ADAPTIVE MANAGEMENT SCHEMES FOR MOBILE AD HOC NETWORKS

by

Sungsoon Cho

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer Science and Engineering)
in The University of Michigan
2009

Doctoral Committee:

Professor John P. Hayes, Chair
Professor Kang G. Shin
Associate Professor Jason N. Flinn
Associate Professor Mingyan Liu

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. John P. Hayes for his guidance in the preparation of this dissertation, and for his confidence and support throughout the course of my doctorate research. I would also like to thank all the members of my committee, Dr. Kang G. Shin, Dr. Jason N. Flinn, and Dr. Mingyan Liu for their time and constructive criticism.

I am indebted to my former advisor, Dr. Chong Nam Chu at Seoul National University. I also thank my pleasant fellow students, Ken Zick, Dae Young Lee, and Chien-Chih Yu. I further wish to express my thanks to former colleagues, Feng Gao, Smita Krishnaswamy, and Jia-Yi Chen.

Finally, I would like to thank my wife Jaeyoon for her support and patience.

PREFACE

Mobile ad hoc networks (MANETs) are wireless communication networks which are of interest because of their flexibility and ease of deployment. MANET nodes are often powered by batteries, and their replacement is usually difficult. Inter-node transmission power thus constrains the network topology and affects the communication efficiency. Furthermore, the network connectivity changes continuously due to mobility. Hence, understanding node mobility and adaptive management of transmission power are essential for successful network operation.

This work addresses the implementation of efficient power-aware MANET management schemes. First, we analyze mathematical models of node movement and propose a metric that quantifies mobility. Existing network control algorithms are usually evaluated using random mobility models. However, since such models employ incompatible mobility parameters, it is hard to compare the performance of different algorithms. It has been shown that the impact of mobility on the network performance is a function of route lifetime. We show that link duration has a nearly invariant relationship with route lifetime regardless of the adopted mobility model, and thus is a good mobility metric.

Second, we investigate the issues of power control and link maintenance. Existing power control schemes are mainly intended for static or pseudo-static networks, and their effectiveness in highly mobile networks has not been demonstrated. We develop a novel algorithm, which adaptively controls transmission power, and reduces communication power

needs by more than 50% compared to existing algorithms with homogeneous transmission range. We also analyze the impact of medium access control on network performance. We show that the widely used RTS/CTS handshake protocol may adversely affect the network throughput when communication power is adjusted to the minimum necessary level. We further present a means to maximize the network throughput.

Third, we investigate the problem of optimally placing base station and relay nodes. Appropriate insertion of such nodes can reduce power consumption and improve network performance. We apply non-linear optimization techniques to node placement, and present distributed node placement techniques which place nodes among radio obstacles to minimize the energy consumption. Simulation results confirm that the efficiency of the proposed algorithms is comparable to that of an existing centralized algorithm.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

SYMBOL TABLE

Symbols	Descriptions
$\|\| \cdot \|\|$	Euclidian norm
α	Radio attenuation exponent
b	Link bandwidth (bps)
$Cost(d)$	Power cost for communication across distance d
CS_{th}	Carrier-sense threshold; the minimum *RSS* that overhearing nodes detect as carrier
d_{CS}	Carrier-sense distance
$d(i, j)$	Topological distance between nodes i and j as the sum of edge weights
$Dist(i, j)$	Euclidian distance between nodes i and j
$E_{acc,i}$	Accumulated energy; the total energy consumed from the source to node i
$E_p(x)$	Artificial potential energy stored in a spring of length x
$\vec{F}_i$	Net artificial force exerted on node i
$Gain(i, j)$	Radio gain from node i to j
λ_{LCR}	Link change rate
$N_{success}$	The number of successful transmissions

Symbols	Descriptions
N_{TX}	The number of transmitting nodes
$Pow(i)$	Transmission power of node i
$RSS(i)$	Received signal strength at node i
$RSS_{\min}$	Receiver's sensitivity; minimum RSS for successful reception
r_{CS}	Ratio between carrier-sense distance and transmission range; $r_{CS} = d_{CS}/r_{TX}$
r_{TX}	Transmission range; often written as r
ρ	Node density
SIR	Signal-to-interference ratio
T_e	Maximum end-to-end throughput
T_{LD}	Link duration
$\tau(i)$	Transmission time of node i; the number of generated bits divided by the bandwidth
$T(i)$	Total transmission time of node i including transmission time for data forwarding
v	Node speed
$w(i, j)$	Traffic rate across link (i, j)
$\vec{x}_i$, $\vec{p}_i$	Location vector of node i
z_0	Minimum SIR for successful reception

CHAPTER 1

Introduction

A mobile ad hoc network (MANET) is a multi-hop, wireless network consisting of a set of interacting hosts or nodes that can move through space. Potential applications of such networks include a group of pedestrians or personal computing devices that can communicate with each other, a group of sensors for observing the environment, and vehicle tracking systems with wireless communication. When a node has to send a message to another node, the sender can either directly transmit the message to the recipient, or transmit the message to its immediate neighbors. Intermediate nodes located between the source and the destination can sometimes serve as routers and relay the message to the final destination.

The operation of MANETs differs from that of traditional wired communication networks in many respects. The network connection structure or topology of a MANET can constantly change due to movement of its nodes. The links between node pairs can be created or deleted by adjusting the transmission power of the nodes. The connection topology is often asymmetric because the transmission powers of a node pair can be different from each other. The communication medium, e.g. air, is usually shared by multiple hosts, and the transmitted data can be garbled if the channel access is not controlled appropriately. For

these reasons, efficient operation of a MANET requires different approaches from those for wired networks, and poses some unique problems. Among them, we will focus on the following major issues: node mobility and its impact on network topology, transmission power control, link maintenance, and medium access control. In this chapter, some of the above problems will be introduced, and relevant prior work will be reviewed.

1.1 Mobile Computer Networks

During the last few decades, computer network technologies have advanced so that most companies, universities, and institutes are now using such networks to store, exchange, and process data within the organization. Moreover, the Internet, the "network of networks", is being used so widely that most people in developed countries can access the Internet for message and file exchange, home banking, or online shopping using e-mail and world wide web protocols [87].

Mobile phones are being pervasively used in all parts of the world. According to the data published by the International Telecommunication Union (ITU) [44], the number of mobile phone subscribers has increased dramatically, and now is higher than the number of fixed phone lines; see Figure 1.1. As of December 2008, the number of mobile phone subscribers has reached 4 billion [17], exceeding a half of the world population. Mobile telephony is rapidly growing in both developing and undeveloped countries. Between 1998 and 2003, according to a Vodafone study, the number of mobile phone subscribers per 100 in South Africa grew from 7.92 to 36.36, and during the same period, the number in Rwanda increased from 0.12 to 2.52 [93]. The same study also reported that 97% of people in Tanzania could access mobile phones, while only 28% could access a landline somewhere in the community. In India, the number of subscribers increased from 1 million

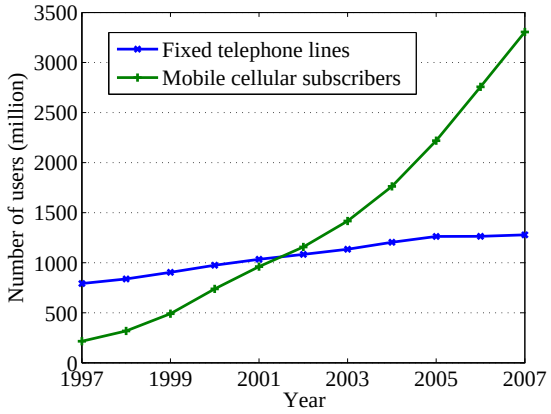


Figure 1.1: Growth of the number of fixed and mobile phone users in the world [source: ITU]

to 315 million between 1999 and 2008 [94].

From the widespread use of the Internet and mobile phones, it seems that use of wireless networks will continue to grow. Recently, so-called smart phones [43] that have high computation performance and can provide users with access to the Internet have come into commercial use, which is making pervasive and ubiquitous computing possible [72, 105]. Small hand-held devices that enable users to store, read, and transfer hundreds of electronic books have also come to market, and are being sold at affordable prices [84]. However, the users of these devices may need to pay a rather high monthly fee to access the Internet. In order to obtain the Internet access without paying the service fee, the users need to connect to nearby open access points, i.e., gateways linked to the Internet.

Recent research has investigated ways to access the Internet using open access points, and demonstrated the operability of those techniques by simulation. Balasubramanian *et al.* [4] proposed opportunistic connection techniques which enable web search in vehicles in urban areas where only intermittent connectivity is available. Banerjee *et al.* [5] proposed an energy-efficient data forwarding architecture based on the use of fixed and battery-powered data centers which store and forward packets according to the connection

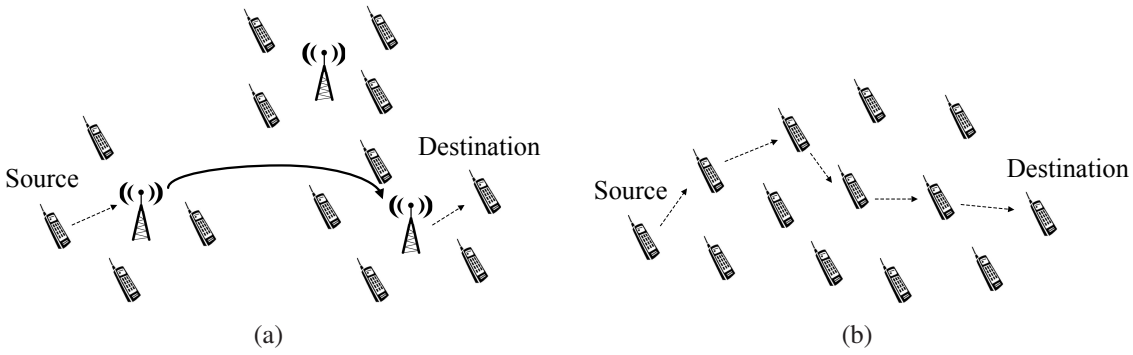


Figure 1.2: Examples of multi-hop networks: (a) cell-based, and (b) ad hoc

availability. However, the most serious limitation with these techniques is that connectivity is available only when open access points exist nearby. For this reason, efficient ways to provide seamless network connectivity are still needed.

Some networks including MANETs employ multi-hop communication over longer distances, where data from a source host is forwarded to the final destination via intermediate hosts. Figure 1.2 illustrates two types of multi-hop network architectures. In the cell-based network of Figure 1.2(a), the source host sends the data to the nearest base station, which is non-mobile and connected to other stations through broadband wires. The base station near the destination receives the data, and sends it to the destination host. However, as shown in Figure 1.2(b), MANETs do not normally require base stations, and the data from the source is forwarded to the destination through other hosts between the source and the destination. The data delivery can be done without network infrastructure if there are sufficiently many wireless hosts. Thus, even when no nearby access point is present, the MANET technology can provide seamless connectivity between hosts, and can enable the access to the Internet when only a few open access points are sparsely located. Furthermore, as MANET hosts communicate mainly with nearby neighbors, the communication distance can become very short, and thus the devices will consume much less power than those with other wireless communication technologies.

In addition to the node-to-node communication, MANETs have various other applications. For example, MANETs can be used for emergency communications. In a disaster area, we cannot expect to have a working network infrastructure, but may still need to provide communications between rescue agents. We can construct a MANET consisting of the wireless devices carried by the agents, and obtain a power-efficient communication network in the absence of a fixed network infrastructure.

To monitor environmental conditions over a large area, a *sensor network* [18, 27] can be used. By spreading a massive number of cheap sensor devices with wireless communication capability over the area to be monitored, we can construct a power-efficient ad hoc network consisting of the sensors, and collect data from the environment continuously. For example, on Great Duck Island in Maine, 32 wireless sensor nodes are deployed [59] in a manner that is accessible and manageable via the Internet. With appropriately assigned operation loads, this habitat monitoring network is expected to last 9 months with two AA batteries. In [90], a mobile sensor deployment project is presented, which delivers wireless sensors to a road using GPS-controlled unmanned aerial vehicles (UAVs). The sensor nodes, controlled by TinyOS [91], constitute a multi-hop communication network, track nearby vehicles passing on the road, and report the tracking data to a base station via a UAV.

The US Department of Transportation is supporting research on intelligent transportation systems via a project called WAVE, wireless access for vehicular environments [97]. In WAVE, vehicles communicate with each other and with roadside access points, and drivers are warned when they are entering unsafe regions or road sections where the traffic is heavy. The WAVE system can also be used to prevent collision between vehicles, and to control traffic signals adaptively. European vehicle manufacturers are aiming to reduce road fatalities, and achieve higher traffic efficiency via an inter-vehicle communication net-

work called C2X Communication [16]. In this project, vehicles are equipped with wireless transceivers, exchange traffic condition messages with each other, and communicate with traffic signals. The C2X Communication system provides drivers with up-to-date road traffic conditions and route recommendations, controls traffic signals in an adaptive way, and positions vehicles accurately for a higher traffic efficiency taking others' positions into account. An airborne network for inter-airplane communication is also a promising MANET application [95]. Instead of relying on instructions from a control tower, airplanes could continuously communicate with each other exchanging information about their identity, condition, location, and speed, thus achieving higher fuel efficiency and safety.

Another potential MANET application is a communication network linking personal electronic devices such as MP3 players, digital cameras, or mobile phones. A *personal area network* (PAN) [34] refers to a network that connects the electronic devices with each other or to the Internet. Its objective is to provide seamless connectivity between devices, typically within a few meters range. Several wireless communication standards are in use or under development for this purpose [34, 42, 60]. The various MANET applications mentioned above are summarized in Table 1.1.

In order to enable efficient inter-host communication within a MANET, we still need to solve the following problems. As most mobile devices are battery-powered, power-

Application	Objective
Emergency networks	To provide connectivity between distant devices where the network infrastructure is unavailable
Sensor networks	To monitor environmental conditions over a large area
Vehicular ad hoc networks	To enable real-time vehicle monitoring and adaptive traffic control
Personal area networks	To provide flexible connectivity between personal electronic devices or home appliances

Table 1.1: Examples of MANET applications

efficient communication is a key problem. The users are mobile, and it is usually infeasible to track all user locations. Thus, the communication links should be adaptively managed in a distributed manner. In addition, it is desirable that communication links between devices can be readily established without the support of network infrastructure. In this dissertation, we investigate several of these problems, and propose novel techniques for power-efficient communication and adaptive link maintenance.

1.2 Terminology and Network Models

This section discusses the formal network models adopted in this dissertation, and introduces the terminology used.

We refer to communication hosts in a network as *nodes*. Two nodes are said to be *connected* if and only if they can directly communicate with each other. The nodes that are connected to node A are called A's *neighbors*. The data forwarding process or link from a node to a neighboring node is called a *hop*. A data delivery path consisting of one or more hops is called a *route*. If we define a graph $G = (V, E)$ such that

$$E = \Big\{(u, v) \mid u,\ v \in V,\ u \neq v,\ u \text{ and } v \text{ are directly connected}\Big\}$$

then G is called the connectivity graph of the network, or the network *topology*. The network topology depends on the node arrangement and the transmission power level assigned at nodes; it changes over time when nodes are mobile.

We now summarize the widely used 5-layer network model [52]. The software and hardware parts for networks are often designed in layers; see Figure 1.3. Data generated in the top application layer at host A are transferred to lower layers, and the bottom physical layer transmits the data to the corresponding physical layer at host B. Then, the data are

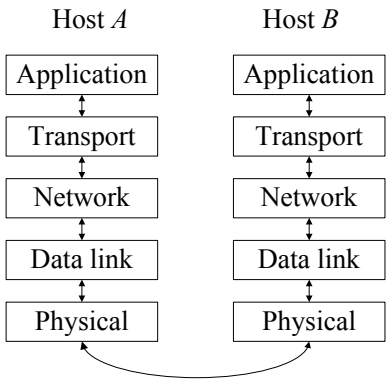


Figure 1.3: Five-layer network model

transferred up to *B*'s top application layer.

The application layer generates and receives communication data according to the network application such as FTP, HTTP, and SMTP [87]. The transport layer has two representative protocols, the transmission control protocol (TCP) and the user datagram protocol (UDP). TCP attempts to avoid data congestion at the intermediate nodes between the source and the destination by controlling the transmission rate at the source. It also guarantees end-to-end communication reliability by exchanging acknowledge (ACK) signals. On the other hand, for some types of data such as real-time audio/video stream, partial data loss is acceptable, and reliable data delivery through retransmission is unnecessary. For such data, UDP is often used, which performs none of the above processes performed by TCP, but provides a finer application-level control. The network layer takes responsibility for constructing multi-hop data paths from sources to destinations, and enables data forwarding from a node to its next hop neighbor. The data link layer enables reliable communication between adjacent nodes. The medium access control (MAC) sublayer located in the link layer coordinates transmissions from different nodes, and reduces signal collision due to interference. The physical layer at the transmitting node converts the data bits to electrical or electromagnetic waves. At the receiving node, the reverse conversion occurs, and the

data bits are sent to upper layers.

In some networks including MANETs, a data stream is first partitioned into blocks, which are transferred across layers [19, 48, 73]. Such blocks are often referred to as messages, segments, or frames depending on the layer under consideration. In this dissertation, we will call those blocks *packets* unless further clarification is necessary. The major properties and research challenges of MANETs will be presented in the following sections, focusing on the network, data link, and physical layers.

1.2.1 Network Layer

The network layer of a MANET is responsible for maintaining appropriate routes from sources to destinations. The network topology can change over time due to node movement, and usually there is no entity that has global knowledge of the network's exact connection status. Hence, the delivery route from source to destination can frequently change, and the nodes need the ability to discover the routes using some appropriate routing protocol.

Routing protocols for MANETs can be classified into two categories: *reactive* and *proactive* [72]. Reactive protocols attempt route discovery only when a node needs to send a message to another. Dynamic source routing (DSR) [47] is a representative reactive protocol. Route discovery in DSR is done as follows; see Figure 1.4. If a source node needs

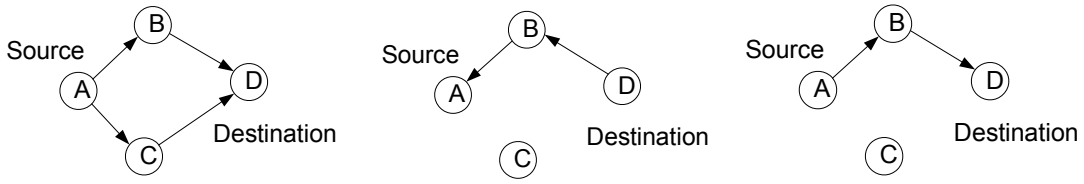


Figure 1.4: Example of route discovery and data delivery in DSR: (a) route request; (b) route reply; and (c) data delivery

to send a message to another node (destination), and if no route from the source to the destination is known, the source first broadcasts a route request message to its neighbors, and the neighbors again broadcast the route request to their neighbors, as in Figure 1.4(a). Such recursive message propagation is called *flooding*. The flooding terminates when the message reaches the whole network. Once the destination node receives the route request, it responds by sending a route reply message to the source through intermediate nodes, as in Figure 1.4(b). The intermediate nodes record the route in their route caches, and reuse the route information when they need to convey a new message to the destination. Once a route is discovered, the source node and intermediate router nodes convey the message through the route, as in Figure 1.4(c).

Route discovery is performed by reactive protocols only when necessary, i.e., when there exists a message to be delivered and no appropriate route to the destination has been discovered recently. If the network needs to deliver a new message, and a route has been discovered and stored in the route cache, the nodes along the route try to reuse the stored route information. However, due to node movement, the network topology may have changed since the last route discovery, in which case a new route may have to be found. Usually route discovery is slow and energy-consuming. The total energy consumption E_{RREQ} of a route request with flooding amounts to

$$E_{RREQ} = N(E_{TX} + k \cdot E_{RX})$$

where N, E_{TX}, E_{RX}, and k denote the number of nodes, the energy consumption due to transmission and reception, and the average number of neighbors, respectively. As can be seen, the total energy consumption can grow very large when N and k are large. Furthermore, during the route discovery operation, the data delivery from the source to the

destination is delayed. Hence, it is preferable that the network uses long-lasting routes rather than frequently changing ones.

Unlike reactive protocols which perform on-demand discovery only when necessary, proactive routing protocols attempt to maintain a valid routing table even when there is no message to be delivered. To do so, in networks with proactive protocols, the nodes periodically broadcast special messages called *beacons* that contain the information about route updates. Once topology change occurs, the nodes that sense the change in connectivity propagate the change information to their neighbors so that all nodes can maintain valid routing tables. In this way, proactive protocols reduce delivery time by eliminating the initial route discovery time for the first data delivery. A drawback of proactive protocols, however, is that they may consume more power than reactive protocols since the nodes must exchange route information to keep the routes up-to-date.

The destination node D in Figure 1.4 may choose node C as its intermediate router (route A-C-D) instead of node B (route A-B-D). Hence, whenever multiple routes are available, it is preferable to choose the route that best meets some objective. Depending on the network applications, the objective could be maximization of data throughput, or minimization of delivery delay or power consumption. Since use of out-of-date routing information can lead to unexpected results such as waste of communication power or data loss, it is usually desirable to choose stable and long-lasting links rather than short-lived ones.

There have been various efforts to improve network performance by allowing the nodes to choose links with a longer lifetime [1, 12, 46, 67, 85, 88]. All the proposed schemes analyze the movement of nodes or the received signal strength, predict the remaining lifetime of links, and enable nodes to select links with a long lifetime. However, serious limitations of the above schemes are that they assume fixed transmission power, and implicitly

require periodic beaconing by nodes. For these reasons, the nodes may consume unnecessary energy, and the severe signal interference between different transmitters can lower the network bandwidth. In Chapter 3, we propose a new method that improves route stability by adaptively adjusting transmission power thereby requiring signal transmission only by nodes that are around the data delivery routes; this effectively reduces the energy consumption.

1.2.2 Data Link Layer

A medium access control (MAC) protocol at the data link layer is responsible for coordinating the channel accesses by multiple hosts. MAC protocols can be classified as fixed-assignment, random-access, and demand-assignment [2, 40]. Demand-assignment protocols can be understood as a compromise between fixed-assignment and random-access. We briefly review three representative MAC protocols: frequency division multiple access (FDMA), time division multiple access (TDMA), and carrier-sense multiple access (CSMA). The first two techniques belong to the class of fixed-assignment protocols, and the third one is a random-access protocol.

In FDMA, the transmitters use independent frequency channels. When a node receives messages, it decodes the incoming signal with respect to frequency. By dividing the communication channel into a spectrum of frequencies, multiple messages do not interfere with each other when the communication medium is shared.

Instead of dividing channels using multiple frequencies, TDMA divides the communication channel into multiple time slots. Each transmitter is assigned its own time slot for transmission, and is allowed to transmit signals only during its time slot. Allocating separate time slots to transmitters ensures that there exists at most one active transmitter at a time, and signal collision is prevented.

The transmitters in random-access protocols contend for channel access; one transmitter is chosen at random and is given the chance to transmit. CSMA [50] is one such random-access protocol. A CSMA transmitter first senses the channel, and if it detects the presence of signals, i.e., if the channel is busy, it does not transmit. Once it determines that the channel is idle, it waits for a random period (the backoff time) before transmission. By forbidding immediate transmission, CSMA reduces signal collision due to simultaneous transmissions by multiple hosts, and improves channel utilization. Further, unlike fixed-assignment protocols, CSMA dynamically schedules transmissions. This results in more efficient communication, especially when the network traffic is sparse.

There are some radio communication techniques that do not quite fit into any of the above MAC schemes. Frequency-hopping spread spectrum (FHSS) is one such technique. Instead of using a fixed frequency channel, an FHSS transmitter keeps switching between different frequency bands according to pseudo-random sequences. The first advantage of FHSS is that as it does not fully rely on any particular channel. The data loss due to interference between different transmitters or other radio-emitting devices is only partial; thus, any lost data can be easily recovered by short and infrequent retransmissions. The second advantage is that FHSS provides more secure communication since its channel is changing randomly, making it hard to intercept a complete message. FHSS is used for Bluetooth communications and military purposes.

1.2.3 Physical Layer

The physical layer of a MANET is responsible for conversion between data bits and radio signals. In free space, the received intensity of a radio signal transmitted from an omnidirectional antenna is inversely proportional to the square of the distance between the transmitter and the receiver. However, when obstacles are present, the radio signal can

be absorbed, reflected, or refracted, and the radio profile with respect to distance becomes complicated. In general, the received signal strength (RSS), or the intensity of the radio signal at the receiver RX, can be expressed as

$$RSS(RX) = Pow(TX) \cdot Gain(TX, RX)$$

where $Pow(TX)$ denotes the transmission power of node TX, and $Gain(TX, RX)$ denotes the radio gain between nodes TX and RX, characterized by the radio environment.

One of the simplest radio propagation models is the log-distance path loss model [29, 77], in which the radio gain $Gain(TX, RX)$ is given by

$$Gain(TX, RX) \propto \Big(\frac{d_0}{Dist(TX, RX)}\Big)^{\alpha}$$

where d_0 denotes the reference distance, and $Dist(TX, RX)$ the distance between TX and RX. The radio attenuation exponent α is chosen according to the radio environment, and usually ranges from 2 to 6. With this model, the radio gain becomes

$$Gain(TX, RX) = \frac{k_0}{Dist(TX, RX)^{\alpha}} \tag{1.1}$$

The following radio model derived from (1.1) is also used [57].

$$Gain(TX, RX) = \frac{1}{k_1 \cdot Dist(TX, RX)^{\alpha} + k_2} \tag{1.2}$$

For successful radio reception, the RSS of an incoming signal should be higher than a minimum threshold, which is characterized by the receiver's sensitivity $RSS_{\min}$. If we assume that the radio gain function $Gain(TX, RX)$ depends on distance only, and is mono-

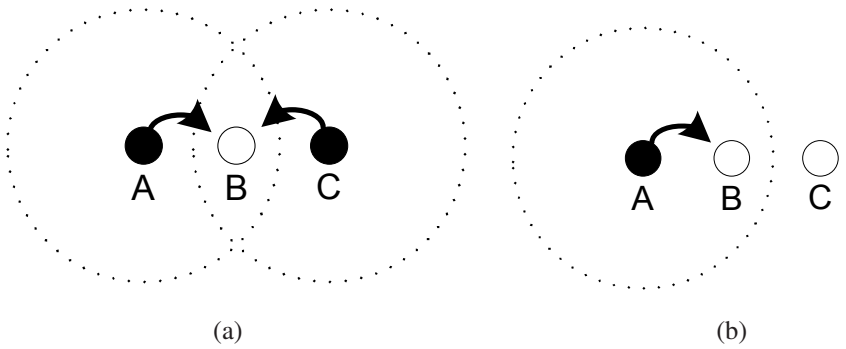


Figure 1.5: Examples of (a) communication failure and (b) successful communication. Black nodes are transmitting and white nodes are not.

tonically decreasing, the transmission range r_{TX} can be defined as the maximum distance at which the RSS exceeds the receiver's sensitivity, or equivalently, $r_{TX} = Dist(TX, RX)$, where $Dist(TX, RX)$ is the solution of

$$\begin{aligned} RSS_{\min} &= Pow(TX) \cdot Gain(TX, RX) \\ &= \frac{Pow(TX)}{k_1 \cdot Dist(TX, RX)^{\alpha} + k_2} \end{aligned}$$

According to this definition of transmission range, the receivers within the transmission range r_{TX} of the transmitter can successfully receive the message.

Hosts in wireless networks share a communication medium, e.g., air, and data can be garbled and lost if multiple hosts are transmitting simultaneously. If MANET nodes are equipped with multiple directional antennas [64], or the MIMO technology is adopted [25], nodes can receive data from multiple transmitters. In this dissertation, we assume that a MANET node has only one omnidirectional antenna, and communicates over a single wireless channel, unless stated otherwise. Hence, a node can receive data from at most one transmitter at a time. Two different criteria for successful communication are extensively used. The first one states that for a successful communication, there should be exactly one

transmitting node within the transmission range of the receiver. Figure 1.5 shows examples of communication success and failure. Black circles represent transmitting nodes, and white circles represent nodes that are not transmitting. Dotted circles show transmission regions within which the receivers are affected by the signals sent from the transmitters. Suppose node A is attempting to send a message to node B. In Figure 1.5(a), node C is also transmitting while A is sending a message to B. The signals at B from A and C collide, and node B cannot receive any message from A or C. On the other hand, as suggested by Figure 1.5(b), node B can successfully receive the message from A since node C stays silent during the communication.

The second criterion for a successful communication is that the signal-to-interference-noise ratio (SINR) should be higher than a minimum threshold, i.e.,

$$SINR(TX, RX) = \frac{Pow(TX) \cdot Gain(TX, RX)}{\sum_{i \neq TX} Pow(i) \cdot Gain(i, RX) + n_{RX}} \geq SINR_{\min}$$

where n_{RX} denotes the noise level at the receiver RX. When the noise level is omitted, SINR reduces to a signal-to-interference ratio (SIR), and the criterion with SIR is defined in a similar way as follows.

$$SIR(TX, RX) = \frac{Pow(TX) \cdot Gain(TX, RX)}{\sum_{i \neq TX} Pow(i) \cdot Gain(i, RX)} \geq z_0 \tag{1.3}$$

During network operation, MANET nodes consume stored energy for communication and computation [57, 74], and the communication power cost can be broken down to transmission and reception power. Let Pow_{TX}, Pow_{RX}, and Pow_{comp} denote transmission, reception, and computation power, respectively. Assuming the radio model (1.2), when a node adjusts its transmission power to yield the received signal strength $RSS_{\min}$ at distance d,

the transmission power Pow_{TX} becomes

$$Pow_{TX} = RSS_{\min} \cdot (k_1 \cdot d^{\alpha} + k_2)$$

If the number of nodes that receive or overhear a transmitted signal is N_o, then the total power consumption Pow_{total} due to a single transmission is given by

$$\begin{aligned} Pow_{total} &= Pow_{TX} + Pow_{RX} + Pow_{comp} \\ &= RSS_{\min} \cdot (k_1 \cdot d^{\alpha} + k_2) + N_o \cdot Pow_{RX} + Pow_{comp} \end{aligned}$$

If the number N_o of receiving or overhearing neighbors is small, the total power cost Pow_{total} can be replaced with the transmission power Pow_{TX} by assuming appropriate k_1 and k_2 values. Thus, when our goal is to minimize the total power consumption of a MANET rather than to evaluate it accurately, we can use Pow_{TX} as the optimization objective instead of Pow_{total}.

1.3 Mobility and Topology Control

In MANETs, the connection topology can change over time due to movement of nodes, and the network topology can critically affect the network operation. Understanding the impact of mobility and network topology on the network performance is important for designing the network control algorithms. This section examines the impact of node mobility, and reviews some representative artificial mobility models and topology control algorithms.

1.3.1 Node Mobility

A MANET's network topology is a function of the location and movement of its nodes. Figure 1.6 shows an example of a change in network topology due to node movement. When node *A* needs to send a message to node *D*, node *B* can serve as a router and deliver the message to the final destination *D*. However, if node locations change over time due to mobility, the route *A*-*B*-*D* may become invalid because some of its constituent links are broken. If such a topology change occurs, the network needs to deliver the message through a new valid route such as *A*-*C*-*D*, as in Figure 1.6(b). As was pointed out in Section 1.2.1, in networks with reactive routing protocols, route discovery is usually a power-consuming process, and causes delay in message delivery. On the other hand, in networks with proactive routing protocols, once topology change occurs, the nodes should propagate the information about connection change in order to maintain valid routing tables, a process that wastes communication power and channel resources.

Analysis of the impact of mobility on network performance usually relies on simulations using artificial random mobility models [7, 8, 102]. Mobility models can be classified into *entity* and *group* mobility models [15]. Entity mobility models describe random, independent movements of nodes, while group mobility models specify correlated movements of a group of nodes. Our study will mainly focus on the entity type, and "mobility model"

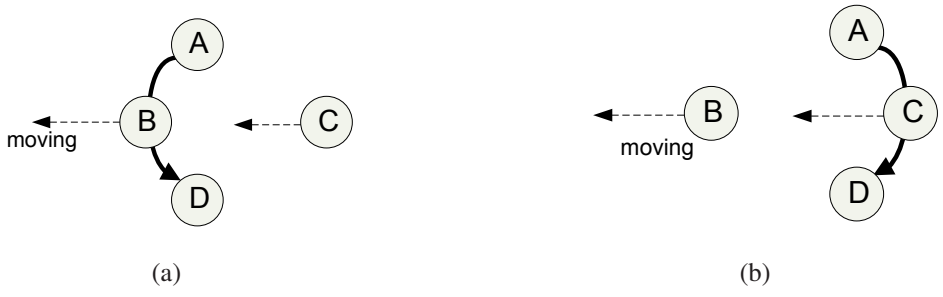


Figure 1.6: Example of topology change due to node movement: (a) route from *A* to *D* through *B*; and (b) route from *A* to *D* through *C*

will refer to entity mobility models. In Section 2.2, we will propose a new mobility model which, unlike most previous ones, allows mathematical analysis, and derive mobility metrics for it, such as link duration and link change rate.

1.3.2 Topology Control

Unlike wired networks, the connection topology in MANETs can be actively controlled by adjusting the transmission power of the nodes. For this reason, transmission power management in MANETs is often called *topology control*. The objectives of topology control include reduction of power consumption and improvement of data throughput. Long-distance communication causes increase in radio power consumption and interference between transceivers, and thus is usually considered undesirable. Conversely, when transmission range is too short, the network connectivity can be lost. Therefore, the transmission range of the nodes should be appropriately assigned so that the power consumption and signal interference are minimized while the network connectivity is maintained. There have been a number of studies as to how transmission range affects the network performance [14, 26, 32, 33] and connectivity [9, 68, 73, 78, 79, 89, 100].

Most topology control algorithms attempt to bound the nodes' maximum degree, i.e., the number of neighbors, and reduce signal interference to achieve the objectives of power conservation and throughput improvement. Topology control algorithms can be grouped according to their methodology [75]. We identify a few representative topology control algorithms (see Table 1.2) and review their approaches.

"Local Information No Topology" (LINT) [76] allows each node to maintain a bounded number of neighbors, e.g., 5 to 7, by incrementally adjusting its transmission power. LINT requires no additional capability from transceivers except for transmission power adjustment itself. However, a drawback of LINT is that though it forms a network that is con-

Topology control algorithms	Methodology
LINT and LILT [76] K-neigh protocol [11]	Bounded number of neighbors
Cone-based topology control [54, 96]	At least one neighbor in every cone
MST-based topology control [55, 56]	Local minimum spanning tree

Table 1.2: Representative topology control algorithms and their methodology

nected with a high probability, it does not guarantee global connectivity. In order to fix this problem, the Local Information Link-State Topology (LILT) method uses global connectivity information from routing tables, and maintains a connected network.

The basic operation of the "K-neigh" protocol [11] is similar to that of LINT, in that it attempts to maintain a bounded number of neighbors. K-neigh assumes that a node can estimate its distance from neighbors by exchanging beacons. Instead of incremental adjustment as in LINT, each node picks an appropriate transmission power so that it can have the prescribed number of neighbors. Using this distance estimation mechanism, the operation of the K-neigh protocol can complete after each node transmits its beacon message just twice, whereas the operation time of LINT varies according to the network condition.

The "Cone-Based Topology Control" algorithm (CBTC) [96] requires that a node can sense the direction of incoming signals. By increasing the transmission power until every cone with angle α has at least one neighbor, CBTC maintains connectivity of the network. It has been shown that when $\alpha < \frac{5}{6}\pi$, CBTC guarantees a connected network without partitioning.

Topology control algorithms based on formation of a minimum spanning tree (MST) [55, 56] construct a globally connected network by merging local minimum spanning trees. In MST-based topology control, the nodes first exchange beacons or control signals at their full transmission power. Then, each node forms a local minimum spanning tree by pruning the heaviest edges, i.e., the edges between neighboring nodes that are located far from

each other, and the network achieves global connectivity. An advantage of the MST-based approach is that it makes no assumption about radio propagation, and can be easily applied to cases where the signal obstruction by physical obstacles is not negligible.

It has been shown that the CBTC and MST-based topology control algorithms result in a network where the maximum node degree is always less than or equal to 6, and so keep the network topology reasonably sparse. By maintaining a sparse network, these topology controls try to conserve communication power and avoid excessive signal interference between nodes.

All the above topology control algorithms attempt to fix the network connection after the topology change occurs, and cannot reduce the number of connection changes caused by node movements. To remove this limitation, we will propose a *proactive* topology control algorithm that adjusts the transmission powers of communicating nodes, prevents frequent link breaks, and improves the communication performance.

1.4 Related Technologies

In this section, a few more wireless communication technologies are discussed briefly. With the growth in the number of wireless communication devices such as mobile phones, personal digital assistants, and vehicle telematics, the demand for radio channels is increasing fast. On the other hand, lots of different communication technologies and protocols are emerging, which may not be compatible with each other. Hence, it becomes attractive for radio hosts to flexibly determine the available radio channels and the corresponding communication technologies in an intelligent way. Cognitive radio [61] is one such effort. Agreeing on so-called formal radio etiquettes, and learning the radio environment, cognitive radio devices flexibly adapt themselves to different channels and protocols, and select

the ones that are most suitable for the current conditions. Advances in software-defined radio (SDR) technology make such flexibility feasible. Though it still requires more research for practical implementation, cognitive radio is expected to lead to the convergence of different radio technologies for more efficient wireless channel utilization.

1.5 Dissertation Outline

This dissertation is organized as follows. Chapter 2 addresses the issue of quantifying node mobility. We propose a simplified mobility model that allows strict derivation of link duration (LD) and link change rate (LCR), and investigate the relation between node mobility and the lifetime of routes. Then, we quantify connection stability by relating it to LD, and show that LD is a good mobility metric, while LCR is unsuitable for the purpose of quantifying node mobility.

In Chapter 3, we introduce a novel power-aware link maintenance (PALM) algorithm, in which mobile nodes adaptively control their transmission power, and discover power-efficient data delivery routes in a distributed manner. We propose the concept of virtual hop-distance, which enables discovery of power-efficient and loop-free routes, and allows overhearing nodes to determine whether modifying an existing route can reduce the total energy consumption. We also analyze the effectiveness of the RTS/CTS handshake from the viewpoint of network throughput and energy efficiency. By simulation experiments, we show that the use of RTS/CTS may adversely affect the throughput of ad hoc networks in which nodes adaptively control their transmission power to the minimum necessary level. Then, we propose a means to tune the hardware parameters of wireless transceivers to maximize the network throughput.

Chapter 4 investigates node placement optimization problems assuming several differ-

ent network structures. We first assume a clustered network structure, and show that the communication power minimization problem is similar to the K-means problem in vector quantization for multimedia data compression [80], which is an NP-hard problem. We prove that the cost function of the clustered network structure is convex. Taking advantage of the convexity of the cost function, we develop an iterative optimization technique. We also present a distributed node placement algorithm that places relay nodes, and constructs a power-efficient network structure in the presence of radio obstruction. We model the wireless network as a mechanical system, and implement a node placement algorithm utilizing the PALM algorithm proposed in Chapter 3.

Finally, Chapter 5 summarizes the major contributions of the research presented in this dissertation, and discusses some of future directions for MANET research.

CHAPTER 2

Impact of Mobility on Network Performance

The performance of mobile ad hoc networks is highly sensitive to changes in node-to-node connections (communication links) caused by node movement. Link instability of this kind has proven very difficult to analyze mathematically, so previous work has relied heavily on simulation. This chapter presents a mathematically tractable model of node motion, the constant velocity model, and uses it to derive a precise relation between mobility and connection stability. Our analysis also allows determination of the appropriate frame length for successful and efficient single-hop communication. We further investigate connection stability in multi-hop communication, and uncover some underlying properties of previously proposed mobility metrics. In particular, we demonstrate that link duration has a nearly invariant relationship with the stability of multi-hop connections for a wide range of mobility models, and thus is an excellent mobility metric.

2.1 Introduction

Maintaining node-to-node connectivity under complex, often random, mobility conditions is a central problem in MANETs. Small changes in the mobility parameters can

drastically affect the network performance. For example, in single-hop communication, if a data packet is too long, mobile receiver nodes may leave the communication region of a transmitter node during a single packet transmission. Once this premature disconnection occurs, the receiver loses part of the message from the transmitter and the communication fails. Mobility also affects multi-hop communication. When a sender is supposed to choose a message delivery route, as in dynamic source routing [47], the connectivity between the sender and the destination may change after route discovery, and it is possible that the route specified by the sender is no longer valid. Clearly, stable link connections are vital for successful data delivery in MANETs.

To investigate the impact of mobility on network operation, previous work has relied mainly on simulations with mobility models which, as discussed in Section 1.3.1, describe the manner in which nodes move through space; see Table 2.1 [7, 15, 35, 45, 47]. These mobility models typically aim to provide an accurate description of the behavior of a network consisting of pedestrians or moving vehicles. It is unlikely that a single mobility model can be applied to all types of networks, and for this reason, new mobility models are being continuously developed. It has proven to be very difficult to analyze the relation between mobility models and network connectivity due to the many parameters affecting the

Mobility models	Key input parameters
Random waypoint [47]	Maximum speed Minimum speed Pause time
Random walk [45]	Maximum speed Minimum speed Movement time or distance
Boundless simulation area [35]	Maximum speed Linear acceleration Angular acceleration

Table 2.1: Examples of random mobility models and their input parameters

network operation, such as node speed, pause time, node density, and transmission range.

We approach the foregoing problems as follows. First, we set up a relatively simple and mathematically tractable mobility model, the constant velocity (CV) model, and derive two mobility metrics for it, link duration (LD) and link change rate (LCR) [53, 69, 86]. Then, we develop an analytic expression for successful packet delivery using single-hop communication. This analysis provides a guide to the appropriate packet length for efficient single-hop communication. We further investigate the relation between LD and multi-hop route stability, which is the key parameter that determines the success of data delivery. We quantify connection stability by the mean residual duration (RD) of routes, also called residual lifetime [1], which measures how long multi-hop routes last under the given mobility conditions. We show that, among previously proposed mobility metrics [3], LCR is unsuitable for estimating link stability because the relation between LCR and RD depends on other network parameters. In contrast, by using the analytic expressions derived from our CV model, we show that RD is a function of LD; i.e., the multi-hop connection stability is primarily determined by single-hop link duration. We also present simulation results which show that LD has a quite consistent relation with RD for a wide range of mobility conditions. Our analysis and simulations confirm that as suggested in [12], LD constitutes a very good, unified metric for the multi-hop link stability of many types of mobility models.

This chapter is organized as follows. Section 2.2 describes the CV model, and derives the corresponding metrics, LCR and LD. In Section 2.3, we derive a relationship between the mobility and the success ratio of single-hop communication. This provides an answer to the question: What is the appropriate packet length for successful and efficient single-hop communication? Section 2.4 investigates the relation between LD and the lifespan of multi-hop routes using analysis and simulation. We compare the simulation results for various mobility models, and show that these models have the same multi-hop route lifespan

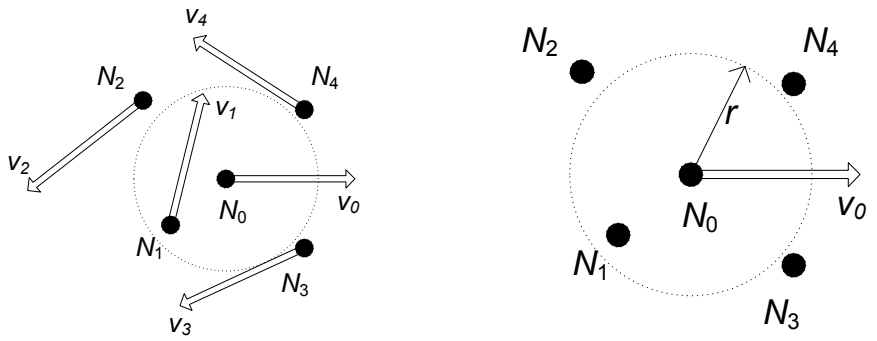
distribution. Section 2.5 discusses the relations between LD, LCR, and RD, and confirms that LD is a good metric for mobility. Section 2.6 summarizes the chapter.

2.2 Constant Velocity Model

We first describe a simplified mobility model, the constant velocity (CV) model, and use it to derive the expected change rate of node-to-node connections, and the expected lifespan of such connections. We assume that nodes are randomly placed on an infinitely large boundless plane with a finite node density ρ. The nodes move linearly at a constant velocity in random directions, but do not change their directions while moving; see Figure 2.1(a).

2.2.1 Link Change Rate

Before dealing with the general case of Figure 2.1(a), we investigate link change rate in a nearly static node arrangement, where only one node is moving, as in Figure 2.1(b). Suppose n nodes are randomly placed with a node density ρ. Only one node N_0 is moving linearly with a constant velocity $\vec{v}_0$; all other nodes are assumed to be stationary. N_0 has a transmission range r, and when the distance from the mobile node N_0 to a static node



(a) All nodes are moving. (b) Only N_0 is moving.

Figure 2.1: Examples of node movements in a MANET

N_i becomes shorter than r, we say a connection link is *generated* between N_0 and N_i. Conversely, as the distance becomes longer than r, we say the link is *broken*. The frequency of such link generations/breaks per unit time is called the link generation/break rate. The *link change rate* (LCR), also called topology change rate [69], is the sum of link generation rate and link break rate. In steady state, the link generation and break rates should be equal, so we first focus on the link generation rate. As the mobile node is travelling on the plane, its circular communication region sweeps the plane with a rate $2rv_0$, and the average number of newly generated links per unit time is $2\rho rv_0$. Hence, the link generation rate is proportional to the mobile node speed.

Returning to the general model of Figure 2.1(a), assume that all nodes are moving at the same constant speed v but in random directions. The link generation rate $2\rho rv_0$ does not hold anymore because the relative speed between nodes is no longer a constant v_0. We derive the relative velocity between nodes as follows. Suppose that the node N_i is passing through the transmission region of N_0, as in Figure 2.2. Let $\vec{v}_i$ denote the velocity of node N_i, where for all i, $||\vec{v}_i|| = v$. The velocity can be expressed as $\vec{v}_i = v(\vec{i}\cos\Theta + \vec{j}\sin\Theta)$ where Θ denotes the movement direction of node N_i, a random variable uniformly distributed

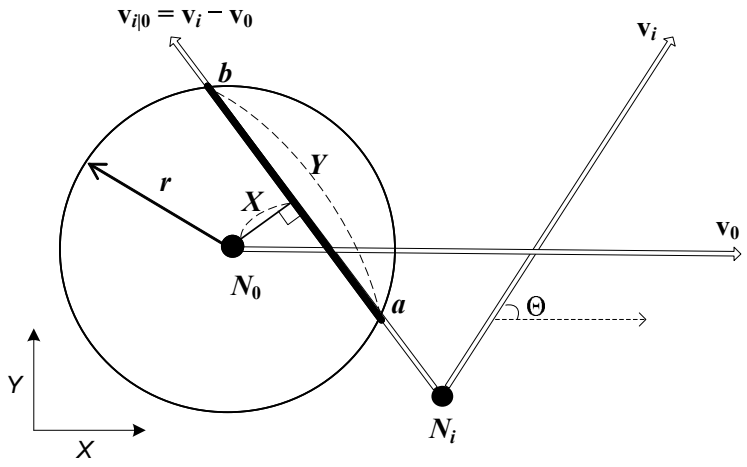


Figure 2.2: Motion of a node N_i passing through the transmission region of another node N_0

between 0 and π. We further assume that N_0 is moving in the positive direction of the X-axis, i.e., $\Theta_0 = 0$. The relative velocity of N_i with respect to N_0 is $\vec{v}_{i|0} = \vec{v}_i - \vec{v}_0$, and its magnitude is $v_{i|0} = v\sqrt{2 - 2\cos\Theta} = 2v\sin\frac{\Theta}{2}$. Thus, the average link generation rate is given by

$$\lambda_{gen} = \int_0^{\pi} 2\rho r v_{i|0}(\theta)\frac{1}{\pi}d\theta = \frac{8}{\pi}\rho r v \tag{2.1}$$

while the average LCR is

$$\lambda_{LCR} = 2\lambda_{gen} = \frac{16}{\pi}\rho r v \tag{2.2}$$

For example, if each node is moving at the speed $v = 10$ m/s with transmission range $r = 10$ m, and the node density is $\rho = 0.02$ m^{-2}, then the link generation/break rates are $\lambda_{gen} = \lambda_{brk} = 5.1$ s^{-1}. Assuming steady state, $\lambda_{LCR} = \lambda_{gen} + \lambda_{brk} = 10.2$ s^{-1}.

2.2.2 Link Duration

Next, we derive the *link duration* (LD) which measures the lifespan of a node-to-node link from the time a receiver enters the communication region of the transmitter to the time the receiver exits the communication region. In other words, LD is the time from link generation to link break, and can be interpreted as a measure of the stability of single-hop connections [3].

Suppose again that nodes are randomly distributed on a plane and moving in random directions. A node N_0 observes N_i passing through the circular transmission region of N_0, as in Figure 2.2, where, as before, $v_{i|0} = 2v\sin\frac{\Theta}{2}$. We can describe the event that N_i passes through the transmission region of N_0 with two parameters (X, Θ). The link duration T_{LD} is given by

$$T_{LD}(X, \Theta) = \frac{Y(X)}{v_{i|0}(\Theta)}$$

where $Y(X) = 2\sqrt{r^2 - X^2}$. The derivation of (2.1) implies that the arrival rate of nodes N_i is proportional to their relative speed with respect to N_0. Hence, the relative frequency of the event (X, Θ) is proportional to $v_{i|0}$, so the joint probability density function $f_{X,\Theta}(x, \theta)$ of (X, Θ) is proportional to $v_{i|0}(\Theta)$, and is given by $f_{X,\Theta}(x, \theta) = v_{i|0}(\theta) / \int_0^{\pi} \int_0^{r} 2v \sin\frac{\theta^*}{2} dx d\theta^* = \frac{1}{2r} \sin\frac{\theta}{2}$. Therefore, the mean LD value $\overline{T}_{LD}$ becomes

$$\overline{T}_{LD} = \int_0^{\pi} \int_0^{r} T_{LD}(x, \theta) f_{X,\Theta}(x, \theta) dx d\theta = \frac{\pi^2}{8}(\frac{r}{v}) \tag{2.3}$$

It should be noted that LD is not the reciprocal of LCR or the link generation rate. LCR is the reciprocal of the time between two successive link changes, whereas LD is defined as the time between link generation and link break. The relation between LD and LCR is discussed in Section 2.5.2.

2.3 Single-hop Communication

In this section, we define and derive the probability of complete transmission, p_{comp}, which measures successful delivery ratio in single-hop communication. Then, we apply the analysis of p_{comp} to determine the appropriate packet length for the given mobility conditions.

2.3.1 Probability of Complete Transmission

Suppose a receiver N_i enters the communication region of a transmitter N_0 at time $t = 0$ and exits the region at time $t = T_{LD}$. For a packet with transmission time T_{comm} to complete communication before node N_i moves out of range, the transmission should start at time t, $0 \leq t \leq T_{LD}(x, \theta) - T_{comm}$. Hence, for given X and Θ, the conditional probability of

complete transmission is

$$p_{comp}(T_{comm}|X,\Theta) = \frac{\max[T_{LD}(X,\Theta) - T_{comm}, 0]}{T_{LD}(X,\Theta)} \tag{2.4}$$

and the total probability of complete transmission is

$$p_{comp}(T_{comm}) = \int_0^{\pi}\int_0^{r} p_{comp}(T_{comm}|x,\theta) g_{X,\Theta}(x,\theta) dx d\theta$$

where $g_{X,\Theta}(x,\theta)$ denotes the joint probability density function of random variables X and Θ. Little's theorem [6] states that the average number of customers in a system is equal to the product of the customers' arrival rate to the system and the average time customers spend in the system. Hence, the joint probability density function $g_{X,\Theta}(x,\theta)$ is proportional to $v_{i|0}(\theta) \cdot T_{LD}(x,\theta) = Y(x)$, and is given by

$$g_{X,\Theta}(x,\theta) = \frac{Y}{\int_0^{\pi}\int_0^{r} Y dx d\theta}$$

For brevity, we define a normalized communication time τ such that

$$\tau \triangleq T_{comm} \cdot \frac{v}{r} = \frac{\pi^2}{8} \cdot \frac{T_{comm}}{\overline{T}_{LD}} \tag{2.5}$$

and calculate

$$p_{comp}(\tau) = \int_0^{\pi}\int_0^{1} p_{comp}(\tau|x,\theta) g_{X,\Theta}(x,\theta) dx d\theta$$

The normalized communication time τ can be interpreted as the ratio of node mobility to communication speed.

First, consider the case $\tau < 1$. We define a function $B(\tau,\theta)$ such that $B(\tau,\theta) = \frac{\sqrt{2-\tau^2+\tau^2\cos\theta}}{\sqrt{2}}$. It can be seen that for all $x < 1$, $T_{LD}(x,\theta) > \tau$ if and only if $0 < x < B(\tau,\theta)$. Using this

fact, we eliminate the max operator in (2.4), and change the range of the inner integral of $p_{comp}(\tau)$ from $(0, 1)$ into $(0, B(\tau, \theta))$. Now we have

$$p_{comp}(\tau) = \int_{\theta=0}^{\pi} \int_{x=0}^{B(\tau,\theta)} p'_{comp}(\tau|x, \theta) g_{X,\Theta}(x, \theta) dx d\theta$$

where $p'_{comp}(\tau|x, \theta)$ is defined as $\frac{T_{LD}(x,\theta)-\tau}{T_{LD}(x,\theta)}$. Replace x and θ with $\alpha = \sin^{-1} x$ and $\beta = \theta/2$, respectively. Then, p_{comp} becomes

$$\begin{aligned} p_{comp}(\tau < 1) &= \frac{8}{\pi^2} \int_{\beta=0}^{\frac{\pi}{2}} \int_{\alpha=0}^{\cos^{-1}(\tau \sin\beta)} (\cos\alpha - \tau \sin\beta) \cos\alpha d\alpha d\beta \\ &= \frac{4}{\pi^2} \int_{0}^{1} \frac{\cos^{-1}(\tau c) - \tau c \sqrt{1 - \tau^2 c^2}}{\sqrt{1 - c^2}} dc \end{aligned} \tag{2.6}$$

In the other case $\tau > 1$, swap the order of integration of x and θ, and eliminate the max operator by changing the integration range of θ in a similar way. This yields

$$p_{comp}(\tau > 1) = \frac{8}{\pi^2} \int_{0}^{1} \frac{c^2 \sin^{-1}(c/\tau) - \tau c(1 - \sqrt{1 - c^2/\tau^2})}{\sqrt{1 - c^2}} dc \tag{2.7}$$

Figure 2.3 compares p_{comp} from (2.6) and (2.7) with simulated results using 100 mobile nodes on a 1×1 plane. The network parameters such as the transmission range r, the node speed v, and the communication time T_{comm} are varied so that τ ranges from 0 to 2.2. Comparison of the simulated and calculated p_{comp} values confirms that the derivation of (2.6) and (2.7) from the CV model is accurate.

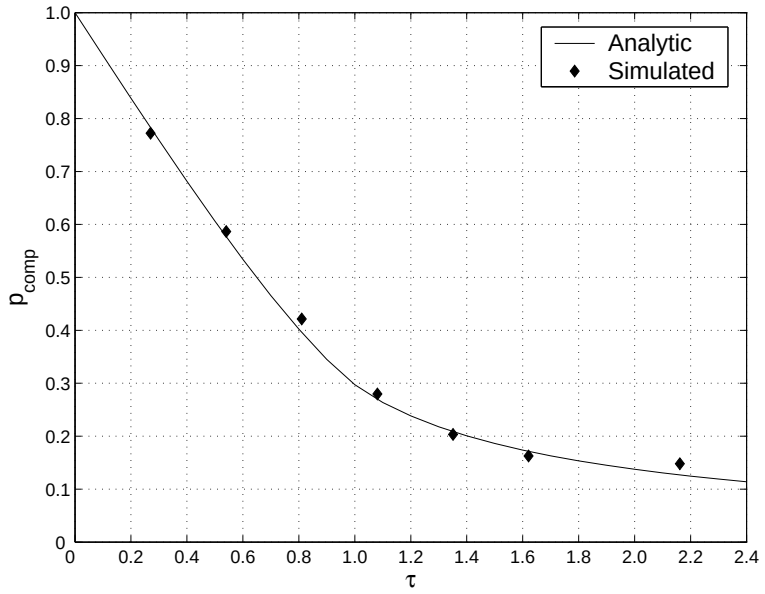


Figure 2.3: Plot of probability of complete transmission p_{comp} vs. normalized communication time τ

2.3.2 Message Frame Length

A message frame usually consists of header and trailer parts of constant length, and a data packet part whose length may vary. For efficient communication, the data packet should be long enough to amortize the cost of the header and the trailer. However, if the frame length grows too long, the corresponding p_{comp} drops, and communication reliability is worsened. Hence, the message frame length should be determined considering both reliability and efficiency. The foregoing analysis of the probability of complete transmission can be used to determine the appropriate frame length.

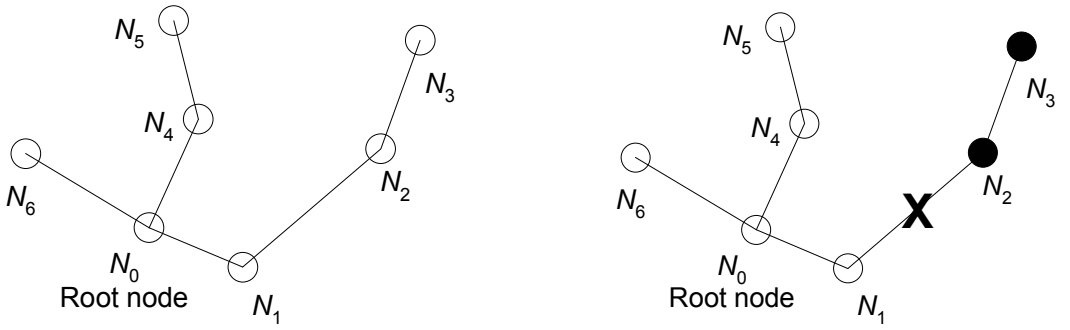
Suppose each node in a MANET has transmission range r = 10 m, and movement speed v = 3 m/s. Single-hop communication often requires additional time T_{add} including backoff and feedback exchange time. Assume that the given network has T_{add} = 100 ms. We want a single-hop communication success ratio of 95% ignoring data loss due to signal interference. Equation (2.6) implies $\tau = 0.061$ for $p_{comp} = 0.95$, and from (2.5), we get

$T_{comm} = \tau \cdot r/v = 0.203$ s. Hence, the frame length including control bits should be no longer than $T_{comm} - T_{add} = 103$ ms for 95% success rate.

2.4 Multi-hop Communication

Although the success ratio of single-hop communication in ad hoc networks is insensitive to mobility when a packet length is not too long, successful data delivery over multi-hop routes critically depends on the connection instability due to node movement. Data delivery over multi-hop routes takes much longer, and hence the probability that the connection changes during the data delivery is not negligible. We define the *mean residual duration* (RD) of a multi-hop route as the mean time from the route discovery to the breaking of the route. Further we define k-RD as the mean residual duration of k-hop routes. A multi-hop route is regarded as broken when any of its constituent single-hop links is broken. RD is the key factor which determines the success of packet delivery over multi-hop routes, and so is an important parameter in MANET design [1]. For instance, when the residual duration of a multi-hop route is 100 ms, data delivery which takes 500 ms over the route will probably fail due to the connectivity change. Hence, RD can be interpreted as an indicator of the stability of multi-hop routes whereas link duration (LD) indicates the stability of single-hop links.

We now investigate the relationship between LD and RD. It is very difficult to derive rigorous expressions for RD using existing mobility models such as random waypoint, random walk, and boundless simulation area. Hence, in this section, we present a series of simulations with these mobility models, which show that there exists a strong correlation between LD and RD. Then, in Section 2.5, we analyze the relationships between LD and RD by using our CV model.



(a) All nodes have routes to N_0.

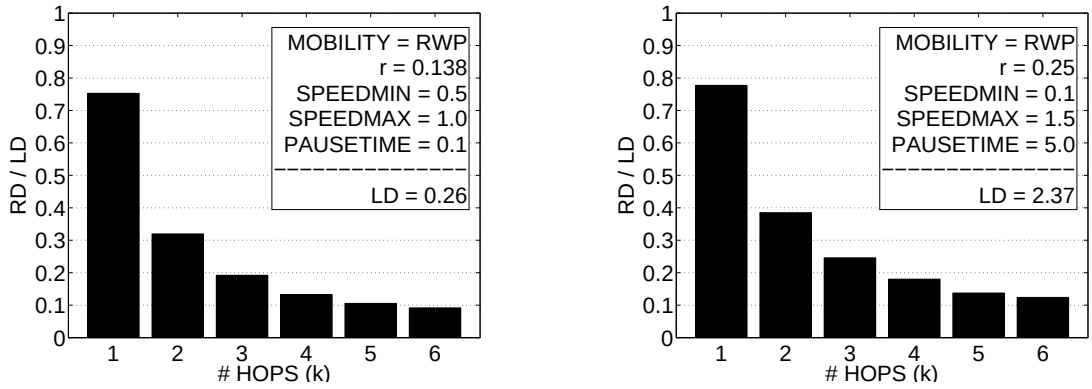
(b) Routes from N_2 and N_3 to N_0 are broken.

Figure 2.4: Examples of multi-hop routes. White nodes have multi-hop routes to N_0 while routes from black nodes are broken.

The simulation is organized as follows. First, 100 mobile nodes are randomly placed on a 1x1 plane. The nodes start moving according to the given mobility model. At time $t = 0$, the simulator computes the shortest, multi-hop path from each node to the root node N_0. Figure 2.4 shows examples of multi-hop connections. White circles correspond to the nodes that have multi-hop routes to N_0, and black circles denote the nodes that have lost their connections to N_0. For instance, in Figure 2.4(a), node N_3 has the shortest path $N_3 \rightarrow N_2 \rightarrow N_1 \rightarrow N_0$ to the root N_0. If the link between N_2 and N_1 breaks at time t (see Figure 2.4(b)), the simulator records that "a 2-hop route from N_2 to N_0 is broken at time t," and "a 3-hop route from N_3 to N_0 is also broken at time t." In other words, when an n-hop route is broken at time t, all routes with later hops that pass along the broken n-hop route are also regarded as broken at time t. This procedure measures the RD of routes from the shortest path discovery to their break times.

2.4.1 Random Waypoint Model

Mobile nodes in the random waypoint (RWP) [47] model behave as follows. First, a node selects a random destination point within a bounded movement area, and moves



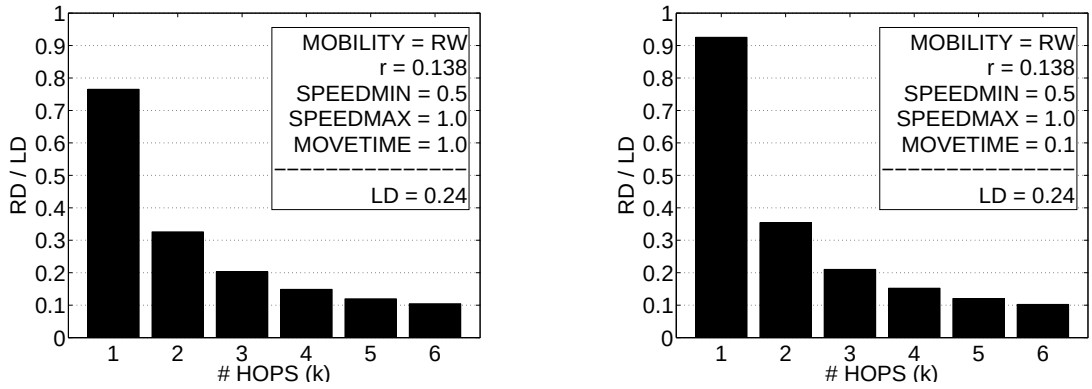
(a) High-mobility case (b) Low-mobility case

Figure 2.5: Simulation results with RWP model

toward it at a random speed v. Once it arrives at the destination, it pauses for a predefined *pausetime*. The speed v is a random variable uniformly distributed between *speedmin* and *speedmax*. By varying r, *speedmin*, *speedmax*, and *pausetime*, we can control the the node movements.

Figure 2.5(a) shows simulation results corresponding to a high-mobility case with mobility parameters: $r = 0.138$, *speedmin* $= 0.5$, *speedmax* $= 1.0$, and *pausetime* $= 0.1$, for which link duration $\overline{T}_{LD} = 0.26$. The RD/LD ratio of 0.32 for 2-hop links, for example, indicates that the average residual duration of a 2-hop link is $0.32 \times \overline{T}_{LD} = 0.0832$.

Figure 2.5(b) corresponds to a low-mobility case: $r = 0.25$, *speedmin* $= 0.1$, *speedmax* $= 1.5$, and *pausetime* $= 5.0$, for which $\overline{T}_{LD} = 2.37$. In this case, due to the relatively long *pausetime*, a considerable number of static nodes are observed, that is, nodes in their *pause* state. Consequently, this case has a larger LD value than the previous case. It can be seen that although the two cases differ in their mobility parameters and LDs, their RD/LD ratios are nearly identical within an error of a few percent. Instead of changing *pausetime* alone, we also varied the *speedmin* and *speedmax* values and obtained nearly the same distribution of RDs.



(a) Long *movetime* case

(b) Short *movetime* case

Figure 2.6: Simulation results with RW model

2.4.2 Random Walk Model

Node movement in the random walk (RW) [45] model is determined by the following rules. First, each node decides the direction in which to move. Once it starts moving, it goes on for a predefined *movetime*, at the end of which it selects a new direction. At every random decision of movement direction, the speed is also randomly chosen from an interval (*speedmin*, *speedmax*). When a node reaches the system boundary, it bounces off the border with an angle equal to the incoming angle, and continues until *movetime* expires.

Figure 2.6(a) shows simulation results for RW with mobility parameters: r = 0.138,

speedmin = 0.5, *speedmax* = 1.0, and *movetime* = 1.0, for which $\overline{T}_{LD} = 0.24$. Note that in this case, *movetime* is sufficiently long, and lots of nodes can move from one border to the opposite border before *movetime* expires, which makes this mobility condition similar to that of our CV model. In fact, *movetime* longer than 0.3 does not affect the distribution of RD because the nodes spend most of their time moving at a constant velocity.

Figure 2.6(b) illustrates a case with a short *movetime* = 0.1. The other parameters are as follows: r = 0.138, *speedmin* = 0.5, and *speedmax* = 1.0, for which $\overline{T}_{LD} = 0.24$. With

these parameters, the RDs show nearly the same relationship with LD as in the previous case. An exception is observed in the single-hop case, where 1-RD is over 90% of LD, while in other cases, 1-RD/LD ranges from 70% to 80%. This issue is reexamined in Section 2.5.3.

2.4.3 Boundless Simulation Area

The boundless simulation area (BSA) [35] model differs from the other models in two respects. First, instead of abrupt velocity and direction changes, the speed and direction of BSA nodes are randomly accelerated as follows. The linear acceleration A is a random variable uniformly distributed between $-A_{\max}$ and $+A_{\max}$, where $A_{\max}$ is a predefined parameter. The angular acceleration ω is a random variable uniformly distributed between $-\omega_{\max}$ and $+\omega_{\max}$. Every Δt, these random values are updated, and the speed and the movement direction of a node are determined as follows.

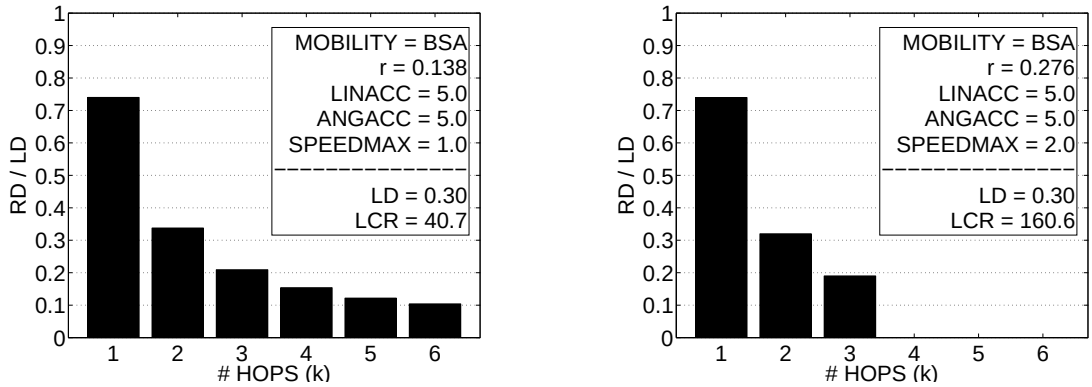
$$v(t + \Delta t) = \max\{\min\{v(t) + A\Delta t, speedmax\}, 0\}$$

$$\phi(t + \Delta t) = \phi(t) + \omega\Delta t$$

Second, BSA has a toroidal space while most other mobility models assume bounded planes. Hence, distance is measured by a toroidal metric defined as follows.

$$D = \{(\min(|x_1 - x_2|, |X_{\max} - x_1 + x_2|)^2 + (\min(|y_1 - y_2|, |Y_{\max} - y_1 + y_2|)^2\}^{\frac{1}{2}}$$

Figure 2.7(a) shows simulation results for BSA with mobility parameters: $r = 0.138$, linear acceleration $A_{\max} = 5.0$, angular acceleration $\omega_{\max} = 5.0$, and $speedmax = 1.0$, for which $\overline{T}_{LD} = 0.30$. It can be seen that even with the toroidal distance metric, the RD/LD ratio has the same distribution as before. Figure 2.7(b) uses parameters $r = 0.276$, linear acceleration $A_{\max} = 5.0$, angular acceleration $\omega_{\max} = 5.0$, and $speedmax = 2.0$, and has



(a) Short transmission range case

(b) Long transmission range case

Figure 2.7: Simulation results with BSA model

$\overline{T}_{LD} = 0.30$. Note that this case has a long transmission range $r = 0.276$, and that the graph connectivity in the BSA model is better than that of the usual planar and bounded models because BSA allows more links around the system "edges." Consequently 4- to 6-hop links rarely appeared in this long transmission range simulation, and are omitted from the plot. Nevertheless, the 1- to 3-RDs show the same distribution as the other mobility models, RWP and RW.

Consider the two network conditions in Figures 2.7(a) and 2.7(b). We obtained the same LDs from these two networks, namely, $\overline{T}_{LD,1} = \overline{T}_{LD,2} = 0.3$. From another set of simulations, we obtained LCRs for the two networks as $\lambda_{LCR,1} = 40.7$ and $\lambda_{LCR,2} = 160.6$, respectively. These two cases have the same LDs but different LCRs. Note that they have the same RDs for 1- to 3-hop routes. Hence, it is clear that LD predicts lifespan of routes more accurately, and so is a better unified mobility metric than LCR.

2.5 Mobility Metric Relationships

In the previous section, we observed that LD has a consistent relation with RD for a wide range of mobility models and parameter assignments. Now, using the CV model, we

discuss what causes the strong correlation between LD and RD. We also investigate the relations between LD, LCR, and RD.

2.5.1 k-RD and LD

First, we ask: What makes LD a better mobility metric than LCR? Suppose at time $t = 0$, a mobile node N_0 observes a neighbor node N_1 within its transmission range, and at time $t = T_0$, the neighbor leaves the range. This exit time $T_0 > 0$ is a random variable, and its cumulative distribution function (cdf) and probability density function (pdf) are $F(t)$, and $f(t)$, respectively, i.e., $p(T_0 \leq t) = F(t) = \int_0^t f(u)du$. On the other hand, we can view the probability of complete transmission $p_{comp}(T_{comm} \cdot v/r)$, which was introduced in Section 2.3.1, as $p(T_0 > T_{comm})$. Hence, the probability that a link remains connected until time t is given by

$$p_{comp}(t \cdot v/r) = 1 - F(t) \tag{2.8}$$

Next, suppose that a k-hop route consists of $k + 1$ mobile nodes $\{N_i\}$, $0 \leq i \leq k$, and each node pair (N_{j-1}, N_j), $1 \leq j \leq k$, is connected by a link L_j. At time $t = 0$, the route is connected, and at time $t = T_1$, the route breaks. Let $G_k(t)$ and $g_k(t)$ denote the cdf and pdf, respectively, of the random variable $T_1 > 0$. Then, the probability that the route remains connected until time t is $1 - G_k(t)$, and k-RD for this route is given by

$$\overline{T}_{RD,k} = \int_0^\infty t \cdot g_k(t)dt \tag{2.9}$$

We assume that the generation/break processes of k links that constitute the k-hop route are mutually independent. This assumption is not strictly true, but it has been shown that the link generation/break processes of links that are far apart are nearly independent, and even the link durations of neighboring links which share a common node have a negligible

correlation [36]. Thus, the probability that the k-hop route remains connected until time t becomes

$$1 - G_k(t) = [1 - F(t)]^k \tag{2.10}$$

From (2.8)-(2.10), we get k-RD as

$$\overline{T}_{RD,k} = \int_0^\infty t \cdot [\frac{d}{dt}\{1 - p_{comp}(t \cdot v/r)^k\}]dt$$

Substituting from (2.3) yields

$$\overline{T}_{RD,k} = \int_0^\infty t \cdot [\frac{d}{dt}\{1 - p_{comp}(\frac{\pi^2}{8}\frac{t}{\overline{T}_{LD}})^k\}]dt$$

Hence, it can be seen that RD is a function of LD rather than LCR, which makes LD a good indicator of multi-hop connection stability.

2.5.2 LD and LCR

Next, we investigate the relationship between LD and LCR. From (2.2) and (2.3), we get

$$\frac{1}{2} \cdot \lambda_{LCR} \cdot \overline{T}_{LD} = \rho\pi r^2 \tag{2.11}$$

which implies that the product of half of LCR and LD equals the average node degree. This relation also follows from Little's theorem, which states that the average number of customers in a system, $\rho\pi r^2$, is equal to the product of the customer arrival rate, $\lambda_{gen} = \lambda_{LCR}/2$, and the average time customers spend in the system, $\overline{T}_{LD}$. Table 2.2 shows examples of $\lambda_{LCR} \cdot \overline{T}_{LD}/2$ values for three mobility models. In all cases, the nominal average node degree $\rho\pi r^2$ is set to 6.0, and it is expected that $\lambda_{LCR} \cdot \overline{T}_{LD}/2 = 6.0$.

The discrepancies between (2.11) and the simulated data in Table 2.2 are due to the

Mobility model	Parameters	Real average node degree
BSA	ρ = 100 r = 0.138 Maximum speed = 1.0 Linear acceleration = 5.0 Angular acceleration = 5.0	5.95
RWP	ρ = 100 r = 0.138 Minimum speed = 0.5 Maximum speed = 1.0 Pause time = 0.1	7.57
RW	ρ = 100 r = 0.138 Minimum speed = 0.5 Maximum speed = 1.0 Movement time = 1.0	5.41
RW	ρ = 400 r = 0.069 Minimum speed = 0.5 Maximum speed = 1.0 Movement time = 1.0	5.80

Table 2.2: Simulation results for real average node degree $\lambda_{LCR}/2 \cdot \overline{T}_{LD}$. Nominal average node degree is set to $\rho\pi r^2 = 6.0$

border effect noted in [8]. Like the CV model, the BSA model has a borderless space in which nodes move, and the simulated values of $\lambda_{LCR} \cdot \overline{T}_{LD}/2$ and $\rho\pi r^2$ are very close. However, in the RWP model, the node density around the center area is higher than that around the system border because nodes tend to cross the center of the system area. In consequence, the actual number of neighbors exceeds $\rho\pi r^2$, and hence, $\lambda_{LCR} \cdot \overline{T}_{LD}/2 > \rho\pi r^2$. On the other hand, the nodes in the RW model have fewer neighbors if they are located around the border. The first simulation result with the RW model shows that $\lambda_{LCR} \cdot \overline{T}_{LD}/2 < \rho\pi r^2$ due to the border effect. If the transmission range is short, as in the second example with RW, the border effect is reduced and $\lambda_{LCR} \cdot \overline{T}_{LD}/2$ gets closer to the nominal average

node degree $\rho\pi r^2$. Therefore, it can be seen that if a mobility model has a uniform node distribution without the border effect, then LD can be easily estimated from LCR using (2.11).

2.5.3 LD and 1-RD

It was pointed out in Section 2.4.2 that when *movetime* is short (see Figure 2.6(b)), the 1-RD value is longer than in other cases. This exceptional behavior results from the fact that short *movetime* causes node motion to appear memoryless. When linear movement is dominant, the nodes that have already spent much time in the transmission region of another node tend to exit the region earlier than ones just entering the region. However, as the node speed grows, and *movetime* decreases, the node movement converges to Brownian motion [28]. In this memoryless movement case, the time a node has already spent in the transmission region of another node becomes independent of the time it will spend in the transmission region later. Hence, the residual duration of single-hop links, that is, 1-RD, tends to be close to the total link duration LD. On the other hand, when the linear or near-linear movement dominates, the node mobility resembles the CV model, and we can observe the relation between 1-RD and LD, as in Section 2.4. Thus, we conclude that we can obtain the consistent ratio of 1-RD to LD as long as the linear movement is dominant, which holds for a wide extent of existing mobility models.

Apart from the memoryless property, $\overline{T}_{RD,1} > \overline{T}_{LD}/2$. This results from the randomness of link duration. Suppose a node N_0 has two types of neighbors $\{N_A\}$ and $\{N_B\}$. The neighbors have an identical arrival rate λ and different link durations T_A and T_B, respectively. LD is measured considering all neighbor arrivals to N_0, and is given by

$$\overline{T}_{LD} = \frac{T_A + T_B}{2}$$

which is the arithmetic mean of T_A and T_B, equally weighted. On the other hand, if N_0 randomly picks one of its existing neighbors and observes the neighbor's duration, the average duration becomes

$$\overline{T}'_{LD} = \frac{\lambda T_A^2 + \lambda T_B^2}{\lambda T_A + \lambda T_B} = \frac{T_A^2 + T_B^2}{T_A + T_B}$$

which is the arithmetic mean of T_A and T_B weighted by the expected number of neighbors of each type, i.e., λT_A and λT_B, respectively. It can be seen that if $T_A \neq T_B$, then the sampled average $\overline{T}'_{LD}$ is always greater than the actual average $\overline{T}_{LD}$, which is called *the inspection paradox* [39]. It is plausible to presume that the RD of a single-hop link is half of $\overline{T}'_{LD}$. Therefore, the RD of single-hop links exceeds $\overline{T}_{LD}/2$.

2.6 Summary

We introduced the constant velocity (CV) model to evaluate the relationship between mobility and connection stability in ad hoc networks. Though simplified, the CV model fairly accurately describes the behavior of single-hop communication, as we verified with simulations.

In addition, we investigated the role of link duration (LD) as a mobility metric and evaluated the lifespan of multi-hop routes with respect to LD. Extensive simulations with various mobility models demonstrate that this relation between LD and the average route lifespan is consistent for a wide range of mobility models. Also we clarified what makes LD a good mobility metric by showing that the lifespan of a multi-hop route is a function of LD.

We also showed that link change rate is not suitable as a unified metric because, unlike LD, its relation with the route lifespan depends on the node density, which may not be

uniform in some mobility models. A recent study by Nayebi *et al.* has observed that the probability distribution of LD for the RWP model is similar to that of our CV model [63]. Furthermore, using the boundless random direction model (BRDM) derived from CV, they showed that the pdf of RWP can be approximated fairly accurately by adding stationary nodes. From these observations, it can be seen that LD with RWP is practically equivalent to that with CV. Therefore, we conclude that LD is a good unified mobility metric for most types of mobile ad hoc networks, and that CV is a useful model for mobility analysis.

Our analysis of the CV model motivates development of new adaptive network control algorithms. When a node senses a high mobility condition, it could reduce the packet length to maintain high success rate for single-hop communication. On the other hand, when too short a lifespan of multi-hop routes is expected, the node might increase its transmission power to get a longer lifespan of multi-hop routes to the destination, and thus achieve satisfactory data delivery rates.

CHAPTER 3

Distributed Power-Aware Link Maintenance

We propose a power-aware link maintenance (PALM) algorithm for mobile ad hoc networks (MANETs) that simultaneously performs transmission power control and route connectivity maintenance. Unlike most topology control algorithms, PALM manages the transmission power of active nodes only, and thus eliminates energy and channel resource waste due to unnecessary beaconing. The basic idea of PALM is that by recording the received signal strength on packets, each node continuously estimates the required transmission power, and adapts to location changes due to mobility. We also introduce some efficient local link repair schemes. When a route change is needed due to node movement, overhearing nodes participate in data forwarding, and locally repair the route before disconnection without the need to propagate route error messages. Through these operations, PALM prevents frequent link breaks due to node mobility, reduces the occurrence of rerouting, and significantly improves network performance. Experimental results with the *ns* simulator confirm that PALM effectively conserves communication energy with modest overhead. In addition, we investigate how MAC parameters such as the carrier-sense threshold and the transmission power for control packets affect the network performance. We evaluate the tradeoff relation between network throughput and energy efficiency. Further experiments

show that the use of the maximum transmission power for control packets may adversely affect the energy efficiency, and when no significant radio obstruction exists in the network space, the carrier-sense multiple access method can have higher energy efficiency than the RTS/CTS handshake.

3.1 Introduction

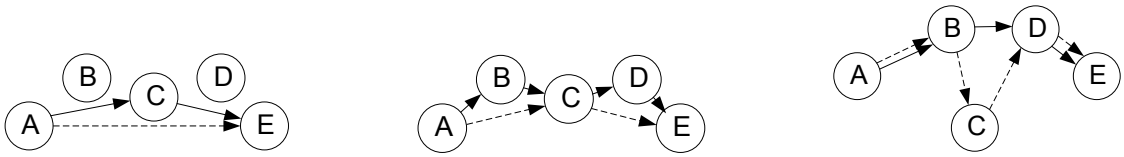
Transmission power control, which is to determine the strength of the radio signal transmitted by wireless transceivers, is a key issue in MANETs for the following reasons. First, as the mobile hosts are usually battery-powered, the power consumption level due to radio transmission affects the network lifetime. Second, the network topology depends on the hosts' transmission power levels, and affects the network performance such as communication throughput and delivery latency. In general, short transmission range is desirable as long as the network connectivity is maintained since it reduces energy consumption and signal collision between transmitters.

Topology controls [11, 55, 76, 96] have long been considered for MANET power management. They attempt to reduce the transmission power while maintaining a connected network topology. However, topology control algorithms have several drawbacks. First, they require periodic beaconing, which wastes energy and channel resources. Second, though they provide network connectivity, the actual connected routes still have to be discovered by the routing layer. Thus, even when the network is connected, frequent rediscovery of connected routes may occur. Third, in order to make the network insensitive to node mobility, topology control algorithms must allow longer transmission range, which may worsen the network performance.

The so-called BASIC power control [49] algorithm has been proposed to mitigate the

above problems. The transmitter computes the minimum required transmission power for data delivery based on received signal strength (RSS) at the receiver. Then, the data (DATA) and acknowledgement (ACK) packets are transmitted at minimum power, while control packets such as request-to-send (RTS) and clear-to-send (CTS) are transmitted at full power. In this way, BASIC reduces the energy consumption without unnecessary beaconing by the nodes that are not transferring data. However, BASIC still has some shortcomings. The nodes must abide by the data delivery routes discovered by the routing layer, which may not be power-efficient. The control packets transmitted at full power may collide with DATA and ACK packets between other nodes. The power control MAC (PCM) protocol [49] mitigates the latter problem by allowing periodic full-power transmission for DATA packets, but the former problem still persists.

In an effort to discover power-efficient routes in a distributed manner without periodic beaconing, the power-aware routing optimization (PARO) algorithm [30, 31] has been proposed. Figure 3.1 shows an example of PARO's routing process; dotted arrows indicate the routes before, and solid after each iteration. PARO assumes that all nodes are within transmission range of each other. At the beginning of the first iteration, the source node A directly sends packets to the destination node E. During the first iteration, node C, for instance, overhears the communication from A to E, and determines that route redirection via itself conserves energy. Then, at the end of the iteration the route becomes ACE. Similarly, during the second iteration, nodes B and D redirect the route, and the route becomes



(a) First iteration (b) Second iteration (c) Redirector removal

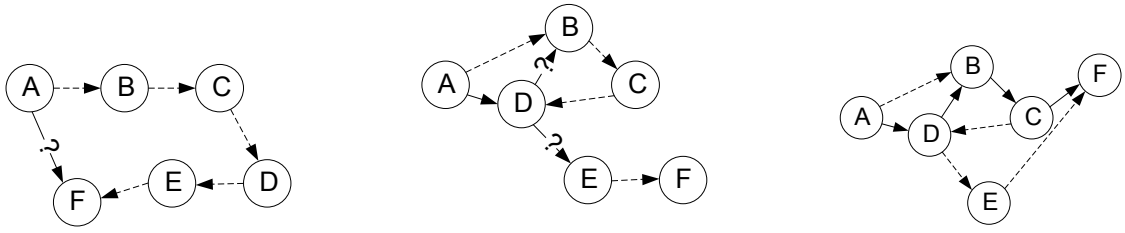
Figure 3.1: Example of routing by PARO

ABCDE, which consumes less energy than *AE* or *ACE*.

Sometimes routes redirected by the above procedure become inefficient, as in the dotted route in Figure 3.1(c), where a shorter route *ABDE* is more power-efficient. PARO removes unnecessary redirectors such as *C* in this example as follows. While *B* is transmitting to *C*, node *D* overhears the communication, and determines that direct delivery from *B* to *D* costs less energy than *BC*, and thus obviously less than *BCD*. Then, *D* transmits a redirect request, and the route changes from *ABCDE* to *ABDE*. Through these redirection and redirector removal procedures, PARO discovers power-efficient routes in a distributed manner, and saves communication energy.

However, PARO still carries the risk of generating inefficient routes. Unlike the obvious case in Figure 3.1(c) where the energy required for the single link *BC* exceeds that for *BD*, if each constituent link in *ABCDEF* in Figure 3.2(a) consumes less energy than the direct link *AF*, node *F* cannot tell which of routes *ABCDEF* and *AF* is more power-efficient. Thus, the route may become unnecessarily long when there exist shortcuts that consume less energy.

A more complicated case is illustrated in Figure 3.2. Suppose PARO has previously discovered a power-efficient route *ABCDEF*, and nodes have moved, as in Figure 3.2(b). On overhearing *A* and *B*, node *D* determines that the redirection *ADB* can reduce energy



(a) Removal of multiple redirectors

(b) Determination of the next hop node

(c) Forbidden redirection

Figure 3.2: Examples of some shortcomings of PARO

consumption between AB. Now the question is: Which node should D assume as its next hop node? In the example of Figure 3.1(c), node D recognizes that it should not use C as its next hop node to avoid creating a loop as C is its previous hop node. However, node D in Figure 3.2(b) cannot tell which node to use for its next hop. In this case, taking B as its next hop node will cause a loop $ADBCDB\cdots$. In fact, node D's decision should depend on which of B and E has the smaller hop distance to the destination. The problem, however, is that hop distance cannot be easily determined by PARO. For instance, the hop distance of node B in Figure 3.1(b) changes due to redirections occurring at later hops such as the redirection of CE. Therefore, in order to avoid loops, PARO should not change the redirectors' next hop nodes except in such cases as Figure 3.1(c). Even when routes such as $ADBCF$ in Figure 3.2(c) are desirable, PARO cannot allow them, and may produce inefficient routes such as $ABCDEF$. Thus, although PARO is capable of discovering routes that are initially power-efficient, its efficiency as a link maintenance scheme for mobile networks is limited.

Consider a MANET consisting of n nodes with node speed v, a fixed transmission range r, and a routing frequency $f_{routing}$. Assuming that the average hop distance is proportional to $\sqrt{n}$ and that the routing algorithm discovers shortest path routes, the mobility analysis in [20] suggests

$$f_{routing} \propto \frac{v\sqrt{n}}{r}$$

which, in turn, implies that efficient link maintenance is particularly important for large and highly mobile networks with short transmission ranges.

We propose a new scheme called power-aware link maintenance (PALM) that resolves the above problems with BASIC and PARO. We first determine a *virtual* hop distance to efficiently assign and measure hop distances in the presence of route redirections. Second, we add an *accumulated energy* field to packets, and enable the removal of multiple

redirectors. Third, we use different transmission power levels for control and broadcast packets to mitigate the problems with BASIC's power control. These new techniques allow PALM to discover more power-efficient routes than PARO and similar algorithms. We also investigate how medium access control (MAC) parameters affect the performance of ad hoc networks in which the transmission power for DATA packets is set at the minimum necessary value. Then, we present means to maximize the network throughput, while maintaining a high energy efficiency.

This chapter is organized as follows. Section 3.2 describes our new PALM algorithm, and Section 3.3 presents its basic procedures. Section 3.4 presents simulation results for PALM, and compares its performance with other algorithms. Section 3.5 investigates how MAC operation affects the network performance by considering the network throughput and the energy efficiency. Finally, Section 3.6 summarizes the chapter.

3.2 Power-Aware Link Maintenance

This section describes the basic operation of the proposed PALM scheme. The medium access control (MAC) in PALM is based on IEEE 802.11. We assume ad hoc on-demand (AODV) [71] as the routing algorithm. We also assume that transceivers are equipped with omnidirectional antennas, and that receivers can detect the received signal strength (RSS). The radio model presented in Section 1.2.3 is also adopted here, i.e.,

$$RSS(RX) = Pow(TX) \cdot Gain(TX, RX) \tag{3.1}$$

where $RSS(RX)$ and $Pow(TX)$ denote the RSS at RX and the transmission power at TX, respectively. In other words, the received signal strength is proportional to the transmission power. We further assume that the radio channel is symmetric, i.e., $Gain(A, B) \equiv$

$Gain(B, A)$. The issue of channel asymmetry is discussed in Section 5.2.

3.2.1 Power Control

PALM's power control method is similar to that of the BASIC power control scheme [49] except for the power levels of RTS/CTS packets. PALM controls transmission power as follows. The transmitter TX records its transmission power level $Pow(TX)$ on outgoing packets, and receivers RX_i, including overhearing nodes, estimate the channel gain $Gain(TX, RX_i)$ from (3.1). After receiving the channel gain $Gain(TX, RX_i)$ from RX_i, TX sends DATA packets at the power level given by

$$Pow(TX, RX_i) = \beta \cdot \frac{RSS_{\min}}{Gain(TX, RX_i)} \tag{3.2}$$

where $RSS_{\min}$ and β denote the minimum RSS for successful reception and a safety factor, respectively. The RSS threshold is determined by the receiver sensitivity. We need a safety factor $\beta > 1$ as transmission power previously estimated from (3.2) may vary due to changes in the distance between nodes. We define the minimum level of the RSS with the safety factor taken into account as

$$RSS_0 \triangleq \beta \cdot RSS_{\min}$$

Replacing $RSS_{\min}$ by RSS_0/β, (3.2) becomes

$$Pow(TX, RX) = \frac{RSS_0}{Gain(TX, RX)} \tag{3.3}$$

If (3.3) exceeds the maximum transmission power $Pow_{\max}$ supported by the transceiver hardware of TX, then TX sets its transmission power at $Pow_{\max}$.

In addition to DATA and ACK packets, IEEE 802.11 utilizes the exchange of RTS and CTS packets. BASIC uses minimum power for DATA and ACK packets and maximum power for RTS and CTS. We propose to improve the communication concurrency by reducing power for control packets, which is discussed later in Section 3.2.5. For now, we assume that minimum power is used for DATA and ACK packets, and relatively greater power for RTS and CTS; the impact of the transmission power of control packets on the network performance will be investigated in depth in Section 3.5. Finally, transmission power for broadcast packets, for which no particular receiver is specified, is determined in a similar manner as for control packets.

3.2.2 Route Redirection

PALM performs route redirection in a similar way to PARO. In Figure 3.3, node i is forwarding DATA packets to its next hop node j, and node k is overhearing them. Solid lines in the figure indicate intended packet flow, and dotted lines, overheard flow. For consistency, we call the nodes that transmit RTS and DATA packets *initiators*, and the ones that transmit CTS and ACK, *responders*, as suggested in [10]. The recipient nodes located

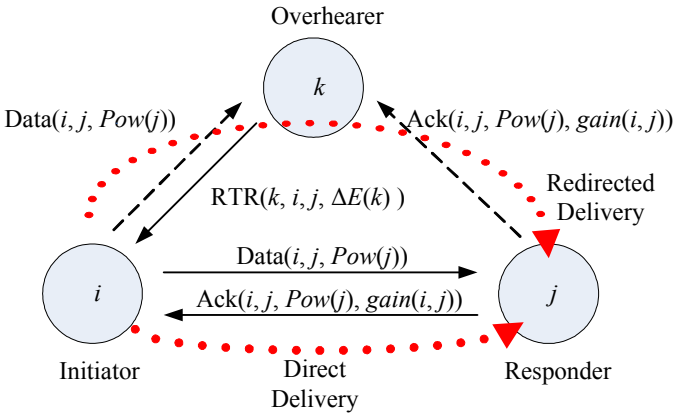


Figure 3.3: Communication between nodes for route redirection

at the end of multi-hop routes are called *final destinations.* The DATA packets contain the initiator address i, the responder address j, and i's transmission power $Pow(i)$. Node j replies with an ACK message containing i, j, $Pow(j)$, and $Gain(i, j)$. The overhearing node k now can determine which of the direct $i \rightarrow j$ and the redirected $i \rightarrow k \rightarrow j$ routes consumes less energy by comparing the following values:

$$E(direct) = \frac{RSS_0}{Gain(i, j)} \cdot T_{TX}$$
$$E(redirected) = \frac{RSS_0}{Gain(i, k)} \cdot T_{TX} + \frac{RSS_0}{Gain(k, j)} \cdot T_{TX}$$

where T_{TX} denotes the transmission time for each packet, which we assume to be a constant. The overhearing node k estimates the energy saving ΔE_k as

$$\Delta E_k = \frac{RSS_0}{Gain(i, j)} - \left(\frac{RSS_0}{Gain(i, k)} + \frac{RSS_0}{Gain(k, j)}\right) \tag{3.4}$$

In order to prevent excessively frequent route redirection with small power improvement, redirection is allowed only when

$$\alpha \cdot \frac{RSS_0}{Gain(i, j)} > \left(\frac{RSS_0}{Gain(i, k)} + \frac{RSS_0}{Gain(k, j)}\right) \quad (\alpha < 1) \tag{3.5}$$

After overhearing ACK, node k sends a request-to-redirect (RTR) packet to node i. Receiving an RTR message, node i modifies its routing table such that its new next hop node becomes k instead of j. The AODV routing layer stores a routing table in the following

format:

$$rTable(fdst_1) = (n_1, h_1)$$

$$rTable(fdst_2) = (n_2, h_2)$$

$$\vdots$$

where $fdst_i$, n_i, and h_i denote the final destination address, the next hop address to reach $fdst_i$, and the number of hops to $fdst_i$, respectively.

Suppose that before redirection, node i has $rTable(f) = (j, h)$ where the final destination is f. On receiving an RTR from node k, node i modifies its routing table to

$$rTable(f) = (k, h)$$

and node k adds an entry

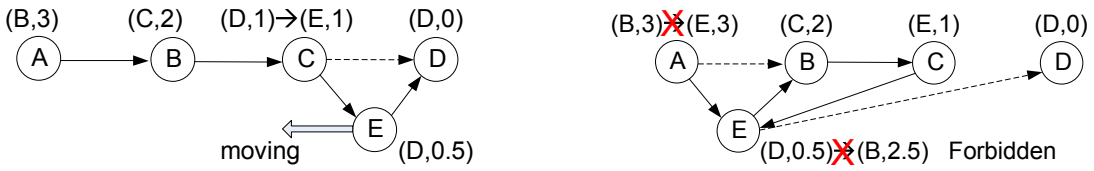
$$rTable(f) = (j, h')$$

to its own routing table. Subsequent packets are delivered along $i \rightarrow k \rightarrow j$ instead of $i \rightarrow j$. Determination of h' for the newly added entry is described in the next section.

There might be more than one overhearing node that sends an RTR to the source. When a node that is to send an RTR message overhears another RTR with a larger ΔE value, it does not send an RTR message, and discards its own RTR.

3.2.3 Loop Avoidance

The foregoing route redirection technique enables nodes to repair routing tables locally without propagating rerouting requests. However, such local repair carries the risk of generating loops in the delivery path. Consider the case in Figure 3.4 where A is the data source,



(a) Allowed redirection

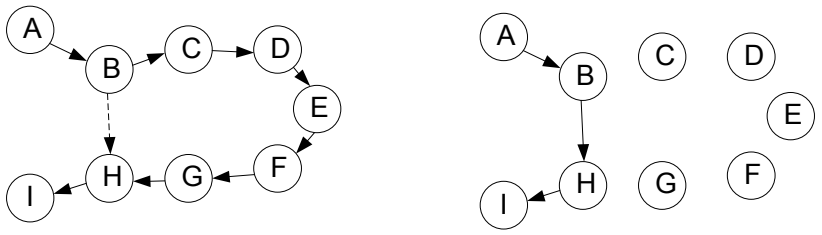
(b) Forbidden redirection

Figure 3.4: Example of loop avoidance

and D the final destination. Node E, which is moving left, determines that redirection of link CD via E conserves energy, sends an RTR to C, and the route is redirected; dotted arrows denote the data path before redirection, and solid ones, after redirection. During redirection, the routing table in C is modified from $(D, 1)$ to $(E, 1)$, and node E records D as the next hop node. Now suppose E has moved to the location depicted in Figure 3.4(b). It determines that redirection of link AB via itself conserves energy, and so it sends an RTR to A. Due to this redirection, A's routing table is modified from $(B, 3)$ to $(E, 3)$, and E has to modify its next hop node to B. After this redirection, a loop $AEBCEB\cdots$ appears.

PALM resolves such route loop problems in the following way. First, we define the *virtual hop distance* to the final destination as a rational number instead of an integer. After a route is discovered, the real-numbered virtual hop distance has the same value as the usual integer-numbered hop distance. When redirecting a link between nodes X and Y with hop distances h_X and h_Y, respectively, the redirector R computes its (virtual) hop distance as $(h_X + h_Y)/2$. By taking this value between h_X and h_Y, we can maintain the invariant condition of *monotonically decreasing hop distance numbers along routes*, without propagating the redirection event to other nodes. For instance, in Figure 3.4(a), node E assumes $(1 + 0)/2 = 0.5$ as its hop distance.

Second, PALM *locks* the links when packets are forwarded along the links or when a link entry is added. Whenever the routing layer refers to a routing table entry, the corresponding link to the next hop node is locked, and it gets *unlocked* after T_{unlock} seconds.



(a) Before warping (b) After warping

Figure 3.5: Example of route warping

Thus, if a link is locked, it implies that the link is along or around an active route. When the redirector node E overhears communication from nodes X to Y, and its recorded hop distance h_E is larger than h_X, it does not send an RTR unless it is unlocked, as this redirection will cause a loop. For instance, in Figure 3.4(b), node E does not send an RTR to A because the hop distance 3 of its new candidate source, node A, is greater than its recorded hop distance 0.5. The conditions for redirection are described in detail in Section 3.3. Hence, by forbidding redirections that generate loops, PALM can continuously construct power-efficient and loop-free routes.

3.2.4 Route Warping

The route redirection techniques described above tend to shorten per-hop distance and increase the number of hops between source and destination. Thus, the route may become inefficient due to location changes as illustrated in Figure 3.5(a). The route from A to I is $ABCDEFGHI$ while shortcuts such as $ABHI$ may be more power-efficient. PALM discovers such shortcuts through the following *route warping* technique.

The transmission energy E_i, consumed by node i to reach node $i+1$, is given by

$$E_i = Pow(i, i+1) \cdot T_{TX} = \frac{RSS_0}{Gain(i, i+1)} \cdot T_{TX} \tag{3.6}$$

where T_{TX} denotes the packet transmission time, which we assume to be a constant. We then define

$$E_{acc,i} \triangleq \sum_{j=0}^{i-1} E_j \quad \text{for } i \geq 1$$

which indicates the accumulated energy consumption from the source node to reach node i. We define $E_{acc,0} = 0$ at the source node $i = 0$.

In PALM, each RTS/DATA packet contains a field $E_{acc,i}$. Node i, upon receiving the packet with $E_{acc,i-1}$ from the previous hop, adds E_{i-1} to the field $E_{acc,i-1}$ to get $E_{acc,i}$.

$$\begin{aligned} E_{acc,i} &= E_{acc,i-1} + E_{i-1} \\ &= E_{acc,i-1} + \frac{RSS_0}{Gain(i-1,i)} \end{aligned} \tag{3.7}$$

Node i records the value $E_{acc,i}$ in its memory, and then forwards packets along with the accumulated energy field.

In Figure 3.5(a), by overhearing RTS and DATA packets from B to C, node H can determine whether route warping is advantageous from the following relation.

$$\begin{aligned} \Delta E_{BH} &= E(ABCDEFGH) - E(ABH) \\ &= E_{acc,H} - \left(E_{acc,B} + \frac{RSS_0}{Gain(B,H)}\right) \end{aligned}$$

If $\Delta E_{BH} > 0$, then the shortcut from B to H is more power-efficient. In that case, node H sends an RTR to B, the long route $ABCDEFGHI$ is warped to $ABHI$, and the route becomes more power-efficient, as in Figure 3.5(b). Note that even when node H cannot overhear DATA packets from B to C, it may be able to overhear the corresponding RTS packets as control packets are transmitted at greater power than DATA/ACK packets. In general, node Y can compute the energy saving ΔE_{XY} with warping from node X to node Y

from the equation:

$$\Delta E_{XY} = E_{acc,Y} - \left(E_{acc,X} + \frac{RSS_0}{Gain(X,Y)}\right)$$

In order to prevent nodes with stale E_{acc} records from sending an RTR, PALM marks as *used* the links along which packets are being forwarded, and as *unused* after T_{unused} seconds.

Consider the route $ABCED$ in Figure 3.4(b). As node E moves left, the direct delivery from C to D becomes more efficient than the redirected delivery CED. Then, node D sends an RTR to C, and the route $ABCED$ is warped, and becomes $ABCD$. After T_{unlock} seconds, node E is unlocked. Finally, E redirects the link AB, and a power-efficient route $AEBCD$ is constructed.

3.2.5 Transmission Power for RTS/CTS

As mentioned in Section 3.2.4, the route redirection techniques of PALM tend to increase the number of hops from source to destination, and thus increases channel contention between nodes. In particular, the long transmission range for RTS/CTS packets may worsen this problem. With the BASIC power control scheme, control packets such as RTS and CTS and broadcast packets use maximum transmission power, whereas DATA and ACK packets are sent at minimum power. Suppose that node 4 in Figure 3.6 is transmitting an RTS packet at full power. All nodes except for node 9 are within node 4's transmission range, and thus cannot transmit during the time specified by the RTS packet although transmission from node 0 to node 1, or from node 7 to node 8 would not affect the communication between nodes 4 and 5.

PALM mitigates the above problem by reducing transmission power for broadcast and control packets as follows. We say a node i regards a neighbor j as *active* if j has been overheard by i within the last T_{act} seconds forming an *active neighbor window*. PALM requires

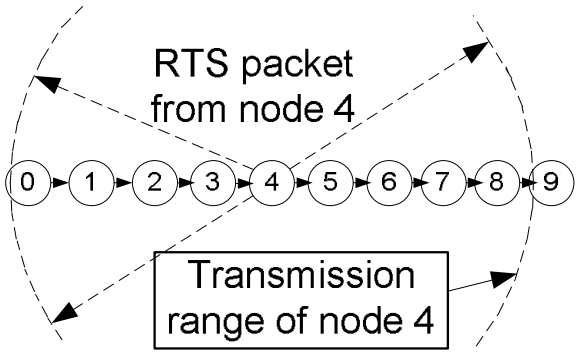


Figure 3.6: Impact of long transmission range of control packets on communication concurrency

each node to transmit to at most k active neighbors, where k is a PALM parameter, *the (maximum) number of active neighbors*. The details of power computation are presented below.

Let $N_{act,i}$ and $Neigh_{act,i}$ denote the number of i's active neighbors, and the set of such neighbors, respectively. Then, the set $Gain_{act,i}$ of *active gains* incident on i can be defined as

$$Gain_{act,i} \triangleq \{Gain(i, j) : j \in Neigh_{act,i}\}$$

We define l and $Gain_l$ as

$$l \triangleq \min(k, N_{act,i})$$

$$Gain_l \triangleq \text{the } l\text{-th largest gain in } Gain_{act,i}$$

Node i computes the transmission power Pow_{bcast} for broadcast packets as

$$Pow_{bcast} = \frac{RSS_0}{Gain_0}$$

and the power $Pow_{ctrl}(j)$ for control packets to node j as

$$Pow_{ctrl}(j) = \frac{RSS_0}{\min(Gain_l, Gain(i, j))}$$

Finally, if $N_{act,i} = 0$, node i transmits the packet at the full power.

The rationale behind the above procedure is as follows. If node i recognizes only a few active neighbors, its control packets do not need to reach further than those neighbors. As the network traffic rate and the number of active neighbors grow, node i must increase the power for control packets to prevent frequent signal collision. However, transmission power should not exceed the level at which each node has k active neighbors. It can be seen that this maximum k-neighbor condition is similar to that with some of existing topology controls [11, 76]; in other words, PALM implicitly performs such topology controls using the above power control procedure without periodic beaconing. This power reduction scheme enables PALM to have higher communication concurrency than BASIC.

3.3 Algorithms Used in PALM

This section presents formal algorithms for the power control and link maintenance techniques described above. These algorithms span from the routing layer to the MAC layer; for simplicity, route discovery procedures are not included. Thus, *sending up* implies packets are sent to the application layer; and *sending down*, that they are sent to the physical layer.

Each packet contains the following fields: transmitter's address *ta*; receiver's address *ra*; transmission power *pow*; final destination address *fdst*; accumulated energy *Eacc*; and (virtual) hop distance of the transmitter to the final destination *hops*. In addition, the CTS and ACK packets contain a parameter *gain* for the channel gain between source and desti-

nation. The RTR packets contain the redirector address *redirector*, and the energy saving *dE*. The physical layer informs the MAC layer of the received signal strength *RSS*. Finally, the routing table contains *locked time* T_{lock} fields as well as the next hop node and the hop distance to the final destination. Thus, the routing table has the format:

$$rTable(fdst_1) = (n_1, h_1, T_{lock,1}, T_{used,1})$$

$$rTable(fdst_2) = (n_2, h_2, T_{lock,2}, T_{used,2})$$

$$\vdots$$

The algorithms that perform route repair and power control functions are described next.

Send (Figure 3.7): Before sending down the packet, the appropriate transmission power is computed based on previously sensed channel gains.

Receive (Figure 3.8): Upon receiving a packet, the channel gain and the transmitter's

```
if pkt is a broadcast packet or a control packet then
    k_a ← the number of active gains ;
    if k_a == 0 then
        pkt.pow ← pow_max ;
        send down pkt ;
        return;
    end
    k_0 ← min(k_a, k) ;
    gain_0 ← k_0-th largest in active gains ;
    if pkt is a control packet then gain ← min(gain_0, gain(myAddr, pkt.ra)) ;
    else gain ← gain_0 ;                          // pkt is a broadcast packet
end
else
    gain ← gain(myAddr, pkt.ra) ;
    pkt.pow ← min(RSS_min/gain, pow_max) ;
end
send down pkt ;
```

Figure 3.7: Function Send(*pkt*)

hop distance are recorded first. Then, if the packet is destined for the node x itself, the accumulated energy is calculated according to (3.7). If the incoming packet is an RTR, then x modifies its routing table according to the redirection information contained in the RTR packet. If the node x is the final destination, it sends the packet to the upper layer; otherwise, it locks the link first, and then forwards the packet to the next hop. Finally, if the packet is an overheard one, x considers route redirection or route warping.

RepairRoute (Figure 3.9): If the link to the next hop is still locked, and the node x's hop distance is smaller than the source's hop distance, warping is not considered as it may form a loop. If 1) warping does not cause a loop, 2) warping reduces energy consumption,

```
gain(pkt.ta, myAddr) ← rss/pkt.pow ;
gain(pkt.ta, pkt.ra) ← pkt.gain ;
hops(pkt.ta) ← pkt.hops ;
if pkt.ra == myAddr then
    // I am the recipient
    if pkt is an RTR packet and pkt_ΔE > recent_ΔE then
        rTable(pkt.fdst).next ← pkt.redirector ;          // Modify my routing table
        rTable(pkt.fdst).T_lock ← current time ;
        recent_ΔE ← pkt.ΔE;
        return;
    end
    E_acc ← pkt.E_acc;
    pkt.E_acc ← pkt.E_acc + pkt.pow/rss ;
    if pkt.fdst == myAddr then send up pkt ;        // I am the final destination
    else
        // Forward the packet to the next hop node
        rt ← rTable(pkt.fdst) ;
        rTable(pkt.fdst).T_lock ← current time ;                    // Lock the link
        rTable(pkt.fdst).T_used ← current time ;
        pkt.ta ← myAddr ;
        pkt.ra ← rt.next ;
        Send( pkt ) ;
    end
end
else RepairRoute(pkt) ;                                  // I am overhearing
```

Figure 3.8: Function Receive(*pkt*)

and 3) no other RTR packets with larger energy saving have been overheard, then x sends to the source an RTR packet with the amount of energy saved. Route redirection is similar except for the way the RTR-sending node modifies its routing table.

$rt \leftarrow rTable(pkt.fdst)$;
if $rt.T_{used} + T_{unused} \leq$ *current time* ***or*** $rt.hops \geq hops(pkt.ra)$ **then**
    goto *REDIRECTION* ; `// A loop may appear; do not warp`
end
WARPING:
$E_1 \leftarrow E_{acc}$;
$E_2 \leftarrow pkt.E_{acc} + RSS_{\min}/gain(pkt.ta, myAddr)$;
$\Delta E \leftarrow E_1 - E_2$;
if $\alpha \cdot E_1 > E_2$ ***and*** *no overheard RTR with larger* ΔE **then**
    new (RTR) ; `// Prepare a new RTR packet`
    $rtr.\Delta E \leftarrow \Delta E$;
    $rtr.fdst \leftarrow pkt.fdst$;
    $rtr.ra \leftarrow pkt.ta$;
    $rtr.ta \leftarrow myAddr$;
    $rtr.redirector \leftarrow myAddr$;
    $rTable(pkt.fdst) \leftarrow (rt.next, rt.hops,$ current time$, rt.T_{used}$;
    Send(rtr) ; `// Send the RTR to the source`
    return;
end
REDIRECTION:
if $rt.T_{lock} + T_{unlock} >$ *current time* ***and*** $rt.hops \leq hops(pkt.ta)$ **then**
    return;
end
$E_1 \leftarrow RSS_{\min}/gain(pkt.ta, pkt.ra)$;
$E_2 \leftarrow RSS_{\min}/gain(pkt.ta, myAddr) + RSS_{\min}/gain(myAddr, pkt.ra)$;
$\Delta E \leftarrow E_1 - E_2$;
if $\alpha \cdot E_1 > E_2$ ***and*** *no overheard RTR with large* ΔE **then**
    new (rtr) ;
    $rtr.\Delta E \leftarrow \Delta E$;
    $rtr.fdst \leftarrow pkt.fdst$;
    $rtr.ra \leftarrow pkt.ta$;
    $rtr.ta \leftarrow myAddr$;
    $rtr.redirector \leftarrow myAddr$;
    $rTable(pkt.fdst) \leftarrow (pkt.ra, (hops(pkt.ta) + hops(pkt.ra))/2,$ current time$, rt.T_{used}$;
    Send(rtr) ;
end

Figure 3.9: Function RepairRoute(pkt)

3.4 Performance Evaluation

This section presents experimental results comparing PALM with a variety of other link-maintenance algorithms.

3.4.1 Simulation Environment

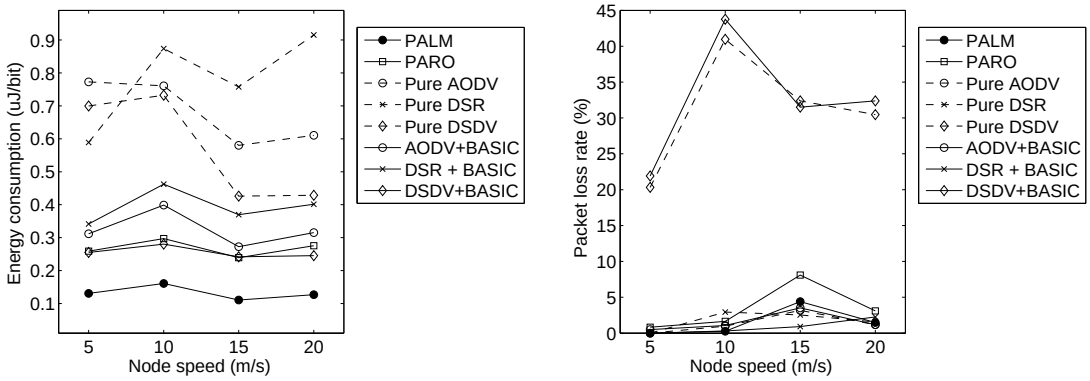
We used the *ns* simulator [92], and modified its AODV, MAC 802.11, and physical layer to implement PALM. A MANET model for the experiments was constructed as follows. The radio channel parameters are taken from [49]: the RSS threshold $RSS_{th} = 0.3652$ nW; the maximum transmission power $Pow_{\max} = 281.8$ mW corresponding to 250 m transmission range; and the channel bandwidth is 2 Mbps. We assume that each node i can take any value between 0 and $Pow_{\max}$ as its transmission power $Pow(i)$. The following PALM-specific parameters are also used: energy saving threshold $\alpha = 0.8$; safety factor $\beta = 1.5$; active neighbor window $T_{act} = 0.5$ s; usage expiration time $T_{unused} = 0.5$ s; and unlock time $T_{unlock} = 0.5$ s. Finally, the *two-ray ground model* [49] was adopted as the radio propagation model.

3.4.2 Simulation Results

First, we conducted simulations with varying node speed. The network consists of 35 mobile nodes randomly distributed on a 700×700 m^2 plane. Source and destination nodes are randomly chosen and a constant-bit-rate (CBR) source sends packets at 80 kbps. Packet size is 1000 bytes, and packet interval time is 0.1 s. The random waypoint model [15] was used with the following parameters: pause time = 0.0; minimum speed = maximum speed = v varying from 5 m/s to 20 m/s. Each simulation was run for 500 s. The maximum number of active neighbors is set to $k = 8$.

Figure 3.10 presents the simulation results with respect to node speed ranging from 5 m/s to 20 m/s. For a fair comparison, simulation results with AODV [71], DSR [47], and DSDV [70] with the BASIC power control scheme are included. We also implemented a variant of PARO with the following modifications. First, the assumption that all nodes are initially within each other's range is relaxed; AODV discovers the initial route, and PARO optimizes the route through route redirection. Second, prioritizing redirection requests by adjusting the backoff interval is not included.

Figure 3.10 shows the energy consumption per bit transmitted by the traffic source and packet loss rate of various algorithms including PALM. It can be seen that BASIC improves the power efficiency of DSR, AODV, and DSDV. In particular, the energy consumption of DSDV with BASIC is comparable to that of PARO. As DSDV continually discovers shortest paths, and the energy consumption for the data traffic is much larger than that for beaconing, the beaconing energy is compensated for by the high energy efficiency of newly discovered routes. However, its packet loss rate is much higher than other algorithms as the periodic beaconing of DSDV interferes with data traffic. Thus, it is clear that periodic beaconing can lead to inefficient channel resource usage. PALM's energy consumption



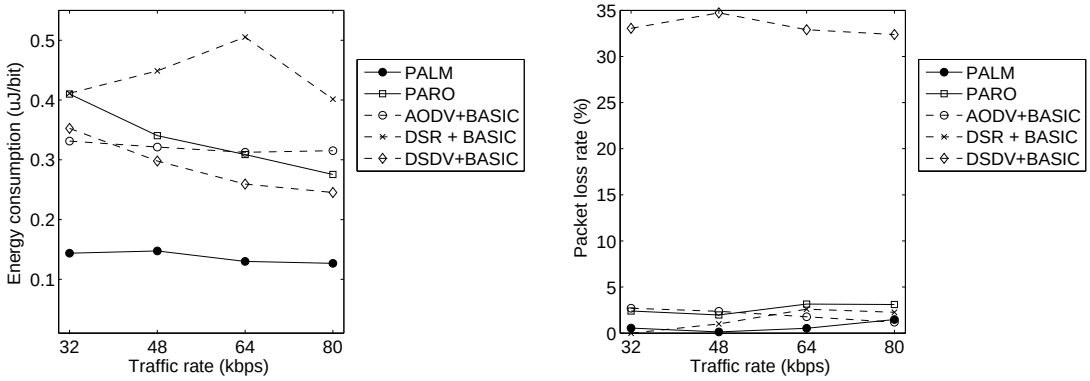
(a) Energy consumption per transmitted bit (b) Packet loss rate

Figure 3.10: Simulation results with varying node speed and a CBR of 80 kbps

is much lower than that of other algorithms including PARO, and its packet loss rate is comparable to others.

Next, we conducted simulations that varied the traffic rate from 32kbps to 80kbps; see Figure 3.11. The packet size is 1000 bytes, and the packet interval ranges from 0.1 s to 0.25 s. The node speed is set to 20 m/s. The energy consumption of DSDV decreases as the traffic rate grows because its beaconing energy is amortized by the increased data traffic. However, DSDV still results in high packet loss rate for the same reason as in the above experiments. On the other hand, PALM produces low packet loss rate while it still consumes much less energy than other algorithms.

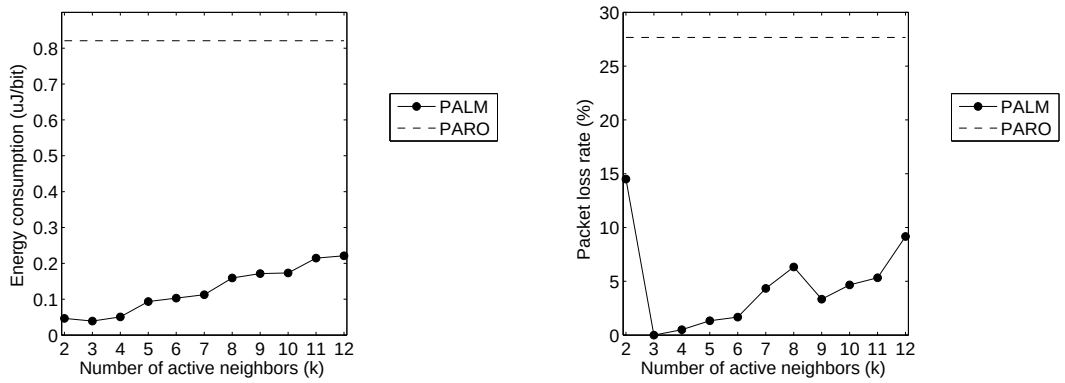
We also measured the impact of the maximum number k of active neighbors on the energy consumption and the packet loss rate, and compared it with PARO's energy consumption and loss rate; note that PARO incorporates the BASIC power control scheme, and expends minimum transmission power for DATA/ACK packets and maximum power for RTS/CTS. The simulation was done over a stationary network to isolate the effect of transmission power. PALM's operation with a stationary network is identical with that of PARO except for the transmission power for control and broadcast packets. To demon-



(a) Energy consumption per transmitted bit

(b) Packet loss rate

Figure 3.11: Simulation results with varying traffic rate and a node speed of 20 m/s



(a) Energy consumption per transmitted bit (b) Packet loss rate

Figure 3.12: Impact of the number k of active neighbors in a stationary network with 100 nodes

strate a congested condition with a denser topology and heavier traffic than the above experiments, we had 100 nodes on a 700×700 m^2 plane and 3 CBR sources each of which transmits packets at 32 kbps.

Figure 3.12(a) shows the energy consumption with respect to k. The energy consumption of PALM increases as k grows, but is still much lower than that of PARO, even for fairly large k. On the other hand, as illustrated by Figure 3.12(b), PALM's packet loss rate is much lower than that of PARO. Thus, it can be seen that transmitting control packets to only a few of the active neighbors is sufficient to prevent signal collision, and considerably improves communication concurrency.

As also observed in [49], using BASIC for power control is particularly undesirable with PARO for two reasons. First, as PARO discovers routes with short per-hop distances, most of the transmission energy might be consumed for control packets rather than for data packets, especially with dense networks. Second, as pointed out in Section 3.2.5, BASIC causes severe channel contention between adjacent transmitters resulting in a high packet loss rate. For the same reasons, limiting transmission power for control packets can significantly reduce energy consumption and channel contention, especially when the

network is dense.

3.5 Impact of MAC Parameters on Performance

Here we investigate how medium access control (MAC) parameters affect the performance of MANETs in which the transmission power for DATA packets is controlled to the minimum value. We investigate the impact of three factors in the MAC layer on the network performance: RTS/CTS handshake, transmission power control, and carrier-sense threshold. The RTS/CTS handshake is used to reduce signal collisions due to simultaneous transmissions by nodes located outside each other's transmission range, which is called the *hidden terminal problem* [10]. It has been recently observed that while the RTS/CTS handshake effectively reduces the hidden terminal problem in local area networks, its effectiveness is limited in ad hoc networks [98, 99, 101].

For the purpose of reducing communication power consumption, the BASIC power control method has long been considered, which adjusts the power level for DATA packets to the minimum necessary value, while the RTS/CTS packets are transmitted at the maximum level [49]. However, recent studies [51, 62, 82, 106] have shown that when the RTS/CTS handshake is used with BASIC, the overall network throughput may become worse than networks without transmission power control, because the RTS/CTS packets transmitted at the maximum power level can easily interfere with other DATA packets transmitted at reduced power levels. In addition, the carrier-sense threshold affects concurrency of network communication, and in consequence, the end-to-end network throughput also depends on the carrier-sense threshold [101].

In the next section, we first review how the RTS/CTS handshake and the carrier-sense threshold affect the performance of ad hoc networks in which all nodes have the same trans-

mission range. Then, we compare the results with the networks in which node transmission power is variable.

3.5.1 MAC with Homogeneous Transmission Power

Let us review how the RTS/CTS handshake and the carrier-sense threshold affect the network throughput in ad hoc networks with homogeneous transmission range. Assume that all nodes have the same transmission range r_{TX}. If a receiver is located within this transmission range, and there is no interfering node, it can successfully receive a message. We also assume the radio attenuation model presented in Section 1.2.3, i.e.,

$$RSS = \frac{Pow}{d^{\alpha}}$$

where Pow, RSS, and d denote the transmitted power, the received signal strength at the receiver, and the distance between the transmitter and the receiver, respectively. The carrier-sense threshold is set to a predetermined value such that overhearing nodes within d_{CS} from the transmitter can sense the carrier. The minimum signal-to-interference ratio (SIR) for a successful signal capture is z_0. Let TR_A and CS_A, denote the transmission and the

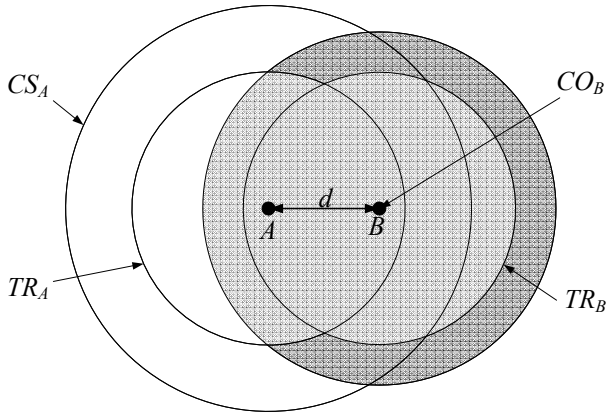


Figure 3.13: Transmission (TR), carrier-sense (CS), and collision (CO) regions

carrier-sense regions of node A, the circular regions centered at A with radii r_{TX} and d_{CS}, respectively. Let CO_B denote the collision region centered at B. The radius d_{CO} of CO_B is given by

$$d_{CO} = z_0^{1/\alpha} \cdot d \tag{3.8}$$

Figure 3.13 illustrates the above regions.

In [99], the effectiveness η of RTS/CTS is defined as

$$\eta \triangleq \frac{Area\Big(CO_B \cap (TR_A \cup TR_B)\Big)}{Area\Big(CO_B\Big)} \tag{3.9}$$

The numerator of (3.9) corresponds to the lightly shaded area in Figure 3.13. Nodes in this area can receive RTS/CTS from A and B, while they might interfere with the communication between A and B if they do not receive RTS/CTS. On the other hand, nodes in the darkly shaded area cannot receive RTS/CTS, and they may cause collisions. As can be seen from (3.8), d_{CO} is determined not only by z_0, but also by the distance d between A and B. As d increases, the RSS at B decreases, and so signal reception by B becomes more susceptible to signal collision. Xu *et al.* [99] have shown that the effectiveness of RTS/CTS decreases as d exceeds $1/z_0^{1/\alpha}$.

Assuming that no RTS/CTS handshake is used, according to the analysis in [101], the maximum end-to-end throughput T_e is given by

$$T_e = \frac{\sqrt{6}Lbd}{(\sqrt[\alpha]{6z_0})^2} \times \frac{1}{d} \quad \text{when} \quad d_{CS} < \sqrt[\alpha]{6z_0}d \tag{3.10}$$

$$T_e = \frac{\sqrt{6}Lb}{d_{CS}^2} \times d \quad \text{when} \quad d_{CS} \geq \sqrt[\alpha]{6z_0}d \tag{3.11}$$

Figure 3.14 presents the maximum throughput T_e computed by (3.10) and (3.11). It shows that for the given carrier-sense distance d_{CS}, there exists an optimal communication dis-

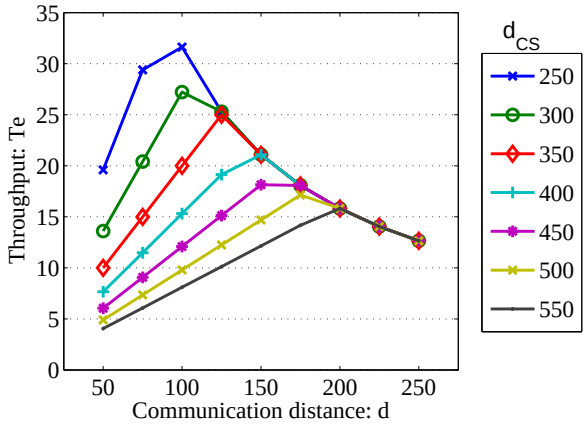


Figure 3.14: Maximum end-to-end throughput T_e as a function of d

tance $d = 1/\sqrt[\alpha]{6z_0} \cdot d_{CS}$ that maximizes the network throughput T_e.

Note that although Figure 3.14 provides an optimality condition for T_e, it is difficult to take advantage of the optimality relation $d = d_{CS}/\sqrt[\alpha]{6z_0}$ because 1) the distance d is determined by the node locations, the routing protocol used, and the network topology, and 2) it is difficult to adjust d_{CS} adaptively for a given d [98, 101]. In the following sections, we will show how to maintain the above optimality relation by using a distributed power control scheme.

3.5.2 CSMA with Variable Transmission Power

Now we investigate the impact of the carrier-sense threshold on MANET performance, where the carrier-sense multiple access (CSMA) method is used. Let us assume that RTS/CTS handshake is not used, and measure the network throughput and energy efficiency. We assume that the network topology is managed by PALM, as described earlier. We also assume that node transmission power is controlled according to the BASIC power control scheme described in Section 3.2.1 such that the received signal strength (RSS) of DATA and ACK packets at the receiver are always maintained at RSS_0. Suppose a trans-

mitter TX and a receiver RX are separated by distance d. The transmitter adjusts its power to $Pow = RSS_0 \cdot d^{\alpha}$ such that the received signal strength at RX is RSS_0. Other nodes within d_{CS} from TX do not transmit because they sense the carrier from TX.

In order to maximize the end-to-end throughput T_e, we need to maximize

$$T_e = b \cdot N_{success}/N_{hops}$$

where b, $N_{success}$, and N_{hops} denote the link bandwidth, the number of concurrent successful transmissions, and the average number of hops from sources to destinations, respectively. Unlike the homogeneous transmission range case in Section 3.5.1, the network topology of interest is already produced by PALM, and we can take N_{hops} as constant. Thus, it is sufficient to focus on maximizing

$$N_{success} = N_{TX} \cdot p_{success} \tag{3.12}$$

where N_{TX} and $p_{success}$ denote the number of concurrent transmissions (including both successful and unsuccessful ones), and the probability of successful transmission, respectively. For this reason, we can use $N_{success}$ as the metric for end-to-end network throughput instead of T_e.

Let us compute the optimal carrier-sense distance d_{CS} via an analysis similar to that in [101]. In order to obtain the maximum throughput, two conditions need to be satisfied: 1) the SIR should be greater than z_0, and 2) the inter-transmitter distance should be minimized. Consider the node configuration in Figure 3.15, which corresponds to the densest arrangement of transmitters. Transmitter TX_0 is sending data to RX_0, and transmissions from TX_1 through TX_6 are also reaching RX_0. For the communication between TX_0 and

RX_0 to succeed, the SIR at RX_0 should be greater than z_0, i.e.,

$$\frac{d^{-\alpha}}{2(D-d)^{-\alpha} + (D-d/2)^{-\alpha} + D^{-\alpha} + (D+d/2)^{-\alpha} + (D+d)^{-\alpha}} > z_0 \tag{3.13}$$

Define $D_{\min}$ as the minimum inter-transmitter distance D that satisfies (3.13). The densest transmitter arrangement occurs when $D = D_{\min}$. We adopt the simplifying assumption in [101] that TX_1 through TX_6 are separated from RX_0 by the same distance $D_{\min}$. By solving

$$\frac{d^{-\alpha}}{6{D_{\min}}^{-\alpha}} = z_0$$

we obtain $D_{\min} = \sqrt[\alpha]{6z_0} \cdot d$.

When CSMA is used, the inter-transmitter distance D depends on the carrier-sense distance d_{CS}. Since a node cannot start transmission when its distance from the nearest transmitting node is less than d_{CS}, the inter-transmitter distance D is always greater than or

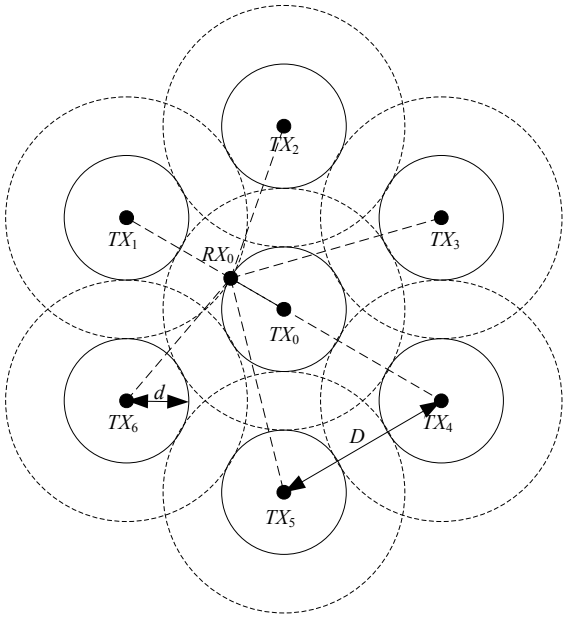


Figure 3.15: Densest arrangement of transmitters

equal to d_{CS}, i.e., $D \geq d_{CS}$. In order to demonstrate the densest transmitter arrangement, we make an additional simplifying assumption that $D \cong d_{CS}$. If d_{CS} grows greater than $D_{\min}$, the inter-transmitter distance D also becomes greater $D_{\min}$, and the number N_{TX} in (3.12) will be reduced. Conversely, if d_{CS} is less than $D_{\min}$, then the interference between different transmitters will reduce $p_{success}$. Thus, it is reasonable to assert that the maximum T_e occurs when

$$d_{CS} \cong D_{\min} = \sqrt[\alpha]{6z_0} \cdot d \tag{3.14}$$

Next, we present analysis and simulation results that measure the impact of the CS threshold on the network throughput and energy consumption. The network structure of interest is produced by PALM, and the transmission power is adjusted such that DATA packets are transmitted at the minimum necessary power. Unlike min-hop routing protocols [51] such as DSR and AODV, which select the routes with the minimum hop distances, PALM produces routes that have the minimum energy cost, but may have longer hop distances. To simplify the simulations, we approximate the topology generated by PALM using the shortest-path graph, in which the edge weights are given by the corresponding power cost. In addition, we assume that ACK packets do not collide with other messages, and the transmitter can tell whether its last DATA packets are successfully received or not.

Since the communication distance and the transmission power are variable, we cannot fix the carrier-sense distance d_{CS}. Instead, we fix the carrier-sense ratio r_{CS}, defined as the ratio between the carrier-sense distance d_{CS} and the communication distance d.

$$r_{CS} \triangleq \frac{d_{CS}}{d}$$

We can set r_{CS} to an intended value by tuning the carrier-sense threshold CS_{th} using

$$r_{CS} = \left(\frac{RSS_0}{CS_{th}}\right)^{1/\alpha} \tag{3.15}$$

We now compute the total energy consumption taking retransmission due to communication failure into account. When a data transmission with the cost E has success probability $p_{success}$, the expected total cost is given by

$$\mathrm{E}[E_{total}] = \sum_{k=1}^{\infty}(1 - p_{success})^{k-1} p_{success} \cdot k \cdot E = E/p_{success}$$

Figure 3.16 presents the pseudocode of the simulation method for estimating the network throughput and total energy consumption for a given CS ratio. If and only if a node is outside the carrier-sense range of all nodes that are already transmitting, it is allowed to transmit. The algorithm incrementally finds such a node in a greedy manner, and turns on its transmission. By repeating this procedure, we can compute the number $N_{success}$ of concurrent successful transmissions and the total energy consumption P_{total}. Note that this algorithm counts the *average* number of concurrent successful transmissions instead of the *maximum*.

We adopt the following parameters for the simulation. One thousand nodes ($N = 1000$) are randomly distributed on a 1×1 plane. The minimum RSS is set to $RSS_0 = 1$, and the radio attenuation exponent is set to $\alpha = 4$. The receiver's capture parameter is set to $z_0 = 10$. In order to eliminate the border effect [9], we assume the toroidal space described in Section 1.3.1. The algorithm of Figure 3.16 is repeated 100 times to compute the average values.

Figure 3.17(a) shows the simulation result for the number N_{TX} of concurrent transmissions. As the carrier sense ratio r_{CS} increases, the increasing carrier-sense distance prevents

Input: $\{V\}$; the set of nodes
Output: N_{TX}, $N_{success}$, $p_{success}$, E_{total}; the number of transmitting nodes, the number of successful transmissions, the success probability, and the total energy consumption

$S \leftarrow \{V\}$;
$E \leftarrow 0$;
Find the shortest path from each node to the sink $s \in S$;
foreach $u \in S \setminus s$ **do**
    $Pow[u] \leftarrow RSS_0 \cdot Dist\left(u, \Pi(u)\right)^{\alpha}$;
    $E \leftarrow E + Pow[u]$;
end
$N_{TX} \leftarrow 0$;
$TX \leftarrow \emptyset$;
while $S \neq \emptyset$ **do**
    Randomly pick one node u from S ;
    $S \leftarrow S \setminus \{u\}$;
    $TX \leftarrow TX \cup \{u\}$;
    $N_{TX} \leftarrow N_{TX} + 1$;
    foreach $v \in S$ **do**
        if $Pow(u)/Dist\left(u, v\right)^{\alpha} > CS_{th}$ **then** $S \leftarrow S \setminus v$;
    end
end
$N_{success} \leftarrow 0$;
foreach $u \in TX$ **do**
    $interference \leftarrow 0$;
    foreach $v \in TX \setminus \{u\}$ **do**
        $interference \leftarrow interference + Pow[v]/Dist\left(\Pi(u), v\right)^{\alpha}$
    end
    if $RSS_0/interference > z_0$ **then**
        $N_{success} \leftarrow N_{success} + 1$;
    end
end
$p_{success} \leftarrow N_{success}/\left|TX\right|$;
$E_{total} \leftarrow E/p_{success}$;

Figure 3.16: Pseudocode of algorithm for measuring the impact of the CS threshold

concurrent transmissions by neighboring nodes, and thus N_{TX} decreases. Figure 3.17(b) shows the success probability. As r_{CS} increases, the success probability also increases because the increased carrier-sense distance reduces the possibility of signal collision.

Figure 3.17(c) illustrates how the network throughput changes with respect to r_{CS}. The

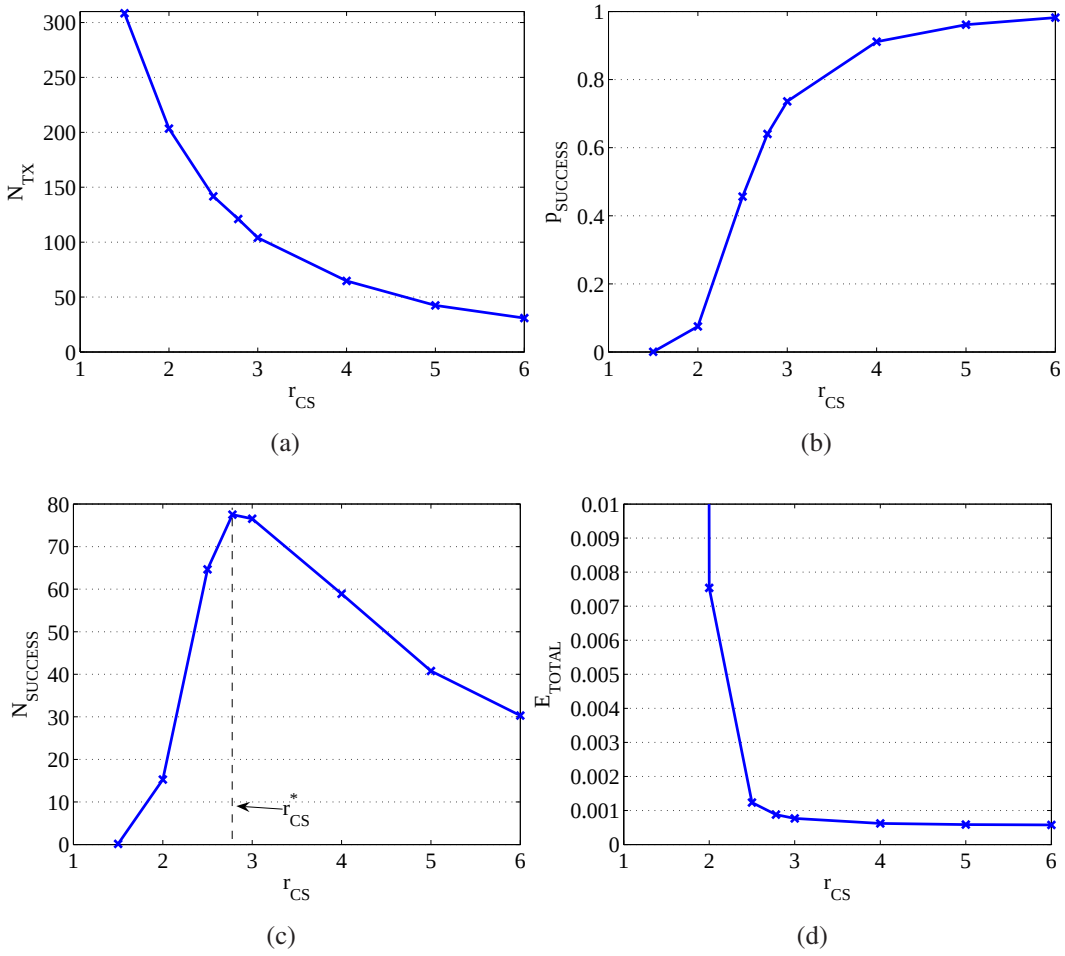


Figure 3.17: Simulation results for CSMA: (a) number of transmissions, (b) success probability, (c) network throughput, and (d) total energy consumption

network throughput is measured by the number $N_{success}$ of successful concurrent transmissions which is computed by (3.12). It can be seen that as r_{CS} increases, $N_{success}$ reaches the maximum point, and then gradually decreases. We define

$$r_{CS}^{*} \triangleq D_{\min}/d = (6z_0)^{1/\alpha} \tag{3.16}$$

The maximum of $N_{success}$ occurs when $r_{CS} \cong r_{CS}^{*}$.

Figure 3.17(d) shows that the total energy consumption E_{total} decreases as r_{CS} increases,

because the increased carrier-sense distance tends to reduce the number of retransmissions. Note that E_{total} drops quickly when r_{CS} is small, and converges to its minimum as r_{CS} grows beyond r^*_{CS}. When $r_{CS} = r^*_{CS}$, the network throughput reaches its maximum value, while the total energy consumption is only marginally greater than its lower bound.

In order to validate the optimality condition $r_{CS} = r^*_{CS}$, simulation results with varying z_0 and α are presented in Figure 3.18. As Figure 3.18(a) shows, the radio attenuation exponent α is reduced to 2. Consequently, the spatial channel reuse is degraded, so the throughput is lower than in Figure 3.17(c). In Figure 3.18(b), the receiver's capture parameter is reduced to 1, i.e., perfect capture is assumed. Consequently, the throughput is improved compared to Figure 3.17(c). In both cases, the maximum throughput occurs when $r_{CS} \cong r^*_{CS}$, and it can be seen that r^*_{CS} both maximizes network throughput, and achieves high energy efficiency.

When the transmission range is fixed as mentioned in Section 3.5.1, it is difficult to maintain an appropriate ratio between the communication distance and the carrier-sense distance. However, when transmission power is controlled by the BASIC scheme, we can

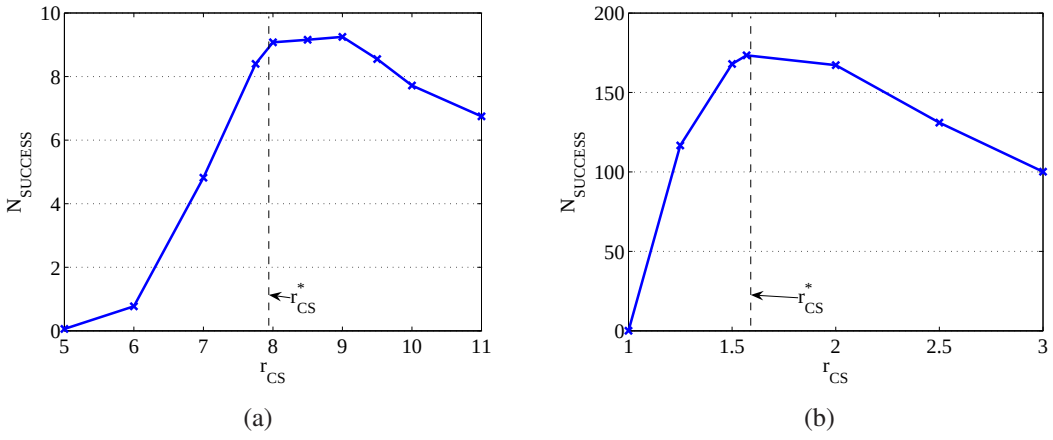


Figure 3.18: The number of concurrent successful transmissions $N_{success}$ as a function of the carrier-sense ratio r_{CS}: (a) $z_0 = 10$, $\alpha = 2$; and (b) $z_0 = 1$, $\alpha = 4$

maintain the optimal carrier-sense ratio $r_{CS} = r^*_{CS}$ by tuning the carrier-sense threshold CS_{th} such that $CS_{th} = RSS_0 / {r^*_{CS}}^{\alpha}$, or equivalently,

$$CS_{th} = \frac{RSS_0}{6 \cdot z_0} \tag{3.17}$$

Since (3.17) depends only on the hardware parameters RSS_0 and z_0, once CS_{th} is set to its optimal value, the above optimality condition can be satisfied regardless of the communication distance d, the transmission power Pow, and the radio attenuation exponent α.

3.5.3 RTS/CTS Handshake with Variable Transmission Power

Here we investigate the BASIC scheme with the RTS/CTS handshake enabled and two variants of BASIC. Then, we evaluate their impact on the network performance. According to the BASIC scheme, RTS/CTS packets are transmitted at the maximum power, and DATA/ACK at the minimum necessary power. Again, we do not consider the effect of ACK packets. We assume that nodes that overhear or sense the carrier of RTS refrain from their own transmission for a sufficiently long time; in consequence, nodes within the CS range of RTS never collide with the ongoing transmission. This assumption emulates the PCM method [49], according to which the DATA packets are transmitted at the minimum power, but transmitters sporadically raise their DATA power to the maximum level for brief periods. We also make a similar assumption about CTS: nodes in the CS range of CTS do not transmit while the ongoing communication continues.

Second, we consider a variant of BASIC, that we call *short-RTS*, in which RTS packets are transmitted at the minimum necessary power. The goal of the RTS/CTS handshake is to protect the on-going communication by preventing transmission of neighboring nodes.

Figure 3.13 shows that the collision region is centered at receivers, and it can be seen that we need to protect the area around receivers rather than transmitters. Thus, it seems reasonable to reduce the power for RTS. Lastly, we consider another variant of BASIC, *short-RTS/CTS*, in which both RTS/CTS packets are transmitted at the minimum necessary power.

In order to measure the impact of the above BASIC variants on the network performance, we use an algorithm similar to that of Figure 3.16 except for the RTS/CTS handshake parts. We set the maximum transmission power $Pow_{\max}$ to $Pow_{\max} = RSS_0 \cdot {d_{\max}}^{\alpha}$, where $d_{\max}$ denotes the largest separation of adjacent nodes in the given shortest-path tree. In other words, $Pow_{\max}$ is the minimum power that can maintain the connectivity of the shortest-path tree. The number of bits of control packets is assumed to be 5% of that of DATA packets, and control packets use the same radio channel and coding as DATA packets. The other simulation parameters are the same as in Section 3.5.2.

Figure 3.19 shows simulation results for the three versions of BASIC. Simulation results for CSMA in Figure 3.17 are also shown here for comparison. Figure 3.19(a) shows that the BASIC variants allow far fewer concurrent transmissions than CSMA due to their RTS/CTS packets. As expected, the short-RTS/CTS variant allows more transmissions than the short-RTS or BASIC schemes. Figure 3.19(b) shows that the short-RTS scheme has a higher success probability than the others when the carrier-sense ratio r_{CS} is small, and that both short-RTS and BASIC have a nearly perfect success rate at $r_{CS} = r_{CS}^{*} \cong 2.78$. Figure 3.19(c) shows that the all variants of BASIC have lower throughput than CSMA. Though short-RTS and BASIC have higher success probabilities, they result in lower throughput because they overly suppress concurrent transmissions, as their low N_{TX} values suggest. Figure 3.19(d) shows that CSMA and short-RTS consume less energy than the other schemes at $r_{CS} = r_{CS}^{*}$.

Figure 3.20 compares the network throughput and total energy consumption of CSMA

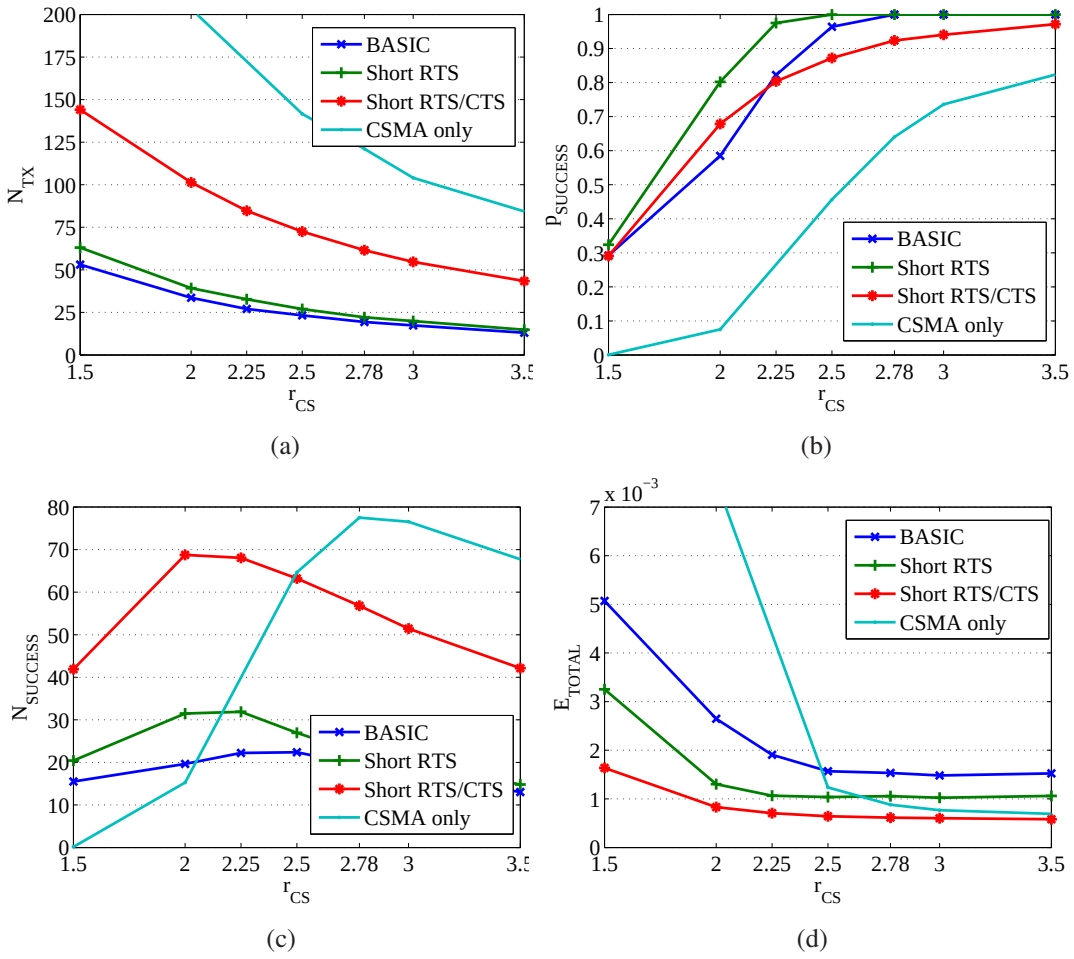


Figure 3.19: Simulation results for RTS/CTS handshake variants; (a) number of transmissions, (b) success probability, (c) network throughput, and (d) total energy consumption

with the BASIC variants. With CSMA only, the maximum throughput $N_{success} = 77.5$ can be achieved at total energy cost $E_{total} = 0.88 \times 10^{-3}$. With the BASIC scheme, the maximum throughput is $N_{success} = 22.4$, only 29% of that of CSMA, and the corresponding total energy consumption amounts to $E_{total} = 1.6 \times 10^{-3}$. It can be seen that for the given energy consumption, CSMA has much higher throughput than any other BASIC variant, and that the BASIC power control scheme has much worse energy efficiency, even lower than the short-RTS/CTS scheme.

Now let us relate the above results to the PALM algorithm. When there exists no signif-

icant radio obstruction except for the path loss with respect to the communication distance, tuning the carrier-sense threshold CS_{th} according to (3.17) maximizes the network capacity. However, in cases where the radio signal from the transmitter cannot reach other nodes due to a radio obstacle, the CSMA method without the RTS/CTS handshake may not work properly. Consider the example in Figure 3.21. Node A is transmitting to B, and there is a radio obstacle between A and C. Without the obstacle, the signal carrier from A would be sensed by C. However, due to the signal obstruction, node C cannot recognize A's signal, and may start its own transmission, resulting in signal collision at node B. In this case, the use of the RTS/CTS handshake that allows node B to transmit CTS to its neighbors prevents

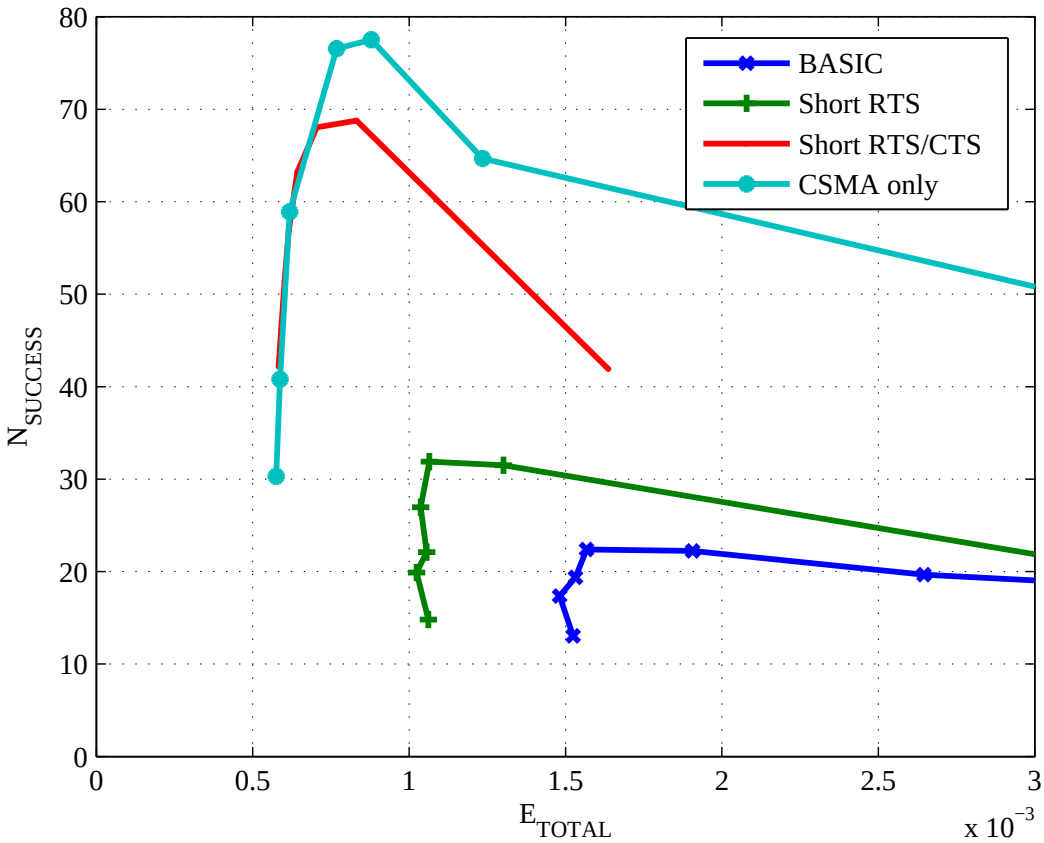


Figure 3.20: Throughput versus energy consumption for different MAC schemes

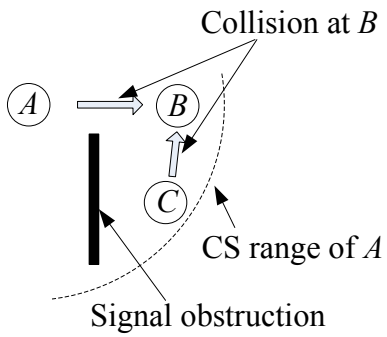


Figure 3.21: Example of CSMA failure

the simultaneous transmission by A and C, and may fix the collision problem. The trade-off relation illustrated in Figure 3.20 suggests that even when the RTS/CTS handshake is adopted, reducing the transmission power for RTS and CTS packets increases the network capacity while maintaining a high energy efficiency. Thus, when radio obstacles are present in the network space, the use of the RTS/CTS handshake with a reduced transmission power could produce a high network capacity with reasonable energy efficiency.

3.6 Summary

We have proposed PALM as a practical link maintenance scheme for MANETs that continuously discovers power-efficient routes without periodic beaconing, and adaptively adjusts transmission power at the same time. Our experiments show that PALM outperforms existing algorithms including PARO, and that PALM's power control is much more efficient than the BASIC power control approach.

The high efficiency of PALM is due to: 1) improved redirection with loop avoidance; 2) the route warping technique; and 3) transmission power control that mitigates the severe channel contention problem of BASIC. In particular, the transmission power limiting technique significantly improves energy efficiency and communication concurrency in dense

networks.

We have further investigated the impact of the carrier-sense distance on the network throughput and energy efficiency, and confirmed through simulations that the optimal CS threshold is given by $CS_{th} = RSS_0/(6 \cdot z_0)$. Since this condition does not depend on the communication distance or the radio attenuation exponent α, once CS_{th} is set to the optimal value, even when the network nodes continuously adjust the transmission power, and the radio channel property changes, the optimality condition can always be satisfied.

We have shown that use of the RTS/CTS handshake with power control protocols such as PALM may adversely affect the network performance. The simulation results with variants of the BASIC power control scheme confirm that the use of maximum power for control packets can reduce the network throughput by 70%, and the total energy consumption with retransmission taken into account is much higher than that with CSMA only. Therefore, when the transmission power for DATA packets in ad hoc networks is controlled to the minimum necessary level through the channel gain feedback between nodes, the RTS/CTS packets with the maximum transmission power may worsen signal collision and increase energy consumption. We also have shown that a high network capacity can be obtained by reducing the transmission power for control packets, and proposed a means to determine the appropriate transmission power for control packets in a distributed manner.

CHAPTER 4

Node Placement Optimization

Energy conservation is a key issue in ad hoc wireless network operation. Placing additional nodes, variously termed base stations, clusterheads, relays, etc., at appropriate locations can substantially lower the power requirements of communicating nodes, thereby reducing overall energy needs. This chapter investigates node placement for energy-constrained wireless networks, and presents node placement algorithms that aim to minimize total energy consumption. We first consider mobile base station (BS) placement problems for hierarchical wireless networks, and develop efficient heuristic solutions for them. We model the placement of multiple BSs as a clustering optimization problem in which BSs and user nodes are treated as clusterheads and cluster members, respectively. We also devise a heuristic that discovers the central area of a multi-hop network, and solves the BS placement problem with multi-hop connectivity. Our simulation results confirm that our methods reduce the energy consumption of wireless networks by up to 55% compared with grid networks. Using the distributed link maintenance algorithm presented in Chapter 3, we devise a distributed relay placement algorithm for the flat network structure that discovers energy-efficient routes, while maintaining connectivity in the presence of radio obstruction. Simulation results show that the power consumption of the distributed implementation is

greater than that of an existing centralized algorithm by only 25%.

4.1 Introduction

Since a MANET requires no pre-installed infrastructure, it is most suitable for such applications as an emergency communications network in a disaster area, a vehicular network, or a sensor network for environmental monitoring. In many MANET applications, the network consists of a large number of small wireless transceivers powered by batteries. It is often very hard to replace their energy sources, so energy efficiency is one of the most important criteria for successful network operation.

Recently, considerable attention has been paid to utilizing random, predictable, or controlled node mobility for conserving communication energy and improving network performance. Zhao *et al.* [104] propose a message-ferrying approach that controls node trajectory, and enables data delivery in sparse networks. Goldenberg *et al.* [29] describe a self-adaptive mobility control method that improves communication performance in a distributed manner. Li and Cassandras devise an iterative node placement scheme [57] that places network nodes at near-optimal locations between sensor nodes and the sink node in a sensor network.

As the unmanned aerial vehicle (UAV) [37], distributed robotics [83], and sensor network technologies advance, constructing MANETs that have mobile nodes with wireless communication capability is becoming more practical. Shi and Hou [81] propose a base station (BS) placement scheme that computes the optimal location of a single mobile BS that is an unmanned autonomous vehicle. They show that when the BS placement problem is constrained, i.e., the candidate BS locations are given, the joint computation of the optimal BS location and flow routes for maximizing the network lifetime can be modeled

as a linear programming problem, and thus can be solved in polynomial time. Then, they devise an unconstrained BS placement algorithm that produces provably $(1 - \epsilon)$-optimal solutions.

Ad hoc networks can be categorized to two classes, hierarchial and flat; Figure 4.1 illustrates both classes. In hierarchical networks, as in Figure 4.1(a), nodes are grouped to clusters, and a BS node or *clusterhead* in each cluster takes responsibility for inter-cluster communication. In flat networks, as in Figure 4.1(b), all nodes are assumed to have the same communication capability, and data from the source user node are delivered to the destination via intermediate user nodes or relays.

We first consider a two-layered hierarchical network structure consisting of user nodes and mobile BSs, which can be moved to any location. We further assume that the BSs have a separate wireless channel for communication between them, and they have sufficiently large energy sources. Consider the following example. Suppose that there are N sensor nodes with wireless transceivers in a forest. We dispatch UAVs with wireless communication capability to assist network communication between the sensors. Where should we place the UAVs to minimize energy consumption of the sensor devices?

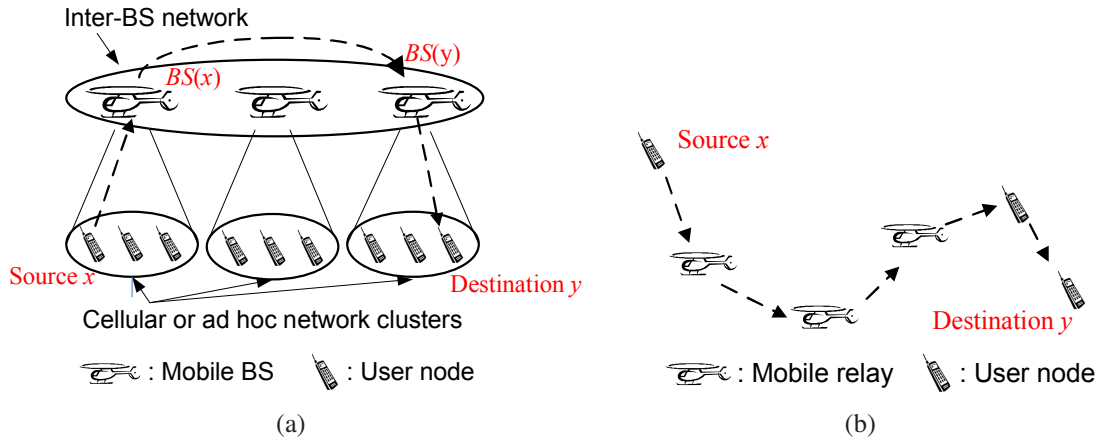


Figure 4.1: Hierarchical and flat network structures: (a) a hierarchical network with user nodes and mobile BSs, and (b) a flat network with user nodes and mobile relays.

We model this problem as a clustering problem composed of unconstrained convex optimization subproblems. We treat BSs as clusterheads, and aim to minimize the energy consumption of *uplink* communication, i.e., communication from user nodes to BSs. As the clustering problem is NP-hard, we develop an efficient heuristic method based on the K-means algorithm [58, 80], which is widely used for determining clusters that minimize the mean squared distance from points to the nearest cluster-means. Simulation results show that our BS placement algorithm produces near-optimal solutions with high probability.

Then, we investigate the relay placement problem for the flat network structure by modeling it as an analogous mechanical system. Our goal in relay placement is to place mobile relays at appropriate locations, and minimize the total power consumption, while maintaining network connectivity. By emulating the artificial forces exerted on mobile relays, and utilizing the PALM algorithm presented in Chapter 3, we solve the relay placement problem in the presence of radio obstruction in a distributed manner.

The rest of the chapter is organized as follows. Section 4.2.1 states the assumptions made about the radio model and the network structure. Section 4.2 presents the proposed BS placement algorithms for hierarchical network structures, and Section 4.3 gives simulation results for these algorithms. Section 4.4 presents the distributed relay placement algorithm for flat wireless networks, and Section 4.5 provides the simulation results. Finally, Section 4.6 summarizes the chapter.

4.2 Base Station Placement Optimization

We investigate base station placement (BSP) optimization for hierarchical networks considering four different network structures. We start from investigating the simplest network structure, and then develop BSP algorithms for more complicated ones. We first

Network structure	Number of base stations	Route connectivity
1-SH	1	Single-hop
K-SH	$K > 1$	Single-hop
1-MH	1	Multi-hop
K-MH	$K > 1$	Multi-hop

Table 4.1: Four different network structures for BSP problems.

consider the network structure with a single BS and single-hop links only (1-SH). Then, we extend it to multiple BSs with single-hop links (K-SH). Next, we extend the 1-SH problem to the multi-hop links case (K-SH). Finally, we solve the multi-BS multi-hop links problem (K-MH). These network structures are summarized in Table 4.1.

4.2.1 Communication and Network Model

Here we describe our assumptions concerning the wireless channel and network traffic. For the node placement problem of the hierarchical networks, we again adopt the radio model presented in Section 1.2.3. The channel gain $Gain(TX, RX)$ from a transmitter TX to a receiver RX is given by

$$Gain(TX, RX) = \frac{1}{k_1 \cdot Dist(TX, RX)^{\alpha} + k_2} \tag{4.1}$$

In other words, the channel gain is solely determined by the distance between the transmitter and the receiver.

We further assume that nodes expend the minimum power for transmission necessary for successful communication by using the BASIC power control scheme [21, 49] described

in Section 3.1. The power cost of the transmitter TX then becomes

$$Cost(TX, RX) = \frac{RSS_{\min}}{Gain(TX, RX)}$$

where $RSS_{\min}$ denotes the receiver's sensitivity, i.e., the minimum received signal strength (RSS) needed for successful communication. Assuming $RSS_{\min} = 1$, we get

$$Cost(TX, RX) = k_1 \cdot Dist(TX, RX)^{\alpha} + k_2 \tag{4.2}$$

We assume that BSs have full information about the locations of all nodes. This can be made possible by letting the BSs initially hover around the nodes for a while, and determine node locations using appropriate localization algorithms. We also allow source and destination nodes to be randomly chosen. Unlike sensor networks in which most nodes send data to a single sink, data can be generated at any node, and can be delivered to any other node, and the source-destination pair can frequently change over time.

We assume the following network architecture, which resembles that of the near-term digital radio (NTDR) network [72, 103]. In Figure 4.1(a), in order for a source node x to send data to a destination y, node x either directly transmits the data to the nearest BS $BS(x)$, or sends it over the shortest multi-hop path to $BS(x)$. Then, $BS(x)$ sends the data to the BS $BS(y)$ of the destination y through a separate communication channel. Eventually, $BS(y)$ sends the data to the final destination y. In this structure, the inter-BS network provides the network backbone, and the clusterhead serves as the gateway.

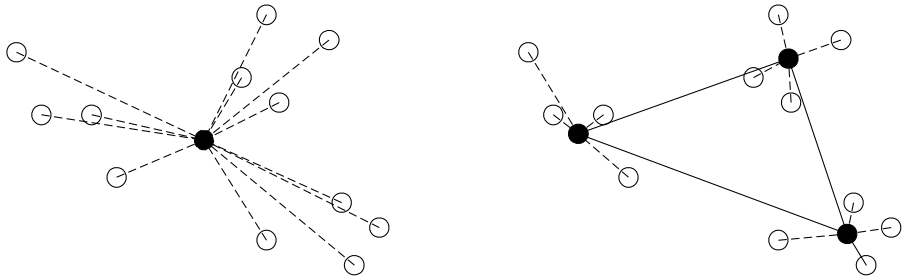
4.2.2 Single Base Station with Single-Hop Links (1-SH)

First, consider the case where user nodes, "nodes" for short, directly communicate via a single BS, as in Figure 4.2(a). Suppose that N nodes are placed at locations denoted by vectors $\vec{p}_i$ $(1 \leq i \leq N)$. Throughout this chapter, the location vector of node p_i is written as $\vec{p}_i$ consistently. The nodes directly communicate with a BS located at $\vec{x}$, and we want to minimize the uplink communication energy consumed by the nodes by controlling the location $\vec{x}$ of the BS. The BSP problem is then: What is the optimal location $\vec{x}^*$?

Let $\tau(p_i)$ denote the transmission time of node p_i, which is the number of bits to be sent by p_i divided by the bandwidth (bit/s) of the transceiver. For simplicity, assume that all nodes have the same bandwidth. Note that the number of bits is measured at the application layer, and thus the transmission time for forwarding other node data is not accounted for; the definition of the transmission time τ will be clarified in Section 4.2.5. As we are now dealing with single-hop connection only, we can tentatively assume that $\tau(p_i)$ is equal to the total transmission time of p_i.

Then, according to the power consumption model in Section 4.2.1, each node p_i consumes

$$E(p_i) = \tau(p_i)(k_1\|\vec{x} - \vec{p}_i\|^\alpha + k_2)$$



(a) (b)

Figure 4.2: Network structure with single-hop communication only: (a) user nodes (white) directly communicate with the single BS (black); (b) user nodes communicate with their nearest BSs.

where $\|\cdot\|$ denotes the Euclidian norm, and the total energy consumption by all user nodes is given by

$$\sum_{i=1}^{N} \tau(p_i)(k_1\|\vec{x} - \vec{p}_i\|^{\alpha} + k_2) \tag{4.3}$$

Thus, the 1-SH BSP problem becomes the following optimization problem.

$$\text{Minimize} \quad \sum_{i=1}^{N} \tau(p_i)(k_1\|\vec{x} - \vec{p}_i\|^{\alpha} + k_2)$$

where BS location $\vec{x}$ is the optimization variable. This is an unconstrained convex optimization problem that can be efficiently solved with existing convex optimization techniques. Proof of its being convex is given next.

4.2.3 Convexity of Cost Function

Here we prove that the following function $f(\vec{x})$ is convex.

$$f(\vec{x}) \triangleq \sum_{i=1}^{n} \tau_i(k_1\|\vec{x} - \vec{p}_i\|^{\alpha} + k_2) \qquad (\alpha \geq 1,\ \tau_i \geq 0,\ k_1 \geq 0,\ k_2 \geq 0)$$

Define $f_i(\vec{x})$ as

$$f_i(\vec{x}) \triangleq \|\vec{x} - \vec{p}_i\|^{\alpha}$$

and consider convexity of $f_i(\vec{x})$ first. According to the definition of convexity, $f_i(\vec{x})$ is convex if and only if the following inequality holds for all $\vec{x}$, $\vec{y}$, and $0 \leq t \leq 1$.

$$tf_i(\vec{x}) + (1-t)f_i(\vec{y}) \geq f_i(t\vec{x} + (1-t)\vec{y}) \tag{4.4}$$

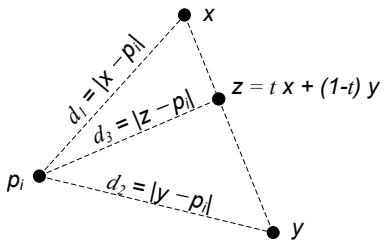


Figure 4.3: Illustration of convexity condition

The geometric illustration in Figure 4.3 suggests that proof of the following inequality

$$td_1^\alpha + (1-t)d_2^\alpha \geq d_3^\alpha \tag{4.5}$$

where d_3 is given by

$$\begin{aligned} d_3 &= \left\|\vec{z} - \vec{p_i}\right\| = \left\|t\vec{x} + (1-t)\vec{y} - \vec{p_i}\right\| \\ &= \left\|t(\vec{x} - \vec{p_i}) + (1-t)(\vec{y} - \vec{p_i})\right\| \end{aligned}$$

is sufficient for proving (4.4).

We divide this problem into four cases. First, consider the case that both d_1 and d_2 are zero, i.e., $d_1 = d_2 = 0$. Then, it is obvious that (4.5) holds. Second, assume $d_1 > 0$ and $d_2 = 0$. Then, we get $d_3 = td_1$, and (4.5) holds. Third, assume $d_1 = 0$, and $d_2 > 0$. Then, we get $d_3 = (1-t)d_2$, and (4.5) holds. Finally, we need to consider the case where both d_1 and d_2 are positive, i.e., $d_1 > 0$ and $d_2 > 0$. Define h_0 such that

$$h_0 \triangleq \frac{1}{d_2^\alpha}\left(td_1^\alpha + (1-t)d_2^\alpha - d_3^\alpha\right)$$

From $\cos\theta \leq 1$,

$$\begin{aligned} d_3 &= \sqrt{t^2 d_1^2 + (1-t)^2 d_2^2 + 2t(1-t)d_1 d_2 \cos\theta} \\ &\leq \sqrt{t^2 d_1^2 + (1-t)^2 d_2^2 + 2t(1-t)d_1 d_2} \\ &= td_1 + (1-t)d_2 \end{aligned}$$

Then, we get

$$\begin{aligned} h_0 &= \frac{1}{d_2^\alpha}\left(td_1^\alpha + (1-t)d_2^\alpha - d_3^\alpha\right) \\ &\geq \frac{1}{d_2^\alpha}\left(td_1^\alpha + (1-t)d_2^\alpha - \left(td_1 + (1-t)d_2\right)^\alpha\right) \\ &= t(\frac{d_1}{d_2})^\alpha + 1 - t - \left(t(\frac{d_1}{d_2}) + 1 - t\right)^\alpha \triangleq h_1(\frac{d_1}{d_2}) \end{aligned}$$

From the following facts,

$$\begin{aligned} &h_1(x) = tx^\alpha + 1 - t - (tx + 1 - t)^\alpha \\ &\frac{d}{dx}h_1(x) \geq 0 \quad \text{if } x \geq 1 \\ &\frac{d}{dx}h_1(x) \leq 0 \quad \text{if } 0 < x \leq 1 \\ &h_1(1) = 0 \end{aligned}$$

it can be seen that the function $h_1(x)$ is always greater than or equal to 0 for all $x > 0$. Thus, we get

$$h_0 = \frac{1}{d_2^\alpha}\left(td_1^\alpha + (1-t)d_2^\alpha - d_3^\alpha\right) \geq h_1(\frac{d_1}{d_2}) \geq 0$$

and

$$td_1^\alpha + (1-t)d_2^\alpha \geq d_3^\alpha$$

Now that (4.5) and consequently (4.4) are proved, it can be seen that the function $f_i(\vec{x})$ is convex. Finally, the function $f(\vec{x})$ is a non-negative sum of convex functions $f_i(\vec{x})$. Therefore, $f(\vec{x})$ is a convex function. □

4.2.4 Multiple Base Stations with Single-Hop Links (K-SH)

Next, we extend the case in Section 4.2.2 to multiple-BS placement. Assume that there are K BSs and N nodes, as in Figure 4.2(b). Each node communicates with its nearest BS, and there is a separate channel for inter-BS communication. We further assume that the BSs have large energy sources relative to those of the nodes. This assumption is reasonable for the applications described in Section 4.1, because 1) the mobile BSs must have energy sources with large capacity for their operation, and 2) it is relatively easy to renew their energy sources compared to those of user nodes. In this scenario, we only need to be concerned about the uplink communication energy consumed by the user nodes.

The energy cost function (4.3) of the 1-SH case does not change its form with respect to the BS location $\vec{x}$. However, such an invariable function cannot be defined for the K-SH case as user nodes change their BSs depending on the BS locations, which makes this problem NP-hard [66]. To circumvent this difficulty, we devise a fast heuristic algorithm with an iterative approach similar to the K-means algorithm.

Define *clusters* $\{C_k\}$ as K disjoint subsets of $P = \{p_i : 1 \leq i \leq N\}$. The K-SH BSP problem can be modeled as the following clustering optimization problem

$$\text{Minimize} \quad \sum_{k=1}^{K} \sum_{i \in C_k} \tau(p_i)(k_1 \|\vec{x}_k - \vec{p}_i\|^\alpha + k_2) \tag{4.6}$$

where clusters C_k and BS locations $\vec{x}_k$ are the optimization variables.

It can be seen that this problem is a generalization of the K-means clustering problem, which is equivalent to the case where $k_0 = 0$, $\alpha = 2$, and $\tau(p_i) = \tau(p_j)$ for all i and j. In order to solve this problem, we propose the K-SH algorithm in Figure 4.4. It first groups nodes into their nearest clusters. Then, it computes BS locations $\vec{x}_k$ by solving

$$\text{Minimize} \quad \sum_{p_i \in C_k} \tau(p_i)(k_1 \|\vec{x}_k - \vec{p}_i\|^\alpha + k_2) \tag{4.7}$$

which is a 1-SH problem. These steps are repeated until no change in clustering occurs.

Convergence of the K-SH algorithm can be proved as follows. First, it is obvious that the energy function is lower-bounded as it should be always non-negative. The topology reconstruction in Step 1 and the BS location adjustment in Step 2 can only decrease the energy function. Thus, the total energy consumption of the network produced by the above operations should always converge.

Optimality of the solution produced by the K-SH algorithm depends on the initial BS locations $\vec{x}_k$, just as in the K-means case [38]. For this reason, in order to obtain satisfactory solutions, K-SH needs to be repeated. In Section 4.2.7, an efficient heuristic method for

Input: $\{\vec{p}_i\}$, K, and $\{\vec{x}_k\}$; node locations, the number of BSs, and the initial guess of BS locations
Output: $\{\vec{x}_k\}$; BS locations

1. Group nodes p_i into clusters such that

$$p_i \in C_k \quad \text{where } k = \arg\min_j \|\vec{p}_i - \vec{x}_j\|$$

2. For each cluster C_k, compute the BS location $\vec{x}_k$ by solving (4.7). Repeat Steps 1 and 2 until no change in clustering occurs.

Figure 4.4: K-SH BSP

making a reasonable guess of BS locations is proposed.

4.2.5 Single Base Station with Multi-Hop Links (1-MH)

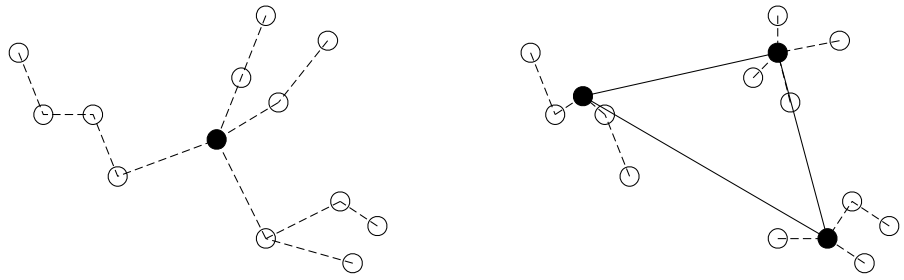
Now we consider the case where nodes can form multi-hop links to a single BS, as in Figure 4.5(a). Again, the cost function changes its form depending on the BS location, and the problem is NP-hard. So we devise an efficient heuristic algorithm similar to that in Section 4.2.4.

Just as in the K-means and the K-SH algorithms, the solution quality of the 1-MH algorithm is determined by the initial clustering. Depending on the initial BS position, 1-MH may produce an unsatisfactory solution by getting stuck at local minima. In order to mitigate the local minimum problem, the 1-MH algorithm searches the central area of the network first since it is likely to contain the optimal BS location.

We define a *centroid* node m of a graph $G = (P, E)$, a centroid for short, as a node that satisfies

$$m = \arg\min_{u \in P} \sum_{v \in P} \tau(v) \cdot d(u, v) \tag{4.8}$$

where $d(u, v)$ denotes the multi-hop distance between nodes u and v, i.e., the sum of edge



(a) (b)

Figure 4.5: Network structure with multi-hop communication: (a) user nodes communicate with the single BS through multi-hop routes; (b) user nodes communicate through multi-hop routes with BSs that minimize communication energy.

weights along the shortest path between u and v. Intuitively, a centroid is a node that can receive data from all other nodes at the minimum energy cost. Note that a centroid node need not be unique, and a graph can have several centroids. In our application, the weight $w(u, v)$ of an edge (u, v) is defined as the communication power between two nodes given by (4.2).

$$w(u, v) \triangleq k_1 \|\vec{u} - \vec{v}\|^{\alpha} + k_2$$

We also define a centroid location $\vec{m}$ as the location of the centroid. We assume that G is a dense graph, i.e., $(u, v) \in E$ for all $u, v \in P$. A centroid can be discovered by an all-pairs shortest path search [23], which in turn, can be done by repeating a single-source shortest path search $N = |P|$ times.

Our proposed 1-MH algorithm is summarized in Figure 4.6. We now describe its details, and apply it to a small example.

Input: $\{\vec{p}_i\}$ and $\vec{x}^0$; node locations, and an initial guess of the BS location
Output: $\vec{x}$; BS location

1. Set the BS location $\vec{x}$ to a centroid location $\vec{m}$; $\vec{x} \leftarrow \vec{m}$.
2. Find the set B_i of neighbors of p_i.
3. Compute the total transmission time $T(p_j)$ for each $p_j \in B_i$.
4. Compute the BS location $\vec{x}^*$ by solving (4.9). Set $\vec{x}$ to one of $\vec{x}^*$ and $\vec{x}^0$, which incurs less energy cost.
5. Find the shortest paths from all nodes to the BS.
6. Find the set B_x of neighbors of the BS.
7. Compute $T(p_j)$ for each $p_j \in B_x$.
8. Compute the BS location $\vec{x}$ by solving (4.10). Repeat Steps 5 through 8 until no change in B_x occurs.

Figure 4.6: 1-MH BSP

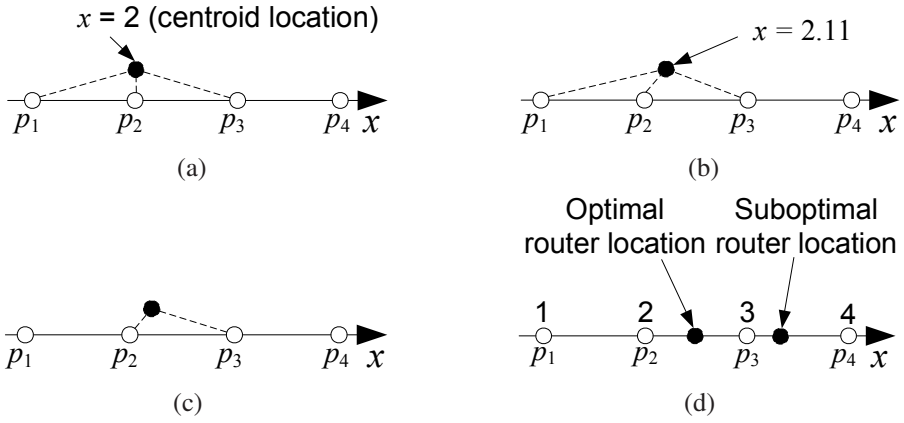


Figure 4.7: Example of 1-MH BSP: the BS location after (a) Steps 1 through 3; (b) Step 4; (c) Steps 5 through 7; and (d) Step 7. The y-coordinate of the BS is always 0.

Step 1 first finds a centroid node p_i and the shortest-paths tree $G_i = (P, E_i)$ rooted at p_i. Then, it sets the BS location $\vec{x}$ to the centroid location $\vec{m}$. Consider the example in Figure 4.7. Four nodes are linearly placed, and the x-coordinates of nodes p_1 through p_4 are 1, 2, 3, and 4, respectively. In this example, either p_2 or p_3 can be selected as the centroid. Assume that p_2 is selected as the centroid. The corresponding shortest paths are given by

$$\Pi_2(p_1) = p_2, \quad \Pi_2(p_2) = p_2, \quad \Pi_2(p_3) = p_2, \quad \Pi_2(p_4) = p_3$$

where $\Pi_k(p_i)$ denotes the predecessor node of p_i, assuming that p_k is the source.

Step 2 builds the neighbor set B_i that contains p_i and the nodes adjacent to p_i.

$$B_i = \{p_i\} \cup \{p_j | (p_i, p_j) \in E_i\}$$

In Figure 4.7(a), nodes p_1 and p_3 are adjacent to the centroid p_2. Thus, the neighbor set B_2 becomes

$$B_2 = \{p_2\} \cup \{p_1, p_3\} = \{p_1, p_2, p_3\}$$

Step 3 computes the total transmission time $T(p_j)$ for each $p_j \in B_i$. For a node p_j adjacent to the centroid p_i, $T(p_j)$ is defined as the sum of the $\tau(p_m)$'s of nodes $\{p_m\}$ in the subtree rooted at p_j. For the centroid p_i itself, $T(p_i)$ is defined as $\tau(p_i)$. In other words,

$$T(p_j) = \begin{cases} \sum_{p_m \in \text{Subtree}(p_j)} \tau(p_m) & \text{if } j \neq i \\ \tau(p_j) & \text{if } j = i \end{cases}$$

The total transmission time $T(p_j)$ is the actual time that node p_j spends for transmitting its own data, and forwarding data from other nodes in the subtree. In order to minimize the total energy consumption, the energy cost function needs to be weighted by $T(p_j)$, not by $\tau(p_j)$. In Figure 4.7(a), the total transmission times are

$$T(p_1) = \tau(p_1), \quad T(p_2) = \tau(p_2), \quad T(p_3) = \tau(p_3) + \tau(p_4)$$

Step 4 computes the new BS location $\vec{x}^*$ by solving

$$\text{Minimize} \quad \sum_{p_j \in B_i} T(p_j)(k_1 \|\vec{x}^* - \vec{p}_j\|^\alpha + k_2) \tag{4.9}$$

which is a 1-SH problem. It is possible that the total energy cost with the new BS location $\vec{x}^*$ is greater than that with the input BS location $\vec{x}^0$. In that case, we discard the new location $\vec{x}^*$, and set $\vec{x}$ to the input location $\vec{x}^0$; otherwise, we adopt the new one, and set $\vec{x}$ to $\vec{x}^*$. This is necessary to guarantee convergence of K-MH in Section 4.2.6, which repeats 1-MH; this will be clarified later. In the example of Figure 4.7, it is obvious that the optimal BS location lies on a line through those four nodes. So we can reduce this BSP problem to a one-dimensional case.

Now let us find the BS location x by solving (4.9). Assume that the energy cost function

is given by $Pow(u, v) = \|\vec{u} - \vec{v}\|^4$; i.e., $k_0 = 0$, $k_1 = 1$, $P_0 = 1$, and $\alpha = 4$; see (4.2). For simplicity, assume unit transmission times; i.e., $\tau(p_i) = 1$ for all i. Then, (4.9) becomes

$$\begin{aligned} \text{Minimize} \quad f_1(x) &= T(p_1) \cdot Pow(p_1, x) + T(p_2) \cdot Pow(p_2, x) + T(p_3) \cdot Pow(p_3, x) \\ &= (x-1)^4 + (x-2)^4 + 2(x-3)^4 \end{aligned}$$

and $f_1(x)$ is minimum at $x \cong 2.11$; see Figure 4.7(b).

Step 5 discovers the shortest paths from the BS to all nodes, and builds the shortest path graph $G_x = (P \cup \{x\}, E_x)$, where the BS x is the root. For the given BS location $x \cong 2.11$ in Figure 4.7(c), the shortest paths to the BS x are given by

$$\Pi_x(p_1) = p_2, \quad \Pi_x(p_2) = x, \quad \Pi_x(p_3) = x, \quad \Pi_x(p_4) = p_3$$

Step 6 finds the set B_x of neighbors of the BS x

$$B_x \triangleq \{p_j | (x, p_j) \in E_x\}$$

In Figure 4.7(c), nodes p_2 and p_3 are adjacent to the BS x. Thus, the neighbor set B_x becomes $\{p_2, p_3\}$.

Step 7 computes the total transmission time $T(p_j)$ for each $p_j \in B_x$ in a manner similar to that in Step 3.

$$T(p_j) = \sum_{p_m \in \text{Subtree}(p_j)} \tau(p_m)$$

In Figure 4.7(c), the total transmission times are

$$T(p_2) = \tau(p_1) + \tau(p_2), \quad T(p_3) = \tau(p_3) + \tau(p_4)$$

Step 8 computes the BS location $\vec{x}$ by solving

$$\text{Minimize} \quad \sum_{p_j \in B_x} T(p_j)(k_1 \|\vec{x} - \vec{p}_j\|^{\alpha} + k_2) \tag{4.10}$$

which is a 1-SH problem. In our example, problem (4.10) becomes

$$\begin{aligned} \text{Minimize} \quad f_2(x) &= T(p_2) \cdot Pow(p_2, x) + T(p_3) \cdot Pow(p_3, x) \\ &= 2(x-2)^4 + 2(x-3)^4 \end{aligned}$$

and $f_2(x)$ is minimum at $x = 2.5$, which is the optimal location. The corresponding total energy consumption $E_1(x)$ is given by

$$\begin{aligned} E_1(x) = \tau(p_1) \cdot Pow(p_1, p_2) &+ \Big(\tau(p_1) + \tau(p_2)\Big) \cdot Pow(p_2, x) \\ &+ \Big(\tau(p_3) + \tau(p_4)\Big) \cdot Pow(p_3, x) + \tau(p_4) \cdot Pow(p_4, p_3) \end{aligned}$$

and $E_1(2.5) = 2.25$.

Steps 5 through 8 are repeated until no improvement is seen. In our example, $x = 2.5$ is the global optimum, and so repetition will not improve the solution any further.

Steps 1 through 4 of 1-MH attempt to move the BS to the central area of the network. Then, Steps 5 through 8 iteratively improve the solution. In order to appreciate the role of Steps 1 through 4, let us apply only Steps 5 through 8 to the example in Figure 4.7. Assume that the initial location of the BS is $x = 3.5$. Then, problem (4.10) becomes

$$\begin{aligned} \text{Minimize} \quad f_3(x) &= \big(\tau(p_1) + \tau(p_2) + \tau(p_3)\big) \cdot Pow(p_3, x) + \tau(p_4) \cdot Pow(p_4, x) \\ &= 3(x-3)^4 + (x-4)^4 \end{aligned}$$

and $f_3(x)$ is minimum at $x \cong 3.41$; see Figure 4.7(d). The corresponding total energy consumption $E_2(x)$ is given by

$$E_2(x) = \tau(p_1) \cdot Pow(p_1, p_2) + \Big(\tau(p_1) + \tau(p_2)\Big) \cdot Pow(p_2, p_3)$$
$$+ \Big(\tau(p_1) + \tau(p_2) + \tau(p_3)\Big) \cdot Pow(p_3, x) + \tau(p_4) \cdot Pow(p_4, x)$$

and $E_2(3.41) \cong 3.21 > E_1(2.5)$. This local minimum cannot be escaped by repeating Steps 5 through 8. Thus, it can be seen that the centroid search technique is effective in providing a good initial BS position for solving the 1-MH problem.

Convergence of this algorithm can be proved in a manner similar to that in Section 4.2.4. A new BS location computed by Step 8 is always better than or equal to the previous one. Thus, the BSP algorithm for the 1-MH problem is guaranteed to converge.

4.2.6 Multiple Base Stations with Multi-Hop Links (K-MH)

Finally, we combine the techniques in Sections 4.2.4 and 4.2.5 to solve the multiple-BS placement problem with multi-hop links, as in Figure 4.5(b). The proposed solution

Input: $\{\vec{p}_i\}$, K, and $\{\vec{x}_k\}$; node locations, the number of BSs, and an initial guess of BS locations
Output: $\{\vec{x}\}$; BS location

1. Find shortest paths from all nodes to all BSs; then, group nodes into clusters such that
$$p_i \in C_k \quad \text{where } k = \arg\min_j \mathrm{d}(\vec{p}_i, \vec{x}_j)$$
2. For each cluster C_k, compute the BS location $\vec{x}_k$ by applying Figure 4.6. Repeat Steps 1 and 2 until no change in clustering occurs.

Figure 4.8: K-MH BSP

method is presented in Figure 4.8. It first searches shortest paths from all nodes to all BSs, and groups nodes into the nearest clusters. Each cluster comprises a 1-MH problem, and thus BS locations $\vec{x}_k$ can be computed by applying the algorithm in Figure 4.6 to the corresponding clusters. These steps are repeated until no change in clustering occurs. The proof of convergence of K-MH is similar to that of K-SH and 1-MH.

4.2.7 Cluster Seeding

As mentioned above, optimality of the proposed BSP algorithms depends on the initial locations of the BSs. For this reason, these algorithms need to be repeated with different initial locations to produce optimal or near-optimal solutions. The number of repetitions needed depends on how the initial clustering seeds are constructed.

A widely used cluster seeding method is to choose seed locations that are far from each other with a high probability [66]. In terms of our BSP problem, a node location $\vec{p}_i$ is chosen as the BS location $\vec{x}_k$ with the probability proportional to the distance squared from $\vec{p}_i$'s nearest BS location.

$$\Pr(\vec{x}_k = \vec{p}_i) = \frac{\min_{1 \le j < k} \|\vec{p}_i - \vec{x}_j\|^2}{\sum_i \left(\min_{1 \le j < k} \|\vec{p}_i - \vec{x}_j\|^2 \right)} \tag{4.11}$$

Considering our cost function (4.2), it is also plausible to adopt the following criterion for seeding instead of (4.11).

$$\Pr(\vec{x}_k = \vec{p}_i) = \frac{\min_{1 \le j < k} \left(\tau(p_i) \cdot Pow(p_i, x_j) \right)}{\sum_i \min_{1 \le j < k} \left(\tau(p_i) \cdot Pow(p_i, x_j) \right)} \tag{4.12}$$

However, one of the drawbacks with the seeding method proposed by Ostrovsky et al. [66] is that it often becomes excessively tolerant to suboptimal seeding because it de-

termines the seeding location through a stochastic decision-making process. When there exists no obvious good clustering, and the probabilities given by (4.11) or (4.12) are nearly comparable for different p_i's, the above seeding method becomes hardly better than random seeding. In order to mitigate this problem, we adopt the following seeding scheme.

$$\vec{x}_k = \vec{p}_i \quad \text{where } i = \arg\max_i \min_{1 \le j < k} \left(\tau(p_i) \cdot Pow(p_i, x_j)\right) \tag{4.13}$$

In this scheme, the first BS seed is randomly chosen. Then, the other BS locations are selected one by one from node locations that expend the most energy. In other words, the farthest node location from the previous seeds is chosen. We call this seeding method *farthest-first* seeding, and its impact is discussed in Section 4.3.2.

4.3 Simulation Results for Base Station Placement

This section presents simulation results with the BSP algorithms described in Sections 4.2.

4.3.1 Simulation Environment

First, we describe the simulation environment and assumptions about the radio channel model. Nodes and BSs are placed on a 1×1 km^2 plane. We assume the following radio parameters, as in Section 3.4: radio attenuation exponent $\alpha = 4.0$; transmission power $Pow = 281.8$ mW for transmission distance $Dist = 250$ m. Accordingly, we set the transmission power coefficients k_1 and k_2 in (4.2) to $k_1 = 7.2141 \cdot 10^{-11}$ W/m^4 and $k_2 = 0$ W, respectively.

4.3.2 Impact of Cluster Seeding

The solution quality of the proposed BSP algorithms as well as K-means depends on the initial clustering seeds. Figures 4.9(a) and 4.9(b) show examples of optimal and suboptimal BSP obtained by the K-SH and the K-means algorithms. White circles denote nodes, and black, BSs; 36 nodes are regularly placed in a grid structure, and the horizontal and the vertical distance between grid points is 200 m. Nine BSs are placed using the K-SH algorithm; in fact, the K-means algorithm produces exactly the same BS locations in this particular case.

The BS locations in Figure 4.9(a) are optimal; each node is $\sqrt{2}/10 \cdot 1000$ m from the nearest BS, and is consuming Pow = 28.9 mW. Consequently, the total power consumption is $\text{Pow}_{\text{tot}} = 36 \cdot 28.9$ mW = 1.04 W. On the other hand, Figure 4.9(b) shows an example of suboptimal BS locations produced by the K-SH and the K-means algorithms. Both optimal and suboptimal solutions can be produced by these algorithms depending on the initial cluster seeding.

Figures 4.10(a) through 4.10(d) contain histograms of solutions produced by (a) ran-

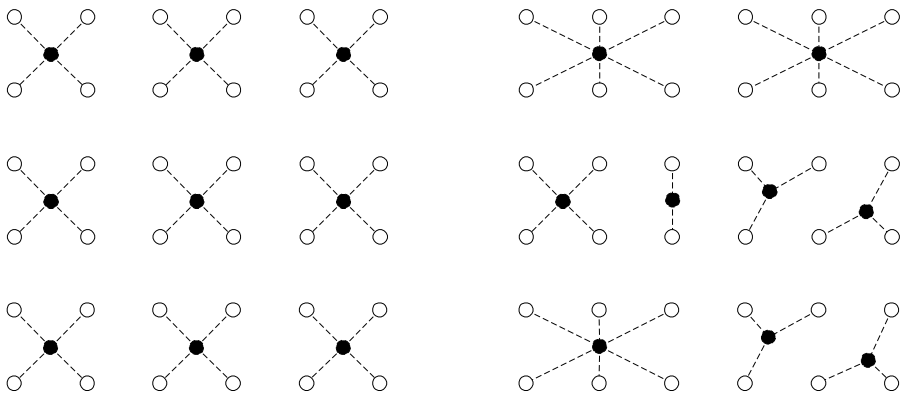


Figure 4.9: Examples of (a) optimal BSP, and (b) suboptimal placement. User nodes (white) are placed in a grid structure, and locations of BSs (black) are computed by K-SH.

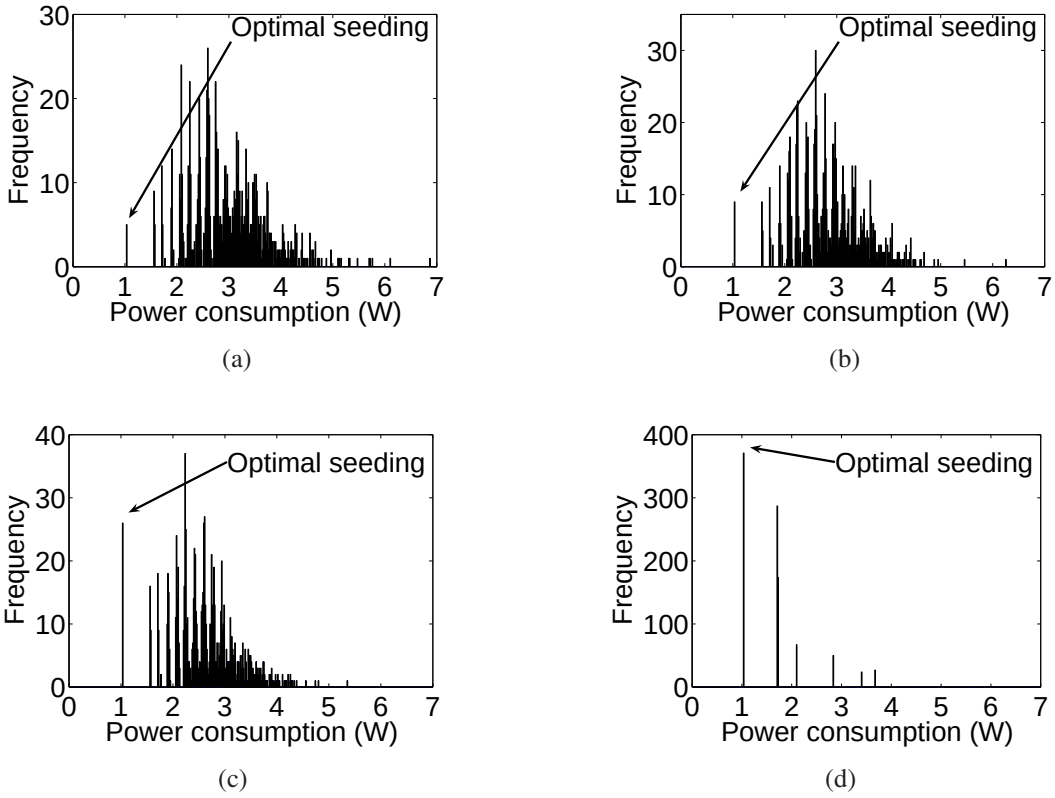


Figure 4.10: Total power consumption after (a) random seeding; (b) the seeding method of [66] (original); (c) the seeding method of [66] (modified); (d) the farthest-first method; resolution of the histogram data is 15 mW.

dom seeding; (b) the seeding method of Ostrovsky *et al.* [66] using (4.11); (c) the modified method (4.12); and (d) the farthest-first method. Each seeding method was repeated 1000 times, and the solution quality was measured by the total transmission power. Note that optimal BSP corresponding to Figure 4.9(a) consumes 1.04 W. According to Figure 4.10(b), the probability that the seeding method of Ostrovsky *et al.* produces an optimal solution is approximately 1%, which is only slightly higher than that of the random seeding (0.4%). Figure 4.10(c) shows the quality of the solution produced by the seeding method using (4.12). The probability (2.7%) that this method produces an optimal solution is slightly higher than those of the two previous methods.

Figure 4.10(d) shows that the farthest-first method produces the optimal solution with

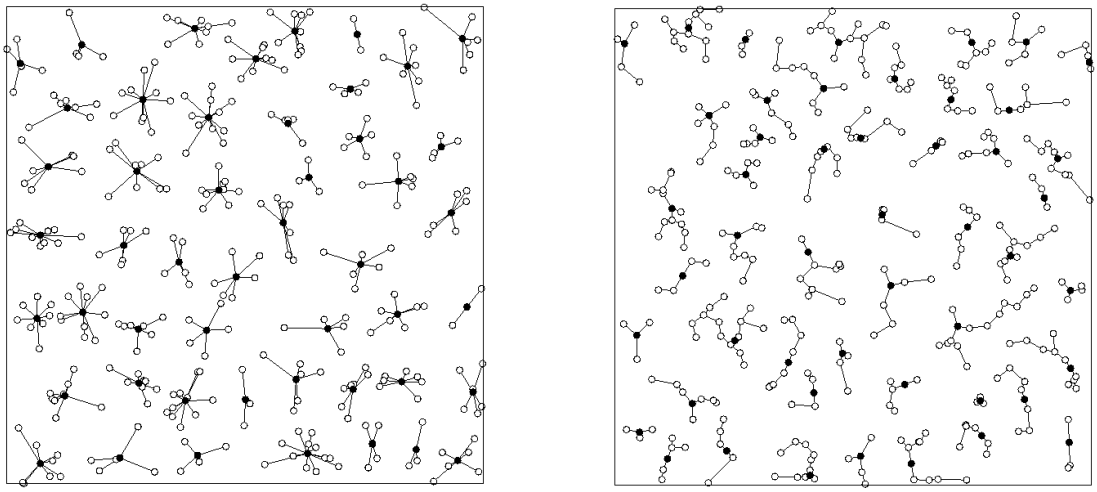
high probability (38%). As mentioned in Section 4.2.7, the method of Ostrovsky *et al.* tends to produce suboptimal solutions, especially when the best clustering mode is only marginally better than the other modes. Conversely, the farthest-first method pursues only the best clustering mode in a greedy manner, and consequently, outperforms the other seeding methods when nodes are almost uniformly distributed. Thus, the farthest-first method was adopted for the remaining experiments.

4.3.3 Base Station Placement for Single-Hop Networks

In order to evaluate efficiency of the K-SH algorithm, we compare the power consumed by two different network structures: a network in which the BS locations are determined by the K-SH algorithm; and a grid network structure which resembles the routing backbone of cellular networks [87]. The number N of nodes is set to 320, and the number K of BSs varies from 9 to 64. In the grid structure, the 1×1 km^2 plane is divided into K squares, and a BS is located at the center of each square. For example, when $K = 4$, four BSs are located at (x, y) = (250 m, 250 m), (250 m, 750 m), (750 m, 250 m), and (750 m, 750 m). For a given K, 1000 different sets of random node locations were considered to compute the average power consumption; for each node arrangement, the K-SH algorithm was repeated 20 times to get the best BS locations.

Assuming that nodes and BSs are randomly distributed, the probability density function $f_d(x)$ of the distance from a node to its nearest BS is given by [24]

$$f_d(x) = 2K\pi x \exp(-K\pi x^2)$$



(a) (b)

Figure 4.11: Simulation results for BSP with (a) K-SH and (b) K-MH. Three hundred and twenty nodes (white) are randomly dispersed, and 49 BSs (black) are placed to minimize energy consumption.

Thus, the average power consumption of each node is

$$\begin{aligned} Pow &= \int_0^\infty (k_1 \cdot x^4 + k_2) f_d(x) dx \\ &= 7.2141 \cdot 10^{-11} \cdot \frac{2}{K^2 \pi^2} \text{ (W)} \end{aligned} \tag{4.14}$$

which is monotonically decreasing with respect to K.

Figure 4.11(a) illustrates a simulation of BSP using the K-SH method. Nodes are linked to each BS (clusterhead) via single-hop connections. Figure 4.12(a) shows the simulation results for the total power consumption computed by K-SH, the grid structure, and the theoretical power consumption from (4.14); note that the vertical scale is logarithmic. It can be seen that both the grid and the K-SH structures consume less power as the number of BSs increases. Also the simulated power consumption by K-SH is in line with the theoretical value, especially when K is large.

Figure 4.12(b) shows the ratio between power consumption by the K-SH and grid struc-

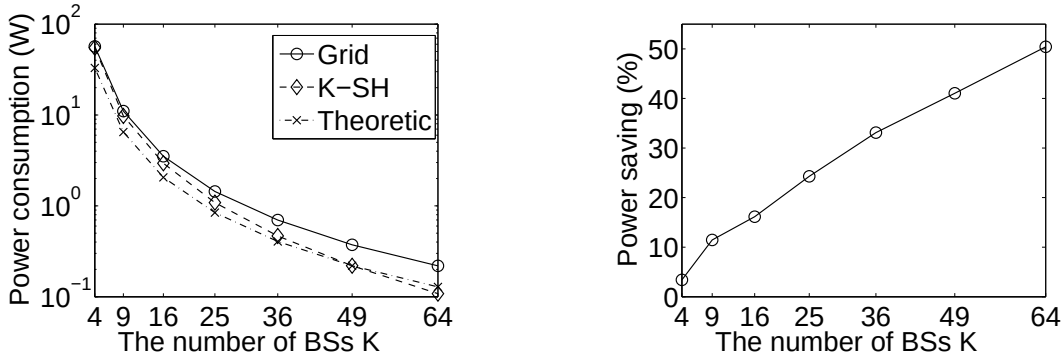


Figure 4.12: Power consumption of a grid network and a K-SH structure; (a) total power consumption by the grid and K-SH; and (b) the power saving obtained by K-SH.

tures. For instance, when $K = 64$, the power saving by the K-SH structure amounts to 51%; i.e., it consumes only 49% of the power consumed by the grid structure. Thus, we conclude that K-SH effectively reduces the total power consumption by determining optimal BS locations, and the relative saving increases as the number of BSs grows.

4.3.4 Base Station Placement for Multi-Hop Networks

Now, we evaluate the efficiency of the K-MH algorithm by comparing its power consumption to that of the grid structure. Again, the number N of nodes is set to 320, and the number K of BSs varies from 9 to 64. Figure 4.11(b) shows an example of BSP using the K-MH method. Nodes are linked to each BS via multi-hop connections.

Assuming that nodes and BSs are randomly distributed, the average power consumption for data from a node to reach its nearest BS is proportional to the number of hops from the node to the BS. In turn, the number of hops is proportional to the Euclidian distance from the node to the BS. Thus, the average power consumption for data from a node to reach a

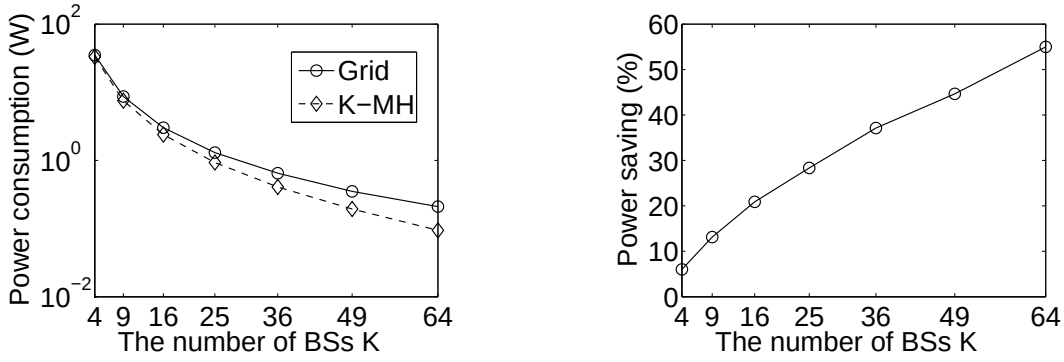


Figure 4.13: Power consumption of a grid network and a K-MH structure; (a) total power consumption by the grid and K-MH; and (b) the power saving obtained by K-MH.

BS is given by

$$Pow \propto \int_0^\infty x f_d(x) dx = \frac{1}{2\sqrt{K}}$$

which is monotonically decreasing with respect to K.

Figure 4.13(a) shows the simulation results for the total power consumed by the K-MH and grid structures. It can be seen that both structures consume less power as the number of BSs increases. Figure 4.13(b) shows the ratio of the power consumption by the K-MH and grid structures. From these results, we conclude that K-MH effectively reduces the total power consumption, and the relative saving increases as the number of BSs grows.

4.4 Distributed Relay Placement Optimization

Now we present a distributed relay placement (DRP) optimization algorithm for energy-efficient ad hoc networks. While the placement algorithms presented in Section 4.2 compute the optimal locations of BSs that act as clusterheads with relatively large energy sources, the algorithms presented in this section attempt to place *mobile relays* that forward data from previous hop nodes to next hop nodes. We also investigate the DRP prob-

lem in the presence of radio obstacles. We devise a mechanical model of a mobile relay's propulsion force, and utilize it to implement the DRP algorithm.

4.4.1 Communication and Network Model

First, we state our assumptions concerning the radio communication and the network node's capabilities. As illustrated in Figure 4.1(b), the nodes consist of user nodes and relays, which may be mobile. For simplicity, we assume that only one user node is assigned as the *sink* node that serves as the final destination of all data generated by other user nodes. In order to distinguish other user nodes from the sink, we refer to non-sink user nodes that send data to the sink as *sources*. Sources send new data to the sink through direct or multi-hop routes. Relays do not generate data, and only forward data from sources to the sink. The network attempts to construct energy-efficient routes using the PALM algorithm introduced in Chapter 3. We again adopt the radio model (1.2) presented in Section 1.2.3.

$$Gain(TX, RX) = \frac{1}{k_1 \cdot Dist(TX, RX)^{\alpha} + k_2}$$

We assume that nodes adjust their transmission power to the minimum necessary level, and the receiver's sensitivity $RSS_{\min}$ is set to 1. Then, the power cost $Cost(d)$ for communication across distance d is given by

$$Cost(d) = k_1 \cdot d^{\alpha} + k_2 \tag{4.15}$$

We also assume that relays can sense the direction of incoming radio signal, and the distance to the transmitter. By sensing the direction and the distance, relays can deduce the relative locations of their neighbors. We further assume that nodes can sense their velocities. For the current network topology, relays identify their immediate neighbors, and move toward locally optimal locations according to an artificial "force", which will be presented

in the next section.

When an object is located between two nodes, their direct communication may be obstructed. We make the following assumptions about such radio obstacles. First, the above radio channel model (4.15) holds only for node pairs between which a line-of-sight communication is available. When no line-of-sight communication is available, the radio gain between two nodes becomes zero. Second, mobile relays have full information about the locations and the shapes of radio obstacles.

4.4.2 Mechanical Analog

In order to formulate and solve the relay placement problem, we model the network as a mechanical system with springs and a viscous damper, which is a widely used approach for solving optimization problems [57]. Specifically, we model the communication energy cost as an artificial potential energy stored in springs, and nodes as objects with unit mass, moving according to the artificial force field, i.e., the negative gradient of an artificial potential function. Movement of objects in this model resembles the progressive solution improvement of the steepest descent method [13] in convex optimization. Thus, through this model, we can obtain a locally optimal solution as the mechanical system converges to an equilibrium point.

Consider the mechanical system in Figure 4.14(a). Let us model the potential energy stored in an artificial spring of length x as

$$E_p(x) = k_1 \cdot x^{\alpha} + k_2$$

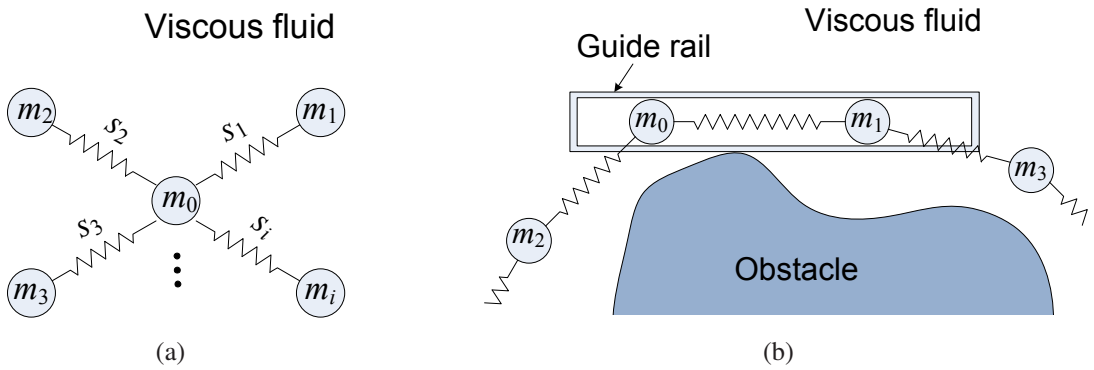


Figure 4.14: MANET structure modeled as a mechanical system in a viscous fluid: (a) connection with spring elements; and (b) connection with spring elements with a guide rail in contact with an obstacle

The total potential energy of the system in Figure 4.14(a) is given by

$$\sum_{i=1}^{n} E_p(x_i) = \sum_{i=1}^{n} (k_1 \cdot x_i^{\alpha} + k_2)$$

Let us also assume that the whole mechanical system is immersed in a viscous fluid. Each spring s_i exerts tensile force $\vec{F}_i(x_i)$ on the mass, the magnitude of which is given by

$$F_i(x_i) = \frac{\partial E_p(x_i)}{\partial x_i}$$

The direction of $\vec{F}_i$ is toward m_i. Mass m_0 also experiences drag force $\vec{F}_v$ due to the viscous fluid [65], which is given by

$$\vec{F}_v = -K_v \dot{\vec{x}}_0$$

where $\vec{x}_0$ and K_v denote node 0's location and the coefficient of viscosity of the fluid. Under these conditions, the system may initially oscillate for a while, and will eventually converge to the equilibrium state in which the total potential energy is minimum.

In addition to spring and damper elements described above, we adopt another mechanical element in order to avoid node movement that could cause disconnection due to radio

obstacles. Consider a *guide rail* element illustrated in Figure 4.14(b). Masses m_0 and m_1 are connected by a spring, and again, immersed in viscous fluid. They can move linearly within the rail, experiencing the tensile force exerted by the spring. We tentatively assume that the guide rail can rotate around the point of contact with the radio obstacle. Masses m_0 and m_1 are also connected to other springs, the obstacle is pushing the guide rail, and in consequence, m_0 and m_1 are being pushed away from the obstacle by the guide rail. According to this model, the masses can move in the direction parallel to the guide rail, but the line-of-sight between them does not intersect the obstacle.

4.4.3 Proposed Relay Placement Algorithm

Based on the above mechanical analog, we design a distributed controller of mobile relays as follows. The sink node determines the area where an additional relay is needed, and dispatches a new mobile relay to that area. On reaching the area, the new relay starts to discover energy-efficient routes, and moves in the direction specified by the mechanical system so that the total energy consumption decreases. Then, the sink again dispatches another relay. These steps continue until the desired number of relays are placed in the network.

Before describing the DRP algorithm in detail, we state some additional definitions and assumptions. We assume that each source s has a data generation rate $b(s)$. We also assume that before additional relays are inserted, there exists a route from each source to the sink, either direct or via multi-hop paths through existing relays.

In the DRP algorithm, each relay performs the following operations repeatedly: (1) route redirection, (2) location sensing, (3) energy cost estimation, and (4) movement control. First, it continuously attempts to discover energy-efficient routes by performing route redirection according to the PALM operations presented in Chapter 3. Second, both sources

and relays continuously broadcast hello messages to the immediate neighbors that are directly connected to themselves. By sensing the incoming hello messages, each relay i can estimate the energy cost to reach its neighbors.

Third, the relay estimates the energy cost around itself, and forwards the estimated value to the next hop node. Each source node records its data generation rate $b(s)$ on the data packets it transmits. By inspecting the source address and $b(s)$ recorded on the received data packets, relay i can determine the data rate $w(i, j)$ across the link (i, j), where j is a neighbor node of i. Thus, the relay can estimate the total energy cost $E_{p,total}(i)$ around itself as follows:

$$E_{p,total}(i) = \sum_{j \in N(i)} w(i, j) \cdot E_p(\|\vec{i} - \vec{j}\|)$$

where $N(i)$ denotes the set of i's neighbors. Source nodes perform a similar computation, and record their addresses and the energy cost on outgoing packets. Relay i inspects the $E_{p,total}(i')$ value recorded on the packet, and if $E_{p,total}(i) > E_{p,total}(i')$, then it records its own address and energy cost on the packet, and forwards it. In consequence, the sink node will receive the address of the node that has the greatest energy cost.

Fourth, each mobile relay controls its movement according to the mechanical model described in Section 4.4.2. Suppose the network system has no radio signal obstruction, and assume that mobile relays are equipped with propulsion devices that can exert driving forces of arbitrary direction and magnitude. The propulsion force $\vec{F}_i$ of relay i that emulates the behavior of the above mechanical system is given by

$$\vec{F}_i = \sum_{j \in N(i)} w_{i,j} \cdot \vec{F}_{i,j} - K_v \cdot \dot{\vec{x}}_i$$

where $N(i)$, $w_{i,j}$ and $\vec{F}_{i,j}$ denote the set of i's neighbor nodes, the number of data bits that pass across link (i, j), and the force exerted on node i by node j, respectively. The direction

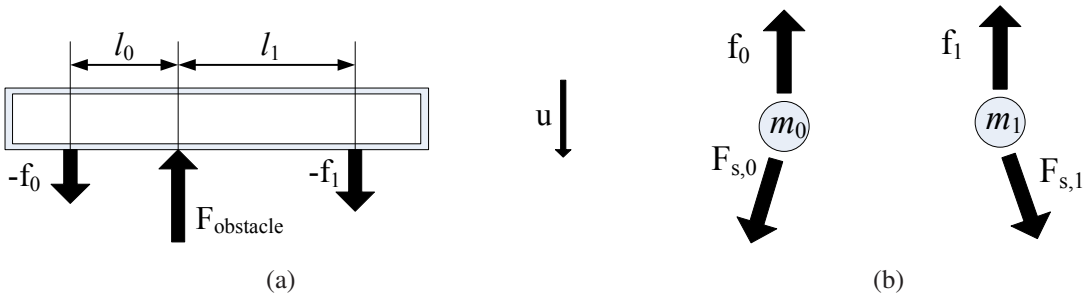


Figure 4.15: Free-body diagrams of (a) a guide rail, and (b) masses m_0 and m_1

of $\vec{F}_{i,j}$ is toward node j, and its magnitude is given by

$$\left\|\vec{F}_{i,j}\right\| = \left.\frac{\partial E_p(x)}{\partial x}\right|_{x=Dist(i,j)}$$

Now consider the case where radio obstacles exist among the network system. Figure 4.15 illustrates the free-body diagrams of a guide rail and masses m_0 and m_1. Let us first compute the reaction forces, $\vec{f_0}$ and $\vec{f_1}$, exerted on m_0 and m_1 by the guide rail. Let $\vec{F}_{s,0}$ and $\vec{F}_{s,1}$ denote the vector sum of all forces exerted on m_0 and m_1 by springs, with link weights taken into account, i.e.,

$$\vec{F}_{s,i} \triangleq \sum_{j \in N(i)} w_{i,j} \cdot \vec{F}_{i,j}$$

Note that $\vec{F}_{s,0}$ and $\vec{F}_{s,1}$ do not contain the reaction force components. We assume that masses m_0 and m_1 have unit value, i.e., $m_0 = m_1 = 1$, and the mass of the guide is negligible. In addition, we define a vector $\vec{u}$ perpendicular to the guide rail with unit magnitude, i.e., $\|\vec{u}\| = 1$. Then, applying Newton's second law to the guide rail in Figure 4.15(a), we get

$$l_0 \cdot \vec{f_0} = l_1 \cdot \vec{f_1} \tag{4.16}$$

The geometry of the guide rail implies

$$\left(\frac{\ddot{\vec{x}}_0}{l_0} + \frac{\ddot{\vec{x}}_1}{l_1}\right) \cdot \vec{u} = \left(\frac{\vec{F}_{s,0} + \vec{f}_0}{l_0} + \frac{\vec{F}_{s,1} + \vec{f}_1}{l_1}\right) \cdot \vec{u} = 0 \tag{4.17}$$

where $(\cdot)$ denotes the dot (inner) product. From (4.16) and (4.17), we get

$$\vec{f}_0 = -\left[\left(\frac{l_1^2}{l_0^2 + l_1^2}\vec{F}_{s,0} + \frac{l_0 l_1}{l_0^2 + l_1^2}\vec{F}_{s,1}\right) \cdot \vec{u}\right] \cdot \vec{u} \tag{4.18}$$

Equation (4.18) suggests that in order to compute the reaction force exerted on m_0 by the guide rail, we need to get the spring force $\vec{F}_{s,1}$ exerted on m_1 and the leverage distances l_0 and l_1, as well as $\vec{F}_{s,0}$. Note that (4.18) holds only when m_1 is incident on only one guide rail that m_0 is incident on. If node m_1 is incident on more than one guide rail, then in order to compute the reaction force $\vec{f}_0$, we need to compute all spring and reaction forces on all masses simultaneously, which makes it difficult to compute the reaction forces in a distributed manner. In order to avoid such complication, we propose two heuristics as follows.

First, we compute the reaction force $\vec{f}_{i,j}$ exerted on m_i due to the guide rail incident on m_i and m_j using (4.18) and assuming that m_j is incident to only one guide rail, i.e.,

$$\vec{f}_{i,j} = -\left[\left(\frac{l_j^2}{l_i^2 + l_j^2}\vec{F}_{s,i} + \frac{l_i l_j}{l_i^2 + l_j^2}\vec{F}_{s,j}\right) \cdot \vec{u}\right] \cdot \vec{u} \qquad \text{(First heuristic for reaction force)} \tag{4.19}$$

Second, instead of computing the reaction force using (4.19), we assume that the guide rail cannot rotate around the contact point. Then, we get

$$\vec{f}_{i,j} = -(\vec{F}_{s,i} \cdot \vec{u}) \cdot \vec{u} \qquad \text{(Second heuristic for reaction force)} \tag{4.20}$$

Note that for the first heuristic computation of the reaction force, m_0 still needs to communicate with m_1 to get l_1 and $\vec{F}_{s,1}$, while the second heuristic does not require such an exchange of data. Finally, we compute reaction forces by either (4.19) or (4.20) if and only if the distance between a guide rail and an obstacle is less than a prescribed critical distance $d_{critical}$.

Now that we have a means to evaluate the reaction forces when radio obstacles exist, we compute the propulsion force $\vec{F}_i$ as

$$\vec{F}_i = \sum_{j \in N(i)} w_{i,j} \cdot \vec{F}_{i,j} + \sum_{j \in Obst(i)} \vec{f}_{i,j} - K_v \cdot \dot{\vec{x}}_i$$

where $Obst(i)$ denotes the set of nodes that are adjacent to i, and the distance between a radio obstacle and the communication link with i is less than $d_{critical}$.

4.5 Simulation Results for Relay Placement

This section provides simulation results for the DRP algorithm. We first consider the network space without radio obstacles, and compare the energy efficiency of the network structure produced by the DRP algorithm with that of Li and Cassandras' algorithm [57], which is described next.

Unlike our algorithm, Li and Cassandras' relay placement algorithm is a centralized scheme with assumptions similar to the mechanical model presented in Section 4.4.3, except for the guide rail element. Its operation is as follows. First, the *bottleneck* node k is chosen, which is defined as

$$k = \arg\max_i \sum_{j \in N(i)} \|\vec{F}(i,j)\|$$

Second, the algorithm investigates all possible ways of linking nodes around the bottleneck

node, the number of which is $3 \cdot 2^m - 2$, where m denotes the number of k's neighbors. Next, it selects the best connectivity with the minimum power cost, and adds a new relay according to the selected connectivity. Then, for the given network connectivity, optimal node locations are determined through the so-called *inner-force* method. The algorithm continues to add relays until the intended number of relay nodes are inserted.

For a fair comparison, we adopt the following radio parameters from [57]:

$$Cost(d) = k_1 \cdot d^{\alpha} + k_2 \tag{4.21}$$

where $\alpha = 4$, $k_1 = 0.001$ pJ/bit·m^4, and $k_2 = 180$ nJ/bit. Each source generates and sends

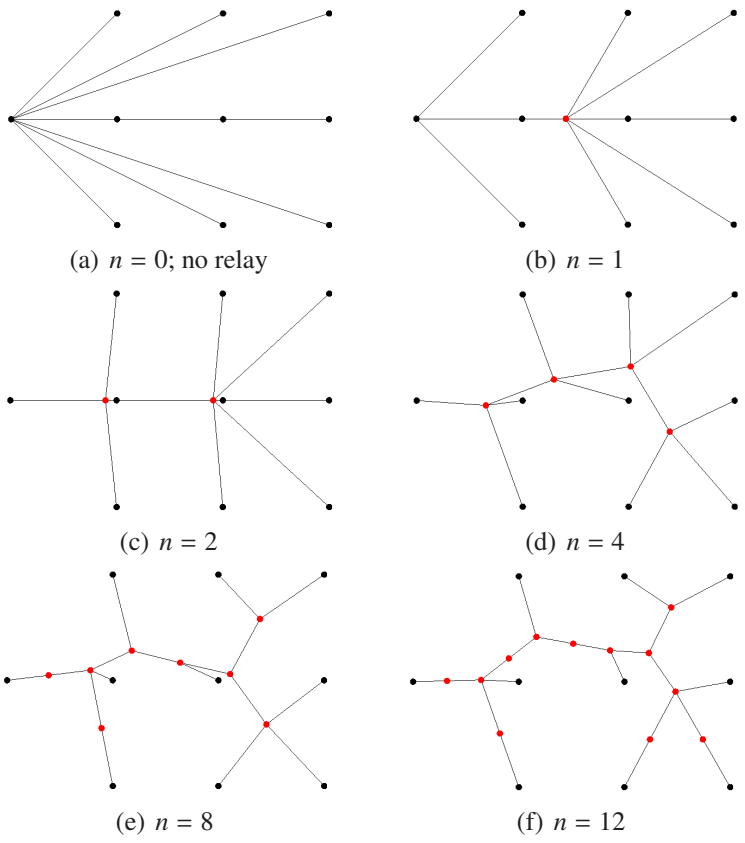


Figure 4.16: Examples of simulation results for relay placement without radio obstacles: n denotes the number of added relays

data to the sink at 1 kbps. Figure 4.16 shows simulation results for the DRP algorithm without radio obstacles. In Figure 4.16(a), nine source nodes are directly connected to the single sink node. The separation between nodes is 300 m.

Figure 4.17 compares the power consumption of our DRP algorithm and that of Li and Cassandras. As Figure 4.17(a) shows, the power consumption of each algorithm decreases as the number of added relays n increases, and the gap between these algorithms also decreases. Let $P_{distributed}$ and $P_{centralized}$ denote the total power consumption of the algorithms. Figure 4.17(b) shows the relative overhead of the power consumption measured as $(P_{distributed} - P_{centralized})/P_{centralized}$. When the number n of added relays is small, the overhead is large, but as n grows, it decreases and becomes as small as 25%. The DRP algorithm constructs a power-efficient network structure in a distributed manner, and its power efficiency is comparable to that of Li and Cassandras' centralized algorithm.

Figure 4.18 illustrates simulation results for relay placement in the presence of radio obstacles. The cost model (4.21) is again adopted with the same parameter values. Nine source nodes (black) are placed in a grid structure on a 2500 m $\times$ 2500 m plane. The sink

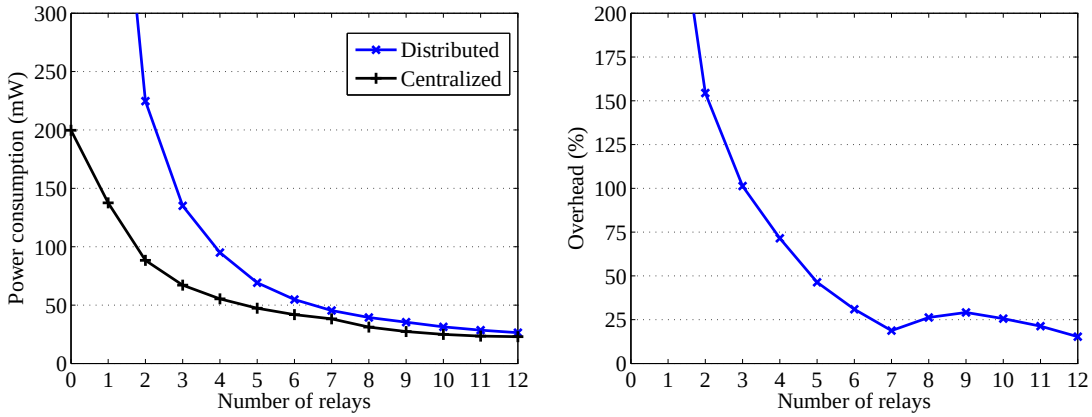
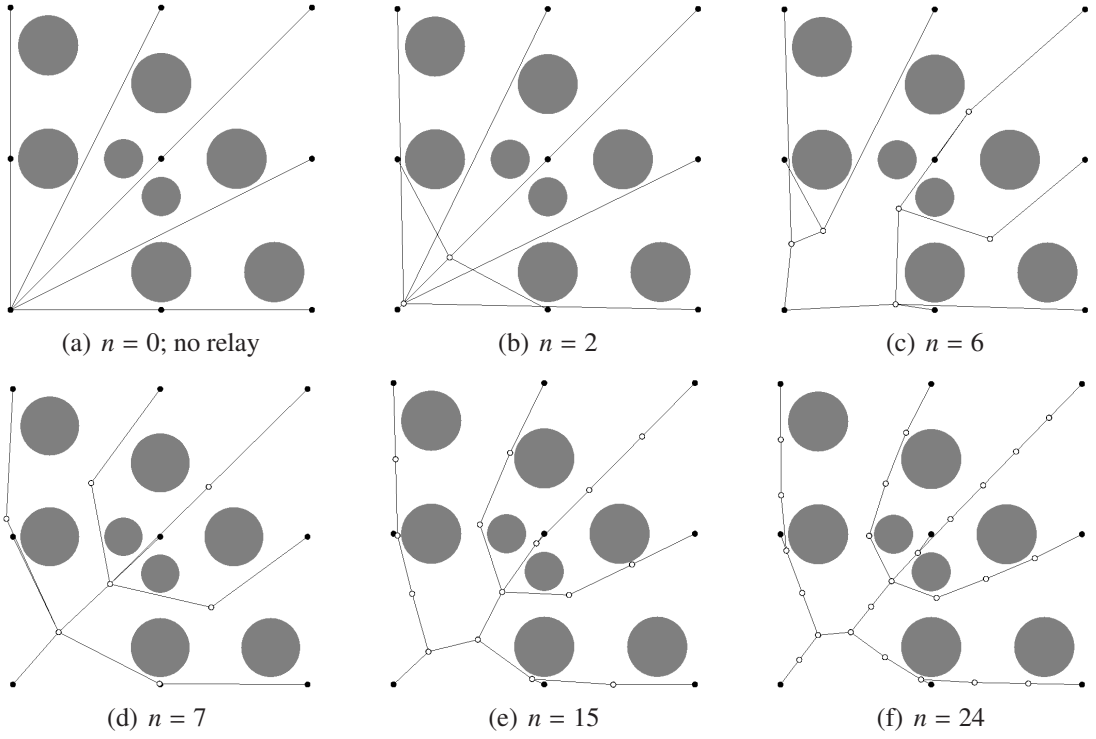


Figure 4.17: Simulation results for the power consumption of centralized and distributed relay placement algorithms: (a) total power consumption, and (b) power consumption overhead due to distributed implementation



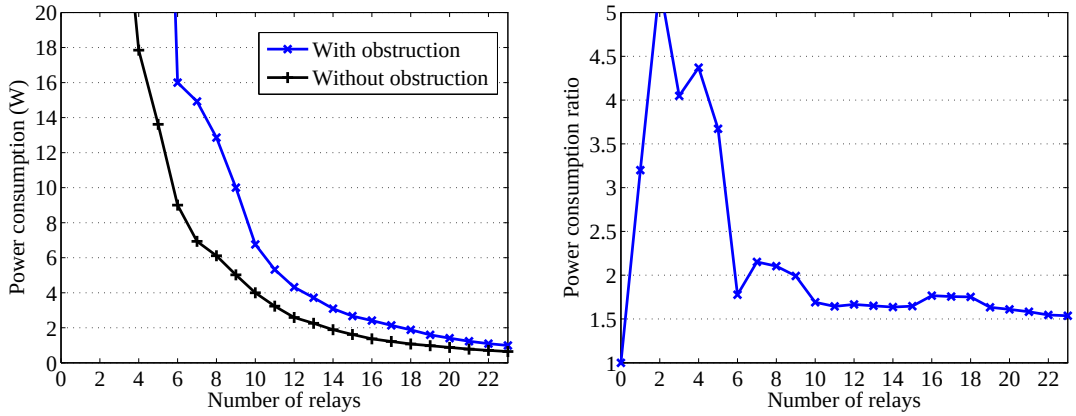
(a) $n = 0$; no relay (b) $n = 2$ (c) $n = 6$

(d) $n = 7$ (e) $n = 15$ (f) $n = 24$

Figure 4.18: Examples of simulation results for relay placement with obstacles; n denotes the number of added relays

node is at the bottom left, and the other eight source nodes are data sources. There are eight radio obstacles (grey disks), and n mobile relays (white) are inserted. It can be seen that the relay arrangement maintains connectivity in the presence of the radio obstacles.

Figure 4.19(a) shows simulation results for the total power consumption obtained by the DRP algorithm in the presence of radio obstructions arranged, as in Figure 4.18. In order to appreciate the effect of radio obstructions, simulation results without radio obstructions are also shown. As expected, the total power consumption monotonically decreases as the number of relays increases. Figure 4.19(b) shows the ratio between the power consumption values with and without radio obstructions. It can be seen that if we add about six relays, the power overhead due to radio obstructions stays almost constant with respect to the case without obstructions. Thus, we conclude that the proposed relay placement algorithm



(a) (b)

Figure 4.19: Simulation results for the power consumption of the proposed DRP algorithm with obstructions: (a) total power consumption, and (b) ratio of power consumption levels with and without obstacles

effectively reduces the total power consumption, even in the presence of obstructions, while maintaining the network connectivity.

4.6 Summary

We have investigated several key problems of placing BS and relay nodes for energy-efficient wireless networks. We modeled multiple-BS placement as a clustering optimization problem, and developed efficient heuristics for single-hop and multi-hop communication. We adopted an iterative approach similar to the K-means algorithm using the fact that the total power consumption by uplinks is a convex function of the BS location. The simulation results confirm that the proposed algorithms can substantially reduce energy consumption by MANETs.

We also have implemented a distributed placement optimization algorithm for mobile relays using the PALM method presented in Chapter 3. This algorithm discovers energy-efficient routes without exchanging global location information among nodes. Our simula-

tion results show that the power efficiency of the proposed distributed algorithm for relay placement is comparable to that of an existing centralized algorithm for relay placement. Furthermore, the proposed algorithm was shown to construct a power-efficient network structure while maintaining connectivity, even in the presence of significant radio obstruction.

CHAPTER 5

Conclusions

This chapter summarizes the major contributions of this dissertation, and discusses some possible directions for future research.

5.1 Thesis Contributions

Recent advances in communication and VLSI technology have made it possible to construct scalable wireless networks from tiny low-cost transceivers. A mobile ad hoc network (MANET) is one such network consisting of a large number of mobile transceiver nodes with no specific interconnection structure. Potential applications of MANETs include natural environment sensing and emergency communication networks in a disaster area.

In such MANET applications, the traffic pattern, node locations, and network topology can continuously change over time. Thus, for efficient operation, the data delivery paths, transmission power, and the network structure need to be appropriately controlled in an adaptive manner while taking node mobility into account.

This dissertation addresses the mathematical analysis of MANETs and the design of adaptive management schemes for power-efficient network operation. The primary contri-

butions of the research are as follows.

- A novel mobility model is proposed, the constant velocity model, which is more tractable than existing models; it is used to derive relations among link duration, link change rate, and the route stability.

- It is shown that link duration has a strong correlation with route stability, which makes it possible to quantify and compare mobility conditions for various mobility models.

- A novel power-aware link maintenance (PALM) algorithm which discovers power-efficient routes in a distributed manner is proposed. PALM consumes significantly less energy than existing mobile networks with constant transmission power.

- It is further shown that PALM needs only active nodes to exchange information about local radio conditions, and does not require all nodes to periodically transmit beacons.

- The effectiveness of the RTS/CTS in MANETs with variable transmission power is evaluated through analysis and simulation, and reasons why RTS/CTS with maximum transmission power can damage the network performance.

- Power-aware mobile base station placement is modeled as a clustering optimization problem, and efficient heuristic algorithms are proposed for solving this problem. It is shown that the algorithms reduce energy consumption by over 50%.

- A new distributed algorithm for mobile relay placement is presented. This algorithm simultaneously discovers energy-efficient routes and relay locations in the presence of radio signal obstruction.

In Chapter 2, we analyzed mathematical models of MANET node movement, and proposed a mobility metric that quantifies how mobile network nodes are. Research on MANETs often relies on simulations with various artificial random mobility models. Since different models adopt different mobility parameters and movement behaviors, it has been hard to quantify network mobility in a consistent way. For the same reason, it has been hard to compare the performance of different network control algorithms if their proponents have chosen different mobility models for evaluation purposes. In our research, we showed that link duration has a nearly invariant relationship with route lifetime, and thus is a good mobility metric.

Chapter 3 investigated the issues of transmission power control and communication link maintenance in the presence of node mobility. Unlike wired networks, the connection topology of MANETs continually changes due to node movement, and can be actively controlled through transmission power adjustment. Changes in transmission range affect communication power consumption and signal interference between transmitters. However, existing topology control schemes are mainly intended for static or pseudo-static networks, and their effectiveness in highly mobile networks has not been verified. We developed a power-aware link maintenance (PALM) algorithm, which adaptively controls the transmission power of MANET nodes, and significantly reduces their communication power needs. We also analyzed the impact of medium access control (MAC) on the network performance. In order to reduce signal collisions between nodes, the RTS/CTS handshake has been widely used with wireless networks. However, it had been previously observed that the effectiveness of the RTS/CTS handshake is limited in ad hoc networks with homogeneous transmission ranges as it reduces the communication concurrency and the network throughput. We showed that the use of the RTS/CTS handshake may adversely affect the network throughput when the nodes adjust the transmission power for DATA packets to the

minimum necessary value. We also presented a means to maximize the network throughput at a reasonable energy cost by tuning the hardware parameters of the node transceivers.

Chapter 4 investigated the base station (BS) placement optimization problem. Appropriate placement of BSs can reduce power consumption by eliminating excessively long links, and can improve network performance by reducing interference. We studied nonlinear and combinatorial optimization techniques, and applied them to the base station placement (BSP) problem. We adopted a cluster-based network structure, and developed BSP algorithms that produce power-efficient network structures. Simulation results confirm that the network structure produced by our algorithms significantly reduces the overall energy consumption. In addition, we presented a distributed relay placement (DRP) technique which places mobile relays among nodes and radio obstacles, and minimizes the total energy consumption. Simulation results show that the proposed DRP algorithm produces a network structure with energy efficiency which is comparable to that produced by an existing centralized relay placement scheme.

5.2 Future Directions

We conclude with some suggestions for future research into power-efficient management techniques for MANETs.

Link maintenance with asymmetric radio channels: PALM currently does not guarantee the existence of reverse routes from the destination to the source, and thus cannot handle TCP-style traffic. If we modify PALM's routing layer such that each node updates a reverse routing table according to the observed forward delivery route, PALM could maintain appropriate reverse routes after redirection, and thus could be used with TCP traffic. In addition, PALM assumes symmetry of the radio channel, just as PARO does. PALM

requires channel symmetry because an overhearing node cannot estimate the channel gain from itself to the responder node, while the gain in the other direction is readily available. However, PALM's key ideas could be applied with minor changes to asymmetric channel conditions by modifying the flow of redirection requests. In the current version of PALM, overhearing nodes send a redirection request to the initiator causing the initiator to change its routing table. The problem with this approach is that the radio gain from the overhearing nodes to the responder is not evaluated, and in consequence, the redirected route might be worse than the previous route when significant channel asymmetry exists. Instead, we could redesign the algorithm so that overhearing nodes send redirection requests to the responder, and the responder passes the redirection request to the initiator, together with ACK or CTS packets. With this two-step redirection process, we could guarantee that the redirected route is always more power-efficient than previous routes.

Measuring the impact of MAC parameters: In our analysis of the impact of the CSMA and the RTS/CTS handshake methods, we assumed that no radio obstruction exists. In practice, there might be significant radio obstructions among the nodes. Moreover, we focused only on the impact of the network parameters on the spatial reuse of the data communication. In many existing MAC protocols such as IEEE 802.11 [41], network nodes control their transmission time using the random backoff method. Experiments with real wireless transceivers might validate the tradeoff relation between the network throughput and the energy efficiency presented in Chapter 3, and could answer the question whether the RTS/CTS handshake should be adopted in such network environments.

Distributed node placement optimization: Though our experiments on node placement optimization focused on two-dimensional cases, the proposed algorithms can be readily adapted to 3-dimensional applications. Furthermore, the techniques used in the node placement for minimum-energy networks could be extended to solve a wider variety of

optimization problems such as: how to maximize the lifetime of the network when nodes have different battery capacities; how to determine optimal node locations without violating the topographic constraints defined by buildings, trees, etc.; and how to maximize the total throughput between nodes. A possible future research direction is to investigate modeling these problems as clustering optimization problems similar to the K-SH or the K-MH problems, taking channel bandwidth and energy capacity into account. For instance, when the remaining battery capacity of node i is given by $C(i)$, then in order to maximize the lifetime of a network with only single-hop connections, we need to solve the following optimization problem

$$\text{Minimize} \quad \max_i \left(\frac{\tau(i)}{C(i)} \cdot (k_1 \|\vec{x} - \vec{p}_i\|^\alpha + k_2) \right)$$

which can readily be converted to a constrained convex optimization problem. In addition, when multi-hop connections exist, we will need to design efficient algorithms that discover appropriate clusterhead locations, like the centroid search algorithm in Section 4.2.5.

In summary, we have analyzed various aspects of MANETs, and proposed novel algorithms that allow a power-efficient network operation in the presence of node mobility and radio obstructions. Used with the many mobile devices now coming to market, algorithms such as these make it possible to construct power-efficient and scalable MANETs, with seamless connectivity between users.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] S. Agarwal, A. Ahuja, J. P. Singh, and R. Shorey, "Route-lifetime assessment based routing (RABR) protocol for mobile ad-hoc networks," in *Proc. Int. Conference on Communications (ICC)*, vol. 3, Jun. 2000, pp. 1697–1701.

[2] A. Andreadis and G. Giambene, *Protocols for High-efficiency Wireless Networks*. Kluwer Academic Publishers, 2003.

[3] F. Bai, N. Sadagopan, and A. Helmy, "Important: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks," in *Proc. Infocom*, vol. 2, Apr. 2003, pp. 825–835.

[4] A. Balasubramanian, Y. Zhou, W. B. Croft, B. N. Levine, and A. Venkataramani, "Web search from a bus," in *Proc. Workshop on Challenged Networks (CHANTS)*, Sep. 2007, pp. 59–66.

[5] N. Banerjee, M. D. Corner, and B. N. Levine, "An energy-efficient architcture for dtn throwboxes," in *Proc. Infocom*, May 2007, pp. 776–784.

[6] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.

[7] C. Bettstetter, "Smooth is better than sharp: A random mobility model for simulation of wireless networks," in *Proc. 4th Int. Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Jul. 2001, pp. 19–27.

[8] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," *IEEE Trans. Mobile Computing*, vol. 2, pp. 257–269, Jul. 2003.

[9] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. the 3rd Int. Symp. on Mobile ad hoc networking & computing*, 2002, pp. 80–91.

[10] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," in *Proc. SIGCOMM*, 1994, pp. 212–225.

[11] D. M. Blough, M. Leoncini, G. Resta, and P. Santi, "The K-neigh protocol for symmetric topology control in ad hoc networks," in *Proc. MobiHoc*, 2003, pp. 141–152.

[12] J. Boleng, W. Navidi, and T. Camp, "Metrics to enable adaptive protocols for mobile ad hoc networks," in *Proc. Int. Conference on Wireless Networks (ICWN)*, Jun. 2002.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[14] M. Burkhart, P. von Rickenbach, R. Wattenhofer, and A. Zollinger, "Does topology control reduce interference?" in *Proc. MobiHoc*, 2004, pp. 9–19.

[15] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," in *Proc. Wireless Communication and Mobile Computing (WCMC)*, vol. 2, 2002, pp. 483–502.

[16] CAR 2 CAR Communication Consortium. http://www.car-to-car.org.

[17] Cellular News. Mobile phone subscribes pass 4 billion mark. http://www.cellular-news.com/story/35298.php?source=newsletter.

[18] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao, "Habitat monitoring: Application driver for wireless communications technology," in *Proc. SIGCOMM Workshop Data Communications in Latin America and the Caribbean*, Apr. 2001.

[19] I. Chlamtac and A. Faragó, "Making transmission schedules immune to topology changes in multi-hop packet radio networks," *IEEE/ACM Trans. Networking*, vol. 2, pp. 23–29, Feb. 1994.

[20] S. Cho and J. P. Hayes, "Impact of mobility on connection stability in ad hoc networks," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, vol. 3, Mar. 2005, pp. 1650–1656.

[21] ——, "Power-aware link maintenance (PALM) for mobile ad hoc networks," in *Proc. Local Computer Networks*, Oct. 2007, pp. 403–410.

[22] ——, "Optimizing router locations for minimum-energy wireless networks," in *Proc. Local Computer Networks*, Oct. 2008, pp. 544–546.

[23] T. H. Cormen, C. E. Leserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. MIT Press, 2001.

[24] N. A. C. Cressie, *Statistics for Spatial Data*. John Wiley & Sons, 1991.

[25] S. Cui, A. J. Goldsmith, and A. Bahai, "Enerfy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1089–1098, Aug. 2004.

[26] E. Dutkiewicz, "Impact of transmit range on throughput performance in mobile ad hoc networks," in *Proc. Int. Conf. Communications (ICC)*, vol. 9, Jun. 2001, pp. 2933–2937.

[27] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Scalable coordination in sensor networks," in *Proc. Mobicom*, Aug. 1999, pp. 263–270.

[28] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *Proc. Infocom*, vol. 1, Mar. 2004, pp. 464–475.

[29] D. K. Goldenberg, J. Lin, A. S. Morse, B. E. Rosen, and Y. R. Yang, "Towards mobility as a network control primitive," in *Proc. MobiHoc*, May 2004, pp. 163–174.

[30] J. Gomez, A. T. Campbell, M. Naghshineh, and C. Bisdikian, "Conserving transmission power in wireless ad hoc networks," in *Proc. Int. Conf. Network Protocols (ICNP)*, 2001, pp. 24–34.

[31] ——, *PARO: Supporting Dynamic Power Controlled Routing in Wireless Ad Hoc Networks*. ACM/Kluwer Journal on Wireless Networks (WINET), Sep. 2003, vol. 9.

[32] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE Trans. Networking*, vol. 10, pp. 477–486, Aug. 2002.

[33] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Information Theory*, vol. 46, pp. 388–404, Mar. 2000.

[34] J. A. Gutierrez, M. Naeve, E. Callaway, M. Bourgeois, V. Mitter, and B. Heile, "IEEE 802.15.4: A developing standard for low-power low-cost wireless pesonal area networks," *IEEE Network*, vol. 15, pp. 12–19, Sep. 2001.

[35] Z. J. Haas, "A new routing protocol for the reconfigurable wireless networks," in *Proc. Int. Conference on Universal Personal Communications (ICUPC)*, vol. 2, Oct. 1997, pp. 562–565.

[36] Y. Han, R. J. La, and A. M. Makowski, "Distribution of path durations in mobile ad-hoc networks - Palm's theorem at work," in *Proc. the 16th ITC Specialist Seminar*, Aug. 2004.

[37] Z. Han, A. L. Swindlehurst, and K. J. R. Liu, "Smart deployment/movement of unmanned air vehicle to improve connectivity in MANET," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, Apr. 2006, pp. 252–257.

[38] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 3rd ed. Springer, 2003.

[39] D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research, Volume I*. McGraw-Hill, 1982.

[40] O. C. Ibe, *Fixed Broadband Wireless Access Networks and Services*. John Wiley & Sons, 2002.

[41] IEEE. IEEE Std 802.11-2007, Local and Metropolitan Area Network, Part 11: Wireless LAN Medium Access Control and Physical Layer Specifications. http://standards.ieee.org/getieee802/download/802.11-2007.pdf.

[42] ——. IEEE Std 802.15.4-2006, Local and Metropolitan Area Network- Specific requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LR-WPANs). http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf.

[43] L. Iftode, C. Borcea, N. Ravi, P. Kang, and P. Zhou, "Smart Phones: An embedded system for universal interactions," in *Proc. Int. Workshop on Future Trends of Distributed Computing Systems*, vol. 4, May 2004, pp. 88–94.

[44] International Telecommunication Union. Key global telecom indicators for the world telecommunication service sector. http://www.itu.int/ITU-D/ict/statistics/at_glance /KeyTelecom99.html.

[45] B. Jabbari, Z. Yong, and F. Hillier, "Random walk modeling of mobility in wireless networks," in *Proc. Vehicular Technology Conference (VTC)*, vol. 1, May 1998, pp. 639–643.

[46] S. Jiand, D. He, and J. Rao, "A prediction-based link availability estimation for mobile ad hoc networks," in *Proc. Infocom*, vol. 3, Apr. 2001, pp. 1745–1752.

[47] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Proc. Mobile Computing*, vol. 353, 1996, pp. 153–181.

[48] J.-H. Ju and V. O. K. Li, "An optimal topology-transparent scheduling method in multihop packet radio networks," *IEEE/ACM Trans. Networking*, vol. 6, pp. 298–306, Jun. 1998.

[49] E.-S. Jung and N. H. Vaidya, "A power control MAC protocol for ad hoc networks," in *Proc. Mobicom*, Sep. 2002, pp. 36–47.

[50] L. Kleinrock, "Packet switching in radio channels: Part I-carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Communications*, vol. 3, pp. 1400–1975, Dec. 1975.

[51] M. Krunz, A. Muqattash, and S.-J. Lee, "Transmission power control in wireless ad hoc networks: Challenges, solutions, and open issues," *IEEE Network*, vol. 18, pp. 8–14, Sep. 2004.

[52] J. F. Kurose and K. W. Ross, *Computer Networking*, 3rd ed. Addison Wesley, 2005.

[53] B.-J. Kwak, N.-O. Song, and L. E. Miller, "A mobility measure for mobile ad hoc networks," in *IEEE Communications Letters*, vol. 7, Aug. 2003, pp. 379–381.

[54] L. Li, J. Y. Halpern, P. Bahl, Y.-M. Wang, and R. Wattenhofer, "Analysis of a cone-based distributed topology control algorithm for wireless multi-hop networks," in *Proc. Symp. Principles of Distributed Computing (PODC)*, 2001, pp. 264–273.

[55] N. Li and J. C. Hou, "FLSS: a fault-tolerant topology control algorithm for wireless networks," in *Proc. Mobicom*, 2004, pp. 275–286.

[56] N. Li, J. C. Hou, and L. Sha, "Design and analysis of an MST-based topology control algorithm," in *Proc. Infocom*, vol. 3, Mar. 2003, pp. 1702–1712.

[57] W. Li and C. G. Cassandras, "A minimum-power wireless sensor network self-deployment scheme," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, vol. 3, Mar. 2005, pp. 1897–1902.

[58] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. Berkeley Symp. Mathematical Statistics and Probability*, vol. 1. University of California Press, 1967, pp. 281–297.

[59] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. Int. Workshop Wireless Sensor Networks and Applications (WSNA)*, Sep. 2002.

[60] P. McDermott-Wells, "What is Bluetooth?" *IEEE Potentials*, vol. 23, pp. 33–35, Dec. 2005.

[61] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in *IEEE Int. Workshop on Mobile Multimedia Communications (MuMuC)*, Nov. 1999, pp. 3–10.

[62] P. C. Nar and E. Cayirci, "PCSMAC: A power controlled sensor-MAC protocol for wireless sensor networks," in *Proc. European Workshop on Wireless Sensor Networks*, Jan. 2005, pp. 81–92.

[63] A. Nayebi, A. Khosravi, and H. Sarbazi-Azad, "The impact of stationary nodes on the performance of wireless mobile networks," in *Proc. Int. Conf. Wireless and Mobile Communications (ICWMC)*, Mar. 2007.

[64] S. Nedevschi, R. K. Patra, S. Surana, S. Ratnasamy, L. Subramanian, and E. Brewer, "An adaptive high performance MAC for long-distance multihop wireless networks," in *Proc. Mobicom*, Sep. 2008, pp. 259–270.

[65] K. Ogata, *System Dynamics*, 4th ed. Prentice Hall, 2003.

[66] R. Ostrovsky, Y. Rabani, L. J. Schulman, and C. Swamy, "The effectiveness of Lloyd-type methods for the k-means problem," in *Proc. Symp. Foundations of Computer Science (FOCS'06)*, 2006, pp. 165–176.

[67] K. Paul, S. Bandyopadhyay, A. Mukherjee, and D. Saha, "Communication-aware mobile hosts in ad-hoc wireless network," in *Proc. Int. Conf. Personal Wireless Communication (ICPWC)*, Feb. 1999, pp. 83–87.

[68] M. D. Penrose, "On k-connectivity for a geometric random graph," *Random Structures & Algorithms*, vol. 15, pp. 145–164, Sep. 1999.

[69] X. Pérez-Costa, C. Bettstetter, and H. Hartenstein, "Toward a mobility metric for comparable & reproducible results in ad hoc networks research," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 4, pp. 58–60, 2003.

[70] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Computer Communication Reviews*, Oct. 1994, pp. 234–244.

[71] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *2nd IEEE Workshop on Mobile Computing Systems and Applications*, 1999, pp. 90–100.

[72] C. E. Perkins, *Ad Hoc Networking*. Addison Wesley, 2001.

[73] T. K. Philips, S. S. Panwar, and A. N. Tantawi, "Connectivity properties of a packet radio network model," *IEEE Trans. Information Theory*, vol. 35, pp. 1044–1047, Sep. 1989.

[74] D. Qiao, S. Choi, A. Jain, and K. G. Shin, "MiSer: An optimal low-energy transmission strategy for IEEE 802.11a/h," in *Proc. Mobicom*, Sep. 2003, pp. 161–175.

[75] R. Rajaraman, "Topology control and routing in ad hoc networks: a survey," *SIGACT News*, vol. 33, no. 2, pp. 60–73, 2002.

[76] R. Ramanathan and R. Rosales-Hain, "Topology control of multihop wireless networks using transmit power adjustment," in *Proc. Infocom*, Mar. 2000, pp. 404–413.

[77] T. Rappaport, *Wireless Communications: Principles and Practices*. Prentice Hall, 1996.

[78] P. Santi and D. M. Blough, "An evaluation of connectivity in mobile wireless ad hoc networks," in *Proc. International Conference on Dependable Systems and Networks (DSN)*, Jun. 2002, pp. 89–98.

[79] P. Santi, D. M. Blough, and F. Vainstein, "A probabilistic analysis for the radio range assignment problem in ad hoc networks," in *Proc. MobiHoc*, Oct. 2001, pp. 212–220.

[80] K. Sayood, *Introduction to Data Compression*, 3rd ed. Morgan Kaufman, 2005.

[81] Y. Shi and Y. T. Hou, "Theoretical results on base station movement problem for sensor network," in *Proc. Infocom*, Apr. 2008, pp. 376–384.

[82] K.-P. Shih and Y.-D. Chen, "CAPC: A collision avoidance power control MAC protocol for wireless ad hoc networks," *IEEE Commnunications Letters*, vol. 9, pp. 859–861, Sep. 2005.

[83] G. T. Sibley, M. H. Rahimi, and G. S. Sukhatme, "Robomote: A tiny mobile robot platform for large-scale ad-hoc sensor networks," in *Proc. Int. Conf. Robotics and Automation*, Sep. 2002, pp. 1143–1148.

[84] S. Sontag, "Re-kindling a love of books," *IEEE Spectrum*, vol. 45, p. 25, Jul. 2007.

[85] W. Su, S.-J. Lee, and M. Gerla, "Mobility prediction and routing in ad hoc wireless networks," *Int. Journal of Network Management*, vol. 11, pp. 3–30, 2001.

[86] D. S. Tan, S. Zhou, J.-M. Ho, J. S. Mehta, and H. Tanabe, "Design and evaluation of an individually simulated mobility model in wireless ad hoc networks," in *Proc. Communication Networks and Distributed Systems Modeling and Simulation*, 2002.

[87] A. Tanenbaum, *Computer Networks*, 4th ed. Prentice Hall, 2003.

[88] O. Tickoo, S. Raghunath, and S. Kalyanaraman, "Route fragility: A novel metric for route selection in mobile ad hoc networks," in *Proc. Int. Conf. Networks (ICON)*, Oct. 2003, pp. 537–542.

[89] O. K. Tonguz and G. Ferrari, "Is the number of neighbors in ad hoc wireless networks a good indicator of connectivity?" *IEEE Int. Zurich Seminar on Communications*, pp. 40–43, Feb. 2004.

[90] University of California, Berkeley. 29 Palms fixed/mobile experiment: Tracking vehicles with a UAV-delivered sensor network. http://www.eecs.berkeley.edu/~pister /29palms0103, 2001.

[91] ——. TinyOS Community Forum. http://www.tinyos.net.

[92] University of Southern California, Information Sciences Institute. The network simulator – ns-2. http://www.isi.edu/nsnam/ns/.

[93] Vodafone. Africa: The impact of mobile phones. http://www.vodafone.com/etc /medialib/attachments/cr_downloads.Par.78351.File.tmp/GPP_SIM_paper_3.pdf.

[94] ——. India: The impact of mobile phones. http://www.vodafone.com/etc/medialib /public_policy_series.Par.56572.File.dat/public_policy_series_9.pdf.

[95] Y. Wang and Y. J. Zhao, "Fundamental issues in systematic design of airborne networks for avation," in *Proc. Aerospace Conf.*, 2006, pp. 1–8.

[96] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang, "Distributed topology control for power efficient operation in multihop wireless networks," in *Proc. Infocom*, vol. 3, Apr. 2001, pp. 1388–1397.

[97] W. Xiand, P. Richardson, and I. Guo, "Introduction and preliminary experimental results of wireless access for vehicular environments (WAVE) systems," in *Int. Conf. Mobile and Ubiquitous Systems: Networking & Services*, Jul. 2006, pp. 1–8.

[98] K. Xu, M. Gerla, and S. Bae, "How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks?" in *Proc. Global Telecommunications Conf. (GLOBECOM)*, vol. 1, Nov. 2002, pp. 72–76.

[99] ——, "Effectiveness of RTS/CTS handshake in IEEE 802.11 based ad hoc networks," *Ad Hoc Network Journal*, vol. 1, pp. 107–123, Jul. 2003.

[100] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, pp. 169–181, Mar. 2004.

[101] C. Yu, K. G. Shin, and L. Song, "Link-layer salvaging for making routing progress in mobile ad hoc networks," in *Proc. MobiHoc*, 2005, pp. 242–254.

[102] D. Yu and H. Li, "Influence of mobility models on node distribution in ad hoc networks," in *Proc. Int. Conf. Communication Technology (ICCT)*, vol. 2, Apr. 2003, pp. 985–989.

[103] J. Zavgren, "NTDR mobility management protocols and procedures," in *Proc. Military Commun. Conf. (MILCOM)*, Nov. 1997.

[104] W. Zhao, O. Ammar, and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," in *Proc. MobiHoc*, May 2004, pp. 187–198.

[105] P. Zheng and L. M. Ni, "Spotlight: The rise of the smart phone," in *Proc. IEEE Distributed Systems Online*, vol. 7, Mar. 2006.

[106] L. Zhitang, F. Jun, N. Wei, W. Li, and C. Yuan, "A novel MAC protocol for wireless ad hoc networks with power control," in *Proc. Int. Conf. Multimedia and Ubiquitous Engineering (MUE)*, Apr. 2007, pp. 347–352.