SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

BACHELOR'S THESIS

论文题目：移动互联网的多播树构造

学生姓名：龚虹宇

学生学号：5110309759

专　　业：信息工程

指导教师：王新兵

学院(系)：电子信息与电气工程学院

上海交通大学学位论文

摘　要

组播是网络中信息的一种重要传输方式。随着移动服务的兴起，移动互联网的组播支撑了无线用户的大量服务需求。传输延迟和能量消耗是网络传输的两项重要的性能指标，延迟很大程度上决定了通信的质量，而节能是绿色科技的核心。已有的研究提出了很多方法来优化其中的一项指标，但是至今我们仍不清楚延迟和能量效率之间的潜在关系。同时，很多研究是在一些具体的应用中提高信息传输的能量效率，但是他们并没有明确在给定条件下可以实现的最小能量消耗，并且也没有给出通用的实现最小能耗的方法。实际中的传输策略的能耗和最优的能耗间的差距也不清楚。在这篇论文中我们聚焦移动互联网，探索对于传输延迟和能量消耗的可行域的一个根本的理解。我们发现在组播中可以以传输延迟为代价节省能量消耗，然后从延迟和能耗的均衡角度来展示网络传输的设计空间。这种均衡也使得我们对于当前设计下组播传输的性能有了更加深入的理解，包括特定限制条件下的能耗目标实现可能性、绿色传输策略的设计等方面。无论是在传输半径可调还是不可调的条件下，我们严格证明了实现最小能耗的组播是一个 NP 难问题。但是，我们通过构造组合优化问题，能够得出最小能耗的理论上界与下界，并且能够在多项式时间内找到一个接近最优能耗的可行组播传输策略。

关键词： 移动网络　延迟　能量

Multicast Tree in Mobile Internet

ABSTRACT

Multicasting is an important transmission patten in networks. With the popularity of mobile services, multicasting in mobile Internet supports large demands of wireless users. Two important performance metrics are transmission delay and energy consumption, where delay determines the communication quality, and energy saving is the core of green technology. Existing works have proposed various methods to optimize the single metric, but up till now we are still unware of the underlying relationship between delay and energy efficiency. Moreover, many researches focus on improving energy efficiency in specific applications, but they do not make the minimum possible energy consumption clear under given conditions. It is also unknown how the minimum energy consumption under given conditions can be achieved and what's the gap of current deployment from the optimality. In this paper, we explore a fundamental understanding on the feasible region of delay and energy efficiency in mobile Internet. We point out that the energy saving in mobile Internet is often at the cost of transmission delay, and present the design space in terms of a dedicated tradeoff between delay and energy consumption. This tradeoff provides us with insights into judging the efficiency of current deployment including transmission policies, the possibility of implementation under specific constraints, designing green scheduling schemes, etc.. We show that it is NP-hard to achieve and even compute the boundary of minimal energy consumption whether the transmission range is fixed or variable. However, we can derive both upper and lower bound of optimal energy consumption by formulating optimization problems. Considering the practical feasibility, we also propose an efficient approximate algorithm with low computational cost but high accuracy with Lagrangian relaxation method.

KEY WORDS: Mobile networks, delay, energy saving

Content

Figure Index

Table Index

Symbols

S	source
D	group of destinations
H	hop or relay node
n	network size
m	the number of destinations
r	transmission range
d	delay constraint
e	edge in the dynamic graph
E	energy consumption in multicasting
J	journey in the dynamic graph

Chapter 1 Introduction and Related Work

1.1 introduction

Mobile Internet consists of many mobile and wireless terminals such as mobile phones and mobile sensors. The mobile part of the Internet often connects physical objects to the Internet, bridging the gap between cyber and physical space. Mobile objects are pervasively distributed, collecting data and impacting the physical environment. Many of these mobile objects are battery-powered, and thus have limited energy supply, which imposes strict constraints on the power consumption of these devices. Battery capacities increase slowly by only 5%-8% per year for given weight, while the increasing complexity of devices indicates a larger demand of energy [1]. Worse still, as the network becomes large scale, energy saving becomes even critical so as to guarantee the system scalability. Energy saving can be addressed in terms of energy harvesting, conservation and consumption [2]. In this paper, I focus on reducing the energy consumption in a wireless network system where nodes can be mobile. Considering that static networks can be regarded as a special case of mobile networks, the analysis can also be applied to the static networks. Multicasting is a one-to-many transmission pattern, which plays a critical role in information dissemination among a group of nodes in the network. Delay and energy consumption are two important measurements of multicasting performance. Existing works mainly focus on specific applications, trying to develop corresponding routing methodologies or algorithms to reduce the energy consumption and extend the lifetime of the devices. However, up till now we are still unaware of a number of fundamental understandings of multicasting performance. For example, we are curious to answer:

- What is the minimum possible energy consumption under given conditions in multicasting?
- How far is a current deployment away from the optimality in terms of energy saving?
- Given certain technology requirements and energy consumption constraints, are we able to design and implement such a system?

To answer the above questions, it is very important for us to understand the feasible region of energy saving, or the boundary of the design space, of mobile network systems. In such systems, energy consumption is mainly caused by data transmission among various nodes. Transmission refers to the process that one node forwards one packet to another node.

If transmission power of each node is fixed and equivalent, the times of transmissions processed can be used to estimate the energy consumed for communication. Intuitively, if we can reduce the times of transmissions, we can save energy; meanwhile, wireless interference caused by redundant transmissions also reduces. Therefore, the key question to saving energy in mobile networks is how to reduce the times of data transmission while guaranteeing data transmission metrics required by specific applications.

One important metric in multicasting is the transmission delay. Formally, we refer to transmission delay as the time needed for transmitting data from a sender to the destination. Energy saving is sometimes satisfied at the cost of transmission delay. For example, source is initially far away from the destination. The source can either wait for a long time to forward the message directly to the destination or engage several relays to pass on the message in a shorter time. Naturally, we can represent the design space in terms of the tradeoff between energy saving and transmission delay. In real systems, requirements on the delay vary a lot. Real-time applications like voice chat or video conference requires strict and tight constraints on delay, while delay-tolerant services like file transfer may allow longer delay. Therefore, we are interested in building up a formal framework for analyzing the relationship between energy consumption and delay, so as to give insights to designing various kinds of applications. In addition, this understanding also provides us with guidelines to design efficient scheduling policies that minimize the energy consumption under given conditions.

In this paper, we formulate this tradeoff as the Minimum Energy Problem in Multicasting (MEPM), aiming to design a scheduling policy under given delay constrains, which achieves minimum energy consumption to support one-to-many transmission. Unfortunately, we prove it is NP-hard whether the transmission range can be adjusted by nodes with the reduction from minimum-weight set cover problem to this problem. We first give the mathematical expression of the Minimum Energy Problem in Multicasting, and then apply Lagrangian relaxation method as a fast approximation. Our algorithm can achieve a near-optimal scheduling policy with low time complexity. Through extensive simulations we show that its performance gap from optimality is limited. Our contributions are:

- We construct a new model to capture the topology of mobile networks in both spatial and temporal dimensions.
- We represent the design space of mobile networks by the tradeoff between transmission delay and energy consumption, and formulate Minimum Energy Problem in Multicasting to capture this tradeoff.
- We prove the NP-hardness for optimal scheduling for MEPM, give an efficient ap-

proximate solution and show its high accuracy by simulations.

The rest of this paper is organized as follows: Section 1.2 introduces current research results related with our work. Chapter 2 presents network, mobility and transmission models, laying the basis of theoretical analysis. Both minimal delay journeys and delay bounded journeys are discussed in Chapter 3. Minimum Energy Problem in Multicasting is formulated in Chapter 4, the case of fixed transmission range is considered in Section 4.1, and the range is allowed to be adjusted to further reduce energy consumption in Section 4.2. The theoretical analysis is evaluated by extensive simulations in Chapter 5, and this paper is concluded with a brief summary.

1.2 Related Work

1.2.1 Multicasting in Mobile Networks

Multicasting can be used in many applications, such as smart grid and multimedia. According to Li *et al.* [3], multicast can be used in Distributed Energy Generator(DEG) voltage control of smart grid, where multicast routing is needed. Li, Dan *et al.* points out that Multicast improves group communications in reducing network traffic and increasing application throughput, both of which are very important in data center networks [4]. Mobile video or mobile TV, is expected to become a popular application for wireless network operators [5]. Video multicast services are commonly provided in these applications. Multicast concepts are designed to be a scalable and efficient solution for local communication needs such as file transfer and even streaming services in the cellular networks into which the device-to-device communication is integrated [6]. Multicasting helps support a seamless multimedia service such as live-video service [7].

1.2.2 Delay and Energy Efficiency

Delay is an important measure of network performance. To our knowledge, most researchers have focused on decreasing the delay. Yang *et al.* show that packet replication technique can improve the multicast delay performance, and they further study the real achievable multicast delay performance in MANETs under a general two-hop relay routing [8].

When the delay constraint is not so tight for network applications, researchers focus their attention on Delay Tolerant Networks (DTN). Wenrui Zhao, *et al.* study multicasting in delay tolerant networks, proposing a new multicast semantic models for DTN environments [9]. Multiple multicast routing algorithm with different routing strategies are developed under the constraints on group membership and data delivery. Christopher M.

Sadler, *et al.* also conduct research on delay tolerant networks, devising computationally-efficient lossless compression algorithms for data transfer [10]. Redundant data is removed from transmission, and significant energy improvements are made. The design issues are discussed about the energy implications on delay tolerant networks.

A store-carry-and-forward (SCF) scheme is came up with to forward data based on its delay sensitivity in relay-engaged cellular networks [11]. Experiments in this work show that a factor more than 30 in energy efficiency can be achieved with the SCF scheme and delay relaxation [11]. It is implied that delay tolerance should be provided for data application so as to achieve the system scalability when the wireless nodes are energy-constrained in networks [12].

With high energy consumption characteristics of future network and slow development of battery capacity, it has always been a hot topic about reaching a higher energy efficiency in many aspects of research fields. Chou *et al.* study the maximum energy-efficient multicast scheduling (MEMS) problem, and they propose genetic algorithm-based multicast scheduling for the approximate solution to this problem [13]. Niu proposes a new framework called traffic-aware network planning and green operation (TANGO) for GREEN networks, aimed at increasing the energy efficiency from the perspective of system, while guaranteeing coverage and optimizing radio resources as well [14]. Some key technologies for the migration to TANGO includes: from always on to always available with BS sleep control; from static to dynamic cell planning with cell zooming; from uniformed to differentiated services with delay-energy trade-off.

Raffaele Bolla, *et al.* provide a twofold contribution to green networking [15]. Their paper explores current data and studies power consumption for next generation networks. It also provides an up-to-date survey on the current state-of-the-art technologies in energy efficiency for static networks, which are regarded as remarkable improvements that can be introduced into today's networking equipment and the future Internet.

Kerry Hinton, *et al.* provide an overview of a network-based model for power consumption in Internet infrastructure [16]. They also present three effective strategies to improve the energy efficiency: requiring equipment to reduce its power consumption when not in use, reducing the processing rate of a device when its work load is low, and improving the energy efficiency of core routers.

1.2.3 Tradeoff between delay and energy

Delay and energy efficiency are both critical issues in wireless networks, but they can't reach optimality most of the time. It is expected that the transmission is reliable with differential delay. Moreover, tight energy consumption constraints are imposed on wireless network where many devices are battery-powered. Delay-bounded Energy-constrained Adaptive Routing (DEAR) problem is extensively studied to satisfy these requirements in wireless communication. Shi Bai *et al.* presented a pseudo-polynomial-time solution to solve a special case of DEAR. They represent the delay as integers and an $(1+a)$-approximation algorithm is proposed to solve the optimization version of the DEAR problem [17]. However, the wireless network model in their works is static and it has pre-existing topology. Mobility brings about more uncertainty to analysis, adding difficulty to the study of delay and energy efficiency in mobile networks.

Andrea J. Goldsmith *et al.* have reported the biggest design challenges in Wireless Ad Hoc Networks are the lack of centralized control, limited node capability, and variability of the links and network topology [18]. But they did not propose a suitable algorithm or solution to figure out such kind of troubles.

Shah *et al.* argue that it is impossible to design a policy with low complexity, high throughput, and low delay for a general network [19]. Specifically, they establish that for certain network models, any policy with favorable delay properties and modest throughput requirements must implicitly solve a computationally hard problem. This implies that under certain widely believed computational hypotheses, any such policy will have high computational requirements. Their study gives us the motivation to study the tradeoff between different performance aspects.

1.2.4 Mobility Prediction

In dynamic networks, the topology is time-varying and at most of the time not known. The indeterminacy of the dynamic network topology imposes more difficulty and complexity to our analysis of the data transmission scheduling. Prediction of motion information of wireless nodes plays a critical role in improving many aspects of the performance in wireless networks. There are already many mature researches on the prediction of node movement, and the accuracy of prediction is shown to be quite accurate. It is assumed in our work that node mobility can be predicted, and we present some techniques on mobility prediction in order to show that our assumption is reasonable.

Amnir Hadachi *et al.* propose an application of an enhanced Markov algorithm in order to predict the mobility of the mobile phone users in cellular network [20].The authors first collect passive location data from the real-world scenarios, which can be divided two types: stream data and call detail records data. Then, they designed a system architecture to process the collected data and to train the model based on trajectories of users. In the mobility prediction algorithm, they use a Markov process which consists of two components: GPA (Global Prediction Algorithm) for users who exist in the training database and

LPA (Local Prediction Algorithm) for users who don't. These two algorithms are modified through the testing of two different association rule: the universal behavior rule and the temporal rule of the users mobility patterns. The latter one adds a time window (also called time restriction) to the algorithm. The two algorithms work together in the following order: if the GPA fails, then the LPA takes over. As a result, they create four other algorithms. Using these algorithms, they can predict the next location of all the users in a cell, including users who do not exist in the training database. The approach can reach an accurate prediction rate up to 95.67%.

Meghanathan discusses a new location prediction based routing (LPBR) protocol for mobile ad-hoc networks (MANETs), and extends the research to the area of multicasting and multi-path routing [21]. The node location prediction is made by the means of Location Update Vector, which is collected by destination node upon every flooding-based route discovery. When the process to discover the minimum-hop route fails through flooding, the destination node runs a minimum-hop path algorithm using the predicted topology. If the predicted minimum-hop route exists in reality, there is no need for expensive flooding-based route discovery. The extensions of LPBR(referred to as NR-MLPBR) to multicasting simultaneously minimizes the number of global broadcast tree discoveries as well as the hop count on every source-destination path.

Kusy *et al.* propose a routing protocol which does not require direct measurements of node location, yet the protocol can predict the best choice of the next relay node for a mobile sink. Such predication can be made several seconds in advance of the optimal transition time with accuracy up to 90% [22]. Gambs *et al.* predict future location of an individual by the means of the observation of one's movement behavior and the recent locations [23]. They extend Mobility Markov Chain (MMC) model to incorporate n previous visited locations, and develop a new algorithm for this model that they called it as n-MMC. Their prediction accuracy ranges from 70% to 95% when $n = 2$. It does not seem to improve the accuracy significantly at the cost of larger overhead when it holds that $n > 2$.

Tran The Son, *et al.* introduced a non GPS-based model to infer the mobility of network nodes. The relative velocity of a node itself is also predicted in Mobile Ad hoc Networks. Their algorithm is based on node degree and the Average Encounter Rate (AER) which can be observed and calculated locally at each node [24]. The proposed model uses distributed algorithm which is independent of the mobility pattern, and thus can be deployed on any routing protocol type. The outcome of prediction is still recorded at high accuracy, and the prediction error is fewer than 10% when nodes are moving at high speeds.

S.M.Mousavi *et al.* propose a protocol in which a node can predict its own future position together with some parameters like future distance between two neighboring nodes [25]. They use an enhanced future distance predictor which adaptively produces the coefficients of a specified estimator using learning automaton. The estimator can even adapt itself to different mobility models, speeds and sampling rates. Using this enhanced predictor in the proposed protocol, they can achieve lower power consumption and lower interference cost in Media Access layer and therefore improve capacity of the network.

Nermin Makhlouf, *et al.* give a algorithm to predict the future movement of wireless nodes in terms of the direction and speed, on the basis of the status of current network [26]. The observation is based on Global Positioning System and directional antennas with no overlapping directions. One node can estimate the position of its destination in the next time slot. It is achieved by using the similarity of triangles, and the current position information of nodes is also applied to the prediction. The prediction accuracy would be increased if more information of the nodes movement pattern was known. In fact high accuracy of mobility prediction depends on complete awareness of the network topology ad mobility behavior of nodes.

Fraser Cadger, *et al.* adopt a new location predication method in Mobile Ad hoc Networks, applying machine learning algorithms such as Decision Trees, Artificial Neural Network, and Support Vector Regression [27]. Different configurations of the algorithms would yield different results. They used the algorithms to predict the environment of the nodes and the future location of mobile devices in the networks.

Chapter 2 Analysis Model

In this chapter, we present network model, mobility model and transmission model to characterize the mobile networks, which is the basis of our analysis in the network delay and multicasting energy consumption in the following chapters.

2.1 Network Model

We model the mobile devices in the wireless system as a group of nodes distributed in a region, and they can interact with each other. Assume that the transmission range of nodes in the network is r. Two nodes u and v can communicate with each other directly if and only if the Euclidean distance between them is no larger than r. However, direct communication does not necessarily guarantee successful data transmission. Given that two nodes are within the transmission range, extra time is needed before initiating data transmission, since it takes time for them to identify each other, prepare data to be transmitted, etc. We call this period of time the *preparation time*, denoted as $\triangle T$. Besides, it also takes time for messages to be received by the receiver, and we call it *transmission time* denoted as T_t. We define the processing time T_p as

$$T_p = \triangle T + T_t. \tag{2–1}$$

If the time interval during which two nodes can communicate directly is less than T_p, then the message cannot be sent completely before they move out of the transmission range. Therefore, this transmission fails. Therefore, the transmission between two nodes is successful only if the time interval when they stay within the transmission range is long enough, at least longer than T_p.

Let V denote the set of all nodes in this time-varying network, and $E \subseteq V \times V$ be the set of possible edges between vertices in V. Since the nodes are moving around and the edges between nodes may appear or disappear over time, we use a *time-varying graph* to describe the dynamics of the network topology. An edge between two nodes exists over the time interval $[t_1, t_2]$ if and only if the messages from one node can be received successfully by the other one during this interval, i.e., the following conditions are satisfied:

$$\begin{cases} l(t) \le r, \forall t \in [t_1, t_2]. \\ T_p \le t_2 - t_1. \end{cases} \tag{2–2}$$

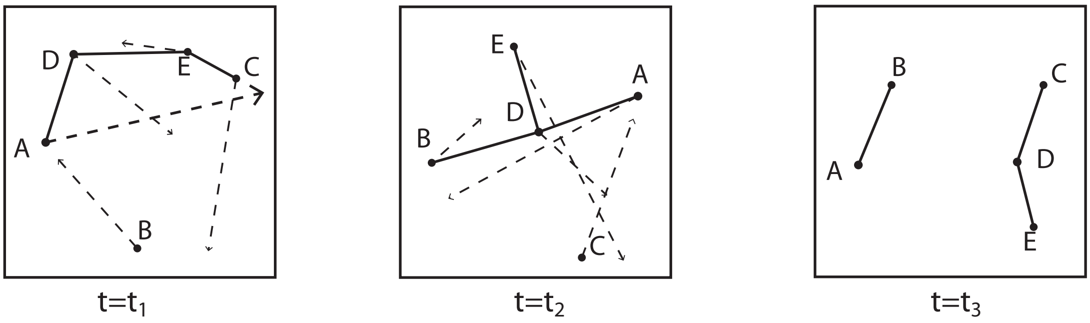


Figure 2–1 Snapshots of a time-varying graph

A time-varying graph is denoted as a tuple $G = \{V, E, \tau, \rho\}$, where τ is the time when we take a snapshot of the time-varying graph, and $\rho : E \times \rho \to \{0, 1\}$, called indicator function, indicating whether an edge exists at a given time τ. There is an edge between nodes A and B in G if and only if the direct communication between them exists for the interval $[\tau, \tau + T_p]$. An example of a time-varying graph is shown in Fig. 2–1. Network topology is shown at three time slots t_1, t_2 and t_3. Five nodes, A, B, C, D and E, are in the network. Solid dots show their current locations, and dashed lines present their movement trail. When a transmission can be successful at time t, an edge exists between these two nodes.

2.2 Mobility Model

In our work, we consider a general mobility model only requiring that the movement of nodes can be predicted. For a given node, we can know where it will be at a given time. Mobility prediction has already been intensively studied and achieved with high accuracy, and some techniques about it are introduced in Section 1.2. Hence it is reasonable to assume that the motion trail of nodes can be predicted in this work.

Movement predication is achievable, and such information is important in scheduling. For example, when the devices are densely located in the networks, redundant messages are likely to be forwarded with no knowledge of node mobility. As a result, more energy is consumed and serious interference might also be caused, affecting the communication quality. Location information of nodes helps avoid forwarding unnecessary packets.

The findings in our work apply to the network systems consisting of mobile or static nodes, and our conclusions are further validated by the simulations in which we use time-varying network model together with three popular theoretical mobility models: i.i.d model, Random Waypoint model and Gauss-Markov model.

2.3 Transmission Model

We study the one-to-many transmission since it is a commonly seen transmission pattern in data dissemination. Assume that the source is denoted as S, and a group of destinations are represented as a vertex set D ($D \subseteq V$). Multiple hops could be engaged as relays for information transmission from one source to multiple destinations. The static networks where both nodes and edges keep unchanged are modeled as graphs. In traditional graph model, the concept of *path* is used to describe the node sequence a message passes through from a sender to a receiver. However, considering that the nodes are moving around, the connection between them are changing with time. Thus the corresponding concept of "path" in mobile networks is related with not only geographical distance but also the temporal factors. Therefore, we introduce the new concept *journey* to mobile networks.

Definition 2.3.1 (*Journey*)**.**
Let J be a journey from A to B, and we define $start(J) = A$ and $end(J) = B$. A journey is a sequence of tuples $j = \{(u_1, t_1), (u_2, t_2), ..., (u_k, t_k)\}$ where $t_i(1 \leq i \leq k)$ is the time when message arrives at u_i. Let e_{u_{i-1},u_i} be an edge between u_{i-1} and u_i. The following requirements are satisfied according to the definition of a journey:

1. $\forall i, j$, $1 \leq i, j \leq k$, *if* $i \neq j$, *then* $u_i \neq u_j$.
2. $\rho\left(e_{u_{i-1},u_i}, t_i\right) = 1, \forall 1 < i \leq k$;
3. $t_i + T_p < t_{i+1}$

Here we define journey j's topological length as the number of hops engaged in J, and it holds that

$$|j| = k. \tag{2–3}$$

Furthermore, $Departure(J)$ and $Arrival(J)$ denote the time t_1 and t_k respectively. The journey's temporal length, δ_j is defined as:

$$\delta_j = Departure(j) - Arrival(j). \tag{2–4}$$

The temporal length shows how much time is needed for a message to be transmitted along this journey. Notice that the journey is directed. The existence of a journey from A to B does not guarantee that a journey exists from B to A in the same time interval.

We say that messages from node A can reach node B if there exists a journey from A to B. In our study of transmission delay in multicasting, two kinds of journeys are discussed: minimal delay journey and delay bounded journey. To enable the transmission from A to B within the minimal delay, we need to find the journey j such that δ_j is minimized. If we want to find a journey such that the transmission delay is upper bounded by d, we need to find journeys which satisfy the inequality $\delta_j \leq d$.

Then we define two operations for journeys: *composition* and *union*, for the simplicity of description in the manipulation of journeys.

Definition 2.3.2.

(Composition) Assume $J_1 = \{(u_{11}, t_{11}), (u_{12}, t_{12}), ..., (u_{1m}, t_{1m})\}$ *and* $J_2 = \{(u_{21}, t_{21}), (u_{22}, t_{22}), ..., (u_{2n}, t_{2n})\}$ *are two journeys in a network, which satisfy the conditions that* $u_{1m} = u_{21}$ *and* $t_{21} > t_{1m} + T_p$. *We define their composition* $J_1 + J_2$ *as* $J_1 + J_2 = \{(u_{11}, t_{11}), ..., (u_{1m}, t_{1m}), (u_{22}, t_{22})..., (u_{2n}, t_{2n})\}$

Definition 2.3.3.

(Union) Let $J_1 = \{(u_{11}, t_{11}), ..., (w, t_{1p}), ..., (u_{1m}, t_{1m})\}$ *and* $J_2 = \{(u_{11}, t_{21}), ..., (w, t_{2q}), ..., (u_{2n}, t_{2n})\}$ *be two journeys. Assume that* $t_{1p} < t_{2q}$. *Two journeys intersect at node* w, *corresponding to the case that the node* u_{11} *forwards messages to* u_{1m} *and* u_{2n} *through the same hop* w. *We union these two journeys, and wipe out redundant transmissions to node* w. *After the union, the message first travels through the journey* $J'_1 = \{(u_{11}, t_{11}), ..., (w, t_{1p})\}$ *if it holds that* $t_{1p} < t_{2q}$. *Then the message is forwarded from* w *to* u_{1m} *and* u_{2n} *along two journeys* J'_2 *and* J'_3. *It is defined that* $J'_2 = \{(w, t_{1p}), ..., (u_{1m}, t_{1m})\}$ *and that* $J'_3 = \{(w, t_{2q}), ..., (u_{2n}, t_{2n})$. *Therefore, the journeys' union operation is*

$$Union(J_1, J_2) = Compostiion(J'_1, J'_2), Composition(J'_1, J'_3) \qquad (2\text{–}5)$$

Chapter 3 Journey in Mobile Networks

The journey is the route along which the messages are forwarded in mobile networks. For delay-sensitive applications, the information should be disseminated as rapidly as possible, and thus the journey with minimal delay should be found. However, in delay-tolerant applications, it is allowed that information is received under the delay constraint, and hence the journey selection is under relaxed delay constraints. Some other factors such as energy saving should be taken into consideration, and the union of journeys is a good way to reduce energy consumption.

We first study the minimal delay journey in this chapter. It is conjectured that tradeoff exists between the delay and energy efficiency. For example, the journeys with prolonged temporal lengths are likely to engage fewer transmissions, and they are more likely to be unioned with other journeys. We also study the algorithm to find the delay-bounded journey. By comparing these two types of journeys, we can explore the relationship between the delay and energy efficiency.

3.1 Finding the Journey with Minimal Delay

We first find the minimal delay journey for data transmission. With the network model, mobility model and transmission model defined above, we explore how to achieve the fastest transmission in this section. We propose Alg. 1 to find a "journey" whose temporal length is minimized from the sender to the destination.

Step 1: Initialization. To find the "fastest journey" from a source S to a destination D, we construct a *Minimal-Delay Reachable Graph* denoted by $G_M = \{V, E\}$. Its vertex set V is the set of all nodes in the network, and edge set E contains all edges on the journeys with the minimal delay. We use $Arrival(i)$ to denote the earliest time when a message from S arrives at the node i. Next we describe how to construct the graph G_M. We first set the edge set $E = \emptyset$, reference time $t_r = 0$, $Arrival(S) = 0$, and $Arrival(i) = \infty$, $\forall i \in V$, and $i \neq S$. Besides we define a vertex set V' to record the nodes that have not found the minimal-delay journey starting from S. Set V' is initialized to be V.

Step 2: Update the delay. Since the motion trail of nodes can be predicted, we could find the earliest time when an edge exists between S and any other node in the network. Suppose that the earliest time of edge existence for node i is t_i, and we set $Arrival(i) = t_i + T_p$ ($\forall$ vertex i in $V - S$). We remove the sender S from V', and set

Algorithm 1 Minimal Delay Journey

Require:

The set of vertices, V;

Vertex set V';

The motion of all vertices;

The source S and the destination set D;

The earliest time when a message from S arrives at the vertex w, $Arrival(w)$;

The previous hop from which the message is forwarded to w, $Parent(w)$;

Ensure:

$Arrival(S) := 0$;

for all vertices w in V and $w \neq S$ **do**

  $Arrival(w) := \infty$;

  $Parent(w) := NULL$;

end for

$V' := V$;

repeat

  for all vertices w in V' **do**

    pick vertex w such that $Arrival(w) = \min\{Arrival(u)|u \in V'\}$;

    remove w from V';

    for all vertices v in V' **do**

      calculate the time needed for v to meet w after the time $Arrival(w)$, T_{wv};

      if $Arrival(v) > Arrival(w) + T_{wv} + T_p$ **then**

        Update $Arrival(v) := Arrival(w) + T_{wv} + T_p$;

        Update $Parent(v) := w$;

      end if

    end for

  end for

until all vertices in D are removed from V'

node S as the parent node of all nodes in V'.

Step 3: Construct the Minimal-Delay Reachable Graph. As we have updated $Arrival(i)$ for all node i in V', we select the node u such that $Arrival(u) = \min\{Arrival(i)|i \in V'\}$. We add a directed edge from S to node u . and set the reference time $t_r = Arrival(u)$. Then we remove u from V', and find the earliest time when an edge exists between u and any other node in V' after the reference time t_r. Suppose that the earliest time for node j is t_j, and it takes $T_{u,j}$ for node u to meet node j after t_r. We take the value of $Arrival(j) = \min\{Arrival(j), t_j + T_{uj} + T_p\}$ $(\forall j \in V')$. If $t_j + T_{uj} + T_p < Arrival(j)$, we set u as the parent node of node j. We repeat the operations above and update both the arrival time and parent node for each node in V'. This algorithm could be terminated when the destination is removed from V'.

Step 4: Finding the Minimal-Delay Journey Now according to the parent nodes recorded by the nodes, we can add directed edges between them to the graph G_M. An example of *Minimal-Delay Reachable Graph* is shown in Fig.(3–1). The network consists of five nodes $A, B, C, Dest$ and E. The sender is A and node $Dest$ is the destination. The journey from A to a destination corresponds to the directed path in the Fig.(3–1), and arrival time of A's messages at each node is shown besides each node.

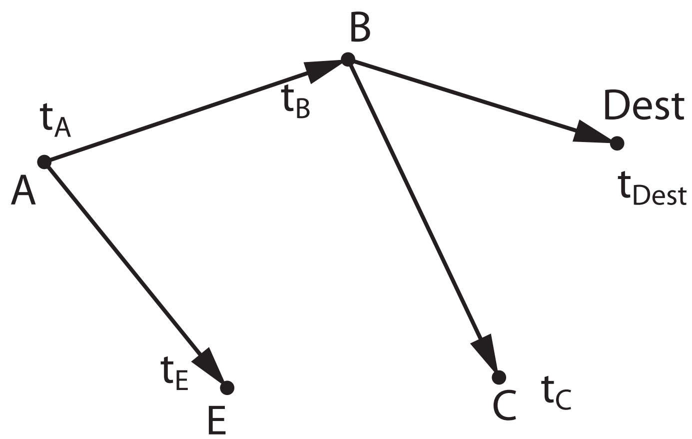


Figure 3–1 Minimal-Delay Reachability Graph

It is easy to find the path $A - B - Dest$ from sender A to the destination in G_M, and accordingly the composition of journey AB and $B - Dest$ is exactly the journey from A to the destination. Arrival time at each node is the earliest time when this node receives messages from S, and $Arrival(destination)$ is the minimal delay between the source-destination pair. Accordingly, the minimal delay journey form A to $Dest$ is $\{(A, t_A), (B, t_B), (Dest, t_{Dest})\}$.

Theorem 3.1.1.

With Algorithm 1, we can find the journey from the given source to the destination with minimal delay.

Proof. Assume that a source S and n destinations exist in the network. In Algorithm 1, we update the delay of each node in V' and remove the one with the minimal delay each time from V'. The minimal-delay journey and the time cost are found for the removed node one by one. In the first iteration, we find a node that S can directly send messages to within minimal delay.

Inductive Assumption: Suppose that minimum-delay journeys from S to m destinations have been found.

Now we consider the node u, which is selected in the $(m+1)$-th iteration. The message can not reach u earlier than messages for m destinations, since we select the node with minimal delay in each iteration. Therefore, the earliest message to u must be forwarded from S or one of m destination. To find the earliest time when the message from S arrives at u, we need to check when u could meet one of m destinations after they have received the messages. By the inductive hypothesis, the minimal-delay journeys have been found for m nodes. We can find the delay from each of m destinations to u, and add the time cost of journeys from S to it. The journey of the minimal time cost is the minimal-delay journey from S to u. This completes our proof. □

Theorem 3.1.2.
The complexity of Algorithm 1 is $O(n^2)$.

Proof. Suppose that it takes one time unit for the estimation of two nodes' meeting time given their mobility track. In the first iteration, the source finds the earliest time when it reaches other $(n-1)$ nodes, and then label the nodes with the arrival time. The time cost in the first iteration is $O(n-1)$. We find the earliest time when the message arrives at one node in each iteration, and mark the corresponding node, indicating that the journey from the source to this node has been determined. In each iteration, one node is marked. Therefore in the $i-th$ iteration, the node which is newly marked in last iteration will find the earliest time when they meet $(n-i)$ remaining unmarked nodes. Hence the time cost of the $i-th$ iteration is $O(n-i)$. The maximal number of iterations are $n-1$, and the total time cost is

$$(n-1) + (n-2) + ... + 1 = n(n-1)/2. \quad (3\text{–}1)$$

Hence in the worst case, the time cost of Algorithm 1 is upper bounded by $O(n^2)$. □

3.2 Delay-bounded journey

In many applications such as Delay Tolerant Networks, it is not compulsory that information should reach the destinations as fast as possible. Instead, messages are expected to be received within a given delay bound. Therefore, we discuss how to find the delay-bounded journeys in this section. Given a source S and a destination D as well as the delay constraint d, we need to find all journeys from S to D with temporal length no larger than d. The algorithm to find all delay-bounded journeys are described in Algo. 2.

Algorithm 2 Delay bounded journey

for each destination D **do**
  Construct a new graph G with vertex set $V = \emptyset$ and edge set $E = \emptyset$
  Add a new vertex labeled with $(S, 0)$ to the G
  repeat
    for any leaf vertex (u, t_u) in G **do**
      for any node w ($w \neq D$) that can receive messages from u at time t_{uw} and $t_u < t_{uw} \leq d$ **do**
        if w does not exist in the path from $(S, 0)$ to (u, t_u) **then**
          Add a new vertex labeled with (w, t_{uw}) to G
          Add an edge between (u, t_u) and this new vertex (w, t_{uw})
        end if
      end for
    end for
  until no new vertex is added to G
  for any leaf node whose label contains D **do**
    record all paths from S to D
  end for
end for

Theorem 3.2.1.
All delay-bounded journeys from S to D can be found with Algo. 2.

Proof. Every time we look for the next hop within the remaining time of the given delay so that we can ensure that the journeys are delay bounded. Besides, it is required that one node can not appear in the same journey more than once, and we avoid the case that messages are forwarded back and forth. It is asserted that all delay-bounded journeys can be found with algorithm 2, and we prove it by contradiction.

Suppose that this is a delay bounded journey $journey = \{(S, t_S), (H_1, t_1), ..., (H_k, t_k), (D_i, t_D)\}$, which can not be found with algorithm 2.

Since H_1 is reachable from S within the delay, (H_1, t_1) becomes the leaf node with parent S in the first iteration. In the second iteration, (H_2, t_2) is one of the leaf nodes with parent (H_1, t_1) in that H_2 is also reachable from H_1 within the delay. Similarly, we can show that (H_3, t_3), ..., (H_k, t_k) can be added to the graph G one by one. Therefore, the journey $\{(S, t_S), (H_1, t_1), ..., (H_k, t_k), (D_i, t_D)\}$ appears in graph G, and it can be found with algorithm 2, which contradicts with the assumption. Hence the assumption is wrong, and all delay bounded journeys can be founded from the source and the given destination. This completes our proof. □

Theorem 3.2.2.
The complexity of Algorithm 2 is $O\left(\frac{(n-1)!}{(n-d-1)!}\right)$ *in the worst case.*

Proof. We analyze the algorithm assuming that the model is discussed in discrete time. Suppose that the nodes change its location in each time slot, and that one time slot is long enough to support the information transmission. In the first iteration, the source will find all nodes that can reach it within the delay bound, and the time cost is $O\,(n-1)$ at most. For all the reachable nodes in the first iteration, we find all possible nodes that can meet them in the second iteration. Such process is repeated, and we will find nodes reachable by the nodes in the previous iteration. Considering that one node will not appear more than once in a journey, we can assert that no more than $O\,(n-i)$ nodes are reachable from the same node in the $i-th$ iteration. Since there are at most d iterations ($d < n$), and the total time cost is

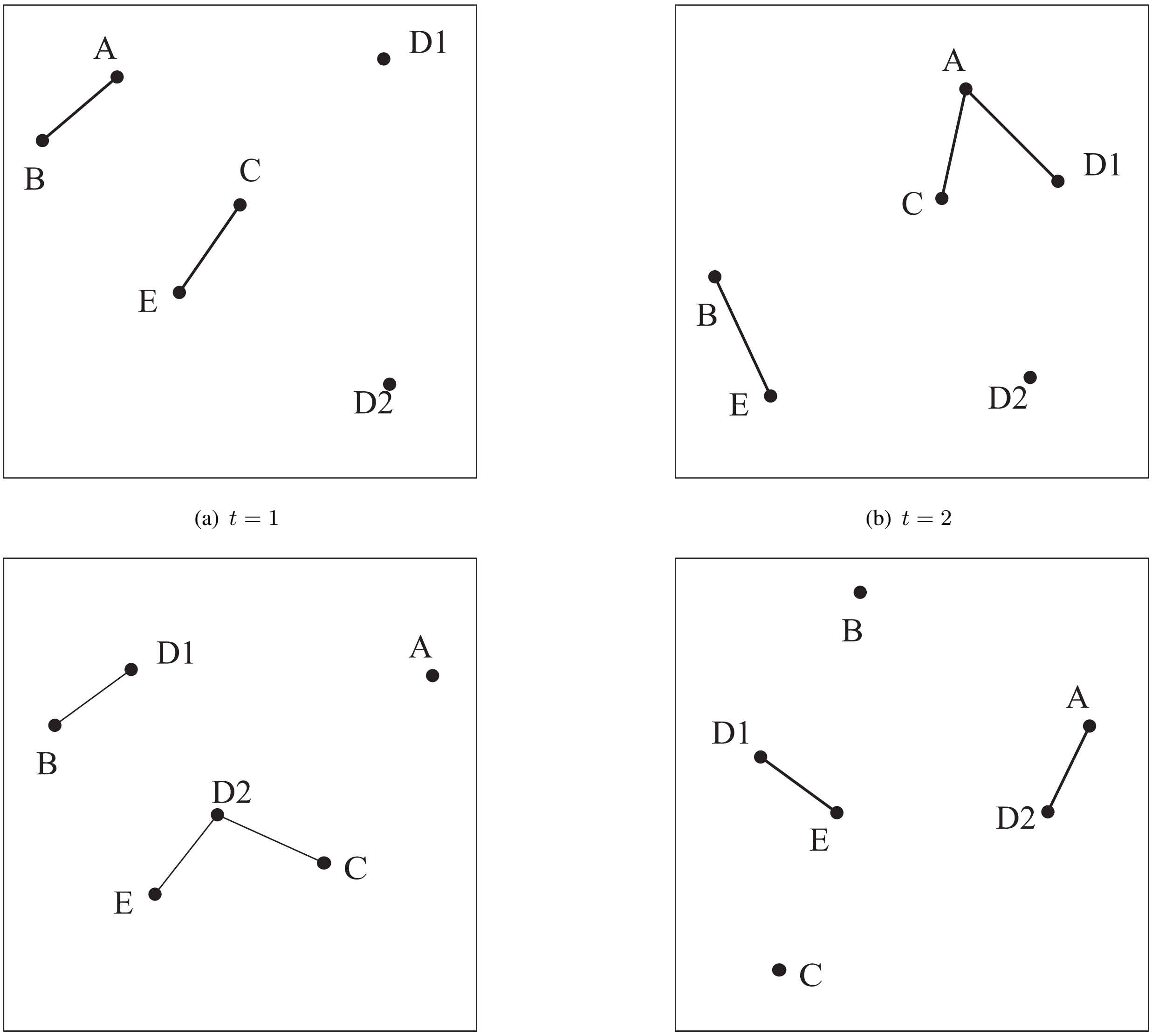
$$(n-1)(n-2)...(n-d) = \frac{(n-1)!}{(n-d-1)!}. \tag{3–2}$$

The time cost of the algorithm to find all delay-bounded journeys is $O\left(\frac{(n-1)!}{(n-d-1)!}\right)$. □

Remark 3.2.1.
When delay constraint $d << n$*, the complexity to find all delay bounded journeys is* $O\left(n^d\right)$*. As we can see, the time cost is increased rapidly with the relaxation of delay constraint.*

3.3 Benefits of Relaxed Delay Constraint

We give a specific example to illustrate our definition of minimal delay journey as well as the delay bounded journey in this section. Moreover, we also want to show the underlying benefits of higher energy efficiency when the delay constraint is relaxed. We



(a) $t = 1$ (b) $t = 2$ (c) $t = 3$ (d) $t = 4$

Figure 3–2

Fig 3–2 An example of energy minimization multicasting when the transmission range is fixed given the network topology at four time slots

assume that the transmission range is fixed and equivalent for all nodes in the network, so we can count the transmissions to estimate the energy consumption. Figure 3–2 shows the network topology at four time slots. There are six nodes, $A, B, C, D1, D2$ and E in the network, where A acts as the source, and $D1, D2$ are chosen as destinations. In each time slot, the edges between two node indicate that one node can forward messages to the other.

If we want multicasting to be finished as soon as possible, we need to find the minimal delay journeys. The minimal delay journey for $D1$ is $\{(A, 0), (D1, 2)\}$, and its temporal length is 2. The minimal delay journey for $D2$ is $\{(A,0),(C,2),(D2,3)\}$, and the temporal length is 3. Three transmissions are engaged in multicasting in total.

Now we relax the delay constraint, and assume that it is set as 4. First we look for all the delay bounded journeys to $D1$ and $D2$. Two journeys exist for $D1$, one is $\{(A,0),(B,1),(D1,3)\}$ with temporal length 3, and the other is $\{(A,0),(D1,2)\}$ with temporal length 2. There are also two journey for $D2$, $\{(A,0),(C,2),(D2,3)\}$ with temporal length 3 and $\{(A,0),(D2,4)\}$ with temporal length 4. To minimize the energy consumption, we choose the journey $\{(A,0),(D1,2)\}$ for destination $D1$, and the journey $\{(A,0),(D2,4)\}$ for $D2$. Only two transmissions are needed with the relaxed delay constraint. It is shown that energy can be saved at the cost of transmission delay.

Chapter 4 Minimum Energy Problem in Multicasting

We have discussed how messages could arrive at the destination as fast as possible in the previous chapter. In many applications, the delay is not the only metric of network communication performance. With the introduction of green technology and green communication, energy efficiency becomes an important concern in the network design. Many applications are delay-tolerant and allow message to arrive within a given delay bound. Based on these considerations, we explore the routing as well as transmission scheduling policies to achieve the minimal energy multicasting under the delay constraint, which is formulated as the Minimum Energy Problem in Multicasting (MEPM).

In this chapter, we study the Minimum Energy Problem in Multicasting in two cases of fixed transmission range and changeable range of network nodes. The problems are first described by the means of mathematical optimization problem, and then their complexity is evaluated. Based on the problem formulation and complexity analysis, we apply Lagrangian multiplier method and propose efficient algorithms to solve them.

4.0.0.1 Lagrangian Relaxation and Dual Problem

We first give a brief introduction of the Lagrangian relaxation method together with the dual problem, which lays the foundation of the following fast approximation of transmission scheduling and estimation of energy bounds for the Minimum Energy Problem in Multicasting.

Suppose that the primal problem P is

$$\min g(\vec{x}) \tag{4–1}$$

$$s.t. f_i(\vec{x}) \leq 0, \forall i \tag{4–2}$$

Lagrangian relaxation is a method in mathematical optimization, providing a simple approximate algorithm for a complex problem which is formulated as a constrained optimization problem. A relaxed algorithm gives an approximation to the solution of the primal problem. The Lagrangian function LP corresponding to P is

$$LP = \min_{w_i} g(\vec{x}) + \sum_i w_i f_i(\vec{x}) \tag{4–3}$$

Here w_i is called the Lagrangian multiplier, and $w_i \geq 0$. The objective function together with the constraints of the primal optimization problem is included in the Lagrangian function. For the inequality constraints in the primal problem, the Lagrangian relaxation method penalizes these violated inequality constraints by the means of Lagrange multipliers [28]. It consists of the weighted sum of the inequality constraints, and the weight is exactly the multiplier it assigns to different constraints.

If an inequality does not hold, the costs are increased as the multiplier acts as the penalty, indicating the deviation from the optimum. Then the original variable will be updated in the opposite direction. If the inequality constraint is satisfied, the multiplier act as the reward, so the costs of the optimization problem can be reduced. Hence the original variable will continue to be updated in the same direction and approach the optimal point. This is the underlying efficacy of multipliers in the Lagrangian relaxation.

The reason for the application of Lagrangian relaxation method is that it has much lower computation costs than solving the primal optimization problem directly. Besides, the Lagrangian relaxation method can return the lower bound for the optimum of the primal problem. When we search for the minimum of Lagrangian function by trying a variety of multipliers, the values of primal variables returned by P gives a feasible solution to primal optimization. The feasible solutions provide the upper bound for the primal problem.

We take one more step and explore the dual problem DP defined on the basis of the Lagrangian function. In fact, the multipliers in L are exactly the dual variables in DP. The dual problem is also an optimization problem which aims to maximize the value of the Lagrangian function and requires that the dual variables should be non-negative. Dual problem DP is defined as:

$$\max L \tag{4–4}$$

$$s.t. w_i \geq 0, \forall i \tag{4–5}$$

Dual problem DP presents a tighter lower bound than Lagrangian function for primal problem P by maximizing the value of Lagrangian function L. The comparison among values of P, LP and DP is:

$$LP \leq DP \leq P \tag{4–6}$$

The duality gap is defined as the difference between the optimal values of the primal problem and the dual problem. The duality gap is zero under some conditions, such as for a convex optimization problem with KKT conditions satisfied. Generally, the gap is non-zero between the two problems. If we start from a feasible but suboptimal point of the primal problem, and all constraints are satisfied by this point. There is a direction in which the variable vector can move so as to further reduce the value of the objective function. These variables approach the boundary of the feasible region described by the constraints, and they can get closer to the optimum. Dual variables are also updated to reach the optimum of the dual problem. By adjusting the dual variables, the lower bound of the primal problem's objective function is also changed.

Dual problem plays the role of providing the highest lower bound for primal problem. The feasible primal values found during the iterations provide the upper bound for the primal problem. Given both upper and lower bound, we can estimate the how far the current solution is from the optimum. What's more, the lower bound offered by the dual problem also plots the theoretical achievable region of the primal problem. We can find the optimum if the lower and upper bounds collide.

4.1 Fixed Transmission Range Minimum Energy Multicasting

In this section, we first study the Minimum Energy Problem in Multicasting assuming that the transmission ranges of all nodes in the network are fixed and equivalent. Tree topology is a popular network architecture since it can avoid redundant transmissions. Through multicasting, one source forwards messages to multiple destinations. Under the assumption of fixed transmission range, the power consumption is linear with the number of transmissions. The goal of Minimum Energy Problem in Multicasting is exactly to minimize the number of transmissions engaged in multicasting.

Given the tree topology, transmission count is the total number of destinations and extra relays if we do not consider the case that multiple nodes can receive messages sent by one node at the same time. Considering that the wireless nodes have broadcast nature, the transmission count is generally no larger than the total number of nodes participating in multicasting.

On the basis of the problem formulation and complexity analysis, we then consider a simple case that only direct and one-hop transmission is engaged in multicasting. Then we give the corresponding solution. Next, we extend our analysis to a more general case that multi-hop transmission is also allowed, and adopt a more complex method called Lagrangian relaxation to solve it.

4.1.1 Problem formulation and complexity

The mathematical expression of Minimum Energy Problem in Multicasting can be written in the form of an optimization problem. The optimization problem consists of one optimization goal together with a series of constraints. The optimization goal can usually be to obtain the minimum or maximum value of an objective expression. The optimization problem requires that the optimal goal should be reached under the condition that all constraints listed should be satisfied.

We first define some variables in the problem formulation and explain their mathematical meanings. Transmission delay bound is denoted as d. Let $I(\bullet)$ be the indicator function, and $I(u,t) = 1$ when node u forwards the message at time t. Similarly, $I(j) = 1$ when the journey j is chosen for transmission from the source to one destination. Variable e_{uvt} indicates the existence of edge from u to v at time t. Let $\delta(j, e_{uvt}) = 1$ if e_{uvt} is on journey j, and it is 0 otherwise. Let J_{D_i} be the set of all delay-bounded journeys from S to the destination D_i. The mathematical formulation of MEPM with fixed transmission range is given below:

$$\min \sum_{u\in V} \sum_{t=1}^{d} I(u,t)$$

$$s.t. I(u,t) \geq \sum_{j\in J_{D_i}} \sum_{v\in V} I(j)\delta(j, e_{uvt}), \forall D_i \in D, \forall u \in V, \forall t \in [1, d] \qquad (4–7)$$

$$\sum_{j\in J_{D_i}} I(j) = 1, \forall D_i \in D \qquad (4–8)$$

$$I(u,t) = \{0, 1\}, \forall u \in V, \forall t \in [1, d] \qquad (4–9)$$

$$I(j) = \{0, 1\}, \forall D_i \in D, \forall j \in J_{D_i} \qquad (4–10)$$

We aim to minimize the energy consumed in multicasting, and the energy consumption is determined by the number of transmissions, so we take it as the optimization goal to minimize the transmission count. Besides, we give four constraints corresponding to requirements in practical applications.

Constraint 4–7 means that a node will act as a relay once a journey containing it is chosen. Constraint 4–8 ensures that the message from S can reach every destination, and that destinations will not receive redundant messages. As is indicated in Constraint 4–9 and 4–10, the choices of node and journey in transmission are expressed by the indicator function $I(\bullet)$ whose value is either 0 or 1. The value of indicator function is 1 if the node participates in information forwarding or the journey is what the message travels along.

Theorem 4.1.1.
Minimum Energy Problem in Multicasting is an NP-hard problem given the fixed transmission range of each node.

Definition 4.1.1 (*Set Cover Problem*)**.**
Given elements $e_1, e_2, ..., e_n$, the union set $U = \{e_1, e_2, ..., e_n\}$ and the union of subsets of U, $S = \{s_1, s_2, ..., s_m\}$ such that $\bigcup_{1\leq i\leq m} s_i = U$. A subset of S is called a set cover C if

and only if $\bigcup_{s \in C} s = U$. *The Minimum Set Cover Problem (Set Cover Problem for short) is to find* C *that minimizes the cardinality of C,* $|C|$ *[29].*

Proof. To prove a problem is an NP-hard problem, we should reduce an NP-hard problem to this problem and show that the reduction takes polynomial time. We first show that set cover problem can be reduced to the Minimum Energy Problem in Multicasting under the assumption of fixed transmission range. Minimum Set Cover problem is classic NP-hard problem, and we give its definition above.

Based on the definition of set cover problem, we let any element e_i in U be destinations D_i in the network, and any element s_i in set cover C be the node H_i neither the source nor the destinations. The correspondence between the subset and the elements are reflected by the relationship between the journey and relay. Suppose that a subset s_i contains k elements e_{i_1}, e_{i_2},..., e_{i_k}, and that we construct k one-hop delay-bounded journeys (source $\rightarrow H_i \rightarrow D_j$) where $j = \{i_1, i_2, ..., i_k\}$ and require that messages can not be forwarded to more than one node at the same time. It is also assumed that no other delay-bounded journeys are available for transmission from the source to destinations. The reduction is finished, and we then show that the solution to MEPM can be transformed to the solution to Minimum Set Cover problem.

Let $n(transmission)$ be the number of transmission which directly determines the energy consumption, and $n(destination)$ be the number of destinations. We suppose *relay* to be the nodes which are neither the source nor the destinations but participants in multicasting, and define $n(relay)$ to be the number of these nodes. Considering that a mobile multicast tree will be constructed among nodes, we have

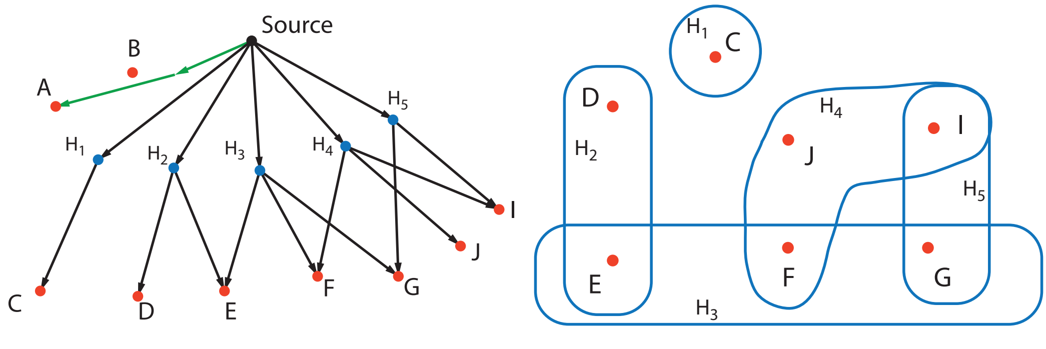
$$n(transmission) = n(destination) + n(relay) \tag{4–11}$$

Once we find the solution to MEPM and minimize the energy consumption in multicasting, the fewest relays are actually selected to forward information to all destinations in D. According to the reduction we have performed, the minimal subsets can be chosen to cover all elements in U. This completes our proof. □

4.1.2 Fast Approximation with One-hop Transmission

When the transmission range is fixed, designing transmission policies to achieve minimal energy consumption is infeasible in practice considering the NP-hardness of MEPM. It is also our motivation to find an approximate but efficient algorithm for the relay selection and transmission scheduling.

We start with a simple case of one-hop transmission. Moreover, we also require that a node only forwards the message to one node at a time slot, e.g., the broadcast nature of



(a) One-hop journeys (b) Destination-Hop sets

Figure 4–1

Fig 4–1 Reduction of the Minimum Energy Problem in Multicasting in Fig. (a) to the Minimum Set Cover Problem in Fig. (b)

wireless nodes is not considered. Inspired by the approximate solution of Minimum Set Cover problem, we find a solution to MEPM in the simple case. Then we will extend our analysis to a more general setting, multi-hop transmission.

We define a one-hop transmission pattern such that the message from the source can pass through no more than one hop, i.e. the source can only communicate with the destination either directly or through only one hop. Here we define the Destination-Hop Set as follows:

Definition 4.1.2 (*Destination-Hop Set*)**.**
For a network with vertex set V and destination set D, we define the Destination-Hop Set of a node v as $H(v)$, and $H(v)$ contains all destinations which can receive messages from the source with v alone as its hop within the delay constraint d. We have $H(v) = \{D_i | D_i \in D, \exists Journey\ J, J = \{(S, t_S), (v, t_v), (D_i, t_{D_i})\}, t_{D_i} \leq d\}$.

To illustrate our definition of one-hop journey and destination-hop set more clearly, we give a specific example in Fig. 4–1. Among all nodes in the network, $A, B, ..., I$ are chosen as destinations, and $H_1, H_2, ..., H_5$ are engaged as relays. The directed edge between two nodes shows that message is forwarded from the start point to the end point. Fig. 4.1(a) shows all one-hop or direct transmission patterns from the source to destinations. Fig. 4.1(b) shows the destination-hop sets corresponding to delay-bounded journeys in Fig. 4.1(a). Since both A and B do not use non-destination nodes as relay, we only consider other destinations. We construct sets for each relay node $H_1, H_2, ..., H_5$, and these sets contain the reachable destinations as their elements.

With the Destination-Hop Set illustrated in Fig. 4–1, we first reduce the Minimum Energy Problem to the Minimum Set Cover problem, which has been studied for a long time. The approximate algorithm dealing with the Minimum Set Cover problem can be adapted to solving the minimum transmission problem. We have given the definition of Minimum Set Cover problem above.

Considering that not all nodes need relays, we cannot reduce Minimum Energy Problem in Multicasting to the Minimum Set Cover problem directly. There are three cases in one-hop transmission pattern:

1. The source can communicate with the destination directly;
2. The source uses one destination as the relay of another destination;
3. The source uses a non-destination node as the relay.

Let set D_A contain all destinations in the first case, and no hop is needed for messages to reach these destinations. Let set D_B contain all destinations in the second case. If the destination which acts as the relay can directly communicate with the destination, then no hop is needed. As for case 3, one hop is needed. Let D be the set of all destinations. We exclude the cases where no hop is needed before performing the reduction, and define the set $U = D - D_A - D_B$. For each node v_i that is neither the source nor a destination, we define a subset D_{v_i} to be $H(v_i)$, the Destination-Hop Set of v_i. When we find a minimum number of subsets $\{D_{v_i}\}$ to cover all elements in U, the corresponding nodes v_i can be selected as relays to minimize the transmissions. We now finish the reduction from the minimum transmission problem to the Minimum Set Cover problem.

Theorem 4.1.2.
The approximation ratio that can be achieved in polynomial time for Minimum Energy Problem in Multicasting is $O(\log m + 1)$, where m is the number of destinations.

Proof. The Set Cover Problem is known to be NP-complete. Assume that k elements in total are contained in the union set U, and that m subsets are available for selection into C. Many algorithms and heuristics have been proposed, among which a greedy algorithm for Set Cover Problem with an approximation ratio of $O(\ln k + 1)$ is commonly used [30]. It is revealed that there exists the tradeoff between approximation ratio and the costs of solutions to Set Cover problem [31]. The approximation ratio that can be achieved in polynomial time for Set Cover Problem in Multicasting is no better than $O(\log |X| + 1)$, where $|X|$ is the total number of elements in the set U.

According to the reduction from Minimum Energy Problem to Set Cover Problem, it is easy to show that the number of subsets in the approximate solution to Minimum Set Cover Problem is exactly the number of extra relays in the solution to Minimum Energy

Problem. Therefore, the approximation ratio achievable in polynomial time for Minimum Energy Problem is also $O(\log m + 1)$. This completes the proof. □

By deriving the destination-hop sets, we can apply the Greedy Set Cover algorithm to the scheduling of MEPM so as to decrease the transmissions and achieve the energy efficiency. The approximate algorithm of Minimum Set Cover problem is described in Alg. 3.

Algorithm 3

Require:

Elements $e_1, e_2, ..., e_n$;

The union set, $U = \{e_1, e_2, ..., e_n\}$;

The union of subsets of U, $S = \{s_1, s_2, ..., s_m\}$ such that $\bigcup_{1 \le i \le m} s_i = U$;

Ensure:

Set Cover C;

```
define coverage function c(C) = |⋃_{s∈C} s|
C ← φ;
while ⋃_{s∈C} s ≠ U do
    choose s ∈ S that maximizes c(C ∪ {s}) − c(C);
    C ← C ∪ s
    S ← S \ s
end while
```

We can select the hops based on the solution to the Minimum Set Cover problem, and then design the scheduling of transmissions, in particular, we need to specify the time when the hops forward data. Multiple nodes might serve as the relay for the same destination, and we should ensure that their Destination-Hop Sets are disjoint so as to eliminate redundant transmissions, i.e., every destination only receives information from no more than one relay.

4.1.3 Fast Approximation with Multi-hop Transmission

We continue to solve the Minimum Energy Problem in Multicasting allowing multi-hop transmission. It should be noticed that the solution to MEPM in the case of multi-hop transmission is not a simple extension of the solution in the case of one-hop transmission. For one-hop transmission, we list relays together with all of their reachable destinations. However, as for multi-hop transmission, more than one hop is engaged in a journey from the source to the destination. Therefore, a large number of permutation and combination of relays exist for a given pair of source and destination. It takes great costs to list all correspondence between destinations and the relays they share in common.

Traveling salesman is an NP-hard Problem, and extensive study shows that it can be efficiently solved by combinatorial optimization method [32]. It thus triggered the discussion in the area of computational complexity about whether a polynomial-time solution can be found to an NP-hard problem. Given the complexity of Minimum Energy Problem in Multicasting and the generality of multi-hop transmission, we thus turn to the study of its optimization problem for the fast solution. We apply Lagrangian relaxation and subgradient method to MEPM assuming that the transmission range is fixed.

4.1.3.1 Lagrange Function

The Lagrangian function is derived from the optimization problem defined above. We assign weight w_{uD_it} to every inequality constraint, and w_{uD_it} is also called the dual variable. Dual variables will appear in the dual problem corresponding to the original optimization problem later.

The Lagrangian function is denoted as L, and it contains two original variables: $I(u,t)$ and $I(j)$.

$$L = \min_{I(u,t),I(j)} \left(\sum_{u \in V} \sum_{t=1}^{d} I(u,t) + \sum_{u\ inV} \sum_{t=1}^{d} \sum_{D_i\ inD} w_{uD_it} \left(I(j)\delta\left(j, e_{uvt}\right) - I(u,t) \right) \right) \tag{4–12}$$

We divide the primary Lagrangian function into two subproblems L_1 and L_2. As is shown below, L_1 is only related with variable $I(u,t)$, and L_2 is related with $I(j)$.

$$L_1 = \min_{I(u,t)} \left(\sum_{u \in V} \sum_{t=1}^{d} \left(I(u,t) - \sum_{D_i \in D} w_{uD_it} I(u,t) \right) \right) \tag{4–13}$$

The value of $I(u,t)$ can be taken as follows to minimize L_1:

$$I(u,t) = \begin{cases} 1, & \sum_{D_i \in D} w_{uD_it} > 1; \\ 0, & \sum_{D_i \in D} w_{uD_it} \leq 1. \end{cases} \tag{4–14}$$

The second Lagrangian subproblem L_2 is:

$$L_2 = \min_{I(j)} \sum_{D_i \in D} \sum_{j \in J_{D_i}} \sum_{u \in V} \sum_{v \in V} \sum_{t=1}^{d} w_{uD_it} I(j)\delta\left(j, e_{uvt}\right) \tag{4–15}$$

To obtain the minimal value of the expression above, we construct a new weighted graph G_{D_i} for each destination D_i. We add all delay-bounded journeys from the source to destination D_i to the graph G_{D_i}, i.e., we add all nodes and edges e_{uvt} existing on these journeys to G_{D_i}. Then we assign w_{uD_it} to any edge e_{uvt} on the journeys. To minimize L_2, we should find the journey with the minimal weight from the source to every destination.

4.1.3.2 Dual Problem

We derive the dual problem from the original Minimum Energy Problem in Multicasting, based on the Lagrangian function L_1 and L_2, assuming that transmission range is fixed. The dual problem is also an optimization problem, and it contains w_{uD_it} as its dual variables. The dual problem is shown like:

$$\begin{aligned} &\max L_1 + L_2 \\ &w_{uD_it} > 0, \forall u \in V, \forall D_i \in D, \forall t \in [1, d] \end{aligned} \tag{4–16}$$

Corresponding to the original problem, the dual problem aims to maximize the Lagrangian function L. The constraints in dual problem is simple compared with those in the original problem, and dual variables are required to be non-negative. We solve the dual problem by using the subgradient method to update the dual variables w_{uD_i}. We denote the step size as Δ which shows how much the dual variable would be changed in each iteration. Every time when w_{uD_it} is updated according to its previous value and the constraints imposed in the original problem, we gradually move towards the maximum value of the objective function.

$$w_{uD_it}(k+1) = \left(w_{uD_it}(k) + \Delta \left(\sum_{j \in J_{D_i}} \sum_{v \in V} I(j) \delta\left(j, e_{uvt}\right) - I(u,t) \right) \right)^+ \tag{4–17}$$

Since the dual variable w_{uD_it} should be non-negative, it is defined that

$$(x)^+ = \begin{cases} x, & x > 0; \\ 0, & x \le 0. \end{cases} \tag{4–18}$$

As we change the value of both original variables, $I(u,t)$ and $I(j)$ as well as update the dual variables w_{uD_it}, the value of objective function in the original problem, i.e., the number of transmissions will be changed in each iteration. At the same time, the value of the objective function in the dual problem is also adjusted. When the value of the dual problem becomes stable, we derive a tight lower bound of energy consumption for the primal problem. In each iteration, we record the value of objective function in the primal problem, and we obtain the feasible scheduling if all constraints are satisfied. The original variables achieving the smallest feasible energy consumption among all iterations is the solution we find to the primal problem. Lastly, we can make the transmission scheduling based on the value of $I(u,t)$. For any $I(u,t)=1$, we let node u forward the messages at the time slot t.

The variable updates happen in many iterations until the objective function's value in the dual problem becomes stable or the maximum iteration are reached. In each iteration, the primal variables are renewed based on the current values of dual variables, and dual variables are changed according to the newly updated primal variables. Recall that the lower bound of the primal problem is achieved by the dual problem, and that the upper bound is achieved by the feasible variables of the primal problem. The solution returned by the primal problem is the result achieved by the Lagrangian relaxation method, and is denoted B_u to present the upper bound of this optimization problem. We denote the optimum of the formulated problem as B_{min}. To evaluate the efficiency of the approximation, the deviation of the solution $error$ satisfies that

$$error = \frac{B_u - B_{min}}{B_{min}}. \tag{4–19}$$

4.1.4 Complexity of the Fast Approximation

The approximate algorithm is proposed for the Minimum Energy Problem in Multicasting to achieve the feasibility of energy saving multicasting. We apply Lagrangian relaxation method, using a little more energy for multicasting with a much less time cost than the optimal scheduling. In this part, we show the theoretical time complexity of the approximate algorithm in table 4–1. The network size is n, and the number of destinations is m. The count of all delay-bounded journeys from the source to a group of destinations is $|J|$, and $|J_D|$ is the maximal number of journeys from the source to any destination. We use $|J_{D_i}|$ to denote the total number of journeys to a given destination D_i, and $|J_D|$ is defined as:

$$|J_D| = \max_{D_i \in D} |J_{D_i}| \tag{4–20}$$

With some known information about the node mobility and delay-bounded journeys, the time cost of approximate algorithm mainly comes from updating original and dual variables. The complexity of variable update in each iteration is shown in table 4–1.

Assume that p iterations are included in this approximate algorithm, and its time complexity T is

$$T = p * O\left(mn^2d^2 + d|J| + n^2d|J|\right) \tag{4–21}$$

$$= O\left(pmn^2d^2\right) \tag{4–22}$$

The number of iterations needed in the approximate algorithm depends on the step size in updating the variables, since the step size affects the convergence speed of the objective function [33]. If the step size is small, it takes many iterations for the optimization to find the optimum. However, the step size should not be too large to miss the optimal point. A more efficient choice of step size is related with the gap between upper and lower bound derived in each iteration. When the gap is large, the step size can be raised up to approach the optimum more rapidly. If the gap is small, the step size needs to be lowered down so that it won't miss the optimal point.

4.2 Variable Transmission Range Minimum Energy Multicasting

In many cases, the transmission ranges of nodes in the networks can be adjusted flexibly for the concern of green communication. Minimum Energy Problem in Multicasting under the condition of fixed transmission range is a special case of that under the condition of variable transmission range. In fact, adjustable range bring the benefits of less energy consumption. Often the transmission range can be reduced so as to exactly cover the next hop for the purpose of energy saving. In this section,we assume that the transmission range can be changed and revisit the Minimum Energy Problem in Multicasting.

4.2.1 Energy Benefits of Transmission Range Adjustment

We display the network topology at two time slots in Fig. 4–2, and the Euclidean distance between nodes are also shown besides the dashed lines in this figure. The successful communication between nodes is determined by the transmission range and their Euclidean distance. As is shown in Fig. 4.2(a), there are four nodes in the network where A is the source, and $D1, D2$ are two destinations.

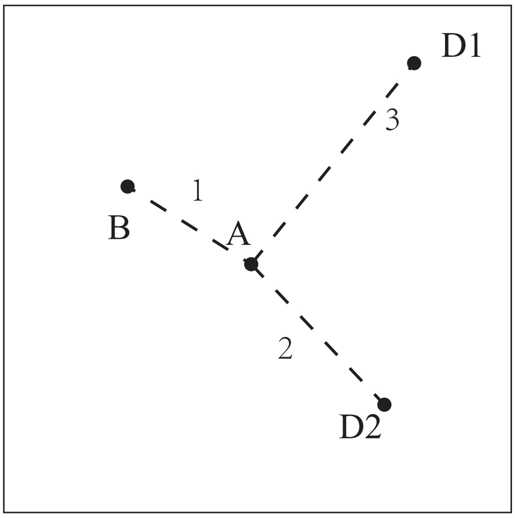
An example of transmission scheduling is applied to multicasting in both cases of variable range and fixed range. By comparing the energy consumption in both cases, we point out that the energy benefits can be brought about by the flexible range adjustment.

Suppose that the relationship between energy consumption E and transmission range r is $E = r^{\alpha}$, and we take $\alpha = 2$. We first assume that the transmission range is fixed as

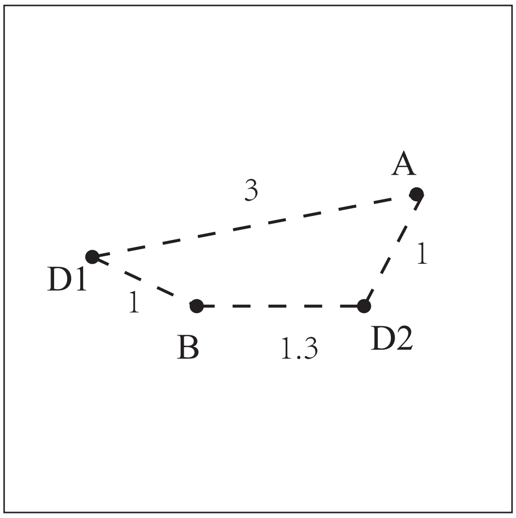
Table 4–1

Table 4–1 Algorithm complexity

Variable	Update complexity of one variable	Maximum number of variables	Update complexity of variables of the same type
$I(u,t)$	mnd	nd	mn^2d^2
$I(j)$	d	$\|J\|$	$d\|J\|$
w_{uD_it}	$n\|J_D\|$	mnd	$n^2d\|J\|$



(a) $t = 1$



(b) $t = 2$

Figure 4–2

Fig 4–2 An example of energy minimization multicasting when the transmission range is variable given the network topology at two time slots

$r = 1.2$, and discuss the minimal energy consumed for multicasting. The delay constraint is set as 2. The delay bounded journey for $D1$ is $\{(A, 0), (B, 1), (D1, 2)\}$, and the journey for $D2$ is $\{(A, 0), (D2, 2)\}$. We choose the journeys for information multicasting to two destinations, and the energy consumption E_{fixed} is $E_{fixed} = 3 \times 1.2^2 = 4.32$.

Now we assume that the transmission range can be adjusted, and the maximal range is $r_{max} = 1.8$. Therefore, all nodes can flexibly change their ranges from 0 to 2. When the delay constraint is 2, the delay bounded journey for $D1$ is $\{(A, 0), (B, 1), (D1, 2)\}$. Two journeys for destination $D2$ are $\{(A, 0), (B, 1), (D2, 2)\}$ and $\{(A, 0), (D2, 2)\}$. We select the journeys $\{(A, 0), (B, 1), (D1, 2)\}$ and $\{(A, 0), (B, 1), (D2, 2)\}$ in order to minimize the energy consumption. The total energy $E_{variable}$ needed in this case is $E_{variable} = 1^2 + 1.3^2 = 2.69$. Clearly, the adjustment of transmission range allows energy saving, which motivates us to explore more flexible multicasting scheduling in the case of variable transmission range.

4.2.2 Complexity Analysis

Intuitively, the Minimum Energy Problem in the case of variable transmission range is of more complexity than the case of fixed transmission range. Firstly, both journey and transmission are directed given the variable transmission range, which adds to the computation workload. In the dynamic graph capturing the characteristics of mobile networks, the journey is directed. If the journey from A to B to C exists, the journey from C to B to

A does not necessarily exist due to node mobility. When we suppose that the transmission range of nodes can be adjusted to minimize the energy consumption, it is very likely that different nodes set different transmission ranges. Unequal ranges lead to the asymmetric communication among nodes. For example, there are two nodes A and B with transmission range r_A and r_B respectively. The Euclidean distance between these two nodes is d_{AB}. Without loss of generality, we suppose that $r_A > d_{AB} > r_B$. The transmission from A to B is successful since r_A is larger than their geographical distance. However, the transmission from B to A fails in that r_B is not large enough to cover node A. It is clear that the transmission is also directed under the condition of variable transmission range.

Secondly, we should consider the mutual influence of range adjustment and transmission scheduling. When a journey is selected for message from the source to the destination, the transmission range of each node needs adjustment and grows to be large enough so as to cover the next hop. Moreover, when the transmission range is reduced below the distance between two hops, the information cannot be forwarded along the journey. Accordingly, a new journey should be selected to ensure that destinations can receive message with the preset transmission range. It can be seen that we should adjust both the transmission range and journey at the same time to ensure the successful mutlicasting. Besides, we should always walk towards minimizing energy cost during our adjustments, and many iterations of adjustments are unavoidable.

Thirdly, we introduce a new variable to our analysis, the transmission range. Given the fixed range, we should only decide the routes for message dissemination together with the specific forwarding time in making transmission scheduling policies. Assuming that range is changeable, we should not only determine the routes and time, but also set the transmission range for each transmitting node at each time slot.

Fourthly, variable transmission range brings about more uncertainty. Mobility results in dynamic network topology, and the connectivity among nodes are time-varying. Furthermore, even if the nodes are static, the connectivity is changing with time in the network as the transmission range is adjusted from time to time. The network behavior should be more dynamic and responsive in the face of uncertain range adjustment.

In Theorem 4.2.1, we prove the NP-hardness of MEPM with variable transmission range using strict and logical complexity analysis.

Theorem 4.2.1.
The Minimum Energy Problem in Multicasting is an NP-hard problem under the assumption that the transmission range can be adjusted at each time slot for all nodes.

Proof. We first show that weighted set cover problem can be reduced to the Minimum

Energy Problem in Multicasting. Since weighted set cover problem is NP-hard, we can prove that minimum transmission problem is also an NP-hard problem. Assume that there is a non-empty set A and $\forall i \in \{1, 2, ..., m\}, A_i \subseteq A$. The weight of the subset A_i is w_i. A set S contains the subsets, and $S = \{A_1, ..., A_m\}$. The minimum-weighted set cover problem is defined to select some subsets in S so that they contain all elements in set A and have the minimum total weight.

Now we establish the correspondence between the two problems in the way similar to the reduction from Minimum Set Cover Problem to MEPM. For each element in A, we create a destination D_i in the destination set D. For each subset A_i, we create a relay H_i accordingly, and we also assume that H_i are engaged in the journeys: source $\rightarrow H_i \rightarrow D_j$ where D_j corresponds to the element in A_i. The weight of A_i is the energy consumed for both the transmissions from the source to H_i and from H_i to the destinations. It is supposed that no other delay bounded journeys exist and messages are forwarded to all destinations at the same time slot by a relay. The prohibition of simultaneous transmission from a relay to all its destinations excludes the probability that one relay consumes energy less than the subset weight when it only communicates with some of the destinations corresponding to the elements in the subset. Besides, it is also assumed that the distances between the relay and destinations are the same. It avoids the case that the energy consumption of one journey is reduced when the relay forwards messages to part of the closer destinations.

The reduction has been completed, and then we continue to show that the solution to MEPM can be transformed into the solution to Minimum Weighted Set Cover Problem. Once a relay is chosen, the energy consumption of the transmissions related with the relay is fixed including the energy needed for the transmission from the source to this relay and the transmission from the relay to the destinations. The selection of relay also determines the journeys from the source to destinations as well as the energy needed for transmission. When the journeys consuming minimal energy from the source to destinations are selected, the minimal weight subsets are those corresponding to the selected relays. Considering that the journeys can reach all destinations in D, the subsets correspondent to the selected relays cover all elements in A. Hence these selected subsets are exactly the minimum-weighted set cover. This completes our proof. □

4.2.3 Problem's Mathematical Formulation

We formulate the MEPM as an optimization problem under the assumption that the transmission range can be adjusted. Here we consider the Minimum Energy Problem in discrete time. Still, we first define some new variables for the clarity of explanation. Let d_{uv}^t be the Euclidean distance between u and v at time slot t, and r_u^t be the transmission range at time slot t. Let $E\left(r_u^t\right)$ be the energy consumed by node u when it forwards messages with transmission range r_u^t. Let $I(e_{uv},t)$ be the indicator function, and $I(e_{uv},t)=1$ if the edge e_{uv} exists at time slot t. Moreover, J_{D_i} is the set of all delay-bounded journeys from the source to the destination D_i, and $I(j)=1$ if the journey j is chosen from the source to the destination. The optimization is formulated as follows.

$$\min \sum_{t=1}^{d}\sum_{u\in V} E\left(r_u^t\right)$$

$$s.t. \sum_{j\in J_{D_i}} I(j)=1, \forall D_i\in D \tag{4–23}$$

$$\sum_{t=1}^{d} I\left(e_{uv},t\right) \geq \sum_{j\in J_{D_i}} I(j)\delta\left(j,e_{uvt}\right), \forall u,v\in V, \forall D_i\in D \tag{4–24}$$

$$\sum_{t=1}^{d}\sum_{u\in V} I\left(e_{uv},t\right)\leq 1, \forall v\in V \tag{4–25}$$

$$d_{uv}^t I\left(e_{uv},t\right)\leq r_u^t, \forall u,v\in V, t\ is\ an\ integer\ in [0,d] \tag{4–26}$$

$$r_u^t\leq R, \forall u\in V, t\ is\ an\ integer\ in [0,d] \tag{4–27}$$

$$I(e_{uv},t)\in\{0,1\}, \forall u,v\in V, t\ is\ an\ integer\ in [0,d] \tag{4–28}$$

$$I(j)\in\{0,1\}, \forall D_i\in D, \forall j\in J_{D_i} \tag{4–29}$$

This optimization problem contains three original variables: r_u^t, $I(j)$ and $I(e_{uv},t)$. The optimization goal is to minimize the energy consumed in all transmissions of multicasting while seven constraints above are satisfied. Constraint 4–23 ensures that every destination can receive one copy of message from the source, and Constraint 4–24 requires that the edge should be chosen once its corresponding journey is selected for transmission. Constraint 4–25 guarantees that every node receives no more than one message so as to avoid redundant transmissions, and Condition 4–26 demands the Euclidean distance between two nodes should be no larger than the transmission range in order to forward information successfully. Due to the broadcast nature of wireless nodes, it is possible that one node can forward messages to multiple nodes at the same time. If this happens, then the sender will set the transmission range that can support the communication with the farthest receiver. Constraint 4–27 imposes the constraint of the maximal transmission range, considering the limited power supply of the mobile devices in practical scenarios. Since $I(\bullet)$ is an indicator function, its value is either 0 or 1.

4.2.4 Fast Approximation

We have given the mathematical expression of the Minimum Energy Problem in Multicasting assuming that the transmission range is adjustable, and we now consider the solution to this optimization problem. We apply the Lagrangian multiplier method and find the corresponding dual problem of MEPM with adjustable transmission range.

4.2.4.1 Lagrangian

We adopt the Lagrangian relaxation method which can give the lower bound of the optimal solution, and the multipliers are the weights assigned to the constraints in the optimizations problem. The multipliers are also the dual variables in the dual problem. The Lagrangian function L is the sum of objective function and the weighted sum of all constraints in the original problem.

$$L\left(w^1_{uvD_i}, w^2_v, w^3_{uvt}, w^4_{ut}\right) = \min_{I(j), I(e_{uv},t), r^t_u} \Bigg(\sum_{t=1}^{d}\sum_{u\in V} E\left(r^t_u\right) + \sum_{u\in V}\sum_{v\in V}\sum_{D_i\in D} w_{uvD_i} \left(\sum_{j\in J_{D_i}} I(j)\delta\left(j, e_{uv}\right) - \sum_{t=1}^{d} I\left(e_{uv}, t\right)\right) + \sum_{v\in V} w^2_v \left(\sum_{t=1}^{d}\sum_{u\in V} I\left(e_{uv}, t\right) - 1\right) + \sum_{t=1}^{d}\sum_{u\in V}\sum_{v\in V} w^3_{uvt}\left(d^t_{uv} I\left(e_{uv}, t\right) - r^t_u\right) + \sum_{t=1}^{d}\sum_{u\in V} w^4_{ut}\left(r^t_u - R\right) \Bigg) \tag{4–30}$$

To find the values of $I(j)$, $I(e_{uv}, t)$ and r^t_u and minimize the Lagrangian function, we divide the complex expression into three sub-problems L_1, L_2 and L_3. The first subproblem L_1 only contains the variable $r(u,t)$, and it is:

$$\begin{aligned} L_1 &= \min_{r^t_u}\left(\sum_{t=1}^{d}\sum_{u\in V} E\left(r^t_u\right) - \sum_{t=1}^{d}\sum_{u\in V}\sum_{v\in V} w^3_{uvt} r^t_u + \sum_{t=1}^{d}\sum_{u\in V} w^4_{ut} r^t_u\right) \\ &= \min_{r^t_u}\left(E\left(r^t_u\right) + \left(w^4_{ut} - \sum_{v\in V} w^3_{uvt}\right) r^t_u\right) \end{aligned} \tag{4–31}$$

Since L_1 is a convex function, and it obtains the minimum when the value of variable is taken so that function's derivative is 0. Therefore the value of $(r^t_u)^*$ satisfies the following equation:

$$E'(r^t_u) + \left(w^4_{ut} - \sum_{v\in V} w^3_{uvt}\right) = 0. \tag{4–32}$$

Here $E'(\bullet)$ is the derivative of energy function.

The second part of the Lagrangian function is L_2 which only contains the original variable $I(e_{uv}, t)$.

$$L_2 = \min_{I(e_{uv},t)} \Big(-\sum_{u\in V}\sum_{v\in V}\sum_{D_i\in D} w^1_{uvD_i} \sum_{t=1}^{d} I(e_{uv}, t) + \sum_{v\in V} w^2_v \sum_{t=1}^{d}\sum_{u\in V} I(e_{uv}, t) .$$
$$+ \sum_{t=1}^{d}\sum_{u\in V}\sum_{v\in V} w^3_{uvt} d^t_{uv} I(e_{uv}, t) \Big)$$
$$= \min_{I(e_{uv},t)} \left(\sum_{t=1}^{d}\sum_{u\in V}\sum_{v\in V} \left(\left(-\sum_{D_i} w^1_{uvD_i} + w^2_v + w^3_{uvt} d^t_{uv} \right) I(e_{uv}, t) \right) \right) . \quad (4–33)$$

Next we need to find the appropriate value of $I(e_{uv}, t)$ in order to minimize the value of Lagrangian function L_2. We can find the minimum of L_2 when the value of $I(e_{uv}, t)$ is chosen in the following manner:

$$I(e_{uv}, t) = \begin{cases} 1, & -\sum\limits_{D_i} w^1_{uvD_i} + w^2_v + w^3_{uvt} d^t_{uv} < 0; \\ 0, & \text{otherwise.} \end{cases} \quad (4–34)$$

The third part of Lagrangian function is L_3, which is only related with variable $I(j)$.

$$L_3 = \min_{I(j)} \left(\sum_{u\in V}\sum_{v\in V}\sum_{D_i\in D} w^1_{uvD_i} \sum_{j\in J_{D_i}} I(j)\delta(j, e_{uv}) \right)$$
$$= \min_{I(j)} \left(\sum_{u\in V}\sum_{v\in V}\sum_{D_i\in D}\sum_{j\in J_{D_i}} w^1_{uvD_i} \delta(j, e_{uv}) I(j) \right) . \quad (4–35)$$

Unlike L_1 and L_2, the minimum of L_3 can not be directly obtained. We first construct a weighted time-varying graph $G_{D_i} = \{V_{D_i}, E_{D_i}\}$ with vertex set V_{D_i} and edge set E_{D_i} for each destination D_i. Vertex set consists of all nodes in the network, and an edge exists between two nodes u and v if and only if $\delta(j, e_{uv}) = 1$. The value of $w^1_{uvD_i}$ is assigned to the weight of edge e_{uv}. To minimize L_3, we just need to find the minimum-weighted journey from the source to D_i. It is possible that the edge weight is a negative value, so we can apply Bellman-Ford algorithm in order to find the minimal value of L_3.

There are some remaining items in Lagrangian function L:

$$-\sum_{v\in V} w^2_v - \sum_{t=1}^{d} R w^4_{ut} \quad (4–36)$$

This part can be regarded as a constant since the variables in Lagrangian L are r^t_u, $I(e_{uv}, t)$ and $I(j)$.

When solving the three Lagrangian functions above, we can further reduce the complexity in solving the Lagrangian by narrowing the range of edge selection e_{uv}. A new variable L_t is defined to be the edge existing at the time slot t when all nodes set the maximal transmission range, and thus L_t contains all possible connections between nodes. In fact, we do not need to take all pairs of nodes in the networks into consideration, and we only consider $e_{uv} \in L_t$ instead.

4.2.4.2 Dual Problem

Minimum Energy Problem in Multicasting is the original problem we target, and now we find its dual problem DP which is defined as:

$$\max_{w^1_{uvD_i}, w^2_v, w^3_{uvt}, w^4_{ut}} (L_1 + L_2 + L_3)$$

$$s.t. w_{uvD_i} \geq 0, \forall u, v \in V, \forall D_i \in D \tag{4–37}$$

$$w^2_v \geq 0, \forall v \in V \tag{4–38}$$

$$w^3_{uvt} \geq 0, \forall u, v \in V, t\ is\ an\ integer\ in [1, d] \tag{4–39}$$

$$w^4_{ut} \geq 0, \forall u \in V, t\ is\ an\ integer\ in [1, d] \tag{4–40}$$

The dual problem DP aims to maximize the Lagrangian function, and the dual variables are required to be non-negative. Then we adopt the subgradient method to update the dual variables in each iteration until the optimality is achieved. We introduce the variable Δ which is the step size reflecting the speed at which dual variables are changed.

For dual variable $w^1_{uvD_i}$, it is updated in the following manner:

$$w^1_{uvD_i}(T+1) = \left(w^1_{uvD_i}(T) + \Delta \left(\sum_{j \in J_{D_i}} I(j) \delta(j, e_{uv}) - \sum_{t=1}^{d} I(e_{uv}, t) \right) \right)^+ \tag{4–41}$$

The update of dual variable w^2_v is below:

$$w^2_v(T+1) = \left(w^2_v(T) + \Delta \left(\sum_{t=1}^{d} \sum_{u \in V} I(e_{uv}, t) - 1 \right) \right)^+ \tag{4–42}$$

For dual variable w^3_{uvt}, it is changed like:

$$w^3_{uvt}(T+1) = \left(w^3_{uvt}(T) + \Delta \left(d^t_{uv} I(e_{uv}, t) - r^t_u \right) \right)^+ \tag{4–43}$$

Dual variable w^4_{ut} is renewed as follows:

$$w^4_{ut}(T+1) = \left(w^4_{ut}(T) + \Delta \left(r^t_u - R \right) \right)^+ \tag{4–44}$$

Recall that the dual variables should be non-negative. In each iteration, we update four types of variables with the equations above. The value of a dual variable is set as the result of the expression when the result in non-negative. Otherwise, the dual variable's value is set as 0.

The minimum energy can be obtained through multiple iterations. The three original variables are updated based on the value of the dual variables, and then the dual variables are also renewed according to the new values of original variables. The value of objective function in the dual problem, i.e., energy consumed for multicasting, gradually converges to a tight lower bound. We also obtain the minimum energy consumption with the fast approximate algorithm among all the feasible transmission scheduling derived in many iterations. Then we can decide the multicast routing and make the transmission scheduling based on the value of original variable r_u^t. For any $r(u,t) \neq 0$, node u sets its transmission range as r_u^t and forwards the messages at the time slot t.

4.2.5 Complexity of Fast Approximation

Table 4–2

Table 4–2 Algorithm complexity

Variable	Update complexity of one variable	Maximum number of variables	Update complexity of variables of the same type
$I(j)$	$d\|J\|$	m	$dm\|J\|$
r_u^t	n	nd	n^2d
$I(e_{uv},t)$	m	n^2d	mn^2d
$w_{uvD_i}^1$	$\|J_D\|+d$	mn^2	$d\|J\|$
w_v^2	nd	n	n^2d
w_{uvt}^3	1	n^2d	n^2d
w_{ut}^4	1	nd	nd

Considering the NP-hardness of Minimum Energy Problem in Multicasting, we reduce the time complexity at the cost of energy consumption. We have proposed the approximate algorithm for this problem given the variable transmission range based on the Lagrangian relaxation method.

Now we analyze the time complexity of this approximate algorithm theoretically. Since the major part of this approximation is the update of original as well as dual variables, we display the complexity of variable update in the table 4–2. We define the following variables for the clarity of explanation, and denote $|J|$ as the total number of delay-bounded journeys, $|J_D|$ as the maximal number of journeys from source to one destination,

d as the delay bound, n as the network size and m as the number of destinations.

If the number of iterations we set is p, the algorithm complexity T is

$$T = p * \left(dm|J| + n^2d + mn^2d + d|J| + n^2d + n^2d + nd\right) \tag{4–45}$$

$$= O\left(pdm\left(|J| + n^2\right)\right). \tag{4–46}$$

The algorithm complexity is $O\left(pdm\left(|J| + n^2\right)\right)$, and it can be reduced when the step size of variable update is appropriately adjusted. The time cost is also related with the network size, the number of destinations as well as delay constraint. Beside, the connectivity of the whole network can also affect the algorithm complexity. If the connectivity is quite high, there are many time-bounded journeys available for information transmission. Hence the large value of $|J|$ may lead to more time costs.

Chapter 5 Simulation

5.1 Simulation

In this part, we perform extensive simulations about multicasting in mobile networks for the following purposes:

- Verify the conjecture of tradeoff between delay and energy consumption;
- Explore the energy efficiency gain brought by mobility;
- Evaluate the efficiency of fast approximations for Minimum Energy Problem in Multicasting in comparison with the optimal energy consumption;
- Compare the energy for multicasting in the case of fixed and variable transmission range.

5.1.1 Mobility Models

To consider the general mobility, three popular mobility models are applied in our simulations: i.i.d model, random waypoint model and Gauss-Markov model. Suppose that all nodes are distributed in a square with edge length of $a = 100$, and their movement obeys the mobility models described below. Following the introduction of mobility model, we also present the node's movement trail in this model.

i.i.d Model: This model assumes that all mobile nodes are independently and uniformly distributed in the simulation area at each time slot. A node's trail is shown in Fig. 5–1.

Random Waypoint Model: It is assumed that the moving direction and/or speed of each node change after a time interval in Random Waypoint Model, and one node chooses a random direction and a speed which is uniformly distributed in $[Min(speed), Max(speed)]$ [34]. The movement trail described by this model is shown in Fig. 5–2.

Gauss-Markov Model: The speed and moving direction are updated after a fixed time interval in Gauss-Markov Model [34]. The movement state of one node at the n^{th} time slot depends on the $(n-1)^{st}$ slot. Let s_n and $\bar{s}$ be the speed at the n^{th} slot and the mean speed, and similarly let d_n and $\bar{d}$ be the moving direction at the n^{th} slot and its mean value during the infinite time interval. The values of $s_{x_{n-1}}$ and $d_{x_{n-1}}$ are random values

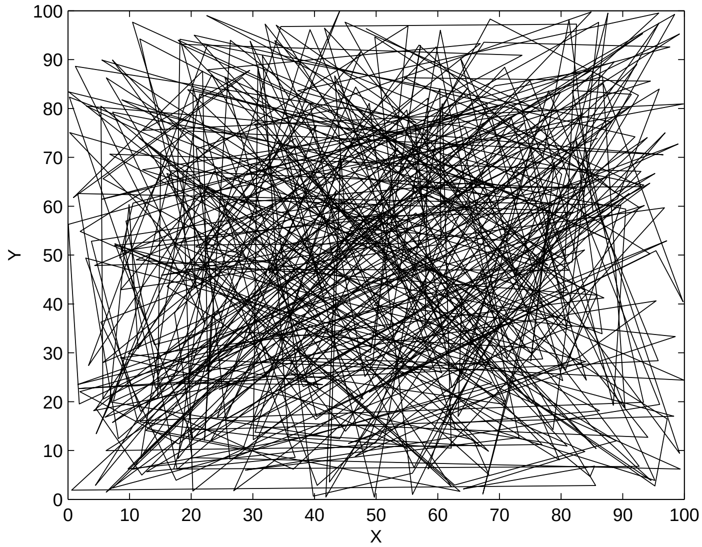


Figure 5–1 Motion Trail with iid model

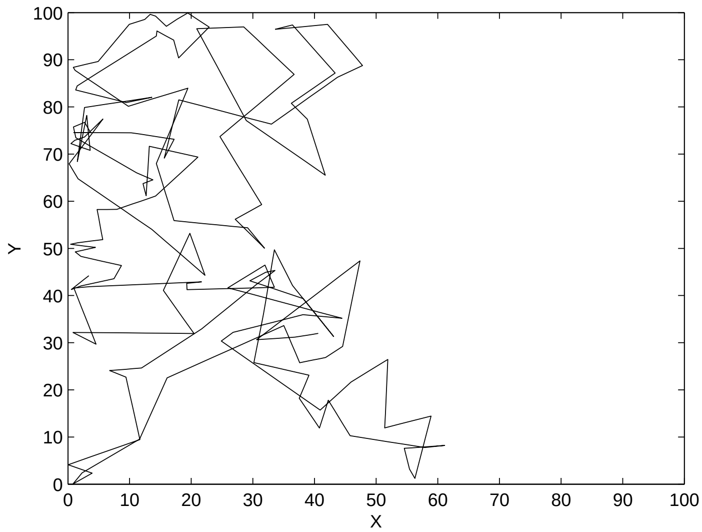


Figure 5–2 Motion Trail with random waypoint model

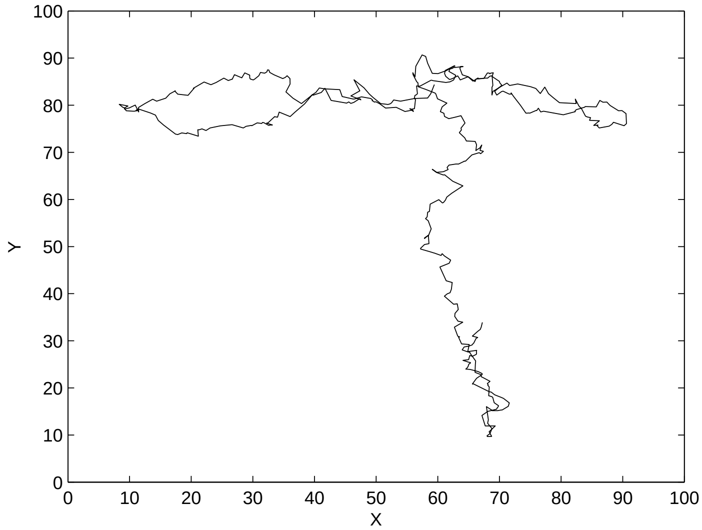


Figure 5–3 Motion Trail with Gauss-Markov model

which are derived from the Gaussian distribution. It is assumed that:

$$s_n = \alpha s_{n-1} + (1-\alpha)\bar{s} + \sqrt{(1-\alpha^2)}s_{x_{n-1}}.$$

$$d_n = \alpha d_{n-1} + (1-\alpha)\bar{d} + \sqrt{(1-\alpha^2)}d_{x_{n-1}}.$$

The node motion trail given the Gauss-Markov model is plotted in Fig. 5–3.

Here we define some parameters in our simulations. Transmission range is denoted as r, delay constraint is denoted as d, and we use an easing factor β to quantize it. It is defined that

$$\beta = \frac{Delay\ constraint}{Minimal\ delay}. \tag{5–1}$$

The relationship between energy consumption $E(r)$ and the transmission range r is defined as:

$$E(r) = r^{\alpha}. \tag{5–2}$$

In general, $\alpha = 2 \sim 4$. We take $r = 2$ in our simulations.

5.1.2 Efficiency Gain Brought by Mobility

We assume that the transmission range is fixed for all nodes, and we apply the Gauss-Markov model, which can most accurately describe node mobility in real scenarios among three models, to simulate the node movement. Let N_{trans} be the number of transmissions,

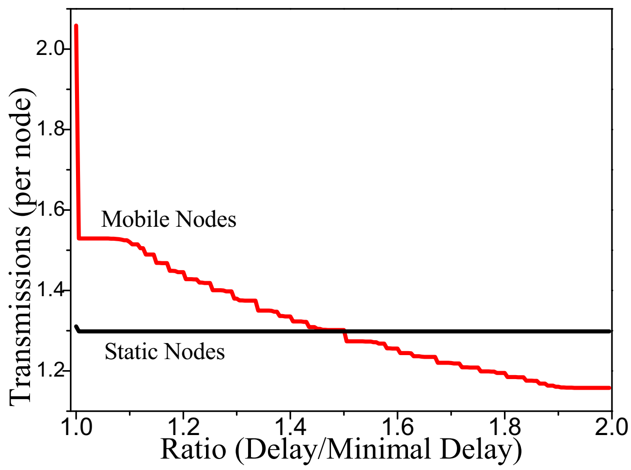


Figure 5–4 Energy comparison between mobile and static networks

and its average is:

$$avg(N_{trans}) = \frac{Total\ number\ of\ transmissions}{The\ number\ of\ destinations}. \tag{5–3}$$

Now we explore how the relaxation of delay bound can improve energy efficiency by exploring the number of necessary transmissions under different delay constraints. Moreover, we also explore how the mobility can bring energy saving by comparing the number of transmissions in both mobile and static networks. In Fig. 5–4, the energy consumption is quite high when the information are multicast when it is transmitted with the shortest delay. Once the delay constraint is relaxed, the energy decreases sharply. Energy efficiency can be further improved when we continue to relax the delay constraint, which indicates the tradeoff between transmission delay and energy efficiency.

In Fig. 5–4, we also compare the energy consumption in static and mobile networks when the same delay constraint is imposed. When the delay constraint is relaxed, the energy consumed for transmission is non-increasing in both cases. If the tolerance of delay is small, static networks performs better in saving energy than mobile networks. The opportunities to forward the messages might be missed by mobile nodes since the links are dynamic and may not support the data transmission. It is likely that messages should pass through more mobile hops than static hops in order to arrive at the destination. The curve of mobile nodes is above that of static nodes when the delay ratio is small. However, when we continue to relax the delay constraint, the benefits brought by the node mobility outweigh the cost brought by its dynamic topology. The reason is that one node is more likely to move close to another node that is previously far from it when longer delay is

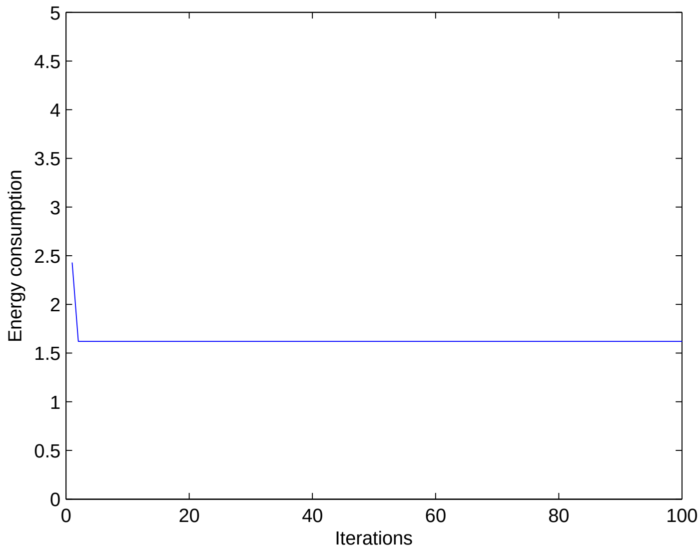


Figure 5–5 Stability of Scheduling with the fixed range

allowed. Geographical distance can be narrowed at the cost of time, and fewer hops are engaged for transmission. Mobile networks have higher energy efficiency than static ones in the case of large delay tolerance.

The intersection of two curves in Fig. 5–4 indicates that mobility brings benefits of energy efficiency when the delay constraint is not so strict. The flat curve of static nodes shows that the increase of delay does not bring too much improvement in energy efficiency. The scheduling to achieve minimum transmissions in static networks is quite fixed even if the delay constraint is relaxed, since its network topology is unchanged.

5.1.3 MEPM with Fixed Transmission Range

Suppose that the network consists of 20 nodes among which 5 nodes are selected as destinations. Three mobility models are applied to simulations and we show the energy consumption of both approximate scheduling as well as the optimal energy consumption. We assume that the transmission range of wireless nodes are unadjustable, and hence the energy consumption for multicasting is only determined by the number of transmissions, N_{trans}.

With the our fast approximation proposed, the transmission scheduling is updated in each iteration, and the energy consumption in dual problem converges gradually when both the original and dual variables become stable. In Fig. 5–5, the change of achievable energy consumption in each iteration is shown. The stable value is the energy consumption achieved with our approximate scheduling.

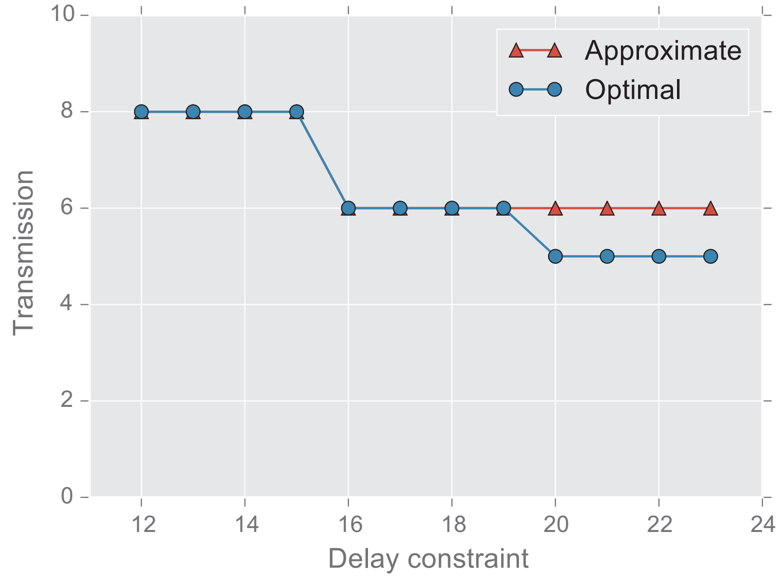


Figure 5–6 Comparison of approximate and optimal scheduling in i.i.d model

5.1.3.1 I.i.d. Model

We generate the node locations at different time slots with i.i.d mobility model, and the location information of 20 nodes is shown in Table 5–1. The first node is designated as the source, and nodes numbered $2-5$ are the destinations. The selections of source and destinations are the same with different mobility models. Due to page limit, we only show the location of multicast members in first eight time slots.

Table 5–1 Node location in i.i.d model

Slot / Point	1	2	3	4	5	6	7	8
1	(78.2310, 89.5810)	(52.5220, 82.6050)	(25.2610, 81.6920)	(80.9720, 39.1120)	(64.8100, 10.1210)	(50.5170, 67.0380)	(49.8970, 76.1840)	(26.3410 11.9570)
2	(41.9820, 12.6240)	(7.9720, 8.0984)	(13.7120, 25.4840)	(99.1510, 93.8830)	(23.9760, 16.8620)	(20.6770, 94.8490)	(72.9630, 58.3030)	(70.2950 36.9040)
3	(75.6360, 21.2050)	(20.3230, 51.9030)	(0.5639, 19.5980)	(95.3630, 6.7298)	(91.6750, 28.1400)	(99.7660, 61.3610)	(70.4700, 12.3740)	(84.7810 23.7850)
4	(28.0680, 32.9410)	(7.4976, 22.4090)	(35.4710, 52.4310)	(88.8430, 94.5670)	(36.3620, 75.1980)	(16.5680, 81.7850)	(41.5620, 2.1085)	(13.0800 9.4700)
5	(60.5690, 91.2070)	(2.8026, 53.3040)	(63.6370, 20.7520)	(3.1199, 59.1690)	(46.9600, 76.0100)	(94.9670, 75.5020)	(19.9610, 62.9960)	(86.7970 7.3353)
6	(84.4050, 51.2580)	(80.7730, 50.7130)	(1.3931, 55.1360)	(28.9800, 12.9880)	(51.1630, 20.8740)	(62.8570, 63.3800)	(70.7620, 6.0367)	(65.2610 60.6300)

The transmission range is set as $r = 8$, and the minimal delay is $d_{min} = 12$ given the node movement information above. We explore the relationship between the number of transmissions and the delay constraint, and we also compare the energy consumed in the approximate and optimal scheduling. As is shown in Fig. 5–6, the approximate scheduling consumes about 6.15% more energy than the optimal scheduling.

5.1.4 Random Waypoint Model

The node locations are generated when 20 nodes are moving according to the random waypoint mobility model. Their location are shown in Table 5–2.

Table 5–2 Node location in random waypoint model

Slot / Point	1	2	3	4	5	6	7	8
1	(66.4330, 32.3600)	(66.4330, 32.3600)	(68.5270, 34.1260)	(70.6200, 35.8930)	(72.7140, 37.6590)	(72.7140, 37.6590)	(73.9810, 38.2230)	(73.9810 38.2230
2	(48.6860, 84.0350)	(48.6860, 84.0350)	(45.5700, 86.6340)	(42.4540, 89.2340)	(39.3380, 91.8330)	(36.2220, 94.4330)	(36.2220, 94.4330)	(33.5160 95.7120
3	(71.1740, 34.3830)	(71.1740, 34.3830)	(72.3810, 35.2930)	(72.3810, 35.2930)	(71.4800, 32.3760)	(70.5790, 29.4590)	(70.5790, 29.4590)	(70.6420 30.7940
4	(72.3950, 66.2540)	(72.3950, 66.2540)	(75.0560, 65.6280)	(77.7160, 65.0020)	(80.3770, 64.3760)	(83.0380, 63.7490)	(83.0380, 63.7490)	(82.1680 65.6750
5	(95.5060, 72.5220)	(95.5060, 72.5220)	(92.9520, 75.0780)	(90.3990, 77.6340)	(87.8450, 80.1910)	(85.2920, 82.7470)	(82.7390, 85.3040)	(82.7390 85.3040
6	(54.5470, 69.6830)	(54.5470, 69.6830)	(50.1400, 67.4990)	(45.7320, 65.3140)	(45.7320, 65.3140)	(45.4070, 67.4320)	(45.0820, 69.5500)	(44.7570 71.6680

The transmission range is set as $r = 27$, and the minimal delay is $d_{min} = 8$ in this case. The relationship between the transmission count and delay is shown in Fig. 5–7. The approximate scheduling exactly achieves the optimal energy consumption assuming the transmission range is fixed.

5.1.4.1 Gauss-Markov Model

We use Gauss-Markov mobility model to show the approximate and optimal scheduling from two dimensions of delay and energy. Their mobility information is presented in Table 5–3.

The transmission range is set as $r = 24$, and the minimal delay is $d_{min} = 12$. No more than 13% energy is consumed in the approximate scheduling, and it has 10% of energy consumption on average when compared with the optimal transmission scheduling.

In the simulation, we take the node location information as the input, and return the transmission counts in both cases of approximate scheduling and optimal scheduling. The statistics above shows the performance of data scheduling with only one group of node locations. Now we perform extensive simulations, using these three mobility models to

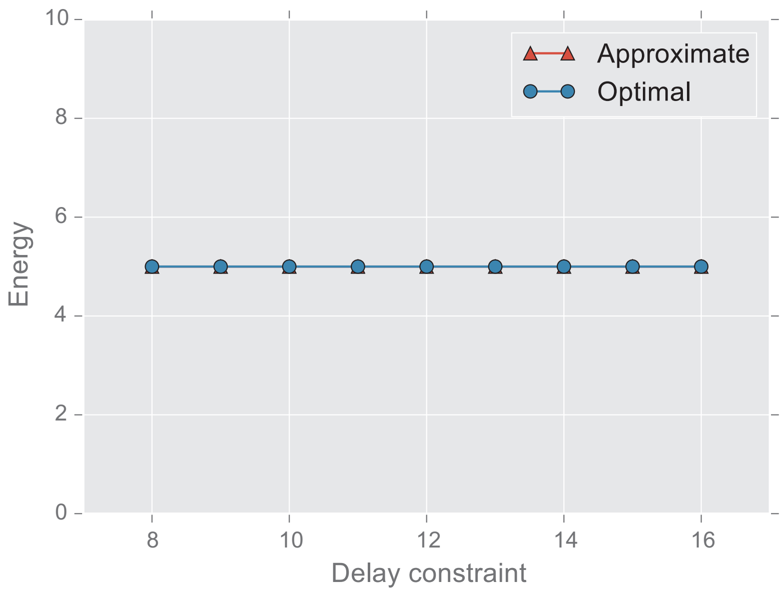


Figure 5–7 Comparison of approximate and optimal scheduling in random waypoint model

Table 5–3 Node location in Gauss-Markov model

Point \ Slot	1	2	3	4	5	6	7	8
1	(18.4220, 20.3010)	(17.9430, 21.5760)	(16.9270, 22.7250)	(15.9110, 22.8430)	(15.4170, 22.3890)	(13.4430, 22.9820)	(12.9850, 23.2660)	(13.8860 23.9760)
2	(70.6740, 64.1590)	(70.6100, 63.9030)	(70.9340, 63.5910)	(70.7100, 63.2490)	(70.5200, 62.8460)	(70.5150, 63.0030)	(69.8640, 63.1930)	(69.6500 63.1830)
3	(74.9110, 71.9960)	(75.7820, 70.3440)	(75.9300, 68.1220)	(76.0070, 67.9470)	(76.2430, 67.6000)	(76.1020, 66.7560)	(75.6920, 66.1490)	(75.9300 65.6440)
4	(32.5620, 41.9630)	(32.5940, 41.9490)	(32.7840, 42.3250)	(33.7610, 42.4010)	(34.1600, 43.8130)	(34.9030, 42.7510)	(36.2310, 42.0750)	(35.5200 40.4100)
5	(47.2950, 39.9940)	(45.8570, 39.6550)	(45.4080, 39.3930)	(44.3080, 39.2910)	(43.9720, 39.0910)	(43.5550, 38.6590)	(42.9590, 38.9810)	(42.2000 37.9190)
6	(33.8560, 45.1030)	(35.1940, 44.4910)	(36.2960, 43.8530)	(36.5820, 43.8820)	(36.5910, 43.8850)	(37.5890, 44.7200)	(37.7000, 44.9410)	(37.8320 45.3440)

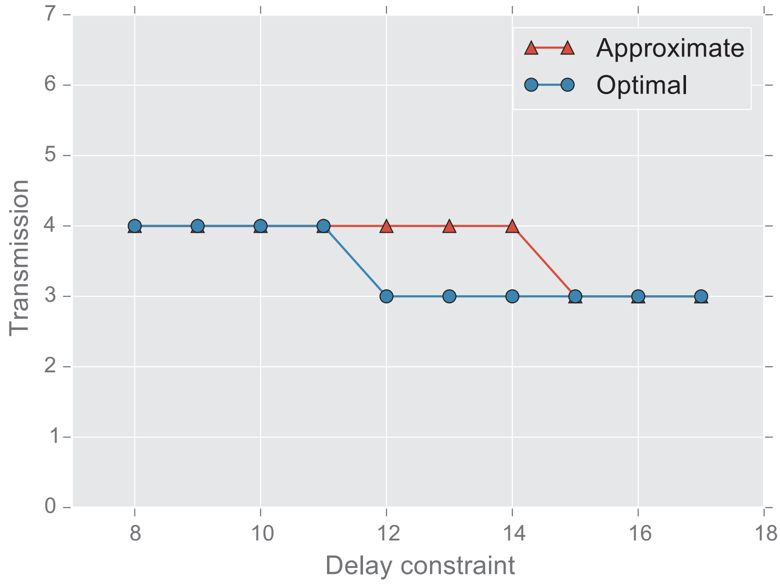


Figure 5–8 Comparison of approximate and optimal scheduling in Gauss-Markov model

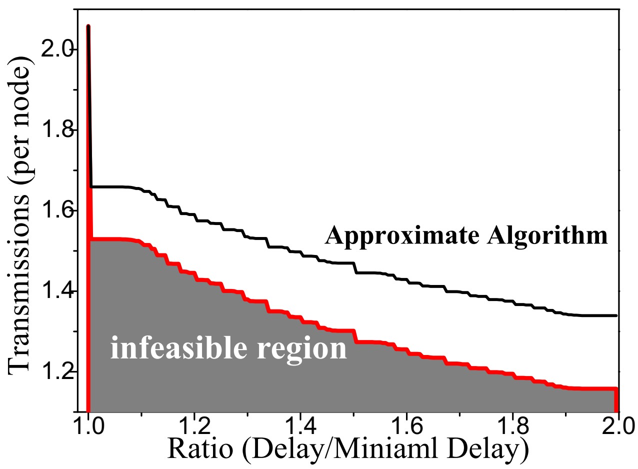


Figure 5–9 Optimal and approximate scheduling for MEPM

generate a large number of inputs. Still, we show how energy and delay influences each other. Fig. 5–9 shows the tradeoff between energy saving and transmission delay and compare the energy efficiency of two multicast scheduling. The red boundary in these figures indicates the minimal transmissions necessary to support multicasting under the delay constraint, and the infeasible region of transmission times is shown by the grey area. Any point with coordinate (x, y) in the infeasible region means that it is impossible to find the routing and transmission scheduling so that only y transmissions are needed on average for messages from the source to reach all destinations within the delay x. The

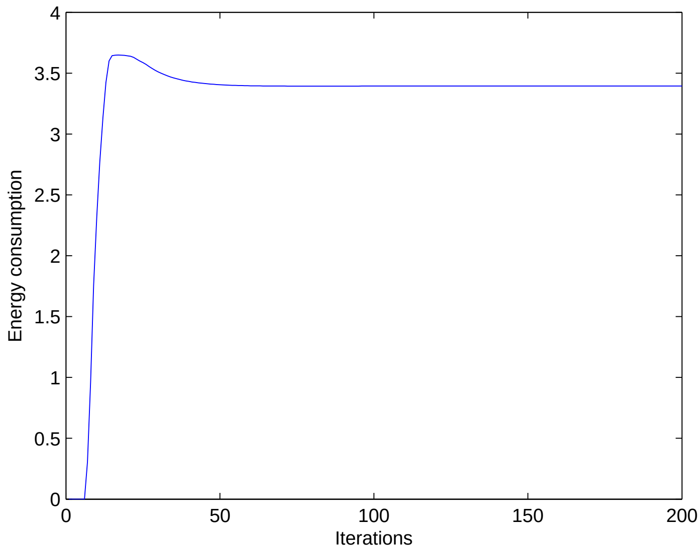


Figure 5–10 Stability of Scheduling with the variable range

energy saving achieved under the delay limitation is plotted by the black curve when our approximate algorithm is applied for transmission scheduling. Statistics show that no more than 16% transmissions are needed in the approximate algorithm in the comparison with the minimal transmissions. Hence it shows that our scheduling is quite efficient in solving the Minimum Energy Problem in Multicasting in the case of fixed range.

5.1.5 MEPM with Variable Transmission Range

A total of 20 nodes are distributed in the network region of 100×100, 5 among which are chosen as destinations. Suppose that the maximum transmission range of each node is R due to the limited energy supply. Now we suppose that the transmission range of wireless nodes can be adjusted provided that it is below the maximum range.

The stable value is the energy consumption achieved with the our approximate scheduling. In each iteration, the energy consumption changes since both original and dual variables are slightly adjusted. When all these variables reach the stability, and the lower bound of energy consumption converges at last. The dynamic adjustment of energy consumption in primal problem is shown in Fig. 5–10.

5.1.5.1 I.i.d Model

The node locations are the same as that in i.i.d model in the case of fixed range. The maximal range is set as $R = 8$. On average, 2.7% more energy is consumed in approximate algorithm than optimum in the i.i.d mobility model when the range is adjustable, and the

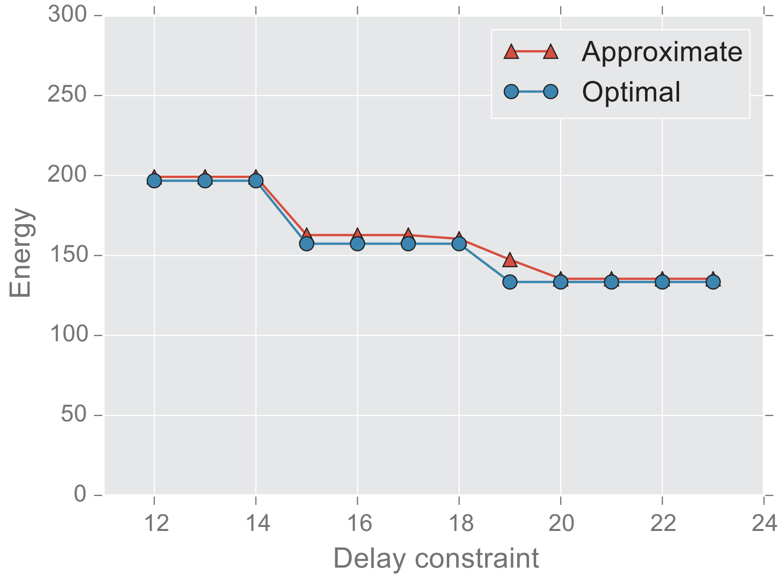


Figure 5–11 Comparison of approximate and optimal scheduling in i.i.d model

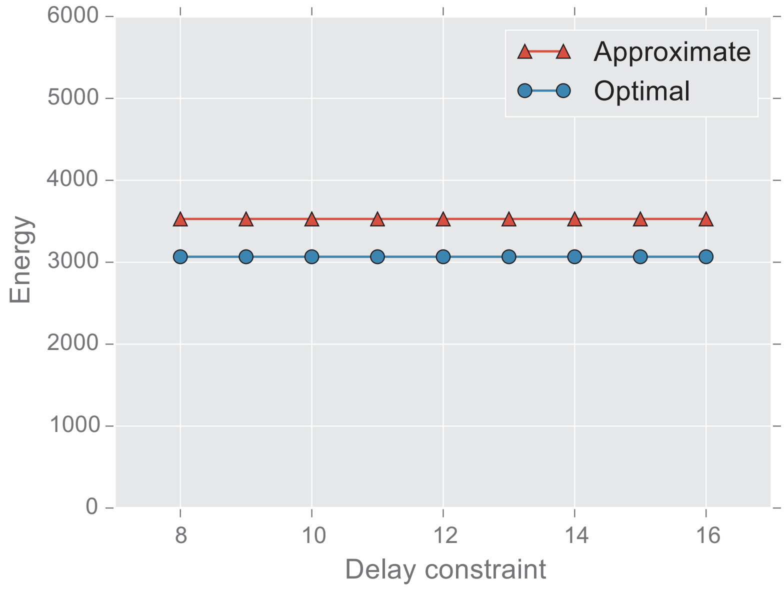


Figure 5–12 Comparison of approximate and optimal scheduling in random waypoint model

extra energy is no more than 3.4% in the worst case.

5.1.5.2 Random Waypoint Model

The node locations are the same as that in random waypoint model, and the maximal range is $R = 27$. Fig. 5–12 gives the comparison of approximate solution and optimum when the Gauss-Markov mobility model is applied. In this simulation, no energy gain is brought with the relaxation of delay constraint. The approximate algorithm costs no more

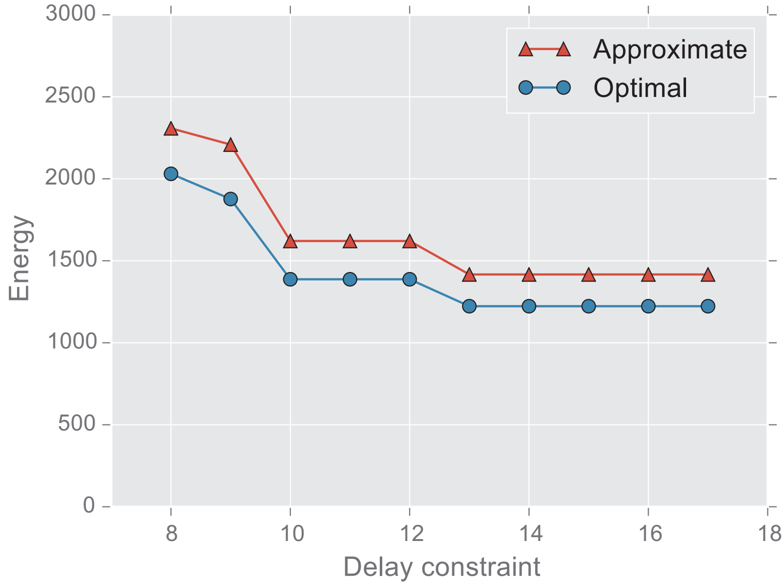


Figure 5–13 Comparison of approximate and optimal scheduling in Gauss-Markov model

than 15% energy according to the statistics obtained from the simulation statistics.

5.1.6 Gauss-Markov model

Nodes' movement follows the Gauss-Markov model, and the maximal range is $R = 24$. The energy consumption is shown in Fig. 5–13 under different delay constraints in two different algorithms for data transmission. The statistics show that the extra energy needed in approximate algorithm is bounded by 17.7%, and is 16% on average. It can be seen that the approximation achieves the scheduling feasibility by consuming tolerable energy.

5.1.7 Comparison of Fixed and Variable Transmission Range

Now let's compare the multicasting performance in two cases of both fixed and variable transmission range. Fig. 5–14 gives the energy comparison when the i.i.d model is applied, Fig. 5–15 adopts the random waypoint model, and Fig. 5–16 explores the energy saving brought by range adjustment with the Gauss-Markov model. In the same model, the node locations at different time slots are the same. The maximal range in the case of variable range is equal to the transmission range in the case of fixed range. Therefore, we can ensure that the energy consumptions are comparable for different range adjustments in the same model.

As is shown in Fig. 5–14, Fig. 5–15 and Fig. 5–16, variable transmission range enables less energy consumption than the fixed range in the same network setting and the

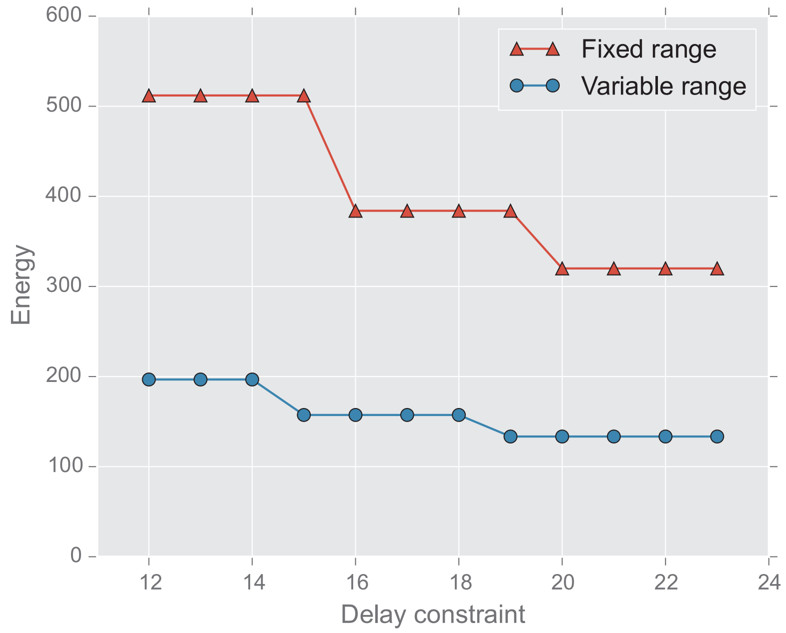


Figure 5–14 Energy comparison of fixed and variable range in i.i.d model

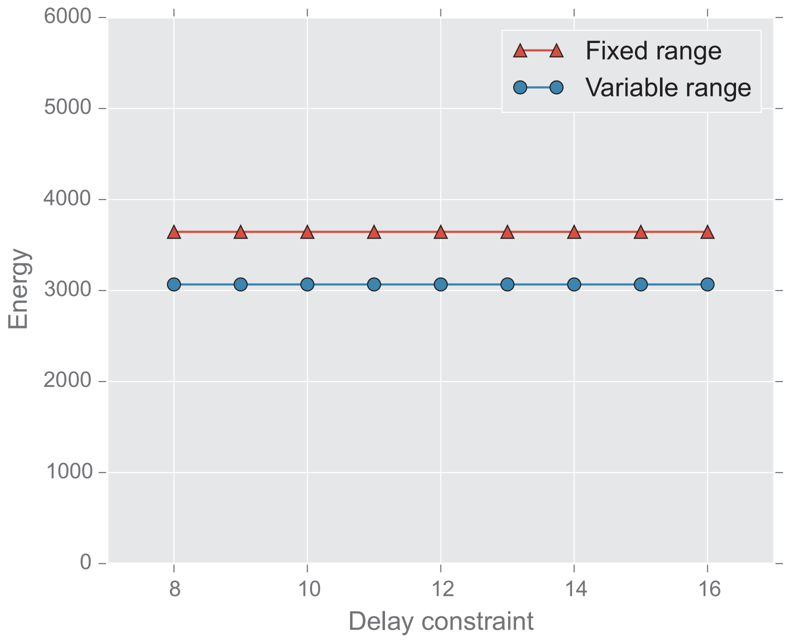


Figure 5–15 Energy comparison of fixed and variable range in random waypoint model

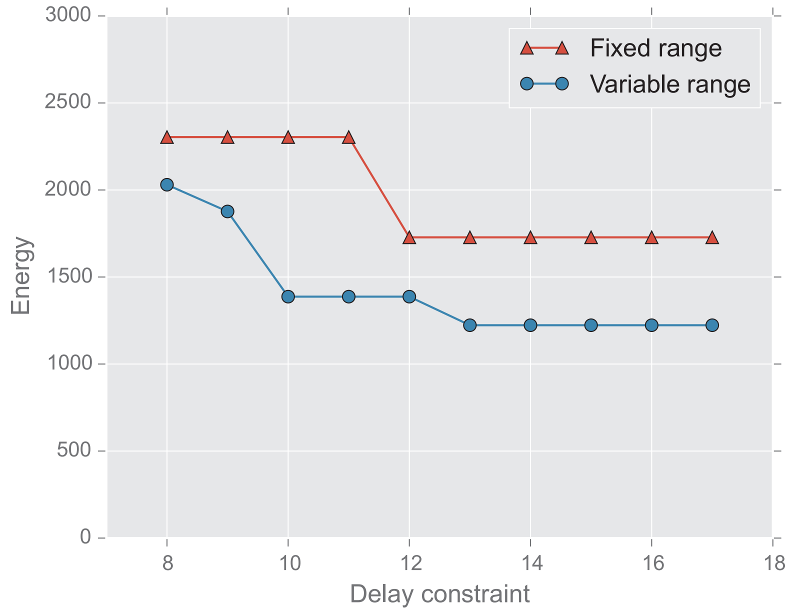


Figure 5–16 Energy comparison of fixed and variable range in Gauss-Markov model

same node mobility model. This is quite intuitive, and the energy consumed in the case of fixed transmission range actually provides the upper bound of energy consumption in the the case of variable transmission range.

The transmission scheduling policy can be made in the case of fixed transmission range. The distance between two nodes directly communicating with each other should be shorter than the transmission range so as to ensure the successful transmission. Therefore, we can further decrease the range in each transmission so that we can reduce energy consumption as well as ensure the successful multicasting. The same transmission policy with reduced transmission range can be the multicasting strategy in the case of variable range. This scheduling also provides an achievable lower bound of energy for the multicast scheduling with variable range.

Summary

In this paper, we study the popular transmission pattern of multicasting, and focus on two important performance metrics: delay and energy consumption in mobile networks. We first propose an algorithm to achieve the minimal-delay transmission among the source and a group of destinations, and show that there is a dedicated tradeoff between energy saving and transmission delay when we try to minimize the energy consumption in multicasting. Relaxing the delay constraint, we explore the feasible region of energy consumption in mobile networks, which is formulated as the Minimum Energy Problem in Multicasting. Two practical cases with either fixed or variable transmission range are taken into consideration, and they are both proved to be NP-hard. The corresponding optimization problems are mathematically formulated, and we apply the Lagrangian relaxation method to give fast approximations to MEPM in both cases. The approximate algorithms are implemented in the simulation part, achieving a near-optimal performance in terms of energy saving and delay constraint. We realize the dynamic control of multicast routing accordingly when the network topology changes with node mobility.

Reference

[1] A. S. K.Lahiri and D.Panigrahi. "*Battery-driven system design: a new frontier in low power design*". In: *Proc. Joint Asia and South Pacific Design Automation Conf./Int. Conf. VLSI Design*, **2002**: 261–267.

[2] P. O.Vermesan P.Friess and et al. "*Internet of things strategic research roadmap*". In: *Internet of Things: Global Technological and Societal Trends*, **2011**: 9–52.

[3] L. L. H. Li and H. V. Poor. "*Multicast routing for decentralized control of cyber physical systems with an application in smart grid*". *Selected Areas in Communications, IEEE Journal*, **2012**, *30*: 1097–1107.

[4] J. Y. D. Li and I. Wu. "*Exploring efficient and scalable multicast routing in future data center networks*". In: *INFOCOM, 2011 Proceedings IEEE*, **2011**: 1368–1376.

[5] Y. G. S. Hua and et al. "*Scalable video multicast in hybrid 3G/ad-hoc networks*". *Multimedia, IEEE Transactions*, **2011**, *13*: 402–413.

[6] T. K. J. Seppala and T. Chen. "*Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks*". In: *Wireless Communications and Networking Conference (WCNC)*, **2011**: 986–991.

[7] Y. Woo and I. Kim. "*Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks*". In: *Vehicle-Based Multicast Communication System for Live-Video*. **2013**.

[8] Y. C. B. Yang and X. Jiang. "*Multicast Delay of Mobile Ad Hoc Networks*". In: *Computing and Networking (CANDAR) in 2014 Second International Symposium*, **2014**: 272–277.

[9] M. A. W. Zhao and E. Zegura. "*Multicasting in delay tolerant networks: semantic models and routing algorithms*". In: *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, **2005**: 268–275.

[10] C. M. Sadler and M. Martonosi. "*Data compression algorithms for energy-constrained devices in delay tolerant networks*". In: *Proceedings of the 4th international conference on Embedded networked sensor systems*, **2006**: 265–278.

[11] V. F. P. Kolios and K. Papadaki. "*Ultra low energy store-carry and forward relaying within the cell*". In: *Vehicular Technology Conference Fall*, **2009**: 1–5.

[12] V. Rodoplu and T. H. Meng. "*Bits-per-joule capacity of energy-limited wireless networks*". In: *Wireless Communications*, **2007**: 857–865.

[13] Z. Chou and Y. Hou. "*Genetic algorithm-based energy efficient multicast scheduling for WiMAX relay networks*". In: *Computing, Networking and Communications (ICNC), 2014 International Conference*, **2014**: 1061–1065.

[14] Z. Niu. "*TANGO: traffic-aware network planning and green operation*". *Wireless Communications, IEEE*, **2011**, *18*: 402–413.

[15] F. D. R. Bolla R. Bruschi and et al. "*Energy efficiency in the future internet: a survey of existing approaches and trends in energy-aware fixed network infrastructures*". *Communications Surveys and Tutorials, IEEE*, **2011**, *13*: 223–244.

[16] M. F. K. Hinton J. Baliga and et al. "*Power consumption and energy efficiency in the internet*". *Network, IEEE*, **2011**, *25*: 6–12.

[17] W. . S. Bai and et al. "*Dear: delay-bounded energy-constrained adaptive routing in wireless sensor networks*". In: *Proc. of IEEE INFOCOM*, **2012**: 1593–1601.

[18] A. Goldsmith and S. Wicker. "*Design challenges for energy-constrained ad hoc wireless networks*". *IEEE Wireless Communications*, **2002**, *9*(4): 8–27.

[19] D. N. T. D. Shah and J. N. Tsitsiklis. "*Hardness of low delay network scheduling*". *Information Theory, IEEE Transactions*, **2011**, *57*: 7810–7817.

[20] O. B. A. Hadachi and et al. "*Cell phone subscribers mobility prediction using enhanced Markov Chain algorithm*". In: *Intelligent Vehicles Symposium Proceedings*, **2014**: 1049–1054.

[21] N. Meghanathan. "*A location prediction based routing protocol and its extensions for multicast and multi-path routing in mobile ad hoc networks*". *Ad Hoc Networks*, **2011**, *9*: 1104–1126.

[22] H. L. B. Kusy and et al. "*Predictive QoS routing to mobile sinks in wireless sensor networks*". In: *Information Processing in Sensor Networks*, **2013**: 1845–1850.

[23] M. O. K. S. Gambs and M. N. del Prado Cortez. "*Next place prediction using mobility markov chains*". In: *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, **2012**: 3.

[24] G. S. T. T. Son H. Le Minh and et al. "*Bayesian model for mobility prediction to support routing in Mobile Ad-Hoc Networks*". In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium*, **2013**: 3186–3190.

[25] M. M. S. M. Mousavi H. R. Rabiee and et al. "*Model based adaptive mobility prediction in mobile ad-hoc networks*". In: *Wireless Communications, Networking and Mobile Computing (WiCom)*, **2007**: 1713–1716.

[26] L. N. N. Makhlouf and P. Vajsar. "*Prediction of movement of wireless nodes in mobile ad-hoc networks*". In: *Cognitive Infocommunications (CogInfoCom), 2013 IEEE 4th International Conference*, **2013**: 115–120.

[27] J. S. F. Cadger K. Curran and et al. "*MANET location prediction using machine learning algorithms*". In: *Wired/Wireless Internet Communication*, **2012**: 174–185.

[28] P. W. M. H. Held and H. D. Crowder. "*Validation of subgradient optimization, Mathematical Programming*". *Mathematical Programming*, **1974**, *6*(1): 62–88.

[29] J. S. Fraser Cadger Cadger Kevin Curran and S. Moffett. *MANET Location Prediction Using Machine Learning Algorithms*. Springer Berlin Heidelberg, **2012**.

[30] Chvatal and Vasek. "*A greedy heuristic for the set-covering problem*". *Mathematics of operations research*, **1979**, *4*(3): 233–235.

[31] Feige and Uriel. "*A threshold of ln n for approximating set cover*". *Journal of the ACM (JACM)*, **1998**, *45*(4): 634–652.

[32] A. H. G. R. K. E. L. Lawler Jan Karel Lenstra and D. B. Shmoys. *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. Wiley, New York, **1985**.

[33] M. S. Bazaraa and H. D. Sherali. "*On the choice of step sizes in subgradient optimization*". *Journal of Operational Research*, **1981**, *7*(2): 380–388.

[34] J. T.Camp and V.Davies. "*A survey of mobility models for ad hoc network research*". *Wireless Communication & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, **2002**, *2*(5): 483–502.

Acknowledgement

My deepest gratitude goes first and foremost to Professor Xinbing Wang , my supervisor, for his constant encouragement and guidance. He has walked me through all the stages of my undergraduate research. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

Secondly, I would like to express my heartfelt gratitude to all the instructors in my undergraduate courses and researches, who lead me into the world of electrical and computer engineering. They contribute a lot to my solid knowledge and deep exploration into this field.

Last my thanks would go to my beloved family for their loving considerations and great confidence in me all through these years. I also owe my sincere gratitude to my friends and my fellow classmates who gave me their help and time in listening to me and helping me work out my problems during the difficult course of the thesis.