Theoretical Results on Base Station Movement Problem for Sensor Networks

Yi Shi Y. Thomas Hou*

The Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA

Abstract

The benefits of using mobile base station to prolong sensor network lifetime have been well recognized. However, due to the complexity of the problem (time-dependent network topology and traffic routing), theoretical performance limit and provably optimal algorithms remain difficult to develop. This paper fills this important gap by contributing theoretical results regarding the optimal movement of a mobile base station. Our main result hinges upon a novel transformation of the joint base station movement and flow routing problem from time domain to space domain. Based on this transformation, we first show if the base station is allowed to be present only on a set of pre-defined points, then we can find the optimal time duration for the base station on each of these points so that the overall network lifetime is maximized. Based on this finding, we show that when the location of the base station is un-constrained (i.e., can move to any point in the two-dimensional plane), we can develop an approximation algorithm for the joint mobile base station location and flow routing problem such that the network lifetime is *guaranteed* to be at least $(1 - \varepsilon)$ of the maximum network lifetime, where ε can be made arbitrarily small depending on required precision.

Keywords

Theory, approximation algorithm, optimization, mobile base station, lifetime, sensor networks.

1 Introduction

The benefits of using mobile base station to prolong sensor network lifetime have been well recognized [15, 28]. Since base station is the sink node for data generated by all the sensor nodes in the network, this approach aims to alleviate the traffic aggregation burden from a *fix* set of sensor nodes near the base station to other sensor nodes in the network, and it is possible to extend the network lifetime significantly.

Indeed, although the practical feasibility of using a mobile base station is still considered far-fetched a few years ago, such capabilities are nowadays reality, thanks to recent breakthrough in unmanned autonomous vehicle (UAV) competition by DARPA's Grand Challenge program [6] and advances in customized robotics for sensors [21]. It has

*Please direct all correspondence to Prof. Tom Hou, The Bradley Department of Electrical and Computer Engineering, 302 Whittemore Hall (0111), Virginia Tech, Blacksburg, VA, 24061, USA. Phone: +1-540-231-2950; Fax: +1-540-231-8292; Email: thou@vt.edu, URL: http://www.ece.vt.edu/thou/index.html.

now become entirely plausible to consider that an unmanned vehicle be employed to serve as a mobile base station for sensor data collection.

Although the potential benefit of using mobile base station to prolong sensor network lifetime is significant, the theoretical difficulty of this problem is enormous. There are two components that are tightly coupled in this problem. First, the location of the base station is now a function of time, i.e., at different time instances, we have a different physical network topology, with the sink node being at different positions. Second, the traffic (or flow) routing behavior may change with both time as well as the location of the base station. As a result, an optimization problem with the objective of maximizing the network lifetime needs to consider both base station location (time-dependent) and flow routing. Due to these difficulties, existing solutions to this problem remain heuristic at best (e.g., [15, 28]) and cannot provide *provably* optimal solution to network lifetime performance.

To fill in this theoretical gap, this paper offers an in-depth study on the network lifetime performance limit when mobile base station is employed. We formulate an optimization problem with base station movement and flow routing as part of the constraints. As a first step, we show that as far as network lifetime objective is concerned, flow routing only needs to be dependent on the base station location, regardless of *when* the base station is present at this location. Further, the specific time instances for the base station to visit a location is not important, as long as the total time duration for the base station to be present at this location is fixed. With this finding, we show how to make a novel transformation from a time-dependent problem formulation to location (space)-dependent problem formulation.

As a second step, we show that when base station is only allowed to be present at a finite set of pre-determined points (called *constrained mobile base station* (C-MB) problem), we can find the optimal time duration for the base station to stay on each of these points (as well as the corresponding flow routing solution) such that the overall network lifetime (i.e., sum of the time durations) is maximized via a single linear program (LP).

Building upon these results, the main result in this paper (Section 5) shows that for the *un-constrained mobile base station* (U-MB) problem, i.e., the base station can be present at any point in the two dimensional plane, we can develop a *provably* $(1-\varepsilon)$ optimal algorithm that provides a solution with network lifetime guaranteed to be at least $(1-\varepsilon)$ of the maximum network lifetime (albeit it is unknown), where ε can be made arbitrarily small depending on required precision. The main idea in this approximation algorithm is to exploit a clever way of dividing the search space into subareas, with each of its link cost having some nice properties that are related to $(1-\varepsilon)$ optimal. A novel idea in the design of $(1-\varepsilon)$ optimal algorithm is to represent each subarea with so-called "fictitious cost point," which is an N-tuple cost vector with each component representing an upper bound of cost to the respective node. As a result, we can apply the LP approach developed for the C-MB problem on the fictitious cost points and develop provably $(1-\varepsilon)$ optimal solution.

The rest of this paper is organized as follows. In Section 2, we describe the network model and introduce the mobile base station problem. In Section 3, we present a novel transformation that enables to transform a time-dependent problem to a space-dependent problem. In Section 4, we develop optimal solution for the constrained mobile base station (C-MB) problem. Section 5 presents the main result of this paper, which is an algorithm for the unconstrained mobile base station (U-MB) problem with provably $(1-\varepsilon)$ optimal network lifetime. Section 6 reviews related work and Section 7 concludes this paper.

2 Network Model and Problem Formulation

2.1 Network Model

We consider a set of nodes $\mathcal{N}$ deployed over a two-dimensional area, with the location of each sensor node $i \in \mathcal{N}$ being at a point (x_i, y_i) . We assume each node generates data at a rate of r_i, although extension to variable bit rate case may also be developed. There is a base station B for the sensor network and it serves as the sink node for all data. Data generated by each sensor node can be relayed via single or multi-hop toward the base station.

We now discuss the energy consumption due to communications (i.e., data transmission and reception). We assume that each node has power control capability.[1] Suppose that node i transmits data to node j with a rate of f_{ij}, then the transmission power at node i can be modeled as [11]

$$p_{ij}^t = c_{ij} \cdot f_{ij} \ , \tag{1}$$

where c_{ij} is the cost on wireless link (i, j) and can be modeled as

$$c_{ij} = \alpha + \beta \cdot d_{ij}^n \ , \tag{2}$$

with α and β being two constant terms and d_{ij} the physical distance between nodes i and j. n is the path loss index and is typically between $2 \le n \le 4$ [18].

The receiving power consumption at sensor node i can be modeled as [11]

$$p_i^r = \rho \sum_{1 \le k \le N}^{k \ne i} f_{ki} \ , \tag{3}$$

where ρ is a constant and f_{ki} is the incoming bit-rate received by sensor node i from sensor node k.

In this theoretical study, we assume a contention-free MAC protocol, where interference from simultaneous transmission can be effectively minimized or avoided. For given rate traffic pattern, a contention-free MAC protocol is fairly easy to design (see, e.g., [23]) and its discussion is beyond the scope of this paper.

Each node $i \in \mathcal{N}$ is initially provisioned with an amount of energy e_i. The base station is not constrained with energy. In this study, network lifetime is defined as the first time instance when any of the sensor nodes runs out of

[1] If a two-tier wireless sensor network architecture (e.g., [7, 12, 14, 17]) is used, then a sensor node in this paper corresponds to the upper tier backbone node that is used for data forwarding.

Table 1: Notation.

General notation	
$\mathcal{N}$	The set of sensor nodes in the network
N	$= \lvert\mathcal{N}\rvert$, the number of sensor nodes in the network
B	Denotes the base station
$(x, y)(t)$	Location of base station B at time t
(x_i, y_i)	Cartesian coordinate of sensor node i
r_i	Bit rate generated at sensor node i
e_i	Initial energy at sensor node i
α, β	Two constant terms in power consumption model for data transmission
ρ	Power consumption coefficient for receiving data
n	Path loss index, $2 \leq n \leq 4$
C_{ij}	Cost for transmitting data from sensor i to sensor j
$c_{iB}(t)$	Cost for transmitting data from sensor i to base station B at time t
$c_{iB}(p)$	Cost for transmitting data from sensor node i to base station B when B is at point p
$g_{ij}(t)$ (or $g_{iB}(t)$)	Flow rate from sensor node i to sensor node j (or base station B) at time t
$f_{ij}(p)$ (or $f_{iB}(p)$)	Flow rate from sensor node i to sensor node j (or base station B) when B is at point p
$W(p)$	Time duration for the base station to be present at point p
C-MB problem specific notation	
M	The number of pre-determined locations
$\psi^*_{\text{C-MB}}$	An optimal solution to the C-MB problem
$T^*_{\text{C-MB}}$	The maximum network lifetime achieved by $\psi^*_{\text{C-MB}}$
U-MB problem specific notation	
ε	Required approximation precision, $\varepsilon > 0$ and $\varepsilon \ll 1$
$\mathcal{A}$	The search space for the base station, which can be the smallest enclosing disk to cover all sensor nodes
$O_{\mathcal{A}}, R_{\mathcal{A}}$	The center and radius of $\mathcal{A}$
M	The number of subareas
$\mathcal{A}_m$	The m-th subarea in the search space
$W(A_m)$	Time duration for the base station to be present in subarea A_m
$c^{\min}_{iB}, c^{\max}_{iB}$	Lower and upper bounds of $c_{iB}(p)$
$C[h]$	$= \alpha(1+\varepsilon)^h$, the transmission cost for the h-th circle
H_i	The required number of circles at sensor node i
$\psi_{\text{U-MB}}$	$(1-\varepsilon)$ optimal solution to the U-MB problem
$T_{\text{U-MB}}$	$(1-\varepsilon)$ optimal network lifetime achieved by $\psi_{\text{U-MB}}$

energy. From (1), (2), and (3), it is easy to understand that the location of the base station and the corresponding flow routing among the nodes will determine energy consumption behavior at each node and thus the network lifetime. Table 1 lists all notation used in this paper.

2.2 Problem Description

The focus of this paper is to investigate how to optimally move a mobile base station to collect real time data in a sensor network so that the network lifetime can be maximized. Note that the network lifetime problem has attracted great interest even for static (fixed) base station problem (see, e.g., [3, 4, 5, 20])

As a first step, we consider the case when the base station is only allowed to be present at a set of pre-determined M positions, denoted as $p_1, p_2, \cdots, p_M$. We call this problem as constrained mobile base station (C-MB) problem. Results on C-MB problem will help us devise solution to the general problem where the base station is allowed to roam anywhere on the two-dimensional plane. We term the latter problem unconstrained mobile base station (U-MB) problem.

Denote $(x, y)(t)$ the position of base station B at time t and T the network lifetime (which is an optimization variable). Then a feasible flow routing solution realizing this network lifetime T must satisfy both flow balance and energy constraints at each sensor node. These constraints can be formally stated as follows. Denote $g_{ij}(t)$ and $g_{iB}(t)$ the data rates from node i to node j and base station B at time t, respectively. Then the flow balance for each sensor node $i \in \mathcal{N}$ at any time $t \in [0, T]$ is

$$\sum_{k \in \mathcal{N}}^{k \neq i} g_{ki}(t) + r_i = \sum_{j \in \mathcal{N}}^{j \neq i} g_{ij}(t) + g_{iB}(t) \ ,$$

i.e., the sum of total incoming flow rates plus self-generated data rate is equal to the sum of total outgoing flow rates at time $t \in [0, T]$. The energy constraint for each sensor node $i \in \mathcal{N}$ is

$$\int_0^T \left[\sum_{k \in \mathcal{N}}^{k \neq i} \rho \cdot g_{ki}(t) + \sum_{j \in \mathcal{N}}^{j \neq i} C_{ij} \cdot g_{ij}(t) + c_{iB}(t) \cdot g_{iB}(t) \right] dt \leq e_i \ ,$$

i.e., total consumed energy due to reception and transmission over time T cannot exceed its initial energy e_i. By (2), we have

$$c_{iB}(t) = \alpha + \beta \left[\sqrt{(x(t) - x_i)^2 + (y(t) - y_i)^2} \right]^n \ ,$$

where (x_i, y_i) are the fixed Cartesian coordinates of node i.

Denote $\mathcal{A}$ the search space for the base station, which can be narrowed down to the smallest enclosing disk for all nodes in the network (see Lemma 3 in Section 5). The general U-MB problem can be formulated as follows.

$$\begin{array}{lll}
\textbf{Max} & T & \\
\textbf{s.t.} & \sum_{1 \leq k \leq N}^{k \neq i} g_{ki}(t) + r_i = \sum_{1 \leq j \leq N}^{j \neq i} g_{ij}(t) + g_{iB}(t) & (1 \leq i \leq N, 0 \leq t \leq T) \\
& \int_0^T \left[\sum_{1 \leq k \leq N}^{k \neq i} \rho \cdot g_{ki}(t) + \sum_{1 \leq j \leq N}^{j \neq i} C_{ij} \cdot g_{ij}(t) + c_{iB}(t) \cdot g_{iB}(t) \right] dt \leq e_i & (1 \leq i \leq N) \\
& c_{iB}(t) = \alpha + \beta \left[\sqrt{(x(t) - x_i)^2 + (y(t) - y_i)^2} \right]^n & (1 \leq i \leq N, 0 \leq t \leq T) \\
& (x, y)(t) \in \mathcal{A} & (0 \leq t \leq T) \\
& T, g_{ij}(t), g_{iB}(t) \geq 0 & (1 \leq i, j \leq N, i \neq j, 0 \leq t \leq T)
\end{array}$$

In the above formulation, the base station location (i.e., $(x, y)(t)$ for $0 \leq t \leq T$) and the corresponding flow routing (i.e., $g_{ij}(t)$ and $g_{iB}(t)$ for $0 \leq t \leq T$) form a joint optimization space for the objective T. This formulation is in the form of *non-polynomial programming*. Since even a simpler non-linear programming problem is NP-hard [10], we conclude that the above formulation is NP-hard.

3 From Time Domain to Space Domain

The difficulty of the formulated problem in last section resides in that base station location $(x, y)(t)$ and flow routing $g_{ij}(t)$ and $g_{iB}(t)$ are all functions of time. In this section, we show that as far as network lifetime performance is concerned, such dependency on time can be relaxed. Specifically, we will show (Theorem 1) that the flow routing only needs to be dependent on the location of the base station and independent of when the base station is present at this location. Further, as long as the total time duration for this location is the same, "when" the base station visits this location is not important.

To start with, we define the indicator function $1^+\{\text{event}\}$ as 1 if event is true and 0 otherwise. Denote $W(p)$ the cumulative time periods for the base station B to be present at location p under a solution φ, i.e.,

$$W(p) = \int_0^T 1^+\{(x, y)(t) = p\}dt \ . \tag{4}$$

We have the following theorem.

Theorem 1 *Denote T^* the maximum network lifetime achieved by an optimal solution φ^* with a base station moving path $(x, y)^*(t)$ and a flow routing $g^*_{ij}(t)$ and $g^*_{iB}(t)$. There exists an equivalent solution $\bar{\varphi}^*$ with $W^*(p)$ and time-independent flow rates $f^*_{ij}(p)$ and $f^*_{iB}(p)$ that yields the same maximum network lifetime. Under $f^*_{ij}(p)$ and $f^*_{iB}(p)$, as long as $W^*(p)$ remain the same, the network lifetime T^* will remain unchanged regardless of the ordering and specific time instances when the base station visits each point p.*

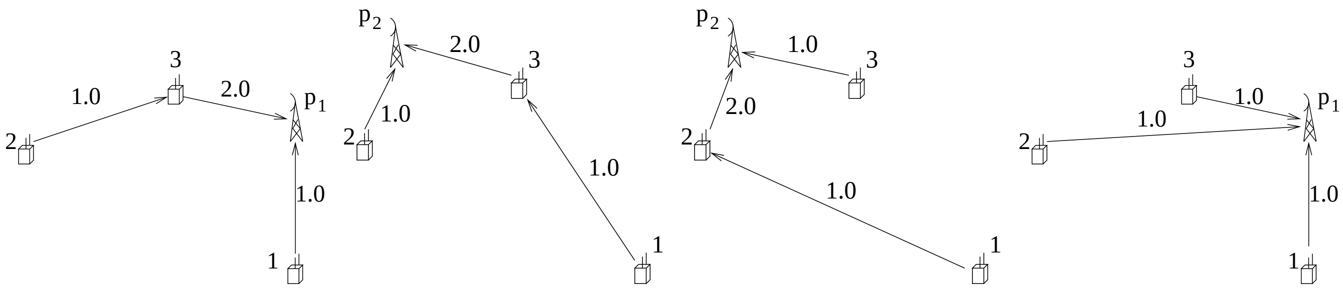
The proof of Theorem 1 is based on the following two lemmas.

Lemma 1 *Given a feasible solution φ with a specific base station moving path $(x, y)(t)$, a time-varying flow routing $g_{ij}(t)$ and $g_{iB}(t)$, and a corresponding network lifetime T, we can find a solution $\bar{\varphi}$ with time-independent flow rates $f_{ij}(p)$ and $f_{iB}(p)$ for each point p in path $(x, y)(t)$ that yield the same network lifetime T by having the base station following the same moving path and setting the flow routing on any point p in this path as*

$$f_{ij}(p) = \frac{\int_0^T g_{ij}(t) \cdot 1^+\{(x, y)(t) = p\}dt}{W(p)} \ ,$$

$$f_{iB}(p) = \frac{\int_0^T g_{iB}(t) \cdot 1^+\{(x, y)(t) = p\}dt}{W(p)} \ .$$

In essence, Lemma 1 focus on flow routing component and shifts the flow routing's dependency from time domain to space domain. We use the following simple example to illustrate this idea.



(a) Flow routing during $[0, 50]$ when the base station is at p_1.

(b) Flow routing during $[50, 90]$ when the base station is at p_2.

(c) Flow routing during $[90, 100]$ when the base station is at p_2.

(d) Flow routing during $[100, 130]$ when the base station is at p_1.

Figure 1: A simple example illustrating time-varying flow routing under φ in Lemma 1.

(a) Flow routing during $[0, 50]$ and $[100, 130]$ when the base station is at p_1.

(b) Flow routing during $[50, 100]$ when the base station is at p_2.

Figure 2: An space-dependent flow routing solution under $\bar{\varphi}$ in Lemma 1.

Example 1 *For illustration purpose, we consider a simple* 3*-node sensor network shown in Fig. 1. Each node generates data at a bit rate of* 1*. We assume all units are normalized appropriately. The network lifetime is* 130 *under a solution* φ*, where the base station stays at* p_1 *during periods* $[0, 50]$ *and* $[100, 130]$*, and stays at* p_2 *during* $[50, 100]$*. The flow routing for each period is given in Fig. 1(a)–(d), with bit rate of each flow marked on the respective flow arrow. Note that the flow routing for periods* $[0, 50]$ *and* $[100, 130]$ *are different, despite that the base station is located at* p_1 *during both periods. Further, during* $[50, 100]$*, the flow routing changes between periods* $[50, 90]$ *and* $[90, 100]$ *while the base station is at* p_2*.*

Under $\bar{\varphi}$*, we show that as far as achieving the same network lifetime, flow routing only needs to be dependent on the base station's location. For example, under* φ*, the total time that the base station stays at* p_1 *is* 80 *(*$= 50 + 30$*). During the time periods when the base station is at* p_1*,* $f_{3B}(t) = 2.0$ *when* $0 \le t \le 50$ *and* $f_{3B}(t) = 1.0$ *when* $100 \le t \le 130$*. Thus, the average rate on link* $(3, B)$ *when the base station stays at* p_1 *is* $(2.0 * 50 + 1.0 * 30)/80 = 1.625$*. We can set* $f_{3B}(p_1) = 1.625$ *under* $\bar{\varphi}$*. Similarly, we obtain the average rate* $f_{23}(p_1) = 0.625$ *and* $f_{2B}(p_1) = 0.375$*, respectively. This lead to a new flow routing shown in Fig. 2(a) for periods* $[0, 50]$ *and* $[100, 130]$*. Following the same token, we can convert the flow routing in Fig. 1(b) and (c) to that in Fig. 2(b), i.e.,* $f_{12}(p_2) = 0.2$*,* $f_{13}(p_2) = 0.8$*,* $f_{2B}(p_2) = 1.2$*, and* $f_{3B}(p_2) = 1.8$*.*

It is easy to verify that flow balance and total data volume transmitted on each link under $\bar{\varphi}$ *remain the same as those under* φ*. Based on (1) and (3), it is easy to verify that the network lifetime under* $\bar{\varphi}$ *also remains the same as that under* φ *(i.e.,* 130*). As a result, we have obtained a solution* $\bar{\varphi}$ *with flow routing only dependent on base station's location.* $\square$

The following proof of Lemma 1 follows the same spirit in the above example.

Proof of Lemma 1. Denote $\mathcal{P}$ the base station moving path $(x, y)(t)$ for $0 \le t \le T$ in solution φ. We let base station follow the same path in solution $\bar{\varphi}$. For the data routing in $\bar{\varphi}$, we can define $f_{ij}(p)$ and $f_{iB}(p)$ for each point $p \in \mathcal{P}$ as follows.

$$f_{ij}(p) = \frac{\int_0^T g_{ij}(t) \cdot 1^+\{(x, y)(t) = p\} dt}{W(p)} \tag{5}$$

$$f_{iB}(p) = \frac{\int_0^T g_{iB}(t) \cdot 1^+\{(x, y)(t) = p\} dt}{W(p)} \tag{6}$$

To show the data routing scheme with $f_{ij}(p)$ and $f_{iB}(p)$ is feasible and $\bar{\varphi}$ has the same network lifetime T, we need to prove that flow balance holds at any point $p \in \mathcal{P}$ and the energy consumption is the same as that in solution φ at time T.

For flow balance when base station is at any point $p \in \mathcal{P}$, we have

$$\begin{aligned}
\sum_{1\le k\le N}^{k\neq i} f_{ki}(p) + r_i &= \sum_{1\le k\le N}^{k\neq i} \frac{\int_0^T g_{ki}(t)\cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} + \frac{\int_0^T r_i \cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} \\
&= \frac{\int_0^T \left[\sum_{1\le k\le N}^{k\neq i} g_{ki}(t) + r_i\right]\cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} \\
&= \frac{\int_0^T \left[\sum_{1\le j\le N}^{j\neq i} g_{ij}(t) + g_{iB}(t)\right]\cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} \\
&= \sum_{1\le j\le N}^{j\neq i} \frac{\int_0^T g_{ij}(t)\cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} + \frac{\int_0^T g_{iB}(t)\cdot 1^+\{(x,y)(t)=p\}dt}{W(p)} \\
&= \sum_{1\le j\le N}^{j\neq i} f_{ij}(p) + f_{iB}(p)
\end{aligned}$$

The first equality holds by Eq. (5). The third equality holds by the flow balance in solution φ. The last equality holds by Eqs. (5) and (6).

For energy consumption at time T, we first have

$$\begin{aligned}
&\int_{\mathcal{P}}\int_0^T c_{iB}(p) f_{iB}(p)\cdot 1^+\{(x,y)(t)=p\}dtdp \\
=\ &\int_{\mathcal{P}} c_{iB}(p) f_{iB}(p)\cdot \int_0^T 1^+\{(x,y)(t)=p\}dtdp \\
=\ &\int_{\mathcal{P}} c_{iB}(p) \frac{\int_0^T g_{iB}(\tau)\cdot 1^+\{(x,y)(\tau)=p\}d\tau}{W(p)}\cdot \int_0^T 1^+\{(x,y)(t)=p\}dtdp \\
=\ &\int_{\mathcal{P}} c_{iB}(p) \int_0^T g_{iB}(\tau)\cdot 1^+\{(x,y)(\tau)=p\}d\tau dp \\
=\ &\int_{\mathcal{P}}\int_0^T c_{iB}(p) g_{iB}(t)\cdot 1^+\{(x,y)(t)=p\}dtdp \\
=\ &\int_{\mathcal{P}}\int_0^T c_{iB}(t) g_{iB}(t)\cdot 1^+\{(x,y)(t)=p\}dtdp \\
=\ &\int_0^T c_{iB}(t) g_{iB}(t) dt
\end{aligned} \tag{7}$$

for $1 \le i \le N$. The second equality holds by Eq. (6). The third equality holds by Eq. (4). Similarly, we have

$$\int_{\mathcal{P}}\int_0^T f_{ij}(p)\cdot 1^+\{(x,y)(t)=p\}dtdp = \int_0^T g_{ij}(t)dt \tag{8}$$

for $1 \le i, j \le N, i \ne j$. Thus, we have

$$\int_{\mathcal{P}} \int_0^T \left[\sum_{1 \le k \le N}^{k \ne i} \rho f_{ki}(p) + \sum_{1 \le j \le N}^{j \ne i} C_{ij} f_{ij}(p) + c_{iB}(p) \cdot f_{iB}(p) \right] \cdot 1^{+}\{(x,y)(t) = p\} dt dp$$

$$= \sum_{1 \le k \le N}^{k \ne i} \rho \int_{\mathcal{P}} \int_0^T f_{ki}(p) \cdot 1^{+}\{(x,y)(t) = p\} dt dp + \sum_{1 \le j \le N}^{j \ne i} C_{ij} \int_{\mathcal{P}} \int_0^T f_{ij}(p) \cdot 1^{+}\{(x,y)(t) = p\} dt dp$$

$$+ \int_{\mathcal{P}} \int_0^T c_{iB}(p) \cdot f_{iB}(p) \cdot 1^{+}\{(x,y)(t) = p\} dt dp$$

$$= \sum_{1 \le k \le N}^{k \ne i} \int_0^T \rho g_{ki}(t) dt + \sum_{1 \le j \le N}^{j \ne i} \int_0^T C_{ij} g_{ij}(t) dt + \int_0^T c_{iB}(t) \cdot f_{iB}(t) dt \le e_i$$

The second equality holds by Eqs. (7) and (8). The last inequality holds due to the energy constraint under φ. Thus, the data routing with $f_{ij}(p)$ and $f_{iB}(p)$ is feasible and has the same network lifetime T. This completes the proof. □

The following lemma further extends Lemma 1 and says that the ordering and specific time instances for the base station to visit a particular point p is not important. As long as $W(p)$ remain the same, the network lifetime remains unchanged. For example, in the transformed solution $\bar{\varphi}$ in Fig. 2, as long as the base station stays at point p_1 for a time period of 80 $(= 50 + 30)$ and point p_2 for a time period of 50, the exact time instances when the base station is present at this location is not important in terms of achieving the same network lifetime.

Lemma 2 *Given a feasible solution φ with a base station moving path $(x, y)(t)$, a flow routing $g_{ij}(t)$ and $g_{iB}(t)$, and a network lifetime T, we can compute $W(p)$ by Eq. (4) and define base station location-dependent flow routing $f_{ij}(p)$ and $f_{iB}(p)$ as in Lemma 1. Under $f_{ij}(p)$ and $f_{iB}(p)$, as long as $W(p)$ remain the same, the network lifetime T will remain unchanged regardless of the ordering and frequency of the base station's presence at each point p.*

Lemma 2 can be easily proved by analyzing the energy consumption behavior at each node over time T. We omit its proof here. Combining Lemmas 1 and 2 and considering the special case that φ is optimal, we have Theorem 1.

Based on Theorem 1, we conclude that as far as network lifetime is concerned, it is sufficient for us to obtain $W(p)$, $f_{ij}(p)$, and $f_{iB}(p)$ values when the base station is at each point p. The specific realization for $(x, y)(t)$ is not important and such realizations are certainly not unique. As a result, we can transform the optimization space from time-dependent functions $(x(t, y)(t), g_{ij}(t), g_{iB}(t))$ to space-dependent functions $(W(p), f_{ij}(p), f_{iB}(p))$. Subsequently, U-MB problem formulation given in Section 2.2 can be reformulated as follows.

Max $\quad T$

s.t. $\quad \int_{\mathcal{A}} W(p) dp = T$

$$\sum_{1 \le k \le N}^{k \ne i} f_{ki}(p) + r_i = \sum_{1 \le j \le N}^{j \ne i} f_{ij}(p) + f_{iB}(p) \qquad (1 \le i \le N, p \in \mathcal{A})$$

$$\int_{\mathcal{A}}\left[\sum_{1\le k\le N}^{k\neq i}\rho\cdot f_{ki}(p)+\sum_{1\le j\le N}^{j\neq i}C_{ij}\cdot f_{ij}(p)+c_{iB}(p)\cdot f_{iB}(p)\right]W(p)dp\le e_i \quad (1\le i\le N) \tag{9}$$

$$T, W(p), f_{ij}(p), f_{iB}(p)\ge 0 \qquad (1\le i,j\le N, i\neq j, p\in\mathcal{A})$$

Note that integration $\int_0^T(\cdot)dt$ (with respect to time) in the original problem formulation (Section 2.2) has been changed to $\int_{\mathcal{A}}(\cdot)dp$ (with respect to space) in the new formulation. This novel transformation will enable us to develop provably approximation algorithm in the space domain, which we will elaborate in Section 5.

4 Optimal Solution to the C-MB Problem

In this section, we show that C-MB problem can be formulated as an LP problem, which can be solved in polynomial time [1]. This result will be useful when we solve the U-MB problem in Section 5.

Recall that in the C-MB problem, the location of base station is limited to a given set of pre-determined locations p_m, $m=1,2,\cdots,M$. Based on the results in Section 3, we need to find optimal time duration $W(p_m)$ and the corresponding flow routing $f_{ij}(p_m)$ and $f_{iB}(p_m)$ when the base station is at each p_m to maximize the network lifetime.[2]

When the base station is at point p_m, $1\le m\le M$, the flow balance for node $i\in\mathcal{N}$ is

$$\sum_{1\le k\le N}^{k\neq i} f_{ki}(p_m)+r_i=\sum_{1\le j\le N}^{j\neq i} f_{ij}(p_m)+f_{iB}(p_m)\ . \tag{10}$$

The energy constraint for node $i\in\mathcal{N}$, at the end of network lifetime T is

$$\sum_{m=1}^{M}\left[\sum_{1\le k\le N}^{k\neq i}\rho\cdot f_{ki}(p_m)+\sum_{1\le j\le N}^{j\neq i}C_{ij}\cdot f_{ij}(p_m)+c_{iB}(p_m)\cdot f_{iB}(p_m)\right]W(p_m)\le e_i\ . \tag{11}$$

Note that for each i and p_m, $c_{iB}(p_m)$ is a constant.

We can formulate C-MB problem as an LP by letting $V_{ij}(p_m)=f_{ij}(p_m)\cdot W(p_m)$ and $V_{iB}(p_m)=f_{iB}(p_m)\cdot W(p_m)$, where $V_{ij}(p_m)$ (or $V_{iB}(p_m)$) can be interpreted as the total data volume from sensor node i to sensor node j (or base station B) when the base station is at p_m. We have

LP(C-MB) **Max** $\quad T$

s.t. $\quad \sum_{m=1}^{M} W(p_m)-T=0$

[2] When $W(p_m)=0$ for some p_m, it means that base station does not visit p_m in this solution.

$$\sum_{1\le k\le N}^{k\neq i} V_{ki}(p_m)+r_i\cdot W(p_m)-\sum_{1\le j\le N}^{j\neq i} V_{ij}(p_m)-V_{iB}(p_m)=0 \qquad (1\le i\le N, 1\le m\le M) \quad (12)$$

$$\sum_{m=1}^{M}\left[\sum_{1\le k\le N}^{k\neq i}\rho\cdot V_{ki}(p_m)+\sum_{1\le j\le N}^{j\neq i} C_{ij}\cdot V_{ij}(p_m)+c_{iB}(p_m)\cdot V_{iB}(p_m)\right]\le e_i \qquad (1\le i\le N) \quad (13)$$

$$T, W(p_m), V_{ij}(p_m), V_{iB}(p_m)\ge 0 \quad (1\le i,j\le N, i\neq j, 1\le m\le M)\ ,$$

where (12) and (13) follow from (10) and (11), respectively. Once we solve the above LP, we can obtain $f_{ij}(p_m)$ and $f_{iB}(p_m)$ by $f_{ij}(p_m)=\frac{V_{ij}(p_m)}{W(p_m)}$ and $f_{iB}(p_m)=\frac{V_{iB}(p_m)}{W(p_m)}$. We summarize the result in this section with the following proposition.

Proposition 1 *C-MB problem can be solved via a single LP in polynomial time.*

The solution to the above LP problem yields the time durations for the base station at each location p_m, $m=1,2,\cdots,M$, and the optimal flow routing when the base station is present at p_m. So far we assume that after base station B completes its stay at p_i, it can move to p_j, $j\neq i$, in zero time. In practice, such travel will take some time. We assume that such travel time is negligible comparing with the time scale of network lifetime (typically several months). It can be shown that if buffering is available at sensor node, the maximum network lifetime can still be achieved. In this case, a node needs to slightly delay its transmission until the base station arrives at its new location.

5 A $(1-\varepsilon)$ Optimal Algorithm to the U-MB Problem

Building upon the results in the previous sections, we are now ready to present the main result of this paper.

5.1 Subareas and Fictitious Cost Points

Motivation. Under the U-MB problem, the base station can move to any point in the two-dimensional plane. Clearly, in order to maximize network lifetime, we would expect that the location for the base station can be narrowed down to the *smallest enclosing disk* (SED) [27], which is a disk with the smallest radius that contains all the sensor nodes in the network. This is formally stated in the following lemma.

Lemma 3 *To maximize the network lifetime for U-MB problem, the base station must stay within the smallest enclosing disk $\mathcal{A}$ that covers all the sensor nodes in the network.*

Proof. We prove this lemma by showing that if the base station location is not always in SED in a solution to U-MB problem, then this solution is not optimal, i.e., the network lifetime can be increased. Assume at time t, the

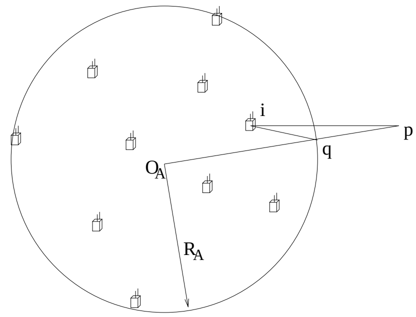


Figure 3: The area of possible base station location.

base station is at location p, which is outside SED $\mathcal{A}$. Denote q the cross point of the border of SED and the segment $[p, O_{\mathcal{A}}]$, where $O_{\mathcal{A}}$ is the center of SED (see Fig. 3). We can revise the location by letting the base station at q. We now analyze the effect on energy consumption after this revision. For any sensor i, since $\angle iqp > \pi/2$, it is clear that $D_{iq} < D_{ip}$. Based on Eqs. (1) and (2), we can save some energy on any node i by this revision. Thus, the network lifetime can be increased. In another word, the base station location must be always in SED in an optimal solution to U-MB problem. □

It has been shown in [16] that SED is unique and can be found in $O(N)$ time, where N is the number of nodes to be covered.

Although we have narrowed down the search space for base station B from an infinite two-dimensional plane to SED $\mathcal{A}$, there still exist infinite number of points in $\mathcal{A}$. To obtain a finite search space, we consider dividing $\mathcal{A}$ into small subareas, $\mathcal{A}_1, \mathcal{A}_2, \cdots$, up to say $\mathcal{A}_M$, i.e.,

$$\mathcal{A} = \bigcup_{m=1}^{M} \mathcal{A}_m \ .$$

For approximation, it is tempting to use a point $q_m \in \mathcal{A}_m$, $m = 1, 2, \cdots, M$, to represent subarea $\mathcal{A}_m$. Indeed, when each subarea is sufficiently small, we may use some point $q_m \in \mathcal{A}_m$ to represent $\mathcal{A}_m$, $m = 1, 2, \cdots, M$, and apply the LP approach in Section 4 on these M points to get a very good solution.

However, such approach is *heuristic* at best and unfortunately does not provide any *theoretical guarantee* on performance. A theoretical question is how the M subareas for $\mathcal{A}$ should be divided such that an algorithm can be developed that yields *provably* $(1-\varepsilon)$ optimal network lifetime performance, i.e., a network lifetime that is guaranteed to be at least $(1-\varepsilon)$ of the optimum. Note that a naive approach such as breaking up $\mathcal{A}$ into uniform square subareas will not provide any properties that can be used to prove network lifetime performance. Further, there does not appear to be any reason why each subarea within $\mathcal{A}$ should be divided of the same shape or size.

Our Approach. Our approach is to examine the energy constraint in (9) and exploit how the location of the base station affect the energy consumption. Note that the location of the base station is *embedded* in the cost parameter

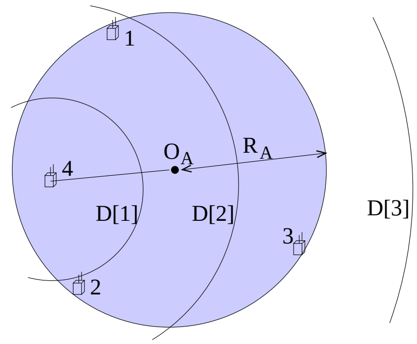


Figure 4: A sequence of circles with increasing costs with center at node 4.

c_{iB}. Thus, to design a $(1-\varepsilon)$ optimal algorithm, we may consider dividing disk $\mathcal{A}$ into subareas, with each subarea to be associated with some nice properties on c_{iB}'s that can be used to prove $(1-\varepsilon)$ optimality.

Denote $O_{\mathcal{A}}$ and radius $R_{\mathcal{A}}$ the origin and radius of the SED $\mathcal{A}$. For each sensor node $i \in \mathcal{N}$, denote $D_{i,O_{\mathcal{A}}}$ the distance from sensor node i to the origin of disk $\mathcal{A}$. Denote $D_{iB}^{\min}$ and $D_{iB}^{\max}$ the minimum and maximum distance between sensor node i and base station B, respectively; denote $C_{iB}^{\min}$ and $C_{iB}^{\max}$ the corresponding minimum and maximum cost between sensor node i and base station B, respectively. Then, since the search space for base station B is now narrowed down to disk $\mathcal{A}$ (see Fig. 4), we have

$$\begin{aligned} D_{iB}^{\min} &= 0 \ , \\ D_{iB}^{\max} &= D_{i,O_{\mathcal{A}}} + R_{\mathcal{A}} \ . \end{aligned}$$

By Eq. (2), we have

$$C_{iB}^{\min} = \alpha \ , \tag{14}$$

$$C_{iB}^{\max} = \alpha + \beta \cdot (D_{iB}^{\max})^n = \alpha + \beta \cdot (D_{i,O_{\mathcal{A}}} + R_{\mathcal{A}})^n \ . \tag{15}$$

Given the range of $d_{iB} \in [D_{iB}^{\min}, D_{iB}^{\max}]$, for each sensor node $i \in \mathcal{N}$, we now show how to divide disk $\mathcal{A}$ into a set of *non-uniform* subareas with the distance of each subarea to sensor node i meeting some properties that can be used to design $(1-\varepsilon)$ optimal algorithm.

Specifically, for each sensor node $i \in \mathcal{N}$, we draw a sequence of circles centered at sensor node i, each with increasing radius $D[1], D[2], \cdots, D[H_i]$ corresponding to costs $C[1], C[2], \cdots, C[H_i]$ that are defined as follows.

$$C[h] = C_{iB}^{\min}(1+\varepsilon)^h = \alpha(1+\varepsilon)^h \quad (1 \le h \le H_i) \tag{16}$$

The number of required circles H_i can be determined by having the last circle in the sequence (with radius $D[H_i]$) to completely include disk $\mathcal{A}$, i.e. $D[H_i] > D_{iB}^{\max}$, or equivalently,

$$C[H_i] \ge C_{iB}^{\max} \ .$$

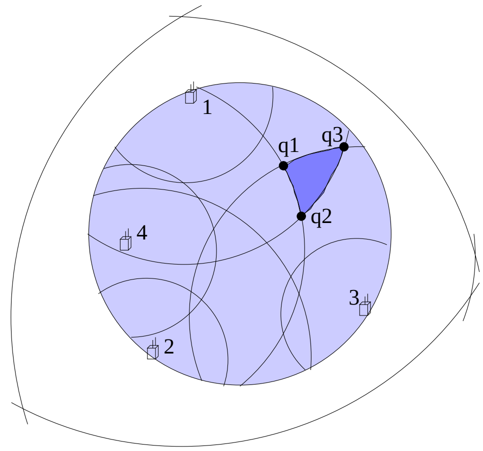


Figure 5: An example of subareas within disk $\mathcal{A}$ that are obtained by intersecting arcs from different circles.

We have

$$H_i = \left\lceil \frac{\ln(C_{iB}^{\max}/C_{iB}^{\min})}{\ln(1+\varepsilon)} \right\rceil = \left\lceil \frac{\ln(1 + \frac{\beta}{\alpha}(D_{i,O_\mathcal{A}} + R_\mathcal{A})^n)}{\ln(1+\varepsilon)} \right\rceil = O\left(\frac{1}{\varepsilon}\right) \ . \tag{17}$$

That is, we have a total of H_i circles with center at sensor node i, each with cost $C[h]$, $h = 1, 2, \cdots, H_i$. These H_i circles provide H_i non-overlapping rings. Now suppose base station B is moved to any point in the h-th ring, $h = 1, 2, \cdots, H_i$. Then we have

$$C[h-1] \leq c_{iB} \leq C[h] \ , \tag{18}$$

where we define $C[0] = C_{iB}^{\min} = \alpha$.

We perform the above process for each sensor node $i \in \mathcal{N}$. The intersecting circles will divide disk $\mathcal{A}$ into a number of non-uniform subareas, with the boundaries of each subarea being either an arc centered at some sensor node $i \in \mathcal{N}$ (with some cost $C[h]$, $1 \leq h < H_i$) or an arc of disk $\mathcal{A}$. As an example, disk $\mathcal{A}$ in Fig. 5 is now divided into 27 subareas.

So what nice properties do we have about dividing $\mathcal{A}$ into these non-uniform subareas? We will show that for a point in each of these subareas, its cost to each sensor node can be *tightly* bounded from both above and below. As a result, this can be exploited to design a $(1-\varepsilon)$ optimal algorithm. Note that for each sensor node $i \in \mathcal{N}$, any subarea $\mathcal{A}_m$ must be within a ring with center at sensor node i. Denote the index of this ring as $h_i(\mathcal{A}_m)$. That is, when the base station B is at any point $p \in \mathcal{A}_m$, we have

$$C[h_i(\mathcal{A}_m) - 1] \leq c_{iB}(p) \leq C[h_i(\mathcal{A}_m)] \tag{19}$$

by (18). Since $\frac{C[h_i(\mathcal{A}_m)]}{C[h_i(\mathcal{A}_m)-1]} = 1 + \varepsilon$ by (16), these two bounds for $c_{iB}(p)$ are very tight.

To use the result in Section 4, we introduce the notion of *fictitious cost point* p_m to represent each subarea $\mathcal{A}_m$, $m = 1, 2, \cdots, M$. The fictitious cost point is an N-tuple vector embodying the upper cost bound for any point within this subarea $\mathcal{A}_m$ to all the N sensor nodes in the network. Specifically, denote the N-tuple cost vector for fictitious cost point p_m (corresponding to subarea $\mathcal{A}_m$) as $[c_{1B}(p_m), c_{2B}(p_m), \cdots, c_{NB}(p_m)]$, with the i-th component $c_{iB}(p_m)$ being

$$c_{iB}(p_m) = C[h_i(\mathcal{A}_m)] \ , \tag{20}$$

where $h_i(\mathcal{A}_m)$ is determined by Eq. (19).

As an example, the fictitious cost point for subarea with corner points (q_1, q_2, q_3) in Fig. 5 can be represented by 4-tuple cost vector $[c_{1B}(p_m), c_{2B}(p_m), c_{3B}(p_m), c_{4B}(p_m)] = [C[2], C[3], C[2], C[3]]$, where the first component $C[2]$ represents an upper bound of cost for any point in this subarea to sensor node 1, the second component $C[3]$ represents an upper bound of cost (which is loose here) for any point in this subarea to sensor node 2, and so forth.

In our design, we use the word "fictitious" to suggest that points p_m, $m = 1, 2, \cdots, M$, may only be used as a bound for the purpose of developing $(1 - \varepsilon)$ optimal algorithm. In reality, p_m may not be mapped to any *physical* point within subarea $\mathcal{A}_m$. This happens when there does not exist a *physical* point in subarea $\mathcal{A}_m$ that can have its costs to all the N sensor nodes equal (one by one) to the respective N-tuple cost vector embodied by p_m *simultaneously*. As an example, comparing corner point q_1 in Fig. 5 and the 4-tuple cost vector of p_m, we have $c_{1B}(q_1) < C[2]$, $c_{2B}(q_1) < C[3]$, $c_{3B}(q_1) = C[2]$, and $c_{4B}(q_1) < C[3]$, i.e., $[c_{1B}(q_1), c_{2B}(q_1), c_{3B}(q_1), c_{4B}(q_1)] \neq [c_{1B}(p_m), c_{2B}(p_m), c_{3B}(p_m), c_{4B}(p_m)]$. So p_m cannot be mapped to q_1. Following the same token, we can verify that p_m cannot be mapped to q_2 or q_3 either. Further, any physical point within this subarea cannot have its costs to the four sensor nodes simultaneously match to each of the respectively component of the 4-tuple embodied by p_m. We have the following property for fictitious cost point p_m.

Property 1 *Denote p_m the fictitious cost point for subarea $\mathcal{A}_m$, $m = 1, 2, \cdots, M$. For any point $p \in \mathcal{A}_m$, we have $c_{iB}(p) \le c_{iB}(p_m) \le (1 + \varepsilon) \cdot c_{iB}(p)$.*

Proof. By (19) and definition of fictitious cost point p_m (see (20)), we have $c_{iB}(p) \le c_{iB}(p_m)$. Further, we have

$$c_{iB}(p_m) = C[h_i(\mathcal{A}_m)] = (1 + \varepsilon) \cdot C[h_i(\mathcal{A}_m) - 1] \le (1 + \varepsilon) \cdot c_{iB}(p) \ ,$$

where the last inequality follows from (19). This completes the proof. □

Now the set of M non-uniform subareas are represented by the M fictitious cost points, with each fictitious cost point having an N-tuple cost vector to all the N sensor nodes in the network. Although a fictitious cost point may not be mapped to any physical point, using these fictitious cost points will aid to design a $(1 - \varepsilon)$ optimal algorithm. Note that for network lifetime problems, we only need to consider the cost terms c_{iB} for $i = 1, 2, \cdots, N$, which is exactly captured by the N-tuple representation for each fictitious cost point. As a result, we can now readily apply

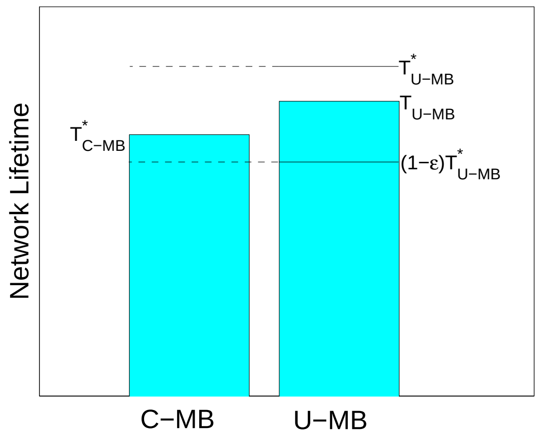


Figure 6: Comparison of network lifetimes under different solutions used to construct $(1-\varepsilon)$ optimal solution.

the LP approach discussed in Section 4 to formulate an optimization problem on these M fictitious cost points. In next section, we will show that an optimal solution to the C-MB problem on fictitious cost points can be used to construct a $(1-\varepsilon)$ optimal solution to the U-MB problem.

5.2 $(1-\varepsilon)$ Optimality

Denote $\psi^*_{\text{U-MB}}$ an optimal solution to the U-MB problem and $T^*_{\text{U-MB}}$ the corresponding maximum network lifetime, both of which are unknown. Our objective is to find a solution to the U-MB problem that has provably $(1-\varepsilon)$ optimal network lifetime. Denote $\psi^*_{\text{C-MB}}$ an optimal solution to the C-MB problem obtained by applying an LP on the M fictitious cost points p_m, $m = 1, 2, \cdots, M$, and $T^*_{\text{C-MB}}$ the corresponding network lifetime.

Our roadmap for theoretical proofs is as follows. In Theorem 2, we will prove that $T^*_{\text{C-MB}} \geq (1-\varepsilon)T^*_{\text{U-MB}}$ (see Fig. 6). Since the optimal solution $\psi^*_{\text{C-MB}}$ corresponding to $T^*_{\text{C-MB}}$ are based on the M fictitious cost points instead of physical points, in Theorem 3 we will further show how to construct a solution $\psi_{\text{U-MB}}$ to U-MB problem based on $\psi^*_{\text{C-MB}}$ and prove that the corresponding network lifetime is $(1-\varepsilon)$ optimal, i.e., $T_{\text{U-MB}} \geq (1-\varepsilon)T^*_{\text{U-MB}}$ (see Fig. 6).

Theorem 2 *For a given $\varepsilon > 0$, define subareas $\mathcal{A}_m$ and fictitious cost points p_m, $m = 1, 2, \cdots, M$, as in Section 5.1. Then we have $T^*_{\text{C-MB}} \geq (1-\varepsilon) \cdot T^*_{\text{U-MB}}$.*

To prove that Theorem 2 is true, we need the following lemma, where subareas $\mathcal{A}_m$ and fictitious cost points p_m, $m = 1, 2, \cdots, M$ are defined the same as in Theorem 2 for a given $\varepsilon > 0$. Denote $W(\mathcal{A}_m)$ the cumulative time period when the base station B is present within subarea $\mathcal{A}_m$ for the U-MB problem. We have

$$W(\mathcal{A}_m) = \int_{\mathcal{A}_m} W(p)dp \ . \tag{21}$$

Lemma 4 *Suppose we have a given solution $\varphi_{\text{U-MB}}$ to the U-MB problem with $W(p)$, $f_{ij}(p)$, $f_{iB}(p)$, and a network lifetime $T_{\text{U-MB}}$. For a given $\varepsilon > 0$, we can always construct a solution $\varphi_{\text{C-MB}}$ to C-MB problem on these fictitious cost points such that network lifetime $T_{\text{C-MB}} \geq (1-\varepsilon) \cdot T_{\text{U-MB}}$ by having the base station spend $W(p_m)$ amount of time on fictitious cost point p_m, where*

$$W(p_m) = (1-\varepsilon) \cdot W(\mathcal{A}_m) \tag{22}$$

and setting the flow routing on p_m as

$$f_{ij}(p_m) = \frac{\int_{\mathcal{A}_m} f_{ij}(p)W(p)dp}{W(\mathcal{A}_m)}, \tag{23}$$

$$f_{iB}(p_m) = \frac{\int_{\mathcal{A}_m} f_{iB}(p)W(p)dp}{W(\mathcal{A}_m)}. \tag{24}$$

Lemma 4 is a powerful result. It states that for any given solution $\varphi_{\text{U-MB}}$ to the U-MB problem, we can find a solution $\varphi_{\text{C-MB}}$ for the set of fictitious cost points (corresponding to a given ε), such that the network lifetime $T_{\text{C-MB}}$ is at least $(1-\varepsilon)$ of $T_{\text{U-MB}}$. The proof of Lemma 4 is given in Appendix.

Now we are ready to prove Theorem 2.

Proof to Theorem 2. Considering the special case of Lemma 4 that the given solution to U-MB problem is an optimal solution with network lifetime $T^*_{\text{U-MB}}$, we can transform it into a solution to C-MB problem on fictitious cost points with network lifetime at least $(1-\varepsilon)T^*_{\text{U-MB}}$, i.e., there is a solution to C-MB problem on fictitious cost points with network lifetime of at least $(1-\varepsilon)T^*_{\text{U-MB}}$. Thus, the solution $\psi^*_{\text{C-MB}}$ to C-MB problem on fictitious cost points must have a network lifetime $T^*_{\text{C-MB}} \geq (1-\varepsilon)T^*_{\text{U-MB}}$. □

Theorem 2 guarantees that the network lifetime obtained from the LP solution on the M fictitious cost points is at least $(1-\varepsilon)$ of $T^*_{\text{U-MB}}$. As discussed in Section 5.1, the N-tuple vector embodied by a fictitious cost point only represents the upper cost bound for any point in the corresponding subarea to the N nodes in the network and a fictitious cost point may not be mapped to a physical point, which is required in the final solution. In the following theorem, we show that if we have an optimal solution to the C-MB problem based on fictitious cost points, we can construct a solution with each point being physically realizable. Further, the network lifetime for this constructed solution is greater than or equal to the maximum network lifetime for the C-MB problem, i.e., $T_{\text{U-MB}} \geq T^*_{\text{C-MB}}$. As a result, this new solution is $(1-\varepsilon)$ optimal.

Theorem 3 *For a given $\varepsilon > 0$, define subareas $\mathcal{A}_m$ and fictitious cost points p_m, $m = 1, 2, \cdots, M$, as discussed in Section 5.1. Given an optimal solution $\psi^*_{\text{C-MB}}$ on these M fictitious cost points with $W^*(p_m)$, $f^*_{ij}(p_m)$, $f^*_{iB}(p_m)$, and corresponding network lifetime $T^*_{\text{C-MB}}$, a $(1-\varepsilon)$ optimal solution $\psi_{\text{U-MB}}$ to U-MB problem can be constructed by having the base station stay in $\mathcal{A}_m$ for*

$$W(\mathcal{A}_m) = W^*(p_m) \tag{25}$$

amount of time with a corresponding flow routing for any point $p \in \mathcal{A}_m$ as

$$
\begin{aligned}
f_{ij}(p) &= f^*_{ij}(p_m) , \qquad (26)\\
f_{iB}(p) &= f^*_{iB}(p_m) . \qquad (27)
\end{aligned}
$$

In Theorem 3, note that in the constructed solution to U-MB problem, when the base station is at *any* point $p \in \mathcal{A}_m$, the flow routing is the same. In other words, the flow routing only depends on the subarea instead of specific point within this subarea.

Proof. In this proof, we will show that $\psi_{\text{U-MB}}$ is feasible, i.e., flow balance holds at any point, and the network lifetime of $\psi_{\text{U-MB}}$ is at least $T^*_{\text{C-MB}}$. Once we proved this, since $T^*_{\text{C-MB}} \geq (1-\varepsilon)T^*_{\text{U-MB}}$ (see Theorem 2), we know that $T_{\text{U-MB}}$ is at least $(1-\varepsilon)T^*_{\text{U-MB}}$, i.e., $\psi_{\text{U-MB}}$ is $(1-\varepsilon)$ optimal.

For flow balance when the base station location p is in any subarea $\mathcal{A}_m$, we have

$$
\sum_{1\leq k\leq N}^{k\neq i} f_{ki}(p) + r_i = \sum_{1\leq k\leq N}^{k\neq i} f^*_{ki}(p_m) + r_i = \sum_{1\leq j\leq N}^{j\neq i} f^*_{ij}(p_m) + f^*_{iB}(p_m) = \sum_{1\leq j\leq N}^{j\neq i} f_{ij}(p) + f_{iB}(p)
$$

The first equality holds by Eq. (26). The second equality holds by the flow balance in solution $\psi^*_{\text{C-MB}}$. The third equality holds by Eqs. (26) and (27). Thus, solution $\psi_{\text{U-MB}}$ is feasible.

For the total consumed energy on node i at time $\sum_{m=1}^{M} W(\mathcal{A}_m) = \sum_{m=1}^{M} W^*(p_m) = T^*_{\text{C-MB}}$, we first have

$$
\begin{aligned}
\int_{\mathcal{A}} c_{iB}(p) f_{ij}(p) W(p)dp &= \sum_{m=1}^{M} \int_{\mathcal{A}_m} c_{iB}(p) f_{ij}(p) W(p)dp \\
&\leq \sum_{m=1}^{M} \int_{\mathcal{A}_m} c_{iB}(p_m) f^*_{ij}(p_m) W(p)dp \\
&= \sum_{m=1}^{M} c_{iB}(p_m) f^*_{ij}(p_m) W(\mathcal{A}_m) \\
&= \sum_{m=1}^{M} c_{iB}(p_m) f^*_{ij}(p_m) W^*(p_m) \qquad (28)
\end{aligned}
$$

for $1 \leq i \leq N$. The second inequality holds Eq. (27) and $c_{iB}(p) \leq c_{iB}(p_m)$ in Property 1. The third equality holds by Eq. (21). The last equality holds by Eq. (25). Similarly, we have

$$
\int_{\mathcal{A}} f_{ij}(p) W(p)dp = \sum_{m=1}^{M} f^*_{ij}(p_m) W^*(p_m) \qquad (29)
$$

for $1 \le i, j \le N$ and $i \ne j$. Thus, we have

$$\begin{aligned}
&\int_{\mathcal{A}} \left[\sum_{1\le k\le N}^{k\ne i} \rho f_{ki}(p)W(p) + \sum_{1\le j\le N}^{j\ne i} C_{ij} f_{ij}(p)W(p) + c_{iB}(p) f_{iB}(p)W(p) \right] dp \\
= &\sum_{1\le k\le N}^{k\ne i} \rho \int_{\mathcal{A}} f_{ki}(p)W(p)dp + \sum_{1\le j\le N}^{j\ne i} C_{ij} \int_{\mathcal{A}} f_{ij}(p)W(p)dp + \int_{\mathcal{A}} c_{iB}(p) f_{iB}(p)W(p)dp \\
\le &\sum_{1\le k\le N}^{k\ne i} \rho \sum_{m=1}^{M} f^*_{ki}(p_m)W^*(p_m) + \sum_{1\le j\le N}^{j\ne i} C_{ij} \sum_{m=1}^{M} f^*_{ij}(p_m)W^*(p_m) + \sum_{m=1}^{M} c_{iB}(p_m) f^*_{iB}(p_m)W^*(p_m) \\
\le &e_i \ .
\end{aligned}$$

The first equality holds by Eqs. (28) and (29). The second equality holds by the energy constraint in solution $\psi^*_{\text{C-MB}}$. Thus, the network lifetime of solution $\psi_{\text{U-MB}}$ is at least $T^*_{\text{C-MB}} \ge (1-\varepsilon)T^*_{\text{U-MB}}$. This completes the proof. □

5.3 Summary of Algorithm and Example

The design of the $(1-\varepsilon)$ optimal algorithm are given in Sections 5.1 and 5.2. We now summarize it into the following algorithm.

Algorithm 1 (A $(1-\varepsilon)$ Optimal Algorithm)

1. Narrow down the search space for base station within smallest enclosing disk $\mathcal{A}$.
2. Within the smallest enclosing disk $\mathcal{A}$, compute the lower and upper cost bounds $C^{\min}_{iB}$ and $C^{\max}_{iB}$ for each node $i \in \mathcal{N}$ by (14) and (15).
3. For a given $\varepsilon > 0$, define a sequence of costs $C[1], C[2], \cdots, C[H_i]$ by (16), where H_i is defined by (17).
4. For each node $i \in \mathcal{N}$, draw a sequence of $(H_i - 1)$ circles corresponding to cost $C[h]$ centered at node i, $1 \le h < H_i$. The intersection of these circles within disk $\mathcal{A}$ will divide $\mathcal{A}$ into M subareas $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_M$.
5. For each subarea $\mathcal{A}_m$, $1 \le m \le M$, define a fictitious cost point p_m, which is represented by N-tuple cost vector $[c_{1B}(p_m), c_{2B}(p_m), \cdots, c_{NB}(p_m)]$, where $c_{iB}(p_m)$ is defined in (20).
6. For the C-MB problem on these M fictitious cost points, apply the LP formulation in Section 4 and obtain an optimal solution $\psi^*_{\text{C-MB}}$ with $W^*(p_m)$, $f^*_{ij}(p_m)$, and $f^*_{iB}(p_m)$.
7. Construct a $(1-\varepsilon)$ optimal solution $\psi_{\text{U-MB}}$ to U-MB problem based on $\psi^*_{\text{C-MB}}$ using the procedure in Theorem 3.

Table 2: Sensor locations, data rate, and initial energy of the example sensor network

Node Index	(x_i, y_i)	r_i	e_i
1	$(0.1, 0.5)$	0.8	390
2	$(1.1, 0.7)$	1.0	400
3	$(0.4, 0.1)$	0.5	130

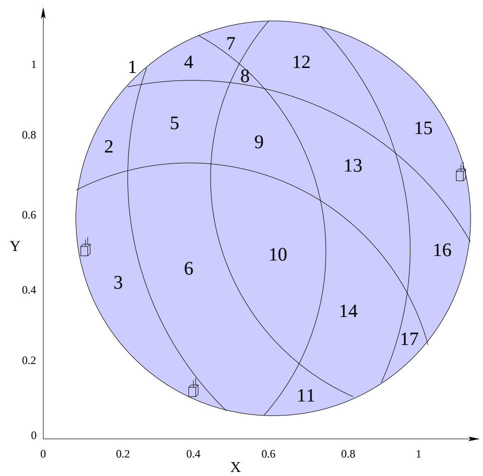


Figure 7: The subareas for the example sensor network.

In the above algorithm, Step 6 has the highest complexity (solving an LP). Since there are $(H_i - 1)$ circles radiating from sensor node $i \in \mathcal{N}$ and one circle for disk $\mathcal{A}$, the total number of subareas M through such intersection of circles is upper bounded by $O([1 + \sum_{i=1}^{N}(H_i - 1)]^2) = O((N/\varepsilon)^2)$. Thus, the LP in Step 6 has polynomial size and the complexity of the above algorithm is polynomial.

Example 2 *To illustrate the steps in Algorithm 1, we solve a small* 3*-sensor network problem as an example. The location, data rate, and initial energy for each sensor are shown in Table 2, where the units of distance, rate, and energy are all normalized. We use* $n = 2$ *in this example and the network setting are* $\alpha = 1$, $\beta = 0.5$ *and* $\rho = 1$ *under normalized units. For illustration, we set* $\varepsilon = 0.2$.[3]

In Step 1, we identify SED $\mathcal{A}$ *with origin* $O_{\mathcal{A}} = (0.61, 0.57)$ *and radius* $R_{\mathcal{A}} = 0.51$ *(see Fig. 7).*

In Step 2, we first identify SED $\mathcal{A}$ *with origin* $O_{\mathcal{A}} = (0.61, 0.57)$ *and radius* $R_{\mathcal{A}} = 0.51$ *(see Fig. 7). Then we have* $D_{i,O_{\mathcal{A}}} = R_{\mathcal{A}} = 0.51$ *for each node* i, $1 \le i \le 3$*. We then find the lower and upper bounds of* c_{iB} *for each node* i *as follows.*

$$C_{iB}^{min} = \alpha = 1$$

[3]In Section 5.4, we use $\varepsilon = 0.05$ for all numerical results.

$$C_{iB}^{max} = \alpha + \beta(D_{i,O_{\mathcal{A}}} + R_{\mathcal{A}})^n = 1 + 0.5 \cdot (0.51 + 0.51)^2 = 1.52$$

In Step 3, for $\varepsilon = 0.2$*, we find*

$$H_i = \left\lceil \frac{\ln(1 + \frac{\beta}{\alpha}(D_{i,O_{\mathcal{A}}} + R_{\mathcal{A}})^n)}{\ln(1+\varepsilon)} \right\rceil = \left\lceil \frac{\ln(1 + \frac{0.5}{1}(0.51 + 0.51)^2)}{\ln(1+0.2)} \right\rceil = 3$$

for each node i, $1 \leq i \leq 3$*, and*

$$\begin{aligned} C[1] &= \alpha(1+\varepsilon) = 1 \cdot (1+0.2) = 1.20 \ , \\ C[2] &= \alpha(1+\varepsilon)^2 = 1 \cdot (1+0.2)^2 = 1.44 \ , \\ C[3] &= \alpha(1+\varepsilon)^3 = 1 \cdot (1+0.2)^3 = 1.73 \ . \end{aligned}$$

In Step 4, we draw circles with centered at each node i, $1 \leq i \leq 3$*, and with cost* $C[h]$, $1 \leq h < H_i = 3$*, to divide the whole disk* $\mathcal{A}$ *into* 17 *subareas* $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_{17}$.

In Step 5, we define a fictitious cost point p_m *for each subarea* $\mathcal{A}_m$, $1 \leq m \leq 17$*. For example, for fictitious cost point* p_1*, we define the* 3*-tuple cost vector as* $[c_{1B}(p_m), c_{2B}(p_m), c_{3B}(p_m)] = [C[1], C[3], C[3]] = [1.20, 1.73, 1.73]$.

In Step 6, we obtain an optimal solution $\psi^*_{\text{C-MB}}$ *to C-MB problem on these* 17 *fictitious cost points by the LP approach discussed in Section 4. We obtain the network lifetime* $T^*_{\text{C-MB}} = 190.37$, $W^*(p_3) = 157.00$, $W^*(p_6) = 33.37$*, and for all other* 15 *fictitious cost points, we have* $W^*(p_m) = 0$ *(i.e., base station will not visit those areas). When the base station is at fictitious cost point* p_3*, the routing is* $f^*_{1B}(p_3) = 1.4$, $f^*_{2B}(p_3) = 1.0$*, and* $f^*_{31}(p_3) = 0.6$*. When the base station is at fictitious cost point* p_6*, the routing is* $f^*_{1B}(p_6) = 0.8$, $f^*_{2B}(p_6) = 1.0$*, and* $f^*_{3B}(p_6) = 0.6$.

In Step 7, we obtain a $(1-\varepsilon)$ *optimal solution* $\psi_{\text{U-MB}}$ *to U-MB problem as follows. Let the base station stay at any point in subarea* $\mathcal{A}_3$ *for* 157.00 *time and stay at any point in subarea* $\mathcal{A}_6$ *for* 33.37 *time. When the base station is at a point* p *in subarea* $\mathcal{A}_3$*, the routing is* $f_{1B}(p) = 1.4$, $f_{2B}(p) = 1.0$*, and* $f_{31}(p) = 0.6$*. When the base station is at a point* p *in subarea* $\mathcal{A}_6$*, the routing is* $f_{1B}(p) = 0.8$, $f_{2B}(p) = 1.0$*, and* $f_{3B}(p) = 0.6$*. The network lifetime for* $\psi_{\text{U-MB}}$ *is greater than or equal to* 190.37 *and is* $(1-\varepsilon)$ *optimal.* □

5.4 Numerical Results

Now we apply the $(1-\varepsilon)$ optimal algorithm for larger sized networks and use numerical results to demonstrate the efficacy of the algorithm. We consider three randomly generated networks consisting of 10, 20, 50, and 100 nodes deployed over a 1×1 square, respectively. The data rate and initial energy for each node are randomly generated between $[0.1, 1]$ and $[50, 500]$, respectively. The units of distance, rate, and energy are all normalized appropriately. The normalized parameters in energy consumption model are $\alpha = \beta = \rho = 1$. We assume the path loss index $n = 2$.

Table 3: Each node's Cartesian coordinates, data generation rate and initial energy for a small 10-node network.

(x_i, y_i)	r_i	e_i	(x_i, y_i)	r_i	e_i
$(0.0, 0.8)$	0.8	150	$(0.6, 0.7)$	0.6	370
$(1.0, 1.0)$	1.0	200	$(0.4, 0.6)$	0.2	420
$(0.3, 0.4)$	0.6	130	$(0.2, 0.9)$	0.6	100
$(0.8, 0.5)$	0.8	460	$(0.9, 0.1)$	0.4	80
$(0.7, 0.3)$	0.3	170	$(0.5, 0.2)$	1.0	150

(a) Locations to be visited by base station B in the $(1-\varepsilon)$ optimal solution

(b) Flow routing when the base station is at particular location.

Figure 8: Results for the small 10-node network.

The required accuracy for approximation algorithm ε is set to $\varepsilon = 0.05$ for all numerical results. That is, we are pursuing a solution with a network lifetime that is at least 95% of the maximum network lifetime.

The network setting (location, data rate, and initial energy for each node) for the 10-node network is given in Table 3. By applying Algorithm 1, we obtain a $(1-\varepsilon)$ optimal network lifetime 142.86, which is guaranteed to be at least 95% of the optimum. In Table 4, we have seven subareas that will be visited by the base station in the $(1-\varepsilon)$ optimal solution. For illustration purpose, we use a point (x, y) within a subarea $\mathcal{A}_m$ to represent the approximate location of this subarea. For example, we use the point $(0.93, 0.96)$ to represent the subarea that contains this point. Table 4 lists the corresponding time duration for the base station to stay in each of these 7 subareas. The flow routing solution when the base station is in each of the 7 subareas is different as expected. For illustration, in Fig. 8(b), we show the flow routing solution when the base station is in subarea containing the point $(0.93, 0.96)$, where a circle represents a sensor node and a star represents a subarea for base station.

It is worth noting that for 95% of accuracy in optimality, there are only 7 subareas for the base station to visit. It turns out that for 20, 50, and 100 node networks, the number of subareas that needs to be visited by the base station

Table 4: $(1-\varepsilon)$ optimal results for the small 10-node network with $\varepsilon = 0.05$.

$\mathcal{A}_m(x,y)$	$W(\mathcal{A}_m)$
$(0.93, 0.96)$	4.28
$(0.88, 0.19)$	3.54
$(0.68, 0.90)$	0.04
$(0.94, 0.89)$	50.36
$(0.69, 0.22)$	44.00
$(0.90, 0.40)$	27.15
$(0.23, 0.86)$	13.49

Table 5: Each node's Cartesian coordinates, data generation rate and initial energy for a 20-node network.

(x_i, y_i)	r_i	e_i	(x_i, y_i)	r_i	e_i
$(0.52, 0.02)$	0.6	480	$(0.29, 0.14)$	0.6	120
$(0.74, 0.76)$	0.3	310	$(0.05, 0.99)$	0.4	60
$(0.95, 0.03)$	0.8	150	$(0.84, 0.06)$	1.0	180
$(0.53, 0.63)$	0.6	220	$(0.99, 0.37)$	0.4	340
$(0.58, 1.00)$	0.4	230	$(0.73, 0.67)$	0.8	220
$(0.48, 0.84)$	0.7	160	$(0.53, 0.27)$	0.5	380
$(0.17, 0.83)$	0.1	380	$(0.57, 0.05)$	0.7	250
$(0.73, 0.39)$	0.1	500	$(0.88, 0.84)$	0.2	240
$(0.36, 0.98)$	0.1	430	$(0.26, 0.12)$	0.9	440
$(0.76, 0.02)$	0.7	500	$(0.71, 0.21)$	0.3	70

Table 6: $(1-\varepsilon)$ optimal results for the small 20-node network with $\varepsilon = 0.05$.

$\mathcal{A}_m(x,y)$	$W(\mathcal{A}_m)$
$(0.28, 0.33)$	2.86
$(0.27, 0.47)$	8.44
$(0.95, 0.86)$	9.62
$(0.80, 0.06)$	5.34
$(0.70, 0.11)$	108.05
$(0.88, 0.05)$	9.92

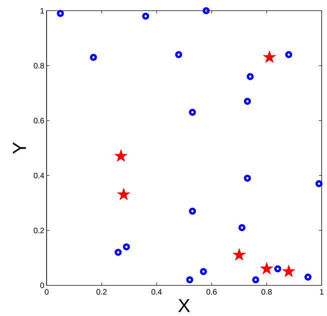


Figure 9: The small 20-node network.

Table 7: $(1-\varepsilon)$ optimal results for the medium 50-node network with $\varepsilon = 0.05$.

$\mathcal{A}_m(x,y)$	$W(\mathcal{A}_m)$
$(0.13, 0.11)$	0.39
$(0.17, 0.64)$	1.47
$(0.32, 0.90)$	26.23
$(0.25, 0.61)$	27.72
$(0.49, 0.12)$	3.43
$(0.94, 0.10)$	9.66
$(0.34, 0.26)$	8.37
$(0.16, 0.30)$	45.03

is still very small (7 subareas for 20-node network, 8 subareas for 50-node network, and 12 subareas for 100-node network). This new observation is not obvious. But it is a good news as it hints that the base station does not need to be involved in frequent movement to achieve near-optimal solution.

The network setting for a small 20-node network (with location, data rate, and initial energy for each of the 20 sensor nodes) is given in Table 5. By applying Algorithm 1, we obtain the $(1-\varepsilon)$ optimal network lifetime 144.23, which is guaranteed to be at least within 95% of the optimal. Again, we use a point (x, y) within the chosen subarea $\mathcal{A}^*$ to represent the approximate location of this subarea. For this particular 20-node network setting, we have 6 subarea in the final solution, with the corresponding time duration in each subarea listed in Table 6.

For the 50-node network, the positions of the nodes are shown in Fig. 10, where a circle represents a sensor node and a star represents an optimal subarea for base station. We omit to list the coordinates of each node due to paper length limitation. The data rate and initial energy for each node are randomly generated between $[0.1, 1]$ and $[50, 500]$, respectively. By applying Algorithm 1, we obtain a $(1-\varepsilon)$ optimal network lifetime 122.30. Again, we use a point (x, y) within a subarea $\mathcal{A}_m$ to represent the approximate location of this subarea. For this particular 50-node network setting, we have 8 subareas that the base station should visit in the final solution, with the corresponding

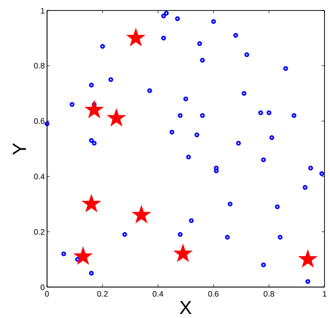


Figure 10: A 50-node network used in numerical investigation.

Table 8: $(1-\varepsilon)$ optimal results for the 100-node network with $\varepsilon = 0.05$.

$\mathcal{A}_m(x, y)$	$W(\mathcal{A}_m)$
$(0.75, 0.83)$	0.15
$(0.86, 0.05)$	0.38
$(0.69, 0.28)$	50.14
$(0.24, 0.04)$	2.10
$(0.59, 0.88)$	0.57
$(0.81, 0.82)$	0.09
$(0.20, 0.53)$	21.21
$(0.91, 0.09)$	0.05
$(0.40, 0.68)$	41.64
$(0.89, 0.19)$	3.16
$(0.61, 0.79)$	23.82
$(0.19, 0.02)$	6.14

time duration in each subarea listed in Table 7.

Finally, we consider a 100-node network shown in Fig. 11. Again, we omit to list the coordinates of each node due to paper length limitation. The data rate and initial energy for each node are again randomly generated between $[0.1, 1]$ and $[50, 500]$, respectively. By applying Algorithm 1, we have a $(1-\varepsilon)$ optimal network lifetime 149.45. For this particular 100-node network setting, we have 12 subareas that the base station should visit in the solution, with the corresponding time duration in each subarea listed in Table 8.

6 Related Work

Energy efficient routing has been an active area of research for sensor network in recent years (see, e.g., [19, 22, 24, 26]). It is now well understood that energy efficient routing differs from lifetime-optimal routing as the former advocates the use of minimum energy-cost path, which may overload nodes along some common shared path,

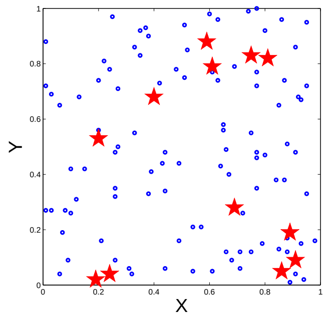


Figure 11: A 100-node network used in numerical investigation.

leading to poor performance in network lifetime.

Routing algorithms to maximize network lifetime has been an active area of research even for fixed base station location (see, e.g., [3, 4, 5, 20] and references therein). The focus is mainly devoted to how to split traffic flow along different routes and how to adjust power level at each node so that some optimal flow routing topology can be set up to maximize network lifetime. These early work have laid foundation on the importance of power control and flow routing topology on network lifetime performance.

There are some recent work on optimal base station placement [8, 17]. The focus on these efforts is on finding an optimal *fixed* position for the base station so that network lifetime can be maximized. However, as pointed out in [15, 28], network lifetime can be substantially increased if the optimization space can be expanded to include movement of the base station during the course of sensor network operation.

Relevant work in the area of mobile base station for network lifetime problems include [2, 9, 15, 25, 28]. In [2, 9, 25], the locations of base station are constrained on a set of "pre-determined" locations. In [28], Younis et al. show that mobile base station can increase network lifetime. In [15], Luo and Hubaux propose to minimize the maximum load on a node among all the nodes in the network, which can be considered as an equivalent problem to maximize network lifetime. The results in [15, 28] are *heuristic*, and thus do not provide any theoretical bound on network lifetime performance.

Note that the mobile base station problem considered in this paper differs fundamentally from the delay-tolerant network (DTN) (see, e.g., [13, 29]). DTN is assumed to experience frequent and long duration partition. The focus is to leverage storage at the intermediate nodes over long period of time and perform intermittent routing "over time" (i.e., delay tolerant) so as to achieve "eventual delivery". Network lifetime is not a major performance objective in the context of DTN.

7 Conclusions

The benefits of employing mobile base station to prolong sensor network lifetime are significant. However, due to the complexity of the problem (time-dependent network topology and traffic routing), provably optimal theoretical results have remained an open research area before this paper. This paper fills in this important gap by contributing a provably optimal algorithm regarding the location of a mobile base station. The foundation of both result hinges upon a novel time-to-space transformation for problem formulation. Based on this transformation, we present an intermediate result which says that when the location of the base station are constrained to be on a set of pre-determined points, then the optimal solution can be obtained via a single LP. Building upon this intermediate result, our main result addressed the general mobile base station problem where the location of the base station is un-constrained. We developed a provably $(1-\varepsilon)$ optimal algorithm for the mobile base station location problem such that the network lifetime is guaranteed to be at least $(1-\varepsilon)$ of the maximum network lifetime, where $\varepsilon > 0$ can be made arbitrarily small depending on required precision.

Appendix – Proof of Lemma 4

We will show solution $\varphi_{\text{C-MB}}$ is feasible, i.e., flow balance holds at each point, and the network lifetime of $\varphi_{\text{C-MB}}$ is at least $(1-\varepsilon)\cdot T_{\text{U-MB}}$.

For flow balance when the base station location is p_m, we have

$$\begin{aligned}
\sum_{1\le k\le N}^{k\neq i} f_{ki}(p_m) + r_i &= \sum_{1\le k\le N}^{k\neq i} \frac{\int_{\mathcal{A}_m} f_{ki}(p)W(p)dp}{W(\mathcal{A}_m)} + \frac{\int_{\mathcal{A}_m} r_i W(p)dp}{W(\mathcal{A}_m)} \\
&= \frac{\int_{\mathcal{A}_m}\left[\sum_{1\le k\le N}^{k\neq i} f_{ki}(p) + r_i\right] W(p)dp}{W(\mathcal{A}_m)} \\
&= \frac{\int_{\mathcal{A}_m}\left[\sum_{1\le j\le N}^{j\neq i} f_{ij}(p) + f_{iB}(p)\right] W(p)dp}{W(\mathcal{A}_m)} \\
&= \sum_{1\le j\le N}^{j\neq i} \frac{\int_{\mathcal{A}_m} f_{ij}(p)dp}{W(\mathcal{A}_m)} + \frac{\int_{\mathcal{A}_m} f_{iB}(p)dp}{W(\mathcal{A}_m)} \\
&= \sum_{1\le j\le N}^{j\neq i} f_{ij}(p_m) + f_{iB}(p_m)\ .
\end{aligned}$$

The first equality holds by Eq. (23). The third equality holds by the flow balance in solution $\varphi_{\text{U-MB}}$. The last equality holds by Eqs. (23) and (24). Thus, solution $\varphi_{\text{C-MB}}$ is feasible.

For the total consumed energy on node i at time $\sum_{m=1}^{M} W(p_m) = \sum_{m=1}^{M}(1-\varepsilon)\cdot W(\mathcal{A}_m) = (1-\varepsilon)\cdot T_{\text{U-MB}}$, we first have

$$\begin{aligned}
\sum_{m=1}^{M} c_{iB}(p_m) f_{iB}(p_m) W(p_m) &= \sum_{m=1}^{M} c_{iB}(p_m) \frac{\int_{\mathcal{A}_m} f_{iB}(p)dp}{W(\mathcal{A}_m)} W(p_m) \\
&= (1-\varepsilon) \sum_{m=1}^{M} c_{iB}(p_m) \int_{\mathcal{A}_m} f_{iB}(p)dp \\
&\leq (1-\varepsilon) \sum_{m=1}^{M} (1+\varepsilon) \int_{\mathcal{A}_m} c_{iB}(p) f_{iB}(p)dp \\
&< \sum_{m=1}^{M} \int_{\mathcal{A}_m} c_{iB}(p) f_{iB}(p)dp \\
&= \int_{\mathcal{A}} c_{iB}(p) f_{iB}(p)dp
\end{aligned} \tag{30}$$

for $1 \leq i \leq N$. The first equality holds by Eq. (24). The second equality holds by Eq. (22). The third inequality holds by $c_{iB}(p_m) \leq (1+\varepsilon)c_{iB}(p)$ in Property 1. Similarly, we have

$$\sum_{m=1}^{M} f_{ij}(p_m) W(p_m) < \int_{\mathcal{A}} f_{ij}(p)dp \tag{31}$$

for $1 \leq i, j \leq N$ and $i \neq j$. Thus, we have

$$\begin{aligned}
&\sum_{m=1}^{M} \left[\sum_{1\leq k\leq N}^{k\neq i} \rho f_{ki}(p_m) W(p_m) + \sum_{1\leq j\leq N}^{j\neq i} C_{ij} f_{ij}(p_m) W(p_m) + c_{iB}(p_m) f_{iB}(p_m) W(p_m) \right] \\
= \; & \sum_{1\leq k\leq N}^{k\neq i} \rho \sum_{m=1}^{M} f_{ki}(p_m) W(p_m) + \sum_{1\leq j\leq N}^{j\neq i} C_{ij} \sum_{m=1}^{M} f_{ij}(p_m) W(p_m) + \sum_{m=1}^{M} c_{iB}(p_m) f_{iB}(p_m) W(p_m) \\
< \; & \sum_{1\leq k\leq N}^{k\neq i} \rho \int_{\mathcal{A}} f_{ki}(p)dp + \sum_{1\leq j\leq N}^{j\neq i} C_{ij} \int_{\mathcal{A}} f_{ij}(p)dp + \int_{\mathcal{A}} c_{iB}(p) f_{iB}(p)dp \\
= \; & \int_{\mathcal{A}} \left[\sum_{1\leq k\leq N}^{k\neq i} \rho f_{ki}(p) + \sum_{1\leq j\leq N}^{j\neq i} C_{ij} f_{ij}(p) + c_{iB}(p) f_{iB}(p) \right] dp \leq e_i
\end{aligned}$$

The first equality holds by Eqs. (30) and (31). The last inequality holds by the energy constraint in solution $\varphi_{\text{U-MB}}$. Thus, the network lifetime of solution $\bar{\varphi}_{\text{C-MB}}$ is at least $(1-\varepsilon)\cdot T_{\text{U-MB}}$. This completes the proof.

References

[1] M.S. Bazaraa, J.J. Jarvis, and H.D. Sherali, *Linear Programming and Network Flows,* second edition, John Wiley & Sons, 1990.

[2] S. Basagni, A. Carosi, E. Melachrinoudis, C. Petrioli, and Z.M. Wang, "A New MILP Formulation and Distributed Protocols for Wireless Sensor Networks Lifetime Maximization," in *Proc. IEEE International Conference on Communications,* pp. 3517–3524, Istanbul, Turkey, June 11–15, 2006.

[3] T.X. Brown, H.N. Gabow, and Q. Zhang, "Maximum flow-life curve for a wireless ad hoc network," in *Proc. ACM MobiHoc,* pp. 128–136, Long Beach, CA, Oct. 4–5, 2001.

[4] J.-H. Chang and L. Tassiulas, "Routing for maximum system lifetime in wireless ad-hoc networks," in *Proc. 37th Annual Allerton Conference on Communications, Control, and Computing,* pp. 22–31, Monticello, IL, Sep. 22–24, 1999.

[5] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE Infocom,* pp. 22–31, Tel Aviv, Israel, March 26–30, 2000.

[6] DARPAR Grand Challenge 2005, URL: `http://www.darpa.mil/grandchallenge/index.html.`

[7] P. Desnoyers, D. Ganesan, and P. Shenoy, "TSAR: A two tier sensor storage architecture using interval skip graphs," in *Proc. ACM SenSys,* pp. 39–50, San Diego, CA, Nov. 2–4, 2005.

[8] A. Efrat, S. Har-Peled, and J. Mitchell, "Approximation algorithms for location problems in sensor networks," in *Proc. IEEE BROADNETS 2005 — Wireless Networking Symposium,* pp. 767–776, Boston, MA, Oct. 3–7, 2005.

[9] S.R. Gandham, M. Dawande, R. Prakash, S. Venkatesan, "Energy efficient schemes for wireless sensor networks with multiple mobile base stations," in *Proc. IEEE Globecom,* pp. 377–381, San Francisco, CA, Dec. 1–5, 2003.

[10] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness,* W. H. Freeman and Company, New York, NY, 1979.

[11] W. Heinzelman, *Application-specific Protocol Architectures for Wireless Networks,* Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, June 2000.

[12] Y.T. Hou, Y. Shi, and H.D. Sherali, "Rate allocation in wireless sensor networks with network lifetime requirement," in *Proc. ACM MobiHoc,* pp. 67–77, Tokyo, Japan, May 24–26, 2004.

[13] S. Jain, K. Fall, R. Patra, "Routing in a delay tolerant network," in *Proc. ACM Sigcomm,* pp. 145–158, Portland, OR, Aug. 30–Sep. 3, 2004.

[14] J. Liu, F. Zhao, P. Cheung, L. Guibas, "Apply geometric duality to energy-efficient non-local phenomenon awareness using sensor networks," *IEEE Wireless Communications,* vol. 11, issue 6, pp. 62–68, Dec. 2004.

[15] J. Luo and J.-P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proc. IEEE Infocom,* pp. 1735–1746, Miami, FL, March 13–17, 2005.

[16] N. Megiddo, "Linear-time algorithm for linear programming in R^3 and related problems," *SIAM J. Computing,* vol. 12, pp. 759–776, 1983.

[17] J. Pan, Y.T. Hou, L. Cai, Y. Shi, and S.X. Shen, "Topology control for wireless sensor networks," in *Proc. ACM Mobicom,* pp. 286–299, San Diego, CA, Sep. 14–19, 2003.

[18] T.S. Rappaport, *Wireless Communications: Principles and Practice,* Prentice Hall, New Jersey, 1996.

[19] V. Rodoplu and T.H. Meng, “Minimum energy mobile wireless networks,” *IEEE Journal on Selected Areas in Communications,* vol. 17, no. 8, pp. 1333–1344, Aug. 1999.

[20] A. Sankar and Z. Liu, “Maximum lifetime routing in wireless ad-hoc networks,” in *Proc. IEEE Infocom,* pp. 1089–1097, Hong Kong, China, March 7–11, 2004.

[21] G.T. Sibley, M.H. Rahimi, G.S. Sukhatme, “Robomote: A tiny mobile robot platform for large-scale ad-hoc sensor networks,” in *Proc. IEEE International Conference on Robotics and Automation,* pp. 1143–1148, Washington, DC, May 11–15, 2002.

[22] S. Singh, M. Woo, and C.S. Raghavendra, “Power-aware routing in mobile ad hoc networks,” in *Proc. ACM/IEEE Mobicom,* pp. 181–190, Dallas, TX, Oct. 25–30, 1998.

[23] K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, “Protocols for self-organizing of a wireless sensor network,” *IEEE Personal Communications Magazine,* vol. 7, pp. 16–27, Oct. 2000.

[24] I. Stojmenovic and X. Lin, “Power-aware localized routing in wireless networks,” *IEEE Trans. on Parallel and Distributed Systems,* vol. 12, no. 11, pp. 1122–1133, Nov. 2001.

[25] Z. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli, “Exploiting sink mobility for maximizing sensor networks lifetime,” in *Proc. IEEE Hawaii International Conference on System Sciences,* Waikoloa, HA, Jan. 3–6, 2005.

[26] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang, “Distributed topology control for power efficient operation in multihop wireless ad hoc networks,” in *Proc. IEEE Infocom,* pp. 1388–1397, Anchorage, AK, April 22–26, 2001.

[27] E. Welzl, “Smallest enclosing disks,” *Lecture Notes in Computer Science (LNCS),* vol. 555, pp. 359–370, 1991.

[28] M. Younis, M. Bangad, and K. Akkaya, “Base-station repositioning for optimized performance of sensor networks,” in *Proc. IEEE Vehicular Technology Conference,* pp. 2956–2960, Orlando, FL, Oct. 4–9, 2003.

[29] W. Zhao, M. Ammar, and E. Zegura, “Controlling the mobility of multiple data transport ferries in a delay-tolerant network,” in *Proc. IEEE Infocom,* pp. 1407–1418, Miami, FL, March 13–17, 2005.