Mobile data collectors in wireless sensor networks

Waleed Al-Salih

A thesis submitted to the School of Computing
in conformity with the requirements
for the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
April, 2009

Your file Votre référence
ISBN: 978-0-494-65438-5
Our file Notre référence
ISBN: 978-0-494-65438-5

Abstract

Recent advances in wireless and sensing technologies have enabled the deployment of large scale Wireless Sensor Networks (WSNs) which have a wide range of scientific and commercial applications. However, due to the limited energy supply of sensor nodes, extending the network lifetime has become crucial for WSNs to deliver their promised benefits. Several proposals have aimed at this objective by designing energy efficient protocols at the physical, medium access, and network layers. While the proposed protocols achieve significant energy savings for individual sensor nodes, they fail to solve topology-related problems. An example of such problems is the bottlenecks around the sink, which is a direct result of multi-hop relaying: sensor nodes around the sink relay data generated all over the network which makes them deplete their energy much faster than other nodes.

A natural solution to this problem is to have multiple mobile data collectors so that the load is distributed evenly among all nodes. We investigate this promising direction for balancing the load and, hence, prolonging the lifetime of the network. We design optimization schemes for routing and placement of mobile data collectors in WSNs. We show, by theoretical analysis and simulations, that our approach has the potential to prolong the lifetime of the network significantly.

Acknowledgements

First of all, I thank ALLAH (God) for his mercy, help and guidance, without which this document would not be possible.

It has been a great honor for me to be supervised by Dr. Selim G. Akl and Dr. Hossam S. Hassanein. Their wide knowledge, guidance, enthusiasm in research, and nice personalities have been a great source of support for me throughout my research. Over the past four years, they have been always accessible and willing to help. I am grateful to Dr. Akl and Dr. Hassanein for the effort they spent to make me enjoy every minute I spent in research activities.

I owe my parents any success I make in my life. Their encouragement and support is a key factor in any achievement I have ever made. I am also indebted to my beloved wife, Nisren, for her continuous encouragement and patience during the course of my Ph.D. program. I also thank my children, Khaled and Salih, for their patience and for the happiness and motivation I always find in their smiles. I want to thank my brothers, Naeif, Abdullah, and Abdulrahman, and my sisters, Maha and Latifa, for their love and support.

I would like to thank all members of the School of Computing at Queens University for their support and friendship.

Special thanks to the Ministry of Higher Education in Saudi Arabia for their financial support.

Statement of Originality

I hereby certify that this Ph.D. thesis is original and that all ideas and inventions attributed to others have been properly referenced.

Contents

List of Figures

List of Tables

List of Acronyms

AUV	Autonomous Unmanned Vehicles
CPU	Central Processing Unit
CS	Carrier Sense
CTS	Clear To Send
DCP	Data Collector Placement
DCPR	Delay-Constrained Placement and Routing
DCR	Data Collector Record
DTPR	Delay-Tolerant Placement and Routing
GG	Gabriel Graph
GPS	Global Positioning System
GPSR	Greedy Perimeter Stateless Routing
HEED	Hybrid, Energy Efficient, Distributed
LEACH	Low Energy Adaptive Clustering Hierarchy
MAC	Medium Access Control
MDS	Minimum Dominating Set
MILP	Mixed Integer Linear Program
MM	Minimizing the Maximum energy
MOR	Maximal Overlapping Region
MR	Maximizing the minimum Residual energy
QoS	Quality of Service
RN	Relay Node
RNG	Relative Neighborhood Graph
RTS	Ready To Send
TDMA	Time Division Multiple Access
TNR	Trajectory Node Record
UCS	Unequal Clustering Size
WSN	Wireless Sensor Network

Chapter 1

Introduction

Advances in wireless communication and embedded microprocessors have brought about the development of small, low-cost sensor nodes, which are able to collect data from the surrounding environment and to communicate using a wireless medium. They report the data they collect to a central host, called the sink node. These tiny devices are battery-operated and, hence, untethered in terms of both power and communication. This enabled a new generation of large-scale networks of untethered, unattended sensor nodes suitable for a wide range of commercial, scientific, health, surveillance, and military applications. This rapidly evolving technology has the potential to revolutionize the way we interact with the physical environment and to facilitate collecting data which have never been available before [2]. However, the drastic energy constraints form a serious threat to the longevity of Wireless Sensor Networks (WSNs). Coping with this challenge and extending the lifetime of WSNs is the main focus of this research.

1.1 Motivations

Due to energy and size limitations, sensor nodes have a limited wireless transmission range. Therefore, sensor nodes, whose separation from the sink node is more than their

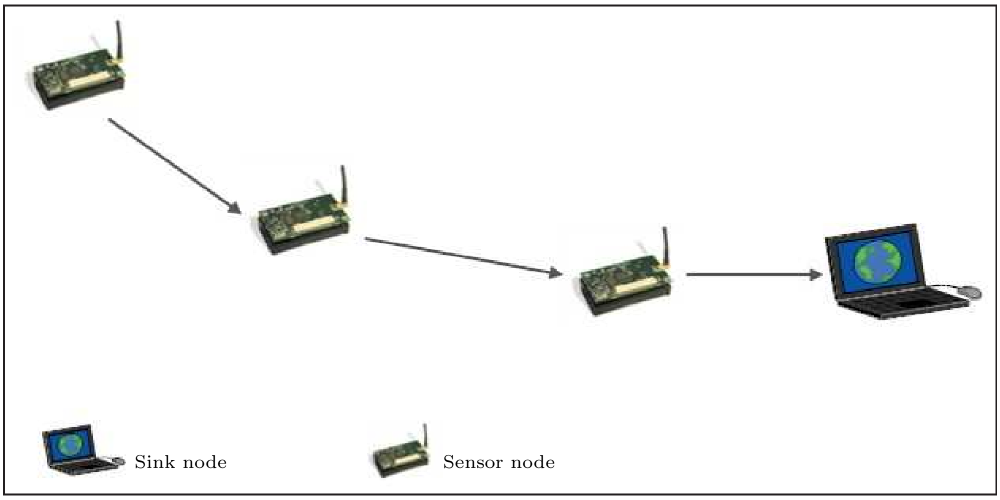

Figure 1.1: Multi-hop relaying.

transmission range, use multi-hop relaying to deliver their data to the sink node. Thereby, a sensor node may function as a router to forward data packets of other nodes to the sink node as shown in Fig. 1.1. However, data transmission is the dominant energy consuming operation in a sensor node; it has been stated in [3] that transmitting 1 Kb over a distance of 100 m would consume the same amount of energy as that of executing 3 million instructions by a local processor (i.e., the cost of sending 1 bit is equivalent to that of executing 3000 instructions). Since wireless communication is the major energy consumer in a sensor node, multi-hop relaying results in an unbalanced energy expenditure over different parts of the network: nodes near the sink become bottlenecks and deplete their energy reserves much faster than nodes distant from the sink node [4] [5] [6][7]. In fact, nodes which are one hop away from the sink node relay data generated all over the network to the sink node and, therefore, they run out of energy much faster than other nodes. Not only does this stop those nodes around the sink from functioning, but it also renders the sink unreachable by other nodes. Fig. 1.2 shows an example of a sensor network where the shaded sensor nodes are expected to have a higher data traffic than other nodes. While existing energy-aware protocols, at the physical, the Medium Access Control (MAC), and the network layers (see Section 2.2), achieve significant energy savings for individual sensor nodes, they fail to solve this topology-related problem.

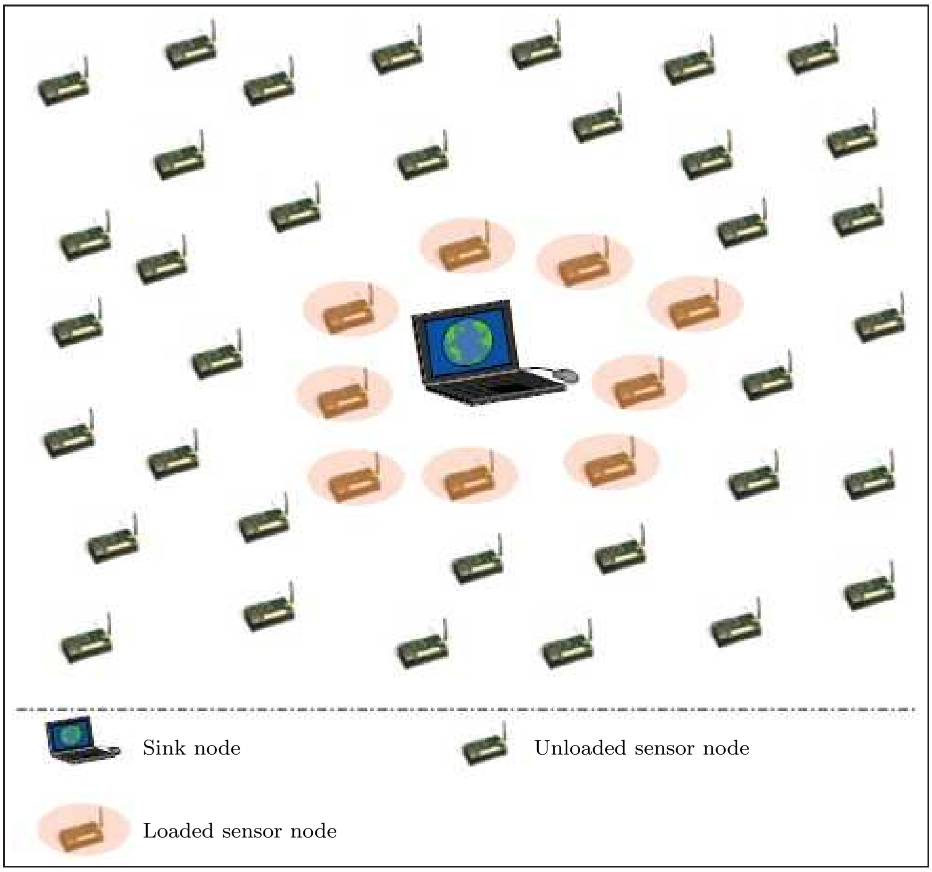

Figure 1.2: Bottlenecks around the sink.

To solve the problem of having a "hotspot" around a stationary sink, we argue for using multiple mobile sinks (which we call data collectors), which change their locations periodically. We address the problem of placing these data collectors in a way that balances the energy expenditure over different parts of the network and, hence, extends its lifetime. We also consider the problem of routing data from the stationary sensor nodes to the mobile data collectors.

With data collector mobility, the network can adapt to changes in data accumulation patterns. For example, a data collector should be located near areas where interesting events are taking place, so that data associated with these events travel over a small number of hops to save energy. The network can also adapt to changes in sensor nodes energy levels. Data collectors should be placed near nodes with relatively good amount of energy. These changes are usually driven by the application itself and by the nature of monitored events.

1.2 Thesis contributions

Our research aims primarily at prolonging the lifetime of a WSN by exploiting mobility of data collectors. We study the optimization problem of maximizing the network lifetime when mobile data collectors are in use. By having the data collectors mobile, the job of being a hotspot node is distributed over different parts of the network at different times and, hence, the relaying load disparity over the network is reduced. However, mobility of data collectors needs to be managed carefully in order to achieve that goal; and that is the focus of this research. To this end, we present novel schemes to place a set of mobile data collectors and to find the routing paths from the stationary sensor nodes to the mobile data collectors in WSNs. The main contributions of this thesis are the following:

1. In general, a data collector can be placed anywhere in the sensing field, which results in an infinite search space for data collectors locations. This has been a deterrent to full utilization of mobile data collectors [8]. We overcome this challenge by discretizing the search space without affecting the quality of the derived solutions.

2. To utilize mobility of data collectors, we divide the network lifetime into equal length rounds (e.g., hours, days, weeks, ... etc) and data collectors are moved to new locations at the beginning of each round. We define and solve the optimization problem of finding data collectors locations and routing paths from sensor nodes to data collectors for each round, such that the minimum residual energy at the end of the round is maximized. We formulate this problem as a Mixed Integer Linear Program (MILP).

3. We extend this optimization problem to an underwater 3D environment in which sensor nodes float at different depths and data collectors roam on the surface of the water.

4. We define and solve a delay-sensitive version of the previous optimization problem; we consider a delay constraint, which is an upper bound on the length of a path

between a sensor node and a data collector. This is useful for real-time applications such as a Tsunami warning system.

5. We propose a distributed routing scheme for WSNs with a mobile data collector moving along a globally known trajectory. The goal is to find routes to deliver delay-tolerant and delay-sensitive data to the data collector. While delay-tolerant data can wait for the data collector to come and pick them up, delay-sensitive data have to be sent to the data collector in its current location. We present a fully distributed routing scheme for such a configuration.

1.3 Document outline

The rest of this document is organized as follows. Chapter 2 presents the relevant background material and surveys previous related work. In Chapter 3, our optimization scheme for delay-tolerant placement and routing in terrestrial WSNs is presented. This includes discretizing the search space of data collectors locations and formulating the problem as a MILP. We present our optimization schemes for delay-tolerant and delay-sensitive placement and routing in underwater WSNs in Chapter 4. We show, in Chapter 4, how to extend the search space discretizing algorithm to handle multiple transmission ranges, delay attributes, and a 3D environment. We also show a linear programming formulation that accommodates delay constraints. In Chapter 5, we present our distributed scheme for routing to a mobile data collector. Finally, Chapter 6 concludes this document by highlighting the main issues in this thesis and outlining some future research directions.

Chapter 2

Background

This chapter presents the background material and surveys previous research related to the work in this thesis. It starts with an introduction to WSNs and their potential applications in Section 2.1. We overview some energy-aware protocols in WSNs in Section 2.2. In Section 2.3, previous work towards extending the network lifetime is classified into two main categories and the main schemes in each category are highlighted. The NP-completeness of the general data collector placement problem is shown in section 2.4. A brief summary is given in Section 2.5.

2.1 Wireless sensor networks

A WSN is composed of a large number of tiny sensor nodes and one or multiple more powerful sink nodes. Sensor nodes collect data from the surrounding environment and deliver the collected data to a sink node. Each sensor node has one or more sensors, a general-purpose Central Processing Unit (CPU) to perform arithmetic and logical operations, and a small amount of storage space. The power is supplied to these untethered sensor nodes through small, non-replenishable batteries. A sensor node has a wireless communication interface through which it can communicate with other nodes in its vicinity. Due to the scarcity of the power reservoir and due to the fact that communication is the dominant power consumer in a sensor node, the transmission range of these devices is limited for power efficiency purposes. Sensor nodes, which are spatially distant from the sink node, can report their data in a multi-hop fashion. A sink node is usually a more powerful device with a virtually unlimited power supply.

Crossbow MICAz motes [1], as an example of sensor nodes, have a low-power microcontroller that is in charge of regular processing, 128 KB of memory, and 512 KB of flash memory. MICAz motes have an IEEE 802.15.4 [9] compatible Radio frequency transceiver with a transmission rate of 250 kbps. A MICAz node has an outdoor transmission range of up to 100 m and an indoor transmission range of up to 30 m. Commercially available MICAz motes have 2X AA batteries which are enough for a week of full load [10]. MICAz mote is shown in Fig. 2.1.

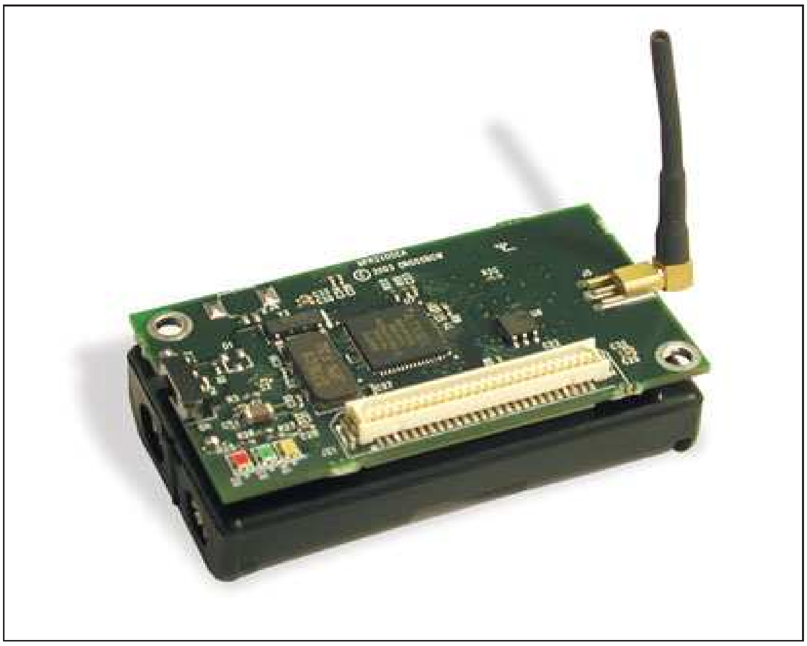

Figure 2.1: Crossbow MICAz node [1].

WSNs have a wide range of potential applications related to scientific, environmental, industrial, and military monitoring. WSNs have been deployed to monitor the seabird nesting environment and behavior in the Great Duck Island [11]. The ZebraNet system is a WSN deployed to support the research of biologists to monitor animal migration in Kenya [12]. WSNs have enabled a new generation of smart environments; an example of such an environment is the Gator Tech smart house in Florida, which is designed to assist elderly and disabled residents [13]. A prototype has been proposed in [14] for using WSNs to monitor drinking water quality. The design and implementation for using WSNs to monitor soil moisture is presented in [15]. An approach for using WSNs in monitoring the health of civil infrastructures (e.g., bridges and highways) is presented in [16]. These are just a few examples of how WSNs can be used to collect important data and facilitate significant services in real life.

2.2 Energy-aware protocols for WSNs

Energy awareness is an essential design issue for WSN communication protocols. In this section, we show how energy is conserved in some energy-aware MAC and routing protocols.

2.2.1 Energy-aware MAC protocols

In communications, a MAC protocol provides a mechanism to control the access of multiple stations (nodes) to a common physical communication channel. According to [17], the major causes of energy waste at the MAC level in WSNs are:

1. Collisions: a node receives more than one packet at the same time.
2. Overhearing: a node receives a packet which is not destined to it.
3. Idle listening: the receiver of a packet is active, yet no packets are being transmitted to it.
4. Overemitting: a node transmits a packet to a destination that is not ready.

Several energy-efficient MAC protocols have been proposed to alleviate the effect of these factors. Collisions are not specific to WSNs and can be dealt with using standard access control mechanisms. TRAMA [18] is a MAC protocol for WSNs that uses Time

Division Multiple Access (TDMA) to deal with collisions. WISEMAC [19] and S-MAC [20] are two protocols that use Carrier Sense (CS). Overhearing and idle listening are alleviated using sleep schedules which allow sensor nodes to turn off their wireless transceiver during sleep periods [21]. To avoid overemitting, a node wishing to communicate with another node initiates their dialogue by sending handshake control packets. The Request to Send / Clear to Send (RTS/CTS) are examples of such control packets which are used in the S-MAC protocol.

2.2.2 Energy-aware routing protocols

Routing is the process of finding paths between pairs of nodes in a network. In energy-aware routing, the quality of a path is estimated based on its effect on the energy reserves of nodes in the network. Tens of energy-aware routing protocols for WSNs have been reported during the last decade [22][23]. These protocols are classified into two categories: proactive and reactive. Proactive protocols are those in which paths are computed in advance. Reactive protocols are those in which a path is discovered only when it is needed. Because reactive protocols incurs high communication overhead for path discovery and setup, it is desirable to use proactive protocols in WSNs [22]. However, some geographic routing protocols (see Section 2.2.3), which are classified as reactive protocols, are able to find routing paths on-the-fly owing to the information nodes have about their geographic locations.

The majority of proactive routing protocols are based on constructing a tree rooted at the sink node [24][25][26]. Once such a tree is built, nodes deliver their data to the sink in a multi-hop fashion by passing them to their parents. The main idea of these protocols can be described as follows. At the deployment stage of the network, the sink initiates the tree construction process by broadcasting a tree construction packet. Each sensor node selects a sensor node, from which it has received the tree construction packet, to be its parent. When a node joins the tree, it rebroadcasts the tree construction packet to its neighbors. The selection of a parent can be made through different policies. An energy-efficient policy could be choosing the node that is closer to the root of the tree in terms of number of hops. Such a policy may reduce the number of transmissions needed to deliver data to the sink.

When the data generation rate of each sensor node is known, linear programming can be used to find the optimal routing in WSNs. In [27], the problem of finding the routing paths that maximize the network lifetime is formulated as a linear program. Such a formulation models a WSN as a flow network whose flow goes from sensor nodes to the sink node. The authors of [28] study the routing problem with splittable and unsplittable traffic. They present two linear programs for minimizing the total consumed energy in each problem.

Hierarchical routing is a class of routing protocols in which sensor nodes are partitioned into clusters. Each cluster has a cluster head and several cluster members. Members of a cluster report their data to the cluster head where data is aggregated and then delivered to the sink. Processing data at cluster heads saves significant amount of energy by removing redundant data, compressing data by computing aggregate functions (e.g., average, maximum, minimum, ... etc), and, thereby, reducing the number of transmissions. The Low Energy Adaptive Clustering Hierarchy (LEACH) is a clustering protocol that utilizes randomized rotation of cluster heads to distribute the energy load among sensor nodes in the network [4]. In LEACH, the system operates in rounds, where each round begins with cluster heads selection followed by steady data transmission. During the cluster heads selection phase, each node n_i generates a threshold $T(i)$ as follows:

$$T(i) = \begin{cases} \dfrac{P}{1 - P * (r \mod \frac{1}{P})} & \text{if } i \in H \\ 0 & \text{otherwise} \end{cases}$$

where P is the desired percentage of cluster heads, r is the current round, and H is the set of sensor nodes that have not been chosen to be cluster heads during the last $\frac{1}{P}$ rounds. At the beginning of each round, each sensor node n_i generates a random number between 0 and 1, and elects itself to be a cluster head if this random number is less than $T(i)$. Cluster heads broadcast advertisement messages to all other nodes. Other nodes listen to advertisement messages, decide on the cluster they wish to belong to, and inform the head of the chosen cluster about their decision. Cluster heads then create transmission schedules to their clusters and start receiving data from members of their clusters. Once a cluster head receives data from all members in its cluster, it performs the required data aggregation and processing, and sends the compressed data directly to the sink. While LEACH represents a clustering protocol that evenly distributes the energy load amongst all nodes, it assumes that all nodes are able to communicate directly with the sink, which does not scale to large-size networks.

The Hybrid, Energy Efficient, Distributed (HEED) clustering [29] differs from LEACH in that it considers the residual energy of different nodes in the selection of cluster heads; nodes with higher residual energy are more likely to be chosen as cluster heads. Moreover, HEED makes use of multi-hop intercluster communication. HEED constructs a connected multi-hop cluster head graph to facilitate this intercluster communication.

2.2.3 Geographic routing

We use geographic routing in parts of the scheme we present in Chapter 5. In this subsection, we give the background material necessary to understand those parts. In geographic routing, it is assumed that sensor nodes know their geographic locations and locations of their immediate neighbors. A node can obtain its location using a Global Positioning System (GPS) or using one of the less expensive localization methods described in [30][31][32]. A routing node forwards a packet based on the geographic location of its destination and locations of immediate neighbors. Geographic routing algorithms can make efficient routing decisions based on local information, i.e., avoiding the overhead of maintaining global topology information. This makes geographic routing suitable for resource constrained WSNs.

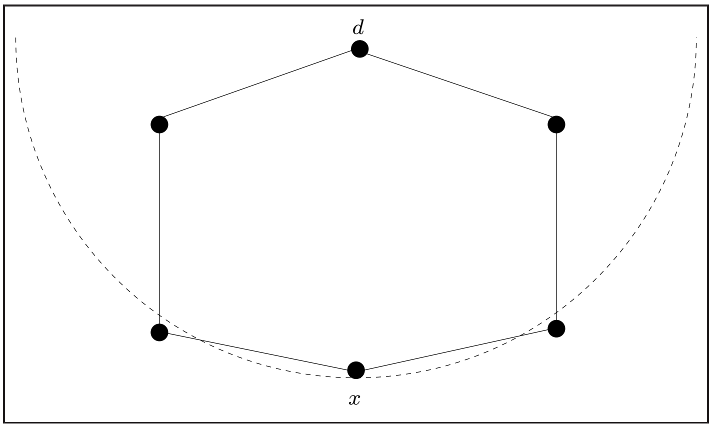

Figure 2.2: Using greedy distance routing, a packet destined to d may get stuck at x.

Greedy algorithms

Because maintaining global information is hard to achieve in WSNs, *locality* is a desired property for any distributed algorithm designed for WSNs. A distributed algorithm is localized if each node makes its decision based on information it has about itself, information about its immediate neighbors, and probably a constant size global information. *Greedy distance routing* [33] is a localized geographic routing protocol in which a routing node forwards a packet to its immediate neighbor which is the closest to the destination. A similar greedy algorithm is *compass routing* [34] in which a routing node x forwards the packet to its immediate neighbor y which minimizes the angle $\angle dxy$, where d is the destination node [35]. While simple, these two algorithms do not guarantee packet delivery; a packet may get stuck at a local minimum node at which no progress can be made towards the destination as shown in the example of Fig. 2.2. Only under global assumptions on the connectivity graph can these algorithms guarantee packet delivery. The greedy distance routing, for instance, guarantees packet delivery if the connectivity graph contains the *Delaunay triangulation* [36] of the sensor nodes. Compass routing guarantees delivery if the connectivity graph is the Delaunay triangulation of the sensor nodes. These assumptions are hard to meet in practice.

Planarization of the connectivity graph

Surprisingly, removing some edges from the connectivity graph is a condition for some routing algorithms to guarantee packet delivery, i.e., it does not get stuck at a local minimum node nor does it get trapped in a cycle. Some routing algorithms guarantee packet delivery if the connectivity graph is planar. The idea is to have a planar subgraph of the connectivity graph, and to apply one of these algorithms to provide guaranteed packet delivery. Fortunately, there are simple localized algorithms for constructing a planar subgraph from an arbitrary connectivity graph. The *Relative Neighborhood Graph* (RNG) [37] and the *Gabriel Graph* (GG) [38] are examples of planar graphs that can be constructed by localized algorithms. The RNG preserves an edge xy if the intersection of the two circles centered at x and y with a radius of the distance between x and y is free of other nodes. The GG preserves an edge xy if the circle, whose diameter is the line segment xy, is free of other nodes. *Perimeter routing* [34] is an example of a routing algorithm that guarantees delivery on planar connectivity graphs. In perimeter routing, when a packet destined to a node d is currently at a node s, the routing process starts with the face beyond s along the segment sd. Note that sd may intersect with one or more boundary segments of that face. The packet traverses the boundary segments of that face and gets back to s. The packet keeps with it the boundary segment kl that intersects with sd at the point furthest from s. When the packet gets back to s, it is sent to k. This process continues until a face that contains d is reached. If all faces of the planar graph are convex, sd intersects with only one boundary segment kl. So there is no need to traverse all boundary segments; the packet can, rather, walk along the boundaries (either clockwise or counterclockwise) until it reaches a segment that intersects with sd, at which point it switches to the next face as shown in Fig. 2.3.

Greedy perimeter stateless routing

Greedy geographic routing algorithms are simple but do not guarantee packet delivery. Other algorithms, such as perimeter routing, guarantee packet delivery but require a

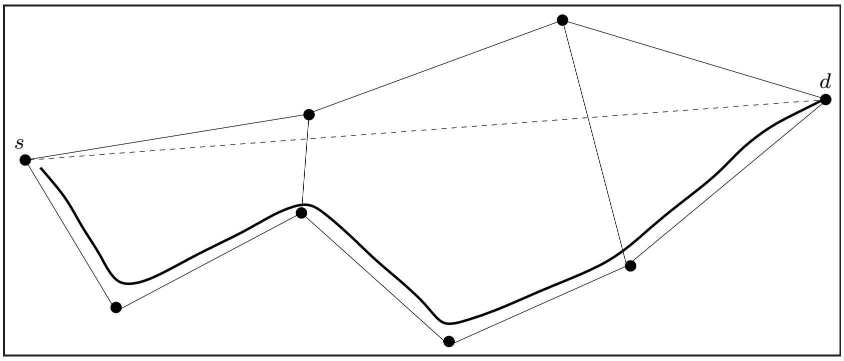

Figure 2.3: Perimeter routing on a convex-faces planar graph.

planar version of the connectivity graph, and use more complex routing algorithms. An interesting observation about WSNs is that they are anticipated to have a huge number of sensor nodes uniformly distributed over a sensing field. Dense deployment makes greedy algorithms work most of the time. However, there might be some holes in the topology of the network which results in some local minimum if a greedy method is used. *Greedy Perimeter Stateless Routing* (GPSR) [39] makes use of this interesting observation. GPSR combines a greedy method on the connectivity graph G and a perimeter method on a planar subgraph G'. GPSR starts with the greedy mode on G. During the greedy mode, each node forwards the packet to the neighbor which is the closest to the destination. If the packet arrives to a local minimum node, GPSR switches to the perimeter mode on G'. During the perimeter mode, the packet transitions to faces of G' closer to the destination. Eventually, the packet will reach a node closer to the destination than the node where the perimeter mode began, and will switch back to the greedy mode. This process continues until the packet reaches its destination. If there is no sensor node at the destination location, the packet will reach the face that contains that location and will be delivered to a sensor node along the perimeter of that face which is the closest to the destination location.

2.3 Device deployment and planning in WSNs

A critical performance measure for WSNs is the network lifetime. A significant amount of research has been done towards prolonging the lifetime of the network. The majority of such research targets energy-efficient routing and MAC protocols [21][20][17][4][23][22]. However, existing energy-aware routing and MAC protocols are not able to solve the topological problem of having bottlenecks around the sink and, hence, have a limited impact on the network lifetime. Accordingly, researchers in the field moved to device provisioning schemes in order to alleviate the effect of topological deficiencies in the network. Proposals in that direction can be classified into two streams: proposals with stationary devices and proposals with mobile devices. This section highlights proposals in these two streams.

2.3.1 Stationary devices

The use of stationary devices for load balancing in WSNs has been achieved through choosing the types, numbers, and locations of different devices in the network. One way to balance the load over the network is to deploy more nodes in areas closer to the sink. The authors in [40] proposed a nonuniform node distribution strategy that divides the sensing field into a number of coronas and gives the ratio in node densities between two consecutive coronas. The work in [41] is another attempt to alleviate the effect of having bottlenecks around the sink by making the density of sensor nodes in an area inversely proportional to its distance to the sink. One problem of this approach is the exponential growth of the total number of sensor nodes in the network. Moreover, with a fixed budget on the total number of sensor nodes, this approach results in a nonuniform sensing coverage over the sensing field.

The authors of [27] studied the problem of routing and placement of multiple sink nodes for maximizing the network lifetime and proposed an approximation algorithm to solve it. However, the time complexity of the proposed scheme is high; it involves solving a number of linear programming instances which is exponential in the number of sink nodes. The authors of [42] defined the following problem: given the power supply of each sensor node and a desired lifetime of the network, find the locations of R sink nodes and the routing paths from sensor nodes to sink nodes, such that the data rate is maximized. The data rate is the number of data units each sensor node sends to a sink node per time unit. When the possible locations of the sink nodes are limited to the exact locations of sensor nodes, the problem is shown to be NP-complete by a reduction from the problem of finding the minimum cardinality dominating set on unit disk graphs [43]. Two heuristic algorithms, a greedy algorithm and a local search algorithm, were proposed for this problem in [42]. It is important to notice that maximizing the data rate for a given lifetime is equivalent to maximizing the lifetime for a fixed data rate.

The work in [44] targets placing a single stationary sink node to collect data from a set of sensor nodes, with variable transmission ranges, when single-hop communication is used. It was shown that finding a sink location that maximizes the network lifetime is equivalent to finding the minimum radius disk that encloses all nodes. An optimal, linear-time algorithm for this problem was presented in [44].

A scheme for placing a number of data collectors to maximize the lifetime of the network and another scheme to find the minimum number of data collectors to meet a desired network lifetime are presented in [45]. Two genetic algorithms for placing a number of data collectors for the objective of minimizing the data collection latency are proposed in [46]. The work in [47] targets the problem of finding the minimum number of data collectors that can meet an upper bound on the maximum data collection latency. Another scheme for placing a number of data collectors in WSNs to achieve minimal latency is proposed in [48].

Some of the work in this literature uses a tier of Relay Nodes (RNs): a RN is a more powerful node that collects data from nearby sensor nodes and sends them to the sink. The Unequal Clustering Size (UCS) was proposed in [49]. In the UCS scheme, a number of RNs are deployed to serve as cluster heads. Each RN collects and aggregates data from sensor nodes in its cluster and sends the aggregated data to the sink. Sending data to the sink is made through multi-hop relaying; every RN forwards its aggregated data to a neighboring RN which is closer to the sink. As a result of this multi-hop relaying, RNs closer to the sink are assigned higher intercluster load than other RNs. The UCS scheme deals with this problem by making the cluster size proportional to the distance to the sink. Thereby, RNs with higher intercluster traffic are serving smaller clusters and, hence, are having a lower intracluster load.

The authors in [50] proposed an energy provisioning approach. The unbalanced load is alleviated by assigning RNs different amounts of energy and by deploying more RNs, if necessary. A heuristic based algorithm is presented to maximize the network lifetime under a certain energy and extra RNs budget. The work in [51] and [52] targets another variation of the problem: finding the minimum number of RNs and their locations to guarantee a desired network lifetime. The scheme presented in [51] is composed of two phases. In the first phase, some RNs are deployed to guarantee the connectivity of sensor nodes. In the second phase, extra RNs are deployed to connect the first phase RNs to the sink and to make all RNs meet the desired network lifetime. In [52], another best effort algorithm to place the second phase RNs is introduced and a theoretical lower bound on the number of the second phase RNs is derived. However, the use of RNs suffers a critical drawback: when sensor nodes use multi-hop relaying to reach the RNs, there will be a load balancing problem among sensor nodes within single clusters; and to make sensor nodes use single-hop transmission, a large number of RNs will be needed and a load balancing problem among RNs will arise.

2.3.2 Mobile devices

The main topological problem in WSNs stems from the fact that the load on a particular sensor node is determined by its distance to the sink: with multi-hop relaying, nodes that are one hop away from the sink relay data generated all over the network, and with single-hop communication, nodes distant from the sink deplete their energy faster than nodes near the sink. Such a problem is hard to alleviate when both the sink and the sensor nodes are stationary. While sensor nodes, which are envisioned to be small in size and low in cost, are hard to be made mobile, it is plausible to have a reasonable number of mobile data collectors to balance the load in the network and to prolong its lifetime. Recently, only few schemes that follow that stream have been proposed.

The work in [5] and [53] assumes the existence of a number of predefined spots where data collectors can be placed; the network lifetime is divided into equal length rounds and data collectors are moved to new locations at the beginning of each round. In [5], the network is modeled as a flow network and the problem of finding the optimal data collectors locations is formulated as a MILP whose objective is either minimizing the maximum energy spent by a single sensor node or minimizing the total consumed energy during the round. We argue that these two objective functions are not really suitable for the placement of mobile data collectors. This is because the optimal solution towards such objectives will not change over time and, hence, locations of data collectors will not be changed. A heuristic scheme, that considers residual energy, is proposed in [53]. However, locations of data collectors are chosen according to local information only: the decision of whether or not a data collector is placed at a given location is made based on the residual energy of sensor nodes that are one hop away from that location. While very simple, this scheme may obtain solutions which are not even close to the optimal one.

The authors in [8] proposed a heuristic for repositioning a data collector to extend the longevity of the network and to improve the data gathering timeliness. That work strives to answer questions about when and how the position of the data collector is changed. A metric of the traffic density and power consumption is used to monitor the status of each sensor node in the network. The decisions of whether or not to move the data collector and to which location it is moved are made based on the status of sensor nodes in the vicinity of the data collector. When such a relocation seems to be beneficial to the network, a local search is made to find a better position. Since the data collector is not moved far from its current location, changes to the routing paths are limited, which reduces the overhead of such a repositioning. In [8], another motive to change the location of the data collector is to improve the timeliness of the data gathering process for real-time applications. Data collector repositioning could be beneficial to meet delay-based Quality of Service (QoS) requirements; for example, getting closer to areas where real-time events are taking place could reduce the delay between the sources of these events and the data collector.

Recently, linear programming has been used to find the optimal routing and placement of a single mobile data collector [54][55]. The problem is defined as follows: given a set of predefined points where the data collector can be located, find the time the data collector stays at each location (i.e., the sojourn time) and the multi-hop routing paths from sensor nodes to the data collector at each location. This important result was found through a nice observation that the order different locations are being visited by the data collector does not affect the network lifetime; what really affects it is the sojourn time and the routing strategy associated with each location. However, this observation does not hold when more than one data collector are available and, hence, this scheme can not be extended to multiple data collectors.

The authors in [56] investigate the problem of uneven relaying load in WSNs. They compare three approaches: nonuniform node densities, multiple sink nodes, and a mobile sink node. Their results indicated that using multiple sink nodes is the most effective solution. Yet, it is the most expensive one. Repositioning the sink node has shown a good performance at a lower cost. However, nonuniform node densities was not recommended as it demands deploying a significant number of extra sensor nodes.

The use of mobile data collectors has been also mentioned in the context of underwater WSNs. The work in [57] involves the hardware and networking software design for using mobile data collectors in underwater WSNs. According to [57], each sensor node collects data and waits for the mobile data collector to become close enough to receive data in a single hop. The authors of [58] try to find the relationship between the number of data collectors and the time needed to harvest all data when data collectors follow random trajectories. Both the work in [57] and that in [58] assume single-hop communication and do not consider the problem of finding a path for the data collector. Moreover, both schemes are not suitable for delay sensitive applications.

2.4 NP-completeness of the data collectors placement problem

For the sake of completeness we show the NP-completeness proof presented in [42] with a minor modification to reflect the objective of maximizing the network lifetime rather than maximizing the data rate. We first make the following definitions:

Definition 1 *For a graph* $G = (V, E)$*, where* V *is the set of vertices and* E *is the set of edges, a dominating set is a subset* D *of* V*, such that for each vertex* $v \in V$ *either* $v \in D$ *or there exists a vertex* $u \in D$ *that shares an edge with* v*.*

Definition 2 *A graph* G *is a unit disk graph if and only if it has the following property: There is an edge between a pair of vertices if and only if their separation is at most 1 distance unit.*

Definition 3 *The Minimum Dominating Set (MDS) problem on unit disk graphs:*
Input: A unit disk graph G*.*
Question: Is there a dominating set of size d*.*

Definition 4 *The Data Collector Placement (DCP) problem:*
Input: Locations of sensor nodes loc_0*,* loc_1*, ...,* loc_{N-1}*; energy supplies of sensor nodes* E_0*,* E_1*, ...,* E_{N-1}*; a data rate* ρ *data units/time unit from every sensor node; a transmission range* r *distance units; and a number of data collectors* R*.*
Question: Is there a placement of R *data collectors at sensor nodes locations and a multi-hop routing strategy that keep all sensor nodes alive for at least* k *time units.*

Two sensor nodes are said to be neighbors if their separation is at most r distance units. Sending a data unit consumes one energy unit, and a sensor node is considered to be alive as long as its residual energy is greater than zero.

Now we state the following theorem.

Theorem 1 *The DCP problem is NP-complete.*

Proof

First, DCP is in NP:

A solution to the DCP problem is in the form of a subset $\mathcal{P}$ of R sensor nodes whose locations are chosen to place the data collectors at, and an assignment of flow values between neighboring sensor nodes. f_{ij}, which denotes the flow from a sensor node n_i to a neighboring sensor node n_j, is the total number of data units to be sent from n_i to n_j. Let F_i^+ denote the sum of the outgoing flow from sensor node n_i, and let F_i^- denote the sum of the incoming flow to sensor node n_i. To certify that a solution is feasible and has a lifetime of at least k time units, one needs to check the following conditions:

1. $\mathcal{P}$ has at most R sensor nodes.
2. If a sensor node n_i is in $\mathcal{P}$, $F_i^+/\rho \geq k$ and $F_i^+ \leq E_i$.
3. If a sensor node n_i is not in $\mathcal{P}$, $(F_i^+ - F_i^-)/\rho \geq k$ and $F_i^+ \leq E_i$.

Since these conditions can be checked in a polynomial time, the DCP is in NP.

Second, DCP is NP-hard:

The NP-hardness of the DCP problem can be shown through a reduction from the MDS problem, which has been shown to be NP-hard in [43]. An instance A of the MDS problem, on a graph G, can be reduced to an instance B of the DCP problem in which there is a sensor node for each vertex in G, every sensor node has a residual energy of 1 energy unit, r is set to 1 distance unit, R is set to d, and ρ is set to 1 data unit/time unit. There exists a dominating set of size at most d for A if and only if there is a placement for the data collectors and a routing strategy with a lifetime of at least 1 time unit for B. This can be shown as follows.

If there is a dominating set of size at most d for A, we can place a data collector at the location of each vertex in the dominating set. Then, by the definition of a dominating set, each sensor node will have a data collector either in its location or at the location of one of its neighbors. Thus, each sensor node can send one data unit directly to a data collector which results in a lifetime of 1 time unit for B.

Now assume that there is a placement for the d data collectors and a routing strategy with a lifetime of at least 1 time unit for B. The data collectors are placed at the locations of a set $\mathcal{P}$ of d sensor nodes. Since each sensor node has only 1 energy unit (which is just enough to send its own data), no sensor node is able to relay a single data unit for any other sensor nodes. Therefore, each sensor node will have a data collector placed not more than 1 distance unit away from it. Since data collectors are placed at sensor nodes locations only, each sensor node must have a data collector placed at its own location or at the location of one of its neighbors. Thus, the set $\mathcal{P}$ form a dominating set for the underlying connectivity graph of the sensor network and, hence, a dominating set of size d for B.

Therefore, the DCP problem is NP-complete. □

When the data collectors are allowed to be placed at any point, rather than at the locations of sensor nodes only, the problem seems to be harder. However, we are not aware of any formal proof for the NP-hardness of that problem.

2.5 Summary

This chapter surveys existing proposals for handling the bottleneck problems in WSNs. These proposals are classified into two categories: proposals with stationary devices and proposals with mobile devices. Proposals in the former category use nonuniform node distribution [40][41], multiple RNs with and without variable energy provisioning

	Data collectors		Objective	Optimal	Multi-hop relaying	Technique
	Multiple	Mobile				
[27]	Yes	No	Lifetime	No	Yes	Linear programming
[45]	Yes	No	Lifetime	No	Yes	Clustering
[42]	Yes	No	Maximum data rate	No	Yes	Greedy algorithm
[44]	No	No	Lifetime	Yes	No	Computational geometry
[5]	Yes	Yes	Lifetime	No	Yes	Linear programming
[53]	Yes	Yes	Lifetime	No	Yes	Heuristic
[8]	No	Yes	Lifetime/ timeliness	No	Yes	Local search
[54][55]	No	Yes	Lifetime	Yes	Yes	Linear programming

Table 2.1: Data collector placement schemes.

[49][50][51][52], or multiple stationary data collectors [27][42]. Proposals in the later category exploit mobility of data collectors to distribute the load over the network and alleviate the bottleneck problem [5][53][8][54][55]. Table 2.1 gives a comparison between different data collector placement schemes discussed in this chapter. Optimal polynomial-time schemes exist only for special settings and assumptions: there is a linear-time optimal algorithm to place a single stationary data collector in a single-hop network with a variable transmission range [44], and there is a linear programming approach to find the optimal routing and placement of a single mobile data collector in a multi-hop network [54][55]. The problem of placing more than one mobile data collectors in a multi-hop network is proven to be NP-complete when data collector locations are limited to the exact locations of sensor nodes [42]. The problem does not seem to be less hard when data collectors can be placed anywhere in the sensing field, yet no formal NP-hardness proof exists in the literature.

Chapter 3

Mobile data collectors in terrestrial WSNs

A major challenge affecting the lifetime of WSNs comes from the unbalanced energy consumption over different parts of the network. In this chapter, we present a mobile data collector placement scheme for extending the lifetime of the network. The lifetime of the network is divided into rounds and data collectors are moved to new locations at the beginning of each round. We define and solve two problems: the on-track placement where data collectors can be placed only along predefined tracks (roads) spanning the sensing field, and the general placement where data collectors may be placed at any point in the sensing field.

We formulate the problems as MILPs and use a linear programming solver (with a constant time limit) to find near-optimal placements of the data collectors and to find routing paths to deliver data to data collectors. Our experiments show that our approach makes a significant extension to the lifetime of the network.

The rest of this chapter is organized as follows. Section 3.1 outlines the proposed approach and highlights its contributions. Section 3.2 describes the model of the system and gives a formal problem definition. In Section 3.3, we present our placement schemes. Section 3.4 shows the experimental results. Finally, Section 3.5 concludes this chapter with a brief summary.

3.1 Schemes outline and contributions

We argue for using multiple mobile data collectors and propose an approach for placing these data collectors in a way that balances the energy expenditure and increases the lifetime of the network. Our approach divides the lifetime of the network into fixed length rounds (e.g., hours, days, or weeks) and moves the data collectors to new locations at the beginning of each round. Our policy of maximizing the lifetime is to maximize the minimum residual energy at the end of each round. Some recently proposed schemes have addressed the issue of mobile sinks. However, some of that work assumes the existence of a set of predefined spots (i.e., points) where data collectors may be placed [5], and some is limited to placing data collectors at the boundaries of the field [53]. To this end, the novel contribution of this work is twofold:

1. We define and solve two placement and routing problems. The first one assumes the existence of predefined tracks (e.g., a road network) spanning the sensing field, and data collectors can be moved over and placed at any point along these tracks. This would be practical in a situation where data collectors are carried on Autonomous Unmanned Vehicles (AUVs) or robots that move along paved roads only. In the second one, a data collector can be placed anywhere in the sensing field.

2. We discretize the search space of data collector locations without affecting the quality of the derived solutions: we devise an algorithm that finds a finite set of relatively small number of points, and we prove the existence of an optimal placement in which each data collector is placed at a point in that set. Since the problem is modeled as a MILP, making the cardinality of this set as small as possible would significantly improve the efficiency and the solution quality.

By formulating the problems as a MILP, an optimal solution can be found. However, this may require an exponential time in the worst case [59]. Therefore, we impose a time