we maintain this high packet reception ratio throughout our algorithm. We note that although the

times larger than the transmission parameter  $b$ , the relays move only once whereas large amounts of data are transmitted. For large enough data chunk sizes, the savings in energy transmission costs compensates for the energy expended to move the nodes resulting in a decrease in total energy consumed.

#### An Illustrative Example

We now describe the main idea of our approach using a simple example. Suppose we have three nodes  $s_1$ ;  $s_2$ ;  $s_3$  located at positions  $x_1$ ;  $x_2$ ;  $x_3$ , respectively (Fig. 1), such that  $s_2$  is a mobile relay node. The objective is to minimize the total energy consumption due to both movement and transmissions. Data storage node  $s_1$  needs to transmit a data chunk to sink  $s_3$  through relay node  $s_2$ . One solution is to have  $s_1$  transmit the data from  $x_1$  to node  $s_2$  at position  $x_2$  and node  $s_2$  relays it to sink  $s_3$  at position  $x_3$ ; that is, node  $s_2$  does not move. Another solution, which takes advantage of  $s_2$ 's mobility, is to move  $s_2$  to the midpoint of the segment

[13]. This will reduce the transmission energy by reducing the distances separating the nodes. However, moving relay node  $s_2$  also consumes energy. We assume the following parameters for the energy  $10^{-7} J$ ;  $b = \frac{1}{4} \times 4 \times 10^{-10} J$ .

In this example, for a given data chunk  $m_i$ , the optimal (a position that we can compute existing sensor network platforms. For CC2420, the authors of [36] studied the transmission power level needed for transmitting packets reliably (e.g., over 95 percent packet reception ratio) over different distances. Let  $E_{\text{tx}}(d)$  be the energy consumed to transmit reliably over distance  $d$ . It can be modeled as

$E_{\text{tx}}(d) = \frac{1}{4} m d a + b d^2$  (precisely). This will minimize the total energy consumption due to both transmission and mobility. For small messages,  $s_2$  moves very little if at all. As the size of the data increases, relay node  $s_2$  moves closer to the midpoint. In this example, it is beneficial to move when the message size exceeds 4 MB. We illustrate in Table 1 the energy savings achieved using our optimal approach and the other two approaches for the relevant range of data sizes. For large enough data

IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 12, NO. 2, FEBRUARY 2013 13 MB), one relay node can reduce total energy consumption by 10 percent compared to the other two approaches. As the data chunk size increases further, the energy savings decrease, and the optimal position converges to the midpoint when the data size exceeds 43 MB. In general, the reduction in energy consumption is higher

The above example illustrates two interesting results. The optimal position of a mobile relay is not the midpoint between the source and sink when both mobility and transmissions costs are taken into consideration. This is in contrast to the conclusion of several previous studies [12], [13] which only account for transmission costs. Second, the optimal position of a mobile relay depends on not only the network topology (e.g., the initial positions of nodes) but also the amount of data to be transmitted. Moreover, as the data chunk size increases, the optimal position converges to the midpoint of  $s_1$  and  $s_3$ . These results are particularly important for minimizing the energy cost of data-intensive WSNs as the traffic load of such networks varies significantly with the sampling rates of nodes and

#### Problem Formulation

In our definitions, we assume that all movements are completed before any transmissions begin. We also assume there are no obstacles that affect mobility or transmissions. In this case, as we show in Section 4.2, the distance moved by a mobile relay is no more than the distance between its starting position and its corresponding position in the evenly spaced configuration which These constraints affect the optimal configuration. often leads to a short delay in mobile relay relocation. Furthermore, we assume that all mobile nodes know their locations either by GPS units mounted on them or a localization service in the network. We focus on the case where all nodes are in a 2D plane  $\mathbb{R}^2$ , but the results apply

Our problem can be described as follows: Given a network containing one or more static source nodes that store data gathered by other nodes, a number of mobile relay nodes and a static sink, we want to find a directed routing tree from the sources to the sink as well as the optimal positions of the mobile nodes in the tree in order to minimize the total energy consumed by transmitting data from the source(s) to the sink and the energy consumed by relocating the mobile relays. The source nodes in our problem formulation serve as storage points which cache the data gathered by other nodes and periodically transmit to the sink, in response to user queries. Such a network architecture is consistent with the

[38]. Our problem formulation also considers the initial positions of nodes and the amount of data that needs to be transmitted from each storage node to the sink. The formal definition

Definition 1 (Optimal mobile relay configuration). Input

Instance:  $S$ , a list of  $n$  nodes  $s_1; \dots; s_n$  in the network;  $O$ , a list of  $n$  locations  $o_1; \dots; o_n$ , where  $o_i$  is the initial position of node  $s_i$  for  $1 \leq i \leq n$ ;  $S_{\text{sources}}$ , a subset of  $S$  representing the source nodes;  $r$ , a node in  $S$ , representing the single sink;  $M_{\text{sources}} = \{f_m | s_m \in S_{\text{sources}}\}$ , a set of data chunk sizes for all sources in  $S_{\text{sources}}$ .

We define  $m_i$ , which we compute later, to be the weight of node  $s_i$  which is equal to the total number of bits to be transmitted by node  $s_i$ . We define a configuration  $hE$ ;  $U_i$  as a pair of two sets:  $E$ , a set of directed arcs  $\delta s_i; s_j$  that represent the directed tree in which all sources are leaves and the sink is the root and  $U$ , a list of locations  $u_1; \dots; u_n$  where  $u_i$  is the transmission position for node  $s_i$  for  $1 \leq i \leq n$ . The cost of a configuration  $hE$ ;  $U_i$  is given by:

$$X = c(hE; U_i) = \sum_{i=1}^n m_i \cdot d(s_i, u_i) + \sum_{(s_i, s_j) \in E} d(s_i, s_j)$$

Output:  $hE$ ;  $U_i$ , an optimal configuration that minimizes the cost  $c(hE; U_i)$ .

#### CENTRALIZED SOLUTION Energy Optimization Framework

The Optimal Mobile Relay Configuration problem is challenging because of the dependence of the solution on multiple factors such as the routing tree topology and the amount of data transferred through each link. For example, when transferring little data, the optimal configuration is to use only some relay nodes at their original positions. As the amount of data transferred increases, three changes occur: the topology may change by adding new relay nodes, the topology may change by changing which edges are used, and the relay nodes may move closer together. In many cases, we may have restrictions such as no mobility for certain relay nodes or we must use a fixed routing tree.

We illustrate how the optimal configuration depends on the amount of data to transfer using the example from Fig. 2a. When there is very little data to transfer, the optimal routing tree  $T_a$  depicted in Fig. 2a uses only some of the relay nodes in their original positions. When the amount of data to transfer from  $s_1$  and  $s_2$  increases to 15 MB, the relay nodes in tree  $T_a$  move to their corresponding positions in tree  $T_b$  of Fig. 2b but the topology does not change. When the amount of data to transfer from  $s_1$  and  $s_2$  is between 25 and 60 MB, the optimal routing tree has a different topology as shown in Fig. 2c. For even larger messages, new trees with even more nodes included are optimal. For example, when the amount of data to be transferred is between 100 and 150 MB, the optimal tree is depicted in Fig. 2d. No existing tree construction strategy handles all these cases. For example, the minimum spanning tree that includes all network nodes has two fundamental problems. It will typically include unneeded nodes, and it typically creates nonoptimal topologies as it focuses only on the current location of nodes as opposed

to where nodes may move to.

We now present a centralized approach to solve OMRC that breaks the problem into three distinct steps: initial tree construction, node insertions, and tree optimization. For each step, we present an algorithm to solve the corresponding subproblem. Our algorithm for initial tree construction is optimal for the static environment where nodes cannot move. However, we can effectively apply the later algorithms if we must start with a different topology. Our greedy heuristic for improving the routing tree topology by adding nodes exploits the mobility of the newly added