On the Performance of Two-Hop Relay Mobile Ad Hoc Networks under Buffer Constraint

by

Jia Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(The School of Systems Information Science)
in Future University Hakodate
September 2016

To my family

ABSTRACT

On the Performance of Two-Hop Relay Mobile Ad Hoc Networks under Buffer Constraint

by

Jia Liu

Mobile ad hoc network (MANET) represents a kind of self-organizing network architecture, which consists of mobile devices communicating with each other over peer-to-peer wireless links without centralized infrastructure. Since MANETs can be deployed and reconfigured rapidly at very low cost, they are appealing for many critical applications, such as disaster relief, emergency rescue, battlefield communications, traffic offloading and cover extension for future 5G networks. To efficiently facilitate the application and commercialization of MANETs, understanding the fundamental performance of such networks is of great importance.

The available performance studies for MANETs suffer from two major limitations. First, they mainly focus on the asymptotic behaviors of network performance as the network size tends to infinity, while the actual achievable performance is largely uninvestigated. Second, to make their analysis tractable, these studies are usually based on the ideal assumption of infinite buffer, which does not hold for a practical MANET. Therefore, it is important to have a thorough study on the actual achievable performance of MANETs under the practical limited-buffer constraint.

For a general MANET with limited-buffer constraint, this thesis is devoted to exploring its actual achievable performance in terms of the throughput, end-to-end (E2E) delay and throughput capacity. We first consider the scenario with only the relay-buffer constraint, where each network node maintains a shared limited relay buffer for storing relay packets of all other nodes. For such a MANET, we develop an efficient theoretical framework to model its dynamic behaviors characterized by the buffer occupancy process, packet source-queuing process and packet delivery process. This theoretical framework is general since it applies to any distributed MAC protocol and any mobility model that leads to the uniform distribution of nodes' locations in steady state. With the help of this framework, we derive the exact expressions for both throughput capacity and expected E2E delay. Case studies are further conducted under two typical network scenarios to demonstrate the application of the proposed theoretical framework.

We then extend our study to the MANETs where both the source buffer and relay buffer are subject to the limited-buffer constraint. Based on the Queuing theory and birth-death chain theory, we develop a general theoretical framework to fully depict the occupancy processes of both source buffer and relay buffer, such that the corresponding stationary occupancy state distributions (QSDs) can be derived. With the help of OSDs, we further obtain the exact expressions of throughput, expected E2E delay and throughput capacity. Finally, extensive simulations and numerical results are presented to demonstrate the efficiency of the proposed theoretical framework and illustrate our theoretical findings. It is expected that the theoretical results developed in this thesis will provide a useful guideline for the practical design and optimization of MANETs.

ACKNOWLEDGEMENTS

Upon accomplishing my three-year doctoral career in Future University Hakodate, I would like to express my sincere thanks to all who provide me help, love and encouragement, which certainly make my experience here become one of the most important and wonderful stages that I will never forget in the rest of my life.

First and foremost, I am greatly indebted to my supervisor Professor Xiaohong Jiang, not only for his continuous guidance and support in my academic research, but also for his serving as my life mentor to teach me a lot of truth in life. During my PhD pursuit in Hakodate, Professor Jiang guided me to deal with various challenges I encountered such that I can finish this thesis. He and his wife, Mrs Li, always gave me countless care.

I owe special thanks to my wife Yang Xu. I am very sorry that we lived separately for three years due to my PhD pursuit. During this period, she always provided endless love and encouragement to me, which make my life meaningful.

I would like to give my sincere gratitude to Professor Osamu Takahashi, who gave me financial support for working as his research assistant. I also want to thank Professor Min Sheng of Xidian University, China, who serves as my co-supervisor and introduced me to Professor Jiang so that I can obtain such a valuable study opportunity.

I would also like to appreciate my thesis committee members, Professor Yuichi Fujino, Professor Hiroshi Inamura and Professor Hideki Satoh for their constructive comments which help me greatly improve the quality of my thesis. They, Professor

Osamu Takahashi and Professor Shiratori Norio of Waseda University also provided me a lot of career support. Moreover, they organized various interesting activities like cherry-blossom viewing in which I felt warmth from a big family. My thanks also go to the research colleagues Juntao Gao, Yin Chen and Yuanyu Zhang; my Japanese teachers Katsuko Takahashi, Keiko Ishikawa and Takako Shikauchi; the university staffs Tooru Yoshida and Satoko Mitobe; my dear friends Aiko Nakamura, Professor Hiroyuki Takamura, Wataru Ohtani, and Ding Yang. It is because of them my life in Japan could be so colorful.

Finally, I want to express my great acknowledgments to my parents and other family members. They always give me unconditional love and support such that I hold the courage to face anything. I love them forever.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

In this chapter, we first introduce the background of mobile ad hoc networks (MANETs), and then we present the motivations and outline of this thesis.

1.1 Background of Mobile Ad Hoc Networks

With the rapid development of wireless communication techniques in past decades, wireless networks, which can break through the connecting limitation of traditional wired networks and provide access service for mobile users, have found extensive applications in our daily life, such as the global deployed cellular networks (GSM, WCDMA, LTE), satellite communications (GPS, audio broadcasting), and wireless local area networks (Wi-Fi, Bluetooth, Zigbee) [1–3]. It is notable that the infrastructure or centralized administration systems (like base stations and satellites) play a core role in these kinds of wireless networks, which makes them highly vulnerable to artificial attacks and nature disasters. Motivated by this, a novel kind of wireless networks with distributed architecture have been proposed recently, termed as mobile ad hoc networks (MANETs).

A mobile ad hoc network can be defined as a collection of self-autonomous mobile devices which communicate with each other via peer-to-peer wireless channels without any support from pre-established infrastructure [4, 5]. In a MANET, each network

node serves not only as a source or destination but also as a relay to help other nodes forward their data, such that the traffic flows in the network can be delivered in a cooperative and distributed way.

Compared with the above existing wireless network architectures, MANETs provide many appealing features such that they have attracted considerable attention from both the academic and industrial communities. First, MANETs can be rapidly deployed at low cost, because they do not rely on the existence of infrastructures which usually incur an extremely high cost, and plenty of mobile devices, like mobile phones, wireless sensors, portable computers can be easily collected to serve as the network nodes. Second, due to their distributed and self-organized nature, MANETs are highly robust in the sense that they can tolerate severe node failure problem. Finally, MANETs can be flexibly extended and quickly reconfigured, since mobile nodes can join, roam around and leave the network freely, and the corresponding reconfiguration information can be quickly spread throughput the network by flooding broadcast.

Thanks to these attractive features of MANETs, they are highly promising for a lot of critical applications, such as military communications, disaster rescue, environment monitoring, daily information exchange, etc. MANETs are also expected to implement the D2D communications for traffic offloading and coverage extension in future cellular networks [6]. Thus, it is believed that MANETs will become an indispensable component among the future heterogeneous wireless network environments.

1.2 Motivations

To facilitate the application and commercialization of MANETs, a thorough understanding on the performance limits of such networks is of great importance [7, 8]. Serving as the two most fundamental performance metrics, throughput and delay have been extensively explored in literature [9–30]. However, the existing perfor-

mance studies for MANETs suffer from two major limitations:

- First, the available performance studies for MANETs mainly focus on the scaling law results [31], while the exact performance analysis remains largely untouched. The term of "scaling law" which also corresponds to "order sense", usually appears together with notations $(\Theta, O, \Omega, o, \omega)$ [32] to characterize the asymptotic behaviors of throughput or delay as the network size tends to infinity. Although scaling law results are helpful to grasp the general performance trend of MANETs, they provide little insight into the exact achievable network performance. In practice, however, a thorough understanding of the exact achievable performance is of more importance for the network design and optimization, and thus is of great interest for network engineers.

- Second, to make the theoretical analysis tractable, they are usually based on some ideal assumptions. In particular, they all assume that the buffer of each network node, which is used for temporarily storing the packets waiting to be sent, has an infinite buffer size. In a practical MANET, however, this assumption can not hold because the buffer size of a mobile device is usually bounded due to both its storage space limitation and computing capability limitation. Therefore, for the practical performance study of MANETs, the constraint on buffer size should be carefully addressed.

To address above limitations and promote a significant progress in the performance study of MANETs, this thesis is devoted to exploring the exact fundamental performance, i.e., achievable throughput, end-to-end (E2E) delay and throughput capacity, of a general MANET with limited-buffer constraint. We first consider a MANET where the relay buffer of each node for storing the packets of other nodes is limited, and develop a theoretical framework to characterize the dynamics in such a MANET, which enables us to derive the exact expressions of throughput capacity

and expected E2E delay. We then extend our results to the MANETs with both source/relay-buffer constraint, and provide theoretical analysis to reveal the insights into the impacts of buffer constraint on network performance.

1.3 Thesis Outline

The remainder of this thesis is outlined as follows:

Chapter II Related Works. In this chapter, we present previous works related to the performance study of MANETs.

Chapter III Preliminaries. This chapter introduces the system models, routing scheme, performance metrics and notations involved in this thesis.

Chapter IV Throughput Capacity of MANETs under Relay-Buffer Constraint. In this chapter, we explore the throughput capacity of MANETs under relay-buffer constraint. To address this technical issue, we develop a theoretical framework to capture the relay buffer occupancy process, such that the closed-form expression of exact throughput capacity to be derived. We also explore the corresponding capacity optimization issue. Finally, extensive simulation and numerical results are provided to validate the efficiency of our framework and to show the impacts of relay-buffer constraint on throughput capacity.

Chapter V End-to-End Delay of MANETs under Relay-Buffer Constraint. In this chapter, we study the end-to-end (E2E) delay performance of MANETs under relay-buffer constraint. Combining the buffer occupancy process analysis proposed in Chapter IV, we apply the fixed-point theory to solve the stationary occupancy state distribution of the relay buffer. Based on this, we develop an absorbing Markov chain model to characterize the packet delivery process, and further derive the exact expressions for the expectations of source-queuing delay, delivery delay and E2E delay. Finally, we present extensive simulation and numerical results to illustrate the efficiency of our delay analysis as well as the impacts of network

parameters on delay performance.

Chapter VI Throughput and Delay of MANETs under General Buffer Constraint. In this chapter, we explore the throughput and delay of MANETs under the general limited-buffer constraint, where each network node maintains a limited source buffer to store its locally generated packets and also a limited shared relay buffer to store relay packets for other nodes. Based on the Queuing theory and birth-death chain theory, we first develop a general theoretical framework to fully depict the source/relay buffer occupancy process in such a MANET. With the help of this framework, we then derive the exact expressions of several key network performance metrics, including achievable throughput, expected E2E delay and throughput capacity. Finally, we provide extensive simulation and numerical results to demonstrate the application and efficiency of our theoretical framework, as well as to illustrate our theoretical findings.

Chapter VII Conclusion. This chapter concludes the whole thesis by summarizing our contributions on the performance study of MANETs under buffer constraint and also discussing potential directions for the future research.

CHAPTER II

Related Works

In this chapter, we introduce the related works of the performance study of MANETs. We first present the studies without buffer constraint, and then present some initial studies with buffer constraint.

2.1 Studies without Buffer Constraint

2.1.1 Scaling Laws

Since the pioneer work of Grossglauser and Tse [9], the scaling laws of throughput capacity and delay-throughput tradeoff have been extensively studied for MANETs under various network scenarios. Grossglauser and Tse first demonstrated that with the help of node mobility, a $\Theta(1)$ per node throughput is achievable in a MANET with the two-hop relay routing scheme (for the definitions of asymptotic notations $(\Theta, O, \Omega, o, \omega)$, please kindly refer to [32]), which indicates that the per node throughput can keep constant as the number of nodes in the MANET tends to infinity. Although Grossglauser and Tse didn't explore the corresponding delay performance, they pointed out that the constant per node throughput is achieved at a cost of large delay.

Neely and Modiano explored in [11] the delay-throughput tradeoff for a cell-

partitioned MANET under the independent and identically distributed (i.i.d) mobility model. They showed that achievable delay-to-throughput ratio is lower bounded as $delay/throughput \geq O(n)$, where n is the number of network nodes. Gamal *et al.* investigated in [12] that the optimal scaling behavior of the delay-throughput tradeoff under a symmetric random walk mobility model, and demonstrated that a $\Theta(n \log n)$ average packet delay is incurred to achieve the $\Theta(1)$ per node throughput there. Sharma *et al.* further explored in [13] the delay-throughput tradeoff under a general and unified mobility model, and indicated that node mobility can not be applied to increase throughput capacity if the delay is below some critical value. The scaling laws of throughput capacity and delay-throughput tradeoff have also been studied under other mobility models, such as Brownian mobility model in [10, 14], restricted mobility model in [15] and correlated mobility model in [19].

The scaling law studies on the performance of MANETs under various network scenarios can be also found in [17, 18, 20–30, 33]. Specifically, the works of [17, 20, 22, 23, 29, 30] explored the scaling laws of MANETs with multicast traffic. The capacity region of MANETs have been studied in [21, 33]. The works of [25, 26] studied the scaling laws of throughput and delay for MANETs with the infrastructure. Recently, the capacity scaling laws of MANETs with the consideration of security performance have been reported in [18, 24, 28]. For a survey on the scaling law results of MANETs, please kindly refer to [31] and the references therein.

2.1.2 Exact Results

To break through the limitation of scaling law results, some preliminary works have been conducted to derive the exact expressions of throughput and delay for MANETs [11, 34–40]. Mergen and Tong [34] derived the throughput capacity in closed-form for the regular Manhattan and ring networks. Neely and Modiano derived in [11] the exact expressions of throughput capacity and expected E2E delay for a cell-partitioned

MANET with 2HR routing scheme and i.i.d. mobility model. Following this line, Gao *et al.* extended the results of [11] to the network scenario with a group-based scheduling scheme in [35], Chen derived the approximations for exact throughput capacity of MANETs with ALOHA protocol in [36] and the exact throughput capacity of intermittently connected mobile networks in [37]. Chen *et al.* also explored in [38] the exact throughput capacity of MANETs with directional antennas.

Regarding the studies on exact delay performance, Jindal and Psounis [39] derived the approximations of E2E delay for MANETs with multi-hop relay routing. For a cell-partitioned MANET with broadcast routing scheme, Gao *et al* [40] developed a theoretical framework to derive the exact expressions of the E2E delay.

2.2 Studies with Buffer Constraint

By now, some initial works have been reported on the performance study of MANETs with the consideration of buffer constraint [41–44]. Specifically, Herdtner and Chong explored in [41] the scaling law of throughput-storage tradeoff for MANETs under the relay-buffer constraint and indicated that the throughput capacity scales as $O(\sqrt{\frac{b}{n}})$, where b denotes the relay buffer size. Subramanian and Fekri [42] explored the throughput capacity in a delay-tolerant network with the relay-buffer constraint and negligible wireless interference. Gao *et al.* [43] considered a MANET with limited source buffer in each node, and derived the corresponding cumulative distribution function of the source delay under a f-cast dispatching scheme. Recently, Fang *et al.* [44] considered the buffer constraint that each network node maintains a limited source buffer and $n-2$ dedicated limited relay buffers for other nodes (one limited relay buffer for one node), and derived the exact expressions of throughput and expected E2E delay for MANETs with a specific routing scheme.

It is notable that in [43], Gao *et al.* only analyzed the dispatching process in source node rather than through the network, thus only the source buffer model is

considered for the calculating of source delay, which serves as a part of the end-to-end delay. In [44], the dedicated relay buffer model makes the relay buffer size tends to infinite as the network size increases. To explore the performance of MANETs under more realistic buffer models, in my thesis we first consider the relay-buffer constraint where each network node maintains a shared limited relay buffer, and then extend to the scenario where both the source buffer and relay buffer are subject to the limited-buffer constraint. The details of our buffer models will be elaborated in Chapter IV and Chapter VI.

CHAPTER III

Preliminaries

In this chapter, we first introduce the general system models and the basic two-hop relay routing scheme for packet delivery. Then we present the fundamental performance metrics and main notations involved in this thesis.

3.1 System Models

3.1.1 Network Model

In this thesis, we consider a time-slotted MANET, which consists of n nodes randomly moving in a torus network area following a "uniform type" mobility model [9]. With such mobility model, the location process of a node is stationary and ergodic with stationary distribution uniform on the network area, and the trajectories of different nodes are independent and identically distributed. It is notable that this "uniform type" mobility model covers many typical mobility models as special cases, such as the i.i.d model [11, 21, 45], random walk model [12], random way-point model [30] and random direction model [46].

Due to the broadcast nature of wireless channel, if two nodes reside in near area, they can not transmit simultaneously since the serious wireless interference will be caused to destroy both their transmitted information. To deal with the wireless

interference and avoid transmission collisions, media access mechanism should be adopted. In this thesis, we mainly consider two media access mechanisms, the cell-based mechanism in Chapter IV, and the classical DCF-style mechanism in Chapter V and Chapter VI. The details of the mechanisms are elaborated in the following chapters.

3.1.2 Traffic Model

The widely-used permutation traffic model [9, 11, 19] is adopted for characterizing the composition of traffic flows in the MANETs. With this traffic model, there are n unicast traffic flows, each node is the source of one traffic flow and also the destination of another traffic flow. More formally, let $\varphi(i)$ denote the destination node of the traffic flow originated from node i, then the source-destination pairs are matched in a way that the sequence $\{\varphi(1), \varphi(2), \cdots, \varphi(n)\}$ is just a derangement of the set of nodes $\{1, 2, \cdots, n\}$. For example, two typical kinds of source-destination pairs can be constituted as follows: $1 \to 2$, $2 \to 3$, $\cdots$, $n-1 \to n$, $n \to 1$; $1 \leftrightarrow 2$, $3 \leftrightarrow 4$, $\cdots$, $n-1 \leftrightarrow n$ (here n is even). The packet generating process at each network node is assumed to be a Bernoulli process [47] with mean rate λ, so that with probability λ a new packet is generated by its source node in each time slot.

3.2 Two-Hop Relay Routing Scheme

Regarding the routing algorithm for packet delivery, we consider the two-hop relay (2HR) scheme which serves as a class of attractive routing algorithms for MANET [48], since it can be implemented easily in a distributed way yet efficient in the sense that it has the capability of achieving the throughput capacity for many important MANET scenarios [9, 11, 35, 36]. Here we introduce the original 2HR scheme proposed in [9], based on which we can develop the improved 2HR schemes according to the specific MANET scenarios.

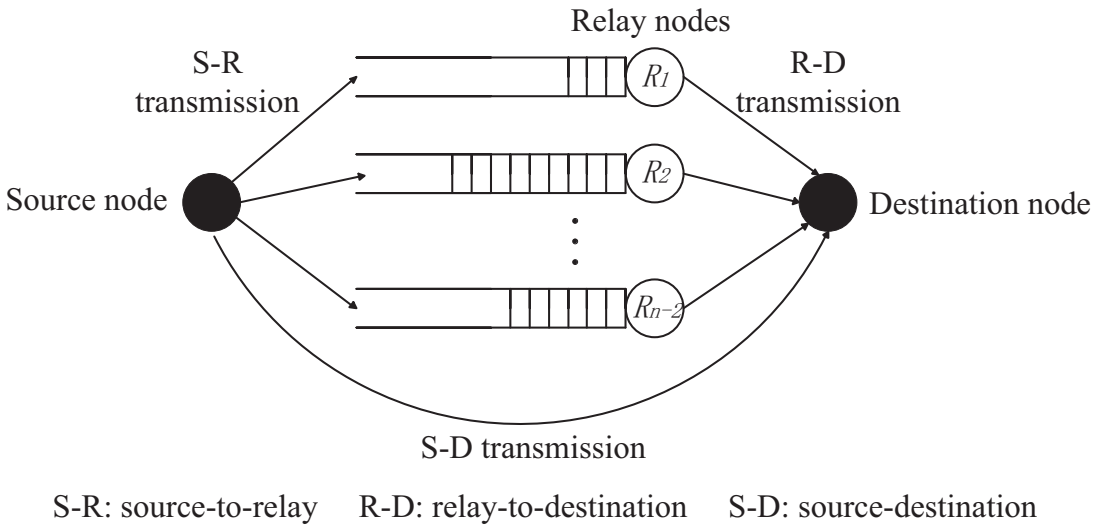


Figure 3.1: Illustration of two-hop relay routing scheme.

The original 2HR routing algorithm can be illustrated in Fig. 3.1. With the 2HR scheme, when a node S gets access to the wireless channel in a time slot, it will transmit a packet directly to its destination node D (S-D transmission) if D is within its transmission range; otherwise with probability 0.5, S selects to transmit a self-generated packet to a relay node R (S-R transmission), or deliver a packet of other nodes to the corresponding destination (R-D transmission). The detailed 2HR routing algorithm can be summarized in Algorithm 1.

Algorithm 1 2HR algorithm

1: **if** The destination D is within the transmission range of S **then**
2: S executes **Procedure** 1.
3: **else if** There exist other nodes within the transmission range of S **then**
4: With equal probability, S selects one node as the receiver.
5: S executes **Procedure** 2 or **Procedure** 3 equally with the receiver.
6: **end if**

3.3 Performance Metrics

The fundamental performance metrics involved in this thesis are defined as follows.

Procedure 1 Source-to-destination (S-D) transmission

1: **if** S has packets in its source queue **then**
2: S transmits the head-of-line (HoL) packet in its source queue to D.
3: S removes the HoL packet from its source queue.
4: S moves ahead the remaining packets in its source queue.
5: **else**
6: S remains idle.
7: **end if**

Procedure 2 Source-to-relay (S-R) transmission

1: **if** S has packets in its source queue **then**
2: S transmits the HoL packet in its source queue to the receiver.
3: S removes the HoL packet from its source queue.
4: S moves ahead the remaining packets in its source queue.
5: **else**
6: S remains idle.
7: **end if**

Procedure 3 Relay-to-destination (R-D) transmission

1: **if** S has packets destined to the receiver **then**
2: S transmits the HoL packet in its corresponding relay queue to the receiver.
3: S removes the HoL packet from this relay queue.
4: S moves ahead the remaining packets in this relay queue.
5: **else**
6: S remains idle.
7: **end if**

Throughput: The *throughput* T of a flow (in units of packets per slot) is defined as the time-average number of packets that can be delivered from its source to its destination.

Throughput Capacity: For the homogeneous network scenario considered in this thesis, the network level *throughput capacity* T_c can be defined as the maximal achievable per flow throughput. Since the total amount of data that can be transmitted by a node in a time slot is normalized to one packet, then we have

$$T_c = \max_{\lambda \in (0,1]} T. \tag{3.1}$$

End-to-end Delay: The *end-to-end (E2E) delay* D of a packet (in units of time

slots) is defined as the time it takes the packet to reach its destination after it is generated by its source, and we use $\mathbb{E}\{D\}$ to denote the expectation of D. It is notable that for the calculation of E2E delay, we only focus on the packets which have been successfully delivered to their destinations, i.e., the dropped packets are not included in the calculation.

3.4 Notations

The main notations of this thesis are summarized in Table 3.1.

Table 3.1: Main notations

Symbol	Quantity
n	number of network nodes
B_s	source buffer size
B_r	relay buffer size
λ	packet generating rate
μ_s	mean service rate of the source queue
T	per-flow throughput
T_c	throughput capacity
T_c^*	optimal throughput capacity
D	end-to-end delay
D_{sq}	source-queuing delay
D_d	delivery delay
L_s	average number of packets in a source buffer
L_s^*	average number of packets in a source buffer conditioned on that the source buffer is not full
L_r^*	average number of packets in a relay buffer conditioned on that the relay buffer is not full
p_{sd}	probability that a node selects to do S-D transmission
p_{sr}	probability that a node selects to do S-R transmission
p_{rd}	probability that a node selects to do R-D transmission
p_o	relay-buffer overflowing probability
$\mathbf{\Pi_s}$	stationary occupancy state distribution of source buffer
$\mathbf{\Pi_r}$	stationary occupancy state distribution of relay buffer
α	transmission control parameter
m	cell-partitioned parameter
ν	transmission range of a node in the MANET with EC-MAC
ε	spatial multiplexing parameter in the MANET with EC-MAC
Δ	guard factor of the protocol model

CHAPTER IV

Throughput Capacity of MANETs under Relay-Buffer Constraint

As a first step towards the practical performance evaluation of MANETs, in this chapter we consider the relay-buffer constraint and develop a general theoretical framework for the exact throughput capacity study. To support efficient operation for MANETs with relay-buffer constraint, we propose an improved 2HR algorithm which incorporates both a transmission control mechanism and a feedback mechanism. For such a MANET, we first present analysis to reveal how its throughput capacity is determined by the relay-buffer overflowing probability (ROP). Based on the birth-death chain model, we then develop a novel theoretical framework to fully characterize the occupancy process of the relay buffer, such that the exact throughput capacity can be derived in closed-form. We further conduct case studies under two typical network scenarios to illustrate the application of our framework, and to explore the corresponding capacity optimization issue.

4.1 Relay-Buffer Constraint

As illustrated in Fig. 4.1, we consider the relay-buffer constraint same as that of previous studies on buffer-limited wireless networks [41, 49], where each node main-

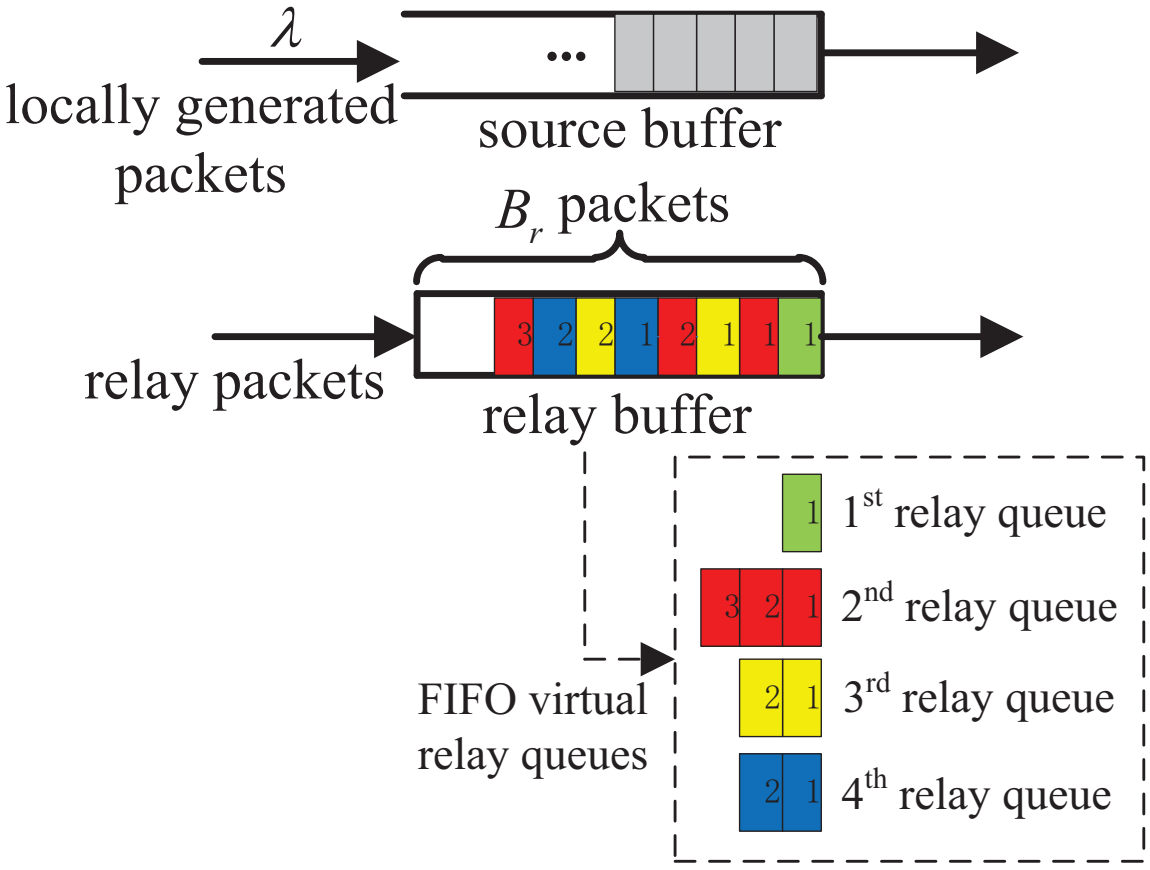


Figure 4.1: Illustration of buffer structure.

tains an infinite source buffer and a limited relay buffer of size B_r. The source buffer is for storing the packets of its own flow (locally generated packets) and works as a first-in-first-out (FIFO) source queue [50], while the relay buffer is for storing packets of all other $n-2$ flows and works as $n-2$ FIFO virtual relay queues (one queue per flow). When a packet of other flows arrives and the relay buffer is not full, a buffer space is dynamically allocated to the corresponding relay queue for storing this packet; once a head-of-line (HoL) packet departs from its relay queue, this relay queue releases a buffer space to the common relay buffer.

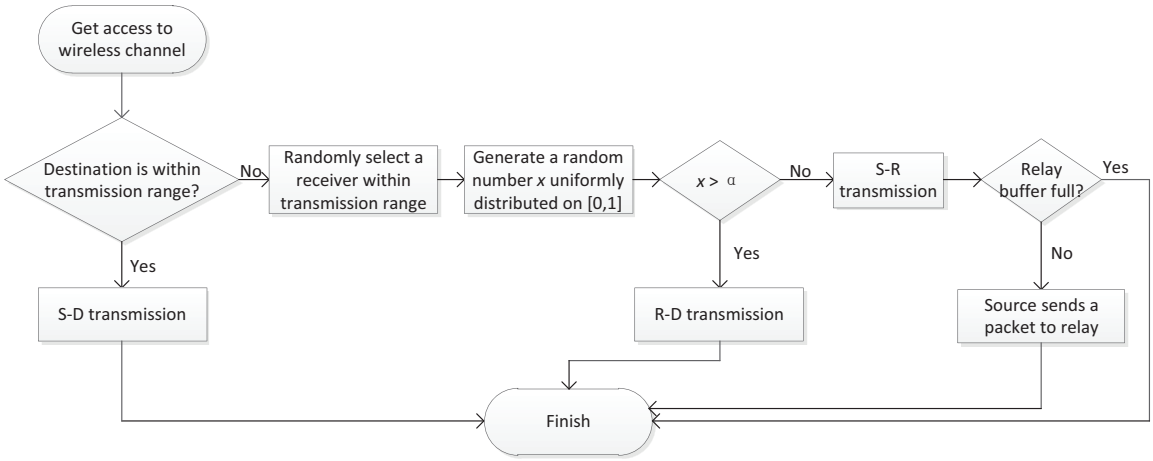


Figure 4.2: Flow chart of 2HR-α routing scheme.

4.2 Routing Scheme

Notice that under the limited relay-buffer constraint, when node S executes the S-R transmission while the relay buffer of the receiver is full, then the transmission will not be successful and the transmitted packet will be lost. To facilitate the operation of MANETs with relay-buffer constraint and improve the throughput performance, we adopt here the 2HR-α scheme as illustrated in Fig. 4.2, which is an extension of the 2HR scheme described in Section 3.2 in the following two aspects.

First, we introduce a parameter α to flexibly control the probability that S selects to conduct S-R transmission, i.e., when S gets access to the wireless channel and its destination node D is not within its transmission range, S selects to transmit a self-generated packet to a relay with probability α, and deliver a packet of other nodes to the corresponding destination with probability $1-\alpha$. Thus, α represents the level of selfishness of a node, from 0 (fully selfless) to 1 (fully selfish), and it is expected that α should be set appropriately according to the network settings to achieve the optimal throughput performance.

Second, to avoid the unnecessary packet loss in S-R transmission, the 2HR-α scheme further adopts a feedback mechanism to confirm the relay-buffer occupancy

state of a receiver. When node S selects to conduct the S-R transmission with a relay node R, R first sends a feedback to S to indicate its relay-buffer state. If the relay buffer of R is not full, S then transmits a packet to R; else S remains idle.

4.3 Throughput Capacity Analysis

For a MANET with relay-buffer constraint and 2HR-α scheme, we use p_{sd}, p_{sr} and p_{rd} to denote the probabilities that in a time slot a node gets access to the wireless channel and selects to execute S-D, S-R and R-D transmission respectively, and use $p_o(\lambda)$ to denote the relay-buffer overflowing probability (ROP) that the relay buffer of a node is full given the packet generating rate λ. With the help of these basic probabilities, we can establish the following theorem regarding the throughput capacity of the network.

Theorem IV.1 *For a MANET with relay-buffer constraint and 2HR-α scheme, its throughput capacity T_c is determined as*

$$T_c = p_{sd} + p_{sr}(1 - p_o(\tilde{\lambda})) \qquad packets/slot, \tag{4.1}$$

where $\tilde{\lambda}$ is the unique solution of the following equation

$$\lambda = p_{sd} + p_{sr}(1 - p_o(\lambda)). \tag{4.2}$$

Proof 1 *To prove the theorem, we first demonstrate that there exists an unique solution $\tilde{\lambda}$ for the equation (4.2), and then show that the throughput is λ when $\lambda < \tilde{\lambda}$, but the throughput is always $\tilde{\lambda}$ when $\lambda \geq \tilde{\lambda}$.*

From our system models in Section 3.1, it is clear that each traffic flow experiences the same service process without priority, so the behavior of each flow is identical and we can focus on a tagged flow in our analysis. Under the 2HR-α routing scheme, the

transmission of a packet from its source to destination involves at most two stages. The first stage is the source-queuing process at its source node, while the second stage is the delivery process through one relay node if the packet is not directly transmitted to its destination.

Regarding the first stage source-queuing process, the source buffer can be modeled as a Bernoulli/Bernoulli queue with arrival rate λ and service rate $\mu_s(\lambda)$ determined as

$$\mu_s(\lambda) = p_{sd} + p_{sr}\left(1 - p_o(\lambda)\right). \tag{4.3}$$

We can easily see that: 1) when $\lambda = 0$, we have $p_o(0) = 0$, so $\mu_s(0) = p_{sd}+p_{sr} > \lambda$; 2) as λ increases, $p_o(\lambda)$ tends to increase, leading to a decrease in $\mu_s(\lambda)$; 3) when $\lambda = p_{sd}+p_{sr}$, we have $p_o(\lambda) > 0$, $\mu_s(\lambda) < \lambda$. Based on these properties of service rate $\mu_s(\lambda)$, we know that there exists an unique $0 < \tilde{\lambda} < p_{sd} + p_{sr}$ such that $\tilde{\lambda} = \mu_s(\tilde{\lambda})$.

Considering a time interval $[0, t]$, we let $m_0(t)$ and $m_1(t)$ denote the number of packets being buffered in all source queues and all relay queues at time slot t, respectively. Since the total number of locally generated packets during this interval is $n\lambda t$, then the throughput T is determined as

$$T = \lim_{t\to\infty} \frac{n\lambda t - m_0(t) - m_1(t)}{n \cdot t}. \tag{4.4}$$

Since the relay buffer of each node has a fixed size B_r, then $\frac{m_1(t)}{n} \leq B_r$ and $\lim_{t\to\infty} \frac{m_1(t)}{n\cdot t} = 0$.

For the case $\lambda < \tilde{\lambda}$, we let L_s denote the queue length of source queue, then its expectation $\mathbb{E}\{L_s\}$ is given by [47]

$$\mathbb{E}\{L_s\} = \frac{\lambda - \lambda^2}{\mu_s(\lambda) - \lambda}. \tag{4.5}$$

Since when $\lambda < \tilde{\lambda}$, we have $\mu_s(\lambda) > \lambda$, so the queue length $\mathbb{E}\{L_s\}$ is bounded in this

case. Thus, $\lim_{t\to\infty} \frac{m_0(t)}{n\cdot t} = 0$ *and* $T = \lambda$.

When $\lambda \geq \tilde{\lambda}$*, then* $\mu_s(\lambda) < \lambda$ *which leads to an increasing number of packets buffered in the source queues. By applying the law of large numbers [51], we have that as* $t \to \infty$

$$\frac{m_0(t)}{t} \xrightarrow{a.s.} n(\lambda - \mu_s(\tilde{\lambda})).$$

Based on (4.4), we then have $T = \tilde{\lambda}$ *when* $\lambda \geq \tilde{\lambda}$.

Thus, the throughput capacity T_c *of the concerned network is determined as*

$$T_c = \mu_s(\tilde{\lambda}) = p_{sd} + p_{sr}\left(1 - p_o(\tilde{\lambda})\right).$$

4.4 Theoretical Framework

The result in Theorem IV.1 indicates that for the throughput capacity analysis of the concerned MANET, we need to determine the ROP $p_o(\lambda)$. To address this issue, in this section we propose our theoretical framework which utilizes a birth-death chain model to depict the complicated occupancy process of a relay buffer, such that the exact expressions of ROP and exact throughput capacity of the concerned MANET can be derived.

4.4.1 Birth-Death Chain Model

Regarding the source queue of a node S, it can be modeled as a Bernoulli/Bernoulli queue [47] with packet arrival rate λ and service rate $\mu_s(\lambda)$, where $\mu_s(\lambda)$ is given by equation (4.3). Due to the reversibility of Bernoulli/Bernoulli queue, the packet departure process of source queue is also a Bernoulli process with rate λ. Regarding the relay buffer in node S, let X_t denote the number of packets in the relay buffer at time slot t, then the occupancy process of the relay buffer can be regarded as a stochastic process $\{X_t, t = 0, 1, 2, \cdots\}$ on state space $\{0, 1, \cdots, B_r\}$. Notice that

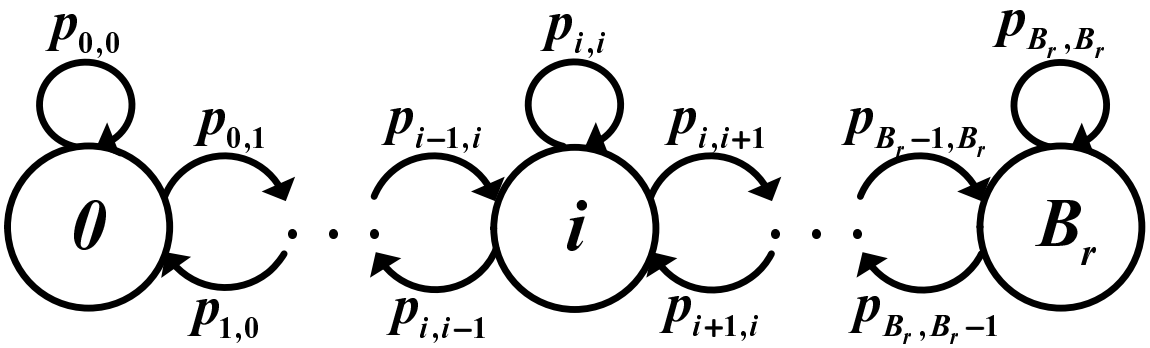


Figure 4.3: State machine of the birth-death chain.

when S serves as a relay in a time slot, the S-R transmission and R-D transmission will not happen simultaneously. Thus, suppose that the relay buffer is at state i in the current time slot, only one of the following transition scenarios may happen in the next time slot:

- i to $i+1$ $(0 \leq i \leq B_r - 1)$: the relay buffer is not full, and a packet arrives at the relay buffer.
- i to $i-1$ $(1 \leq i \leq B_r)$: the relay buffer is not empty, and a packet departures from the relay buffer.
- i to i $(0 \leq i \leq B_r)$: no packet arrives at and departures from the relay buffer.

Let $p_{i,j}$ denote the one-step transition probability from state i to state j $(0 \leq i, j \leq B_r)$, then the occupancy process $\{X_t, t = 0, 1, 2, \cdots\}$ can be modeled as a birth-death chain as illustrated in Fig. 4.3. Let $\pi_r(i)$ denote the probability that there are i packets occupying the relay buffer in the stationary state, the stationary occupancy state distribution (OSD) of the relay buffer $\mathbf{\Pi}_r = [\pi_r(0), \pi_r(1), \cdots, \pi_r(B_r)]$ is determined as

$$\mathbf{\Pi}_r \cdot \mathbf{P} = \mathbf{\Pi}_r, \tag{4.6}$$

$$\mathbf{\Pi}_r \cdot \mathbf{1} = 1, \tag{4.7}$$

where $\mathbf{P}$ is the one-step transition matrix of the birth-death chain defined as

$$\mathbf{P} = \begin{bmatrix} p_{0,0} & p_{0,1} & & & \\ p_{1,0} & p_{1,1} & p_{1,2} & & \\ & \ddots & \ddots & \ddots & \\ & & p_{B_r,B_r-1} & p_{B_r,B_r} \end{bmatrix}, \tag{4.8}$$

and $\mathbf{1}$ is a column vector of size $(B_r + 1) \times 1$ with all elements being 1.

Notice that $p_{0,0} = 1 - p_{0,1}$, $p_{B_r,B_r} = 1 - p_{B_r,B_r-1}$ and $p_{i,i} = 1 - p_{i,i-1} - p_{i,i+1}$ for $0 < i < B_r$, the expressions (4.6)−(4.8) indicate that to derive $\mathbf{\Pi}_r$, we need to determine the one-step transition probabilities $p_{i,i+1}$ and $p_{i,i-1}$.

Lemma 1 *For the birth-death chain in Fig. 4.3, its one-step transition probabilities $p_{i,i+1}$ and $p_{i,i-1}$ are determined as*

$$p_{i,i+1} = \rho_s(\lambda) \cdot p_{sr}, \ 0 \leq i \leq B_r - 1, \tag{4.9}$$

$$p_{i,i-1} = p_{rd} \cdot \frac{i}{n - 3 + i}, \ 1 \leq i \leq B_r, \tag{4.10}$$

where $\rho_s(\lambda) = \dfrac{\lambda}{\mu_s(\lambda)} = \dfrac{\lambda}{p_{sd} + p_{sr}(1 - p_o(\lambda))}$.

Proof 2 *The proof is given in Appendix A.1.*

4.4.2 Derivation of Throughput Capacity

Based on above birth-death chain based framework, we now provide analysis on the exact throughput capacity T_c, as summarized in following theorem.

Theorem IV.2 *For a concerned MANET with n mobile nodes, where each node is allocated with a relay buffer of fixed size B_r and the 2HR-α scheme is adopted for*

packet delivery, the throughput capacity T_c is determined as

$$T_c = p_{sd} + p_{sr}\left(1 - \frac{\mathrm{C}_{B_r} \cdot \beta^{B_r}}{\sum_{k=0}^{B_r} \mathrm{C}_k \cdot \beta^k}\right), \tag{4.11}$$

where $\mathrm{C}_k = \binom{n-3+k}{k}$ *and* $\beta = \frac{p_{sr}}{p_{rd}} = \frac{\alpha}{1-\alpha}$.

Proof 3 *By substituting (4.9) and (4.10) into (4.6) and (4.7), we can see that the stationary OSD of the relay buffer is determined as*

$$\pi_r(i) = \frac{\mathrm{C}_i \cdot \beta^i \cdot \rho_s(\lambda)^i}{\sum_{k=0}^{B_r} \mathrm{C}_k \cdot \beta^k \cdot \rho_s(\lambda)^k}, \tag{4.12}$$

It is notable that the relay buffer overflows when it is at state B_r, then the critical self-mapping function for $p_o(\lambda)$ is constructed as

$$p_o(\lambda) = f\left(p_o(\lambda)\right) = \pi_r(B_r) = \frac{\mathrm{C}_{B_r} \cdot \beta^{B_r} \cdot \rho_s(\lambda)^{B_r}}{\sum_{k=0}^{B_r} \mathrm{C}_k \cdot \beta^k \rho_s(\lambda)^k}. \tag{4.13}$$

From Theorem IV.1 we know that as λ approaches $\tilde{\lambda}$, $\rho_s(\lambda)$ tends to 1. Substituting $\rho_s(\tilde{\lambda}) = 1$ into (4.13), we have

$$p_o(\tilde{\lambda}) = \frac{\mathrm{C}_{B_r} \cdot \beta^{B_r}}{\sum_{k=0}^{B_r} \mathrm{C}_k \cdot \beta^k}. \tag{4.14}$$

The formula (4.11) then follows by substituting (4.14) into (4.1).

Based on Theorem IV.2, we have the following corollaries (See A.2 for the proofs).

Corollary 1 *For a network with $n \geq 3$, its throughput capacity T_c increases as relay buffer size B_r grows.*

Corollary 2 *With the setting of $\alpha = 0.5$, i.e., each node executes S-R and R-D*

transmission with equal probability, T_c is determined as

$$T_c = p_{sd} + p_{sr}\frac{B_r}{n-2+B_r} \tag{4.15}$$

Corollary 3 *When the relay buffer size B_r tends to infinity, the throughput capacity T_c is determined as*

$$T_c|_{B_r\to\infty} = \begin{cases} p_{sd} + p_{sr}, & \alpha \le 0.5 \\ p_{sd} + p_{rd}, & \alpha > 0.5 \end{cases} \tag{4.16}$$

4.5 Case Studies

The results in Theorem IV.2 indicate that by applying our theoretical framework to evaluate the throughput capacity of a MANET with relay-buffer constraint, we only need to determine the basic probabilities p_{sd}, p_{sr} and p_{rd}, which are further related to the specific network configurations. To demonstrate the application of our framework for throughput capacity analysis, in this section we provide case studies under two typical network scenarios, i.e., the cell-partitioned MANETs with local scheduling based MAC (LS-MAC) and with equivalence class based MAC (EC-MAC). The corresponding throughput capacity optimization issue will be also explored.

4.5.1 Cell-Partitioned MANET with LS-MAC

We first consider a cell-partitioned MANET with LS-MAC, which is widely adopted in available works [11, 13, 22, 23]. As illustrated in Fig. 4.4, the whole network area is evenly partitioned into $m \times m$ non-overlapping cells. In each time slot one cell supports only one transmission between two nodes within it, and the concurrent transmissions in different cells will not interference with each other. Regarding the media access control, at the beginning of each time slot, each cell first check whether

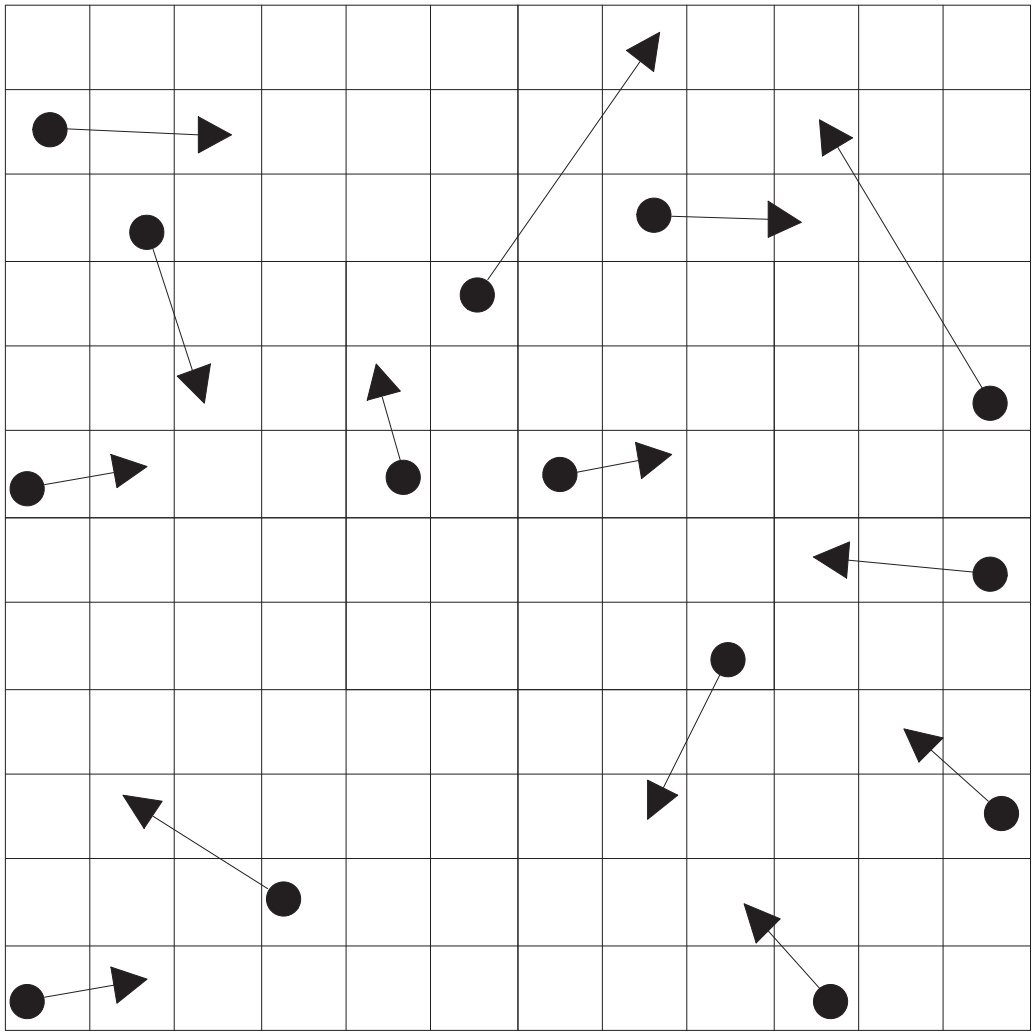


Figure 4.4: A snap of a cell-partitioned MANET.

there exist source-destination pairs within it. If there exist such pairs, this cell uniformly randomly designates one source node to access the wireless channel; otherwise this cell uniformly randomly designates any one node to access the wireless channel.

We use p_0 and p_1 to denote the probabilities that there are at least two nodes in a cell and there is at least one source-destination pair in a cell, respectively. Based

on the results of [11], we then have

$$p_0 = 1 - \left(1 - \frac{1}{m^2}\right)^n - \frac{n}{m^2}\left(1 - \frac{1}{m^2}\right)^{n-1}, \tag{4.17}$$

$$p_1 = 1 - \left(1 - \frac{1}{m^4}\right)^{n/2}. \tag{4.18}$$

At a time slot, the total transmission opportunity in the network is $m^2 \cdot p_0$, which is shared equally by all nodes, so we have

$$n \cdot (p_{sd} + p_{sr} + p_{rd}) = m^2 \cdot p_0. \tag{4.19}$$

Similarly, we have

$$n \cdot p_{sd} = m^2 \cdot p_1. \tag{4.20}$$

Combining with $\frac{p_{sr}}{p_{rd}} = \frac{\alpha}{1-\alpha}$, we have

$$p_{sd} = \frac{1}{d} p_1, \tag{4.21}$$

$$p_{sr} = \frac{\alpha}{d}(p_0 - p_1), \tag{4.22}$$

$$p_{rd} = \frac{1-\alpha}{d}(p_0 - p_1), \tag{4.23}$$

where $d = \frac{n}{m^2}$ denotes the node density.

Substituting (4.21) and (4.22) into (4.11), we can see that the throughput capacity of the cell-partitioned MANET with LS-MAC is determined as

$$T_c = \frac{1}{d} p_1 + \frac{\alpha}{d}(p_0 - p_1)\left(1 - \frac{\mathrm{C}_{B_r} \cdot \beta^{B_r}}{\sum_{i=0}^{B_r} \mathrm{C}_i \cdot \beta^i}\right). \tag{4.24}$$

It is notable from Corollary 3 that when $\alpha = 0.5$ and $B_r \to \infty$, then (4.24) is reduced to the capacity result in [11], i.e., $T_c = \frac{p_0 + p_1}{2d}$.

It is notable from formula (4.24) that by adjusting the control parameter α, we

can achieve the optimal throughput capacity T_c^*, which is defined as the maximum value of T_c optimized over α, i.e., $T_c^* = \max_{\alpha\in[0,1]} T_c$. Regarding T_c^* and corresponding α^* for the cell-partitioned MANET with LS-MAC, we have the following theorem.

Theorem IV.3 *For a concerned MANET with LS-MAC, its optimal throughput capacity T_c^* is determined as*

$$T_c^* = \frac{1}{d}p_1 + \frac{p_0 - p_1}{d(1+\gamma^*)}\frac{h(\gamma^*)}{h(\gamma^*) + \mathrm{C}_{B_r}}, \tag{4.25}$$

and the corresponding optimal transmission ratio α^* is given by $\alpha^* = \frac{1}{1+\gamma^*}$, where

$$h(\gamma) = \sum_{i=0}^{B_r-1} \mathrm{C}_i \cdot \gamma^{B_r-i}, \tag{4.26}$$

$h'(\gamma)$ is the derivative of $h(\gamma)$, and r^* is determined by the following equation

$$h(\gamma^*)[h(\gamma^*) + \mathrm{C}_{B_r}] = (1+\gamma^*)\mathrm{C}_{B_r}h'(\gamma^*). \tag{4.27}$$

Proof 4 *We define $\gamma = \frac{1-\alpha}{\alpha}$ (i.e., $\alpha = \frac{1}{1+\gamma}$, $\beta = \frac{1}{\gamma}$), and $g(\gamma) = (1+\gamma)\left(1 + \frac{\mathrm{C}_{B_r}}{h(\gamma)}\right)$. From (4.24) we can see that the optimal throughput capacity T_c^* is determined as*

$$\begin{aligned} T_c^* &= \max_{\alpha\in[0,1]} T_c \\ &= \frac{1}{d}p_1 + \frac{1}{d}(p_0 - p_1) \cdot \max_{\alpha\in[0,1]} \left\{ \alpha \left(1 - \frac{\mathrm{C}_{B_r} \cdot \beta^{B_r}}{\sum_{i=0}^{B_r} \mathrm{C}_i \cdot \beta^i} \right) \right\} \\ &= \frac{1}{d}p_1 + \frac{p_0 - p_1}{d}\frac{1}{\min_{\gamma\geq 0} g(\gamma)}. \end{aligned} \tag{4.28}$$

We can see that: 1) $g(\gamma)$ is an elementary function [52], so it is continuous and differentiable on the interval $\gamma \in (0,\infty)$; 2) $\lim_{\gamma\to 0} g(\gamma) \to \infty$, $\lim_{\gamma\to\infty} g(\gamma) \to \infty$ and $g(\gamma) > 0$. According to the Extreme Value Theorem [53], there must exists $0 < \gamma^* < \infty$ such that $0 < g(\gamma^*) \leq g(\gamma)$ for $\forall\gamma \in (0,\infty)$ and $g'(\gamma^*) = 0$, so

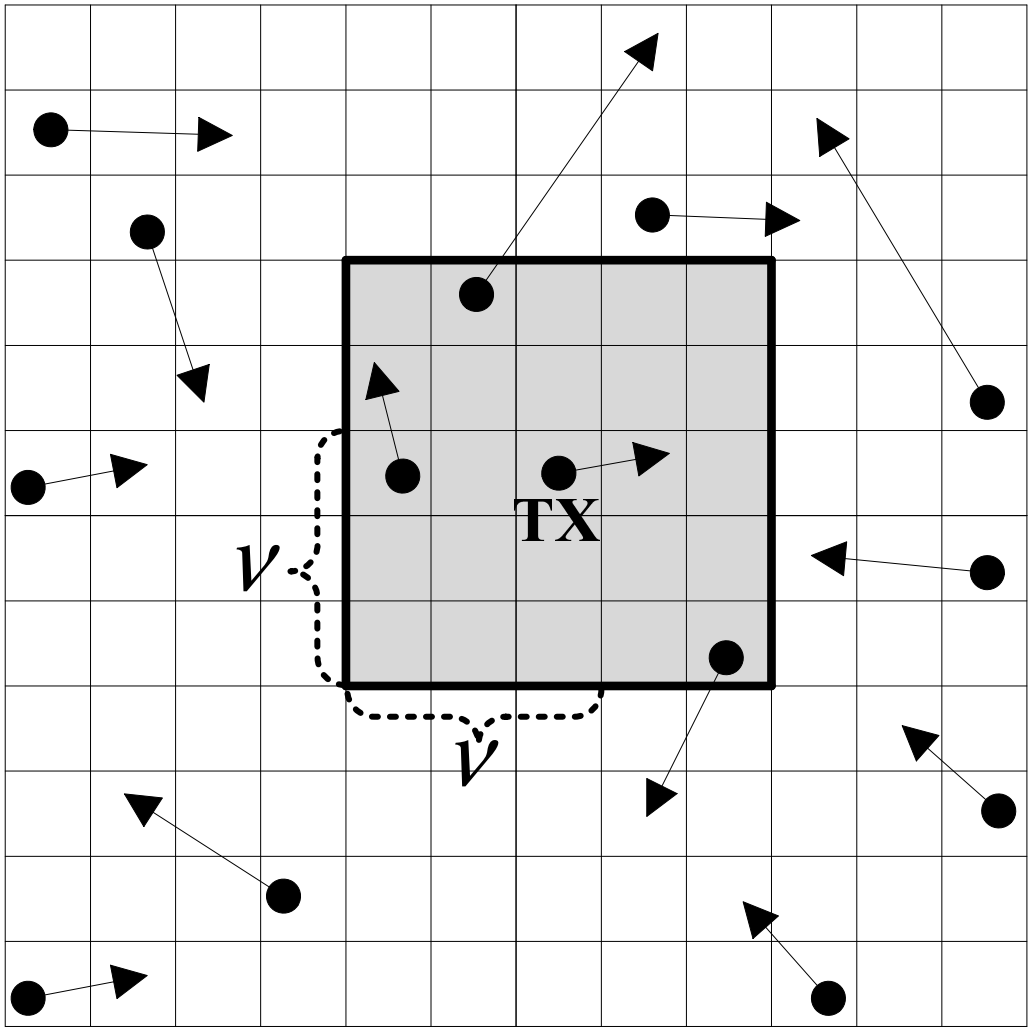


Figure 4.5: Transmission range of a node.

equation (4.27) follows. Then formula (4.25) follows by substituting γ^* into (4.28).

Based on Theorem IV.3 we have the following corollary.

Corollary 4 *For any settings of n and B_r, $\alpha^* < 0.5$; when $B_r \to \infty$, $\alpha^*|_{B_r \to \infty} = 0.5$ and $T_c^*|_{B_r \to \infty} = \frac{p_0+p_1}{2d}$.*

Proof 5 *See A.3 for the proof.*

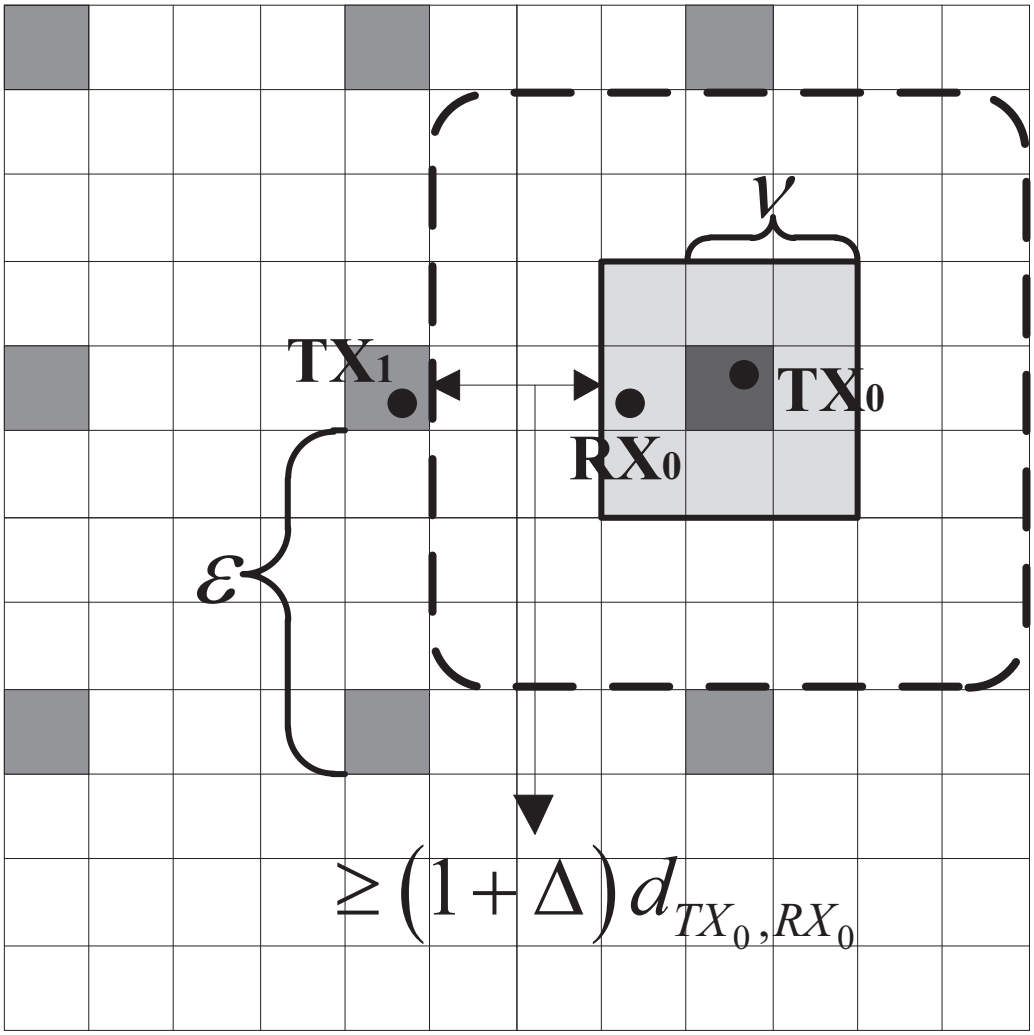


Figure 4.6: Illustration of an equivalence class (all the cells with gray color belong to the same EC).

4.5.2 Cell-Partitioned MANET with EC-MAC

We further consider a more general cell-partitioned MANET which applies a flexible transmission range and the EC-MAC [16, 17, 19, 35, 41, 54]. As shown in Fig. 4.5, the transmission range of a transmitter TX covers a set of cells which have a horizontal and vertical distance of no more than $\nu - 1$ cells away from its own cell. To prevent simultaneous transmissions from interfering with each other, the EC-MAC is adopted. As illustrated in Fig. 4.6 that with EC-MAC, all cells are divided into different ECs, where any two cells in the same EC have a horizontal and vertical distance

of some multiple of ε cells. Thus, the MANET contains in total ε^2 ECs and each EC contains $J = \lfloor m^2/\varepsilon^2 \rfloor$ cells. ECs alternatively becomes active every ε^2 time slots, and each active cell of an active EC allows only one node in it (if any) to conduct data transmission. Suppose that at time slot t, a transmitter TX_0 in an active cell will transmit a packet to its receiver RX_0, in order to ensure the transmission successful, according to the Protocol Model [55] it should satisfy that

$$d_{TX_1,RX_0} \geq (1+\Delta) d_{TX_0,RX_0}, \tag{4.29}$$

where TX_1 denotes a concurrent transmitter in any one of the other active cells, $d_{i,j}$ denotes the distance between nodes i and j, and Δ is a guard factor. Then we have

$$\varepsilon - \nu \geq (1+\Delta)\sqrt{2}\nu. \tag{4.30}$$

To enable as many concurrent transmissions to be scheduled as possible while avoiding interference among these transmissions, ε should be set as

$$\varepsilon = \min\{\lceil (1+\Delta)\sqrt{2}\nu + \nu \rceil, m\}. \tag{4.31}$$

Similar to media access scheme in previous subsection, we consider that each active cell dominate the media access control to determine which node in it as the transmitter. Given a time slot and an active cell c, let p_3 denote the probability that there are at least one node within c and another node within the transmission range of c, and p_4 denote the probability that there are at least one source-destination pair within the transmission range of c and for each of such pair(s), at least one of its two

nodes is within c, then we have

$$p_3 = \frac{1}{m^{2n}}[m^{2n} - (m^2-1)^n - n(m^2-l)^{n-1}], \tag{4.32}$$

$$p_4 = \frac{1}{m^{2n}}[m^{2n} - (m^4-2l+1)^{n/2}], \tag{4.33}$$

where $l = (2\nu-1)^2$. Notice also that $p_{sd} = \frac{J}{n}p_4$, $p_{sr} = \frac{\alpha J}{n}(p_3 - p_4)$ and $p_{rd} = \frac{(1-\alpha)J}{n}(p_3 - p_4)$. By substituting these results into (4.11), the throughput capacity of a cell-partitioned MANET with EC-MAC is then determined as

$$T_c = \frac{J}{n}p_4 + \frac{\alpha J}{n}(p_3 - p_4)\left(1 - \frac{\mathrm{C}_{B_r}\cdot\beta^{B_r}}{\sum_{i=0}^{B_r}\mathrm{C}_i\cdot\beta^i}\right). \tag{4.34}$$

We can see that when $\alpha = 0.5$ and $B_r \to \infty$, then (4.34) is reduced to the capacity result in [35], i.e., $T_c = \frac{J(p_3+p_4)}{2n}$. Based on the proof similar to that of Theorem IV.3, we have the following corollary regarding the optimal throughput capacity T_c^* of the MANET with EC-MAC.

Corollary 5 *For a concerned MANET with EC-MAC, its optimal throughput capacity T_c^* is determined as*

$$T_c^* = \frac{J}{n}p_1 + \frac{J(p_3-p_4)}{n(1+\gamma^*)}\frac{h(\gamma^*)}{h(\gamma^*)+\mathrm{C}_{B_r}}, \tag{4.35}$$

and the corresponding optimal transmission ratio α^* is given by $\alpha^* = \frac{1}{1+\gamma^*}$, where $h(\gamma)$ and γ^* are determined by (4.26) and (4.27), respectively.

4.6 Simulation Results

In this section, we first provide the simulation results to validate our theoretical framework for the throughput capacity analysis of MANETs with relay-buffer constraint, and then apply our theoretical results to illustrate the performance of such

networks.

4.6.1 Simulation Settings

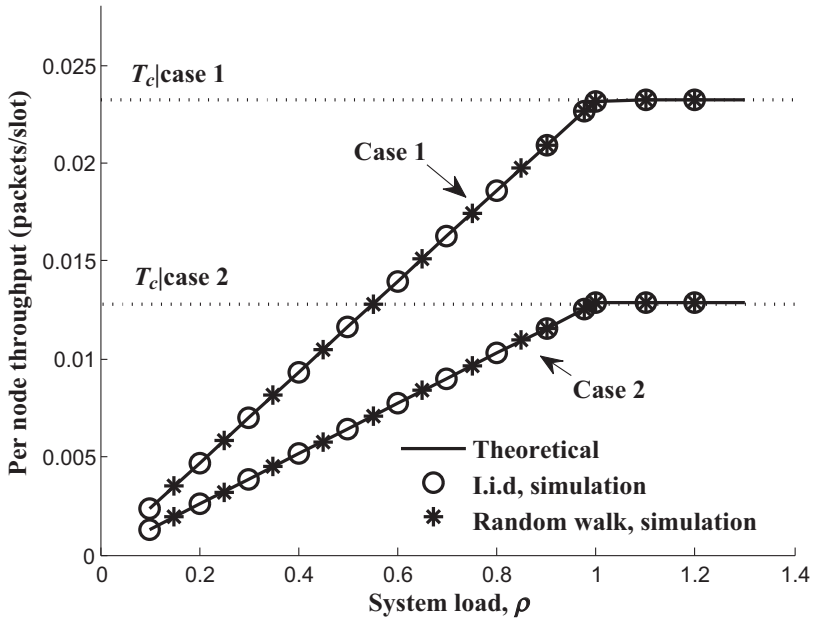
For the validation of our framework, a dedicated C++ simulator was developed to simulate the behaviors of a cell-partitioned MANETs with both the LS-MAC and EC-MAC [56]. The i.i.d mobility model [11, 21, 45] and random walk model [12] were implemented in the simulator. Under the i.i.d model, at the beginning of each time slot, each node independently and uniformly selects a cell among all m^2cells and stays in it until the end of this time slot. Under the random walk model, at the beginning of each time slot, every node independently selects a cell among its current cell and its 8 adjacent cells with equal probability 1/9, then stays in it until the end of this time slot.

Two network scenarios of $(n = 72, m = 6, B_r = 5, \alpha = 0.5)$ and $(n = 200, m = 10, B_r = 8, \alpha = 0.3)$ are considered in the simulation, where we set $\nu = 1$ and $\Delta = 1$ for the MANET with EC-MAC [57][1]. To simulate the throughput, we focus on a specific node and count its received packets over a period of 2×10^8 time slots, and then calculate the averaged number of packets this node can receive per time slot. The system load ρ is defined as $\rho = \lambda/T_c$, and T_c is given by (4.24) and (4.34) for the LS-MAC and EC-MAC, respectively.

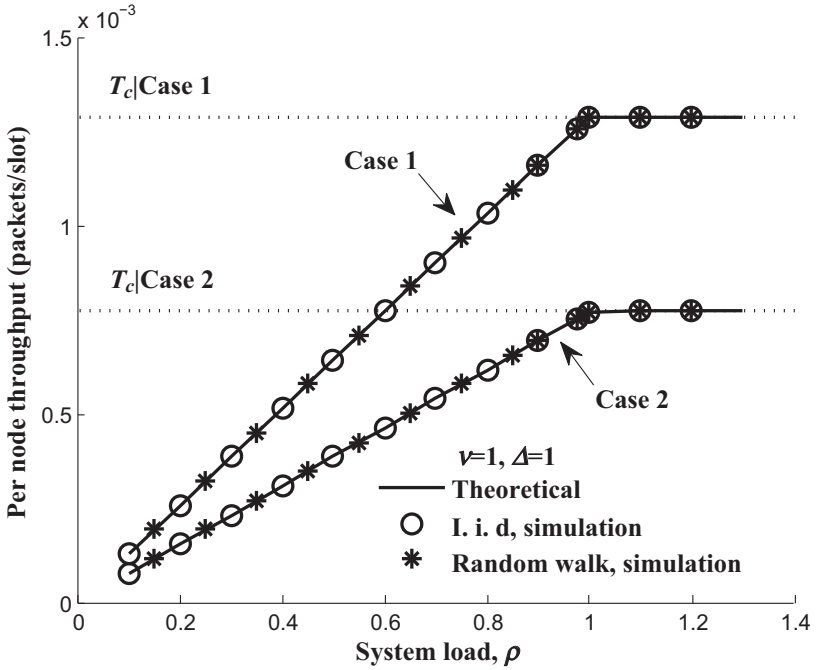
4.6.2 Validation of Theoretical Throughput Capacity Results

To validate the throughput capacity results (4.24) and (4.34), we provide plots of throughput versus system load ρ in Fig. 4.7. It can be observed from Fig. 4.7 that the simulation results agree well with the theoretical ones under both LS-MAC and EC-MAC, indicating that our framework is highly efficient in capturing the throughput behaviors of concerned buffer-limited MANETs. We can see from Fig. 4.7 that just as

[1]The simulation settings can be flexibly adjusted in our simulator.



(a) Throughput performance under the LS-MAC.



(b) Throughput performance under the EC-MAC.

Figure 4.7: Throughput performance of the MANETs with LS-MAC and EC-MAC. Case 1: $n = 72, m = 6, B_r = 5, \alpha = 0.5$. Case 2: $n = 200, m = 10, B_r = 8, \alpha = 0.3$.

Theorem IV.1 predicates that for a concerned MANET with relay-buffer constraint, its throughput increases linearly with ρ when $\rho \leq 1$ and then keeps as a constant T_c

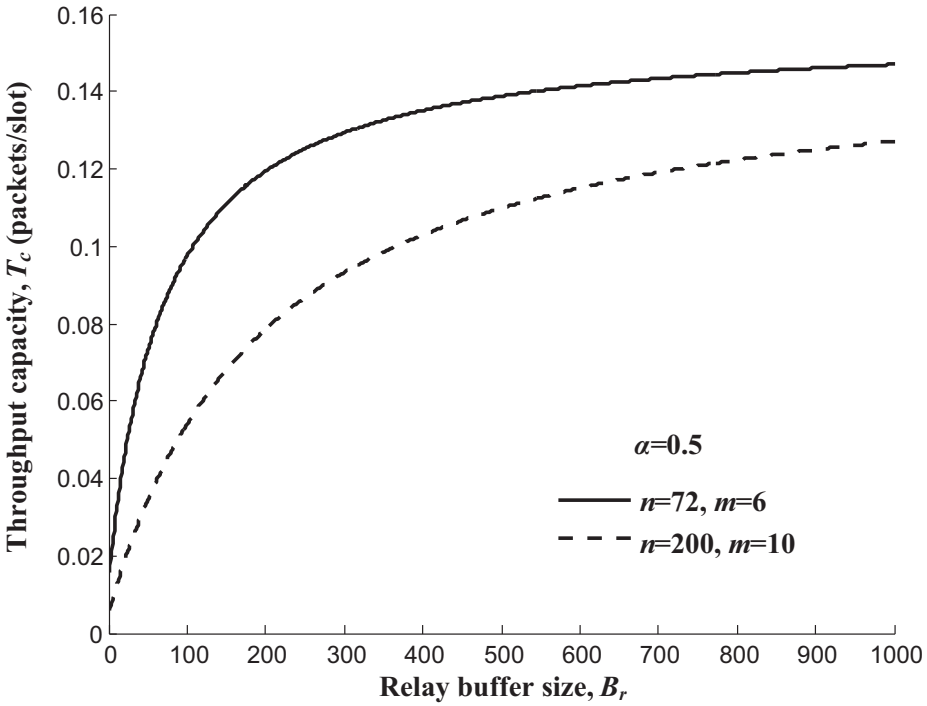


Figure 4.8: Throughput capacity T_c vs. relay buffer size B_r.

determined by (4.1) when $\rho > 1$.

4.6.3 Discussions

With the help of our theoretical results, we illustrate here the impacts of network parameters on the throughput capacity. Notice that for a concerned MANET its overall throughput behavior under the LS-MAC is very similar to that under the EC-MAC, so we only consider the LS-MAC here for illustration.

We first summarize in Fig. 4.8 how throughput capacity T_c varies with relay buffer size B_r under two network scenarios of $(n = 72, m = 6)$ and $(n = 200, m = 10)$, where α is fixed as 0.5. We can see from Fig. 4.8 that just as discussed in Corollary 1, the throughput capacity of such a MANET can be improved by adopting a larger relay buffer. A careful observation of Fig. 4.8 shows that as B_r increases the capacity T_c first increases quickly and then gradually converges to a constant determined by Corollary 3. This observation indicates we can determine a suitable buffer size B_r ac-

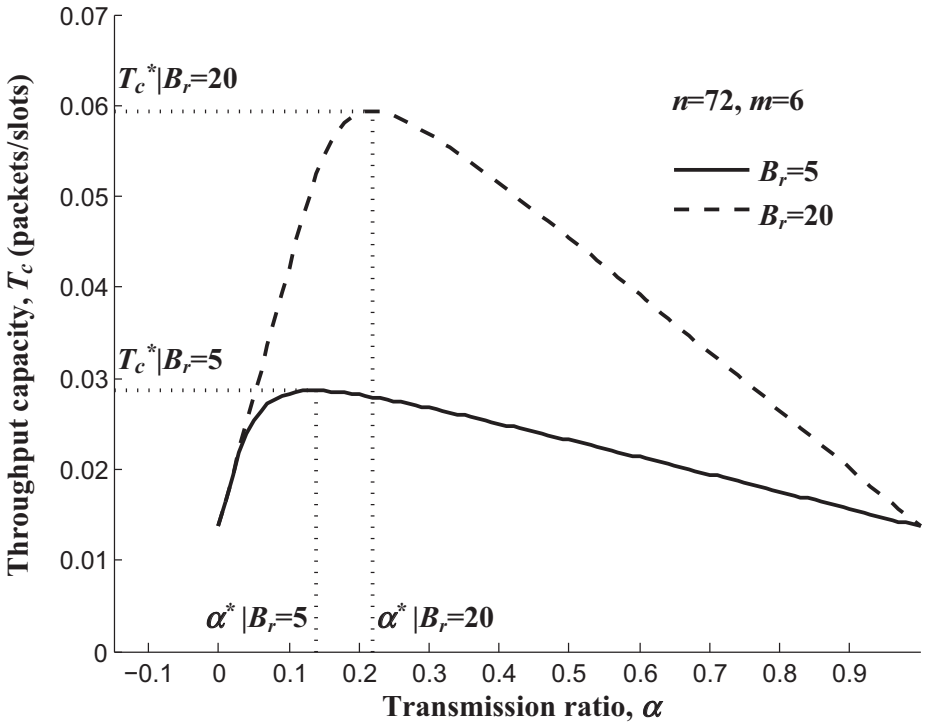
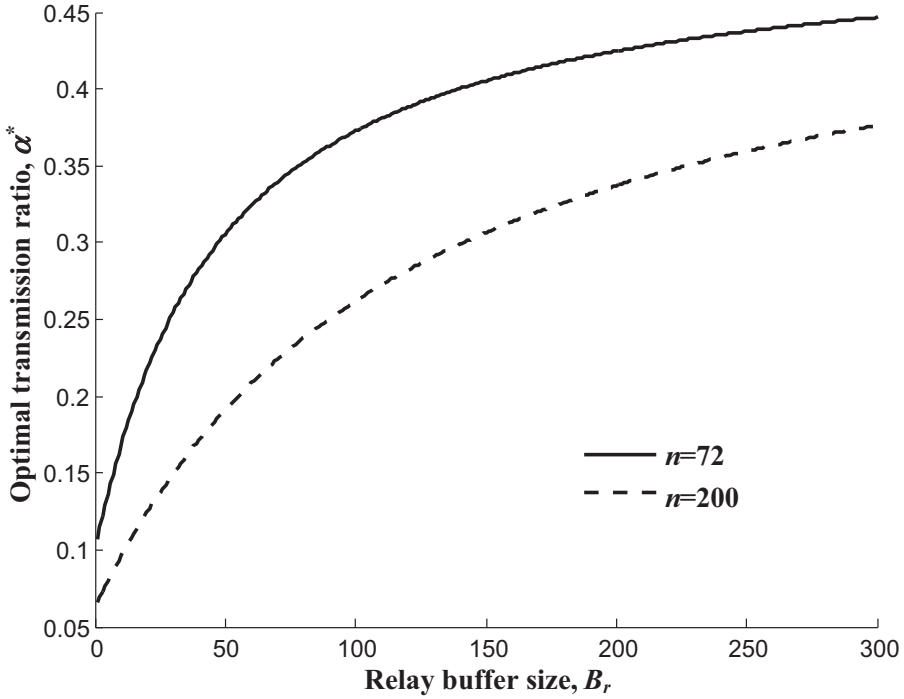


Figure 4.9: Throughput capacity T_c vs. transmission ratio α.

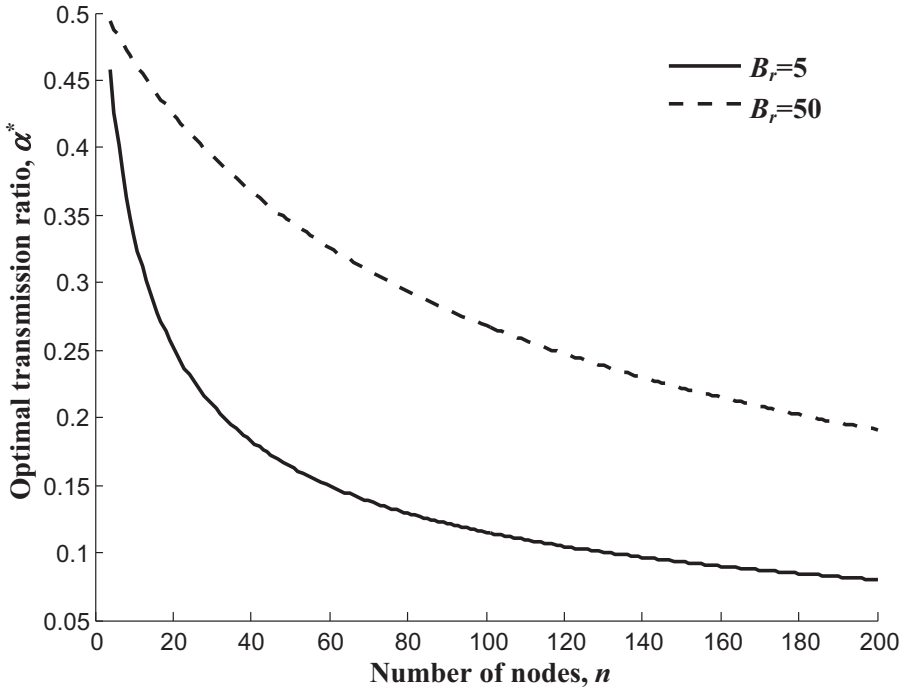
cording to the requirement on network capacity such that a graceful trade-off between capacity performance and buffer cost can be achieved.

To illustrate the optimal throughput capacity, we show in Fig. 4.9 the impact of transmission ratio α on throughput capacity T_c under the settings of $n = 72$, $m = 6$ and $B_r = \{5, 20\}$. We can see from Fig. 4.9 that under a given setting of B_r, as α increases T_c first increases and then decreases, and just as discussed in Theorem IV.3 that there exists an optimal α^* to achieve the optimal throughput capacity T_c^* . This is mainly due to the reason that the effects of α on T_c are two folds. On one hand, a larger α will lead to a higher probability of conducting S-R transmission; on the other hand, a larger α will result in a higher ROP thus a lower opportunity of conducting the S-R transmission. As a summary, in order to improve the throughput performance of a buffer-limited MANET, nodes should cooperate with each other, and they should be neither too selfish nor too selfless.

Based on the results of Theorem IV.3, we illustrate in Fig. 4.10 how the opti-



(a) Optimal transmission ratio α^* vs. relay-buffer size B_r.



(b) Optimal transmission ratio α^* vs. number of nodes n.

Figure 4.10: Optimal transmission ratio α^* vs. B_r and n.

mal transmission ratio α^* is related to B_r and n. We can see that just as proved in Corollary 4 that α^* increases as B_r grows while it decreases as n grows, and the

optimal transmission ratio never exceeds 0.5. These behaviors indicate that in a network with the fixed number of nodes n, if we upgrade the capacity of each node by adopting a larger relay buffer, we should accordingly allocate a higher probability for S-R transmission (i.e., nodes should be more selfish), to achieve the optimal throughput capacity. On the other hand, when the relay buffer size of each node is fixed, if we increase the scale of the network by accommodating more nodes, we should accordingly increase the probability of R-D transmission (i.e., nodes should be more selfless), to release the relay buffer space and thus guarantee the optimal throughput capacity there.

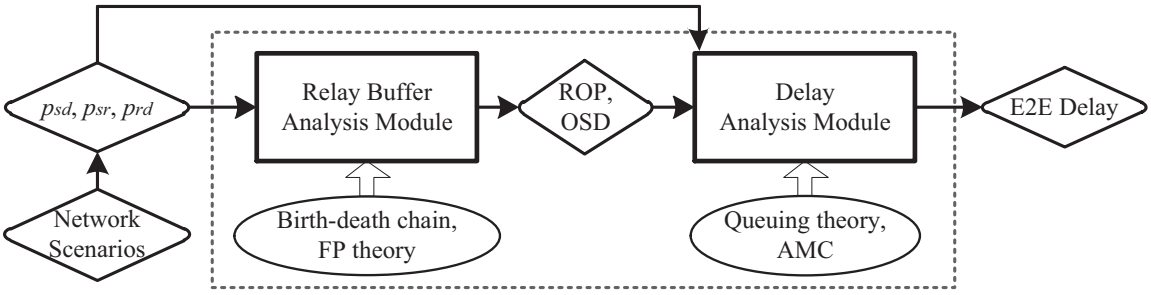
4.7 Summary

In this chapter, we first revealed the inherent relationship between the throughput capacity and ROP in a MANET with relay-buffer constraint, and then developed a theoretical framework to fully characterize the buffer occupancy process. Based on this framework, we derived the throughput capacity in closed-form and further conducted cases studies under two typical network scenarios. It is expected the theoretical framework developed in this chapter will be also helpful for exploring the throughput capacity of buffer-limited MANETs under other mobility models and other network scenarios. An interesting finding of this chapter is that for throughput capacity optimization in such MANETs, the optimal setting of transmission ratio in the 2HR-α scheme there increases with the relay buffer size but decreases with the network size, and it never exceeds 0.5.

CHAPTER V

End-to-End Delay of MANETs under Relay-Buffer Constraint

In this chapter we consider the MANETs with relay-buffer constraint and explore the corresponding E2E delay performance. Based on the theoretical framework developed in Chapter IV, we first adopt the fixed-point theory for the numerical evaluation of the overflowing probability and the stationary occupancy state distribution (OSD) of the relay buffer. With the help of stationary OSD of the relay buffer, we then derive the exact expression of expected E2E delay by modeling the packet source-queuing delay and delivery delay respectively. The packet source-queuing delay is characterized by a Bernoulli/Bernoulli queuing model and the packet deliver delay is characterized by an absorbing Markov chain model. Case studies are also provided under the two typical network scenarios, while we adopt a fully distributed media access scheme in this chapter. Finally, extensive simulation and numerical results are presented to illustrate the efficiency of our delay analysis as well as the impacts of network parameters on delay performance.



E2E Delay Modeling for MANETs with Relay-Buffer Constraint

FP: fixed-point; ROP: relay-buffer overflowing probability; OSD: occupancy state distribution; AMC: absorbing Markov chain

Figure 5.1: Illustration of E2E delay modeling for MANET with relay-buffer constraint.

5.1 Problem Formulation

In this chapter, we continue to consider the MANETs with relay-buffer constraint described in Section 4.1. Notice that the value of parameter α in 2HR-α routing scheme does not affect the development of theoretical framework for buffer occupancy process modeling, without loss of generality, this chapter focuses on the 2HR scheme with the feedback mechanism (i.e., α is fixed as 0.5). The packet source-queuing delay and delivery delay are two important performance metrics which help us derive the E2E delay in this chapter, so we present their formal definitions as follows.

Source-queuing Delay: the *source-queuing delay* D_{sq} is defined as the time it takes a packet to move to the HoL in the source queue after it is generated by its source node.

Delivery Delay: the *delivery delay* D_d is defined as the time it takes a packet to reach its destination after it moves to the HoL in the source queue.

Based on above definitions, we can see clearly that the E2E delay of a packet is just the sum of its source-queuing delay and delivery delay, i.e., $D = D_{sq} + D_d$. Thus, we can derive the packet E2E delay by analyzing the source-queuing delay and delivery delay respectively. Fig. 5.1 illustrates the structure of E2E delay modeling for

the MANETs with relay-buffer constraint, which consists of the relay buffer analysis module and delay analysis module. In relay buffer analysis module, the birth-death chain model is applied to characterize the relay buffer occupancy process, and the fixed-point theory [58] is further applied to solve the ROP numerically and obtain the stationary occupancy state distribution (OSD) of relay buffer recursively. With the help of ROP and OSD, the delay analysis module applies the queuing theory [47] and absorbing Markov chain [59] to analyze the packet source-queuing delay and delivery delay, respectively, such that the exact expression of packet E2E delay can be finally derived.

5.2 ROP and OSD Analysis

Based on the theoretical framework developed in Chapter IV, we obtain the critical self-mapping function of the relay-buffer overflowing probability $p_o(\lambda)$ as follow.

$$p_o(\lambda) = f\left(p_o(\lambda)\right) = \frac{\mathrm{C}_{B_r} \cdot \rho_s(\lambda)^{B_r}}{\sum_{k=0}^{B_r} \mathrm{C}_i \cdot \rho_s(\lambda)^k}, \tag{5.1}$$

where $\rho_s(\lambda) = \dfrac{\lambda}{\mu_s(\lambda)} = \dfrac{\lambda}{p_{sd} + p_{sr}(1 - p_o(\lambda))}$.

It is notable that the self-mapping function in (5.1) is a contraction mapping [58], and given a packet generating rate λ, it contains no other unknown quantities except $p_o(\lambda)$. According to Banach fixed-point theorem, there exists a unique fixed-point for the self-mapping function. The unique fixed-point is just $p_o(\lambda)$ and it can be searched by the fixed-point iteration. The detailed fixed-point iteration for searching $p_o(\lambda)$ is summarized in Algorithm 2.

We let $\pi_r(i)$ denote the probability that there are i packets occupying the relay buffer in the stationary state. Thus the stationary OSD of the relay buffer $\mathbf{\Pi}_r =$

Algorithm 2 Fixed-point iteration

Require:

Basic network parameters $\{n, B_r, \lambda, p_{sd}, p_{sr}, p_{rd}\}$;

Ensure:

Relay buffer overflow probability $p_o(\lambda)$;

1: Set $x_1 = 0$ and $i = 1$;
2: **while** $(x_i - x_{i-1} \geq 10^{-6}) \vee (i = 1)$ **do**
3: $\quad i = i + 1$;
4: $\quad \mu_s(\lambda) = p_{sd} + p_{sr} \cdot (1 - x_{i-1})$;
5: $\quad \rho_s(\lambda) = \frac{\lambda}{\mu_s(\lambda)}$;
6: $\quad x_i = \frac{\mathrm{C}_{B_r} \rho_s(\lambda)^{B_r}}{\sum_{k=0}^{B_r} \mathrm{C}_k \rho_s(\lambda)^k}$;
7: **end while**
8: $p_o(\lambda) = x_i$;
9: **return** $p_o(\lambda)$;

$[\pi_r(0), \pi_r(1), \cdots, \pi_r(B_r)]$ can be determined recursively as

$$\pi_r(B_r) = p_o(\lambda), \tag{5.2}$$

$$\pi_r(i) = \pi_r(B_r) \cdot \rho_s(\lambda)^{i-B_r} \cdot \frac{\mathrm{C}_i}{\mathrm{C}_{B_r}}, \qquad 0 \leq i < B_r. \tag{5.3}$$

5.3 Delay Analysis

5.3.1 Source-queuing Delay

We first analyze the source-queuing delay of a packet. Regarding the source buffer of a node, since in every time slot a new packet is generated with probability λ and a service opportunity arises with probability $\mu_s(\lambda)$ being determined as $\mu_s(\lambda) = p_{sd} + p_{sr}(1 - p_o(\lambda))$, the source buffer can be modeled as a Bernoulli/Bernoulli queue. The ROP $p_o(\lambda)$ is obtained by the fixed-point iteration algorithm, and in the following analysis we use p_o and μ_s to represent $p_o(\lambda)$ and $\mu_s(\lambda)$ respectively if there is no ambiguous.

We know from the theoretical analysis of Chapter IV that the throughput capacity

of the concerned MANET is

$$T_c = p_{sd} + p_{sr}\frac{B_r}{n-2+B_r}, \tag{5.4}$$

and $\lambda < \mu_s$ if $\lambda < T_c$, $\lambda \geq \mu_s$ if $\lambda \geq T_c$. We let L_s denote the average queue length of the source buffer. When $\lambda \geq \mu_s$, $L_s \to \infty$, such that the expected source-queuing delay $\mathbb{E}\{D_{sq}\}$ tends to infinity. When $\lambda < \mu_s$, L_s can be determined as [47]

$$L_s = \frac{\lambda - \lambda^2}{\mu_s - \lambda}. \tag{5.5}$$

According to the Little's Law [50], the mean service time of a packet in its source buffer[1] $\mathbb{E}\{D_s\}$ is given by

$$\mathbb{E}\{D_s\} = \frac{L_s}{\lambda} = \frac{1-\lambda}{\mu_s - \lambda}. \tag{5.6}$$

Thus, the expected source-queuing delay is determined as

$$\mathbb{E}\{D_{sq}\} = \mathbb{E}\{D_s\} - \frac{1}{\mu_s} = \frac{\lambda(1-\mu_s)}{\mu_s(\mu_s - \lambda)}. \tag{5.7}$$

5.3.2 Delivery Delay and E2E Delay

We present the following theorem regarding the expected E2E delay of the concerned MANET with relay-buffer constraint.

Theorem V.1 *For the concerned MANET with number of nodes n, relay buffer size B_r and packet generating rate λ ($\lambda < T_c$), the expected delivery delay $\mathbb{E}\{D_d\}$ and the*

[1]The service time of a packet in its source buffer is defined as the time it takes a packet to depart from its source buffer after it is generated by its source node.

expected E2E delay $\mathbb{E}\{D\}$ *of a packet are determined as*

$$\mathbb{E}\{D_d\} = \frac{1 + (n - 2 + L_r^*)(1 - p_o)}{\mu_s}, \tag{5.8}$$

$$\mathbb{E}\{D\} = \frac{1 - \lambda}{\mu_s - \lambda} + \frac{(n - 2 + L_r^*)(1 - p_o)}{\mu_s}, \tag{5.9}$$

where L_r^* *denotes the expected number of packets in the relay buffer under the condition that the relay buffer is not full, and* L_r^* *is given by*

$$L_r^* = \frac{\sum_{i=0}^{B_r-1} i\mathrm{C}_i \cdot \rho_s^i}{\sum_{i=0}^{B_r-1} \mathrm{C}_i \cdot \rho_s^i}. \tag{5.10}$$

Proof 6 *We focus on a packet* p *which is the HoL packet of the source queue at time slot* t*, then in the next time slot,* p *will be delivered to its destination with probability* p_{sd}*, be forwarded to a relay node with probability* $p_{sr} \cdot (1 - p_o)$*, and still stay in the source queue with probability* $1 - \mu_s$*. Thus, the delivery process of packet* p *can be modeled as an absorbing Markov chain as illustrated in Fig. 5.2, where* S*,* R *and* D *denote the states that* p *is in source queue, forwarded to a relay, and delivered to its destination, respectively. We use* $\overline{X}_S$ *and* $\overline{X}_R$ *to denote the average transition time slots from the initial state* S *and transient state* R *to the absorbing state* D*, respectively. Then we have*

$$\overline{X}_S = 1 + \overline{X}_S \cdot (1 - \mu_s) + \overline{X}_R \cdot p_{sr}(1 - p_o), \tag{5.11}$$

$$\overline{X}_S = \frac{1 + \overline{X}_R \cdot p_{sr}(1 - p_o)}{\mu_s}. \tag{5.12}$$

Let $\pi_r^*(i)$ $(0 \leq i < B_r)$ *denote the probability that there are* i *packets in the relay buffer conditioned on that the relay buffer is not full, then we have*

$$\pi_r^*(i) = \frac{\pi_r(i)}{1 - \pi_r(B_r)} = \frac{\mathrm{C}_i \rho_s^i}{\sum_{k=0}^{B_r-1} \mathrm{C}_k \cdot \rho_s^k}, \tag{5.13}$$

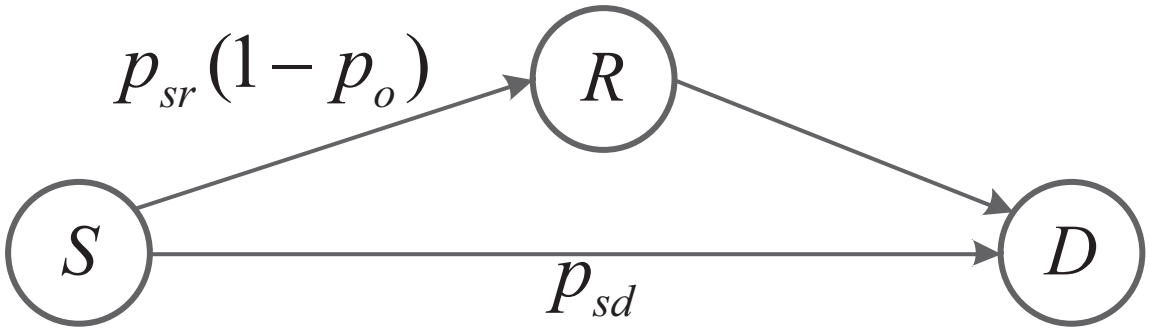


Figure 5.2: Illustration of absorbing Markov chain model for packet deliver process.

and L_r^* *is given by*

$$L_r^* = \sum_{i=0}^{B_r-1} i \cdot \pi_r^*(i) = \frac{\sum_{i=0}^{B_r-1} i\mathrm{C}_i \cdot \rho_s^i}{\sum_{i=0}^{B_r-1} \mathrm{C}_i \cdot \rho_s^i}. \tag{5.14}$$

Due to the symmetry of relay queues in a relay buffer (because the buffered packets are destined to each of the $n-2$ *destinations with equal probability), the mean number of packets in one relay queue under the condition that the relay buffer is not full is* $L_r^*/(n-2)$*. Meanwhile, it is notable that the location of each node is stationary and ergodic with stationary distribution uniform on the network area, thus when a relay node conducts the R-D transmission with probability* p_{rd} *in a time slot, it will deliver a packet for each of the* $n-2$ *traffic flows with equal probability. Thus, the service rate of each relay queue in the relay buffer is* $\frac{p_{rd}}{n-2}$*. Then we have*

$$\overline{X}_R = \left(1 + \frac{L_r^*}{n-2}\right) \cdot \left(\frac{p_{rd}}{n-2}\right)^{-1} = \frac{n-2+L_r^*}{p_{rd}} \tag{5.15}$$

Substituting the results of (5.15) and (5.14) into (5.12), the average transition time slots from the initial state S *to the absorbing state* D *is determined as*

$$\overline{X}_S = \frac{1 + (n-2+L_r^*)(1-p_o)}{\mu_s}. \tag{5.16}$$

Notice that the expected delivery delay of a packet is just the average transition time

slots from the initial state S to the absorbing state D, thus we have $\mathbb{E}\{D_d\} = \overline{X}_S$, the result (5.8) follows, and then the result (5.9) follows from $\mathbb{E}\{D\} = \mathbb{E}\{D_{sq}\} + \mathbb{E}\{D_d\}$.

Based on Theorem V.1, we can further extend our delay results to the infinite buffer scenario (i.e., $B_r \to \infty$), which is shown in the following corollary.

Corollary 6 *Considering the relay buffer size tends to infinity ($B_r \to \infty$), then $\mathbb{E}\{D_d\}$ and $\mathbb{E}\{D\}$ are determined as*

$$\mathop{\mathbb{E}\{D_d\}}_{B_r\to\infty} = \frac{1}{p_{sd} + p_{sr}} + \frac{n-2}{p_{sd} + p_{sr} - \lambda}, \tag{5.17}$$

$$\mathop{\mathbb{E}\{D\}}_{B_r\to\infty} = \frac{n-1-\lambda}{p_{sd} + p_{sr} - \lambda}. \tag{5.18}$$

Proof 7 *See Appendix B.1 for the proof.*

Notice that when $B_r \to \infty$, the result of Corollary 6 is coincident with the expression of E2E delay derived in [11], where the relay buffer size is assumed to be infinite.

5.4 Case Studies

In this section, we conduct case studies to illustrate the application of our E2E delay modeling in the MANETs with relay-buffer constraint. For a given network scenario, the corresponding p_{sd}, p_{sr} and p_{rd} should be determined first, then with the inputs of these probabilities, by sequentially executing the relay buffer analysis module and delay analysis module, our delay modeling framework finally returns the delay results.

Similar to the network scenarios in Section 4.5, here we also consider the typical cell-partitioned MANETs with LS-MAC and EC-MAC. Notice that in Section 4.5, the media access control is scheduled by a cell, in this section, however, we consider

the classical DCF-style mechanism [60, 61], which is a fully distributed media access scheme scheduled by each node. More formally, at the beginning of each time slot, a node which is eligible to access the wireless channel[2] randomly selects an initial value from $[0, CW]$ (CW represents the contention widow) and starts to count down. If this node does not hear any broadcasting message until its back-off counter becomes 0, it broadcasts a message to claim itself as the transmitter; otherwise it stops its back-off counter since some other node has claimed as the transmitter.

With the DCF-style mechanism for media access control, we then determine the corresponding probabilities p_{sd}, p_{sr} and p_{rd} of the MANETs with LS-MAC and EC-MAC respectively, which are provided in the following lemmas (See Appendix B.2 for the proofs.).

Lemma 2 *For the concerned cell-partitioned MANET with LS-MAC, the probabilities p_{sd}, p_{sr} and p_{rd} are given by*

$$p_{sd} = \frac{m^2}{n} - \frac{m^2-1}{n-1} + \frac{m^2-1}{n(n-1)}\left(1-\frac{1}{m^2}\right)^{n-1}, \tag{5.19}$$

$$p_{sr} = p_{rd} = \frac{1}{2}\left\{\frac{m^2-1}{n-1} - \frac{m^2}{n-1}\left(1-\frac{1}{m^2}\right)^{n} - \left(1-\frac{1}{m^2}\right)^{n-1}\right\}. \tag{5.20}$$

Lemma 3 *For the concerned cell-partitioned MANET with EC-MAC, the probabilities p_{sd}, p_{sr} and p_{rd} are given by*

$$p_{sd} = \frac{1}{\varepsilon^2}\left\{\frac{\Gamma - \frac{m^2}{n}}{n-1} + \frac{m^2-1-(\Gamma-1)n}{n(n-1)}\left(1-\frac{1}{m^2}\right)^{n-1}\right\}, \tag{5.21}$$

$$p_{sr} = p_{rd} = \frac{1}{2\varepsilon^2}\left\{\frac{m^2-\Gamma}{n-1}\left(1-\left(1-\frac{1}{m^2}\right)^{n-1}\right) - \left(1-\frac{\Gamma}{m^2}\right)^{n-1}\right\}, \tag{5.22}$$

where $\Gamma = (2\nu-1)^2$.

[2]For the MANET with LS-MAC, each node in the network is eligible to access the wireless channel; for the MANET with EC-MAC, the node in an active cell is eligible to access the wireless channel.

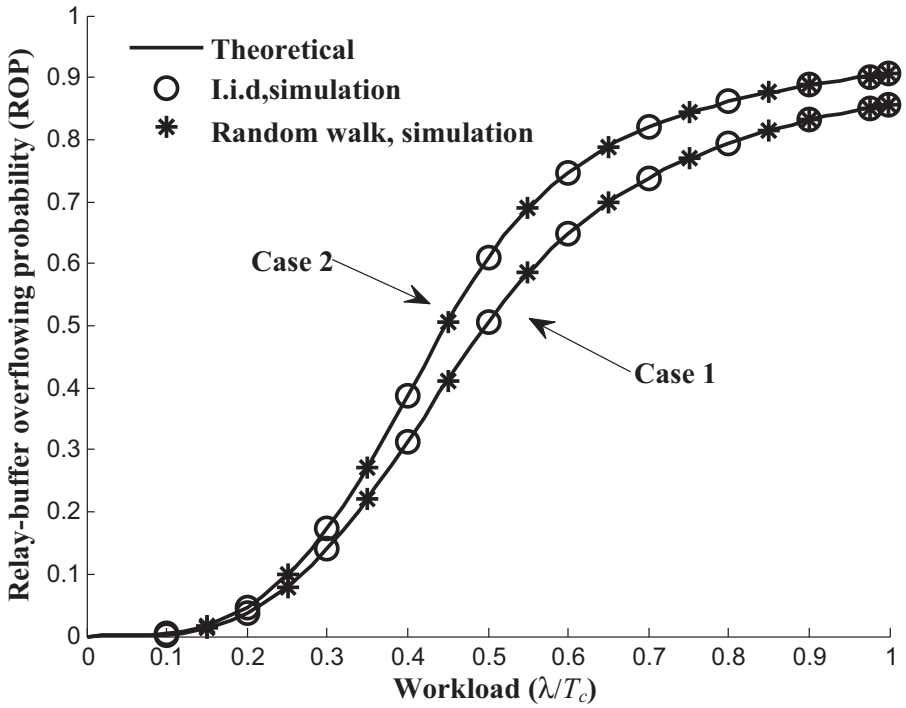
Given the number of nodes n and relay buffer size B_r, substituting formulas (5.19) and (5.20) (resp. (5.21) and (5.22)) into formula (5.4), we first determine the throughput capacity T_c of such a MANET. Then, with any packet generating rate $\lambda < T_c$, we substitute formulas (5.19) and (5.20) (resp. (5.21) and (5.22)) into Algorithm 2 to determine the corresponding ROP p_o, and μ_s can be further determined by $\mu_s = p_{sd} + p_{sr}(1 - p_o)$. Substituting λ, p_o and μ_s into formulas (5.7), (5.8) and (5.9), we finally obtain the expectations of source-queuing delay, delivery delay and E2E delay respectively for the cell-partitioned MANET with LS-MAC (resp. EC-MAC).

5.5 Simulation Results

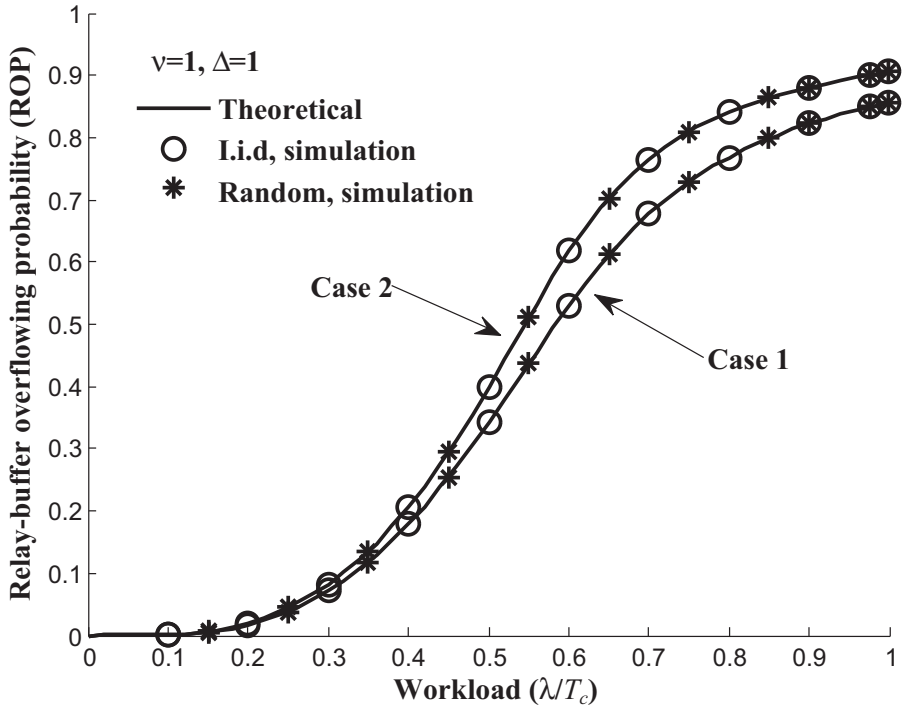
In this section, we first conduct simulations to validate our E2E delay modeling for MANETs with relay-buffer constraint, then provide discussions about the impacts of network parameters on delay performance.

5.5.1 Simulation Settings

For the validation of our delay modeling and theoretical results, a specific C++ simulator was developed to simulate the packet generating, queuing and delivery processes in a cell-partitioned MANET [62], where the network settings, including the relay buffer size B_r, number of nodes n, partition parameter m, packet generating rate λ and the mobility model can be flexibly adjusted to simulate the network performance under various scenarios. For the network scenario with EC-MAC, we set $\nu = 1$ and $\Delta = 1$ [57]. The duration of each task of simulation is set to be 2×10^8 time slots, and we only collect data from the last 80% of the time slots in each task (the system will be in the steady state with high probability), to ensure the accuracy of simulated results.

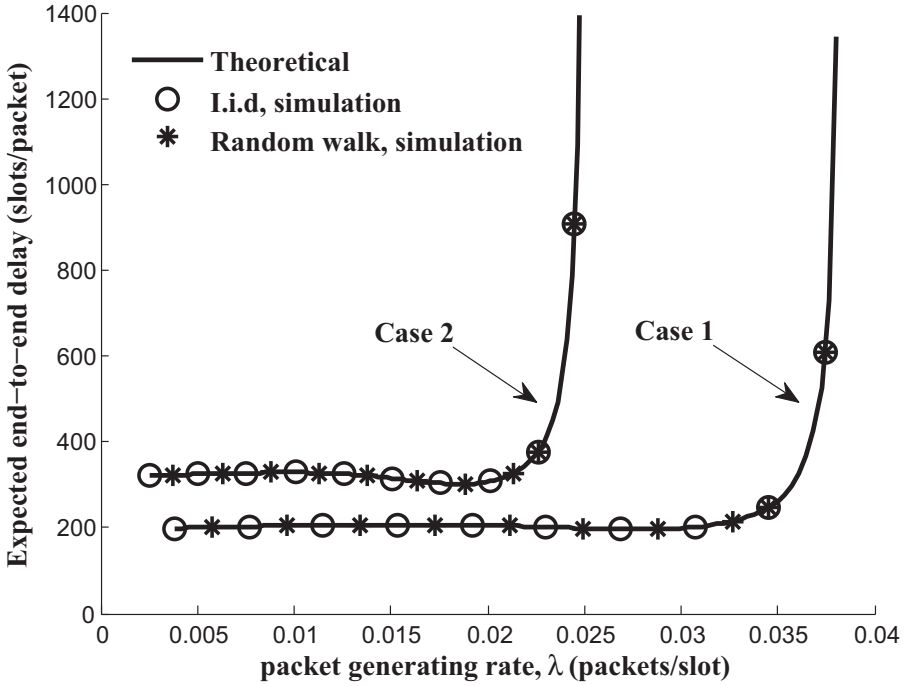


(a) ROP of the MANET with LS-MAC.

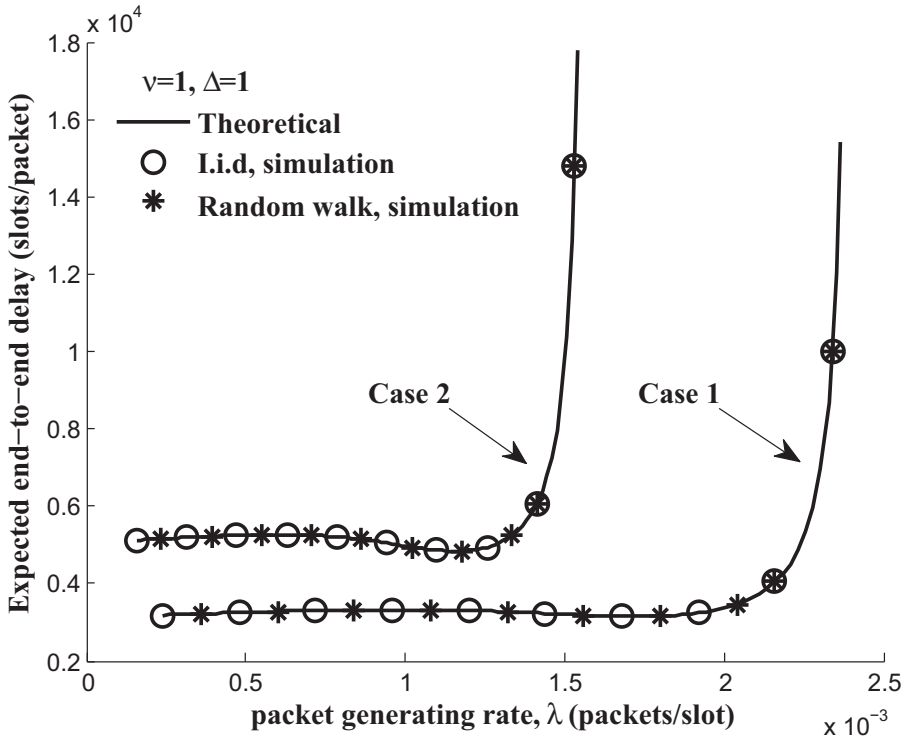


(b) ROP of the MANET with EC-MAC.

Figure 5.3: Theoretical and simulated ROP performance. Case 1: $n = 32, m = 4, B_r = 5$. Case 2: $n = 50, m = 5, B_r = 5$.



(a) Expected E2E delay of the MANET with LS-MAC.



(b) Expected E2E delay of the MANET with EC-MAC.

Figure 5.4: Theoretical and simulated E2E delay performance. Case 1: $n = 32, m = 4, B_r = 5$. Case 2: $n = 50, m = 5, B_r = 5$.

5.5.2 Validation of Theoretical Delay Results

First, we provide plots of the theoretical and simulated ROP performance under two network scenarios in Fig. 5.3, and for each scenario we consider two cases (case 1: $n = 32, m = 4, B_r = 5$, and case 2: $n = 50, m = 5, B_r = 5$) and two mobility models (the i.i.d mobility model and the random walk model). The workload is defined as λ/T_c. We can see from Fig. 5.3 that the simulation results match nicely with the theoretical ones for all the cases, which indicates that our modeling is highly efficient in depicting the occupancy behaviors of the relay buffer in the buffer-limited MANETs.

Then, with the same network settings, we provide plots of the theoretical and simulated E2E delay results in Fig. 5.4. It is observed from Fig. 5.4 that all the simulation results can match the corresponding theoretical curves very nicely, indicating that: 1) our delay modeling is highly efficient for the delay evaluation in the MANETs with relay-buffer constraint; 2) the theoretical framework is very general since it can be applied to various network scenarios. Another observation of Fig. 5.4 is that the packet E2E delay increases sharply as the packet generating rate λ approaches a specific value (e.g., under LS-MAC and case 1, the value is around 0.038), which serves as an intuitive impression of its corresponding throughput capacity T_c.

5.5.3 Discussions

With the help of our E2E delay modeling framework, we explore how the network parameters affect the the delay performance of a buffer-limited MANET. Without loss of generality, we consider here a cell-partitioned MANET with LS-MAC.

We first summarize in Fig. 5.5 that how the expected delivery delay $\mathbb{E}\{D_d\}$ varies with the workload λ/T_c. A very interesting observation is that under the relay-buffer constraint ($B_r = 5$ and $B_r = 20$), as the workload increases, $\mathbb{E}\{D_d\}$ first increases to a maximum and then decreases. This is due to the reason that the effects of

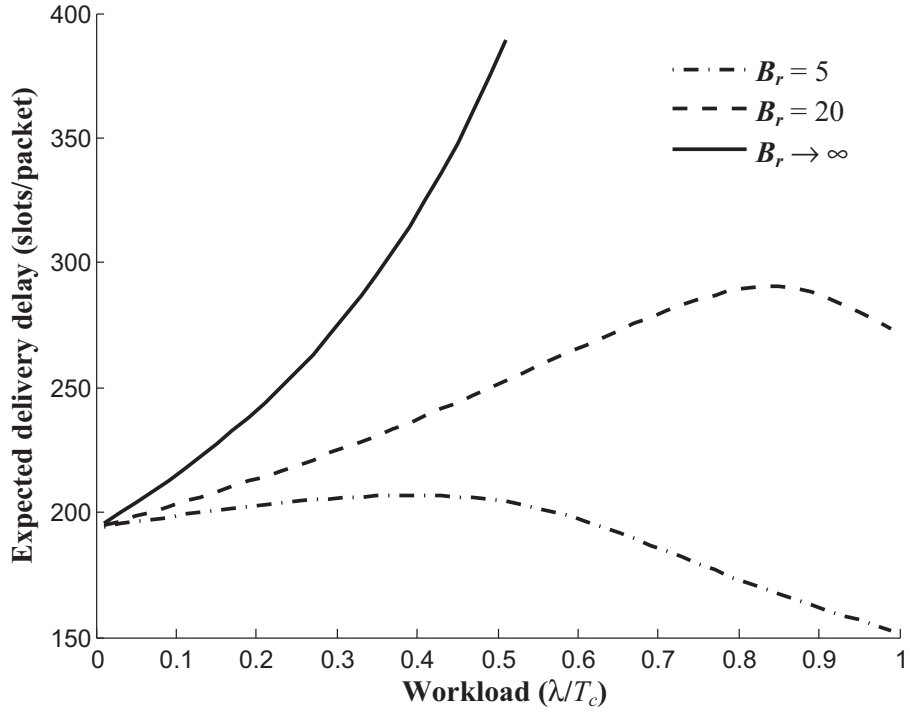


Figure 5.5: Delivery delay vs. workload (λ/T_c) under different settings of relay buffer size. $n = 32, m = 4$.

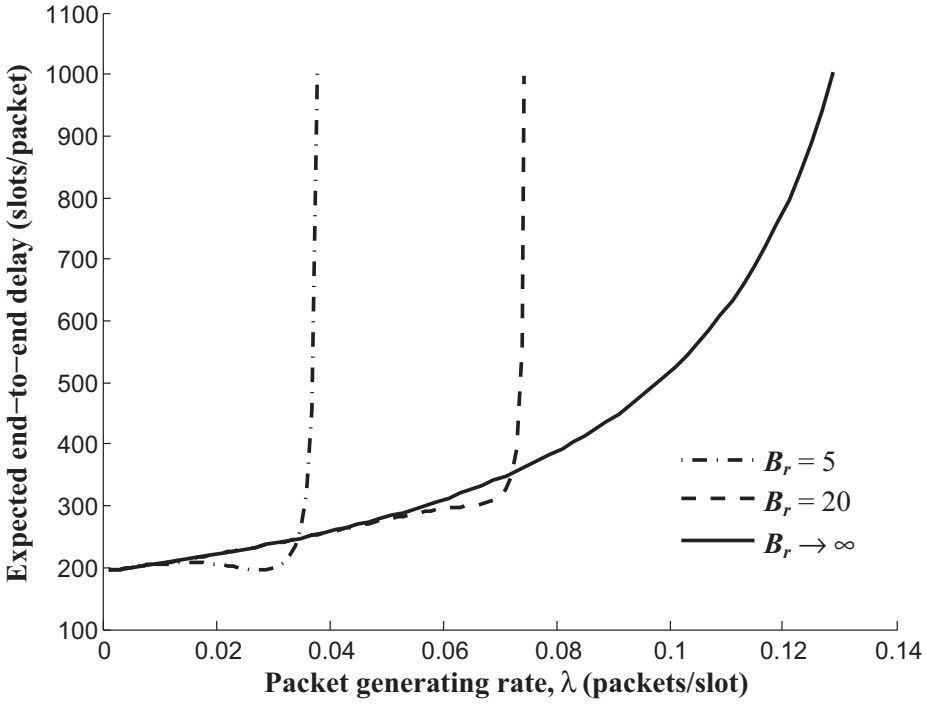


Figure 5.6: E2E delay vs. packet generating rate λ under different settings of relay buffer size. $n = 32, m = 4$.

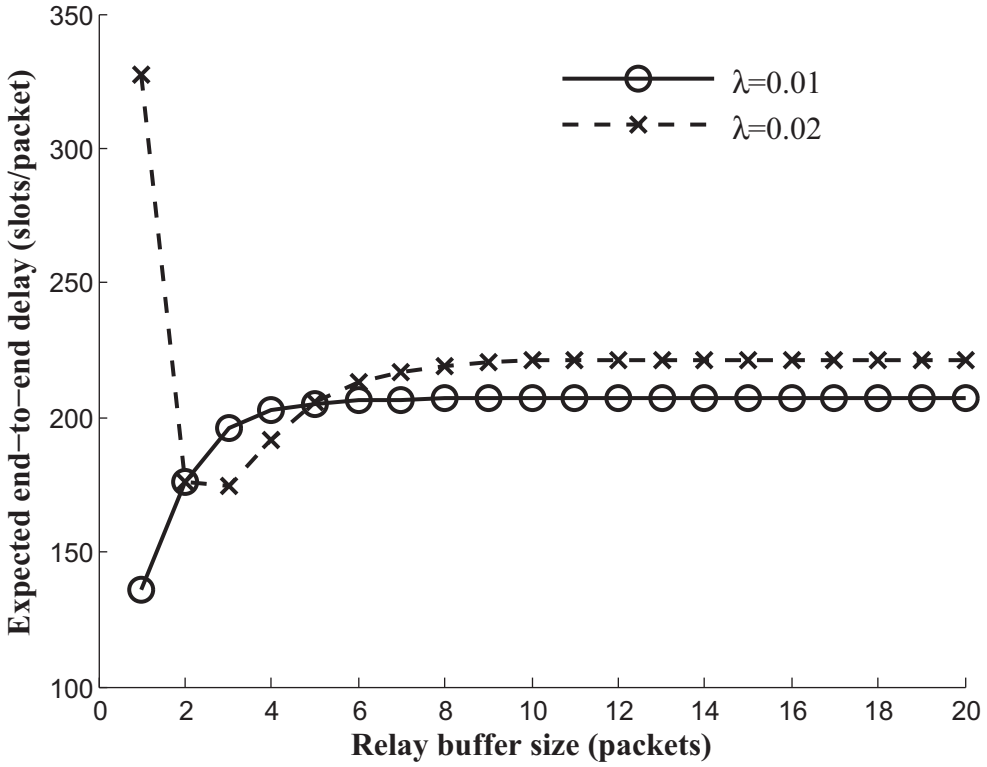


Figure 5.7: E2E delay vs. relay buffer size. $n = 32, m = 4$.

workload on $\mathbb{E}\{D_d\}$ are two folds. On one hand, a larger workload will lead to a longer relay queue length, which further leads to a higher delay in the relay queue; on the other hand, a larger workload will lead to a higher ROP, which further leads to a lower probability that a packet to be delivered by a two-hop way, such that $\mathbb{E}\{D_d\}$ decreases. Since the latter effect, the delivery delay under a small relay buffer is lower than that under a large one.

Fig. 5.6 shows the relationship between the expected E2E delay $\mathbb{E}\{D\}$ and packet generating rate λ. We can see that under the relay-buffer constraint, as λ increases, $\mathbb{E}\{D\}$ does not increase all the time because the delivery delay will decrease when λ exceeds a specific value, however when λ approaches the corresponding throughput capacity, $\mathbb{E}\{D\}$ increases sharply because the source-queuing delay tends to infinity. It also can be seen that when λ is small, $\mathbb{E}\{D\}$ under $B_r = 5$ is smaller than that under $B_r = 20$, since both of the source-queuing delay under the two settings are

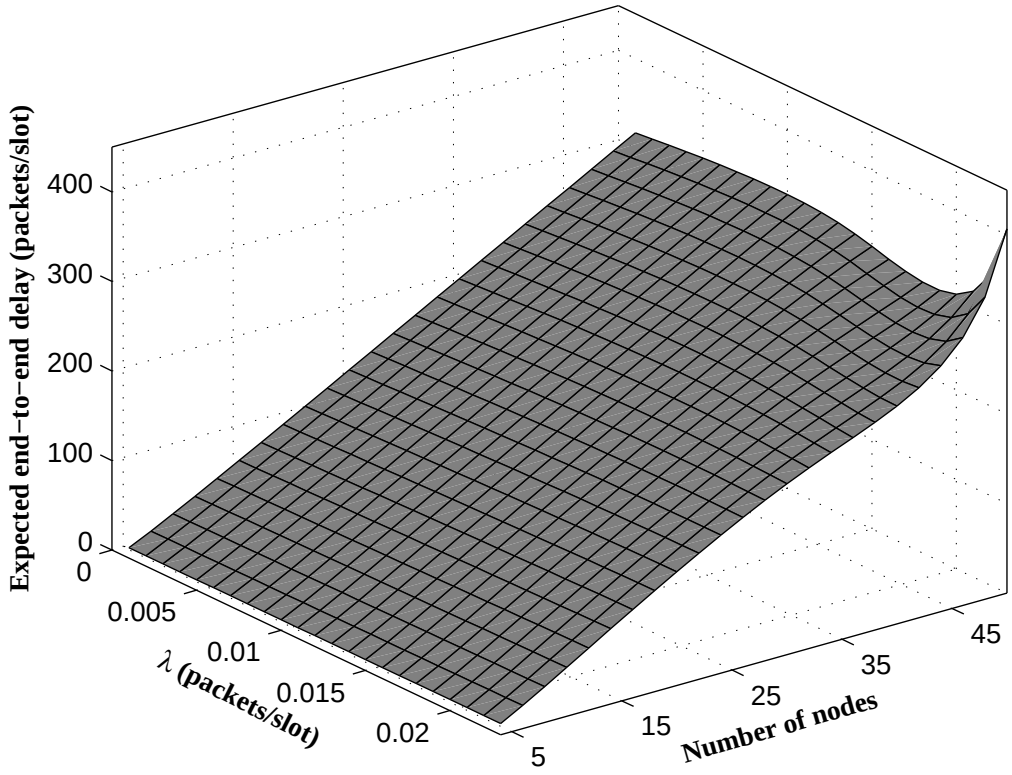


Figure 5.8: E2E delay vs. packet generating rate λ and number of nodes n. $B_r = 5$.

small, but a small relay buffer can lead to a small delivery delay. However, with λ getting larger and larger, $\mathbb{E}\{D\}$ under $B_r = 5$ finally exceeds that under $B_r = 20$, and tends to infinity earlier. It indicates that increasing the relay buffer size can ensure the E2E delay limited for a larger region of packet generating rate.

We illustrate in Fig. 5.7 how $\mathbb{E}\{D\}$ varies B_r under the settings of $(n = 32, m = 4, \lambda = \{0.01, 0.02\})$. According to formula (5.4), $T_c = 0.0227$ when $B_r = 1$, and T_c increases as B_r increases. Thus, for $\lambda = 0.01$ which is much smaller than 0.0227, $\mathbb{E}\{D\}$ increases as B_r increases and finally tends to a constant 206.92 which can be determined by formula (5.18). While for $\lambda = 0.02$ which is very close to the T_c under $B_r = 1$, $\mathbb{E}\{D\}$ under $B_r = 1$ is very large. With B_r increasing, the corresponding T_c increases, leading to the $\mathbb{E}\{D\}$ first decreases, then increases and finally tends to a

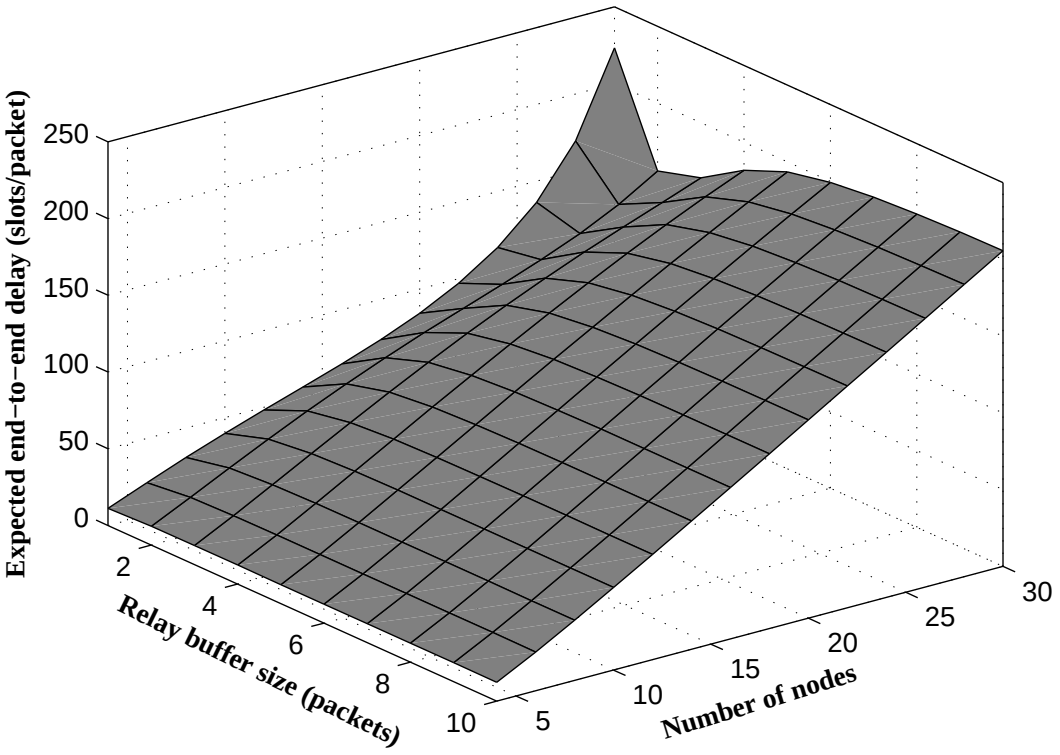


Figure 5.9: E2E delay vs. relay buffer size B_r and number of nodes n. $\lambda = 0.02$.

constant 221.65.

We further illustrate in Fig. 5.8 and Fig. 5.9 how $\mathbb{E}\{D\}$ be influenced by n, where another parameters of the 3D meshes are λ and B_r respectively, and the ratio of n to the number of cells keeps as 2. We can see that the variations of $\mathbb{E}\{D\}$ with n are complicated, but in general $\mathbb{E}\{D\}$ increases as n increases. A more careful observation is that when n increases, $\mathbb{E}\{D\}$ first increases almost linearly when λ is much smaller than T_c, then increases quickly when λ approaches T_c. For example, these behaviors can be found in Fig. 5.8 under $\lambda = 0.23$ and in Fig. 5.9 under $B_r = 1$.

5.6 Summary

This chapter developed a very general E2E delay modeling for the MANETs with relay-buffer constraint. We first applied the fixed-point theory to numerically solve the overflowing probability and the stationary OSD of a relay buffer. Then, based on a Bernoulli/Bernoulli queuing model and an absorbing Markov chain model, we analyzed the source-queuing delay and delivery delay of a packet respectively. Case studies are further conducted under a fully distributed DCF-style media access mechanism. Finally, we provided extensive simulations to demonstrate the efficiency and application of our delay modeling, and discussed some interesting theoretical findings about the impacts of network parameters on delay performance.

CHAPTER VI

Throughput and Delay of MANETs under General Buffer Constraint

In this chapter, we extend our results to the more practical MANETs with a general buffer constraint, i.e., not only the relay-buffer constraint but also the source-buffer constraint are considered for the MANET performance study. Notice that in the previous two chapters, packet loss is avoided by the feedback mechanism so that the throughput is equal to the packet generating rate λ when $\lambda \leq T_c$. Under the general buffer constraint, however, packet loss is inevitable since both the source buffer and relay buffer are limited. Thus, the achievable throughput under any given packet generating rate should also be carefully addressed. Based on the Queuing theory and birth-death chain theory, we first develop a general theoretical framework to fully depict the source/relay buffer occupancy process in such a buffer-limited MANET under both the scenarios with and without feedback. With the help of this framework, we then derive the exact expressions of achievable throughput, throughput capacity, and expected E2E delay. We also provide the related theoretical analysis to reveal some important properties of the network performance. Finally, we present extensive simulation and numerical results to demonstrate the efficiency of our theoretical framework and illustrate our theoretical findings.

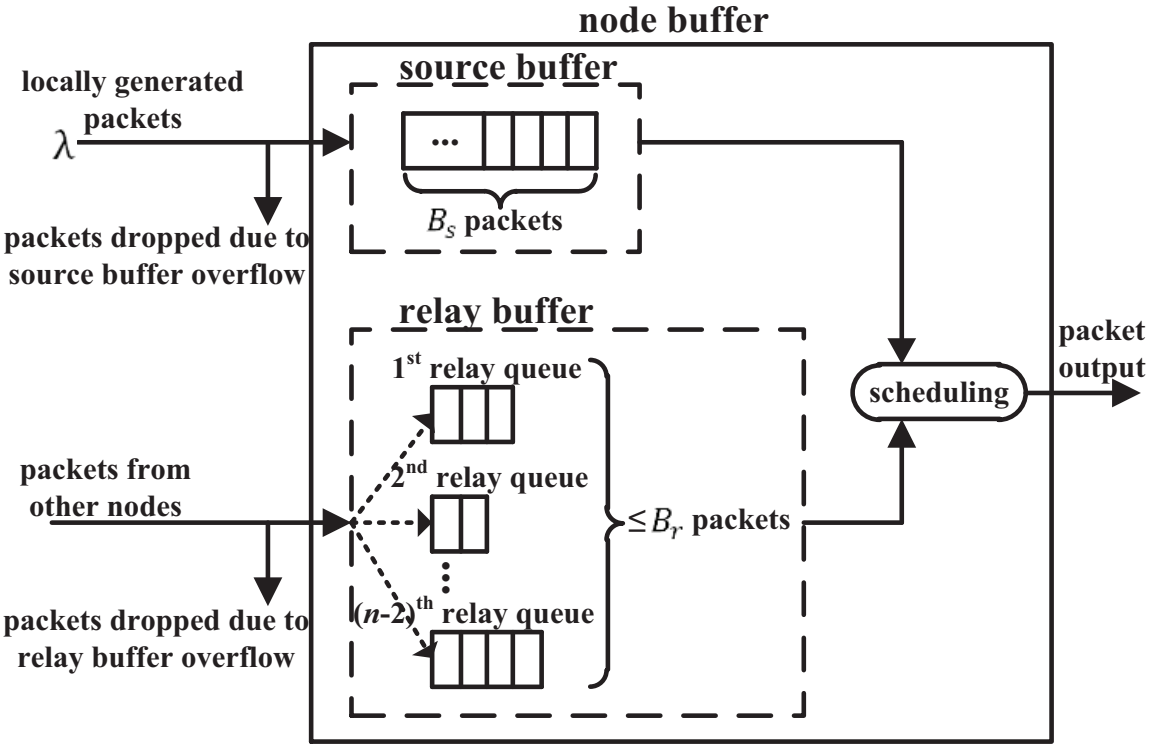


Figure 6.1: Illustration of general limited-buffer constraint.

6.1 General Buffer Constraint and Problem Formulation

As illustrated in Fig. 6.1, we consider a general limited-buffer constraint, where a node is equipped with a limited source buffer of size B_s and a limited relay buffer of size B_r. The source buffer is for storing the packets of its own flow (locally generated packets) and works as a FIFO (first-in-first-out) source queue, while the relay buffer is for storing packets of all other $n-2$ flows and works as $n-2$ FIFO virtual relay queues (one queue per flow). When a packet of other flows arrives and the relay buffer is not full, a buffer space is dynamically allocated to the corresponding relay queue for storing this packet; once a head-of-line (HoL) packet departs from its relay queue, this relay queue releases a buffer space to the common relay buffer. It is notable that this buffer constraint is general in the sense it covers all the buffer constraint assumptions adopted in available works as special cases, like the infinite buffer assumption [9, 11–

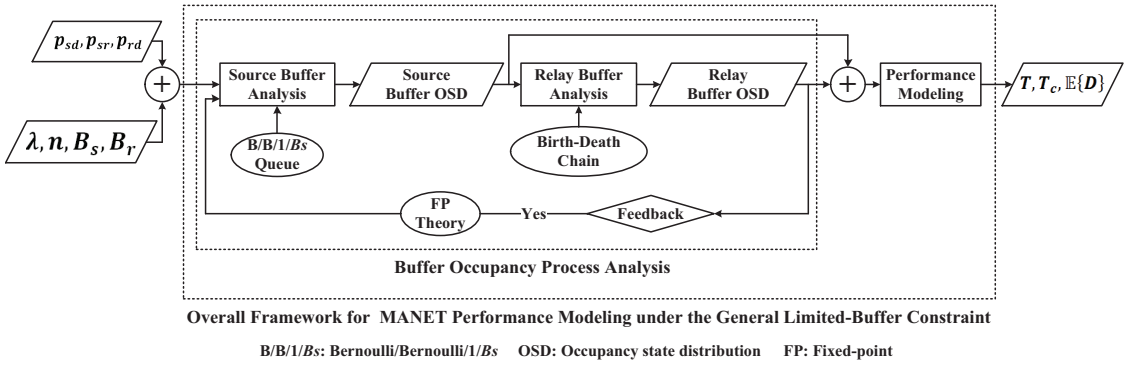


Figure 6.2: Illustration of overall framework for MANET performance modeling under the general limited-buffer constraint.

13, 22, 23, 29] ($B_s \to \infty$, $B_r \to \infty$), limited source buffer assumption [43] ($0 \leq B_s < \infty, B_r \to \infty$), and limited relay buffer assumption [41, 63, 64] ($B_s \to \infty, 0 \leq B_r < \infty$).

We show in Fig. 6.2 our overall theoretical framework for MANET performance modeling under the general buffer constraint. We can see from Fig. 6.2 that for the performance modeling of T, T_c and $\mathbb{E}\{D\}$, the key issue there is to determine the occupancy state distributions (OSDs) for both the source and relay buffers based on the basic parameters of $\{n, B_s, B_r, \lambda, p_{sd}, p_{sr}, p_{rd}\}$. In particular, due to the different arrival/departure processes associated with the source buffer and relay buffer, a Bernoulli/Bernoulli/1/B_s (B/B/1/B_s) queuing model[1] is applied to characterize the packet occupancy process in source buffer, while a birth-death chain is applied to model the complex packet occupancy process in relay buffer. Finally, the fixed-point (FP) theory is applied to deal with the coupling issue between the occupancy processes of source buffer and relay buffer under the scenario with feedback.

[1]A B/B/1/B_s queue refers to that both the packet arrival and departure are Bernoulli processes, the number of server is 1 and the buffer size is B_s.

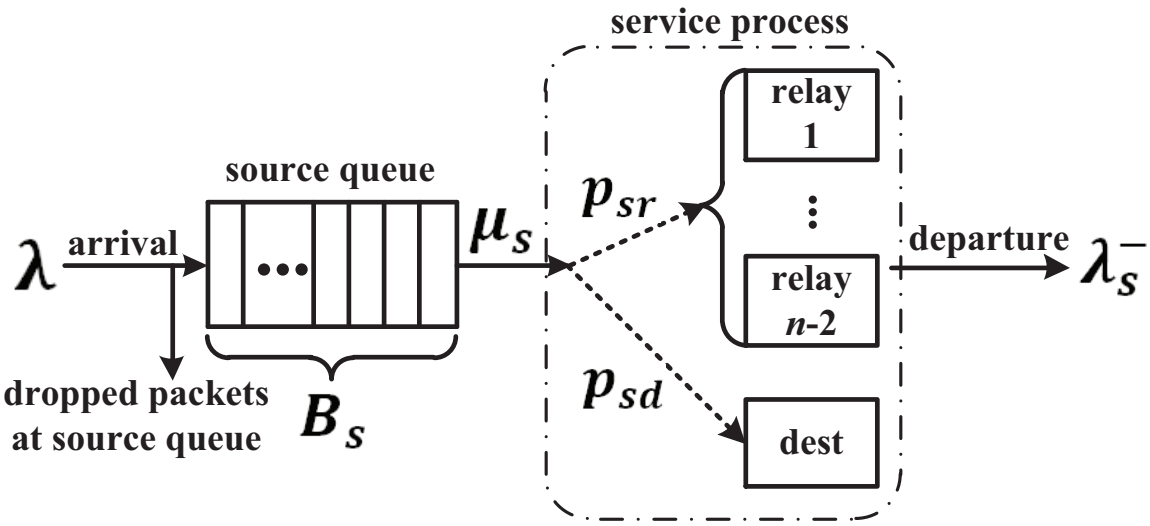


Figure 6.3: Bernoulli/Bernoulli/1/B_s queuing model for source buffer.

6.2 Buffer Occupancy Process Analysis

In this section, we conduct the occupancy process analysis for both the source and relay buffers to determine their OSDs, which helps us to derive the performance metrics of T, T_c and $\mathbb{E}\{D\}$. Without loss of generality, we focus on a tagged node S, and consider the scenarios without and with feedback separately.

6.2.1 OSDs under the Scenario without Feedback

6.2.1.1 OSD of Source Buffer

Regarding the source buffer of node S, since in every time slot a new packet is generated with probability λ and a service opportunity arises with probability μ_s being determined as

$$\mu_s = p_{sd} + p_{sr}, \tag{6.1}$$

the occupancy process of source buffer can be modeled as a B/B/1/B_s queue as illustrated in Fig. 6.3.

Let $\pi_s(i)$ denote the probability that there are i packets occupying the source

buffer in the stationary state, then the stationary OSD of the source buffer $\mathbf{\Pi}_s = [\pi_s(0), \pi_s(1), \cdots, \pi_s(B_s)]$ can be determined as [47]

$$\pi_s(i) = \begin{cases} \dfrac{1}{1-\lambda} H^{-1}, & i = 0 \\ \dfrac{1}{1-\lambda} \dfrac{\tau^i}{1-\mu_s} H^{-1}, & 1 \leq i \leq B_s \end{cases} \tag{6.2}$$

where

$$\tau = \frac{\lambda(1-\mu_s)}{\mu_s(1-\lambda)}, \tag{6.3}$$

and H is the normalizing constant. Notice that $\mathbf{\Pi}_s \cdot \mathbf{1} = 1$, where $\mathbf{1}$ is a column vector of size $(B_s + 1) \times 1$ with all elements being 1, we have

$$\pi_s(i) = \begin{cases} \dfrac{\mu_s - \lambda}{\mu_s - \lambda \cdot \tau^{B_s}}, & i = 0 \\ \dfrac{\mu_s - \lambda}{\mu_s - \lambda \cdot \tau^{B_s}} \dfrac{1}{1-\mu_s} \tau^i. & 1 \leq i \leq B_s \end{cases} \tag{6.4}$$

6.2.1.2 OSD of Relay Buffer

We continue to analyze the occupancy process of the relay buffer in node S. Similar to the theoretical framework established in Section 4.4, we let X_t denote the number of packets in the relay buffer at time slot t, then the occupancy process of the relay buffer $\{X_t, t = 0, 1, 2, \cdots\}$ can be modeled as a birth-death chain as illustrated in Fig. 4.3. Let $\pi_r(i)$ denote the probability that there are i packets occupying the relay buffer in the stationary state, the stationary OSD of the relay buffer $\mathbf{\Pi}_r = [\pi_r(0), \pi_r(1), \cdots, \pi_r(B_r)]$ is determined as

$$\mathbf{\Pi}_r \cdot \mathbf{P} = \mathbf{\Pi}_r, \tag{6.5}$$

$$\mathbf{\Pi}_r \cdot \mathbf{1} = 1, \tag{6.6}$$

where $\mathbf{P}$ is the one-step transition matrix of the birth-death chain defined as

$$\mathbf{P} = \begin{bmatrix} p_{0,0} & p_{0,1} & & & \\ p_{1,0} & p_{1,1} & p_{1,2} & & \\ & \ddots & \ddots & \ddots & \\ & & p_{B_r,B_r-1} & p_{B_r,B_r} \end{bmatrix}, \tag{6.7}$$

and $\mathbf{1}$ is a column vector of size $(B_r + 1) \times 1$ with all elements being 1.

Notice that $p_{0,0} = 1 - p_{0,1}$, $p_{B_r,B_r} = 1 - p_{B_r,B_r-1}$ and $p_{i,i} = 1 - p_{i,i-1} - p_{i,i+1}$ for $0 < i < B_r$, the expressions (6.5)−(6.7) indicate that to derive $\mathbf{\Pi}_r$, we need to determine the one-step transition probabilities $p_{i,i+1}$ and $p_{i,i-1}$. Regarding the calculation of $p_{i,i-1}$, from Lemma 1 we have

$$p_{i,i-1} = p_{rd} \cdot \frac{i}{n-3+i}, \ 1 \leq i \leq B_r. \tag{6.8}$$

Regarding the calculation of $p_{i,i+1}$, we can see that $p_{i,i+1}$ is actually equal to the packet arrival rate λ_r of the relay buffer, so we just need to determine λ_r for the evaluation of $p_{i,i+1}$. When S serves as a relay, all other $n-2$ nodes (except S and its destination) may forward packets to it. When one of these nodes sends out a packet from its source buffer, it will forward the packet to S with probability $\frac{p_{sr}}{\mu_s(n-2)}$. This is because with probability $\frac{p_{sr}}{\mu_s}$ the packet is intended for a relay node, and each of the $n-2$ relay nodes are equally likely. Thus,

$$p_{i,i+1} = \lambda_r = (n-2)\lambda_s^- \cdot \frac{p_{sr}}{\mu_s(n-2)}, \tag{6.9}$$

where λ_s^- denotes the packet departure rate of a source buffer. Due to the reversibility of the B/B/1/B_s queue, the packet departure process of the source buffer is also a

Bernoulli process with its departure rate λ_s^- being determined as

$$\lambda_s^- = \mu_s \left(1 - \pi_s(0)\right). \tag{6.10}$$

Then we have

$$p_{i,i+1} = \lambda_r = p_{sr} \cdot (1 - \pi_s(0)),\ 0 \leq i \leq B_r - 1. \tag{6.11}$$

By substituting (6.11) and (6.8) into (6.5) and (6.6), we can see that the stationary OSD of the relay buffer is determined as

$$\pi_r(i) = \frac{\mathrm{C}_i (1 - \pi_s(0))^i}{\sum_{k=0}^{B_r} \mathrm{C}_k (1 - \pi_s(0))^k},\ 0 \leq i \leq B_r \tag{6.12}$$

where $\mathrm{C}_i = \binom{n-3+i}{i}$.

6.2.2 OSDs under the Scenario with Feedback

Under the scenario with feedback, node S cannot execute a S-R transmission when the relay buffer of its intended receiver is full (with overflowing probability $\pi_r(B_r)$), causing the correlation between the OSD analysis of source buffer and that of relay buffer. It is notable, however, the overflowing probability $\pi_r(B_r)$ only affects the service rate μ_s of the source buffer and the arrival rate at the relay buffer, while the occupancy processes of the source buffer and relay buffer can still be modeled as the B/B/1/B_s queue and the birth-death chain respectively. Thus, based on the similar analysis as that in Section 6.2.1, we have the following corollary.

Corollary 7 *For the network scenario with feedback, the OSD $\mathbf{\Pi}_s$ of the source buffer and the OSD $\mathbf{\Pi}_r$ of the relay buffer are determined as (6.4) and (6.12), where τ is given by (6.3), and the service rate μ_s of the source buffer is evaluated as*

$$\mu_s = p_{sd} + p_{sr} \cdot (1 - \pi_r(B_r)). \tag{6.13}$$

Proof 8 *The proof is given in Appendix C.1.*

Corollary 7 indicates that for the evaluation of OSDs $\mathbf{\Pi}_s$ and $\mathbf{\Pi}_r$, we need to determine the relay-buffer overflowing probability $\pi_r(B_r)$. From formula (6.12) we have

$$\pi_r(B_r) = \frac{\mathrm{C}_{B_r}(1-\pi_s(0))^{B_r}}{\sum\limits_{k=0}^{B_r} \mathrm{C}_k(1-\pi_s(0))^k}, \tag{6.14}$$

where

$$\pi_s(0) = \frac{\mu_s - \lambda}{\mu_s - \lambda \cdot \tau^{B_s}} = \frac{\mu_s - \lambda}{\mu_s - \lambda \cdot \left(\frac{\lambda(1-\mu_s)}{\mu_s(1-\lambda)}\right)^{B_s}}. \tag{6.15}$$

We can see from (6.13)−(6.15) that (6.14) is actually an implicit function of $\pi_r(B_r)$, which can be numerically solved by the following fixed-point iteration algorithm.

Algorithm 3 Fixed-point iteration under general buffer constraint

Require:

Basic network parameters $\{n, B_s, B_r, \lambda, p_{sd}, p_{sr}, p_{rd}\}$;

Ensure:

Relay-buffer overflowing probability $\pi_r(B_r)$;

1: Set $x_1 = 0$ and $i = 1$;
2: **while** $(x_i - x_{i-1} \geq 10^{-6}) \vee (i = 1)$ **do**
3: $\quad i = i + 1$;
4: $\quad \mu_s = p_{sd} + p_{sr} \cdot (1 - x_{i-1})$;
5: $\quad \tau = \frac{\lambda(1-\mu_s)}{\mu_s(1-\lambda)}$;
6: $\quad \pi_s(0) = \frac{\mu_s - \lambda}{\mu_s - \lambda \cdot \tau^{B_s}}$;
7: $\quad x_i = \frac{\mathrm{C}_{B_r}(1-\pi_s(0))^{B_r}}{\sum\limits_{k=0}^{B_r} \mathrm{C}_k(1-\pi_s(0))^k}$;
8: **end while**
9: $\pi_r(B_r) = x_i$;
10: **return** $\pi_r(B_r)$;

6.3 Throughput and Delay Analysis

With the help of OSDs of source buffer and relay buffer derived in Section 6.2, this section focuses on the performance analysis of the concerned buffer-limited MANET

in terms of its throughput, expected E2E delay and throughput capacity.

6.3.1 Throughput and Expected E2E Delay

Regarding the throughput and expected E2E delay of a MANET with the general limited-buffer constraint, we have the following theorem.

Theorem VI.1 *For a concerned MANET with n nodes, packet generating rate λ, source buffer size B_s and relay buffer size B_r, its per flow throughput T and expected E2E delay $\mathbb{E}\{D\}$ are given by*

$$T = p_{sd}\left(1-\pi_s(0)\right) + p_{sr}\left(1-\pi_s(0)\right)\left(1-\pi_r(Br)\right), \tag{6.16}$$

$$\mathbb{E}\{D\} = \frac{1+L_s^*}{\mu_s} + \frac{(n-2+L_r^*)(1-\pi_r(B_r))}{p_{sd}+p_{sr}(1-\pi_r(B_r))}, \tag{6.17}$$

where L_s^* (resp. L_r^*) denotes the expected number of packets in the source buffer (resp. relay buffer) under the condition that the source buffer (resp. relay buffer) is not full, which is determined as

$$L_s^* = \frac{\tau - B_s\tau^{B_s} + (B_s-1)\tau^{B_s+1}}{(1-\tau)(1-\tau^{B_s})}, \tag{6.18}$$

$$L_r^* = \frac{1}{1-\pi_r(B_r)}\sum_{i=0}^{B_r-1} i\pi_r(i), \tag{6.19}$$

and μ_s is determined by (6.1) and (6.13) for the scenarios without and with feedback respectively, τ, $\pi_s(0)$ and $\mathbf{\Pi}_r$ are determined by (6.3), (6.4) and (6.12), respectively.

Proof 9 *Let T_1 and T_2 denote the packet delivery rates at the destination of node S through the one-hop transmission and the two-hop transmission respectively, then we*

have

$$T_1 = \lambda_s^- \cdot \frac{p_{sd}}{\mu_s}, \tag{6.20}$$

$$T_2 = \lambda_s^- \cdot \frac{p_{sr}\left(1 - \pi_r(B_r)\right)}{\mu_s}, \tag{6.21}$$

where λ_s^- *denotes the packet departure rate of source buffer of S. Substituting (6.10) into (6.20) and (6.21), then (6.16) follows from* $T = T_1 + T_2$.

Regarding the expected E2E delay $\mathbb{E}\{D\}$*, we focus on a tagged packet p of node S and evaluate its expected source-queuing delay* $\mathbb{E}\{D_{sq}\}$ *and expected delivery delay* $\mathbb{E}\{D_d\}$*, respectively. For the evaluation of* $\mathbb{E}\{D_{sq}\}$ *we have*

$$\mathbb{E}\{D_{sq}\} = \frac{L_s^*}{\mu_s}. \tag{6.22}$$

Let $\pi_s^*(i)$ *(*$0 \le i \le B_s - 1$*) denote the probability that there are i packets in the source buffer conditioned on that the source buffer is not full, then* $\pi_s^*(i)$ *is determined as [47]*

$$\pi_s^*(i) = \frac{\lambda}{(1-\lambda)^2}\tau^i \cdot H_1^{-1},\ 0 \le i \le B_r - 1 \tag{6.23}$$

where H_1 *is the normalizing constant. Since* $\sum_{i=1}^{B_s-1} \pi_s^*(i) = 1$*, we have*

$$\pi_s^*(i) = \frac{1-\tau}{1-\tau^{B_s}}\tau^i,\ 0 \le i \le B_r - 1.$$

Then L_s^* *is given by*

$$L_s^* = \sum_{i=0}^{B_s-1} i\pi_s^*(i) = \frac{\tau - B_s\tau^{B_s} + (B_s-1)\tau^{B_s+1}}{(1-\tau)(1-\tau^{B_s})}.$$

After moving to the HoL in its source buffer, packet p will be sent out by node S with mean service time $1/\mu_s$*, and it may be delivered to its destination directly or*

forwarded to a relay. Let $\mathbb{E}\{D_r\}$ denote the expected time that p takes to reach its destination after it is forwarded to a relay, then we have

$$\mathbb{E}\{D_d\} = \frac{1}{\mu_s} + \frac{T_1}{T_1 + T_2} \cdot 0 + \frac{T_2}{T_1 + T_2} \cdot \mathbb{E}\{D_r\}. \tag{6.24}$$

Based on the OSD $\mathbf{\Pi}_r$, L_r^* is given by (6.19). Due to the symmetry of relay queues in a relay buffer, the mean number of packets in one relay queue is $L_r^*/(n-2)$, and the service rate of each relay queue is $p_{rd}/(n-2)$. Thus, $\mathbb{E}\{D_r\}$ can be determined as

$$\mathbb{E}\{D_r\} = \left(\frac{L_r^*}{n-2} + 1\right) \Big/ \left(\frac{p_{rd}}{n-2}\right). \tag{6.25}$$

Substituting (6.25) into (6.24), then (6.17) follows from $\mathbb{E}\{D\} = \mathbb{E}\{D_{sq}\} + \mathbb{E}\{D_d\}$.

Based on the results of Theorem VI.1, we can establish the following corollary (See Appendix C.2 for the proof).

Corollary 8 *For a concerned MANET with the general limited-buffer constraint, adopting the feedback mechanism improves its throughput performance.*

6.3.2 Throughput Capacity and Limiting Throughput/Delay

To determine the throughput capacity T_c, we first need the following lemma (See Appendix C.3 for the proof).

Lemma 4 *For a concerned MANET with the general limited-buffer constraint, its throughput T increases monotonically as the packet generating rate λ increases.*

Based on Lemma 4, we can establish the following theorem on throughput capacity.

Theorem VI.2 *For a concerned MANET with n nodes, source buffer size B_s and relay buffer size B_r, its throughput capacity T_c is given by*

$$T_c = p_{sd} + p_{sr}\frac{B_r}{n-2+B_r}. \tag{6.26}$$

Proof 10 *Lemma 4 indicates that*

$$T_c = \max_{\lambda\in(0,1]} T = \lim_{\lambda\to 1} T. \tag{6.27}$$

From (6.3), (6.4) and (6.12) we can see that

$$\begin{aligned}
&\lim_{\lambda\to 1}\tau = \lim_{\lambda\to 1}\frac{\lambda(1-\mu_s)}{\mu_s(1-\lambda)} \to \infty,\\
&\lim_{\lambda\to 1}\pi_s(0) = \lim_{\lambda\to 1}\frac{\mu_s-\lambda}{\mu_s-\lambda\cdot\tau^{B_s}} = 0. &(6.28)\\
&\lim_{\lambda\to 1}\pi_r(B_r) = \frac{\mathrm{C}_{B_r}}{\sum_{k=0}^{B_r}\mathrm{C}_k} = \frac{n-2}{n-2+B_r}. &(6.29)
\end{aligned}$$

Combining (6.16), (6.27), (6.28) and (6.29), the expression (6.26) then follows.

Based on the Theorem VI.1 and Theorem VI.2, we have the following corollary regarding the limiting T and $\mathbb{E}\{D\}$ as the buffer size tends to infinity (See Appendix C.4 for the proof).

Corollary 9 *For a concerned MANET, its throughput increases as B_s and/or B_r increase, and as B_s and/or B_r tend to infinity, the corresponding limiting T and $\mathbb{E}\{D\}$ are determined as (6.28) and (6.29) respectively, where $\rho_s = \min\{\frac{\lambda}{\mu_s}, 1\}$.*

$$T = \begin{cases} p_{sd} \cdot \rho_s + p_{sr} \cdot \dfrac{\sum_{k=0}^{B_r-1} \mathrm{C}_k \rho_s^{k+1}}{\sum_{k=0}^{B_r} \mathrm{C}_k \rho_s^{k}}, & B_s \to \infty \quad (6.30a) \\ (p_{sd} + p_{sr})(1 - \pi_s(0)), & B_r \to \infty \quad (6.30b) \\ \min\{\lambda, p_{sd} + p_{sr}\}. & B_s \to \infty, B_r \to \infty \quad (6.30c) \end{cases}$$

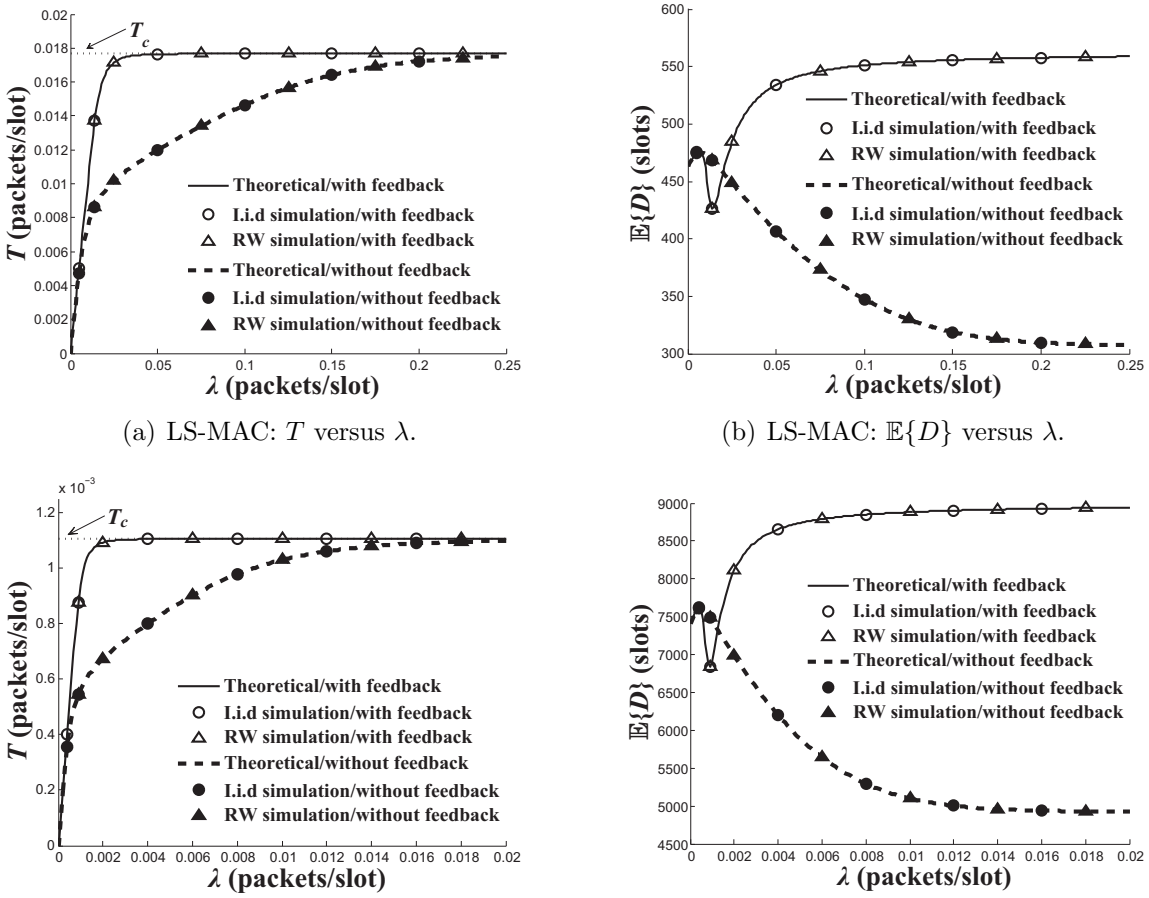
$$\mathbb{E}\{D\} = \begin{cases} \infty, & B_s \to \infty, \lambda \geq \mu_s \quad (6.31a) \\ \dfrac{1-\lambda}{\mu_s - \lambda} + \dfrac{(n-2+L_r^*)(1-\pi_r(B_r))}{p_{sd}+p_{sr}(1-\pi_r(B_r))}, & B_s \to \infty, \lambda < \mu_s \quad (6.31b) \\ \dfrac{n-2+\pi_s(0)\cdot(1+L_s^*)}{\pi_s(0)\cdot(p_{sd}+p_{sr})}, & B_r \to \infty \quad (6.31c) \\ \dfrac{n-1-\lambda}{p_{sd}+p_{sr}-\lambda}, & B_s \to \infty, B_r \to \infty, \lambda < \mu_s \quad (6.31d) \end{cases}$$

We can see from the Theorem VI.2 that the throughput capacity of the concerned MANET is the same for both the scenarios with and without feedback, and it is mainly determined by its relay buffer size B_r. The Corollary 9 indicates that our throughput and delay results of (6.16) and (6.17) are general in the sense that as B_s tends to infinity, they reduce to the results in [63, 64], while as both B_s and B_r tend to infinity, they reduce to the results in [11].

6.4 Simulation Results

6.4.1 Simulation Settings

To validate our theoretical framework for MANET performance modeling, a simulator was developed to simulate the packet generating, packet queuing and packet delivery processes in the cell-partitioned MANETs with LS-MAC and EC-MAC [65], and the DCF-style mechanism is adopted for the media access control. Each simula-



(a) LS-MAC: T versus λ.

(b) LS-MAC: $\mathbb{E}\{D\}$ versus λ.

(c) EC-MAC: T versus λ.

(d) EC-MAC: $\mathbb{E}\{D\}$ versus λ.

Figure 6.4: Performance validation.

tion task runs over a period of 2×10^8 time slots, and we only collect data from the last 80% of time slots to ensure the system is in the steady state. In the simulator, the i.i.d mobility model and random walk model have been implemented.

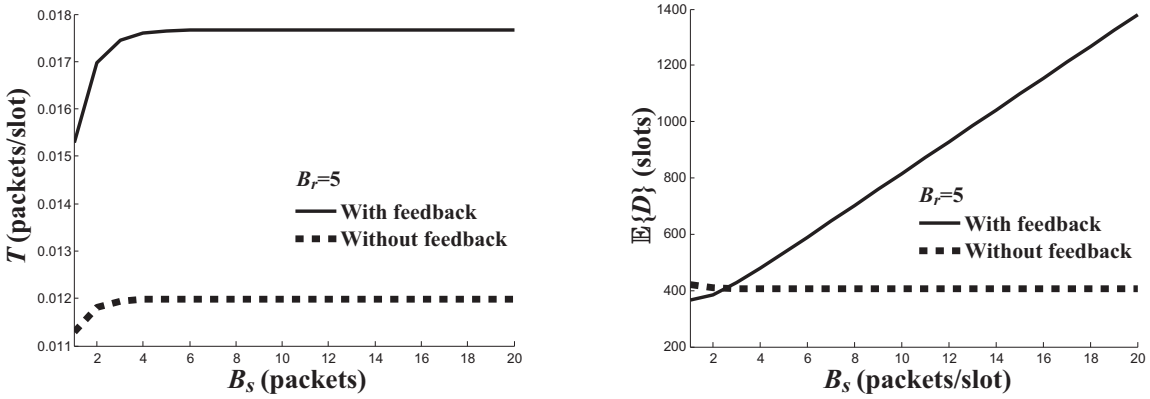
6.4.2 Validation of Throughput and Delay Results

We summarize in Fig. 6.4 the theoretical/simulation results for throughput and delay under the above two network scenarios, respectively. For each scenario we consider the network setting of $(n = 72, m = 6, B_s = 5, B_r = 5)$, and for the scenario with the EC-MAC protocol we set $\nu = 1$ and $\Delta = 1$ there [57]. Notice that the theoretical results here are obtained by substituting (5.19) and (5.20) (resp. (5.21)

and (5.22) into the theoretical framework in Fig. 6.2.

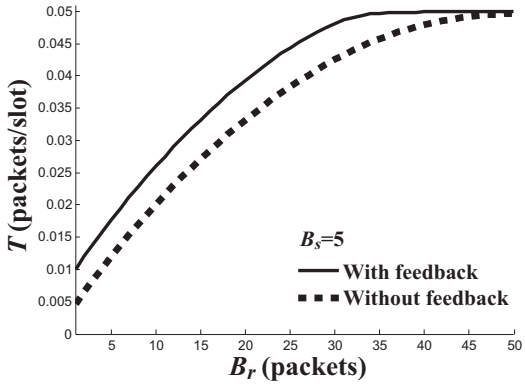
Fig. 6.4 shows clearly that the simulation results match well with the theoretical ones for all the cases considered here, which indicates that our theoretical framework is applicable to and highly efficient for the performance modeling of different buffer-limited MANETs. We can see from Fig. 6.4(a) and Fig. 6.4(c) that for a MANET with LS-MAC or EC-MAC, as the packet generating rate λ increases, the per flow throughput T increases monotonically and finally converges to its throughput capacity T_c, which agrees with the conclusions of Lemma 4 and Theorem VI.2. Another interesting observation of Fig. 6.4(a) and Fig. 6.4(c) is that just as predicated by Corollary 8 and Theorem VI.2, although adopting the feedback mechanism usually leads to a higher throughput, it does not improve the throughput capacity performance.

Regarding the delay performance, we can see from Fig. 6.4(b) and Fig. 6.4(d) that in a MANET with either LS-MAC or EC-MAC, the behavior of expected E2E delay $\mathbb{E}\{D\}$ under the scenario without feedback is quite different from that under the scenario with feedback. As λ increases, in the scenario without feedback $\mathbb{E}\{D\}$ first slightly increases and then decreases monotonically, while in the scenario with feedback $\mathbb{E}\{D\}$ first slightly increases, then decreases somewhat and finally increases monotonically. The results in Fig. 6.4 indicate that although adopting the feedback mechanism leads to an improvement in per flow throughput, such improvement usually comes with a cost of a larger E2E delay. This is because that the feedback mechanism can avoid the packet dropping at a relay node, which contributes to the throughput improvement but at the same time makes the source/relay buffers tend to be more congested, leading to an increase in delay.

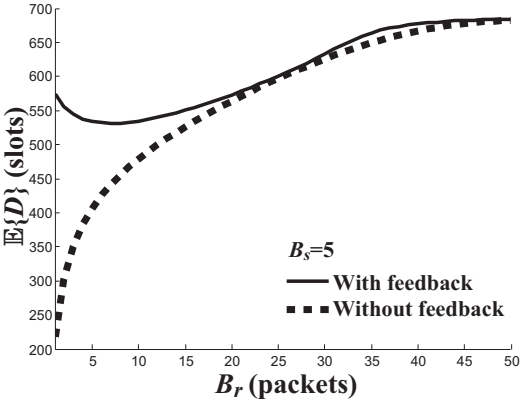


(a) T versus B_s.

(b) $\mathbb{E}\{D\}$ versus B_s.



(c) T versus B_r.



(d) $\mathbb{E}\{D\}$ versus B_r.

Figure 6.5: Throughput and delay versus B_s and B_r for the network setting of ($n = 72$, $m = 6$, $\lambda = 0.05$).

6.4.3 Discussions

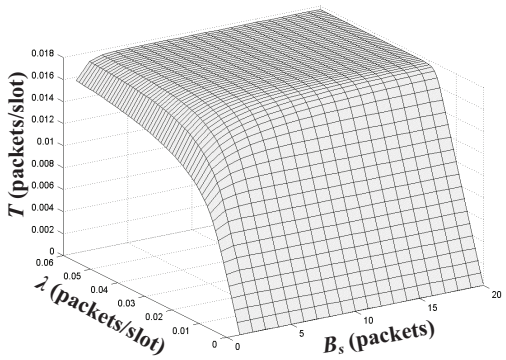
Based on the proposed theoretical framework, we further presents extensive numerical results to illustrate the potential impacts of buffer constraint on network performance. Notice from Section 6.4.2 that the performance behaviors of the LS-MAC are quite similar to that of the EC-MAC, in the following discussions we only focus on a MANET with the LS-MAC.

6.4.3.1 Throughput and E2E Delay

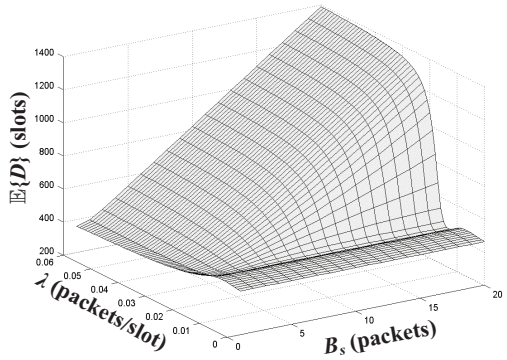
We first summarize in Fig. 6.5 how T and $\mathbb{E}\{D\}$ vary with B_s and B_r under the setting of ($n = 72$, $m = 6$, $\lambda = 0.05$). About the throughput performance, we can see

from Fig.6.5(a) and Fig.6.5(c) that just as predicated by Corollary 9 and Corollary 8, T increases as either B_s or B_r increases, and the feedback mechanism can lead to an improvement in T. It is interesting to see that as B_s increases, T under the two scenarios without and with feedback converges to two distinct constants determined by (6.28a). As B_r increases, however, T under the two scenarios finally converges to the same constant determined by (6.28b). Regarding the delay performance, Fig. 6.5(b) shows that as B_s increases, $\mathbb{E}\{D\}$ under the scenario without feedback quickly converges to a constant determined by (6.29b), while $\mathbb{E}\{D\}$ under the scenario with feedback monotonically increases to infinity, which agrees with the result of (6.29a). We can see from Fig. 6.5(d) that with the increase of B_r, however, $\mathbb{E}\{D\}$ under the scenario without feedback monotonically increases, while $\mathbb{E}\{D\}$ under the scenario with feedback first decreases and then increases. Similar to the throughput behavior in Fig. 6.5(c), Fig. 6.5(d) shows that as B_r increases $\mathbb{E}\{D\}$ under the two scenarios also converges to the same constant determined by (6.29c).

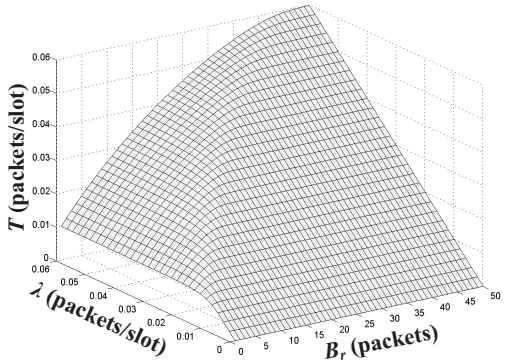
The results in Fig. 6.5 indicate that B_s and B_r have different impacts on the network performance in terms of T and $\mathbb{E}\{D\}$. In particular, as B_s increases, a notable performance gap between the scenarios without and with feedback always exist, where the throughput gap converges to a constant but the corresponding delay gap tends to infinity. As B_r increases, however, the performance gap between the two scenarios tends to decrease to 0, which implies that the benefits of adopting the feedback mechanism are diminishing in MANETs with a large relay buffer size. A further careful observation of Fig. 6.5 indicates that although we can improve the throughput by increasing B_s or B_r, it is more efficient to adopt a large B_r rather than a large B_s for such improvement. For example, under the scenario without feedback, Fig. 6.5(a) shows that by increasing B_s from 1 to 20, T can be improved from 0.0113 to 0.0120 (with an improvement of 6.19%); while Fig. 6.5(c) shows that by increasing B_r from 1 to 20, T can be improved from 0.0046 to 0.0332 (with an improvement of



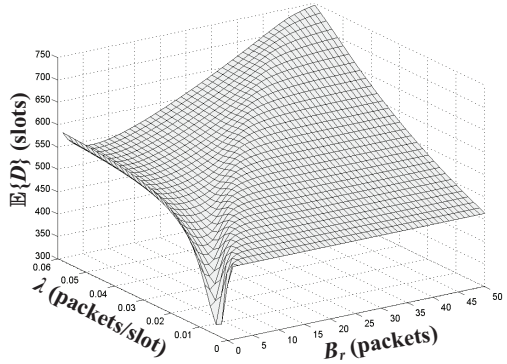
(a) T versus (λ, B_s), $B_r = 5$.



(b) $\mathbb{E}\{D\}$ versus (λ, B_s), $B_r = 5$.



(c) T versus (λ, B_r), $B_s = 5$.



(d) $\mathbb{E}\{D\}$ versus (λ, B_r), $B_s = 5$.

Figure 6.6: Throughput and delay versus (λ, B_s) and (λ, B_r) for the network setting of ($n = 72$, $m = 6$).

621.74%).

To further illustrate how the impacts of buffer size on network performance are dependent on packet generating rate λ, we focus on a MANET with feedback and summarize in Fig. 6.6 how its throughput and delay vary with λ and (B_s, B_r). We can see from Fig. 6.6(a) and Fig. 6.6(c) that although in general we can improve T by increasing either B_s or B_r, the degree of such improvement is highly dependent on λ. As λ increases, the throughput improvement from B_r monotonically increases, while the corresponding improvement from B_s first increases and then decreases. Fig. 6.6(a) and Fig. 6.6(c) also show that as λ increases, T under different settings of B_s finally converges to the same constant (i.e., T_c given by (6.26)), while T under a given setting of B_r converges to a distinct constant of T_c, which monotonically increases as

B_r increases.

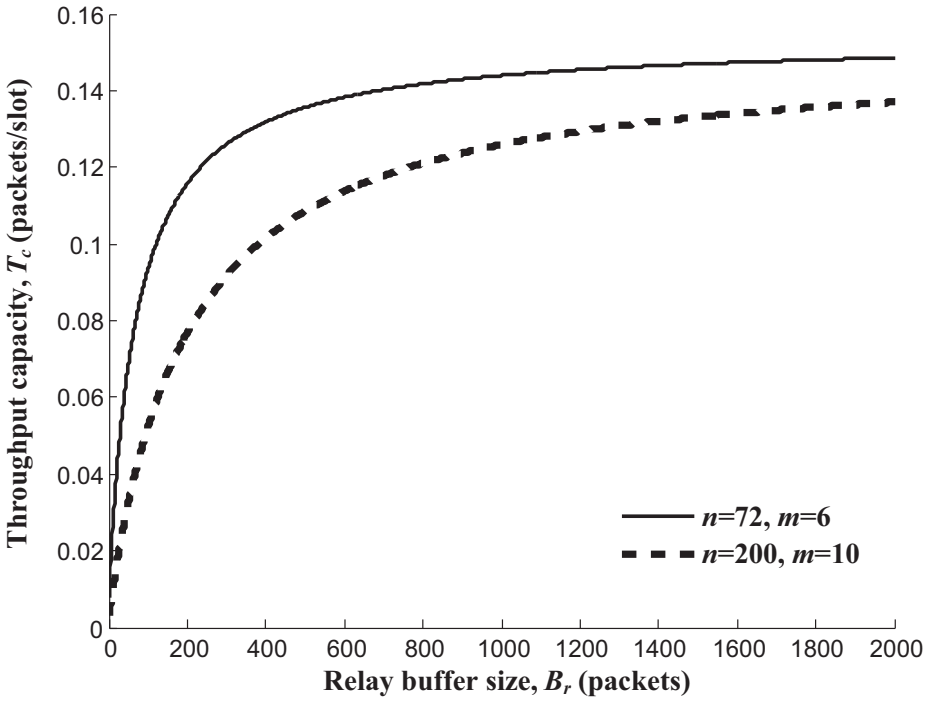
Regarding the joint impacts of λ and B_s on delay performance, we can see clearly from Fig. 6.6(b) that just as discussed in Corollary 9, there exists a threshold of λ beyond which $\mathbb{E}\{D\}$ will increases to infinity as B_s increases, while for a given λ less than the threshold, $\mathbb{E}\{D\}$ almost keeps as a constant as B_s increases. About the joint impacts of λ and B_r on delay performance, Fig. 6.6(d) shows that for a given setting of λ, there also exists a threshold for B_r, beyond which $\mathbb{E}\{D\}$ almost keeps as a constant as B_r increases. It is interesting to see that such threshold for B_r and the corresponding delay constant tend to increase as λ increases. The results in Fig. 6.6(d) imply that a bounded $\mathbb{E}\{D\}$ can be always guaranteed in a MANET as long as its source buffer size is limited.

6.4.3.2 Throughput Capacity

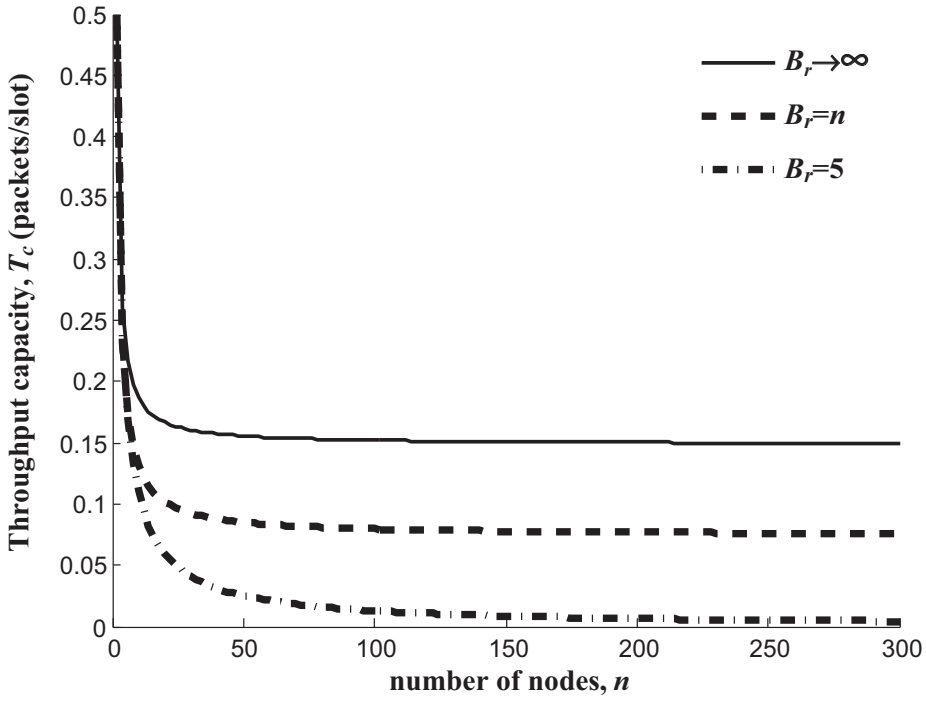
We summarize in Fig. 6.7(a) how throughput capacity T_c varies with relay buffer size B_r, where two network settings of $(n = 72, m = 6)$ and $(n = 200, m = 10)$ are considered. Fig. 6.7(a) shows that as B_r increases, T_c first increases quickly and then gradually converges to a constant $p_{sd} + p_{sr}$ being determined by (6.26). This observation indicates that although the throughput capacity can be improved by adopting a larger relay buffer, in practical network design the relay buffer size should be set appropriately according to the requirement on network capacity such that a graceful tradeoff between network performance and networking cost can be achieved.

It can be observed from Fig. 6.7(a) that T_c is also dependent on the number of nodes n, which motivates us to further explore the scaling law of throughput capacity in such a buffer-limited MANET. Based on (6.26), (5.19) and (5.20), the asymptotic throughput capacity is given by

$$\lim_{n\to\infty} T_c = \frac{1 - e^{-d} - de^{-d}}{2d} \frac{B_r}{n - 2 + B_r}, \tag{6.32}$$



(a) T_c vs B_r.



(b) T_c vs n.

Figure 6.7: Throughput capacity T_c versus relay buffer size B_r and number of nodes n.

where $d = n/m^2$.

From (6.32) we can see that as d tends to either 0 or infinity, T_c tends to 0, while if d is fixed, T_c scales as $\Theta(B_r/n)$ as both n and m^2 scale up. It is notable that in [41] an upper bound of throughput (with the notation O) was proposed for a MANET with limited relay buffer, however, the scaling law developed here is an achievable one (with the notation Θ), which indicates that to achieve a non-vanishing throughput capacity in a MANET with the general limited-buffer constraint, the relay buffer size B_r should grow at least linearly with the number of nodes n. Based on (6.26), we plot in Fig. 6.7(b) that how T_c scales with n under three typical buffer settings, i.e., B_r is fixed as a constant (5 here), $B_r = n$ and $B_r \to \infty$. We can see from Fig. 6.7(b) that in general T_c decreases as n increases, and T_c vanishes to 0 when B_r is fixed, while it converges to a non-zero constant when $B_r = n$ or $B_r \to \infty$.

6.4.3.3 Two-Hop Relay VS. Multi-Hop Relay

As mentioned in Section 3.2, we adopt the 2HR scheme since it can be implemented easily in a distributed way yet efficient in the sense that it has the capability of achieving the throughput capacity for many important MANET scenarios. The more important point is that, under the limited-buffer constraint the multi-hop relay scheme is considered to be inefficient mainly due to the following two reasons.

- If we consider the no redundancy case, that is to say, when a relay node A meets another relay node B, relay A just forwards the original packet to relay B and do not maintain this packet any more. It would be very inefficient because in a statistical view, the probability that relay A meets the destination is equal to the probability that relay B meets the destination. Thus, the transmission from relay A to relay B is meaningless and wastes the precious transmission opportunity.

- If we consider the redundancy case, that is to say, when a relay node A meets another relay node B, relay A forwards the original packet to relay B and meanwhile maintains the copy of this packet. It also would be very inefficient in the sense that it will cause heavy packet dropping problem because of the limited-buffer constraint.

6.5 Summary

This chapter explored the performance modeling for MANETs under the general limited-buffer constraint. In particular, a complete and general theoretical framework was developed to capture the inherent buffer occupancy behaviors in such a MANET, which enables the exact expressions to be derived for some fundamental network performance metrics, like the achievable throughput, expected E2E delay and throughput capacity. Some interesting conclusions that can be drawn from this study are: 1) In general, adopting the feedback mechanism can lead to an improvement in the throughput performance, but such improvement comes with the cost of a relatively large delay; 2) For the purpose of throughput improvement, it is more efficient to adopt a large relay buffer rather than a large source buffer; 3) The throughput capacity is dominated by the relay buffer size (rather than source buffer size) and the number of nodes; 4) To ensure that a buffer-limited MANET is scalable in terms of throughput capacity, its relay buffer size should grow at least linearly with the number of network nodes.

CHAPTER VII

Conclusion

This chapter summarizes the thesis and discusses the future research directions.

7.1 Summary of the Thesis

In this thesis, we studied the actual achievable performance of MANETs under limited-buffer constraint. The main contributions are summarized as follows.

- We first considered the relay-buffer constraint and studied the throughput capacity under a general MANET scenario with the 2HR-α scheme. For such a MANET, we analyzed the relationship between its throughput capacity and the relay-buffer overflowing probability. Based on the birth-death chain model, we developed a general theoretical framework to fully characterize the occupancy process of the relay buffer, which applies to any distributed MAC protocol and any mobility model that leads to the uniform distribution of node's locations in steady state. With the help of the proposed theoretical framework, we derived the throughput capacity of such a MANET in closed-form. Based on the closed-form expression, we further demonstrated that the throughput capacity can be improved by adjusting the transmission scheme, and revealed that how set the optimal transmission control parameter according to the relay buffer size.

- We next developed a theoretical modeling for the delay performance study in MANETs with relay-buffer constraint. Combining the buffer occupancy process analysis, we utilized the fixed-point theory to solve the relay-buffer overflowing probability under any given packet generating rate, and obtained the corresponding stationary occupancy state distribution. Based on these, we applied the Bernoulli/Bernoulli queuing model to compute the expected source-queuing delay, and developed an absorbing Markov chain model to characterize the packet deliver process for the delivery delay evaluation, such that the exact expression of the expected E2E delay can be derived.

- We finally extended our results to the MANET scenario with a general buffer constraint, where both the source buffer and relay buffer are limited, and both the transmission schemes with and without feedback are considered. We developed a B/B/1/B_s queuing model and a birth-death chain model to analyze the occupancy processes of source buffer and relay buffer, respectively, and applied the fixed-point theory to deal with the coupling issue under the scenario with feedback, such that the occupancy state distributions of source buffer and relay buffer are obtained. Based on these, we derived the exact expressions of achievable throughput, end-to-end delay and throughput capacity, and revealed some important features of the concerned MANETs. The theoretical findings proposed in the thesis are expected to provide some useful insights into the practical MANET design, implementation and optimization.

7.2 Future Works

The potential research directions to extend this thesis are summarized as follows.

- The performance study of MANETs conducted in this thesis still relied on some ideal assumptions. For example, we considered a time-slotted system and

assumed that once a node gets access to the wireless channel, it can transmit a fixed amount of data to its receiver. However, the time evolves continuously, and the distance between a pair of nodes and the corresponding meeting time (dominated by the speeds and moving directions of the nodes) significantly affect the data can be transmitted. Moreover, the time-varying channel fading could lead to the transmission failure even though some node contends for the transmission opportunity successfully. Although removing these simplifications will make the performance study of MANET a highly challenging problem, it always serves as a very appealing future direction and it is really worth making progress step by step.

- In this thesis, we considered that the packet generating rate of each network node is the same, which represents a kind of homogeneous traffic pattern. With the homogeneous traffic pattern, the network level performance reduces to the per flow performance such that we can focus on any one flow to conduct analysis. However, under the heterogeneous traffic pattern, the stably supportable packet generating rate of each node constitute a network level throughput region. Therefore, the studies of network throughput region and the corresponding delay performance of MANETs with heterogeneous traffic pattern are of great interest.

- It is notable that in this thesis, only the "fresh" packets could be dropped, i.e., packet loss only occurs when a packet is generated by its source node while the source buffer is full, or a packet is delivered to a relay node while the relay buffer is full. However, in some real-time applications such as the battlefield communications, the information contained in a packet has a period of validity, exceeding which the information may not be useful anymore. Moreover, in some cases, different packets have different priorities according to the importance of

the contained information. As a result, under the limited-buffer constraint, when the buffer resource is fully occupied, designing an efficient packet dropping strategy could improve the network performance. Therefore, combining the buffer constraint with the considerations of packet lifetime and priority will provide a more realistic model for MANETs and can be a very interesting research direction.

APPENDICES

APPENDIX A

Proofs of Chapter IV

A.1 Proof of Lemma 1

Based on the transition scenarios, we can see $p_{i,i+1}$ is actually equal to the packet arrival rate λ_r of the relay buffer, so we just need to determine λ_r for the evaluation of $p_{i,i+1}$. When S serves as a relay, all other $n-2$ nodes (except S and its destination) may forward packets to it. When one of these nodes sends out a packet from its source buffer, it will forward the packet to S with probability $\frac{p_{sr}(1-p_o(\lambda))}{\mu_s(n-2)}$. This is because with probability $\frac{p_{sr}p_o(\lambda)}{\mu_s}$ the packet is intended for a relay node, and each of the $n-2$ relay nodes are equally likely. Due to the reversibility of the Bernoulli/Bernoulli queue, the packet departure process of the source buffer is also a Bernoulli process with rate λ. Thus, we have

$$\begin{aligned}\lambda_r \cdot (1-p_o(\lambda)) + 0 \cdot p_o(\lambda) &= (n-2)\lambda \cdot \frac{p_{sr}\,(1-p_o(\lambda))}{\mu_s(\lambda)} \Big/ (n-2) \\ &= \rho_s(\lambda) p_{sr}\,(1-p_o(\lambda))\,, \end{aligned} \tag{A.1}$$

$$p_{i,i+1} = \lambda_r = \rho_s(\lambda) \cdot p_{sr}.$$

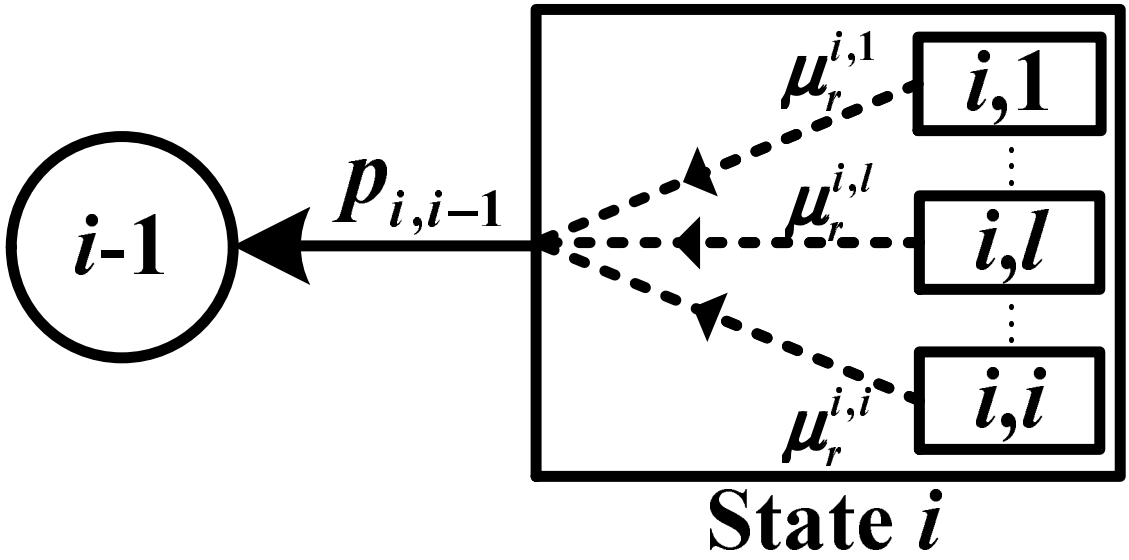


Figure A.1: Illustration of state decomposition.

Regarding the evaluation of transition probability $p_{i,i-1}$, it is notable that $p_{i,i-1}$ just corresponds to the service rate μ_r^i of the relay buffer when it is at state i. To determine μ_r^i, we further decompose the state i ($i > 0$) into i sub-states $\{(i,l), 1 \leq l \leq i\}$ as illustrated in Fig. A.1, where l denotes the number of non-empty relay queues in the relay buffer. Let $\mu_r^{i,l}$ denote the service rate of the relay buffer when it is at sub-state (i,l), and let $P_{l|i}$ denote the probability that the relay buffer is at sub-state (i,l) conditioned on that the relay buffer is at state i, we then have

$$\mu_r^i = \sum_{l=1}^{i} P_{l|i} \cdot \mu_r^{i,l}. \tag{A.2}$$

We first derive the term $\mu_r^{i,l}$ in (A.2). Notice that with probability p_{rd} the node S conducts a R-D transmission, and it will equally likely choose one of the $n-2$ nodes (expect S and its destination) as its receiver. Thus, when there are l non-empty relay queues in the relay buffer, the corresponding service rate $\mu_r^{i,l}$ is determined as

$$\mu_r^{i,l} = l \cdot \frac{p_{rd}}{n-2}. \tag{A.3}$$

To determine the conditional probability $P_{l|i}$, we utilize the following Occupancy technique [66]. Considering the relay buffer at state i, where each of these i buffered packets may be destined for any one of the other $n-2$ nodes, the number of all possible cases N_i is determined as

$$N_i = \binom{n-3+i}{i}. \tag{A.4}$$

Considering the condition that these i packets are destined for only l different nodes, then the number of possible cases $N_{l|i}$ is determined as

$$N_{l|i} = \binom{n-2}{l} \cdot \binom{(l-1)+(i-l)}{i-l}. \tag{A.5}$$

Since each of these cases occurs with equal probability, according to the *Classical Probability* $P_{l|i}$ is then determined as

$$P_{l|i} = \frac{N_{l|i}}{N_i} = \frac{\binom{n-2}{l} \cdot \binom{i-1}{i-l}}{\binom{n-3+i}{i}}. \tag{A.6}$$

It can be easily verified that $\sum_{l \leq i} P_{l|i} = 1$.

Substituting (A.3) and (A.6) into (A.2), $p_{i,i-1}$ is determined as

$$p_{i,i-1} = \mu_r^i = p_{rd} \cdot \frac{i}{n-3+i},\ 1 \leq i \leq B_r.$$

A.2 Proofs of Corollaries 1, 2 and 3

Proof of Corollary 1: Let $s_k = \frac{C_k \cdot \beta^k}{\sum_{i=0}^{k} C_i \cdot \beta^i}$, then

$$\frac{s_{k+1}}{s_k} = \frac{\mathrm{C}_{k+1}\beta^{k+1}\sum_{i=0}^{k}\mathrm{C}_i\cdot\beta^i}{\mathrm{C}_k\beta^k\sum_{i=0}^{k+1}\mathrm{C}_i\cdot\beta^i}$$
$$= \frac{\sum_{i=0}^{k}(n-2+k)\mathrm{C}_i\cdot\beta^{i+1}}{1+k+\sum_{i=0}^{k}(k+1)\mathrm{C}_{i+1}\cdot\beta^{i+1}},$$

Since

$$(k+1)\mathrm{C}_{i+1} = (k+1)\frac{n-2+i}{i+1}\cdot\mathrm{C}_i,$$

and

$$(k+1)(n-2+i) - (i+1)(n-2+k) = (n-3)(k-i) \geq 0,$$

then

$$(k+1)\mathrm{C}_{i+1} \geq (n-2+k)\mathrm{C}_i,$$
$$\frac{s_{k+1}}{s_k} < 1,$$

Substituting the result into (4.11), then Corollary 1 follows.

Proof of Corollary 2: When $\alpha = 0.5$, then $\beta = 1$, and (4.11) is simplified as

$$T_c = p_{sd} + p_{sr}\left(1 - \frac{\mathrm{C}_B}{\sum_{i=0}^{B_r}\mathrm{C}_i}\right). \tag{A.7}$$

Since

$$\sum_{i=0}^{B_r}\mathrm{C}_i = \frac{1}{(n-3)!}\left[(n-3)\times(n-4)\cdots\times 1\right.$$
$$\left.+\ (n-2)\times\cdots\times 2+\cdots+(n-3+B_r)\times\cdots\times(B_r+1)\right]$$
$$= \frac{1}{(n-3)!}\cdot\frac{(n-2+B_r)\times\cdots\times(B_r+1)}{n-2}$$
$$= \binom{n-2+B_r}{B_r}, \tag{A.8}$$

substituting (A.8) into (A.7), then Corollary 2 the follows.

Proof of Corollary 3: For the case $\alpha = 0.5$, since $\lim_{B_r \to \infty} \frac{B_r}{n-2+B_r} = 1$, substituting it into (4.15) we have

$$\underset{\alpha=0.5, B_r \to \infty}{T_c} = p_{sd} + p_{sr}.$$

For the case $\alpha < 0.5$, we have $\beta < 1$ and

$$\begin{aligned}
\sum_{i=0}^{B_r} \mathrm{C}_i \cdot \beta^i &= \frac{1}{(n-3)!} \times \big[(n-3) \times \cdots \times 1 \times \beta^0 + (n-2) \times \cdots \times 2 \times \beta^1 \\
&\quad + \cdots + (n-3+B_r) \times \cdots \times (B_r+1) \times \beta^{B_r}\big] \\
&= \frac{1}{(n-3)!} \cdot \left(\beta^{n-3} + \beta^{n-2} + \cdots + \beta^{n-3+B_r}\right)^{(n-3)} \\
&= \frac{1}{(n-3)!} \cdot \left(\sum_{i=0}^{n-3+B_r} \beta^i\right)^{(n-3)} \\
&= \frac{1}{(n-3)!} \left(\frac{1-\beta^{n-2+B_r}}{1-\beta}\right)^{(n-3)}, \qquad \text{(A.9)}
\end{aligned}$$

where $f(\beta)^{(k)}$ denotes the k-th order derivative of $f(\beta)$. Since

$$\lim_{B_r \to \infty} 1 - \beta^{n-2+B_r} = 1,$$

we have

$$\begin{aligned}
\lim_{B_r \to \infty} \sum_{i=0}^{B_r} \mathrm{C}_i \cdot \beta^i &= \frac{1}{(n-3)!} \left(\frac{1}{1-\beta}\right)^{(n-3)} \\
&= \frac{1}{(1-\beta)^{n-2}}, \qquad \text{(A.10)}
\end{aligned}$$

and then

$$\begin{aligned}
\lim_{B_r \to \infty} \mathrm{C}_{B_r} \beta^{B_r} (1-\beta)^{n-2} &\leq \lim_{B_r \to \infty} (B_r + n)^n \beta^{B_r} \\
&\leq \lim_{B_r \to \infty} 2^n {B_r}^n \beta^{B_r}. \qquad \text{(A.11)}
\end{aligned}$$

Since

$$\lim_{x\to\infty} x^n\beta^x = \lim_{x\to\infty} \frac{x^n}{\frac{1}{\beta}^x} = \lim_{x\to\infty} \frac{n!}{(-\ln\beta)^n \cdot \frac{1}{\beta}^x} = 0,$$

substituting it into (4.11) we have

$$\underset{\alpha<0.5, B_r\to\infty}{T_c} = p_{sd} + p_{sr}$$

For the case $\alpha > 0.5$, we have $\beta > 1$ and

$$\begin{aligned} 1 - \frac{\mathrm{C}_{B_r}\cdot\beta^{B_r}}{\sum_{i=0}^{B_r}\mathrm{C}_i\cdot\beta^i} &= \frac{\sum_{i=0}^{B_r-1}\mathrm{C}_i\cdot\beta^i}{1+\beta\sum_{i=0}^{B_r-1}\mathrm{C}_{i+1}\cdot\beta^i} \\ &= \frac{1}{\frac{1}{\sum_{i=0}^{B_r-1}\mathrm{C}_i\cdot\beta^i} + \beta\cdot\frac{\sum_{i=0}^{B_r-1}\mathrm{C}_{i+1}\cdot\beta^i}{\sum_{i=0}^{B_r-1}\mathrm{C}_i\cdot\beta^i}}. \end{aligned}$$

Since

$$\begin{aligned} \lim_{B_r\to\infty} \frac{1}{\sum_{i=0}^{B_r-1}\mathrm{C}_i\cdot\beta^i} &= 0, \\ \lim_{B_r\to\infty} \frac{\sum_{i=0}^{B_r-1}\mathrm{C}_{i+1}\cdot\beta^i}{\sum_{i=0}^{B_r-1}\mathrm{C}_i\cdot\beta^i} &= 1, \end{aligned}$$

then

$$\lim_{\beta>1, B_r\to\infty} 1 - \frac{\mathrm{C}_{B_r}\cdot\beta^{B_r}}{\sum_{i=0}^{B_r}\mathrm{C}_i\cdot\beta^i} = \frac{1}{\beta}. \tag{A.12}$$

Substituting it into (4.11) we have

$$\begin{aligned} \underset{\alpha>0.5, B_r\to\infty}{T_c} &= p_{sd} + p_{sr}\frac{1}{\beta} \\ &= p_{sd} + p_{sr}\frac{\alpha}{1-\alpha} = p_{sd} + p_{rd}. \end{aligned}$$

A.3 Proof of Corollary 4

Considering $\gamma \in (0, 1]$, the first order derivative of $g(\gamma)$ is

$$g'(\gamma) = \frac{1}{h(\gamma)^2} \cdot \{\underbrace{h(\gamma)[h(\gamma) + \mathrm{C}_{B_r}] - (1+\gamma)\mathrm{C}_{B_r} h'(\gamma)}_{(a)}\}.$$

For $\forall n > 3$, when $B_r = 1$, (a) is determined as

$$(a) = \gamma(\gamma + n - 2) - (1+\gamma)(n-2) = (\gamma^2 - 1) - (n-3) \leq 0.$$

When $B_r = k$, we assume that

$$(a) = h_k(h_k + \mathrm{C}_k) - (1+\gamma)\mathrm{C}_k h'_k \leq 0,$$

where h_k and h'_k are the abbreviations of $h(\gamma)$ and $h'(\gamma)$ under $B_r = k$, respectively. When $B_r = k+1$, we have

$$\begin{aligned}
(a) &= h_{k+1}(h_{k+1} + \mathrm{C}_{k+1}) - (1+\gamma)\mathrm{C}_{k+1} h'_{k+1} \\
&= \gamma \cdot (h_k + \mathrm{C}_k) \cdot [\gamma(h_k + \mathrm{C}_k) + \mathrm{C}_{k+1}] \\
&- (1+\gamma) \cdot \mathrm{C}_{k+1} \cdot [h_k + \gamma h'_k + \mathrm{C}_k] \\
&= \underbrace{\gamma^2 h_k(h_k + \mathrm{C}_k)}_{(b_1)} + \gamma^2 \mathrm{C}_k(h_k + \mathrm{C}_k) + \underbrace{\gamma \mathrm{C}_{k+1} h_k}_{(c_1)} \\
&+ \underbrace{\gamma \mathrm{C}_k \mathrm{C}_{k+1}}_{(d_1)} - \underbrace{(1+\gamma)\mathrm{C}_{k+1} h_k}_{(c_2)} - \underbrace{\gamma(1+\gamma)\mathrm{C}_{k+1} h'_k}_{(b_2)} \\
&- \underbrace{(1+\gamma)\mathrm{C}_k \mathrm{C}_{k+1}}_{(d_2)}.
\end{aligned}$$

Since

$$\begin{aligned}(b_1) - (b_2) &= \gamma[\gamma h_k(h_k + \mathrm{C}_k) - (1+\gamma)\mathrm{C}_{k+1}h'_k] \\ &< \gamma[h_k(h_k + \mathrm{C}_k) - (1+\gamma)\mathrm{C}_k h'_k] \leq 0,\end{aligned}$$

combining (c_1),(c_2) and (d_1),(d_2) we have

$$\begin{aligned}(a) &< \gamma^2\mathrm{C}_k(h_k + \mathrm{C}_k) - \mathrm{C}_{k+1}h_k - \mathrm{C}_k\mathrm{C}_{k+1} \\ &= (h_k + \mathrm{C}_k)(\gamma^2\mathrm{C}_k - \mathrm{C}_{k+1}) < 0.\end{aligned}$$

According to the above mathematical induction, we can conclude that $g'(\gamma) < 0$ for $\gamma \in (0, 1)$ and $g'(1) \leq 0$. Thus, $g(\gamma)$ monotonically decreases when $\gamma \in (0, 1]$, so we know that $\gamma^* > 1$ and $\alpha^* = \frac{1}{1+\gamma^*} < 0.5$.

For the limiting case $B_r \to \infty$, from (4.16), (4.22) and (4.23) we can easily see that $\alpha^*|_{B_r\to\infty} = 0.5$ and $T_c^*|_{B_r\to\infty} = \frac{p_0+p_1}{2d}$.

APPENDIX B

Proofs of Chapter V

B.1 Proof of Corollary 6

We let $F(\rho_s)$, $G(\rho_s)$ denote the sums of infinite series $\sum_{i\geq 0} \mathrm{C}_i \rho_s^i$ and $\sum_{i\geq 0} i\mathrm{C}_i \rho_s^i$, respectively. Notice that $F(\rho_s)$ is the Taylor series expansion of $(1-\rho_s)^{2-n}$, then we have

$$F(\rho_s) = \frac{1}{(1-\rho_s)^{n-2}}, \tag{B.1}$$

$$G(\rho_s) = \rho_s \cdot F'(\rho_s) = (n-2)\frac{\rho_s}{(1-\rho_s)^{n-1}}. \tag{B.2}$$

Further we have

$$\lim_{B_r\to\infty} L_r^* = \frac{G(\rho_s)}{F(\rho_s)} = (n-2)\frac{\rho_s}{1-\rho_s}, \tag{B.3}$$

and

$$\lim_{B_r\to\infty} p_o = \lim_{B_r\to\infty} \mathrm{C}_{B_r}\cdot\rho_s^{B_r}\cdot(1-\rho_s)^{n-2} \tag{B.4}$$

$$\leq \lim_{B_r\to\infty}(B_r+n)^n\rho_s^{B_r} \leq \lim_{B_r\to\infty} 2^n B_r^n\rho_s^{B_r} \tag{B.5}$$

$$= \lim_{B_r\to\infty}\frac{n!\rho_s^{B_r}}{(-\ln\rho_s)^n} = 0, \tag{B.6}$$

where (B.6) is obtained by utilizing the L'Hôpital's rule recursively.

Substituting (B.3) and (B.6) into Theorem V.1, we can obtain (5.17) and (5.18) directly.

B.2 Proofs of Lemma 2 and Lemma 3

For a cell-partitioned MANET with LS-MAC, the event that node S conducts a S-D (resp. S-R or R-D) transmission in a time slot can be divided into the following sub-events: (1) D is (resp. is not) in the same cell with S; (2) other k out of $n-2$ nodes are in the same cell with S, while the remaining $n-2-k$ nodes are not in this cell; (3) S contends for the wireless channel access successfully. Thus we have

$$\begin{aligned}
p_{sd} &= \sum_{k=0}^{n-2}\binom{n-2}{k}(\frac{1}{m^2})^{k+1}(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{1}{k+2}\\
&= \sum_{k=0}^{n-2}\binom{n-1}{k+1}(\frac{1}{m^2})^{k+1}(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{1}{k+2}\\
&- \sum_{k=0}^{n-3}\binom{n-2}{k+1}(\frac{1}{m^2})^{k+1}(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{1}{k+2}\\
&= \frac{m^2}{n}\left\{1-(1-\frac{1}{m^2})^n\right\} - (1-\frac{1}{m^2})^{n-1}\\
&- \frac{m^2-1}{n-1}\left\{1-(1-\frac{1}{m^2})^{n-1}\right\} + (1-\frac{1}{m^2})^{n-1}\\
&= \frac{m^2}{n} - \frac{m^2-1}{n-1} + (\frac{m^2-1}{n-1} - \frac{m^2-1}{n})(1-\frac{1}{m^2})^{n-1},
\end{aligned}$$

and

$$p_{sr} = p_{rd} = \frac{1}{2}\sum_{k=1}^{n-2}\binom{n-2}{k}(\frac{1}{m^2})^k(1-\frac{1}{m^2})^{n-1-k}\cdot\frac{1}{k+1}$$
$$= \frac{1}{2}\left\{\frac{m^2-1}{n-1} - \frac{m^2}{n-1}(1-\frac{1}{m^2})^n - (1-\frac{1}{m^2})^{n-1}\right\}$$

For a cell-partitioned MANET with EC-MAC, by applying the similar approach and algebraic operations we have

$$p_{sd} = \frac{1}{\varepsilon^2}\left\{\sum_{k=0}^{n-2}\binom{n-2}{k}(\frac{1}{m^2})^{k+1}(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{1}{k+2}\right.$$
$$\left.+\sum_{k=0}^{n-2}\binom{n-2}{k}(\frac{1}{m^2})^{k+1}(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{4v^2-4v}{k+1}\right\}$$
$$= \frac{1}{\varepsilon^2}\left\{\frac{\Gamma-\frac{m^2}{n}}{n-1} + \frac{m^2-1-(\Gamma-1)n}{n(n-1)}(1-\frac{1}{m^2})^{n-1}\right\},$$

and

$$p_{sr} = p_{rd} = \frac{1}{2\varepsilon^2}\frac{m^2-\Gamma}{m^2}\cdot\left\{\sum_{k=1}^{n-2}\binom{n-2}{k}(\frac{1}{m^2})^k(1-\frac{1}{m^2})^{n-2-k}\cdot\frac{1}{k+1}\right.$$
$$\left.+\sum_{k=1}^{n-2}\binom{n-2}{k}(\frac{\Gamma-1}{m^2})^k(\frac{m^2-\Gamma}{m^2})^{n-2-k}\right\}$$
$$= \frac{1}{2\varepsilon^2}\left\{\frac{m^2-\Gamma}{n-1}(1-(1-\frac{1}{m^2})^{n-1}) - (1-\frac{\Gamma}{m^2})^{n-1}\right\}.$$

APPENDIX C

Proofs of Chapter VI

C.1 Proof of Corollary 7

For the network scenario with feedback, node S cannot execute a S-R transmission when the relay buffer of its intended receiver is full (with overflowing probability $\pi_r(B_r)$), thus the service rate μ_s of source buffer of node S is given by

$$\mu_s = p_{sd} + p_{sr} \cdot (1 - \pi_r(B_r)).$$

Based on the similar analysis as that in Section 6.2.1, the OSD $\mathbf{\Pi}_s$ of source buffer here can also be determined by expression (6.4), and the one-step transition probabilities of the birth-death chain of relay buffer can be determined as

$$\begin{aligned} p_{i,i+1} &= \lambda_r, \\ p_{i,i-1} &= p_{rd} \cdot \frac{i}{n-3+i}, \end{aligned}$$

where λ_r denotes the packet arrival rate of the relay buffer when the relay buffer is

not full. Regarding the evaluation of λ_r, we have

$$\lambda_r \cdot (1 - \pi_r(B_r)) + 0 \cdot \pi_r(B_r) = (n-2)\lambda_s^{\text{-}} \cdot \frac{p_{sr}(1 - \pi_r(B_r))}{\mu_s(n-2)}, \tag{C.1}$$

$$\Rightarrow \lambda_r = \lambda_s^{\text{-}} \frac{p_{sr}}{\mu_s} = p_{sr} \cdot (1 - \pi_s(0)), \tag{C.2}$$

where $\lambda_s^{\text{-}}$ denotes the packet departure rate of a source buffer, and (C.2) follows from (6.10). Notice that the transition probabilities here are the same as that under the scenario without feedback, thus the OSD $\mathbf{\Pi}_r$ of the relay buffer here can also be determined by expression (6.12).

C.2 Proof of Corollary 8

From expressions (6.1) and (6.13), we can see that for a given packet generating rate λ, the service rate μ_s of the source buffer under the scenario with feedback is smaller than that under the scenario without feedback. From (6.4) we have

$$\begin{aligned}
\frac{\partial \pi_s(0)}{\partial \mu_s} &= \frac{\mu_s - \lambda\tau^{B_s} - \left(1 - \lambda B_s \tau^{B_s - 1} \frac{\partial \tau}{\partial \mu_s}\right)(\mu_s - \lambda)}{(\mu_s - \lambda\tau^{B_s})^2} \\
&= \frac{\lambda - \lambda\tau^{B_s} - B_s \frac{\lambda(\mu_s - \lambda)}{\mu_s(1-\mu_s)} \tau^{B_s}}{(\mu_s - \lambda\tau^{B_s})^2} \\
&= \frac{\lambda(\mu_s - \lambda)^2}{(\mu_s - \lambda\tau^{B_s})^2 \cdot \mu_s^2 \cdot (1 - \lambda)} \cdot \sum_{i=1}^{B_s - 1} \left(1 + \frac{i}{1 - \mu_s}\right) \tau^i > 0,
\end{aligned} \tag{C.3}$$

which indicates that $\pi_s(0)$ under the scenario with feedback is smaller than that under the scenario without feedback.

We let $r = \frac{1}{1 - \pi_s(0)}$ and substitute r into (6.16), then T can be expressed as

$$T = p_{sd} \cdot \frac{1}{r} + p_{sr} \cdot \frac{1}{g(r)}, \tag{C.4}$$

where $g(r) = r \cdot \left(1 + \frac{\mathrm{C}_{B_r}}{h(r)}\right)$ and $h(r) = \sum_{i=0}^{B_r-1} \mathrm{C}_i r^{B_r-i}$. Regarding the derivative of $g(r)$ we have

$$g'(r) = \frac{1}{h(r)^2} \underbrace{\{h(r)(h(r) + \mathrm{C}_{B_r}) - r\mathrm{C}_{B_r} h'(r)\}}_{(a)}, \tag{C.5}$$

where

$$\begin{aligned}
(a) &= \sum_{i=0}^{B_r-1} \mathrm{C}_i r^{B_r-i} \cdot \sum_{i=0}^{B_r-1} \mathrm{C}_i r^{B_r-i} - \mathrm{C}_{B_r} \sum_{i=0}^{B_r-1} (B_r - i)\mathrm{C}_i r^{B_r-i} \\
&= \sum_{i=1}^{B_r} \mathrm{C}_{B_r-i} r^i \cdot \sum_{i=0}^{B_r} \mathrm{C}_{B_r-i} r^i - \sum_{i=1}^{B_r} i\mathrm{C}_{B_r}\mathrm{C}_{B_r-i} r^i \\
&= \sum_{i=1}^{B_r} \left(\sum_{j=0}^{i-1} \mathrm{C}_{B_r-j} r^j \mathrm{C}_{B_r-i+j} r^{i-j} - i\mathrm{C}_{B_r}\mathrm{C}_{B_r-i} r^i \right) \\
&+ \sum_{i=B_r+1}^{2B_r} \sum_{j=i-B_r}^{B_r} \mathrm{C}_{B-j} r^j \mathrm{C}_{B-i+j} r^{i-j} \\
&> \sum_{i=1}^{B_r} \left(\sum_{j=0}^{i-1} \mathrm{C}_{B_r-j}\mathrm{C}_{B_r-i+j} - i\mathrm{C}_{B_r}\mathrm{C}_{B_r-i} \right) r^i > 0,
\end{aligned} \tag{C.6}$$

here (C.6) follows because that $\mathrm{C}_{B_r-j}\mathrm{C}_{B_r-i+j} > \mathrm{C}_{B_r}\mathrm{C}_{B_r-i}$ for $0 < j < i$.

We can see from (C.3) that $\pi_s(0)$ increases as μ_s increases, and from (C.4)$-$(C.6) that T increases as $\pi_s(0)$ decreases. Thus, we can conclude that T under the scenario with feedback is larger than that under the scenario without feedback, which indicates that adopting the feedback mechanism improves the throughput performance.

C.3 Proof of Lemma 4

For the scenario without feedback, we know from (6.4) that

$$\begin{aligned}\frac{\partial \pi_s(0)}{\partial \lambda} &= \frac{-\mu_s + \lambda\tau^{B_s} + \left(\tau^{B_s} + \lambda B_s \tau^{B_s-1}\frac{\partial \tau}{\partial \lambda}\right)(\mu_s - \lambda)}{(\mu_s - \lambda\tau^{B_s})^2} \\ &= \frac{-\mu_s + \mu_s\tau^{B_s} + B_s\frac{\mu_s-\lambda}{1-\lambda}\tau^{B_s}}{(\mu_s - \lambda\tau^{B_s})^2} \\ &= \frac{-(\lambda-\mu_s)^2}{(\mu_s - \lambda\tau^{B_s})^2 \cdot (1-\lambda)^2 \cdot \mu_s} \cdot \sum_{i=1}^{B_s} i\tau^{i-1} < 0. \end{aligned} \tag{C.7}$$

Thus, as λ increases, $\pi_s(0)$ decreases which leads to an increase in T (refer to the analysis in Appendix C.2).

For the scenario with feedback, as λ increases, the MANET tends to be more congested with a larger $\pi_r(Br)$. Thus, we know from (6.13) that the corresponding μ_s decreases, and then from (C.3) that $\pi_s(0)$ decreases, leading to an increase in T.

C.4 Proof of Corollary 9

From an intuitive point of view, a larger buffer implies that more packets can be stored and packet loss can be reduced, thus a higher throughput can be achieved. More formally, from (6.4) we have

$$\pi_s(0)|_{B_s=K+1} - \pi_s(0)|_{B_s=K} = \frac{\lambda\tau^K(\mu_s-\lambda)(\tau-1)}{(\mu_s - \lambda\tau^{K+1})(\mu_s - \lambda\tau^K)} < 0, \tag{C.8}$$

where (C.8) follows since $\tau > 1$ when $\lambda > \mu_s$ and $\tau < 1$ when $\lambda < \mu_s$. Then we can conclude that as B_s increases, $\pi_s(0)$ decreases, leading to an increase in T.

Let $r = \frac{1}{1-\pi_s(0)}$ and substituting r into (6.12), then we have

$$\begin{aligned}
&\pi_r(B_r)|_{B_r=K+1} - \pi_r(B_r)|_{B_r=K} \\
=&\frac{\mathrm{C}_{K+1}r^{-K-1}}{\sum_{i=0}^{K+1}\mathrm{C}_i r^{-i}} - \frac{\mathrm{C}_K r^{-K}}{\sum_{i=0}^{K}\mathrm{C}_i r^{-i}} \\
=&\frac{\underbrace{\mathrm{C}_{K+1}r^{-K-1}\sum_{i=0}^{K}\mathrm{C}_i r^{-i} - \mathrm{C}_K r^{-K}\sum_{i=0}^{K+1}\mathrm{C}_i r^{-i}}_{(b)}}{\sum_{i=0}^{K+1}\mathrm{C}_i r^{-i}\cdot\sum_{i=0}^{K}\mathrm{C}_i r^{-i}},
\end{aligned}$$

where

$$\begin{aligned}
(b) =& \mathrm{C}_{K+1}r^{-K-1}\sum_{i=0}^{K}\mathrm{C}_i r^{-i} - \mathrm{C}_K r^{-K}\sum_{i=1}^{K+1}\mathrm{C}_i r^{-i} - \mathrm{C}_K r^{-K} \\
<& \sum_{i=0}^{K}\left(\mathrm{C}_{K+1}\mathrm{C}_i - \mathrm{C}_K\mathrm{C}_{i+1}\right) r^{-k-i-1} < 0.
\end{aligned}$$

Then we can conclude that as B_r increases, $\pi_r(B_r)$ decreases, leading to an increase in T (refer to expression (6.16)).

Regarding the infinite source buffer (i.e., $B_s \to \infty$), $\tau \geq 1$ when $\lambda \geq \mu_s$, and we have

$$\begin{aligned}
&\lim_{B_s\to\infty}\pi_s(0) = \lim_{B_s\to\infty}\frac{\mu_s - \lambda}{\mu_s - \lambda\tau^{B_s}} = 0, \\
&\lim_{B_s\to\infty} T = p_{sd} + p_{sr}\cdot\frac{\sum_{k=1}^{B_r}\mathrm{C}_{k-1}}{\sum_{k=0}^{B_r}\mathrm{C}_k} = p_{sd} + p_{sr}\frac{B_r}{n-2+B_r} = T_c.
\end{aligned}$$

According to the Queuing theory [47], for a Bernoulli/Bernoulli queue (i.e., the buffer size is infinite), its queue length tends to infinity when the corresponding arrival rate is equal to or larger than the service rate. Thus, we have $L_s^* \to \infty$, which leads that

$\mathbb{E}\{D_{sq}\} \to \infty$ and $\mathbb{E}\{D\} \to \infty$.

When $\lambda < \mu_s$, $\tau < 1$, and we have

$$\lim_{B_s \to \infty} \pi_s(0) = \lim_{B_s \to \infty} \frac{\mu_s - \lambda}{\mu_s - \lambda \tau^{B_s}} = 1 - \frac{\lambda}{\mu_s},$$

$$\lim_{B_s \to \infty} T = p_{sd} \cdot \frac{\lambda}{\mu_s} + p_{sr} \cdot \frac{\sum_{k=0}^{B_r - 1} \mathrm{C}_k (\frac{\lambda}{\mu_s})^{k+1}}{\sum_{k=0}^{B_r} \mathrm{C}_k (\frac{\lambda}{\mu_s})^k}.$$

Based on the analysis in Theorem VI.1, L_s^* is determined as

$$\lim_{B_s \to \infty} L_s^* = \lim_{B_s \to \infty} \frac{1 - \tau}{1 - \tau^{B_s}} \sum_{i=0}^{B_s - 1} i \tau^i = \frac{\tau}{1 - \tau}. \tag{C.9}$$

Substituting (C.9) into (6.17) we obtain (6.31b).

Regarding the infinite relay buffer (i.e., $B_r \to \infty$), from (6.12) and (6.19) we have

$$\begin{aligned} \lim_{B_r \to \infty} \pi_r(B_r) &= \lim_{B_r \to \infty} \mathrm{C}_{B_r} (1 - \pi_s(0))^{B_r} \cdot \pi_s(0)^{n-2} && \text{(C.10)} \\ &\leq \lim_{B_r \to \infty} (B_r + n)^n (1 - \pi_s(0))^{B_r} \\ &\leq \lim_{B_r \to \infty} 2^n B_r^n (1 - \pi_s(0))^{B_r} \\ &= \lim_{B_r \to \infty} \frac{2^n n! (1 - \pi_s(0))^{B_r}}{(\ln \frac{1}{1 - \pi_s(0)})^n} = 0, && \text{(C.11)} \end{aligned}$$

$$\lim_{B_r\to\infty} L_r^* = \frac{\sum\limits_{k\geq 0} k\mathrm{C}_k(1-\pi_s(0))^k}{\sum\limits_{k\geq 0} \mathrm{C}_k(1-\pi_s(0))^k}$$

$$= \frac{-(1-\pi_s(0))\cdot\left(\sum\limits_{k\geq 0} \mathrm{C}_k(1-\pi_s(0))^k\right)'}{\sum\limits_{k\geq 0} \mathrm{C}_k(1-\pi_s(0))^k}$$

$$= -(1-\pi_s(0))\cdot\left(\pi_s(0)^{2-n}\right)'\cdot\pi_s(0)^{n-2} \tag{C.12}$$

$$= \frac{(n-2)(1-\pi_s(0))}{\pi_s(0)}, \tag{C.13}$$

where (C.10) and (C.12) follow since $\sum\limits_{k\geq 0} \mathrm{C}_k(1-\pi_s(0))^k$ is just the Taylor-series expansion [67] of $\pi_s(0)^{2-n}$, and (C.11) follows from the L'Hôpital's rule [67]. Substituting (C.11) into (6.16) we obtain (6.30b), and substituting (C.11) and (C.13) into (6.17) we obtain (6.31c).

Regarding the MANET without buffer constraint (i.e., $B_s \to \infty$ and $B_r \to \infty$), we can directly obtain (6.30c) and (6.31d) by combining the corresponding results of the infinite source buffer scenario and the infinite relay buffer scenario.

BIBLIOGRAPHY

BIBLIOGRAPHY

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

[2] A. Goldsmith, *Wireless Communications.* Cambridge University Press, 2005.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[4] C. E. Perkins, *Ad Hoc Networking.* Addison-Wesley Professional, 2000.

[5] R. Ramanathan and J. Redi, "A brief overview of ad hoc networks: challenges and directions," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 20–22, 2002.

[6] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, 2014.

[7] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 12, pp. 94–101, 2008.

[8] A. Goldsmith, M. Effros, R. Koetter, M. Medard, and L. Zheng, "Beyond Shannon: The quest for fundamental performance limits of wireless ad hoc networks," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 195–205, 2011.

[9] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, 2002.

[10] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay tradeoff in wireless networks," in *Proc. IEEE INFOCOM*, 2004.

[11] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad-hoc mobile networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1917–1936, 2005.

[12] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks-part I: The fluid model," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2568–2592, 2006.

[13] G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 981–992, 2007.

[14] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with brownian mobility," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 2777–2784, 2006.

[15] J. Mammen and D. Shah, "Throughput and delay in random wireless networks with restricted mobility," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1108–1116, 2007.

[16] M. Franceschetti, O. Dousse, D. Tse, and P. Thira, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.

[17] S. Shakkottai, X. Liu, and R. Srikant, "The multicast capacity of large multihop wireless networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1691–1700, 2010.

[18] S. Vasudevan, D. Goeckel, and D. Towsley, "Security-capacity trade-off in large wireless networks using keyless secrecy," in *Proc. ACM MobiHoc*, 2010, pp. 21–30.

[19] D. Ciullo, V. Martina, M. Garetto, and E. Leonardi, "Impact of correlated mobility on delay-throughput performance in mobile ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1745–1758, 2011.

[20] Q. Peng, X. Wang, and H. Tang, "Heterogeneity increases multicast capacity in clustered network," in *Proc. IEEE INFOCOM*, 2011, pp. 703–711.

[21] R. Urgaonkar and M. J. Neely, "Network capacity region and minimum energy function for a delay-tolerant mobile ad hoc network," *IEEE/ACM Trans. Netw.*, vol. 19, no. 4, pp. 1137–1150, 2011.

[22] X. Wang, W. Huang, S. Wang, J. Zhang, and C. Hu, "Delay and capacity tradeoff analysis for motioncast," *IEEE/ACM Trans. Netw.*, vol. 19, no. 5, pp. 1354–1367, 2011.

[23] Y. Wang, X. Chu, X. Wang, and Y. Cheng, "Optimal multicast capacity and delay tradeoffs in MANETs: A global perspective," in *Proc. IEEE INFOCOM*, 2011, pp. 640–648.

[24] C. Zhang, Y. Song, Y. Fang, and Y. Zhang, "On the price of security in large-scale wireless ad hoc networks," *IEEE/ACM Trans. Netw*, vol. 19, no. 2, pp. 319–332, 2011.

[25] D. M. Shila and Y. Cheng, "Ad hoc wireless networks meet the infrastructure: Mobility, capacity and delay," in *Proc. IEEE INFOCOM*, 2012, pp. 3031–3035.

[26] X. Chen, W. Huang, X. Wang, and X. Lin, "Multicast capacity in mobile wireless ad hoc network with infrastructure support," in *Proc. IEEE INFOCOM*, 2012, pp. 271–279.

[27] W. Huang and X. Wang, "Capacity scaling of general cognitive networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1501–1513, 2012.

[28] O. O. Koyluoglu, C. E. Koksal, and H. E. Gamal, "On secrecy capacity scaling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3000–3015, 2012.

[29] Y. Qin, X. Tian, W. Wu, and X. Wang, "Mobility weakens the distinction between multicast and unicast," *IEEE/ACM Trans. Netw.*, 2015, to appear.

[30] S. Zhou and L. Ying, "On delay constrained multicast capacity of large-scale mobile ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5643–5655, 2015.

[31] N. Lu and X. Shen, "Scaling laws for throughput capacity and delay in wireless networks - A survey," *IEEE Commun. Surv. Tutorials*, 2013.

[32] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and S. Clifford, *Introduction to Algorithms*. MIT press, 2001.

[33] S. Toumpis and A. J. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 736–748, 2003.

[34] G. Mergen and L. Tong, "Stability and capacity of regular wireless networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1938–1953, 2005.

[35] J. Gao, J. Liu, X. Jiang, O. Takahashi, and N. Shiratori, "Throughput capacity of MANETs with group-based scheduling and general transmission range," *IEICE Trans. Commun.*, vol. 96, no. 7, pp. 1791–1802, 2013.

[36] Y. Chen, Y. Shen, X. Jiang, and J. Li, "Throughput capacity of ALOHA MANETs," in *Proc. IEEE ICCC*. IEEE, 2013, pp. 71–75.

[37] Y. Chen, Y. Shen, J. Zhu, and X. Jiang, "Capacity and delay-throughput tradeoff in ICMNs with Poisson contact process," *Wireless Networks*, vol. 21, no. 8, pp. 2453–2466, 2015.

[38] Y. Chen, J. Liu, X. Jiang, and O. Takahashi, "Throughput analysis in mobile ad hoc networks with directional antennas," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1122–1135, 2013.

[39] A. Jindal and K. Psounis, "Contention-aware performance analysis of mobility-assisted routing," *IEEE Trans. Mobile Comput.*, vol. 8, no. 2, pp. 145–161, 2009.

[40] J. Gao and X. Jiang, "Delay modeling for broadcast-based two-hop relay manets," in *Proc. IEEE WiOpt*, 2013, pp. 357–363.

[41] J. D. Herdtner and E. K. P. Chong, "Throughput-storage tradeoff in ad hoc networks," in *Proc. IEEE INFOCOM*, 2005, pp. 2536–2542.

[42] R. Subramanian and F. Fekri, "Analysis of multiple-unicast throughput in finite-buffer delay-tolerant networks," in *Proc. IEEE ISIT*, 2009, pp. 1634–1638.

[43] J. Gao, Y. Shen, X. Jiang, and J. Li, "Source delay in mobile ad hoc networks," *Ad Hoc Networks*, vol. 24, pp. 109–120, 2015.

[44] Y. Fang, Y. Zhou, X. Jiang, and Y. Zhang, "Practical performance of MANETs under limited buffer and packet lifetime," *IEEE Systems Journal*, 2016 (in press).

[45] P. Li, Y. Fang, J. Li, and X. Huang, "Smooth trade-offs between throughput and delay in mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 3, pp. 427–438, 2012.

[46] P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of random direction models," in *Proc. IEEE INFOCOM*, 2005, pp. 1897–1907.

[47] H. Daduna, *Queueing Networks with Discrete Time Scale: Explicit Expressions for the Steady State Behavior of Discrete Time Stochastic Networks*. Springer Verlag, 2001.

[48] H. Wei and R. D. Gitlin, "Two-hop-relay architecture for next-generation WWAN/WLAN integration," *IEEE Wireless Commun.*, vol. 11, no. 2, pp. 24–30, 2004.

[49] L. B. Le, E. Modiano, and N. B. Shroff, "Optimal control of wireless networks with finite buffers," *IEEE/ACM Trans. Netw.*, vol. 20, no. 4, pp. 1316–1329, 2012.

[50] T. G. Robertazzi, *Computer Networks and Systems: Queueing Theory and Performance Evaluation*. Springer Science & Business Media, 2012.

[51] W. Feller, *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, 2008.

[52] J. Muller, *Elementary Functions*. Springer, 1997.

[53] H. J. Keisler, *Elementary Calculus: An Infinitesimal Approach*. Courier Dover Publications, 2012.

[54] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1041–1049, 2004.

[55] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[56] J. Liu and Y. Xu, "C++ simulator for throughput performance in buffer-limited manets," 2015, available: http:/jliuyxu.blogspot.jp/.

[57] S. McCanne and S. Floyd, "The Network Simulator - ns-2," [Online]. Available: http://www.isi.edu/nsnam/ns/, 1997.

[58] A. Granas and J. Dugundji, *Fixed Point Theory*. New York: Springer Verlag, 2003.

[59] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. D. Van Nostrand Princeton, NJ, 1960.

[60] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, 2000.

[61] G. Bianchi and I. Tinnirello, "Remarks on IEEE 802.11 DCF performance analysis," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 765–767, 2005.

[62] J. Liu and Y. Xu, "C++ simulator for delay performance in buffer-limited manets," 2015, [Online]. Available: http:/jliuyxu.blogspot.jp/.

[63] J. Liu, M. Sheng, Y. Xu, J. Li, and X. Jiang, "End-to-end delay modeling in buffer-limited MANETs: A general theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 498–511, 2016.

[64] ——, "On throughput capacity for a class of buffer-limited MANETs," *Ad Hoc Networks*, vol. 37, pp. 354–367, 2016.

[65] J. Liu and Y. Xu, "C++ simulator: performance modeling for MANETs under general limited buffer constraint," [Online]. Available: https://www.researchgate.net/profile/Jia_Liu100, 2015. DOI: 10.13140/RG.2.1.1266.8248.

[66] H. Stark and J. W. Woods, *Probability and Random Processes with Applications to Signal Processing*. Prentice Hall, 2002.

[67] J. Stewart, *Calculus*. Cengage Learning, 2011.

Publications

Jounal Articles

[1] Jia Liu, Min Sheng, Yang Xu, Jiandong Li and Xiaohong Jiang. End-to-end delay modeling in buffer-limited MANETs: A general theoretical framework. *IEEE Transactions on Wireless Communications*, 15(1): 498–511, January 2016.

[2] Jia Liu, Min Sheng, Yang Xu, Jiandong Li and Xiaohong Jiang. On throughput capacity for a class of buffer-limited MANETs. *Ad Hoc Networks*, 37(2): 354-367, February 2016.

[3] Yang Xu, Jia Liu, Yulong Shen, Xiangning Li and Xiaohong Jiang. On throughput capacity of large-scale ad hoc networks with realistic buffer constraint. *Wireless Networks*, 2016. In press, DOI:10.1007/s11276-015-1146-2.

[4] Yongzhi Wang, Yulong Shen, Xiaopeng Jiao, Tao Zhang, Xu Si, Ahmed Salema, and Jia Liu. Exploiting content delivery networks for covert channel communications. *Computer Communications*, 2016. In press, DOI:10.1016/j.comcom.2016.07.011.

Conference Papers

[5] Jia Liu, Min Sheng, Yang Xu, Hongguang Sun, Xijun Wang and Xiaohong Jiang. Throughput capacity of two-hop relay MANETs under finite buffers. IEEE 25th Annual International Symposium on Indoor, and Mobile Radio Communication (PIMRC), Washington DC, USA, September 2014.

[6] Jia Liu, Yang Xu, and Xiaohong Jiang. End-to-end delay in two hop relay MANETs with limited buffer. Second International Symposium on Computing and Networking (CANDAR), Shizuoka, Japan, December 2014.

[7] Yang Xu, Jia Liu, Yulong Shen, Xiaohong Jiang and Tarik Taleb. Security/QoS-aware route selection in multi-hop wireless ad hoc networks. IEEE International Conference on Communications, Kuala Lumpur, Malaysia, May 2016.