國立臺灣大學資訊管理研究所碩士論文

指導教授：林永松　博士

無線感應器網路系統生存時間最大化之高效能路由演算法

An Energy-efficient Routing Algorithm for the Maximization of System Lifetime in Wireless Sensor Networks

研究生：葉耿宏　撰

中華民國九十三年七月

無線感應器網路系統生存時間最大化之高效能路由演算法

An Energy-efficient Routing Algorithm for the Maximization of System Lifetime in Wireless Sensor Networks

本論文係提交國立台灣大學

資訊管理學研究所作為完成碩士

學位所需條件之一部份

研究生：葉耿宏　撰

中華民國九十三年七月

謝詞

研究所兩年的求學生涯轉眼間告一段落，當中受到了許多人的指導與鼓勵。最感謝的是指導教授林永松老師，從論文題目的構想啟發、方向的擬定到論文的撰寫，老師無不給予最大的協助。在整個論文撰寫的過程中，老師嚴謹的研究態度、學富五車的專業知識、待人處世的翩翩風度，無一不是我極力效法學習的典範。此外，在論文口試的過程中，承蒙本所莊裕澤所長、蔡益坤老師提供了許多寶貴的意見與指正，予我受益良多，也使得本篇論文能夠更加完善，在此要對各位老師致上最誠摯的謝意。

感謝博士班旭成學長及佩玲學姊在我研究的過程中不斷地給予建議和幫助，讓我能夠更順利的完成此篇論文。感謝與我一起努力的同學們，瑩珍、俊達、育先、明立、大鈞、翔騰、坤威、閔元和榮耀在論文撰寫期間相互的切磋與勉勵，因為有你們的陪伴，辛苦的論文之路走來並不孤單。同時我也要感謝好友筱琪、能瑩、凱貞、雅茹、鵬元和柏琪，謝謝你們不時的為我加油打氣給我力量，倘若沒有你們，本篇論文也是無法如此順利完成的。感謝所有幫助過我，使我成長的人。

最後，謹將本篇論文獻給我最敬愛的父母及家人，衷心感謝父親葉碩先生及母親陳鳳仙女士多年浩瀚的教養之恩，在我求學的生涯過程中，不斷給予我鼓勵與支持，使我得以在今日順利完成碩士學位論文，在此與我的父母共同分享這份喜悅。

葉耿宏 謹識
于台大資訊管理研究所
民國九十三年七月

論文摘要

論文題目：無線感應器網路系統生存時間最大化之高效能路由演算法

作　　者：葉耿宏　　　　　　　　　　　　　　　　九十三年七月

指導教授：林永松　博士

無線感應器網路是近年來相當熱門的研究主題。由於感應器技術的進步，促成了無線感應器網路快速蓬勃的發展。無線感應器網路可以被廣泛的使用在許多不同領域的應用上，例如：衛生醫療、軍事國防、環境偵測等等。雖然無線感應器網路能夠提供許多有價值的應用，但同時，爲了能夠有效達成這些應用，也出現了很多的問題和挑戰需要去解決。其中最大的問題莫過於如何在能源資源極有限的無線感應器網路中，以有效率的方式去使用系統能源以延長整個系統的生存時間。

本論文針對此一問題所採行之解決方法是先利用數學規劃的方式將問題數學模式化為一數學最佳化問題，其目標函數為最大化系統之整體生存時間。另外，為了使所求得的問題解能夠更符合現實環境的狀況，我們進一步改進了原先系統生存時間的定義，把系統涵蓋率的條件考慮進來，並同時考量部分應用對於資訊即時性的需求。透過該改進過後的系統生存時間定義，我們可以求得更佳的路由演算法和整體系統生存時間。

由於該問題的本質為一非線性混合整數問題，具有相當的複雜度和困難度。爲了解決此一複雜的問題，本論文採用以拉格蘭日鬆弛法為基礎的方法來處理，因為該法優越的特性，使得我們不但能夠非常有效率的求得該問題解，同時也解決了系統路由的問題，得到一高效能(energy-efficient)

的路由演算法。

關鍵詞：無線感應器網路、高效能、路由演算法、數學規劃、最佳化、拉格蘭日鬆弛法

THESIS ABSTRACT

GRADUATE INSTITUTE OF INFORMATION MANAGEMENT

NATIONAL TAIWAN UNIVERSITY

NAME：KENG-HUNG YEH **MONTH/YEAR：JULY/2004**

ADVISER：DR. YEONG-SUNG LIN

AN ENERGY-EFFICIENT ROUTING ALGORITHM FOR

THE MAXIMIZATION OF SYSTEM LIFETIME IN

WIRELESS SENSOR NETWORKS

The wireless sensor network has become a popular research topic in recent years. Advances in sensor node technology have enabled the rapid development of wireless sensor networks that can be used in various application areas, such as healthcare, the military, and the environment. Although there are many invaluable applications for wireless sensor networks, there are also a lot of emerging problems and challenges that need to be solved, at the same time. The biggest problem is how to efficiently use energy resources to prolong the overall system lifetime of such highly energy-constrained wireless sensor networks.

Our solution to this problem is to design an energy-efficient routing algorithm. We use a mathematical programming technique to formulate the issue as a combinatorial optimization problem, where the objective function is to maximize the system lifetime. To make it more realistic, we modify the definition of the system lifetime by considering the coverage constraint and time-critical demand of some applications. We can then derive a better routing algorithm to obtain a maximal system lifetime of a sensor network that is much closer to the real environment.

Because the optimization problem itself is highly complicated and difficult, we use Lagrangean Relaxation method to solve it. Due to the method's remarkable properties, we are able to solve this complicated optimization problem efficiently, and obtain an energy-efficient routing algorithm at the same time.

Keywords: Wireless Sensor Network, Energy-efficient, Routing Algorithm, Mathematical Programming, Optimization, Lagrangean Relaxation.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

The wireless sensor network has become a popular research topic in recent years. Advances in sensor node technology, such as wireless communications and electronics, have enabled the development of extremely small, low cost, and low powered sensing devices that are equipped with programmable computing, multiple parameter sensing and wireless communication capabilities.

A wireless sensor network is composed of a large number of wireless sensor nodes that are densely scattered in a sensor field as shown in Figure 1-1. Each of the scattered sensor nodes has the capability to collect and route data back to the Sink via a multi-hop architecture. The Sink can communicate with the task manager node via the Internet or satellite, while users can also make inquiries and retrieve data of interest to them by this mechanism.

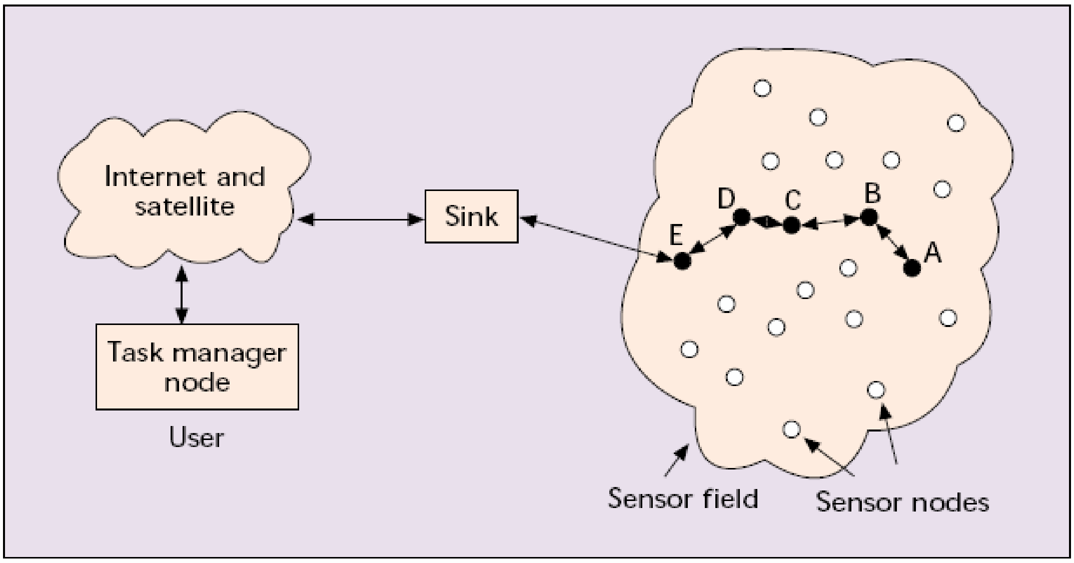

Figure 1-1 A Wireless Sensor Network's Communication Architecture [1]

The tiny wireless sensor nodes, which are the most significant elements of a wireless sensor network, have four basic components, namely: a sensing unit, a processing unit a communication unit (transceiver) and a power unit, as shown in Figure 1-2.

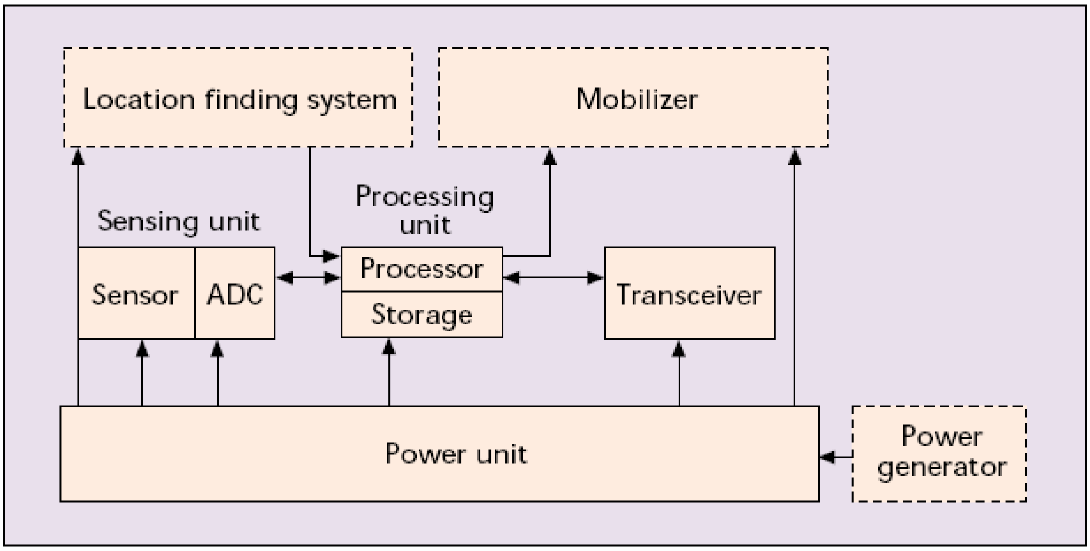

Figure 1-2 Components of A Wireless Sensor Node [1]

The sensing unit is usually composed of two subunits: sensors and analog-to-digital converters (ADCs). The analog signals produced by the sensors, based on the observed phenomenon or event, are converted into digital signals by the ADC. The signals are then fed into a processing unit, which is generally associated with a small storage unit that manages the sensor node as it collaborates with other nodes to carry out the assigned sensing tasks. A communication unit connects the node to the network. One of the most important components of a sensor node is the power unit which provides energy to other components. There are also other application-dependent components, such as a location finding system, a mobilizer and a power generator. Most of the wireless sensor network routing techniques and sensing tasks require extremely accurate information about the location of each sensor node. Thus, it is common that a sensor node has a location finding system. A mobilizer may sometimes be needed to move a sensor node when it is required to carry out assigned tasks. A power generator is used to support power units such as solar cells.

The protocol stack of the Sink and the sensor nodes is shown in Figure1-3. The protocol stack combines power and routing awareness, integrates data with networking protocols, communicates power efficiently through the wireless medium, and promotes the cooperative efforts of sensor nodes. The protocol stack consists of: a physical layer, data link layer, network layer, transport layer, application layer, power management plane, mobility management plane, and task management plane.

The physical layer addresses the needs of simple but robust modulation, transmission, and receiving techniques. Since the environment is noisy and sensor nodes can be mobile, the medium access control (MAC) protocol must be power-aware and able to minimize collision with neighbors' broadcasts. The network layer takes care of routing the data supplied by the transport layer. The transport layer helps to maintain the flow of data if the sensor networks application requires it. Depending on the sensing tasks, different types of application software can be built and used in the application layer. In addition, the power, mobility, and task management planes monitor the power, movement, and task distribution among the sensor nodes. These planes enable the sensor nodes to coordinate the sensing task and lower overall power consumption.

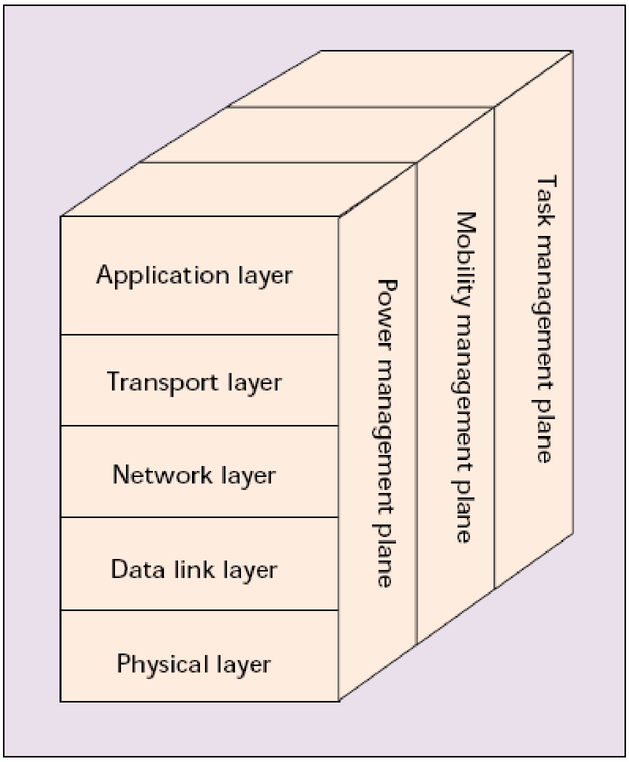

Figure 1-3 Wireless Sensor Network's Protocol Stack [1]

The power management plane manages a sensor node's use of its power. For example, to avoid getting duplicated messages, a sensor node may turn off its receiver after receiving a message from one of its neighbors. Also, when the power level of the sensor node is low, the sensor node broadcasts the situation to its neighbors that it cannot participate in routing messages. The remaining power is reserved for sensing. The mobility management plane detects and registers the movement of sensor nodes, so a route back to the user is always maintained, and the sensor nodes can keep track of who their neighbor sensor nodes are. By knowing who their neighbors are, the sensor nodes can balance their power and task usage. The task management plane balances and schedules the sensing tasks given to a specific region. Not all sensor nodes in that region are required to perform the sensing task at the same time. As a result, some sensor nodes perform the task more than others, depending on their power level. These management planes are needed so that sensor nodes can work together in a power-efficient way to route data in a mobile sensor network, and share resources between sensor nodes.

Wireless sensor networks have a number of advantages over wired networks, such as ease of deployment (reducing installation costs), extended range (a network of tiny sensor nodes can be distributed over a wider region), fault-tolerance (the failure of one node does not affect the network operation), self-organization (the nodes can have the capability to reconfigure themselves), mobility (since these wireless sensor nodes are equipped with batteries, they can be mobile).

The above advantages ensure a wide range of applications for wireless sensor networks. In the military, for example, the rapid deployment, self-organization, and fault tolerance characteristics of wireless sensor networks make them a very promising sensing technique for military command, control, communications, computing, intelligence, surveillance, reconnaissance, and targeting systems. Meanwhile, in healthcare, sensor nodes can also be deployed to monitor patients and assist disabled patients, and in industry, sensor nodes can be used for factory instrumentation. In a large metropolis, sensor nodes can be deployed to monitor traffic density and road conditions. In engineering, sensor nodes can be used to monitor building structures. In the environment, sensor nodes can be used to monitor changes in forests and oceans, and can also make agricultural techniques more precise. Some other commercial applications include managing inventories, monitoring product quality, and assisting rescue efforts in disasters.

Realizing these and other wireless sensor network applications, requires ad hoc wireless networking techniques. Though many protocols and algorithms have been proposed, most are not well suited to the unique features and application requirements of wireless sensor networks. We illustrate the distinctions between wireless sensor networks and ad hoc networks as follows.

- The number of sensor nodes in a sensor network can be many times higher than the nodes in an ad hoc network.
- Sensor nodes are densely deployed and prone to failure.
- The topology of a sensor network changes very frequently.
- Sensor nodes mainly use a broadcast communication paradigm, whereas most ad hoc networks are based on point-to-point communications.

- Sensor nodes are limited in power, computational capacities, and memory.
- Sensor nodes may not have global identification (ID) because of the large amount of overhead and the large number of sensors.
- Adjacent nodes may have similar data. Therefore, rather than sending data separately from each node to the requesting node, it is desirable to aggregate similar data and send it.

Many researchers are actively engaged in developing schemes that fulfill these requirements to enable a variety of applications. This is also the goal of this paper.

1.2 Motivation

The advances in sensor node technology have enabled the rapid development of wireless sensor networks. As described in the previous section, there are many invaluable applications for wireless sensor networks. At the same time, there are also a lot of emerging problems and challenges which need to be solved. The biggest problem is how to efficiently use energy resources in this kind of highly energy-constrained wireless sensor networks.

Energy optimization in wireless sensor networks is very complex because it not only involves reducing the energy consumption of a single sensor node, but also requires maximizing the system lifetime of an entire network. Many papers have already been published on this problem. Some of them focus on an individual sensor node to develop new technology to decrease the energy consumption of each of its operations [13][14][16]. Others, design an energy-efficient routing protocol from the perspective of an entire network [3][15][18], and still others calculate the system lifetime by given environmental conditions and system parameters [2][4][17].

Although there are many energy-efficient routing algorithms, formulating the problem as a combinatorial optimization problem, where the objective function is to maximize system lifetime, is relatively rare. This motivates us to apply Lagrangean Relaxation, combined with the subgradient method, to solve this optimization problem. To the best of our knowledge, no previous research has adopted this approach. Due to the remarkable properties of Largrangean Relaxation, we are able to solve this complicated optimization problem efficiently, and obtain an energy-efficient routing algorithm.

In addition, system lifetime in most of the above papers is defined as the time interval from the point that a sensor network starts its operation until the point the first sensor node fails. This kind of definition is not rational in practice and can be improved.

To make the definition more realistic, we modify it by considering the coverage constraint and time-critical demand for some applications of wireless sensor networks. By using this modified definition, we can derive a better routing algorithm to obtain the maximal system lifetime of wireless sensor networks, which is much closer to the real environment.

1.3 Literature Survey

1.3.1 Routing in Wireless Sensor Networks

Wireless sensor networks can be classified on the basis of their mode of operation and functionality, and the type of target applications. Accordingly, wireless sensor networks are classified into two types:

- **Proactive networks:** The sensor nodes in this network periodically sense the environment and transmit the data of interest. Thus, they provide a snapshot of the relevant parameters at regular intervals. They are well-suited to applications that require periodic data monitoring.
- **Reactive networks:** In this scheme the sensor nodes react immediately to sudden and drastic changes in the value of a sensed attribute. As such, they are well- suited to time critical applications.

Because wireless sensor networks are highly application dependent, once the type of network is decided, an adaptive routing protocol has to be designed. And based on the topology of wireless sensor networks, there are two alternative approaches have been considered: flat routing and hierarchical routing. We discuss them separately below.

1.3.1.1 Flat Routing

In [10], the authors describe the directed diffusion paradigm for designing wireless sensor networks. A wireless sensor network is data-centric, and its application to query dissemination and processing has been demonstrated as follows. The query is disseminated (flooded) throughout the network and then the gradients are set up. The gradients indicate the ‘goodness’ of the different possible next hops and are used to forward sensor node data to users. They are used to direct data satisfying the query toward the requesting node. Data starts flowing toward the requesting nodes from multiple nodes from multiple paths. A small number of paths can be reinforced to prevent further flooding. This type of information retrieval is only well-suited to persistent queries, where requesting nodes are expecting data that satisfies a query for an interval of time. This makes it unsuitable for historical or one-time queries as it is not worth setting up gradients for queries that employ the path only once. Also, this type of data collection doesn’t fully exploit the feature of wireless sensor networks that adjacent nodes have similar data.

In [8], the authors present a family of adaptive protocols called SPIN (Sensor Protocol for Information via Negotiation) that efficiently disseminates information among sensor nodes in an energy-constrained wireless sensor network. Nodes running a SPIN communication protocol name their data using high-level data descriptors, called meta-data. They use meta-data negotiations to eliminate the transmission of redundant data throughout the network. SPIN enables a user to query any node and get the required information immediately. These protocols make use of the property that nearby nodes have similar data and thus distribute only the data that other nodes don’t have. These protocols work proactively and distribute the information all over the network, even when a user does not request any data.

In [14], the authors propose a practical guideline that advocates a uniform resource utilization to prolong system lifetime. The authors also propose a number of practical gradient-based routing algorithms that are inspired by the concept of gradients. The gradients indicate the ‘goodness’ of the different possible next hops and are used to forward sensor data to users. For example, when a sensor node detects that its energy reserve has dropped below a certain threshold, it discourages other sensor nodes from sending data to it by decreasing its gradient. This achieves the goal of utilizing resources uniformly and prolongs system lifetime effectively.

In [3], the authors formulate the routing problem with the objective function of maximizing system lifetime given the sets of origin and destination nodes and the information generation rates at the origin nodes. They propose an algorithm to select the routes and the corresponding power levels such that the time until the batteries of the nodes runs out is maximized. In [18], the authors extend the energy conserving routing model presented by [3] to a network where some of the sensor nodes have a very low data rate, as well as limited battery capacity. Both of these papers define the system lifetime as the length of time until the first battery drains out.

1.3.1.2 Hierarchical Routing

Before discussing hierarchical routing, we first consider the partial network structure shown in Figure 1-4. Each cluster has a cluster head which collects data from its cluster members, aggregates it and sends it to the Sink, or an upper level cluster head. For example, nodes 1.1.1, 1.1.2, 1.1.3, 1.1.4, 1.1.5 and 1.1 form a cluster with node 1.1 as the cluster head. Accordingly, there exist other cluster heads such as 1.2, 1 etc. These cluster heads, in turn, form a cluster with node 1 as their cluster head. This pattern is repeated to form a hierarchy of clusters with the uppermost cluster nodes reporting directly to the Sink, which forms the root of this hierarchy and supervises the entire network.

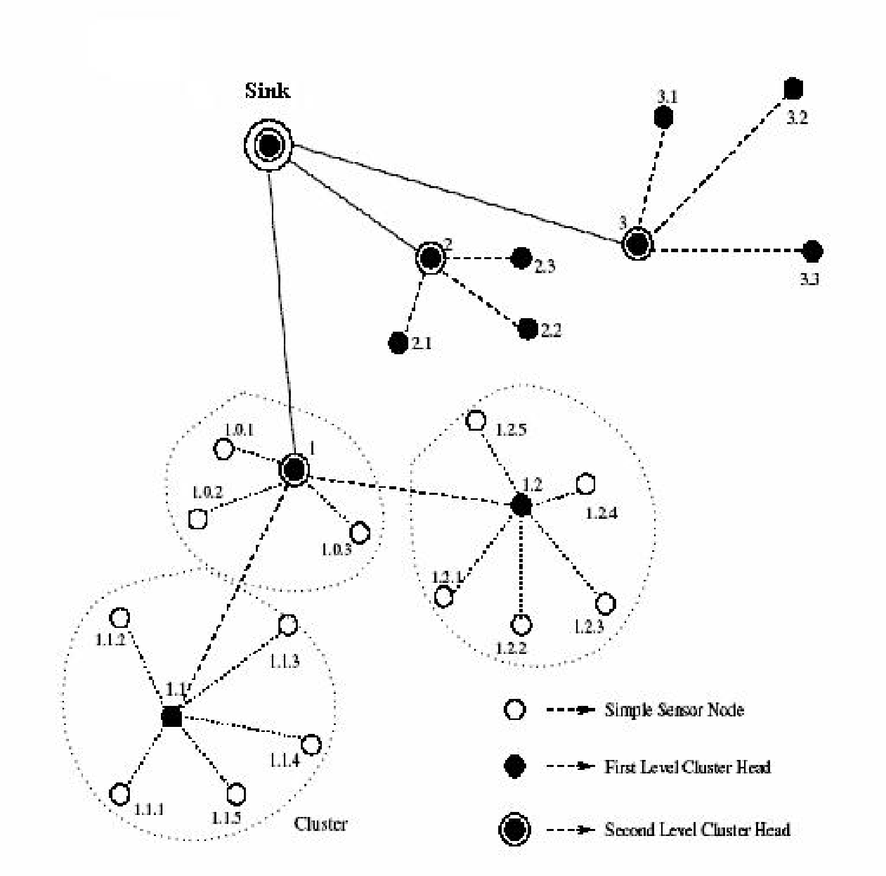

Figure 1-4 Hierarchical Clustering [11]

The main features of this architecture are as follows:

- All the sensor nodes need to transmit only to their immediate cluster head, thus saving energy.
- Only the cluster head needs to perform additional computations on the data. So, energy is again conserved.
- The cluster members are mostly adjacent to each other and have similar data. Since the cluster heads aggregate similar data, aggregation is said to be more effective.

- Cluster heads at higher levels in the hierarchy need to transmit data over correspondingly larger distances. Combined with the extra computations they perform, they end up consuming energy faster than the other nodes. In order to evenly distribute this consumption, all the nodes take turns as the cluster head.
- Since only the cluster heads need to know how to route the data toward their own cluster head or the Sink, the complexity of data routing is reduced.

In [7], the authors propose LEACH (Low-Energy Adaptive Clustering Hierarchy), a clustering-based protocol that minimizes energy dissipation in wireless sensor networks. LEACH is a good approximation of a proactive network protocol. Its key features are: 1) localized coordination and control for cluster set-up and operation; 2) randomized rotation of the cluster heads and the corresponding clusters; and 3) local compression to reduce global communication. LEACH outperforms classical clustering algorithms by using adaptive clusters and rotating cluster-heads, which allow the energy requirements of the system to be distributed among all the sensor nodes. In addition, LEACH is able to perform local computation in each cluster to reduce the amount of data that must be transmitted to the Sink. This achieves a large reduction in the energy dissipation, because computation is much cheaper than communication.

In [11], the authors introduce a new energy efficient protocol, TEEN (Threshold sensitive Energy Efficient sensor Network protocol) for reactive networks. It means time critical data can reach the user almost instantaneously.

It also provides some parameters for users to control the trade-off between energy efficiency and accuracy, based on their needs. The main drawback of this protocol is that it is not well suited for applications where the user needs to get data on a regular basis. In [12], the same authors of [11] further propose a hybrid protocol, APTEEN (Adaptive Periodic Threshold-sensitive Energy Efficient Sensor Network Protocol) which is both well suited for proactive and reactive networks.

1.3.2 Bounding System Lifetime

In [2], the authors propose an optimal role assignment approach, which, in principle, permits derivation of bounds for networks with arbitrarily complex capabilities. However, the computational costs of such derivations may be prohibitive. This paper shows that for several practically useful scenarios, including wireless sensor networks with a specified topology that allows aggregation, this approach in fact leads to polynomial time bound derivation.

In [17], the authors propose an analytical model to estimate and evaluate the network lifetime in a randomly deployed multi-hop wireless sensor network. In this paper the network lifetime is defined as the time interval from the point that a sensor network starts its operation until the point that loss of communication to the Sink by all sensor nodes occurs. In most cases, the operation of the wireless sensor networks is completely disrupted if and only if all of the nodes that can directly communicate with the Sink expire. Consequently, the lifetime of these nodes is more critical to the network lifetime. Thus, we can derive the network lifetime by calculating the lifetime of these critical sensor nodes.

In [4], the authors propose a model to estimate a clustering-based proactive heterogeneous wireless sensor network with two types of sensors equipped with different battery power. In addition, in [18] the authors also derives a bound on the network lifetime.

Chapter 2 Problem Formulation

The goal of this paper is to derive an energy efficient routing algorithm to maximize the system lifetime of wireless sensor networks. We define the system lifetime as the time interval from the point that a sensor network starts its operation until the point that the information about the occurrence of any of events can not be delivered to the Sink. This is the coverage constraint in this paper. With regard to some time-critical applications, if the information aobut the occurrence of any of events cannot be delivered to the Sink within a predefined time interval, namely the hop constraint, the system isn't alive.

At first, the sensor nodes are densely scattered in a sensor field. From the radius of sensing and communication of individual sensor nodes, the topology of wireless sensor networks can be depicted as shown in Figure 2-1.

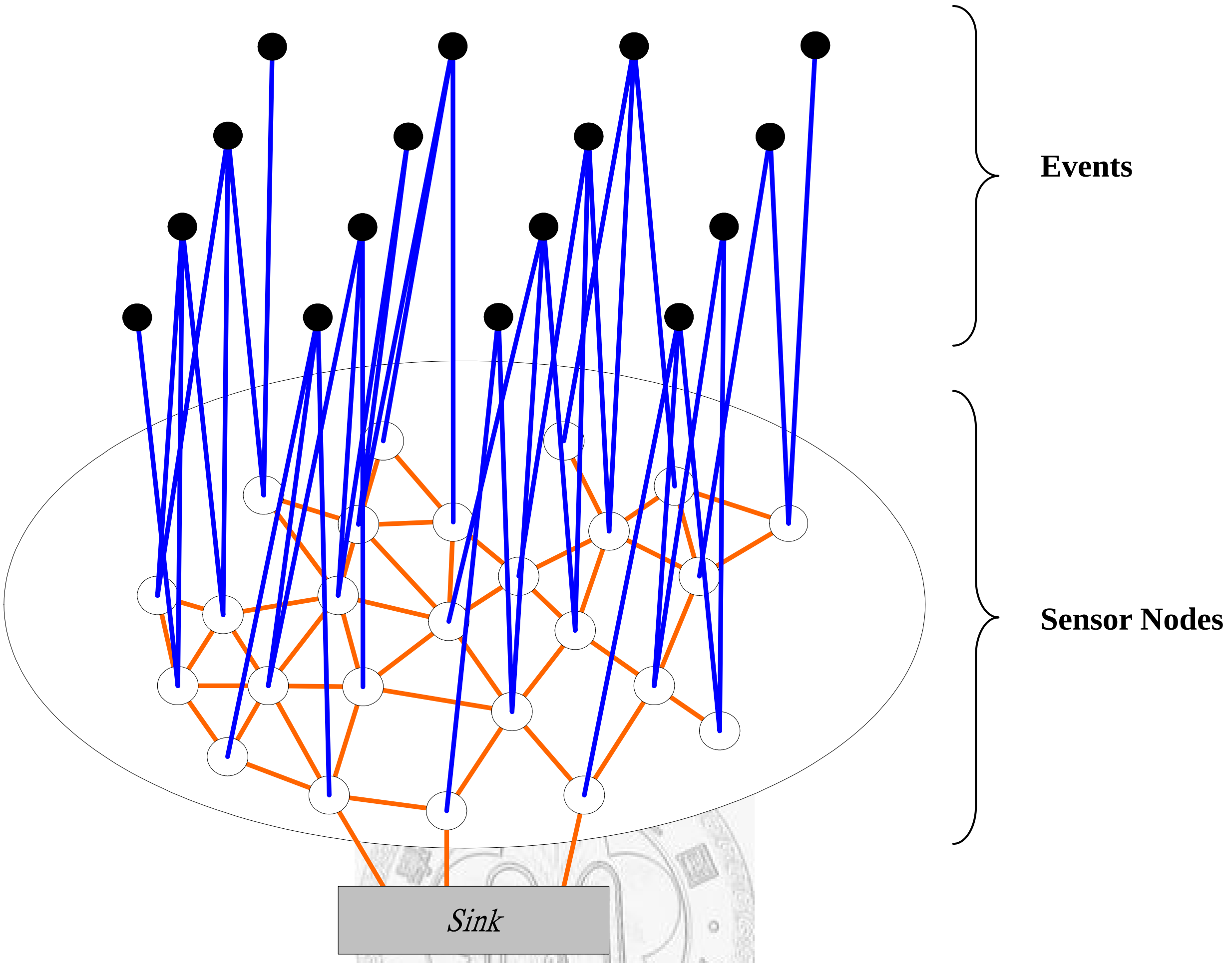

Figure 2-1 Topology of Wireless Sensor Networks

Problem assumption:
- Once an event occurs, the sensor nodes sense this event and deliver this information to the Sink instantly.
- In the beginning, the system is alive.
- All sensor nodes and the Sink are fixed.
- A sensor node dies when it runs out its energy.

Table 2-1 Problem Assumptions

2.1 Model 1

Although this model only describes the time interval from the point that a sensor network starts its operation until the point that the first sensor node runs out its energy, we can still use this model iteratively to obtain the system lifetime and satisfy our modified definition of system lifetime.

2.1.1 Problem Descriptions

Given:

- A set of events.
- A set of nodes including sensor nodes and the Sink.
- A set of links in the nodes.
- A set of out-links and in-links of each sensor node.
- A set of sensor nodes for each event, so that each of the sensor nodes not only senses the event, but also has at least one path to the Sink.
- A set of events for each sensor node, so that the node not only senses each of the events, but also delivers information about the occurrence of these events to the Sink.
- A set of sensor nodes which are one hop away from the Sink.
- Number of flows per unit time caused by each event.
- A set of candidate paths from each sensor node to the Sink.
- Link capacity on each link and nodal capacity on each sensor node.
- Hop constraint for each event.
- Initial energy of each sensor node.

- Energy needed for each sensor node to execute its routine operations, as well as sending and receiving a unit of flow.

Objective:

To maximize the system lifetime of wireless sensor networks.

Subject to:

- Coverage constraint — the information about the occurrence of any of events can be delivered to the Sink by sensor nodes.
- Routing constraint — for each sensor node, once it senses the occurrence of an event, it only selects one path to send data back to the Sink.
- Hop constraint — for each event, paths which are selected to deliver the information about the occurrence of that event are limited by a predefined hop constraint to satisfy the time-critical demand of some applications.
- Link capacity constraint — the total flow on each link can't exceed its capacity.
- Nodal capacity constraint — the total flow passed through a sensor node can't exceed its capacity.

To determine:

1. Routing of wireless sensor networks
2. Total flow on each link which is caused by events.
3. Maximal system lifetime of wireless sensor networks.

Table 2-2 Problem Description for Model 1

2.1.2 Notations

Given Parameters	
Notation	**Description**
E	The set of events.
N	The set of sensor nodes and the Sink.
L	The set of links in *N*.
L_n^+	The set of out-links of sensor node *n*. ($\forall n \in N-\{Sink\}$)
L_n^-	The set of in-links of sensor node *n*. ($\forall n \in N-\{Sink\}$)
A_e	The set of sensor nodes that any one of them not only senses event *e*, but has at least one path to the Sink. ($\forall e \in E$)
B_n	The set of events that can be sensed and the information about the occurrence of each of them that can be delivered to the Sink by sensor node *n*. ($\forall n \in N-\{Sink\}$)
D	The set of sensor nodes which are one hop away from the Sink.
f_e	Number of flows per unit time caused by occurrence of event *e*. ($\forall e \in E$)
P_n	The set of paths which are from sensor node *n* to the Sink. ($\forall n \in N-\{Sink\}$)
δ_{pl}	1 if path *p* uses link *l*; otherwise 0. $\delta_{pl}=\{1,0\}$ ($\forall p \in P_n$, $\forall l \in L$)
C_l	flow capacity on link *l*, the upper bound of flow per unit time on link *l*. ($\forall l \in L$)
U_n	flow capacity on sensor node *n*, the upper bound of flow per unit time on sensor node *n*. ($\forall n \in N-\{Sink\}$)
H_e	Hop constraint for event *e*. ($\forall e \in E$)
E_n	Initial energy of sensor node *n*. ($\forall n \in N-\{Sink\}$) ($E_n > 0$)

E_{sent}	Energy consumption of a sensor node when sending a unit of flow.
E_{recv}	Energy consumption of a sensor node when receiving a unit of flow.
ε	Energy consumption per unit time of a sensor node when operating some routine operations, such as sensing. ($\varepsilon > 0$)

Table 2-3 Notation of Given Parameters for Model 1

Decision Variables	
Notation	**Description**
x_{enp}	1 if path *p* is selected from event *e* through sensor node *n* to the Sink; otherwise 0. ($\forall e \in E, \forall n \in A_e, \forall p \in P_n$)
M_{el}	The total flow per unit time on link *l* caused by the occurrence of event *e*. ($\forall e \in E, \forall l \in L$)
T	System lifetime.

Table 2-4 Notation of Decision Variables for Model 1

Formulation

Objective function:

$$Z_{IP1} = \max T \qquad \text{(IP1)}$$

For the convenience of later problem solving, we transform the original problem (IP1) into another equivalent minimization problem as follows.

$$Z_{IP2} = \min -T \qquad \text{(IP2)}$$

subject to:

$$\sum_{p \in P_n} x_{enp} = 1 \qquad \forall e \in E, \forall n \in A_e \qquad (1)$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e \in E, \forall n \in A_e, \forall p \in P_n \qquad (2)$$

$$\sum_{p \in P_n} \sum_{l \in L} x_{enp} \cdot \delta_{pl} \leq H_e \qquad \forall e \in E, \forall n \in A_e \tag{3}$$

$$\sum_{n \in A_e} \sum_{p \in P_n} x_{enp} \cdot \delta_{pl} \cdot f_e \leq M_{el} \qquad \forall e \in E, \forall l \in L \tag{4}$$

$$\sum_{e \in E} M_{el} \leq C_l \qquad \forall l \in L \tag{5}$$

$$\sum_{e \in E} \sum_{l \in L_n^+} M_{el} + \sum_{e \in E} \sum_{l \in L_n^-} M_{el} \leq U_n \qquad \forall n \in N - \{Sink\} \tag{6}$$

$$0 \leq M_{el} \leq |A_e| \cdot f_e \qquad \forall e \in E, \forall l \in L \tag{7}$$

$$\sum_{e \in E} \sum_{l \in L_n^+} M_{el} \cdot E_{sent} + \sum_{e \in E} \sum_{l \in L_n^-} M_{el} \cdot E_{recv} + \varepsilon \leq \frac{E_n}{T} \qquad \forall n \in N - \{Sink\} \tag{8}$$

$$\min_{n \in N - \{Sink\}} \left(\frac{E_n}{\frac{1}{2} \cdot \left(U_n + \sum_{e \in B_n} f_e \right) \cdot E_{sent} + \frac{1}{2} \cdot \left(U_n - \sum_{e \in B_n} f_e \right) \cdot E_{recv} + \varepsilon} \right) \leq T \leq \frac{\max_{n \in D} E_n}{\sum_{n \in N - \{Sink\} - D} \sum_{e \in B_n} f_e \cdot \frac{1}{|D|} \cdot \left(E_{sent} + E_{recv} \right) + \min_{n \in D} \sum_{e \in B_n} f_e \cdot E_{sent} + \varepsilon} \tag{9}$$

Explanation of Constraints	
(1)(2)(3)	These three constraints, which includes the hop constraint, means that a sensor node n in A_e only selects one path to the Sink.
(4)	This constraint describes how $\{M_{el}\}$ are calculated. Once the routing, i.e. $\{x_{enp}\}$, of the entire network is determined, the flow caused by events will move along the selected paths, so that $\{M_{el}\}$ can be obtained.
(5)	This constraint is the link capacity constraint on each link l.
(6)	This constraint is the nodal capacity constraint on each sensor node n.
(7)	This constraint describes $\{M_{el}\}$'s bounds. Where the $\{M_{el}\}$'s lower bounds are straightforward, while their upper bounds are limited by $|A_e|$ and f_e, because link l is at most shared by $|A_e|$ different paths from event e.
(8)	This constraint describes that the system lifetime T is less or equal to the all lifetime of sensor node n.

(9)	This constraint describes T's bounds. Due to the nodal capacity constraint we can calculate each sensor node's minimal lifetime and the T's lower bound equals the lifetime of the sensor node which has the minimal minimal-lifetime among all sensor nodes. In addition, due to the sensor nodes which are one hop away from the Sink are much more critical than other sensor nodes to the system lifetime, they are in chare of relaying almost all flows on the entire network to the Sink, so we can obtain the T's upper bound by finding the sensor node in D has the maximal lifetime.

Table 2-5 Explanation of Constraints for Model 1

2.2 Model 2

This model describes the time interval from the point that a sensor network starts its operation until the point that the information about the occurrence of any of events can not be delivered to the Sink, and it satisfies our modified definition. We can use this model to obtain the system lifetime and satisfy our modified definition of system lifetime.

2.2.1 Problem Descriptions

Given:

- A set of events.
- A set of nodes including sensor nodes and the Sink.
- A set of links in the nodes.
- A set of out-links and in-links of each sensor node.
- A set of sensor nodes for each event, so that each of the sensor nodes not only senses the event, but also has at least one path to the Sink.
- A set of events for each sensor node, so that the node not only senses each of the events, but also delivers information about the occurrence of these events to the Sink.
- Number of flows per unit time caused by each event.
- A set of predefined candidate paths from each sensor node to the Sink.
- A set of sensor nodes on each path.
- Link capacity on each link and nodal capacity on each sensor node.
- Hop constraint for each event.

- Initial energy of each sensor node.
- Energy needed for each sensor node to execute its routine operations, as well as sending and receiving a unit of flow.

Objective:

To maximize the system lifetime of wireless sensor networks.

Subject to:

- Coverage constraint — information about the occurrence of any of the events can be delivered to the Sink by sensor nodes.
- Routing constraint — for each sensor node, once it senses the occurrence of an event, it only selects one path to send data back to the Sink.
- Hop constraint — for each event, paths which are selected to deliver the information about the occurrence of that event are limited by a predefined hop constraint to satisfy the time-critical demand of some applications.
- Link capacity constraint — the total flow on each link can't exceed its capacity.
- Nodal capacity constraint — the total flow passed through a sensor node can't exceed its capacity.

To determine:

1. Routing of wireless sensor networks
2. Total flow on each link which is caused by events.
3. Maximal system lifetime of wireless sensor networks.

Table 2-6 Problem Description for Model 2

2.2.2 Notations

Given Parameters	
Notation	**Description**
E	The set of events.
N	The set of sensor nodes and the Sink.
L	The set of links in *N*.
L_v^+	The set of out-links of sensor node *v*. （$\forall v \in N-\{Sink\}$）
L_v^-	The set of in-links of sensor node *v*. （$\forall v \in N-\{Sink\}$）
A_e	The set of sensor nodes that any one of them not only senses event *e*, but has at least one path to the Sink. （$\forall e \in E$）
B_v	The set of events that can be sensed and the information about the occurrence of each of them that can be delivered to the Sink by sensor node *v*. （$\forall v \in N-\{Sink\}$）
f_e	Number of flows per unit time caused by the occurrence of event *e*.（$\forall e \in E$）
P_n	The set of *W* predefined paths from sensor node *n* to the Sink. （$\forall n \in N-\{Sink\}$）
δ_{pl}	1 if path *p* uses link *l*; otherwise 0. $\delta_{pl}=\{1,0\}$（$\forall p \in P_n$, $\forall l \in L$）
V_p	The set of sensor nodes on path *p*. （$\forall p \in P_n, \forall n \in N-\{Sink\}$）
C_l	Flow capacity on link *l*, the upper bound of flow per unit time on link *l*. （$\forall l \in L$）
U_v	Flow capacity on sensor node *v*, the upper bound of flow per unit time on sensor node *v*.（$\forall v \in N-\{Sink\}$）
H_e	Hop constraint for event *e*.（$\forall e \in E$）

E_v	Initial energy of sensor node *v*. ($\forall v \in N-\{Sink\}$) ($E_v > 0$)
E_{sent}	Energy consumption of a sensor node when sending a unit of flow.
E_{recv}	Energy consumption of a sensor node when receiving a unit of flow.
ε	Energy consumption per unit time of a sensor node when operating some routine operations, such as sensing. ($\varepsilon > 0$)

Table 2-7 Notation of Given Parameters for Model 2

Decision Variables	
Notation	**Description**
x_{enp}	1 if path *p* is selected from event *e* through sensor node *n* to the Sink, otherwise 0. ($\forall e \in E, \forall n \in A_e, \forall p \in P_n$)
M_{el}	The total flow per unit time on link *l*, caused by the occurrence of event *e*. ($\forall e \in E, \forall l \in L$)
T_v	Lifetime of sensor node *v*. ($\forall v \in N-\{Sink\}$)
T	System lifetime.

Table 2-8 Notation of Decision Variables for Model 2

Formulation

Objective function:

$$Z_{IP3} = \max T \qquad \text{(IP3)}$$

subject to:

$$\sum_{p \in P_n} x_{enp} = 1 \qquad \forall e \in E, \forall n \in A_e \qquad (1)$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e \in E, \forall n \in A_e, \forall p \in P_n \qquad (2)$$

$$\sum_{p \in P_n} \sum_{l \in L} x_{enp} \cdot \delta_{pl} \le H_e \qquad \forall e \in E, \forall n \in A_e \qquad (3)$$

$$\sum_{n \in A_e} \sum_{p \in P_n} x_{enp} \cdot \delta_{pl} \cdot f_e = M_{el} \qquad \forall e \in E, \forall l \in L \qquad (4)$$

$$\sum_{e \in E} M_{el} \le C_l \qquad \forall l \in L \qquad (5)$$

$$\sum_{e \in E} \sum_{l \in L_v^+} M_{el} + \sum_{e \in E} \sum_{l \in L_v^-} M_{el} \le U_v \qquad \forall v \in N - \{Sink\} \qquad (6)$$

$$0 \le M_{el} \le |A_e| \cdot f_e \qquad \forall e \in E, \forall l \in L \qquad (7)$$

$$T_v = \frac{E_v}{\sum_{e \in E} \sum_{l \in L_v^+} M_{el} \cdot E_{sent} + \sum_{e \in E} \sum_{l \in L_v^-} M_{el} \cdot E_{recv} + \varepsilon} \qquad \forall v \in N - \{Sink\} \qquad (8)$$

$$\frac{E_v}{\frac{1}{2} \cdot \left(U_v + \sum_{e \in B_v} f_e \right) \cdot E_{sent} + \frac{1}{2} \cdot \left(U_v - \sum_{e \in B_v} f_e \right) \cdot E_{recv} + \varepsilon} \le T_v \le \frac{E_v}{\sum_{e \in B_v} f_e \cdot E_{sent} + \varepsilon} \qquad \forall v \in N - \{Sink\} \qquad (9)$$

$$T \le \max_{n \in A_e, p \in P_n} x_{enp} \min_{v \in V_p} T_v \qquad \forall e \in E \qquad (10)$$

Explanation of Constraints	
(1)(2)(3)	These three constraints, which includes the hop constraint, means that a sensor node n in A_e only selects one path to the Sink.
(4)	This constraint describes how $\{M_{el}\}$ are calculated. Once the routing, i.e. $\{x_{enp}\}$, of the entire network is determined, the flow caused by events will move along the selected paths, so that $\{M_{el}\}$ can be obtained.
(5)	This constraint is the link capacity constraint on each link l.
(6)	This constraint is the nodal capacity constraint on each sensor node v.
(7)	This constraint describes $\{M_{el}\}$'s bounds. Where $\{M_{el}\}$'s lower bounds are straightforward, while their upper bounds are limited by $\lvert A_e \rvert$ and f_e, because link l is at most shared by $\lvert A_e \rvert$ different paths from event e.

(8)	This constraint describes how $\{T_v\}$ are calculated. For each sensor node v, its lifetime T_v equals its initial energy divided by its total energy consumption per unit time.
(9)	This constraint describes $\{T_v\}$'s bounds. Their upper bounds are straightforward, and their lower bounds can be calculated by considering the nodal capacity for each sensor node v.
(10)	This constraint describes the definition of the system lifetime. At first, we group all sensor nodes into several groups by different event e. Further, we divide each group into $\|A_e\|$ subgroups. The lifetime of each subgroup equals the minimal lifetime of sensor node in each subgroup, and the lifetime of each group equals the maximal lifetime among the subgroups that belong to it. Finally, the system lifetime is the maximal lifetime among the groups.

Table 2-9 Explanation of Constraints for Model 2

For the convenience of later problem solving, the preceding formulation can be reformulated as follows, where K and O are constants.

Let $S_{enp} = \min_{v \in V_p} T_v \quad, \forall e \in E, \ \forall n \in A_e, \ \forall p \in P_n$

Let $R_e = \max_{n \in A_e, p \in P_n} x_{enp} \cdot S_{enp} \quad, \forall e \in E$

Formulation

Objective function:

$$Z_{IP4} = \max T - \sum_{e \in E} K \cdot R_e + \sum_{e \in E} \sum_{n \in A_e} \sum_{p \in P_n} O \cdot S_{enp} \qquad \text{(IP4)}$$

For the convenience of later problem solving, we further transform the problem (IP4) into another equivalent minimization problem as follows.

$$Z_{IP5} = \min -T + \sum_{e\in E} K \cdot R_e - \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n} O \cdot S_{enp} \qquad \text{(IP5)}$$

subject to:

$$\sum_{p\in P_n} x_{enp} = 1 \qquad \forall e\in E, \forall n\in A_e \qquad (1)$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (2)$$

$$\sum_{p\in P_n}\sum_{l\in L} x_{enp}\cdot\delta_{pl} \le H_e \qquad \forall e\in E, \forall n\in A_e \qquad (3)$$

$$\sum_{n\in A_e}\sum_{p\in P_n} x_{enp}\cdot\delta_{pl}\cdot f_e = M_{el} \qquad \forall e\in E, \forall l\in L \qquad (4)$$

$$\sum_{e\in E} M_{el} \le C_l \qquad \forall l\in L \qquad (5)$$

$$\sum_{e\in E}\sum_{l\in L_v^+} M_{el} + \sum_{e\in E}\sum_{l\in L_v^-} M_{el} \le U_v \qquad \forall v\in N-\{Sink\} \qquad (6)$$

$$0 \le M_{el} \le |A_e|\cdot f_e \qquad \forall e\in E, \forall l\in L \qquad (7)$$

$$\sum_{e\in E}\sum_{l\in L_v^+} M_{el}\cdot E_{sent} + \sum_{e\in E}\sum_{l\in L_v^-} M_{el}\cdot E_{recv} + \varepsilon = \frac{E_v}{T_v} \qquad \forall v\in N-\{Sink\} \qquad (8)$$

$$\frac{E_v}{\frac{1}{2}\cdot\left(U_v + \sum_{e\in B_v} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_v - \sum_{e\in B_v} f_e\right)\cdot E_{recv} + \varepsilon} \le T_v \le \frac{E_v}{\sum_{e\in B_v} f_e\cdot E_{sent} + \varepsilon} \qquad \forall v\in N-\{Sink\} \qquad (9)$$

$$S_{enp} \le T_v \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n, \forall v\in V_p \qquad (10)$$

$$\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v + \sum_{e\in B_v} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_v - \sum_{e\in B_v} f_e\right)\cdot E_{recv} + \varepsilon}\right) \le S_{enp} \le \min_{v\in V_p}\left(\frac{E_v}{\left(f_e + \sum_{e\in B_v} f_e\right)\cdot E_{sent} + \varepsilon}\right) \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (11)$$

$$x_{enp}\cdot S_{enp} \le R_e \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (12)$$

$$\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v + \sum_{e\in B_v} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_v - \sum_{e\in B_v} f_e\right)\cdot E_{recv} + \varepsilon}\right)\right] \le R_e \le \max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e + \sum_{e\in B_v} f_e\right)\cdot E_{sent} + \varepsilon}\right)\right] \qquad \forall e\in E \qquad (13)$$

$$T \le R_e \qquad \forall e \in E \tag{14}$$

$$\min_{e\in E}\left\{\max_{n\in A_e,\, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\right\} \le T \le \min_{e\in E}\left\{\max_{n\in A_e,\, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]\right\} \tag{15}$$

Chapter 3 Solution Approaches

3.1 Lagrangean Relaxation Method

In the 1970s [5], Lagrangean Relaxation method was used for scheduling and solving general integer programming problems. It is a flexible approach that can provide proper solutions for many problems, and has become one of the best tools for solving optimization problems such as integer programming, linear programming combinatorial optimization, and non-linear programming. The method has several advantages. For example, it can decompose complex mathematical models in many different ways into some stand-alone subproblems, which can then be solved optimally, using any proper algorithm [5].

In addition, Lagrangean Relaxation allows us to determine the boundary of our objective function, thus we can use it to implement heuristic algorithms to obtain feasible solutions. It is a flexible solution strategy that permits modelers to exploit the underlying structure in any optimization problem by relaxing complicating constraints. This method permits us to "pull apart" models by removing constraints and place them in the objective function with associated Lagrangean multipliers. The optimal value of the relaxed problem is always a lower bound (for minimization problems) on the objective function value of the problem. To obtain the sharpest lower bound, we need to choose the best multiplier (Lagrangean multiplier problem) so that the optimal value of the Lagrangean subproblem is as large as possible. Although we can solve the Lagrangean multiplier problem in a variety of ways, the subgradient optimization technique is the most popular technique for dealing with the issue [5][6].

Figure 3-1 illustrates Lagrangean Relaxation, while Figure 3-2 gives a detailed explanation of Lagrangean Relaxation procedures.

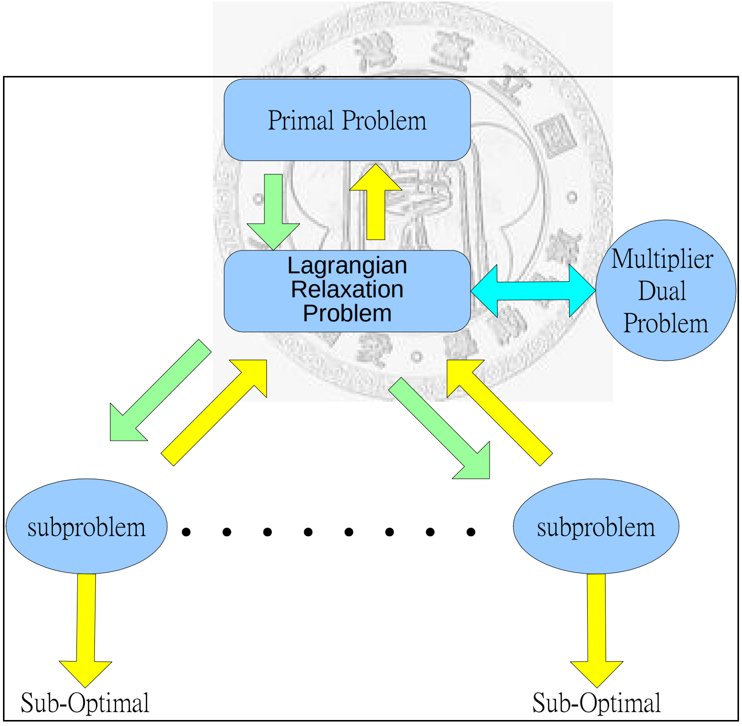

Figure 3-1 An Illustration of Lagrangean Relaxation

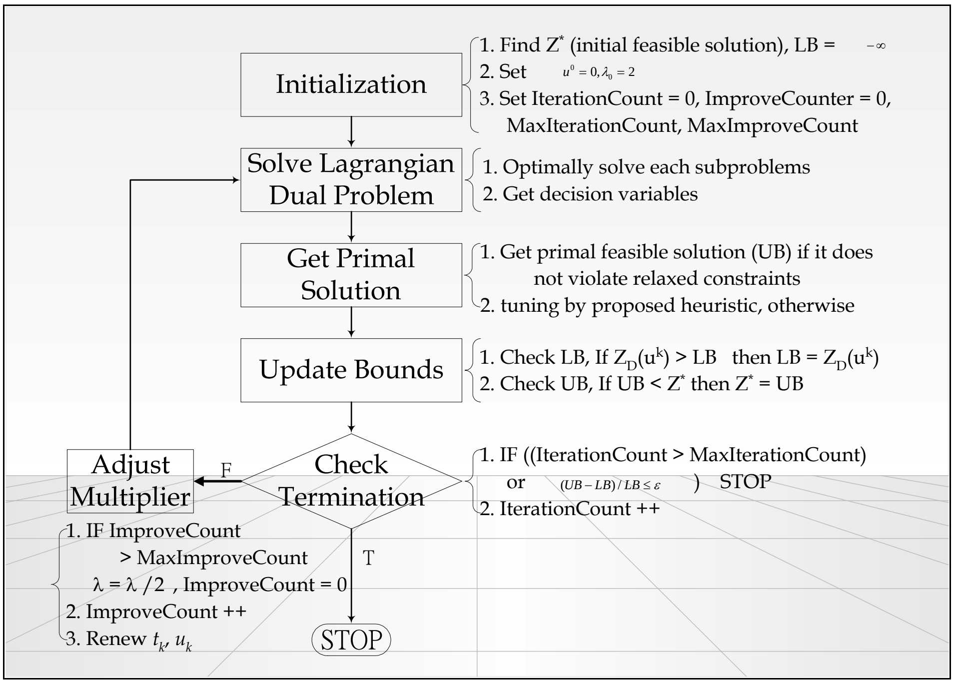

Figure 3-2 Lagrangean Relaxation Procedures

3.2 Model 1

3.2.1 Solution Approach

By using Lagrangean Relaxation method, we can transform the primal problem (IP2) into the following Lagrangean Relaxation problem (LR1) where constraints (4), (6) and (8) are relaxed.

3.2.2 Lagrangean Relaxation

For a vector of non-negative Lagrangean multipliers, a Lagrangean Relaxation problem of IP2 is given by optimization problem (LR1):

$$Z_{d1}(\alpha,\beta,\gamma)$$

$$= \min -T + \sum_{e\in E}\sum_{l\in L}\alpha_{el}\cdot\left(\sum_{n\in A_e}\sum_{p\in P_n} x_{enp}\cdot\delta_{pl}\cdot f_e - M_{el}\right)$$

$$+ \sum_{n\in N-\{Sink\}}\beta_n\cdot\left(\sum_{e\in E}\sum_{l\in L_n^+} M_{el} + \sum_{e\in E}\sum_{l\in L_n^-} M_{el} - U_n\right)$$

$$+ \sum_{n\in N-\{Sink\}}\gamma_n\cdot\left(\sum_{e\in E}\sum_{l\in L_n^+} M_{el}\cdot E_{sent} + \sum_{e\in E}\sum_{l\in L_n^-} M_{el}\cdot E_{recv} + \varepsilon - \frac{E_n}{T}\right) \qquad \text{(LR1)}$$

subject to:

$$\sum_{p\in P_n} x_{enp} = 1 \qquad \forall e\in E, \forall n\in A_e \qquad (1)$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (2)$$

$$\sum_{p\in P_n}\sum_{l\in L} x_{enp}\cdot\delta_{pl} \le H_e \qquad \forall e\in E, \forall n\in A_e \qquad (3)$$

$$\sum_{e\in E} M_{el} \le C_l \qquad \forall l\in L \qquad (5)$$

$$0 \le M_{el} \le \left|A_e\right| \cdot f_e \qquad\qquad \forall e \in E, \forall l \in L \qquad (7)$$

$$\min_{n \in N-\{Sink\}} \left(\frac{E_n}{\frac{1}{2}\cdot\left(U_n + \sum_{e \in B_n} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_n - \sum_{e \in B_n} f_e\right)\cdot E_{recv} + \varepsilon} \right) \le T \le \frac{\max_{n \in D} E_n}{\sum_{n \in N-\{Sink\}-D} \sum_{e \in B_n} f_e \cdot \frac{1}{|D|}\cdot\left(E_{sent} + E_{recv}\right) + \min_{n \in D} \sum_{e \in B_n} f_e \cdot E_{sent} + \varepsilon} \qquad (9)$$

Where α, β, γ are the vectors of $\{\alpha_{el}\}$, $\{\beta_n\}$, $\{\gamma_n\}$ and α, β, γ are the Lagrangean multipliers and α, β, γ ≥ 0. To solve LR1, we can decompose it into the following three independent and easily solvable optimization subproblems.

Subproblem 1-1 (related to decision variable T)

$Z_{sub1-1}(\gamma)$

$$= \min - T - \sum_{n \in N-\{Sink\}} \gamma_n \cdot \frac{E_n}{T}$$

$$= \max T + \sum_{n \in N-\{Sink\}} \gamma_n \cdot \frac{E_n}{T}$$

subject to:

$$\min_{n \in N-\{Sink\}} \left(\frac{E_n}{\frac{1}{2}\cdot\left(U_n + \sum_{e \in B_n} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_n - \sum_{e \in B_n} f_e\right)\cdot E_{recv} + \varepsilon} \right) \le T \le \frac{\max_{n \in D} E_n}{\sum_{n \in N-\{Sink\}-D} \sum_{e \in B_n} f_e \cdot \frac{1}{|D|}\cdot\left(E_{sent} + E_{recv}\right) + \min_{n \in D} \sum_{e \in B_n} f_e \cdot E_{sent} + \varepsilon} \qquad (9)$$

The graph of this subproblem is a convex curve. The maximal value occurs at an intersection point on either side of T's bound. Namely, either

$$T = \min_{n \in N-\{Sink\}} \left(\frac{E_n}{\frac{1}{2}\cdot\left(U_n + \sum_{e \in B_n} f_e\right)\cdot E_{sent} + \frac{1}{2}\cdot\left(U_n - \sum_{e \in B_n} f_e\right)\cdot E_{recv} + \varepsilon} \right) \quad \text{or}$$

$$T = \frac{\max_{n \in D} E_n}{\sum_{n \in N - \{Sink\} - D} \sum_{e \in B_n} f_e \cdot \frac{1}{|D|} \cdot \left(E_{sent} + E_{recv}\right) + \min_{n \in D} \sum_{e \in B_n} f_e \cdot E_{sent} + \varepsilon}$$

can obtain the maximal value.

Subproblem 1-2 (related to decision variable x_{enp})

$$Z_{sub1-2}(\alpha)$$

$$= \min \sum_{e \in E} \sum_{n \in A_e} \sum_{p \in P_n} \sum_{l \in L} \alpha_{el} \cdot x_{enp} \cdot \delta_{pl} \cdot f_e$$

$$= \min \sum_{e \in E} \sum_{n \in A_e} \sum_{p \in P_n} x_{enp} \cdot \sum_{l \in L} (\alpha_{el} \cdot f_e) \cdot \delta_{pl}$$

subject to:

$$\sum_{p \in P_n} x_{enp} = 1 \qquad \forall e \in E, \forall n \in A_e \qquad (1)$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e \in E, \forall n \in A_e, \forall p \in P_n \qquad (2)$$

$$\sum_{p \in P_n} \sum_{l \in L} x_{enp} \cdot \delta_{pl} \le H_e \qquad \forall e \in E, \forall n \in A_e \qquad (3)$$

This subproblem is composed of $|E|$ hop constrained shortest path problems for each event e, where $\alpha_{el} \cdot f_e$ is the link cost of link l. By using the Bellman-Ford algorithm, we can optimally solve these problems, and then properly set $\{x_{enp}\}$ to 1, if path p is selected, otherwise set it to 0.

Subproblem 1-3 (related to decision variable M_{el})

$$Z_{sub1-3}(\alpha, \beta, \gamma)$$

$$= \min - \sum_{e \in E} \sum_{l \in L} \alpha_{el} \cdot M_{el} + \sum_{n \in N - \{Sink\}} \beta_n \cdot \left(\sum_{e \in E} \sum_{l \in L_n^+} M_{el} + \sum_{e \in E} \sum_{l \in L_n^-} M_{el} \right)$$

$$+\sum_{n\in N-\{Sink\}}\gamma_n\cdot\left(\sum_{e\in E}\sum_{l\in L_n^+}M_{el}\cdot E_{sent}+\sum_{e\in E}\sum_{l\in L_n^-}M_{el}\cdot E_{recv}\right)$$

$$=\min-\sum_{e\in E}\sum_{n\in N-\{Sink\}}\sum_{l\in L_n^+}\alpha_{el}\cdot M_{el}+\sum_{e\in E}\sum_{n\in N-\{Sink\}}\sum_{l\in L_n^+}\left(\beta_n+\gamma_n\cdot E_{sent}\right)\cdot M_{el}$$

$$+\sum_{e\in E}\sum_{n\in N-\{Sink\}}\sum_{l\in L_n^-}\left(\beta_n+\gamma_n\cdot E_{recv}\right)\cdot M_{el}$$

$$=\min\sum_{e\in E}\sum_{n\in N-\{Sink\}}\sum_{l\in L_n^+}\left(-\alpha_{el}+\beta_n+\gamma_n\cdot E_{sent}\right)\cdot M_{el}+\sum_{e\in E}\sum_{n\in N-\{Sink\}}\sum_{l\in L_n^-}\left(\beta_n+\gamma_n\cdot E_{recv}\right)\cdot M_{el}$$

subject to:

$$\sum_{e\in E}M_{el}\le C_l \qquad \forall l\in L \qquad (5)$$

$$0\le M_{el}\le\left|A_e\right|\cdot f_e \qquad \forall e\in E,\forall l\in L \qquad (7)$$

To solve this subproblem we first need to determine all the coefficient values of $\{M_{el}\}$. For each event e, by putting the start point sensor node of link l into the first term of this subproblem's objective function and then putting the end point sensor node of link l into the second term of this subproblem's objective function, we can obtain the coefficient value of M_{el} on link l. Thus, we can calculate all the coefficient values of $\{M_{el}\}$. We can then determine $\{M_{el}\}$'s values to obtain the optimal solution to this subproblem. If a M_{el}'s coefficient value is positive, we set its value to 0, otherwise we sort the M_{el}s by their coefficient values from small to large. Later, we will determine these M_{el}'s values in this order, and to further satisfy constraint (5), i.e. $\sum_{e\in E}M_{el}\le C_l$ for each link l, we determine the M_{el}'s value one by one. Before determining each of them, we should check the residual capacity on link l. If it is greater than or equal to $\left|A_e\right|\cdot f_e$, we set its value to $\left|A_e\right|\cdot f_e$, then decrease the residual capacity on link l by $\left|A_e\right|\cdot f_e$, otherwise we set its value to the remaining residual capacity on link l, and decrease the residual capacity on link l to 0.

3.2.3 The Dual Problem and the Subgradient Method

According to the weak Lagrangean duality theorem [6], for any α, β, $\gamma \geq 0$, $Z_{d1}(\alpha,\beta,\gamma)$ is a lower bound on Z_{IP2}. The following dual problem (D1) is then constructed to calculate the tightest lower bound.

Dual Problem (D1):

$$Z_{D1} = \max Z_{d1}(\alpha,\beta,\gamma) \qquad \text{(D1)}$$

subject to:

$$\alpha,\beta,\gamma \geq 0$$

The most popular method for solving the dual problem is the subgradient method [9]. Let g be a subgradient of $Z_{d1}(\alpha,\beta,\gamma)$. Then, at the k th iteration of the subgradient optimization procedure, the multiplier vector $\pi=(\alpha,\beta,\gamma)$ is updated by $\pi^{k+1}=\pi^k+t^k g^k$. The step size t^k is determined by $t^k=\lambda_k \frac{Z^*_{IP2}-Z_{d1}(\pi_k)}{\left\|g^k\right\|^2}$. Z^*_{IP2} is the best upper bound on the primal objective function value after the k th iteration obtained from heuristic solutions. λ_k is a constant between 0 and 2.

3.3 Model 2

3.3.1 Solution Approach

By using Lagrangean Relaxation method, we can transform the primal problem (IP5) into the following Lagrangean Relaxation problem (LR2) where constraints (4), (6), (8), (10), (12) and (14) are relaxed.

3.3.2 Lagrangean Relaxation

For a vector of non-negative Lagrangean multipliers, a Lagrangean Relaxation problem of IP5 is given by optimization problem (LR2):

$$
\begin{aligned}
& Z_{d2}(\alpha,\beta,\gamma,\theta,\mu,\omega) \\
&= \min -T + \sum_{e\in E} K\cdot R_e - \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n} O\cdot S_{enp} \\
&\quad + \sum_{e\in E}\sum_{l\in L}\alpha_{el}\cdot\left(\sum_{n\in A_e}\sum_{p\in P_n} x_{enp}\cdot\delta_{pl}\cdot f_e - M_{el}\right) \\
&\quad + \sum_{v\in N-\{Sink\}}\beta_v\cdot\left(\sum_{e\in E}\sum_{l\in L_v^+} M_{el} + \sum_{e\in E}\sum_{l\in L_v^-} M_{el} - U_v\right) \\
&\quad + \sum_{v\in N-\{Sink\}}\gamma_v\cdot\left(\sum_{e\in E}\sum_{l\in L_v^+} M_{el}\cdot E_{sent} + \sum_{e\in E}\sum_{l\in L_v^-} M_{el}\cdot E_{recv} + \varepsilon - \frac{E_v}{T_v}\right) \\
&\quad + \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\sum_{v\in V_p}\theta_{enpv}\cdot\left(S_{enp} - T_v\right) \\
&\quad + \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\mu_{enp}\cdot\left(x_{enp}\cdot S_{enp} - R_e\right) \\
&\quad + \sum_{e\in E}\omega_e\cdot\left(T - R_e\right) \qquad \text{(LR2)}
\end{aligned}
$$

subject to:

$$\sum_{p\in P_n} x_{enp} = 1 \qquad \forall e\in E, \forall n\in A_e \tag{1}$$

$$x_{enp} = 0 \text{ or } 1 \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \tag{2}$$

$$\sum_{p\in P_n}\sum_{l\in L} x_{enp}\cdot\delta_{pl} \le H_e \qquad \forall e\in E, \forall n\in A_e \tag{3}$$

$$\sum_{e\in E} M_{el} \le C_l \qquad \forall l\in L \tag{5}$$

$$0 \le M_{el} \le |A_e|\cdot f_e \qquad \forall e\in E, \forall l\in L \tag{7}$$

$$\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon} \le T_v \le \frac{E_v}{\sum_{e\in B_v} f_e\cdot E_{sent}+\varepsilon} \qquad \forall v\in N-\{Sink\} \tag{9}$$

$$\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right) \le S_{enp} \le \min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right) \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \tag{11}$$

$$\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right] \le R_e \le \max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right] \qquad \forall e\in E \tag{13}$$

$$\min_{e\in E}\left\{\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\right\} \le T \le \min_{e\in E}\left\{\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]\right\} \tag{15}$$

Where α, β, γ, θ, μ and ω are the vectors of $\{\alpha_{el}\}$, $\{\beta_v\}$, $\{\gamma_v\}$, $\{\theta_{enpv}\}$, $\{\mu_{enp}\}$ and $\{\omega_e\}$ respectively; α, β, γ, θ, μ and ω are the Lagrangean multipliers; and β, θ, μ, ω ≥ 0. To solve LR2, we can decompose it into the following three independent and easily solvable optimization subproblems.

Subproblem 2-1 (related to decision variable T , R_e)

$$Z_{sub2-1}(\mu,\omega)$$

$$= \min -T + \sum_{e\in E} K\cdot R_e - \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\mu_{enp}\cdot R_e + \sum_{e\in E}\omega_e\cdot\left(T-R_e\right)$$

$$= \min\left(\sum_{e\in E}\omega_e - 1\right)\cdot T + \sum_{e\in E} R_e\cdot\left(K - \sum_{n\in A_e}\sum_{p\in P_n}\mu_{enp} - \omega_e\right)$$

subject to:

$$\max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\le R_e\le \max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right] \quad \forall e\in E \qquad (13)$$

$$\min_{e\in E}\left\{\max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\right\}\le T\le \min_{e\in E}\left\{\max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]\right\} \qquad (15)$$

This subproblem can be optimally solved by the following steps:

1. First, determine R_e 's value for each event e . If its coefficient value $\left(K - \sum_{n\in A_e}\sum_{p\in P_n}\mu_{enp} - \omega_e\right)\ge 0$, then we set R_e 's value to $\max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]$; otherwise, we set its value to $\max_{n\in A_e,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]$.

2. To determine T 's value. If its coefficient value $\left(\sum_{e\in E}\omega_e - 1\right)\ge 0$, we set its

value to $\min_{e\in E}\left\{\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\right\}$;

otherwise, we set its value to $\min_{e\in E}\left\{\max_{n\in A_e, p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]\right\}$.

Subproblem 2-2 (related to decision variable x_{enp}, S_{enp})

$$Z_{sub2-2}(\alpha,\theta,\mu)$$

$$=\min -\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n} O\cdot S_{enp}+\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n} x_{enp}\cdot\sum_{l\in L}(\alpha_{el}\cdot f_e)\cdot\delta_{pl}$$
$$-\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\sum_{v\in V_p}\theta_{enpv}\cdot S_{enp}+\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\mu_{enp}\cdot x_{enp}\cdot S_{enp}$$

$$=\min\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\left(x_{enp}\cdot\sum_{l\in L}(\alpha_{el}\cdot f_e)\cdot\delta_{pl}+\left(\mu_{enp}\cdot x_{enp}+\sum_{v\in V_p}\theta_{enpv}-O\right)\cdot S_{enp}\right)$$

subject to:

$$\sum_{p\in P_n} x_{enp}=1 \qquad \forall e\in E, \forall n\in A_e \qquad (1)$$

$$x_{enp}=0 \text{ or } 1 \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (2)$$

$$\sum_{p\in P_n}\sum_{l\in L} x_{enp}\cdot\delta_{pl}\le H_e \qquad \forall e\in E, \forall n\in A_e \qquad (3)$$

$$\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\le S_{enp}\le\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right) \qquad \forall e\in E, \forall n\in A_e, \forall p\in P_n \qquad (11)$$

This subproblem can be optimally solved as follows:

First, we should determine $\{x_{enp}\}$'s values. To solve $\{x_{enp}\}$, we don't have to try all combinations of $\{x_{enp}\}$. It is considerably costly. Thus, we further decompose this searching problem into $\sum_{e\in E}|A_e|$ easily solvable searching subproblems, after which, due to the constraint (1) and the predefined W candidate paths for each sensor node, we only need to try W combinations to solve each searching subproblem. To determine the best combination to each searching subproblem, we should choose the combination that makes $\sum_{p\in P_n}\left(x_{enp}\cdot\sum_{l\in L}(\alpha_{el}\cdot f_e)\cdot\delta_{pl}+\left(\mu_{enp}\cdot x_{enp}+\sum_{v\in V_p}\theta_{enpv}-O\right)\cdot S_{enp}\right)$ the smallest. With regard to S_{enp}'s value, if its coefficient value $\left(\mu_{enp}\cdot x_{enp}+\sum_{v\in V_p}\theta_{enpv}-O\right)\geq 0$ we set its value to $\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v}f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v}f_e\right)\cdot E_{recv}+\varepsilon}\right)$; otherwise, we set its value to $\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v}f_e\right)\cdot E_{sent}+\varepsilon}\right)$.

Subproblem 2-3 (related to decision variable M_{el})

$$Z_{sub2-3}(\alpha,\beta,\gamma)$$

$$=\min-\sum_{e\in E}\sum_{l\in L}\alpha_{el}\cdot M_{el}+\sum_{v\in N-\{Sink\}}\beta_v\cdot\left(\sum_{e\in E}\sum_{l\in L_v^+}M_{el}+\sum_{e\in E}\sum_{l\in L_v^-}M_{el}\right)$$

$$+\sum_{v\in N-\{Sink\}}\gamma_v\cdot\left(\sum_{e\in E}\sum_{l\in L_v^+}M_{el}\cdot E_{sent}+\sum_{e\in E}\sum_{l\in L_v^-}M_{el}\cdot E_{recv}\right)$$

$$
\begin{aligned}
&= \min -\sum_{e\in E}\sum_{v\in N-\{Sink\}}\sum_{l\in L_v^+}\alpha_{el}\cdot M_{el} + \sum_{e\in E}\sum_{v\in N-\{Sink\}}\sum_{l\in L_v^+}\left(\beta_v+\gamma_v\cdot E_{sent}\right)\cdot M_{el} \\
&\qquad + \sum_{e\in E}\sum_{v\in N-\{Sink\}}\sum_{l\in L_v^-}\left(\beta_v+\gamma_v\cdot E_{recv}\right)\cdot M_{el} \\
&= \min \sum_{e\in E}\sum_{v\in N-\{Sink\}}\sum_{l\in L_v^+}\left(-\alpha_{el}+\beta_v+\gamma_v\cdot E_{sent}\right)\cdot M_{el} + \sum_{e\in E}\sum_{v\in N-\{Sink\}}\sum_{l\in L_v^-}\left(\beta_v+\gamma_v\cdot E_{recv}\right)\cdot M_{el}
\end{aligned}
$$

subject to:

$$\sum_{e\in E} M_{el} \le C \qquad \forall l \in L \qquad (5)$$

$$0 \le M_{el} \le |A_e|\cdot f_e \qquad \forall e \in E, \forall l \in L \qquad (7)$$

To solve this subproblem we first need to determine all the coefficient values of $\{M_{el}\}$. For each event e, by putting the start point sensor node of link l into the first term of this subproblem's objective function and then putting the end point sensor node of link l into the second term of this subproblem's objective function, we can obtain the coefficient value of M_{el} on link l. Thus, we can calculate all the coefficient values of $\{M_{el}\}$. We can then determine $\{M_{el}\}$'s values to obtain the optimal solution to this subproblem. If a M_{el}'s coefficient value is positive, we set its value to 0, otherwise we sort the M_{el}s by their coefficient values from small to large. Later, we will determine these M_{el}'s values in this order, and to further satisfy constraint (5), i.e. $\sum_{e\in E} M_{el} \le C_l$ for each link l, we determine the M_{el}'s value one by one. Before determining each of them, we should check the residual capacity on link l. If it is greater than or equal to $|A_e|\cdot f_e$, we set its value to $|A_e|\cdot f_e$, then decrease the residual capacity on link l by $|A_e|\cdot f_e$, otherwise we set its value to the remaining residual capacity on link l, and decrease the residual capacity on link l to 0.

Subproblem 2-4 (related to decision variable T_v)

$$Z_{sub2-4}(\gamma,\theta)$$
$$= \min - \sum_{v\in N-\{Sink\}} \gamma_v \cdot \frac{E_v}{T_v} - \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\sum_{v\in V_p} \theta_{enpv}\cdot T_v$$
$$= \max \sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n}\sum_{v\in V_p} \theta_{enpv}\cdot T_v + \sum_{v\in N-\{Sink\}} \gamma_v \cdot \frac{E_v}{T_v}$$

subject to:

$$\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon} \le T_v \le \frac{E_v}{\sum_{e\in B_v} f_e\cdot E_{sent}+\varepsilon} \quad \forall v\in N-\{Sink\} \quad (9)$$

To solve this subproblem, we need to obtain T_v's maximal value for each sensor node v. This occurs at an intersection point on either side of T_v's bound. Either $\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}$ or $\frac{E_v}{\sum_{e\in B_v} f_e\cdot E_{sent}+\varepsilon}$ can obtain T_v's maximal value for each sensor node v.

3.3.3 The Dual Problem and the Subgradient Method

According to the weak Lagrangian duality theorem [6], for any β, θ, μ, ω ≥ 0, $Z_{D2}(\alpha,\beta,\gamma,\theta,\mu,\omega)$ is a lower bound on Z_{IP5}. The following dual problem (D2) is then constructed to calculate the tightest lower bound.

Dual Problem (D2):

$$Z_{D2} = \max Z_{d2}(\alpha,\beta,\gamma,\theta,\mu,\omega) \quad \text{(D2)}$$

subject to:

$$\beta,\theta,\mu,\omega \geq 0$$

The most popular method for solving the dual problem is the subgradient method [9]. Let g be a subgradient of $Z_{d2}(\alpha,\beta,\gamma,\theta,\mu,\omega)$. Then, at the k th iteration of the subgradient optimization procedure, the multiplier vector $\pi = (\alpha,\beta,\gamma,\theta,\mu,\omega)$ is updated by $\pi^{k+1} = \pi^k + t^k g^k$. The step size t^k is determined by $t^k = \lambda_k \frac{Z_{IP5}^* - Z_{d2}(\pi_k)}{\left\| g^k \right\|^2}$. Z_{IP5}^* is the best upper bound on the primal objective function value after the k th iteration obtained from heuristic solutions. λ_k is a constant between 0 and 2.

Chapter 4 Getting Primal Feasible Solutions

By using Lagrangean Relaxation and the subgradient method as our tools to solve these subproblems, we not only obtain a theoretical lower bound of primal feasible solution, but also get some hints to help us find our primal feasible solution in each iteration of solving the dual problem.

Since some constraints of the primal optimization problem are relaxed by Lagrangean Relaxation, we cannot guarantee that the results of dual problems will be a feasible solution to the primal problem. If the decision variables calculated satisfy the relaxed constraints, then a primal feasible solution is found. Otherwise, a modification to such infeasible primal solutions is necessary to obtain primal feasible solutions.

Therefore, it is necessary to apply additional heuristics to obtain a primal feasible solution. We now give the details of the heuristics..

4.1 Heuristic for Model 1

To find primal feasible solutions for Model 1, solutions to the Lagrangean Relaxation problems are considered.

4.1.1 Reroute Heuristic

$\{x_{enp}\}$, which describe the routing of the entire network, is the most important factor in finding primal feasible solutions. Once $\{x_{enp}\}$ are determined, $\{M_{el}\}$ can be calculated, and T can also be obtained. The solution set of $\{x_{enp}\}$ obtained from Subproblem 1-2 may not be a feasible solution to the primal problem. We can make this infeasible solution become feasible by designing a reroute heuristic. The steps of the reroute heuristic are as follows:

1. Based on $\{x_{enp}\}$, we can calculate $\{M_{el}\}$, then, check whether there are existing links which violate the link capacity. If so, these links are sorted according to the amount of "exceeding flow" (aggregate flow － link capacity), to find the most congested link, i.e. the link with the largest amount of "exceeding flow". If not, go to 7.
2. We analyze the flow on the most congested link to find which events the flow is coming from. Then sort these events by the values of $\alpha_{el} \cdot f_e$ in Subproblem 1-2 and find which event has the largest value of $\alpha_{el} \cdot f_e$. There are probably several selected paths from the event with the largest value of $\alpha_{el} \cdot f_e$ passing through the most congested link. We remove the selected path which has the shortest lifetime sensor node on it, and decrease

the flow from each link on that path by f_e.

3. We then reroute by using the Bellman-Ford algorithm. The link cost of the congested link and the link which has the existing total flow on it, plus the f_e, are larger than its link capacity are set to infinite; other links' costs equal the existing total flow on them. By setting link cost in this manner, we can guarantee that the rerouted path doesn't make an additional link that violates the link capacity, and thereby achieve the goal of effectively balancing the flow on the entire network. Then check if the result of the Bellman-Ford algorithm can successfully find a new path to replace the removed path. If so, we update $\{x_{enp}\}$, and return to 1.

4. So far, $\{x_{enp}\}$ only satisfy the link capacity, but $\{x_{enp}\}$ still have a probability to violate the nodal capacity. So we first check whether the $\{x_{enp}\}$ can result in a node that violates its nodal capacity. If so, go to 5; otherwise, we do find a feasible solution to the primal problem. Then we can further calculate $\{M_{el}\}$ and obtain T by the feasible $\{x_{enp}\}$.

5. We find the most congested node which has the largest amount of "exceeding flow" (aggregate flow − nodal capacity) among all nodes. And analyze the flow on the most congested node to find which events the flow is coming from. Then find the event with the largest value of f_e, and randomly pick one selected path that passes through the congested node and remove it.

6. We then use the Bellman-Ford algorithm to reroute the removed path, in the same way that the link cost is set in Step 5. Further, to avoid the rerouted path causing another node to violate its nodal capacity, we check each node

to see whether its existing flow on it, plus f_e, is larger than its nodal capacity. If so, we set all its in-links' costs to infinite. Then Check if the result of the Bellman-Ford algorithm can successfully find a new path to replace the removed path. If so, we update $\{x_{enp}\}$, then return to 4.

4.2 Heuristic for Model 2

To find primal feasible solutions for Model 2, solutions to the Lagrangean Relaxation problems are considered.

4.2.1 Reroute Heuristic

The solution set of $\{x_{enp}\}$ obtained from Subproblem 2-2 may not be a feasible solution to the primal problem. We can make this infeasible solution become feasible by designing a reroute heuristic. The steps of the reroute heuristic are as follows:

1. Based on $\{x_{enp}\}$, we can calculate $\{M_{el}\}$, then, check whether there are existing links which violate the link capacity. If so, these links are sorted according to the amount of "exceeding flow" (aggregate flow － link capacity), to find the most congested link, i.e. the link with the largest amount of "exceeding flow". If not, go to 7.
2. We analyze the flow on the most congested link to find which events the flow is coming from. Then sort these events by the values of $\alpha_{el} \cdot f_e$ in Subproblem 2-2 and find which event has the largest value of $\alpha_{el} \cdot f_e$.

There are probably several selected paths from the event with the largest value of $\alpha_{el} \cdot f_e$ passing through the most congested link. We remove the selected path which has the shortest lifetime sensor node on it, and decrease the flow from each link on that path by f_e.

3. Then we check the remaining $W-1$ paths. If the path can satisfy the link capacity constraint of each link on it after adding e f_e to it, we put it into a "feasible candidate paths" set; otherwise we omit it.
4. We choose the path which has the largest value of S_{enp} in the "feasible candidate paths" set to be the new selected path, then update $\{x_{enp}\}$ and return to 1.
5. So far, $\{x_{enp}\}$ only satisfy the link capacity, but $\{x_{enp}\}$ still have a probability to violate the nodal capacity. So we first check the $\{x_{enp}\}$ whether can result in a node that violates its nodal capacity. If so, go to 6, otherwise, we do find a feasible solution to the primal problem. Then we can further calculate $\{M_{el}\}$ and obtain T by the feasible $\{x_{enp}\}$.
6. We find the most congested node which has the largest amount of "exceeding flow" (aggregate flow － nodal capacity) among all nodes. And analyze the flow on the most congested node to find which events the flow is coming from. Then find the event with the largest value of f_e, and randomly pick one selected path that passes through the congested node and remove it.
7. Then we check remaining $W-1$ paths individually. If the path can satisfy the link and the nodal capacity constraints of each link and sensor node on it after adding f_e to it, we put it into a "feasible candidate paths" set;

otherwise we omit it.

8. We choose the path that has the largest value of S_{enp} in the "feasible candidate paths" to be the new selected path, then update $\{x_{enp}\}$ and return to 5.

Chapter 5 Computational Experiments

In order to prove that our heuristics are efficient, we implement two simple algorithms to compare with our heuristics.

5.1 Simple Algorithm for Model 1 (SA 1)

We sort events from large to small according to their f_e values. Then, in this order, we select a path for each sensor node individually that can sense the sorted event. We can select the paths, i.e. determine the values of $\{x_{enp}\}$, by using the Bellman-Ford algorithm, where the link cost of each link l is the inverse of its end point sensor node's residual energy. Once we select a path for a sensor node, we add flow on the selected path and update the relevant links' costs. Before we select a path for another sensor node, however, we check if each link l can satisfy the link capacity constraint after adding extra f_e to it. If not, we set its link cost to infinite. We further check if each sensor node can satisfy the nodal capacity constraint after adding extra $2 \cdot f_e$ to it (flow-in + flow-out); otherwise, we set all its in-links' costs to infinite.

5.2 Simple Algorithm for Model 2 (SA 2)

We select a path for a sensor node by checking if path p has the largest value of S_{enp} within the predefined W candidate paths for the sensor node. If so, we set its x_{enp} value to 1; otherwise we set it to 0. Once we select a path for a sensor node, we add flow on the selected path and update the relevant S_{enp} value. We then select a path for another sensor node.

5.3 Parameters and Scenarios of the Experiment

To test the performance of these algorithms, we design three experimental scenarios: (1) 25 events and 25 sensor nodes; (2) 25 events and 50 sensor nodes; (3) 25 events and 100 sensor nodes. All events and sensor nodes are randomized in a 50m $\times$ 50m sensor field, and the Sink is located 5m away under the sensor field. The sensing radius and communication radius of an individual sensor node are 10m and 20m respectively. These parameters and scenarios for Model 1 and Model 2 used in our experiments are listed in Table 5-1 and Table 5-2 below respectively.

Parameters	Values
Number Of Iterations	Firs stage: 2000 Other stages: 1000
Non-improvement Counter	Scenario 1: 50

	Scenario 2: 70 Scenario 3: 90
Initial Upper Bound	SA 1
Initial Scalar Step Size	2
Initial Multipliers	0
f_e (kbit/sec)	(0,2]
C_l (kbit/sec)	20
U_n (kbit/sec)	100
H_e (hop)	5
E_n (J)	1
E_{sent} (J/kbit)	0.9×10^{-4}
E_{recv} (J/kbit)	0.5×10^{-4}
ε (J/sec)	0.1×10^{-4}

Table 5-1 Experimental Parameters for Model 1

Parameters	Values
Number Of Iterations	1000
Non-improvement Counter	100
Initial Upper Bound	SA 2
Initial Scalar Step Size	2
Initial Multipliers	0
f_e (kbit/sec)	(0,2]
C_l (kbit/sec)	20
U_v (kbit/sec)	100

H_e (hop)	5
E_v (J)	1
E_{sent} (J/kbit)	0.9×10^{-4}
E_{recv} (J/kbit)	0.5×10^{-4}
ε (J/sec)	0.1×10^{-4}
W (path)	Scenario 1: 5 Scenario 2: 10
K	1
O	9

Table 5-2 Experimental Parameters for Model 2

The initial upper bounds for Models 1 and 2 are set according to the result of each model's respective simple algorithm. If the simple algorithm for Model 1 can not find a feasible solution to the primal problem, we set it to 0. Meanwhile, if the simple algorithm for Model 2 can not find a feasible solution to the primal problem, we set it to $-T+\sum_{e\in E} K\cdot R_e-\sum_{e\in E}\sum_{n\in A_e}\sum_{p\in P_n} O\cdot S_{enp}$ where

$$T=\min_{e\in E}\left\{\max_{n\in A_e,\,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)\right]\right\},$$

$$R_e=\max_{n\in A_e,\,p\in P_n}\left[\min_{v\in V_p}\left(\frac{E_v}{\left(f_e+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\varepsilon}\right)\right]$$ for each $e\in E$ and

$$S_{enp}=\min_{v\in V_p}\left(\frac{E_v}{\frac{1}{2}\cdot\left(U_v+\sum_{e\in B_v} f_e\right)\cdot E_{sent}+\frac{1}{2}\cdot\left(U_v-\sum_{e\in B_v} f_e\right)\cdot E_{recv}+\varepsilon}\right)$$ for all $\{S_{enp}\}$. The multipliers are all set to 0 initially.

The $\{f_e\}$ are randomized in $(0,2]$ kbit/sec, then the $\{C_l\}$, $\{U_n\}$(or $\{U_v\}$) and $\{H_e\}$ are 20 kbit/sec, 100 kbit/sec and 5 hops respectively, where the values of $\{C_l\}$ refer to the standard of IEEE802.15.4 (20, 40, 256 kbit/sec). All sensor nodes start with an initial energy of 1J, then E_{sent}, E_{recv} and ε are respectively set to 1.4×10^{-4} J/kbit, 0.5×10^{-4} J/kbit and 0.1×10^{-4} J/sec, where values of E_{sent} and E_{recv} are calculated by adopting a simple model from [7] (see Figure 5-1 and Table 5-3). Thus, to transmit a k bit message a distance d using this radio model, the radio expends: $E_{Tx}(k,d) = E_{elec}\cdot k + \varepsilon_{amp}\cdot k\cdot d^2$, and to receive this message, the radio expends: $E_{Rx}(k) = E_{elec}\cdot k$. In Model 2, the W candidate paths from each sensor node are set to 5 and 10 in Scenario 2 and Scenario 3 respectively, and the constants K and O in the objective function are set to 1 and 9.

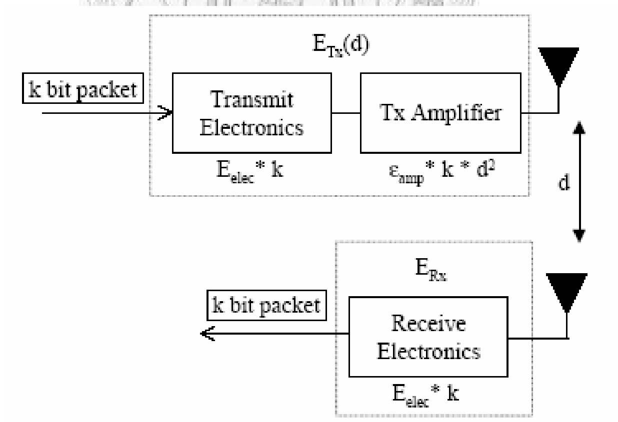

Figure 5-1 First Order Radio Model [7]

Operation	Energy Dissipated
Transmitter Electronics ($E_{Tx-elec}$) Receiver Electronics ($E_{Rx-elec}$) ($E_{Tx-elec} = E_{Rx-elec} = E_{elec}$)	50 nJ/bit
Transmit Amplifier (ϵ_{amp})	100 pJ/bit/m^2

Table 5-3 Radio Characteristics [7]

5.4 Experimental Results

5.4.1 Experimental Results for Model 1

The experimental results obtained from seven randomly generated cases for each scenario for Model 1 are listed below. The SA_1 value means the system lifetime obtained from the first stage of Model 1 using SA 1 and equals the system lifetime of the original definition. The LR_1 value is the optimal system lifetime obtained from the first stage of Model 1 using LR. The gap defined by $(UB - LB) \times 100\%$, where UB and LB is an optimal solution to the primal problem (IP) and dual problem (D), respectively, in the LR process. Gap_1 is the gap obtained from the first stage of Model 1. The improvement ratio of SA_1 is calculated by $(LR_1 - SA_1)/ SA_1$. The LR_T is the system lifetime of our modified definition obtained by using LR, and the system lifetime improvement ratio is calculated by $(LR_T - LR_1)/ LR_1$.

Scenario 1:

Case	SA_1	LR_1	Gap_1	SA_1 Imp. ratio	LR_T	T Imp. ratio
1	385.6	495.2	9.8%	28.4%	496.7	0.3%
2	765.2	1056	33.1%	38.0%	1188	12.5%
3	417.1	702.7	30.5%	68.5%	785.2	11.7%
4	670.6	752.4	20.8%	12.2%	752.4	0.0%
5	408.5	480.9	12.0%	17.7%	491.5	2.2%
6	333.6	624.2	24.3%	87.1%	736	17.9%
7	357.7	501.7	15.4%	40.3%	522.5	4.1%

Table 5-4 Experimental Results of Scenario 1 for Model 1

Scenario 2:

Case	SA_1	LR_1	Gap_1	SA_1 Imp. Ratio	LR_T	T Imp. ratio
1	360.7	500.8	28.6%	38.8%	566.9	13.2%
2	364.1	487.5	28.6%	33.9%	572.1	17.3%
3	363.5	473.3	16.8%	30.2%	515.7	9.0%
4	327.4	413.4	26.3%	26.3%	444.9	7.6%
5	370.4	541.7	19.1%	46.2%	604.6	11.6%
6	378.7	426.6	26%	12.7%	474.9	11.3%
7	358.2	472.4	22.3%	31.9%	495.7	4.9%

Table 5-5 Experimental Results of Scenario 2 for Model 1

Scenario 3:

Case	SA_1	LR_1	Gap_1	SA_1 Imp. Ratio	LR_T	T Imp. ratio
1	335.1	388.5	30.2%	15.9%	476.4	22.6%
2	366.3	397.9	36.0%	8.6%	444.6	11.7%
3	291.0	351.1	13.6%	20.6%	359.6	2.4%
4	355.2	386.8	30.2%	8.9%	461.2	19.2%
5	379.0	428.4	33.8%	13.0%	549.1	28.2%
6	342.7	387.3	31.7%	13.0%	406.9	5.1%
7	376.7	416.6	28.0%	10.6%	489.3	17.5%

Table 5-6 Experimental Results of Scenario 3 for Model 1

5.4.2 Experimental Results for Model 2

The experimental results obtained from seven randomly generated cases for each scenario for Model 2 are listed below. The SA value means the system lifetime obtained from Model 2 by using SA 2, and the LR value is the optimal system lifetime obtained from Model 2 using LR. The Gap defined by $(UB - LB) \times 100\%$, where UB and LB is an optimal solution to the primal problem (IP) and dual problem (D), respectively, in the LR process. The improvement ratio of SA is calculated by (LR- SA)/ SA.

Scenario 2:

Case	SA	LR	Gap	SA Imp. ratio
1	399.7	447.1	82.0%	11.9%
2	0	383.9	80.1%	∞ %
3	528.8	528.8	78.2%	0.0%
4	625.1	670	76.4%	7.2%
5	401.6	474.4	73.4%	18.1%
6	397.2	461.7	82.2%	16.2%
7	392.7	464.7	83.0%	18.3%

Table 5-7 Experimental Results of Scenario 2 for Model 2

Scenario 3:

Case	SA	LR	Gap	SA Imp. ratio
1	381.4	419.4	76.8%	10.0%
2	365.3	374.7	81.9%	2.5%

3	0	426.3	82.7%	∞%
4	372	404.8	82.9%	8.8%
5	423.1	482.9	79.6%	14.1%
6	500.8	500.8	76.6%	0.0%
7	438.2	466.5	80.9%	6.5%

Table 5-8 Experimental Results of Scenario 3 for Model 2

The average system lifetimes of our modified definition obtained from different scenarios for Models 1 and 2 are listed in Table 5-9 and depicted in Figure 5-2.

	Model 1		Model 2
	LR_1	LR_T	LR
Scenario 2	473.7	525.0	490.1
Scenario 3	393.8	455.3	439.3

Table 5-9 Comparison of Average System Lifetimes of Different Models

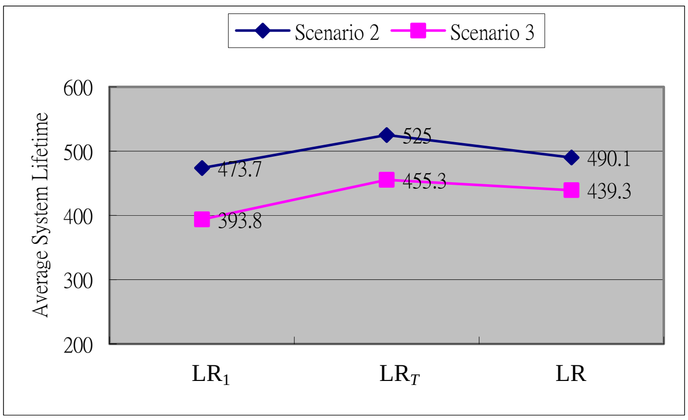

Figure 5-2 Comparison of Average System Lifetimes of Different Scenarios

5.4.3 Computation Time

Model	Average Computation Time (sec)		
	Scenario 1	Scenario 2	Scenario 3
Model 1	178.1	481.7	2388.6
Model 2	X	131.4	1580.1

Table 5-10 Computation Time of Different Models

According to Table 5-10, the computation times increase when the number of sensor nodes scattered in a sensor field increase. Obviously, the computation times in Scenario 3 are much higher than in Scenario 2. This is because, in Model 1, we have to execute more stages when the number of sensor nodes grows. In Model 2, however, we have to use an exhaustive search method to solve Subproblem 2-2. As there are more candidate paths for each sensor node to the Sink once the number of sensor nodes grows, the needed computation time increases rapidly.

5.5 Discussion of Results

According to our experimental results, it is clear that the results of LR are much better than SA. This is because we do not use any informative parameters in SA. In LR method, however, we fully use these informative parameters, which are obtained from solving the subproblems of LR. We use multipliers as the link cost when trying to reroute or find a better path to effectively achieve the purpose of balancing the total flow on the entire network. The values of multipliers are tuned by gradient method iteration by iteration and clearly respond to the situation of an individual link. When a link is over- loaded, the value of that link will increase iteration by iteration, so it will not be a good choice for constructing a path.

In addition, the system lifetime obtained from our modified definition is much longer than the original definition, which is rather underestimated, and the result of Model 1 is better than Model 2.

Further, we can use linear programming relaxation to explain the gap. Because of the integer constraint, there will be a bound between the lower bound and the result of heuristic. We call this the duality gap. If we eliminate the integer constraint, for example, a sensor node can select more than one path to transmit the flow to the Sink, and the total flow on the entire network will be more balanced. In addition, the corresponding calculated system lifetime, i.e. the lower bound, will be higher.

Chapter 6 Summary and Future Work

6.1 Summary

In this paper, our work emphasizes deriving an energy efficient routing algorithm to maximize the system lifetime of wireless sensor networks. We modify the original definition of system lifetime, and thereby obtain a system lifetime that is much closer to the real environment.

We propose two models and formulate both as combinatorial optimization problems where the objective function is to maximize the system lifetime. We then use Lagrangean Relaxation, combined with the subgradient method, to solve these problems. While applying this methodology, we relax some complicated constraints in the primal problem, which makes it much easier to solve. Then, we further decompose this problem (LR) into several independent subproblems. We analyze these subproblems and optimally solve them, and develop several heuristics to obtain the primal feasible solution.

We implement the algorithms in C code, and test them by using three well-designed scenarios. Our experimental results are very satisfactory because, by applying our algorithms, the system lifetime is greatly prolonged.

6.2 Future Work

In this paper, we haven't considered "data aggregation" behavior in our models. To obtain a more precise system lifetime of wireless sensor networks, data aggregation should be further considered.

In addition, because of the uncertainty of wireless communication channels, we can further discuss the fault tolerance issue in wireless sensor networks. For example, we can require sensor nodes to deliver one more copy of a message along different paths to the Sink in order to increase the probability of successful delivery.

Finally, we should focus on the distributed implementation of our proposed algorithms and consider other ways to prolong system lifetime, such as improving the sensor node so that it can adaptively change its sensing radius and communication radius. Another improvement would be to make the sensor node mobile.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, Vol. 40, no. 8, pp. 102-114, Aug. 2002.

[2] M. Bhardwaj and A. P. Chandrakasan, "Bounding the lifetime of sensor networks via optimal role assignments," *Proc. of INFOCOM 2002*, Vol. 3, pp. 1587-1596, June 2002.

[3] J. H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," *Proc. of INFOCOM 2000*, Vol. 1, pp. 22-31, March 2000.

[4] E. J. Duarte-Melo and M. Liu, "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks," *Proc. of GLOBECOM '02*, Vol. 1, pp. 21-25, Nov. 2002.

[5] M. L. Fisher, "The Lagrangian relaxation method for solving integer programming problems", *Management Science*, vol. 27, pp.1-18, 1981.

[6] A. M. Geoffrion, "Lagrangean relaxation and its use in integer programming", *Math. Programming Study*, vol. 2, pp.82-114, 1974.

[7] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," *Proc. of the 33rd Annual Hawaii International Conference on System Sciences*, pp. 3005-3014, Jan. 2000.

[8] W. R. Heinzelman, J. Kulik and H. Balakrishnan, "Adaptive Protocols for Information Dissemination in Wireless Sensor Networks," *Proc. of MobiCom'99*, Aug. 1999.

[9] M. Held, *et al.*, "Validation of subgradient optimization," *Math. Programming*, vol. 6, pp. 62-88, 1974.

[10] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," *Proc. of the 6th annual ACM/IEEE Conference on Mobile Computing and Networking (MOBICOM)*, pp.56-67, Aug. 2000.

[11] A. Manjeshwar and D. P. Agrawal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," *Proc. of 15th International Parallel and Distributed Processing Symposium*, pp. 2009-2015, April 2001.

[12] A. Manjeshwar and D. P. Agrawal, "APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks," *Proc. International of Parallel and Distributed Processing Symposium (IPDPS 2002)*, pp.195-202, April 2002.

[13] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *ACM Communications*, Vol. 43, no. 5, pp.51-58, May 2000.

[14] V. Raghunathan, C. Schurgers, S. Park and M. B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing Magazine*, Vol. 19, no. 2, pp. 40-50, March 2002.

[15] C. Schurgers and M. B. Srivastava, "Energy efficient routing in wireless sensor networks," *IEEE Military Communications Conference (MILCOM 2001)*, Vol. 1, pp. 357-361, Oct. 2001.

[16] M. A. M. Vieira, D. C. da Silva Jr., C. N. Coelho. Jr. and J. M. da Mata, "Survey on wireless sensor network devices," *Proc. of ETFA '03*, Vol. 1, pp. 537-544, Sept. 2003.

[17] J. Zhu and S. Papavassiliou, "On the energy-efficient organization and the

lifetime of multi-hop sensor networks," *IEEE Communications Letters*, Vol. 7, no. 11, pp. 537-539, Nov. 2003.

[18] G. Zussman and A. Segall, "Energy efficient routing in ad hoc disaster recovery networks," *Proc. of INFOCOM 2003*, Vol. 1, pp. 682-691, March/April 2003.

簡　　歷

姓　　名：葉耿宏

出 生 地：台灣省台南市

出 生 日：中華民國六十九年六月二十五日

學　　歷：八十七年九月至九十一年六月

國立中央大學資訊管理學系

九十一年九月至九十三年七月

國立台灣大學資訊管理學研究所

地　　址：桃園市經國路 601 號 17 樓之二