Opportunistic Routing in Multihop Wireless Networks: Capacity, Energy Efficiency, and Security

by
Kai Zeng

A Dissertation
Submitted to the Faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
In partial fulfillment of the requirements for the
Degree of Doctor of Philosophy
in
Electrical and Computer Engineering

July, 2008

Approved:

Prof. Wenjing Lou ECE Department Dissertation Advisor	Prof. Kaveh Pahlavan ECE Department Dissertation Committee
Prof. Donald R. Brown ECE Department Dissertation Committee	Prof. Robert E. Kinicki CS Department Dissertation Committee

Prof. Fred J. Looft
ECE Department Head

Abstract

Opportunistic routing (OR) takes advantages of the spatial diversity and broadcast nature of wireless networks to combat the time-varying links by involving multiple neighboring nodes (forwarding candidates) for each packet relay. This dissertation studies the properties, energy efficiency, capacity, throughput, protocol design and security issues about OR in multihop wireless networks.

Firstly, we study geographic opportunistic routing (GOR), a variant of OR which makes use of nodes' location information. We identify and prove three important properties of GOR. The first one is on prioritizing the forwarding candidates according to their geographic advancements to the destination. The second one is on choosing the forwarding candidates based on their advancements and link qualities in order to maximize the expected packet advancement (EPA) with different number of forwarding candidates. The third one is on the concavity of the maximum EPA in respect to the number of forwarding candidates. We further propose a local metric, EPA per unit energy consumption, to tradeoff the routing performance and energy efficiency for GOR. Leveraging the proved properties of GOR, we propose two efficient algorithms to select and prioritize forwarding candidates to maximize the local metric.

Secondly, capacity is a fundamental issue in multihop wireless networks. We propose a framework to compute the end-to-end throughput bound or capacity of OR in single/multirate systems given OR strategies (candidate selection and prioritization). Taking into account wireless interference and unique properties of OR, we propose a new method of constructing transmission conflict graphs, and we introduce the concept of concurrent transmission sets to allow the proper formulation of the maximum end-to-end throughput problem as a maximum-flow linear programming problem subject to the transmission conflict constraints. We also propose two OR metrics: *expected medium time* (EMT) and *expected advancement rate* (EAR), and the corresponding distributed and local rate and candidate set selection schemes, the Least Medium Time OR (LMTOR) and the Multirate Geographic OR (MGOR). We further extend our framework to compute the capacity of OR in multi-radio multi-channel systems with dynamic OR strategies. We study the necessary

and sufficient conditions for the schedulability of a traffic demand vector associated with a transmitter to its forwarding candidates in a concurrent transmission set. We further propose an LP approach and a heuristic algorithm to obtain an opportunistic forwarding strategy scheduling that satisfies a traffic demand vector. Our methodology can be used to calculate the end-to-end throughput bound of OR in multi-radio/channel/rate multihop wireless networks, as well as to study the OR behaviors (such as candidate selection and prioritization) under different network configurations.

Thirdly, protocol design of OR in a contention-based medium access environment is an important and challenging issue. In order to avoid duplication, we should ensure only the "best" receiver of each packet to forward it in an efficient way. We investigate the existing candidate coordination schemes and propose a "fast slotted acknowledgment" (FSA) to further improve the performance of OR by using a single ACK to coordinate the forwarding candidates with the help of the channel sensing technique. Furthermore, we study the throughput of GOR in multi-rate and single-rate systems. We introduce a framework to analyze the one-hop throughput of GOR, and provide a deeper insight on the trade-off between the benefit (packet advancement, bandwidth, and transmission reliability) and cost (medium time delay) associated with the node collaboration. We propose a local metric named *expected one-hop throughput* (EOT) to balance the benefit and cost.

Finally, packet reception ratio (PRR) has been widely used as an indicator of the link quality in multihop wireless networks. Many routing protocols including OR in wireless networks depend on the PRR information to make routing decision. Providing accurate link quality measurement (LQM) is essential to ensure the right operation of these routing protocols. However, the existing LQM mechanisms are subject to malicious attacks, thus can not guarantee to provide correct link quality information. We analyze the security vulnerabilities in the existing link quality measurement (LQM) mechanisms and propose an efficient broadcast-based secure LQM (SLQM) mechanism, which prevents the malicious attackers from reporting a higher PRR than the actual one. We analyze the security strength and the cost of the proposed mechanism.

To my lovely wife, Jie, and my parents

Acknowledgements

I would like to express my deep and sincere gratitude to my advisor, Professor Wenjing Lou, who is most responsible for helping me complete this dissertation as well as the challenging research that lies behind it. Her wide knowledge and her logical way of thinking have been of great value for me. She was always there to meet and talk about my ideas, to proofread and mark up my papers, and to ask me good questions to help me think through my problems. Without her encouragement and constant guidance, I could not have finished this dissertation. She has set an example as a smart, hardworking, persistent and passionate researcher that I can only hope to emulate.

Besides my advisor, I am thankful to the rest of my dissertation committee: Professor Kaveh Pahlavan, Professor Donald R. Brown, and Professor Robert E. Kinicki, who asked me good questions and gave insightful comments on my work.

I want to thank my friend Peng Ni in the Mathematical department for his help on proving the concavity of the EPA of opportunistic routing. We also shared much fun in the basketball court. I am thankful to Prof. Brigitte Servatius in the Mathematical department for spending her time discussing mathematical puzzles with me.

I am grateful to Professor Yuguang Fang of University of Florida for his consistent encouragement for me to do better in my research. I wish to thank Prof. Guoliang Xue for his comments on the proof of the containing property in opportunistic routing.

I want to thank Dr. Hongqiang Zhai and Dr. Jianfeng Wang in the Wireless Communications and Networking Department of Philips Research North America for their advice on my research. I am specially grateful for Dr. Zhai's collaboration when we study the capacity of opportunistic routing.

I wish to thank my previous colleague Dr. Kui Ren and other labmates: Zhenyu Yang, Shucheng Yu, and Ming Li for creating an intellectual and enjoyable atmosphere in the lab. I am especially indebted to Zhenyu Yang, who has put great effort on helping me implement, simulate, and test our research ideas.

I am greatly indebted to my father, Xiangyuan Zeng, and my mother, Guiwen Hou. Thank you, mom and dad, for giving me life in the first place, for educating me to be a

honest and responsible person, and for unconditional support and encouragement to pursue my interests.

Most importantly, I am specially indebted to my wife, Jie Yang, for staying with me abroad and supporting me all the time. Jie left her family and friends, and rejected a job offer in China three years ago to come to America living with me in Worcester. I am grateful to Jie for her love and all the sacrifices she has made for me.

I am grateful for the financial support of the National Science Foundation through grants CNS-0626601, CNS-0716306 and CNS-0746977. I want to also thank Worcester Polytechnic Institute and ECE department for their fellowship and financial assistance. I also wish to thank AirSprite Technologies for their financial support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Multi-hop wireless networks, such as mobile ad hoc networks (MANETs), wireless sensor networks (WSNs), and wireless mesh networks (WMNs), have received increasing attention in the past decade due to their broad applications and the easy deployment at low cost without relying on existing infrastructure [4,5,18,14,26,6,44]. Network protocol design in such networks presents a great challenge mainly due to the following reasons. First, an important feature of wireless networks is the time-varying channel caused by wireless channel propagation effects, mainly multipath fading, which can result in large fluctuations in signal strength and therefore intermittent link behavior. Second, wireless link is a "soft" concept. The property and quality of a link may vary with the transmission power, transmission rate, distance and path loss between two nodes. Third, since the wireless medium is broadcast in nature, the transmission on one link may interfere with the transmissions on other neighboring links. Fourth, wireless embedded devices, such as sensors, are typically battery powered. The lifetime of the battery imposes a limitation on the operation hours of the network. Energy efficiency has been a critical concern in wireless sensor

network protocol design.

Traditional routing protocols [50, 34, 49] for multihop wireless networks have followed the concept of routing in wired networks by abstracting the wireless links as wired links, and find the shortest, least cost, or highest throughput path(s) between a source and destination. Since most routing protocols rely on the consistent and stable behavior of individual links, the intermittent behavior of wireless links can result in poor performance such as low packet delivery ratio and high control overhead. One the other hand, this abstraction ignores the unique broadcast nature and spacial diversity of the wireless medium.

In a wireless network, when a packet is unicast to a specific next-hop node of the sender at the network layer, all the neighboring nodes in the effective communication range of the sender may be able to overhear the packet at the physical layer. It's possible that some of the neighbors may have received the packet correctly while the designated next-hop node did not. Based on this observation, a new routing paradigm, known as **opportunistic routing** (OR) [41, 77, 9, 84, 85, 57, 28, 3, 11] has recently been proposed. OR integrates the network and MAC layers. Instead of picking one node to forward a packet to, the network layer selects a set of candidate nodes to forward a packet to and at the MAC layer one node is selected dynamically as the actual forwarder based on the instantaneous wireless channel condition and node availability at the time of transmission. Opportunistic routing takes advantages of the spatial diversity and broadcast nature of wireless communications and is an efficient mechanism to combat the time-varying links. OR improves the network throughput [9, 28, 68, 70, 69] and energy efficiency [84, 67] compared to traditional routing.

Performance of OR depends on several key issues. The first **key issue** is the selection of forwarding candidates. Although involving all the neighbors with smaller cost to the destination is seemingly the most effective way, the overhead is expected

to grow with the increase of the number of forwarding candidates. In dense networks, this overhead might potentially be even higher than cost incurred due to repeated transmissions [58]. The prioritization of the candidates is the second **key issue** that affects the performance. In general, we want to forward the packet along the "shortest" path. The lower priority forwarding candidates are essentially the backup to the node that is on the "shortest" path. However, due to the opportunistic nature, the "distance" from a certain node to its multihop away destination will no longer be the same as that obtained by traditional shortest path routing. The path cost also depends on the spatial diversity opportunities along the path. How to quantify and incorporate the spatial diversity opportunities in OR has not been well understood. The third **key issue** is candidate coordination in the MAC layer which ensures the multiple receivers of a packet to agree upon a next-hop forwarder in a distributed fashion [41, 77, 84, 85, 80, 19, 60].

Although opportunistic routing has shown its effectiveness in achieving better energy efficiency [84, 85] and higher throughput [9] than traditional routing, there are still many important issues in OR remained unanswered or not well understood. First, none of the existing works provides a thorough understanding of how well the opportunistic routing can perform and how the selection of the forwarding candidate set will affect the routing efficiency. Questions, such as " a) how many and which neighbor nodes should be involved in the local forwarding? "; and " b) What are the selection criteria and how do they affect the relay priority among the forwarding candidates? ", remain unanswered. Second, there is a lack of theoretical analysis on the throughput bounds achievable by OR. Third, one of the current trends in wireless communication is to enable devices to operate using multiple transmission rates. For example, many existing wireless networking standards such as IEEE 802.11a/b/g include this multi-rate capability. The inherent rate-distance trade-off of multi-rate transmissions has shown its impact on the throughput performance of traditional

routing [8, 75, 74]. Generally, low-rate communication covers a long transmission range, while high-rate communication must occur at short range. It is intuitive to expect that this rate-distance trade-off will also affect the throughput of OR. Because different transmission ranges also imply different neighboring node sets, this results in different spacial diversity opportunities. The rate-distance-diversity trade-offs in OR are not well studied. Furthermore, existing OR coordination schemes have some inherent inefficiencies such as high time delay and potential duplicate forwarding, etc. Improperly designed coordination scheme will aggravate these problems and even overwhelm the potential gain provided by OR. It is necessary to design more efficient candidate coordination schemes. Finally, most state-of-the-art OR protocols [9,68] rely on link quality (packet reception ratio) information to select and prioritize forwarding candidates. It is important to accurately measure the link quality in order to make OR operate optimally. However, the existing link quality measurement mechanisms are subject to malicious attacks. Thus they may not be able to provide accurate link quality information for OR.

This dissertation carries out a comprehensive study on the capacity, energy efficiency, throughput, and security issues in OR, and the associated multi-rate, candidate selection, prioritization, and coordination problems. Our goal is to fully understand the principles, the tradeoffs, the gains of the node collaboration and its associated cost to provide insightful analysis and guidance to the design of more efficient routing protocols.

1.2 Related Work

1.2.1 Opportunistic Routing

Opportunistic routing exploits the spacial diversity of the wireless medium by involving a set of forwarding candidates instead of only one in traditional routing. This

improves the reliability and efficiency of packet relay. Some variants of opportunistic routing, such as ExOR [9] and opportunistic any-path forwarding [82, 81], rely on the path cost information or global knowledge of the network to select candidates and prioritize them. In the least cost opportunistic routing (LCOR) [24], it needs to enumerate all the neighboring node combinations to get the least cost OR paths. Another variant of OR is geographic opportunistic routing (GOR) [84, 57, 28] which uses the location information of nodes to define the candidate set and relay priority. In GeRaF [84], the next-hop neighbors of the current forwarding node are divided into sets of priority regions with nodes closer to the destination having higher relay priorities. Similar to [84], in [57], the network layer specifies a set of nodes by defining a forwarding region in space that consists of the candidate nodes and the data link layer selects the first node available from that set to be the next hop node. [28] discussed three suppression strategies of contention-based forwarding to avoid packet duplication in mobile ad hoc networks. However, there is no theoretical work on determining the end-to-end throughput bounds of OR. It is not well understood how the selection and prioritization of the forwarding candidates will affect the routing efficiency.

1.2.2 Geographic Routing

Owing to its scalability, statelessness, and low maintenance overhead, geographical routing is considered as an efficient paradigm for data forwarding in multi-hop wireless ad hoc and sensor networks. Early works [27, 36, 40] on geographic routing exploit the concept of maximum advancements towards the destination to route packets in a greedy manner. However, recent empirical measurements [21, 78] have proved that the unit disk connectivity model, on which these solutions are based, often fails in real settings. More recent works on geographic routing are focused on lossy channel situations. Seada, et al. [56] articulated the distance–hop energy trade-off for

geographic routing. They concluded that packet advancement timing packet reception ratio, the EPA, is an optimal metric for making localized geographic routing decisions in lossy wireless networks with ARQ (Automatic Repeat reQuest) mechanisms, and is also a good metric for Non-ARQ scenarios. Zorzi and Armaroli also independently proposed the same link metric [83]. Lee et al. [42] presented a more general framework called normalized advance (NADV) to normalize various types of link cost such as transmission times, delay and power consumption. Unfortunately, NADV only applies to geographic routing which involves a single forwarding candidate and cannot be directly used for geographic opportunistic routing.

1.2.3 Capacity of Multi-hop Wireless Networks

The theoretical capacity study on multi-hop wireless networks mainly focuses on two directions. One is on the asymptotic bounds of the network capacity [31, 30]. These works study the capacity trend with regard to the size of a wireless network under specific assumptions or scenarios. Another direction on wireless network capacity is to compute the exact performance bounds for a given network. Jain *et al* [32] proposed a framework to calculate the throughput bounds of traditional routing between a pair of nodes by adding wireless interference constraints into the maximum flow formulations. Zhai and Fang [74] studied the path capacity of traditional routing in a multi-rate scenario. Our work falls into this direction. However, distinguished from the previous works, we propose a method to compute the end-to-end throughput bounds of opportunistic routing, which is different from the traditional routing in that we construct the transmitter (associated with multiple forwarding candidates) based conflict graph instead of link conflict graph to capture the local broadcast nature of OR. Our framework can be used as a tool to calculate the end-to-end throughput bound of different OR variants, and is an important theoretical foundation for the performance study of OR. There has been recent work [39, 7, 76] on capacity bound

computation in multi-radio multi-channel networks. However, they are all based on the assumption of using traditional routing at the network layer, where one transmitter can only deliver traffic to one receiver.

1.2.4 Multi-rate Routing

Multirate wireless networks have started attracting research attention recently. In [23], Draves, Padhye and Zill proposed to use the weighted cumulative expected transmission time (WCETT) as a routing metric. In [8], Awerbuch, Holmer and Rubens adopted the medium time metric (MTM). In [75], Zhai and Fang studied the impact of multirate on carrier sensing range and spatial reuse ratio and demonstrated that the bandwidth distance product and the end-to-end transmission delay (the same as the medium time) are better routing metrics than the hop count. They also proposed the metric of interference clique transmission time to achieve a high path throughput in [74]. However, these metrics or protocols are proposed for routing on a fixed path following the concept of the traditional routing. In this dissertation, we propose a framework to compute the end-to-end throughput bound of OR for different OR schemes in multi-radio multi-channel and multi-rate networks. The throughput bound derived in this paper is the upper bound of the achievable throughput of the proposed and investigated OR schemes. We also study the impact of the protocol overhead and multi-rate capability on the performance of GOR under contention-based medium access protocols.

1.2.5 Energy-aware Routing

Energy-aware routing has received significant attention over the past few years [59,16, 43,35]. Woo et al. [59] proposed five energy aware metrics such as *maximizing time to partition* and *minimizing maximum node cost*. These are important metrics for energy efficient routing. However, it is difficult to directly implement them in a local

algorithm when even the global version of the same problem is NP-complete. Chang et al. [16] proposed a class of flow augmentation algorithms and a flow redirection algorithm which balance the energy consumption rates among the nodes in proportion to the energy reserves. The limitation of this approach is that it requires the prior knowledge of the information generation rates at the origin nodes. Li et al. [43] proposed an "online" power-aware routing and a zone based routing which maximize the network lifetime without knowing the message generation rate. Following [43], another "online" routing algorithm was proposed in [35] that aims to maximize the total number of successfully delivered messages. In this dissertation, we study the energy efficiency of OR to tradeoff the routing performance and energy efficiency in terms of maximizing the bit advancement per unit energy consumption.

1.2.6 Link Quality Measurement

The existing LQM mechanisms proposed in the literature [21,38,54] can be generally classified into three types: active, passive, and cooperative [38]. For broadcast-based active probing [21], each node periodically broadcasts hello/probing packets, and its neighbors record the number of received packets to calculate the packet reception ratios (PRRs) from the node to themselves. In passive probing [38], the real traffic generated in the network is used as probing packets without introducing extra overhead. For cooperative probing [38], a node overhears the transmissions of its neighbor to estimate the link quality from the neighbor to itself. However, for any of the existing LQM mechanisms, the inherent common fact is that a node's knowledge about the forward PRR from itself to its neighbor is informed by the neighbor. Since multihop wireless networks are generally deployed in an ad hoc style or in untrusted environments, nodes may be compromised and act maliciously. This receiver-dependent measurement opens up a door for malicious attackers to report a false measurement result and disturb the routing decision for all the PRR-based protocols.

1.3 Thesis Overview

The contents of each chapter are described as follows.

Chapter 2 of this dissertation identifies and proves the principles and properties of the local behavior of GOR. We first generalize the definition of expected packet advancement (EPA) as proposed in geographic routing [56,42] but apply it to arbitrary number of forwarding candidates in OR. Then we prove that giving candidate closer to the destination higher relay priority maximizes the EPA. We further unveil that though involving more forwarding candidates increases the maximum EPA, the gained EPA becomes marginal when we keep doing so. We also show the consistency between EPA and reliability.

Chapter 3 of this dissertation studies the energy efficiency of GOR. We propose a new metric, EPA per unit energy consumption, which balances the packet advancement, reliability and energy consumption of geographic opportunistic routing (GOR). By leveraging the proved principles in Chapter 2, we then propose two efficient algorithms which select a feasible candidate set that maximizes this local metric. We validate our analysis results by simulations, and justify the effectiveness of the new metric by comparing the performance of our GOR with those of the existing geographic and opportunistic routing schemes.

Chapter 4 of this dissertation proposes the concept of concurrent transmission set which captures the transmission conflict constraints of OR. Then, for a given network with given opportunistic routing strategy (i.e., forwarder selection and prioritization), we formulate the maximum end-to-end throughput problem as a maximum-flow linear programming problem subject to the constraints of transmitter conflict. The solution of the optimization problem provides the performance bound of OR. The proposed method establishes a theoretical foundation for the evaluation of the performance of different variants of OR with various forwarding candidate selection, prioritization policies, and transmission rates. We also propose two OR metrics: *expected*

medium time (EMT) and *expected advancement rate* (EAR), and the corresponding distributed and local rate and candidate set selection schemes, one of which is Least Medium Time OR (LMTOR) and the other is Multirate Geographic OR (MGOR). Simulation results show that for OR, by incorporating our proposed multirate OR schemes, systems operating at multi-rates achieve higher throughput than a system operating at any single rate. Several insights of OR are observed: 1) the end-to-end capacity gained decreases when the number of forwarding candidates is increased; 2) there exists a node density threshold, higher than which 24Mbps GOR performs better than 12Mbps GOR, and lower than which, vice versa.

Chapter 5 of this dissertation extends the framework in Chapter 4 to compute the capacity of opportunistic routing between two end nodes in single/multi-radio/channel multihop wireless networks by allowing dynamic forwarding strategies. We study the necessary and sufficient conditions for the schedulability of a traffic demand vector associated with a transmitter to its forwarding candidates in a concurrent transmission set. We further proposed an LP approach and a heuristic algorithm to obtain an opportunistic forwarding strategy scheduling that satisfies a traffic demand vector. Our methodology can not only be used to calculate the end-to-end throughput bound of OR and TR in multi-radio/channel multihop wireless networks, but also can be used to study the OR behaviors (such as candidate selection and prioritization) in multi-radio multi-channel systems. Leveraging our analytical model, we find that OR can achieve comparable or even better performance than TR by using less radio resource.

Chapter 6 of the dissertation investigates the state-of-the-art candidate coordination schemes in OR and proposes a new scheme "fast slotted acknowledgment" (FSA) to further improve the efficiency of OR, which adopts single ACK to confirm the successful reception and suppress other candidates' attempts of forwarding the data packet with the help of channel sensing technique. We confirm the benefit

of our scheme by simulation. The results show that FSA can decrease the average end-to-end delay up to 50% when the traffic is relatively light and can improve the throughput up to 20% under heavy traffic load where the other coordination schemes are already unable to deliver all the data packets. The simulation also validates that FSA can achieve similar performance as ideal coordination where relay priority can be ensured and duplicate packet forwarding is avoided.

Chapter 7 of this dissertation studies the impact of multi-rate, candidate selection, prioritization, and coordination on the throughput of GOR under a contention-based medium access scenario. Based on our analysis, we propose a new local metric, the *expected one-hop throughput* (EOT), to characterize the trade-off between the packet advancement and one-hop packet forwarding time under different data rates. We further propose a rate adaptation and candidate selection algorithm to approach the local optimum of this metric. Simulation results show that MGOR incorporating the proposed algorithm achieves better throughput and delay performance than the corresponding opportunistic routing and geographic routing at any single rate under contention-based medium assess mechanisms.

Chapter 8 of the dissertation investigates the existing link quality measurement mechanisms, and analyzes their security vulnerabilities. A common inherent fact in all the existing LQM mechanisms are receiver-dependent measuring. That is, a node's knowledge about the forward packet reception ratio (PRR) from itself to its neighbors is informed by its neighbors. We then propose a broadcast-based secure LQM mechanism that prevents a neighboring node from maliciously claiming a higher measurement result. Our mechanism has very low computation, storage, and communication overhead. Thus, it can be implemented in resource-constraint sensor networks as well as mesh networks. Our Secure Link Quality Measurement (SLQM) mechanism can be easily applied to unicast-based and cooperative LQM with slight modifications.

1.4 Thesis Contribution

The main contributions of this dissertation are listed as follows:

- Chapter 2
  - We generalize the the definition of EPA for an arbitrary number of forwarding candidates which follow a specific priority rule to relay the packet in OR.
  - Through theoretical analysis, we prove that the maximum EPA can only be achieved by giving higher relay priorities to the forwarding candidates closer to the destination. This proof convinces us that given a forwarding candidate set, the relay priority among the candidates is only relevant to the advancement achieved by the candidate to the destination, but irrelevant to the packet delivery ratio between the transmitter and the forwarding candidate. The analysis result is the upper bound of the EPA that any GOR can achieve.
  - We find that given a set of M nodes that are available as next-hop neighbors, the candidate set achieving the maximum EPA with r ($r \leq M-1$) nodes is contained in at least one candidate set achieving the maximum EPA with $r+1$ nodes.
  - We prove that the maximum EPA of selecting r ($r \leq M$) nodes is a strictly increasing and concave function of r. This property indicates that although getting more forwarding candidates involved in GOR will increase the maximum EPA, the extra EPA gained by doing so becomes less significant.
- Chapter 3
  - We investigate the energy efficiency of GOR and propose two localized candidate selection algorithms with $\mathbf{O}(M^3)$ and $\mathbf{O}(M^2)$ running time in

the worst case respectively and $\Omega(M)$ in the best case, where M is the number of available next-hop neighbors of the transmitter. The algorithms efficiently determine the optimal forwarding candidate set with respect to the EPA per unit of energy consumption.

  - We propose an energy-efficient geographic opportunistic routing (EGOR) framework applying the node selection algorithms to achieve the energy efficiency. Simulation results show that EGOR achieves better energy efficiency than geographic routing and blind opportunistic protocols in all the cases while maintaining very good routing performance. Our simulation results also show that the number of forwarding candidates necessary to achieve the maximum energy efficiency is mainly affected by the reception to transmission energy ratio but not by the node density under a uniform node distribution. Only a very small number of forwarding candidates (around 2) are needed on average. This is true even when the energy consumption of reception is far less than that of transmission.

- Chapter 4

  - We propose a new method of constructing transmission conflict graphs, and present a methodology for computing the end-to-end throughput bounds (capacity) of OR. We formulate the maximum end-to-end throughput problem of OR as a maximum-flow linear programming problem subject to the transmission conflict constraints and effective forwarding rate on each link. To the best of our knowledge, this is the first theoretical work on capacity problem of OR for multihop and multirate wireless networks.
  - We propose two metrics for OR under multirate scenario, one is *expected medium time* (EMT), and the other is *expected advancement rate* (EAR). Based on these metrics, we propose the distributed and local rate and

candidate selection schemes: least medium time OR (LMTOR) and multi-rate GOR (MGOR), respectively.

  - We show that OR has great potential to improve the end-to-end throughput under different settings, and our proposed multi-rate OR schemes achieve higher throughput bound than any single-rate GOR.
  - We observe some insights of OR: 1) the end-to-end capacity gained decreases when the number of forwarding candidates is increased. When the number of forwarding candidates is larger than 3, the throughput almost remains unchanged. 2) there exists a node density threshold, higher than which 24Mbps GOR performs better than 12Mbps GOR, and lower than which, vice versa. The threshold is about 5.5 and 10.9 neighbors per node on 12Mbps for line and square topologies, respectively.

- Chapter 5

  - We propose a unified framework to compute the capacity of opportunistic routing between two end nodes in single/multi-radio/channel multihop wireless networks by allowing dynamic forwarding strategies.
  - We study the necessary and sufficient conditions for the schedulability of a traffic demand vector associated with a transmitter to its forwarding candidates in a concurrent transmission set.
  - We propose an LP approach and a heuristic algorithm to obtain an opportunistic forwarding strategy scheduling that satisfies a traffic demand vector.
  - Leveraging our analytical model, we find that OR can achieve comparable or even better performance than TR by using less radio resource.

- Chapter 6

  - We propose a new scheme "fast slotted acknowledgment" for candidate coordination in OR, which adopts single ACK to confirm the successful reception and suppress other candidates' attempts of forwarding the data packet with the help of channel sensing technique.
  - Simulation shows that FSA can decrease the average end-to-end delay by up to 50% when the traffic is relatively light and can improve the throughput by up to 20% under heavy traffic load where the other coordination schemes are already unable to delivery all the data packets.
  - The simulation also validates that FSA can achieve similar performance as ideal coordination where relay priority can be ensured and duplicate packet forwarding is avoided.

- Chapter 7

  - We investigate the impact of transmission rate and forwarding strategies (candidate selection, prioritization, and coordination) on throughput of OR under a contention-based medium access scenario.
  - We propose a local metric, *expected one-hop throughput* (EOT), to characterize the trade-off between the packet advancement and one-hop packet forwarding time under different data rates.
  - We propose a rate adaptation and candidate selection algorithm to approach the local optimum of this metric.
  - We propose a multi-rate link quality measurement mechanism.
  - We show that MGOR incorporating our algorithm achieves better throughput and delay performance than the corresponding opportunistic routing and geographic routing operating at any single rate, which indicates that EOT is a good local metric to achieve high end-to-end throughput and low delay for MGOR.

- Chapter 8
  - We analyze the security vulnerabilities in the existing LQM mechanisms and propose an efficient broadcast-based secure LQM (SLQM) mechanism, which prevents the malicious receiver from reporting a higher PRR than the actual one.
  - We analyze the security strength, the cost and applicability of the proposed mechanism.

1.5 System Model and Assumptions

We consider a multi-hop wireless network with N nodes arbitrarily located on a plane. Each node n_i $(1 \leq i \leq N)$ can transmit a packet at J different rates $R^1, R^2, ..., R^J$. When $J = 1$, it is reduced to a single rate system. We say there is a **usable** directed link l_{ij} from node n_i to n_j, when the **packet reception ratio** (PRR), denoted as p_{ij}, from n_i to n_j is larger than a non-negligible positive threshold p_{td} (p_{td} is set as 0.1). The PRR we consider is an average value of the link quality in a long-time scale (e.g. in tens of seconds). There exist several link quality measurement mechanisms [21, 2, 38, 54] to obtain the PRR on each link. In this dissertation, for all the analysis, we assume that there is no power control scheme and the PRR on each link is independent. We define the **effective transmission range** L_j at rate R^j $(1 \leq j \leq J)$ as the sender-receiver distance at which the PRR equals p_{td}.

The basic module of opportunistic routing is shown in Fig. 1.1. Assume node n_i is forwarding a packet to a sink/destination n_d. n_d is out of n_i's effective transmission range. We denote the set of nodes within the effective transmission range of node n_i as the **neighboring node set** $\mathcal{N}_i$ of node n_i. Note that, for different transmission rates, the corresponding effective transmission ranges are different. Thus, we have different neighboring node sets of node n_i, and the PRR on the same link may be

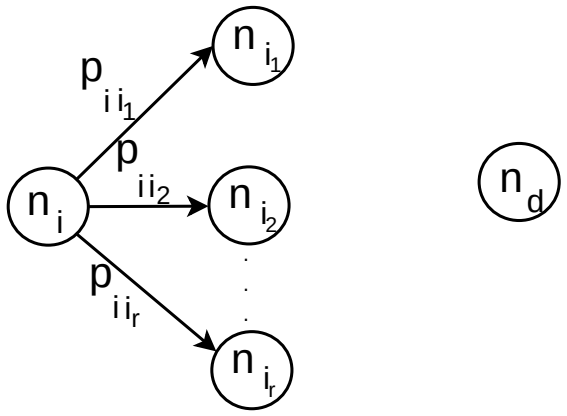


Figure 1.1: Node n_i is forwarding a packet to a remote destination n_d with a chosen forwarding candidate set $\mathcal{F}_i$ at some transmission rate.

different at different rates. We define the set $\mathcal{F}_i := \langle n_{i_1}, n_{i_2}...n_{i_r} \rangle$ shown in Fig. 1.1, as **forwarding candidate set**, which is a subset of $\mathcal{N}_i$ and includes all the nodes selected to be involved in the local opportunistic forwarding based on a particular candidate selection strategy. $\mathcal{F}_i$ is an ordered set, where the order of the elements corresponds to their priority in relaying a received packet.

For GOR, we assume each node is aware of the location information[1] of itself, its one-hop neighbors and the destination. Given a transmitter n_i, one of its forwarding candidates n_{i_q}, and the destination n_d, we define the **packet advancement** d_{i_q} in Eq. (1.1), which is the Euclidean distance between the transmitter and destination subtracting the Euclidean distance between the candidate n_{i_q} and the destination. This definition represents the advancement in distance made toward the destination when n_{i_q} forwards the packet sent by n_i.

$$d_{ii_q} = dist(n_i, n_d) - dist(n_{i_q}, n_d) \tag{1.1}$$

[1]The node location information can be obtained by prior configuration, by the Global Positioning System (GPS) receiver, or through some sensor self-configuring localization mechanisms such as [13, 55].

For GOR, since we are only interested in the neighbors which give positive advancement to the destination, we denote the set of those neighbors as $\mathcal{C}_i$, the **available next-hop node set**. Note that, for GOR, $\mathcal{F}_i$ is a subset of $\mathcal{C}_i$.

The opportunistic routing works by the sender node n_s forwarding the packet to the nodes in its forwarding candidate set $\mathcal{F}_s$. One of the candidate nodes continues the forwarding based on their relay priorities – If the first node in the set has received the packet successfully, it forwards the packet towards the destination while all other nodes suppress duplicate forwarding. Otherwise, the second node in the set is arranged to forward the packet if it has received the packet correctly. Otherwise the third node, the fourth node, etc. A forwarding candidate will forward the message only when all the nodes with higher priorities fail to do so. When no forwarding candidate has successfully received the packet, the sender will retransmit the packet if retransmission is enabled. The sender will drop the packet when the number of retransmissions exceeds the limit. The forwarding reiterates until the packet is delivered to the destination. Several MAC protocols have been proposed in [84, 9, 28, 86] to coordinate the forwarding candidates and ensure the relay priority among them. In this dissertation, for all the analysis, we assume the relay priority can be perfectly realized. So there is no duplicate packet forwarding due to imperfect candidate coordination. We will show in Chapter 6 that it is a realistic assumption when our proposed candidate coordination scheme is used.

For capacity analysis in Chapter 4 and 5, we assume that packet transmissions at the individual nodes can be finely controlled and carefully scheduled by an omniscient and omnipotent central entity. So here we do not concern ourselves with issues such as MAC contention or coordination overhead that may be unavoidable in a distributed network. This is a very commonly used assumption for such theoretical study [32, 74].

1.6 Table of Abbreviations

ACK	ACKnowledgement
ARQ	Automatic Repeat reQuest
CTS	Concurrent Transmission Set
EAR	Expected Advancement Rate
EMT	Expected Medium Time
EPA	Expected Packet Advancement
EOT	Expected One-hop Throughput
ExOR	Extremely Opportunistic Routing
FSA	Fast Slotted Acknowledgement
GeRaF	Geographic Random Forwarding
GOR	Geographic Opportunistic Routing
GPS	Global Positioning System
LCOR	Least Cost Opportunistic Routing
LMTOR	Least Medium Time Opportunistic Routing
LP	Linear Programming
LQM	Link Quality Measurement
MAC	Medium Assess Control
MANET	Mobile Ad hoc NETwork
MGOR	Multirate Geographic Opportunistic Routing
NADV	Normalized ADVance
OR	Opportunistic Routing
PRR	Packet Reception Ratio
SLQM	Secure Link Quality Measurement
TR	Traditional Routing
WCETT	Weighted Cumulative Expected Transmission Time
WSN	Wireless Sensor Network

Table 1.1: Abbreviations used

Chapter 2

Understanding Geographic Opportunistic Routing

This chapter analyzes the principles of the local behavior of GOR. We first generalize the definition of expected packet advancement (EPA) for arbitrary number of forwarding candidates which follow a specific priority rule to relay the packet in OR. Through theoretical analysis, we prove that the maximum EPA can only be achieved by giving the forwarding candidates closer to the destination higher relay priorities. This **relay priority rule** convinces us that given a forwarding candidate set, the relay priority among the candidates is only relevant to the advancement achieved by the candidate to the destination, but irrelevant to the packet delivery ratio between the transmitter and the forwarding candidate. The analysis result is the upper bound of the EPA that any GOR can achieve. We further prove that, given a set of M nodes that are available as next-hop neighbors, a subset of the available next-hop neighbors with r ($r < M$) nodes achieving the maximum EPA is contained in a subset with more nodes achieving the maximum EPA. Leveraging the **containing property**, we unveil that the maximum EPA of selecting r ($r \leq M$) nodes is a strictly **increasing** and **concave** function of r. This property indicates that although getting more for-

warding candidates involved in GOR will increase the maximum EPA, the extra EPA gained by doing so becomes less significant. It also implies the consistency between EPA and reliability. These principles of GOR will help us analyze the capacity of OR in Chapter 4 and 5, and design efficient local candidate selection and prioritization algorithms for achieving energy and throughput efficiency in Chapter 3 and 7, respectively.

In this chapter, since we mainly focus on the local behavior of GOR, for a given transmitter n_i, we abbreviate its **forwarding candidate set** $\mathcal{F}_i$ as $\mathcal{F}$, and its **available next-hop node set** $\mathcal{C}_i$ as $\mathcal{C}$. Note that, $\mathcal{F}$ is an ordered subset of $\mathcal{C}$, which is a set of all the neighbors that are geographically closer to the destination than the transmitter n_i. n_i's neighbor n_{i_q}, its advancement to the destination d_{ii_q}, and the PRR p_{ii_q} on link l_{ii_q} are simplified as i_q, d_q, and p_q, respectively. We denote the number of nodes in $\mathcal{F}$ as r, and the number of nodes in $\mathcal{C}$ as M. Redefine $\mathcal{F} := \langle i_1, ..., i_r \rangle$, and $\mathcal{C} := \{i_1, ..., i_M\}$. Note that, the subscript of i only represents the sequence number of each node in set $\mathcal{F}$ and $\mathcal{C}$, and two nodes having the same subscript in $\mathcal{F}$ and $\mathcal{C}$ are not necessarily the same node. For example, i_1 in $\mathcal{F}$ does not necessarily indicate the same node as i_1 in $\mathcal{C}$. Without loss of generality, we assume all the nodes in $\mathcal{C}$ and $\mathcal{F}$ are descending ordered according to the advancement s.t. given nodes i_m and i_n, we have $d_m > d_n$, $\forall\ m < n$.

2.1 EPA Generalization

Let $\pi(\mathcal{F}) = \langle i_{\pi_1}, i_{\pi_2}, ..., i_{\pi_r} \rangle$ be one permutation of nodes in $\mathcal{F}$, and the order indicates that nodes will attempt to forward the packet with priority $i_{\pi_1} > i_{\pi_2} > ... > i_{\pi_r}$. We define the EPA for the ordered forwarding candidate set $\pi(\mathcal{F})$ in Eq. (2.1)

$$\mathrm{EPA}(\pi(\mathcal{F})) = \sum_{k=1}^{r} d_{\pi_k} p_{\pi_k} \cdot \prod_{n=0}^{k-1} \overline{p}_{\pi_n} \tag{2.1}$$

where $\overline{p}_{\pi_n} = 1 - p_{\pi_n}$ and $\overline{p}_{\pi_0}$:=1. The physical meaning of Eq. (2.1) is the expected packet advancement achieved by GOR in one transmission using the ordered forwarding candidate set $\pi(\mathcal{F})$. The EPA metric accurately indicates the relationship between the packet advancement and candidate selection and prioritization. Note that when $r = 1$, Eq. (2.1) degenerates to the "distance×PRR" proposed in geographic routing [56, 42].

2.2 Principles of Local Behavior of GOR

2.2.1 EPA Strictly Increasing Property

Intuitively, increasing the number of forwarding candidates would result in a larger EPA. We present Lemma 2.2.2 to confirm this intuition.

Definition 2.2.1. *Define EM($\mathcal{C}, r$) be the maximum EPA (defined in Eq. (2.1)) achieved by selecting r forwarding candidates from $\mathcal{C}$.*

Lemma 2.2.2. ***(Strictly increasing property)*** *EM($\mathcal{C}, r$) is a strictly increasing function of r.*

Proof. Assume $1 \leq m < n \leq M$, and without loss of generality, let $\mathcal{A}$=$\langle i_1, i_2, ..., i_m \rangle$ be the ordered node set achieving EM($\mathcal{C}, m$) with forwarding priority $i_1 > ... > i_m$. We then select a subset with $n-m$ nodes from the remaining node set $\{i_{m+1}, i_{m+2}, ..., i_M\}$, say $\mathcal{B}$=$\langle i_{m+1}, ..., i_n \rangle$. Assume we retain the relay priority of the m nodes in $\mathcal{A}$ unchanged and give the nodes in $\mathcal{B}$ lower priorities than those in $\mathcal{A}$. Then in $\mathcal{B}$, we give the nodes with smaller subscripts higher relay priorities. So we have

$$\mathrm{EM}(\mathcal{C}, n) \geq \mathrm{EPA}(\langle i_1, ..., i_n \rangle) = \mathrm{EM}(\mathcal{C}, m) + \mathrm{EPA}(\langle i_{m+1}, ..., i_n \rangle) \prod_{k=1}^{m} \overline{p}_k > \mathrm{EM}(\mathcal{C}, m)$$

□

Lemma 2.2.2 basically indicates that the more nodes get involved in GOR, the larger the EPA can be. The maximum EPA can be obtained by involving all the nodes in $\mathcal{C}$. Then, how to prioritize the candidates to maximize the EPA? We answer this question in the following section.

2.2.2 Relay Priority Rule

Theorem 2.2.3 identifies the upper bound of EPA and the corresponding relay priority rule.

Theorem 2.2.3. ***(Relay priority rule)*** *EM($\mathcal{F}, |\mathcal{F}|$) can only be obtained by giving the node closer to the destination higher relay priority. That is*

$$\mathrm{EM}(\mathcal{F}, |\mathcal{F}|) = \sum_{k=1}^{r} d_k p_k \cdot \prod_{n=0}^{k-1} \overline{p}_n \tag{2.2}$$

where $\overline{p}_0 := 1$.

Proof. We proof Theorem 2.2.3 by induction on r, the size of $\mathcal{F}$.

First, when r = 1, obviously Eq. (2.2) holds.

Next, we assume Eq. (2.2) holds for r = N (N≥1). When $|\mathcal{F}|$ = N+1, $\mathcal{F}$ can be divided into $\mathcal{F}_1$=$\mathcal{F} - \{i_m\}$ with N nodes and $\mathcal{F}_2 = \{i_m\}$ with 1 node. Then

$$\mathrm{EM}(\mathcal{F}, |\mathcal{F}|) = \max_{1 \le m \le N+1} \{ \sum_{k=1}^{m-1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w + \sum_{k=m+1}^{N+1} d_k p_k \frac{\prod_{w=0}^{k-1} \overline{p}_w}{\overline{p}_m} + d_m p_m \frac{\prod_{w=0}^{N+1} \overline{p}_w}{\overline{p}_m} \}$$

Thus we only need to prove for any integer m $(1 \le m \le N)$,

$$\begin{aligned} A &:= \sum_{k=1}^{m-1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w + \sum_{k=m+1}^{N+1} d_k p_k \frac{\prod_{w=0}^{k-1} \overline{p}_n}{\overline{p}_m} + d_m p_m \frac{\prod_{w=0}^{N+1} \overline{p}_w}{\overline{p}_m} \\ &< B := \sum_{k=1}^{N+1} d_k p_k \prod_{w=0}^{k-1} \overline{p}_w \end{aligned}$$

Subtracting A from B, we have

$$B - A = \frac{1}{\overline{p}_m} \sum_{k=m+1}^{N+1} (d_m - d_k) p_m p_k \prod_{w=0}^{k-1} \overline{p}_w > 0$$

Then the Eq. (2.2) holds for r = N+1. So it holds for any r ($r \geq 1$). □

Theorem 2.2.3 indicates that when a forwarding candidate set is chosen, the maximum EPA can only be achieved by assigning the relay priority to each node based on their distances to the destination. That is, the furthest node should try to forward the packet first; if it failed (i.e., did not receive the packet correctly), the second furthest node should try next, and so on. The analysis result is the upper bound of the EPA that any GOR can achieve.

Based on the **Relay priority rule**, next, we will identify and prove two important principles about the maximum EPA. First, we look at the characteristics of the forwarding candidates that are selected to achieve $\mathrm{EM}(\mathcal{C}, r)$ with various sizes r. We prove the **Containing property** for those node sets. Following that, the **Concavity** of the function $\mathrm{EM}(\mathcal{C}, r)$ is proved.

2.2.3 Containing Property of Feasible Candidate Set

Let $\mathcal{F}_r^*$ be a feasible ordered node set that achieves the $\mathrm{EM}(\mathcal{C}, r)$, we have the following containing property of $\mathcal{F}_r^*$'s.

Lemma 2.2.4. ***(Containing property)*** *Given the available next-hop node set $\mathcal{C}$ with M nodes, $\forall\ \mathcal{F}_{r-1}^*$, $\exists\ \mathcal{F}_r^*$, s.t.*

$$\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^* \quad \forall\ 1 \leq r \leq M \tag{2.3}$$

Proof. Let $\mathcal{A}=\langle a_1, ..., a_M\rangle$[1] be an ordered node set with M nodes, and $\mathcal{B}=\langle b_1, ..., b_N\rangle$ with N nodes. $\mathcal{B} \subset \mathcal{A}$ and $b_N = a_M$. For any node $q \notin \mathcal{A}$ with $d_q < d_{a_M}$, we have

$$\mathrm{EPA}(\langle \mathcal{A}, q\rangle) - \mathrm{EPA}(\langle q, \mathcal{A}\rangle) > \mathrm{EPA}(\langle \mathcal{B}, q\rangle) - \mathrm{EPA}(\langle q, \mathcal{B}\rangle) \tag{2.4}$$

We then prove Lemma 2.2.4 by induction on r.

[1]For simplicity, we denote node using its subscript in this proof.

First, for arbitrary N, when $r = 1$, as $\mathcal{F}_0^* = \emptyset$, and $\mathcal{F}_1^* \neq \emptyset$, it is obvious that the containing property holds.

Then, we assume $\forall\ \mathcal{F}_{m-1}^*$, $\exists$ an $\mathcal{F}_m^*$, s.t. $\mathcal{F}_{m-1}^* \subset \mathcal{F}_m^*$, when r=m ($m \geq 1$). We first prove for any feasible $\mathcal{F}_m^*$ and $\mathcal{F}_{m+1}^*$, the first node in $\mathcal{F}_{m+1}^*$ can not be the nodes from the second place to the last place in $\mathcal{F}_m^*$, that is $(m+1)_1 \neq m_i$, $\forall\ 2 \leq i \leq m$.

We prove this by contradiction. Assume $(m+1)_1 = m_i$. Let node $(m+1)_j$ be the first node in $\mathcal{F}_{m+1}^*$ but not in $\mathcal{F}_m^*$. We have

$$\mathrm{EPA}(\mathcal{F}_m^*) \geq \mathrm{EPA}(\mathcal{F}_{m+1}^* - \{(m+1)_j\}) \tag{2.5}$$

then,

$$\mathrm{EPA}(\langle (m+1)_j, \mathcal{F}_m^* \rangle) \geq \mathrm{EPA}(\langle (m+1)_j, \mathcal{F}_{m+1}^* - \{(m+1)_j\} \rangle) \tag{2.6}$$

Assume $(m+1)_{j-1} = m_l$, and according to inequality (2.4), we have

$$\Delta 1 > \Delta 2 \tag{2.7}$$

where

$$\begin{aligned} \Delta 1 &:= \mathrm{EPA}(\langle m_1, ..., m_l, (m+1)_j, m_{l+1}, ..., m_m \rangle) - \mathrm{EPA}(\langle (m+1)_j, \mathcal{F}_m^* \rangle) \\ &= \mathrm{EPA}(\langle m_1, ..., m_l, (m+1)_j \rangle) - \mathrm{EPA}(\langle (m+1)_j, m_1, ..., m_l \rangle) \end{aligned} \tag{2.8}$$

$$\begin{aligned} \Delta 2 &:= \mathrm{EPA}(\mathcal{F}_{m+1}^*) - \mathrm{EPA}(\langle (m+1)_j, \mathcal{F}_{m+1}^* - \{(m+1)_j\} \rangle) \\ &= \mathrm{EPA}(\langle (m+1)_1, ..., (m+1)_{j-1}, (m+1)_j \rangle) \\ &\quad - \mathrm{EPA}(\langle (m+1)_j, (m+1)_1, ..., (m+1)_{j-1} \rangle) \end{aligned} \tag{2.9}$$

Then combining with inequality (2.6), we get

$$\mathrm{EPA}(\langle m_1 ... m_l, (m+1)_j, m_{l+1} ... m_m \rangle) > \mathrm{EPA}(\mathcal{F}_{m+1}^*) \tag{2.10}$$

The inequality (2.10) contradicts with the fact that $\mathrm{EPA}(\mathcal{F}_{m+1}^*)$ is the largest EPA achieved by selecting m+1 nodes. So the assumption $(m+1)_1 = m_i$ is wrong, then $(m+1)_1$ can not be m_i, $\forall\ 2 \leq i \leq m$. So there are two cases for $(m+1)_1$:

1)$(m+1)_1 \neq m_1$. Then $\langle (m+1)_1, \mathcal{F}^*_m \rangle$ should be one $\mathcal{F}^*_{m+1}$.

2)$(m+1)_1 = m_1$. By the inductive hypothesis, we have $\mathcal{F}^*_m - \{m_1\} \subset \langle (m+1)_2, ..., (m+1)_{m+1} \rangle$, then $\mathcal{F}^*_m \subset \mathcal{F}^*_{m+1}$.

From the induction above, we know for arbitrary N, we have $\forall\ \mathcal{F}^*_{r-1}$, $\exists\ \mathcal{F}^*_r$ s.t. $\mathcal{F}^*_{r-1} \subset \mathcal{F}^*_r$, $\forall\ 1 \leq r \leq M$. □

Lemma 2.2.4 indicates that an $r-1$-node set that achieves EM$(\mathcal{C}, r-1)$ is a subset of at least one of the feasible r-node sets that achieve EM$(\mathcal{C}, r)$. It also implies that the increasing of the maximum EPA consists with the increasing of the transmission reliability.

2.2.4 Concavity of Maximum EPA

Following Lemma 2.2.4, we have the concave property of EM$(\mathcal{C}, r)$ as in Theorem 2.2.5.

Theorem 2.2.5. ***(Concavity of maximum EPA)***
EM($\mathcal{C}$*,r+1)−EM(*$\mathcal{C}, r$*)<EM(*$\mathcal{C}, r$*)−EM(*$\mathcal{C}$*,r−1),* $\forall\ r$*, s.t.* $1 \leq r < N$.

Proof. According to Lemma 2.2.4, assume $\mathcal{F}^*_{r+1} - \mathcal{F}^*_r = \{i_k\}$, and $\mathcal{F}^*_r - \mathcal{F}^*_{r-1} = \{i_j\}$. There are two cases for d_k and d_j.

1) $d_k > d_j$. Then $\mathcal{F}^*_{r+1}$, $\mathcal{F}^*_r$ and $\mathcal{F}^*_{r-1}$ can be represented as

$$\mathcal{F}^*_{r+1} = \langle \mathcal{A}_1, i_k, \mathcal{A}_2, i_j, \mathcal{A}_3 \rangle,\ \mathcal{F}^*_r = \langle \mathcal{A}_1, \mathcal{A}_2, i_j, \mathcal{A}_3 \rangle,\ \mathcal{F}^*_{r-1} = \langle \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \rangle$$

where $\mathcal{A}_i$ $(1 \leq i \leq 3)$ is ordered node set and can be $\emptyset$.

We have

$$\mathrm{B} := \mathrm{EPA}(\mathcal{F}^*_r) - \mathrm{EPA}(\langle \mathcal{A}_1, i_k, \mathcal{A}_2, \mathcal{A}_3 \rangle) \geq 0 \tag{2.11}$$

Then,

$$\begin{aligned} &[\mathrm{EPA}(\mathcal{F}^*_r) - \mathrm{EPA}(\mathcal{F}^*_{r-1})] - [\mathrm{EPA}(\mathcal{F}^*_{r+1}) - \mathrm{EPA}(\mathcal{F}^*_r)] \\ &= \mathrm{B} + \overline{p}_{\mathcal{A}_1} \overline{p}_{\mathcal{A}_2} p_k p_j (d_j - \mathrm{EPA}(\mathcal{A}_3)) > 0 \end{aligned}$$

where $\overline{p}_{\mathcal{A}_i}$ is the probability of none of nodes in $\mathcal{A}_i$ receiving the packet correctly.

2) $d_k < d_j$. Similarly, with

$$\mathrm{B} := \mathrm{EPA}(\mathcal{F}_r^*) - \mathrm{EPA}(\langle \mathcal{A}_1, \mathcal{A}_2, i_k, \mathcal{A}_3 \rangle) \geq 0 \tag{2.12}$$

we can derive that

$$\begin{aligned} &[\mathrm{EPA}(\mathcal{F}_r^*) - \mathrm{EPA}(\mathcal{F}_{r-1}^*)] - [\mathrm{EPA}(\mathcal{F}_{r+1}^*) - \mathrm{EPA}(\mathcal{F}_r^*)] \\ &= \mathrm{B} + \overline{p}_{\mathcal{A}_1} \overline{p}_{\mathcal{A}_2} p_k p_j (d_k - \mathrm{EPA}(\mathcal{A}_3)) > 0 \end{aligned}$$

From the analysis above, we know $\mathrm{EM}(\mathcal{C}, r)$ is a concave function of r. □

Combining Lemma 2.2.2 and Theorem 2.2.5, we know that giving an available next-hop node set $\mathcal{C}$ with N nodes, the maximum EPA of selecting r $(1 \leq r \leq N)$ nodes is a ***strictly increasing and concave function*** of r. This means that although the maximum EPA keeps increasing when more nodes get involved, the speed of the increase slows down. When many nodes are involved, the gained extra EPA becomes marginal.

2.2.5 Reliability Increasing Property

Following the **Containing property** in Lemma 2.2.4, we have the **Reliability increasing property** in Corollary 2.2.6.

Denote $\mathcal{F}_r^* = \langle i_{r_1}, i_{r_2}, ..., i_{r_r} \rangle$. Define the one-hop reliability $P_{\mathcal{F}_r^*}$ in Eq. (2.13) which is the probability of at least one node in $\mathcal{F}_r^*$ correctly receiving the packet sent by node i for one transmission.

$$P_{\mathcal{F}_r^*} = 1 - \prod_{n=0}^{r} (1 - p_{r_n}) \tag{2.13}$$

where $p_{r_0} := 0$.

Define $P^*(r)$ in Eq. (2.14) which is the maximum one-hop reliability achieved by one of the feasible $\mathcal{F}_r^*$'s.

$$P^*(r) = \max_{\forall \mathcal{F}_r^*} \{P_{\mathcal{F}_r^*}\} \tag{2.14}$$

Corollary 2.2.6. *$P^*(r)$ defined in Eq. (2.14) is an increasing function of r.*

Proof. The proof is straightforward following Lemma 2.2.4. Assume one $\mathcal{F}_r^*$ achieves $P^*(r)$, then $\exists\, \mathcal{F}_{r+1}^*$ s.t. $\mathcal{F}_{r+1}^* \supset \mathcal{F}_r^*$. According to the definitions of $P^*(r)$ in Eq. (2.14) and the one-hop reliability in Eq. (2.13), we have

$$P^*(r+1) \geq P_{\mathcal{F}_{r+1}^*} > P_{\mathcal{F}_r^*} = P^*(r)$$

So $P^*(r)$ is an increasing function of r. □

Corollary 2.2.6 indicates that the maximum one-hop reliability corresponding to the forwarding candidate set that maximizes the EPA also increases when more forwarding candidates are involved. The increasing of the maximum EPA implies increasing of the reliability. Therefore, the EPA is a good metric for balancing the packet advancement and reliability.

2.3 Conclusions

In this chapter, we generalized the definition of EPA for an arbitrary number of forwarding candidates in GOR. Through theoretical analysis, we first showed that the maximum EPA can only be achieved by following a relay priority rule – giving the forwarding candidates closer to the destination higher relay priorities when a forwarding candidate set is given. We gave the analytical result of the upper bound of the EPA that any GOR can achieve. We found that the node set achieving the maximum EPA with r nodes is contained in at least one node set achieving the maximum EPA with $r+1$ nodes. We also showed that giving an available next-hop neighbor set with M nodes, the maximum EPA achieved by selecting r nodes is a strictly increasing and concave function of r and we show how the candidates should be selected to achieve the maximum EPA. We further show that the increasing of the maximum EPA is consistent with the increasing of the one-hop reliability. These

unveiled properties of the local behavior of GOR will enable us to design efficient local routing metric and candidate selection and prioritization algorithms to approach the global optimum performance.

Chapter 3

Energy Efficiency of Geographic Opportunistic Routing

Wireless sensor networks (WSNs) are characterized by multihop lossy wireless links and severely resource constrained nodes. Among the resource constraints, energy is probably the most crucial one since sensor nodes are typically battery powered and the lifetime of the battery imposes a limitation on the operation hours of the sensor network. Unlike the microprocessor industry or the communication hardware industry, where computation capability or the line rate has been continuously improved (regularly doubled every 18 months), battery technology has been relatively unchanged for many years. Energy efficiency has been a critical concern in wireless sensor network protocol design. Researchers are investigating energy conservation at every layer in the traditional protocol stack, from the physical layer up to the network layer and application layer.

Among the energy consumption factors, communication has been identified as the major source of energy consumption and costs significantly more than computation in WSNs [51]. Opportunistic routing has shown its advantage on energy efficiency [84,67] comparing to traditional routing. However, the existing opportunistic routing schemes

like GeRaF [84] typically include all the available next-hop neighbors as forwarding candidates, which does not lead to optimal energy efficiency.

In this chapter, we propose an energy-efficient geographic opportunistic routing (EGOR) framework which is based on opportunistic routing but more judiciously selects a subset of the available next-hop neighbors as the forwarding candidates to strike a good balance between the packet advancement and energy cost. The analysis on how to achieve maximum EPA in Chapter 4 provides us useful insights on the selection of the forwarding candidate set. Based on which, we propose two localized candidate selection algorithms with $\mathbf{O}(M^3)$ and $\mathbf{O}(M^2)$ running time in the worst case respectively and $\Omega(M)$ in the best case, where M is the number of available next-hop neighbors of the transmitter. The algorithms efficiently determine the optimal forwarding candidate set with respect to the EPA per unit of energy consumption. The performance of EGOR is justified through extensive simulations and comparisons with those of the existing geographic routing and opportunistic routing schemes. The simulation results show that EGOR strikes a good balance between energy consumption and routing efficiency in terms of EPA, and achieves the best energy efficiency among the three schemes in all the cases. Our simulation results also show that under a realistic lossy channel model, the best energy efficiency can be achieved with only a very small number of forwarding candidates (around 2), even when the energy consumption of reception is negligible to that of transmission.

The rest of the chapter is organized as follows. We formulate the EGOR problem in Section 3.1. Two efficient localized candidate selection algorithms are proposed in Section 3.2. In Section 3.3, we propose and analyze our EGOR scheme. Simulation results are presented in Section 3.4. Conclusions are drawn in Section 3.5.

3.1 Problem Formulation

3.1.1 Energy Consumption Model

Here we do not assume the promiscuous mode in which every node "overhears" the transmission within its range. Instead, being energy efficient, we assume nodes/sensors only listen to the transmissions intended to themselves. To achieve this, a second low power radio [63] can be used to wake up nodes that should participate in the EGOR or to inform the neighbors who (including nodes giving negative advancement) are not selected as forwarding candidates to shut down their data radios. Nodes can also only read the header of packets for early rejection [56]. For simplicity, we also ignore the energy consumption of the control packets[1], as usually control packets are much smaller than data packets. We only consider the energy consumption of packet transmission and reception[2]. So the total energy consumption for one opportunistic forwarding attempt is:

$$E_t(r) = E_{tx} + r \cdot E_{rx} \tag{3.1}$$

where E_{tx} and E_{rx} are the packet transmission and reception energy consumption, respectively. Recall that r is the number of candidates in the forwarding candidate set $\mathcal{F}$.

[1]For different MAC protocols, the energy consumption of control packets may be different. However, the energy consumption is likely a non-decreasing function of the number of forwarding candidates. So ignoring it will not affect the upper bound analysis of the energy efficiency in this paper.

[2]In sensor networks, the energy consumption of reception is comparable to that of transmission [22], so is non-negligible.

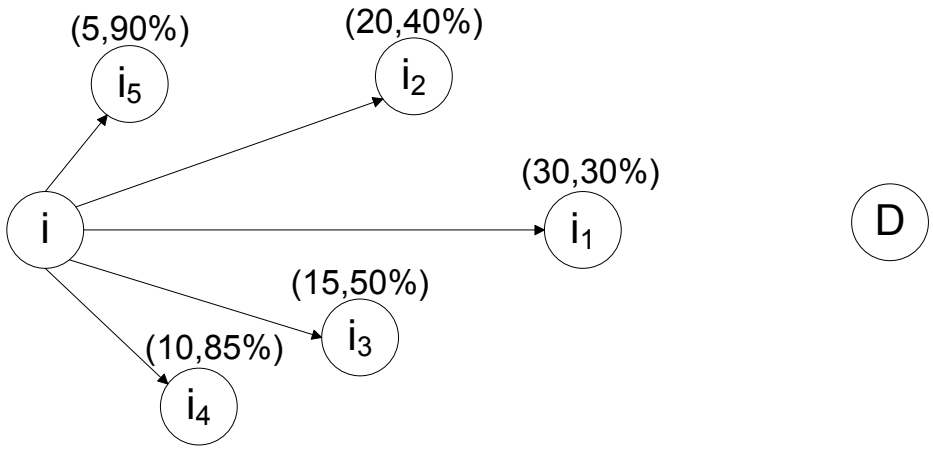


Figure 3.1: Example in which node i is forwarding a packet to a remote destination D.

3.1.2 Trade-off between EPA and Energy Consumption

As we have proved in Lemma 2.2.2 that the more nodes get involved in GOR, the larger the EPA can be. So the GOR that involves all the nodes in $\mathcal{C}$ will achieve the largest EPA. This fact has been implicitly used in the existing opportunistic routing approaches. However, it is not always the most energy efficient way to forward packets by involving all the nodes in $\mathcal{C}$. As from Eq. (3.1), we know one transmission from the transmitter is accompanied by r receptions of the r forwarding candidates, so involving all the nodes in $\mathcal{C}$ consumes the most energy. On the other hand, conventional geographic routing involving only one forwarding candidate has the least energy cost of one transmission and one reception, but it achieves the least EPA per hop. This indicates lower routing efficiency as more hops (transmissions) might be necessary to reach the final destination. Clearly there is a trade-off between the per-hop routing efficiency and the overall energy efficiency.

This trade-off is illustrated in Fig. 3.2 which is corresponding to the example in Fig. 3.1 by assuming $E_{tx} = 1$ unit of energy, $E_{rx} = 0.5$ unit. Note that although $\mathrm{EM}(\mathcal{C}, r)$ and $E_t(r)$ are both strictly increasing function of r, the ratio $\frac{\mathrm{EM}(\mathcal{C},r)}{E_c(r)}$ reaches

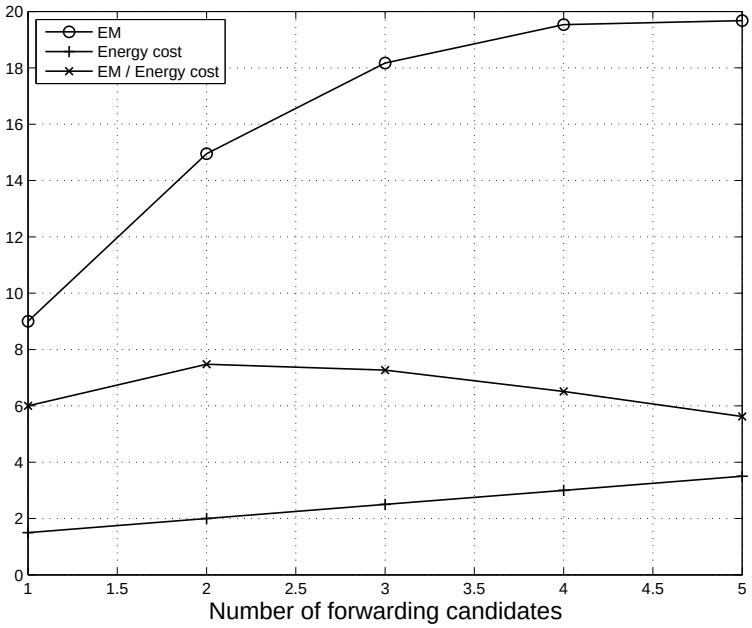


Figure 3.2: EM, energy cost and their ratio as functions of number of forwarding candidates

its maximum at $r = 2$, and the corresponding ordered node set is $\langle i_1, i_4 \rangle$ with node i_1 having higher relay priority than i_4.

Based on the analysis above, we propose a new local metric which aims to strike a good balance between the routing efficiency and energy efficiency. The new metric is denoted as $G(\pi(\mathcal{F}))$ and defined in Eq. (3.2) as follows.

$$G(\pi(\mathcal{F})) = \frac{\mathrm{EPA}(\pi(\mathcal{F})) \cdot L_{pkt}}{E_t(|\mathcal{F}|)} \tag{3.2}$$

where L_{pkt} is the packet length in bits, and $\mathrm{EPA}(\pi(\mathcal{F}))$ is defined in Eq. (2.1). If unit of $\mathrm{EPA}(\pi(\mathcal{F}))$ is meter, and $E_t(|\mathcal{F}|)$ is Joule, the unit of $G(\pi(\mathcal{F}))$ is bmpJ. The physical meaning of $G(\pi(\mathcal{F}))$ is the expected bit advancement to the destination by consuming one Joule of energy per packet forwarding attempt.

Now, our goal is to find a way to find a $\mathcal{F}^*$ which maximizes the metric $G(\mathcal{F}^*)$,

which can be formulated as the following optimization problem

$$\mathcal{F}^* = \text{argmax}_{\mathcal{F} \subseteq \bigcup_{S \subseteq 2^{\mathcal{C}}} Sym(S)} G(\mathcal{F}) \tag{3.3}$$

where $2^{\mathcal{C}}$ is the powerset of $\mathcal{C}$ and $Sym(S)$ is the set of all the permutations of S. Solving this optimization problem needs to answer the following two questions: a) How many and which nodes should be involved in the local forwarding? b) What priority should they follow to forward a packet?

3.2 Efficient Localized Node Selection Algorithms

3.2.1 Reformulate the Node Selection Optimization Problem

We know that when the number of neighbors involved in GOR is given, the denominator of the function $G(\pi(\mathcal{F}))$ defined in Eq. (3.2) is fixed, then maximizing $G(\pi(\mathcal{F}))$ is equivalent to maximize its numerator. So we can find the suboptimal solution for each $r = 1, 2, \cdots, N$, then get a global optimal solution by picking the largest one of the suboptimal solutions. From this analysis, also as the packet length L_{pkt} is fixed, combining Eq. (3.1), the optimization problem in (3.3) is equivalent to

$$Maximize \;\; M(r) := \frac{\text{EM}(\mathcal{C},r)}{E_{tx}+r \cdot E_{rx}} \;\; s.t. \;\; 1 \leq r \leq |\mathcal{C}| \tag{3.4}$$

We now introduce the following Corollary that can help us solve this optimization problem more efficiently.

Corollary 3.2.1. ***(Local maximum of M(r) is global maximum)*** *Given the available next-hop node set $\mathcal{C}$ with $|\mathcal{C}| = M$ ($M \geq 1$), the receiving energy consumption $E_{rx} > 0$ and transmission energy consumption $E_{tx} > 0$, the local maximum of the objective function $M(r)$ defined in (3.4) is the global maximum. That is, if $M(k-1) < M(k)$ and $M(k) \geq M(k+1)$ ($1 \leq k \leq M$), $M(k) \geq M(k+n)$, $\forall\ 1 \leq n \leq M-k$.*

Proof. As

$$M(k) \geq M(k+1)$$

that is

$$\frac{\mathrm{EM}(\mathcal{C},k)}{E_{tx}+k\cdot E_{rx}} \geq \frac{\mathrm{EM}(\mathcal{C},k+1)}{E_{tx}+(k+1)E_{rx}} \Rightarrow n \times \frac{\mathrm{EM}(\mathcal{C},k+1)-\mathrm{EM}(\mathcal{C},k)}{\mathrm{EM}(\mathcal{C},k)} \leq n \times \frac{E_{rx}}{E_{tx}+k\cdot E_{rx}} \tag{3.5}$$

Since $\mathrm{EM}(\mathcal{C}, r)$ is concave and positive, we have

$$\frac{\mathrm{EM}(\mathcal{C},k+n)-\mathrm{EM}(\mathcal{C},k)}{\mathrm{EM}(\mathcal{C},k)} \leq n \times \frac{\mathrm{EM}(\mathcal{C},k+1)-\mathrm{EM}(\mathcal{C},k)}{\mathrm{EM}(\mathcal{C},k)} \tag{3.6}$$

From inequality (3.5) and (3.6), we have

$$\frac{\mathrm{EM}(\mathcal{C},k+n)-\mathrm{EM}(\mathcal{C},k)}{\mathrm{EM}(\mathcal{C},k)} \leq n \times \frac{E_{rx}}{E_{tx}+k\cdot E_{rx}} \Rightarrow \frac{\mathrm{EM}(\mathcal{C},k)}{E_{tx}+k\cdot E_{rx}} \geq \frac{\mathrm{EM}(\mathcal{C},k+n)}{E_{tx}+(k+n)E_{rx}}$$

that is $M(k) \geq M(k+n)$, $\forall\ 1 \leq n \leq M-k$. □

3.2.2 Efficient Node Selection Algorithms

3.2.2.1 Algorithm based on Lemma 2.2.4 and Corollary 3.2.1

Based on the **Containing property** in Lemma 2.2.4, a straightforward way to find an optimal node set containing r nodes is to add a new node into the optimal node set containing r-1 nodes. Furthermore, when a local maximum is found, it is the global maximum based on Corollary 3.2.1. The algorithm GetM-A in Table 3.1 finds an optimal forwarding candidate set $\mathcal{F}^*$ and the corresponding energy efficiency value M^* of the objective function defined in (3.4). Note that $\mathcal{F}^*$, $\mathcal{F}_c^*$ and $\mathcal{F}$ are all ordered node sets with nodes closer to the destination having higher relay priorities. For feasible sets having the same maximum EPA, we choose the one that achieves higher one-hop reliability (line 6).

It's not difficult to find an algorithm to calculate $\mathrm{EPA}(\mathcal{F})$ (in line 5) in $\mathbf{O}(|\mathcal{F}|)$ running time. Then the algorithm GetM-A costs $\mathbf{O}(M^3)$ running time in the worst case, and in the best case it only costs $\Omega(M)$.

GetM-A($\mathcal{C}$,E_{tx},E_{rx})

1 $M^* \leftarrow M_c^* \leftarrow 0$; $A^* \leftarrow A_c^* \leftarrow 0$;

2 $\mathcal{F}^* \leftarrow \mathcal{F}_c^* \leftarrow \emptyset$; $P_{\mathcal{F}^*} \leftarrow P_{\mathcal{F}_c^*} \leftarrow 0$; $\mathcal{B} \leftarrow \mathcal{C}$; /* Initialization */

3 **while** ($\mathcal{B} \neq \emptyset$) **do** /* $\mathcal{B}$ is the remained node set */

4 **for** each node $i_j \in \mathcal{B}$

5 $\mathcal{F} \leftarrow \mathcal{F}^* \bigcup \{i_j\}$; $P_{\mathcal{F}} \leftarrow 1 - (1 - P_{\mathcal{F}^*})(1 - p_j)$; $A \leftarrow \text{EPA}(\mathcal{F})$;

6 **if** $A > A_c^*$ || ($A = A_c^*$ & $P_{\mathcal{F}} > P_{\mathcal{F}_c^*}$)

7 $A_c^* \leftarrow A$; $\mathcal{F}_c^* \leftarrow \mathcal{F}$; $P_{\mathcal{F}_c^*} \leftarrow P_{\mathcal{F}}$;

8 **end for**

9 $M_c^* \leftarrow A_c^*/(E_{tx} + |\mathcal{F}_c^*| \cdot E_{rx})$;

10 **if** $M_c^* \leq M^*$ /* Local maximum is found */

11 **return**(M^*, $\mathcal{F}^*$);

12 **else**

13 $\mathcal{B} \leftarrow \mathcal{C} \setminus \mathcal{F}_c^*$; $A^* \leftarrow A_c^*$; $M^* \leftarrow M_c^*$; $\mathcal{F}^* \leftarrow \mathcal{F}_c^*$; $P_{\mathcal{F}^*} \leftarrow P_{\mathcal{F}_c^*}$;

14 **end while**

15 **return**(M^*, $\mathcal{F}^*$);

Table 3.1: Pseudocode of finding the maximum energy efficiency value M^* and an optimal forwarding candidate set $\mathcal{F}^*$ based on Lemma 2.2.4

Round 1		Round 2		Round 3	
$\mathcal{F}$	EPA($\mathcal{F}$)	$\mathcal{F}$	EPA($\mathcal{F}$)	$\mathcal{F}$	EPA($\mathcal{F}$)
$\langle i_1 \rangle$	9	$\langle i_1, i_2 \rangle$	14.6	$\langle i_1, i_2, i_4 \rangle$	18.17
$\langle i_2 \rangle$	8	$\langle i_1, i_3 \rangle$	14.25	$\langle i_1, i_3, i_4 \rangle$	17.225
$\langle i_3 \rangle$	7.5	$\langle i_1, i_4 \rangle^*$	14.95	$\langle i_1, i_4, i_5 \rangle$	15.4225
$\langle i_4 \rangle$	8.5	$\langle i_1, i_5 \rangle$	12.15		
$\langle i_5 \rangle$	4.5				
M(1)=6		$M^*(2) = 7.475$		M(3)=7.268	

Table 3.2: The procedure of finding the maximum energy efficiency value M^* and an optimal forwarding candidate set $\mathcal{F}^*$ by applying the algorithm GetM-A on the example in Fig. 3.1 with $E_{tx} = 1$ and $E_{rx} = 0.5$

Table 3.2 shows the procedure of finding the M^* and an $\mathcal{F}^*$ by applying the algorithm GetM-A on the example in Fig. 3.1 with $E_{tx} = 1$ and $E_{rx} = 0.5$. The procedure runs from Round 1 to Round 3, and in each round it runs from the top to the bottom. In the first round, $\langle i_1 \rangle$ is found as the node achieves the maximum EPA by selecting one forwarding candidate; in the second round, $\langle i_1, i_4 \rangle$ is found as the optimal node set by selecting two forwarding candidates; in the third round, $\langle i_1, i_2, i_4 \rangle$ is found as the optimal node set by selecting three forwarding candidates, and $M(3) < M(2)$; so searching is terminated, and $M(2)$ is the maximum energy efficiency value and $\langle i_1, i_4 \rangle$ is an optimal forwarding candidate set.

3.2.2.2 Dynamic programming algorithm

We now propose another efficient dynamic programming algorithm which is not based on the **Containing property**, and only costs $\mathbf{O}(M^2)$ in the worst case and $\Omega(M)$ in the best case.

Recall that nodes i_j's $(1 \leq j \leq M)$ in $\mathcal{C}$ are ordered according to the advancements

GetM-B($\mathcal{C}$,E_{tx},E_{rx})

1 **for** $i \leftarrow 1$ **to** N

2 EM$(\mathcal{C}_{i+1}, 0) \leftarrow 0$; $F_{(i+1,0)} \leftarrow \emptyset$; $P_{(i+1,0)} \leftarrow 0$;

3 **end for**

4 $M^* \leftarrow 0$; $\mathcal{F}^* \leftarrow \emptyset$; EM$(\mathcal{C}_{N+1}, 1) \leftarrow 0$; $P_{(N+1,1)} \leftarrow 0$;

5 **for** $r \leftarrow 1$ **to** N

6 **for** $q \leftarrow$ N-r+1 **down to** 1

7 $A \leftarrow d_q p_q + (1 - p_q)EM(\mathcal{C}_{q+1}, r - 1)$;

8 $P \leftarrow 1 - (1 - P_{(q+1,r-1)})(1 - p_q)$;

9 **if** $A >$ EM$(\mathcal{C}_{q+1}, r)$ || $(A =$ EM$(\mathcal{C}_{q+1}, r)$ & $P > P_{(q+1,r)})$

10 EM$(\mathcal{C}_q, r) \leftarrow A$; $F_{(q,r)} \leftarrow F_{(q+1,r-1)} \bigcup \{i_q\}$; $P_{(q,r)} \leftarrow P$;

11 **else**

12 EM$(\mathcal{C}_q, r) \leftarrow$ EM$(\mathcal{C}_{q+1}, r)$; $F_{(q,r)} \leftarrow F_{(q+1,r)}$;

13 $P_{(q,r)} \leftarrow P_{(q+1,r)}$;

14 **end for**

15 $M(r) \leftarrow$ EM$(\mathcal{C}_1, r)/(E_{tx} + r \cdot E_{rx})$;

16 **if** $M(r) \leq M^*$ /* Local maximum is found */

17 **return**(M^*, $\mathcal{F}^*$)

18 **else**

19 $M^* \leftarrow M(r)$; $\mathcal{F}^* \leftarrow F_{(1,r)}$;

20 **end for**

21 **return**(M^*, $\mathcal{F}^*$);

Table 3.3: Pseudocode of dynamic programming algorithm finding the maximum energy efficiency value M^* and an optimal forwarding candidate set $\mathcal{F}^*$

	Round 1		Round 2		Round 3	
q	$F_{(q,1)}$	EM($\mathcal{C}_q$, 1)	$F_{(q,2)}$	EM($\mathcal{C}_q$, 2)	$F_{(q,3)}$	EM($\mathcal{C}_q$, 3)
1	$\langle i_1 \rangle$	9	$\langle i_1, i_4 \rangle^*$	14.95	$\langle i_1, i_2, i_4 \rangle$	18.17
2	$\langle i_4 \rangle$	8.5	$\langle i_2, i_4 \rangle$	13.1	$\langle i_2, i_3, i_4 \rangle$	15.05
3	$\langle i_4 \rangle$	8.5	$\langle i_3, i_4 \rangle$	11.75	$\langle i_3, i_4, i_5 \rangle$	12.0875
4	$\langle i_4 \rangle$	8.5	$\langle i_4, i_5 \rangle$	9.175		
5	$\langle i_5 \rangle$	4.5				
	M(1)=6		$M^*(2) = 7.475$		M(3)=7.268	

Table 3.4: The procedure of finding the maximum energy efficiency value M^* and an optimal forwarding candidate set $\mathcal{F}^*$ by applying the algorithm GetM-B on the example in Fig. 3.1 with $E_{tx} = 1$ and $E_{rx} = 0.5$

as $d_1 > d_2 > ... > d_M$. Denote the set $\langle i_q, i_{q+1}, ..., i_M \rangle$ $(1 \leq q \leq M)$ as $\mathcal{C}_q$. Following the denoting, $\mathcal{C}_1 = \mathcal{C}$. According to the **Relay priority rule** in Theorem 2.2.3 and the definition of EM($\mathcal{C}, r$), we then have,

$$\text{EM}(\mathcal{C}_q, r) = \begin{cases} 0 & r = 0 \ or \ M - q + 1 < r; \\ Max\{d_q p_q + (1 - p_q)\text{EM}(\mathcal{C}_{q+1}, r-1), \text{EM}(\mathcal{C}_{q+1}, r)\} & \text{Otherwise.} \end{cases} \tag{3.7}$$

EM($\mathcal{C}, r$) $(1 \leq r \leq M)$ can be efficiently calculated by applying Eq. (3.7) recursively using dynamic programming [20].

The pseudocode of the dynamic programming algorithm GetM-B is given in Table 3.3, where $|\mathcal{C}| = M$, $F_{(q,r)}$ is the ordered node set corresponding to EM($\mathcal{C}_q, r$), $P_{(q,r)}$ is the corresponding one-hop reliability, and d_i's are sorted in descending order ($d_1 > d_2 ... > d_M$) (1≤k≤M). We also choose the feasible set that achieves higher one-hop reliability when two feasible sets have the same EPA (line 9). Based on Corollary 3.2.1, if a local maximum is found (line 16), the searching is terminated and the optimal solution is returned (line 17). The algorithm GetM-B costs $\mathbf{O}(M^2)$ running

time in the worst case and $\Omega(M)$ in the best case.

Table 3.4 shows the procedure of finding the M^* and an $\mathcal{F}^*$ by applying the algorithm GetM-B on the example in Fig. 3.1 with $E_{tx} = 1$ and $E_{rx} = 0.5$. The procedure runs from Round 1 to Round 3, and in each round it runs from the bottom to the top. It can be seen that although it finds the same M^* (M(2)) and $\mathcal{F}^*$ ($\langle i_1, i_4 \rangle$) as in Table 3.2, most of the tested node sets are different from the ones in Table 3.2.

3.3 Energy-efficient Geographic Opportunistic Routing (EGOR)

The EGOR that applies the local forwarding candidates selection algorithms GetM-A or GetM-B to get $\mathcal{F}^*$ is described in Table 3.5, where node i (i $\neq$ destination D) is routing a packet and the forwarding candidates i_j's are trying to relay the packet collaboratively. Here we do not consider any mechanism to route around voids (when $\mathcal{C} = \emptyset$). If the packet gets stuck due to no node being available for forwarding, it is dropped (line 14 in Procedure A). Mechanisms such as FACE routing [12] or perimeter forwarding in GPSR [36] can be applied here to deal with the communication void problem but it is beyond the scope of this paper. Retransmission limitation is applied. If the retransmission number (RN) reaches the limitation (RL) (line 11 in Procedure A), the packet will also be dropped (line 12 in Procedure A). It is worthy to mention that there is a last hop behavior (line 7 and 8 in Procedure A) in EGOR. When the sink D is in the available next-hop node set $\mathcal{C}$, we calculate the forwarding set by eliminating D from $\mathcal{C}$. Because sink D is not energy constrained, its receiving energy cost should not be counted when maximizing the energy efficiency. After calculating the $\mathcal{F}^*$, D should be added into the forwarding candidate set, since the packet always has a chance to reach D whatever the link quality from i to D is when D is the neighbor of i.

Procedure A

run by transmitter i:

1 $RN \leftarrow 0$

2 **while** $(RN \leq RL)$ **do**

3 **if** $\mathcal{C} \neq \emptyset$

4 **if** $D \notin \mathcal{C}$

5 Get $\mathcal{F}^*$ from $\mathcal{C}$, broadcast the packet to the nodes in $\mathcal{F}^*$.

6 **else**

7 Get $\mathcal{F}^*$ from $\mathcal{C} \setminus \{D\}$,

8 broadcast the packet to the nodes in $\mathcal{F}^* \bigcup D$.

9 **if** None of candidates received packet correctly

10 $RN \leftarrow RN + 1$

11 **if** $RN > RL$

12 Drop the packet

13 **else break**

14 **else**

15 Drop the packet

Procedure B

run by forwarding candidate i_j receiving the packet correctly:

1 **if** $(i_j \neq \mathrm{D})$

2 **if** No candidates having higher priorities received packet correctly

3 i_j becomes the actual forwarder and runs Procedure A

4 **else**

5 i_j drops the packet

6 **else** The packet is arriving at D and routing is terminated.

Table 3.5: The procedure of EGOR when node i is forwarding the packet

3.4 Performance evaluation

We evaluate the performance of EGOR through extensive Monte-Carlo simulations. We compare EGOR with the geographic routing which only has one forwarding candidate that achieves the maximum EPA, and the opportunistic routing which involves all the available next-hop nodes as forwarding candidates. When the packet gets stuck, all the three protocols just drop the packet. Various situations are simulated by varying node densities, transmission to reception energy ratios and retransmission limits.

3.4.1 Simulation Setup

Channel Model: To simulate a realistic channel model for lossy WSNs, we use the log-normal shadowing path loss model derived in [87]:

$$PRR(L_f, d) = (1 - \frac{1}{2} exp^{-\frac{\gamma(d)}{2\times 0.64}})^{8\rho L_f} \tag{3.8}$$

where d is the transmitter-receiver distance, $\gamma(d)$ is the Signal to Noise Ratio (SNR), ρ is the encoding ratio and L_f is the frame length in bytes. This model considers several environmental and radio parameters[3], such as the path-loss exponent (α) and log-normal shadowing variance (σ) of the environment, and the modulation and encoding schemes of the radio. This particular equation resembles a MICA2 mote [22], which has data rate of 19.2 kbps, and the noise bandwidth 30 kHz. Non-coherent FSK and Manchester are used as the modulation and encoding schemes (ρ=2), respectively. The environmental parameters are set to $\alpha = 3.5$ and $\sigma = 4$.

Energy Model: The energy consumption is obtained by multiplying the power consumption and the packet transmission time. The transmission power consumption P_{tx} is the summation of the power of power amplifier (P_{PA}) and electronic (P_{elec}),

[3]Please refer to [87] for a complete description of the model.

and P_{rx} is equal to P_{elec}. We assume that P_{PA} is proportional to the transmit power P_{trans}. Then P_{tx} is

$$P_{tx} = P_{PA} + P_{elec} = \frac{P_{trans}}{\eta} + P_{elec} \tag{3.9}$$

where η is the PA power efficiency which is set to be 0.3 in our simulation.

Evaluation Metrics: We define the following metrics to evaluate the performance of the three protocols.

- *Packet Delivery Ratio (PDR)*: percentage of packets sent by the source that actually reach the sink. This is a measure for reliability.

- *Energy efficiency* $\eta(S, D)$: this metric is measured in bit-meters per Joule (bmpJ). It is calculated as in Eq. (3.10),

$$\eta(S, D) = \frac{L_{pkt} \cdot N_r \cdot Dist(S, D)}{E_{total}} \tag{3.10}$$

  where N_r denotes the number of packets received at the destination, L_{pkt} is the packet length in bits, and E_{total} is the (transmission and reception) energy consumed by all the nodes involved in the routing procedure excluding the sink. We account for the distance factor, because the energy efficiency is indeed relevant to the distance between the communication pair due to the lossy property of multi-hop wireless links in WSNs.

- *Hop count*: it is measured as the number of hops a successfully delivered packet travels from source to destination.

The simulated sensor network has stationary nodes uniformly distributed in a 60×60 m^2 square region, with nodes having identical fixed transmission power of 0 dbm. The frame length is fixed on 50 bytes with preamble of 20 bytes. The source and the sink node are fixed at two corners across the diagonal of the square area.

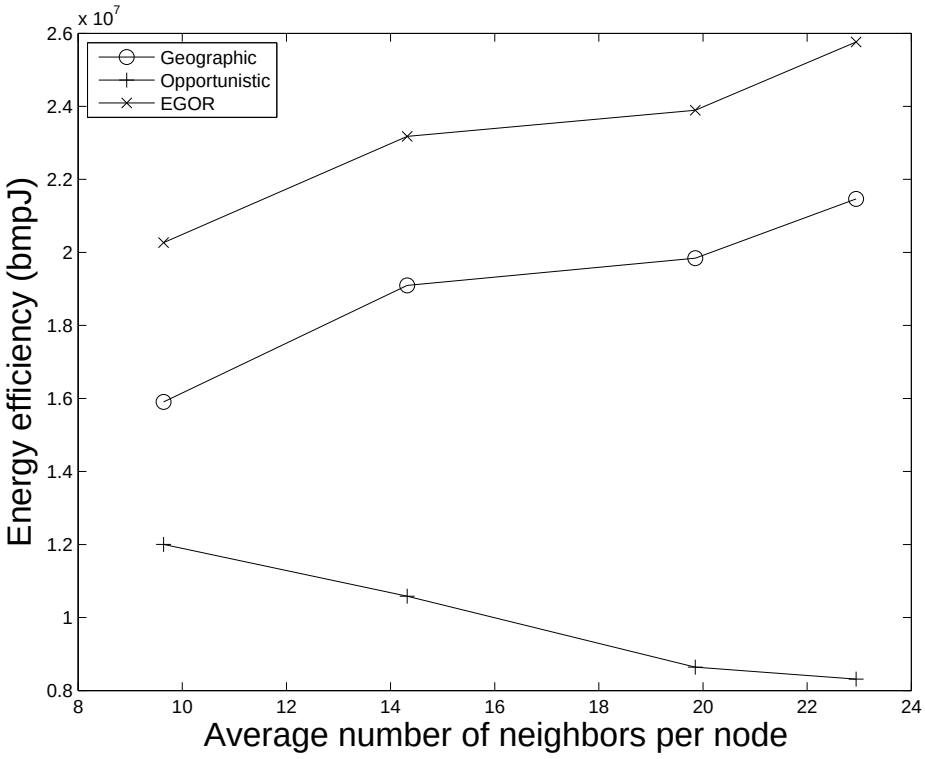


Figure 3.3: Energy efficiency vs network density

All simulations are run for 5000 iterations. For each iteration, node locations are randomly re-assigned and PRRs between nodes are re-calculated.

3.4.2 Simulation results and analysis

3.4.2.1 Impact of node density

We use different node numbers $(100, 144, 196, 225)$ to achieve various node densities corresponding to the average number of neighbors per node of $9.5, 14, 20, 23$ respectively. The reception power consumption is fixed on 2mw, so the reception to transmission energy ratio is $\frac{3}{8}$. There is no retransmission allowed.

Fig. 3.3 shows that EGOR achieves better energy efficiency than the other two routing protocols. This result can be explained as following: for every forwarding

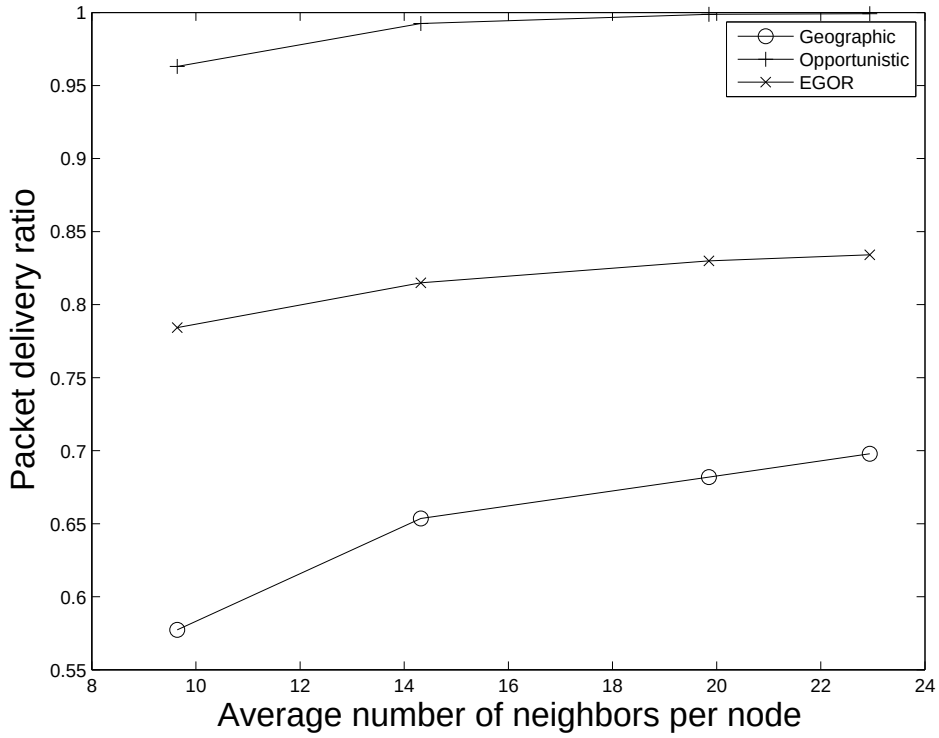


Figure 3.4: Packet delivery ratio vs network density

decision, EGOR chooses the forwarding set that maximizes the EPA per unit energy consumption. This local optimal behavior can achieve a good global performance under a uniformly randomly distributed node deployment, where any intermediate-hop forwarding can be viewed as a similar new first-hop packet forwarding. Statistically, every hop may make similar progress in such a homogeneous environment. The overall energy cost for successfully delivering one packet from source S to destination D can be approximated by $\frac{Dist(S,D)}{\text{EPA}/Local\ energy\ cost}$, so when we maximize the numerator, the total energy consumption is minimized. Opportunistic routing involving all the available next-hop nodes in the routing has the worst performance, since it has the lowest EPA/*Local energy cost* ratio.

Another observation from Fig. 3.3 is that the energy efficiency of EGOR and geographic routing is increased as the network becomes denser, while opportunistic

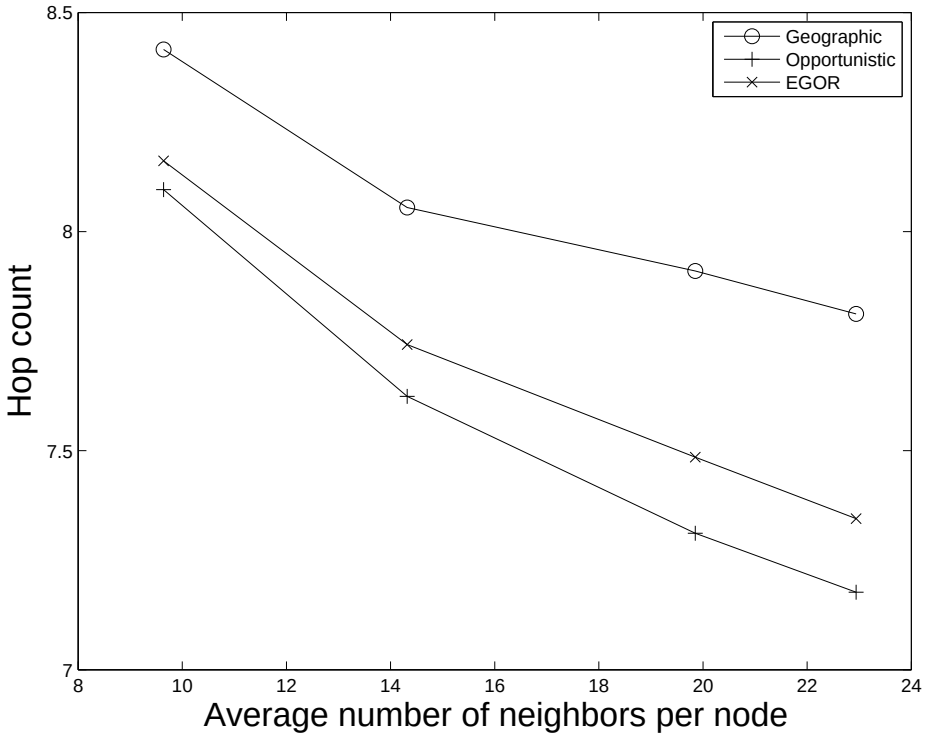


Figure 3.5: Hop count vs network density

routing shows the opposite trend. This result is related to the PDR performance shown in Fig. 3.4 and the hop count performance shown in Fig. 3.5. As we can see, the PDR of the opportunistic routing remains as high as nearly 1 under all the different node densities. That is to say, higher node density does not bring much gain to opportunistic routing on successfully delivering packets. Although the hop count of opportunistic routing is decreased when node density increases, the energy consumption due to unnecessarily involving more nodes in forwarding overwhelms the benefit of hop count decreasing. For geographic routing, the PDR is increased and hop count is also decreased when the network is denser, so the energy efficiency is increased. For EGOR, PDR is not increased much but hop count is decreased when node density increases, so the energy efficiency of EGOR is also increased.

Fig. 3.5 also shows that the hop counts decrease when network is denser for all

the three protocols, since involving more available next-hop nodes brings more chance for packets to make larger advancement when nodes are uniformly distributed. Opportunistic routing has the smallest hop count under all the node densities, while geographic routing has the largest hop count. Hop count of EGOR is between two of them. From the Theorem 2.2.2 in Section 3.1, we know that the maximum EPA by choosing r nodes from a given set is strictly increasing with r. So opportunistic routing archives largest EPA by selecting all the available next-hop nodes as forwarding candidates, geographic routing gets the smallest EPA by choosing only one forwarding candidate, and EGOR archives larger EPA than geographic routing and smaller EPA than opportunistic routing by selecting some (not all) available next-hop nodes as forwarding candidates. Actually, in this simulation, EGOR selects 1.2 nodes on average as the candidate forwarders under each node density. This observation suggests that under such settings, only a few nodes are necessary in order to take advantage of opportunistic routing efficiently. Adding one more node in forwarding can get much better energy efficiency and reliability than geographic routing. Involving all the available next-hop nodes in opportunistic routing is an energy wasteful method.

3.4.2.2 Impact of Reception to Transmission Energy Ratio (RTER)

We study the the performance of the three protocols under different RTERs in this section. In the simulations, no retransmission is allowed and the available next-hop node set size is 6.7 on average. The reception power consumption is varied from 10^{-3} to 10 mw, so the corresponding RTER is in range $[3 \times 10^{-4}, 0.75]$.

Fig. 3.6 shows the energy efficiency of EGOR is always the best of the three protocols and the RTER is a crucial parameter affecting the energy efficiency of the opportunistic routing. There is a watershed on RTER, smaller than which the energy efficiency of the opportunistic routing is better than the geographic routing, while

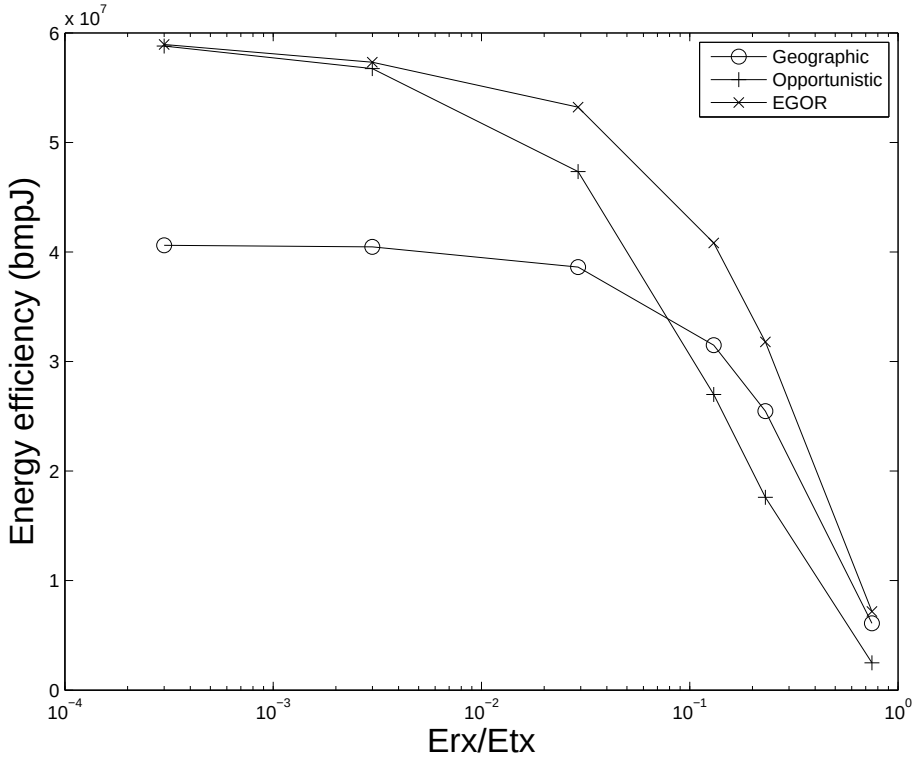


Figure 3.6: Energy efficiency vs reception to transmission power ratio

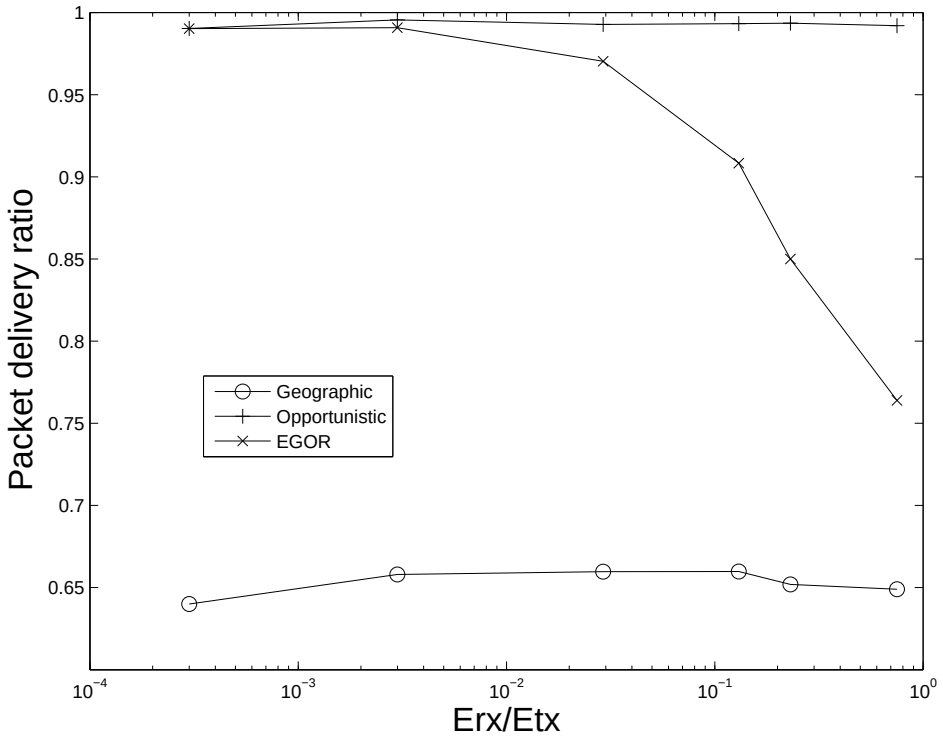


Figure 3.7: Packet delivery ratio vs reception to transmission power ratio

greater than which the geographic routing surpasses the opportunistic routing. The reason is that when the energy consumption of reception is negligible to the transmission, the opportunistic routing achieves larger EPA than the geographic routing while consumes nearly the same energy as the geographic routing. So opportunistic routing is more energy efficient. However, when the energy consumption of reception is comparable to transmission, involving all the available next-hop nodes in the opportunistic routing consumes much more energy than the geographic routing, and the cost of the increased energy consumption overwhelms the benefit of the increased EPA. Thus, the energy efficiency of the opportunistic routing is less than the geographic routing when RTER is greater than the watershed. For these two protocols, RTER does not affect the forwarding candidate(s) selecting criteria, so it does not affect the PDR.

Fig. 3.7 shows the results that the PDR of the geographic routing and opportunistic routing does not change according to the RTER. For EGOR, the PDR decreases when RTER increases, because EGOR takes the energy consumption into account. When the reception energy cost increases, fewer nodes are selected as forwarding candidates, then the packet is more likely to be lost without retransmission. An interesting observation here is that, even when RTER is very small, EGOR only selects a very small number of nodes as the forwarding candidate, but achieves nearly the same energy efficiency as opportunistic routing. For example, when RTER = 0.03%, EGOR only selects 2.2 forwarding candidates on average, while has the same energy efficiency and PDR as the opportunistic routing which selects 6.7 candidates on average. This result again suggests that only a small number of nodes need to be involved in opportunistic routing to achieve a good balance between energy efficiency and routing efficiency.

The hop count performance shown in Fig. 3.8 indicates that the RTER does not affect the hop count of the opportunistic routing and geographic routing, the reason is as the same as the PDR performance of these two protocols. For EGOR, the hop count increases after the RTER is larger than 10% because fewer forwarding candidates are selected and the EPA is decreased.

3.4.2.3 Impact of retransmission limit

In this section we study how the retransmission limit affects the performance of the three protocols. The reception power consumption is fixed on 2mw and the the available next-hop node set size is also 6.7 on average .

Intuitively, increasing retransmission limit will increase the reliability, say PDR. Fig. 3.9 exactly shows this trend for all the three protocols. It is worthy to mention that the benefit of increasing retransmission limit (can be seen as the slopes of the curves) for the opportunistic routing is trivial but for the geographic routing is ob-

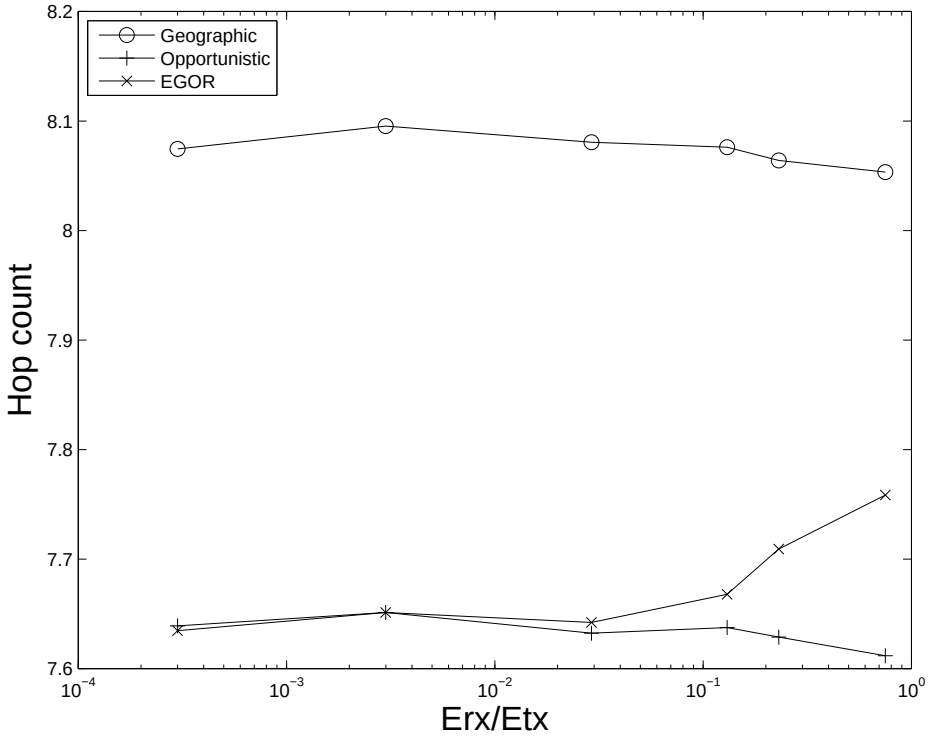


Figure 3.8: Hop count vs reception to transmission power ratio

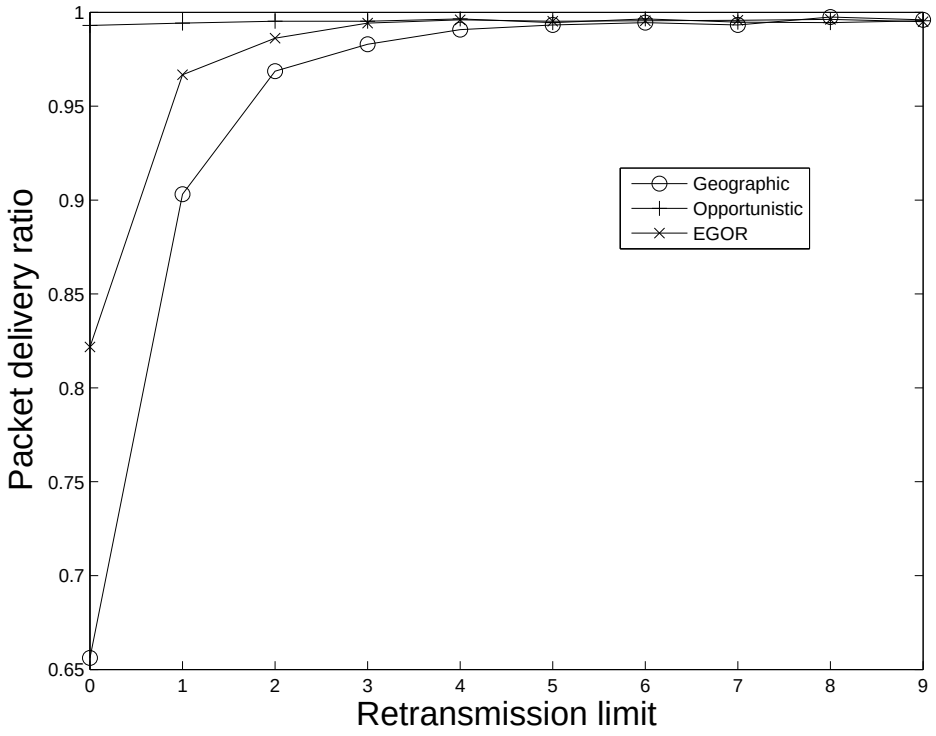


Figure 3.9: Packet delivery ratio vs retransmission limit

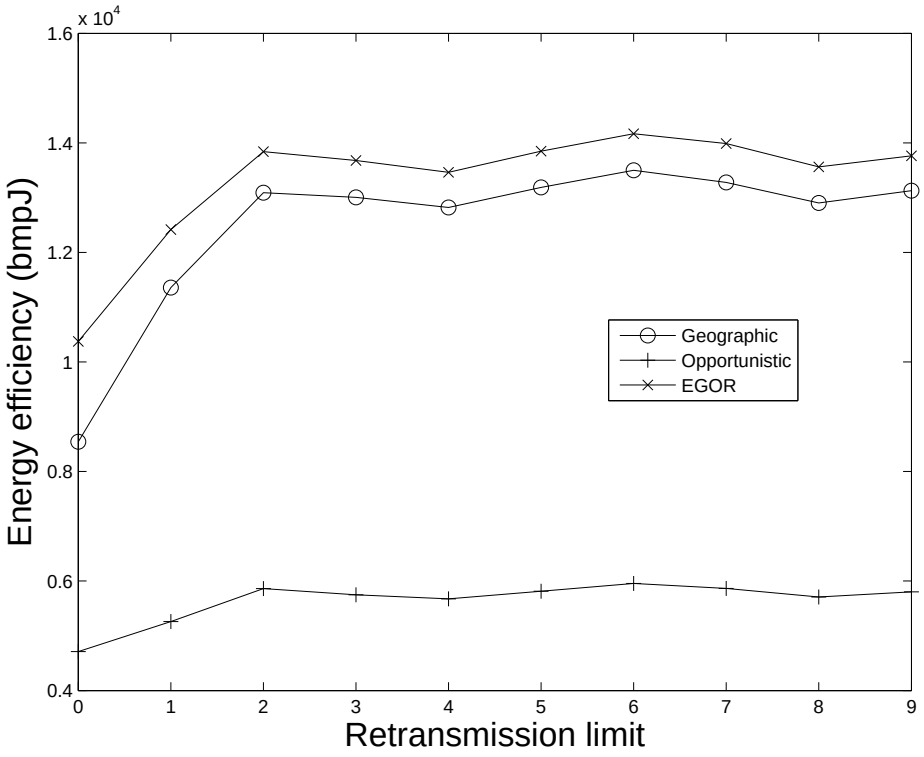


Figure 3.10: Energy efficiency vs retransmission limit

vious (especially when retransmission limit is less than 4). The reason is that the opportunistic routing has already achieves high PDR (nearly 1) by involving all the available next-hop nodes in forwarding even when there is no retransmissions allowed. For the geographic routing, however, there is only one next-hop node involving in the forwarding, then the packet is more likely to be lost in one transmission than the opportunistic routing. For EGOR, the PDR increasing rate is less than that of the geographic routing because EGOR already achieves higher PDR than the geographic routing when there is no retransmissions allowed. When retransmission limit is larger than 1, the PDR gains become less and less for both EGOR and the geographic routing, and when the limit is larger than 3, the PDRs of both are approaching to 1.

Fig. 3.10 shows that for opportunistic routing, the energy efficiency is not changed much according to the change of retransmission limit. The reason is that the retrans-

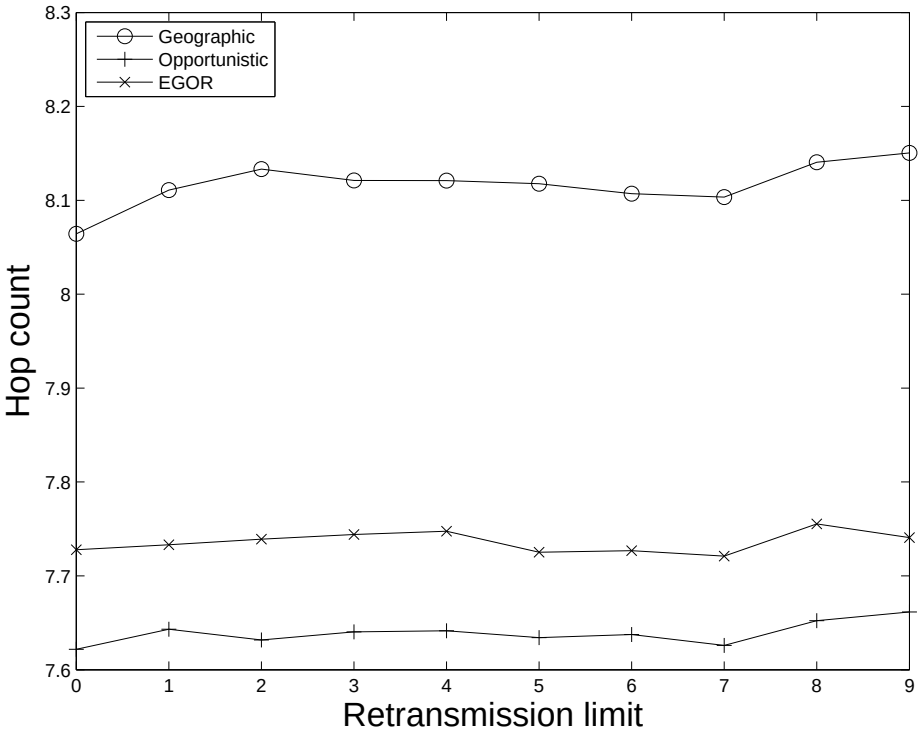


Figure 3.11: Hop count vs retransmission limit

mission does not play a role for the PDR in opportunistic routing, and almost the same packets can be delivered to the destination whether retransmission is allowed or not. For geographic routing and EGOR, the retransmission does play a role for the energy efficiency when retransmission limit is less than 3. As we have analyzed retransmission affects the PDR, especially from allowing no retransmission to one and from one to two. When retransmission limit is larger than 3, the energy efficiency of these two protocols does not change much as the PDRs are already approaching to 1.

As retransmission does not affect EPA much, Fig. 3.11 shows that the hop count remains almost the same when the retransmission limit varies for each of the three protocols.

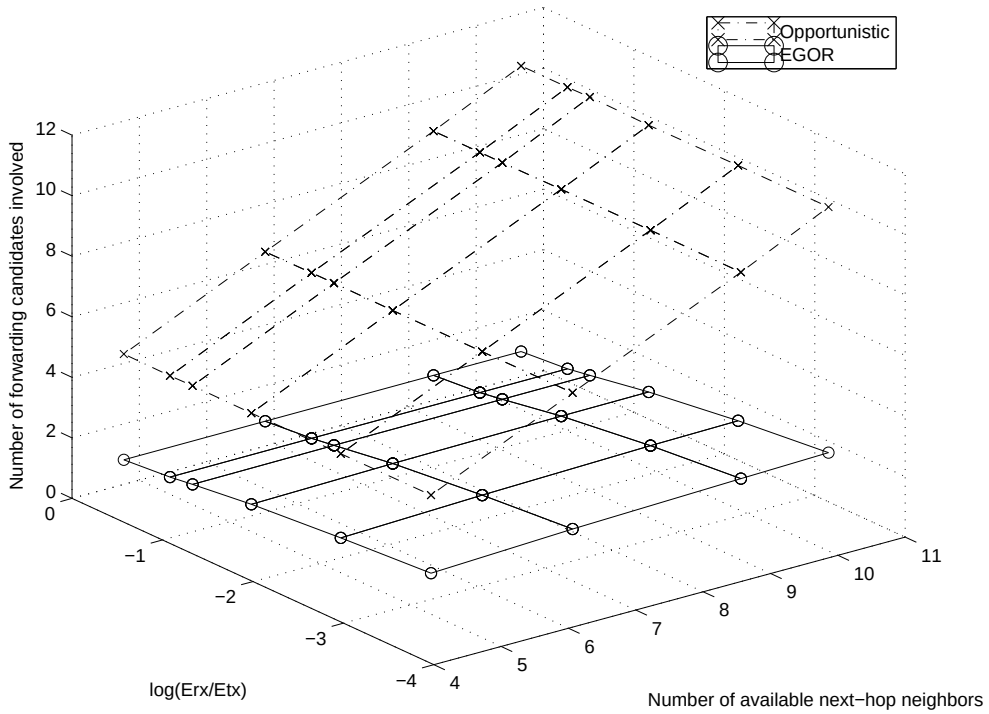


Figure 3.12: Number of forwarding candidates involved under different node densities and RTERs

3.4.2.4 Number of local forwarding candidates involved

We study the number of local forwarding candidates involved for opportunistic routing and EGOR under various node densities and RTERs. In the simulations, we use different node numbers (100,144,196,225) to achieve various node densities corresponding to the available next-hop candidate set sizes of $4.6, 6.7, 9.2, 10.5$ on average respectively. The RTER is in range $[3 \times 10^{-4}$,0.75].

Fig. 3.12 shows the simulation result. The opportunistic routing uses all the available next-hop nodes as the forwarding candidates, while EGOR only uses a very small number of forwarding candidates (around 2 or fewer). For example, even when the RTER is as small as 0.03%, EGOR only chooses 2.2 forwarding candidates on average under various node densities. This means even when the energy consumption

of reception is far less than that of transmission, in order to achieve maximum energy efficiency, we still only need to involve 2 forwarding candidates. Fig. 3.12 also shows that the number of forwarding candidates is affected mainly by the RTER but not by the node density under a uniform node distribution. For example when RTER=0.75, the forwarding candidates in EGOR are nearly unchanged as 1.1 under different node densities. This is an important result as it indicates that involving more forwarding candidates will not bring much more expected packet advancement. Only a small number of forwarding candidates are sufficient to strike a good balance between EPA and energy consumption. This is a very desirable result, because the cost incurred due to assuring one final forwarder from multiple forwarding candidates at the MAC layer are expected to grow when involving more forwarding candidates [58]. Involving fewer candidates introduces less rendezvous and contention cost.

3.4.2.5 Concavity of maximum EPA and its slope

In this section, we study the concavity of the maximum EPA function and its slope in one hop under various node densities. The nodes are uniformly distributed and the next-hop available node number is various from 6 to 12. From Fig. 3.13 we can see that the maximum EPA increases when the number of the forwarding candidates increases, and when nodes are denser, the EPA are larger. A very interesting result is that under different node densities, the slopes of each curve in Fig. 3.13 are nearly the same, which is shown in Fig. 3.14. Notice that, when the forwarding candidate number is 3, the slope is already decreased to below 0.01. When the number of forwarding candidates is larger than 4, the slope is near to zero. Fig. 3.13 and 3.14 manifest that no matter what the node density is, the EPA gain of involving more forwarding candidates becomes very small when the number of forwarding candidates is larger than 3. These results are consistent with Fig. 3.12 where the optimal energy efficiency is achieved when the number of forwarding candidates is around 2.2.

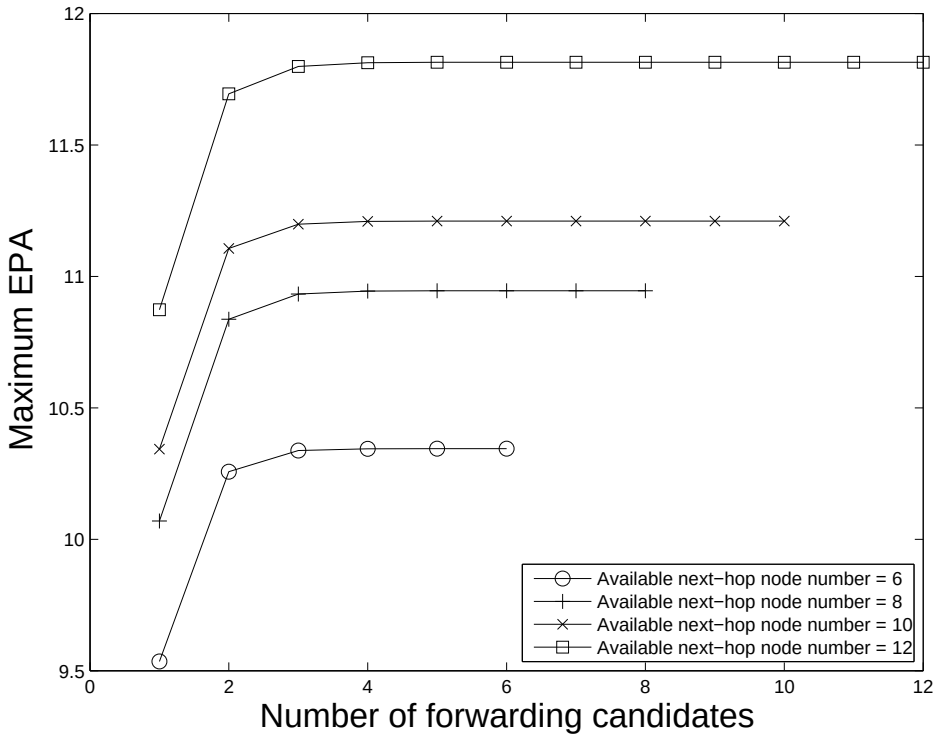


Figure 3.13: Maximum EPA vs forwarding candidate number under different node densities

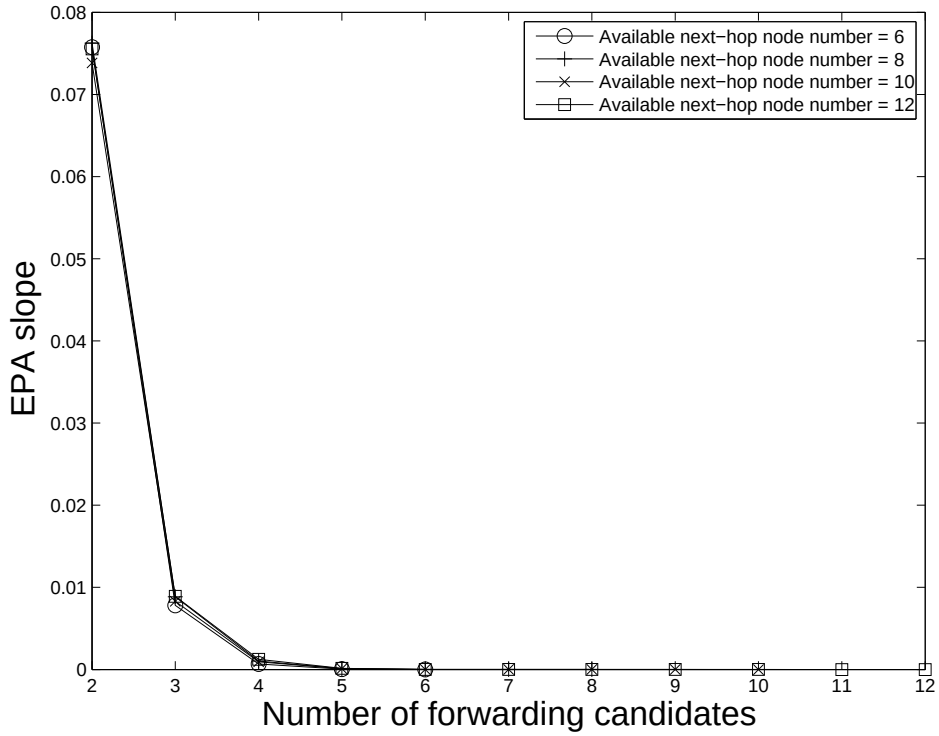


Figure 3.14: EPA slope under different node densities

3.5 Conclusion

In this chapter, we studied the geographic opportunistic routing strategy with both routing and energy efficiencies as the major concerns. We propose a new routing metric which evaluates EPA per unit of energy consumption so that the energy efficiency can be taken into consideration in routing. By leveraging the proved findings in Chapter 2, we propose two localized candidate selection algorithms with $\mathbf{O}(M^3)$ and $\mathbf{O}(M^2)$ running time in the worst case, respectively, and $\Omega(M)$ in the best case, where M is the number of available next-hop neighbors of the transmitter. The algorithms efficiently determine the forwarding candidate set that maximizes the proposed new metric for energy efficiency, namely, the EPA per unit of energy consumption. We further propose an EGOR framework applying the node selection algorithms to achieve the energy efficiency. The performance of EGOR is studied through extensive simulations and compared with those of the existing geographic routing and opportunistic routing protocols. The results show that EGOR achieves the best energy efficiency among the three protocols in all the cases while maintaining very good routing performance. Our simulation results also show that the number of forwarding candidates necessary to achieve the maximum energy efficiency is mainly affected by the reception to transmission energy ratio but not by the node density under a uniform node distribution. Although the EPA can be maximized by involving the most number of nodes in GOR, in terms of energy efficiency, only a very small number of forwarding candidates (around 2) are needed on average. This is true even when the energy consumption of reception is far less than that of transmission.

Chapter 4

End-to-end Throughput Bounds given Opportunistic Routing

The existing works on OR mainly focused on a single-rate system. Researchers have proposed several candidate selection and prioritization schemes to improve throughput or energy efficiency. However, there is a lack of theoretical analysis on the performance limit or the throughput bounds achievable by OR. In addition, one of the current trends in wireless communication is to enable devices to operate using multiple transmission rates. For example, many existing wireless networking standards such as IEEE 802.11a/b/g include this multi-rate capability. The inherent rate-distance trade-off of multi-rate transmissions has shown its impact on the throughput performance of traditional routing [8, 75, 74]. Generally, low-rate communication covers a long transmission range, while high-rate communication must occur at short range. It is intuitive to expect that this rate-distance trade-off will also affect the throughput of OR. Because different transmission ranges also imply different neighboring node sets, which results in different spacial diversity opportunities. These rate-distance-diversity trade-offs will no doubt affect the throughput of OR, which deserves a careful study. To the best of our knowledge, there is no existing work addressing the throughput

problem of OR in a multi-rate network.

In this chapter, we bridge these two gaps by studying the throughput bound of OR and the performance of OR in a multi-rate scenario. First, for OR, we propose the concept of concurrent transmission sets which captures the transmission conflict constraints of OR. Then, for a given network with given opportunistic routing strategy (i.e., forwarder selection and prioritization), we formulate the maximum end-to-end throughput problem as a maximum-flow linear programming problem subject to the constraints of transmitter conflict. The solution of the optimization problem provides the performance bound of OR. The proposed method establishes a theoretical foundation for the evaluation of the performance of different variants of OR with various forwarding candidate selection, prioritization policies, and transmission rates. We also propose two OR metrics: *expected medium time* (EMT) and *expected advancement rate* (EAR), and the corresponding distributed and local rate and candidate set selection schemes, one of which is Least Medium Time OR (LMTOR) and the other is Multi-rate Geographic OR (MGOR). Simulation results show that for OR, by incorporating our proposed multi-rate OR schemes, systems operating at multi-rates achieves higher throughput than systems operating at any single rate. Several insights of OR are observed: 1) the end-to-end capacity gained decreases when the number of forwarding candidates is increased; 2) there exists a node density threshold, higher than which 24Mbps GOR performs better than 12Mbps GOR, and lower than which, vice versa.

The rest of this chapter is organized as follows. We propose the framework of computing the throughput bounds of OR in Section 4.1. Section 4.2 studies the impact of multi-rate capability and forwarding strategy on the throughput bound of OR. We then propose the OR metrics, and rate and candidate selection schemes for multi-rate systems in Section 4.3. Simulation results are presented and analyzed in Section 4.4. Conclusions are drawn in Section 4.5.

4.1 Computing Throughput Bound of OR

The first fundamental issue to address is the maximum end-to-end throughput when OR is used. Any traffic load higher than the throughput capacity is not supported and even deteriorates the performance as a result of excessive medium contention. The knowledge of throughput capacity can be used to reject any excessive traffic in the admission control for real-time services. It can also be used to evaluate the performance of different OR variants. Furthermore, the derivation of throughput of OR may suggest novel and efficient candidate selection and prioritization schemes.

In this section we present our methodology to compute the throughput bound between two end nodes in a given network with a given OR strategy (i.e., given each node's forwarding candidate set, node relay priority, and transmission/broadcast rate at each node). We first introduce two concepts, transmitter based conflict graph and concurrent transmission set, which are used to represent the constraints imposed by the interference among wireless transmissions in a multi-hop wireless network. We then present methods for computing bounds on the optimal throughput that a network can support when OR is used.

4.1.1 Transmission Interference and Conflict

Wireless interference is a key issue affecting throughput. Existing wireless interference models generally fall into two categories: *protocol model* and *physical model* [31]. Under the protocol model, a transmission is considered successful when both of the following conditions hold: 1) The receiver is in the effective transmission range of the transmitter; and 2) No other node that is in the carrier sensing range of the receiver is transmitting. This kind of protocol model requires only the receiver to be free of interference. To model a 802.11 like bidirectional communications, we can extend the protocol model by adding the requirement of interference free also at the transmitter

side. Under the physical model, for a successful transmission, the aggregate power at the receiver from all other ongoing transmissions plus the noise power must be less than a certain threshold so that the SNR requirement at the ongoing receiver is satisfied. In this paper, we use the term "**usable**" to describe a link when it is able to make a successful transmission based on either the protocol model or the physical model. When two (or more) links are not able to be usable at the same time, they are having a "**conflict**".

Link conflict graphs have been used to model such interference [32, 74]. As shown in Fig. 4.2(b), in a link conflict graph, each vertex corresponds to a link in the original connectivity graph. There is an edge between two vertices if the corresponding two links may not be active simultaneously due to interference (e.g., having a "conflict"). However, this link-based conflict graph cannot be directly applied to study capacity problem of OR networks because by the nature of opportunistic routing, for one transmission, throughput may take place on any one of the links from the transmitter to its forwarding candidates. The throughput dependency among multiple outgoing links from the same transmitter makes the subsequent maximum-flow optimization problem very difficult (if it is still possible). Therefore, in this paper, we propose a new construction of conflict graph to facilitate the computation of throughput bounds of OR. Instead of creating link conflict graph, we study the conflict relationship by transmitters (or nodes) associated with their forwarding candidates. As shown in Fig. 4.2(c), in the node conflict graph, each vertex corresponds to a node in the original connectivity graph. Each vertex is associated with a set of links, e.g., the links to its selected forwarding candidates. There is an edge (conflict) between two vertices if the two nodes cannot be transmitting simultaneously due to a conflict caused by one or more unusable links as we will define in Section 4.1.2.

4.1.2 Concurrent Transmission Sets

We define the concepts of **concurrent transmission sets** (**CTS**'s) for OR as follows. These concepts capture the impact of interference of wireless transmissions and OR's opportunistic nature. They are the foundation of our method of computing the end-to-end throughput.

1) **Conservative CTS:** According to a specific OR policy, when one node is transmitting, the packet is broadcast to all the nodes in its forwarding candidate set. The links from a transmitter to all its forwarding candidates are defined as links associated with the transmitter. We define a conservative CTS (CCTS) as a set of transmitters, when all of them are transmitting simultaneously, all links associated with them are still usable. If adding any one more node into a CCTS will result in a non-CCTS, the CCTS is called a maximum CCTS.

The conservative CTS actually requires all the opportunistic receivers to be interference-free for one transmission. This is probably true for certain protocols [28] where RTF (Request To Forward) and CTF (Clear To Forward) control packets are used to clear certain ranges within transmitter and forwarding candidates or confirm a successful reception. But this is a stricter requirement than necessary and will only give us a lower bound of end-to-end capacity. We define the following greedy CTS to compute the maximum end-to-end throughput.

2) **Greedy CTS:** In order to maximize the throughput, we permit two or more transmitters to transmit at the same time even when some links associated with them become unusable. The idea is to allow a transmitter to transmit as long as it can deliver some throughput to one of the next-hop forwarder(s). Therefore, we define a greedy CTS as a set of transmitters, when all of them are transmitting simultaneously, at least one link associated with each transmitter is usable. If adding any one more node into a GCTS will result in changes in the usability status of any link associated with nodes in that set, the GCTS is called a maximum GCTS.

4.1.3 Effective Forwarding Rate

After we find a CTS, we need to identify the capacity on every link associated with a node in the CTS. We introduce the concept of **effective forwarding rate** on each link associated with a transmitter according to a specified OR strategy. Assume node n_i's forwarding candidate set $\mathcal{F}_i = \langle n_{i_1}, n_{i_2}...n_{i_r} \rangle$, with relay priorities $n_{i_1} > n_{i_2} > ... > n_{i_r}$. Let ψ_q denote the indicator function on link l_{ii_q} when it is in a particular CTS: $\psi_q = 1$ indicating link l_{ii_q} is usable, and $\psi_q = 0$ indicating that link l_{ii_q} is not usable. Then the effective forwarding rate of link l_{ii_q} in that particular CTS is defined in Eq. (4.1):

$$\widetilde{R}_{ii_q} = R_i \cdot \psi_q \cdot p_{ii_q} \prod_{k=0}^{q-1} (1 - \psi_k \cdot p_{ii_k}) \tag{4.1}$$

where R_i is the broadcast rate of transmitter i, and $p_{ii_0} := 0$. $p_{ii_q} \prod_{k=0}^{q-1}(1 - \psi_k \cdot p_{ii_k})$ is the probability of candidate n_{i_q} receiving the packet correctly but all the higher-priority candidates not. Note that the candidate (with $\psi_q = 0$), which is interfered by other transmissions, is not involved in the opportunistic forwarding, and has no effect on the effective forwarding rate from the transmitter to lower-priority candidates, as $(1 - \psi_k \cdot p_{ii_k}) = 1$.

In a conservative CTS, all the receptions are interference-free. Therefore, in each CCTS, every link associated with a transmitter is usable, i.e. $\psi = 1$, and the effective forwarding rate on each link is non-zero. And the effective forwarding rate for a particular link remains same when the link is in a different CCTS. The effective forwarding rate indicates that according to the relay priority, only when a usable higher forwarding candidate did not receive the packet correctly, a usable lower priority candidate may have a chance to relay the packet if it received the packet correctly. Note that this definition generalizes the effective rate for unicast in traditional routing, that is, when there is only one forwarding candidate, the effective forwarding rate

reduces to the unicast effective data rate.

While for the greedy mode, some link(s) associated with one transmitter may become unusable, thus having zero effective forwarding rate. Furthermore, the effective forwarding rate on the links may be different when they are in different GCTS's. To indicate this possible difference, we use $\widetilde{R}^{\alpha}_{ii_w}$ to denote the effective forwarding rate of link l_{ii_w} when it is in the α^{th} GCTS.

4.1.4 Lower Bound of End-to-End Throughput of OR

Assume we have found all the maximum CCTS's $\{T_1, T_2...T_M\}$ in the network. At any time, at most one CTS can be scheduled to transmit. When one CTS is scheduled to transmit, all the nodes in that set can transmit simultaneously. Let λ_α denote the time fraction scheduled to CCTS T_α $(1 \leq \alpha \leq M)$. Then the maximum throughput problem can be converted to an optimal scheduling problem that schedules the transmission of the maximum CTS's to maximize the end-to-end throughput. Therefore, considering communication between a single source, n_s, and a single destination, n_d, with opportunistic routing, we formulate the maximum achievable throughput problem between the source and the destination as a linear programming problem corresponding to a maximum-flow problem under additional constraints in Fig. 4.1.

In Fig. 4.1, f_{ij} denotes the amount of flow on link l_{ij}, $\mathbf{E}$ is a set of all links in the connected graph G, and $\mathbf{V}$ is the set of all nodes. The maximization states that we wish to maximize the sum of flow out of the source. The constraint (4.2) represents flow-conservation, i.e., at each node, except the source and the destination, the amount of incoming flow is equal to the amount of outgoing flow. The constraint (4.3) states that the incoming flow to the source node is 0. The constraint (4.4) indicates that the outgoing flow from the destination node is 0. The constraint (4.5) restricts the amount of flow on each link to be non-negative. The constraint (4.6) says there is no flow from the node to the neighboring nodes that are not selected as the

$$Max \sum_{l_{si} \in \mathbf{E}} f_{si}$$

$$s.t.$$

$$\sum_{l_{ij} \in \mathbf{E}} f_{ij} = \sum_{l_{ji} \in \mathbf{E}} f_{ji} \quad \forall\ n_i \in \mathbf{V} - \{n_s, n_d\} \tag{4.2}$$

$$\sum_{l_{is} \in \mathbf{E}} f_{is} = 0 \tag{4.3}$$

$$\sum_{l_{di} \in \mathbf{E}} f_{di} = 0 \tag{4.4}$$

$$f_{ij} \geq 0 \quad \forall\ l_{ij} \in \mathbf{E} \tag{4.5}$$

$$f_{ij} = 0 \quad \forall\ l_{ij} \in \mathbf{E}, n_j \notin \mathcal{F}_i \tag{4.6}$$

$$\sum_{\alpha=1}^{M} \lambda_\alpha \leq 1 \tag{4.7}$$

$$\lambda_\alpha \geq 0,\ 1 \leq \alpha \leq M \tag{4.8}$$

$$f_{ij} \leq \sum_{n_i \in T_\alpha,\ n_j \in \mathcal{F}_i,\ 1 \leq \alpha \leq M} \lambda_\alpha \widetilde{R}_{ij}^{\alpha} \quad \forall\ l_{ij} \in \mathbf{E} \tag{4.9}$$

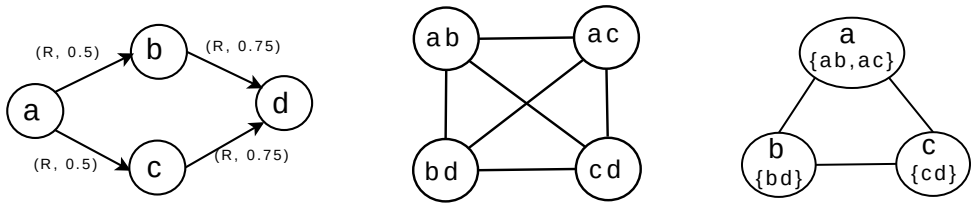
Figure 4.1: LP formulations to optimize the end-to-end throughput of OR

forwarding candidates of it. The constraint (4.7) represents at any time, at most one CTS will be scheduled to transmit. The constraint (4.8) indicates the scheduled time fraction should be non-negative. The constraint (4.9) states the actual flow delivered on each link is constrained by the total amount of flow that can be delivered in all activity periods of the OR modules which contain this link.

The key difference of our maximum flow formulations from the formulations for traditional routing in [32, 74] lies in the methodology we use to schedule concurrent transmissions. With the construction of concurrent transmission sets, we are able to schedule the transmissions based on node set (with each node associated with a set of forwarding candidates) rather than link set in traditional routing. When we schedule a transmitter, we effectively schedule the links from the transmitter to its forwarding candidates at the same time according to OR strategy. While for traditional routing, any two links sharing the same sender cannot be scheduled simultaneously. When a packet is not correctly received by the intended receiver but opportunistically received by some neighboring nodes of the sender, traditional routing will retransmit that packet instead of making use of the correct receptions on the other links. OR takes advantage of the correct receptions. That's why OR achieves higher throughput than traditional routing. Our proposed model accurately captures OR's capability of delivering throughput opportunistically.

A Simple Example: Next, we give an example to show how our formulation helps us to find the end-to-end throughput bound of OR, and we compare this result with the maximum throughput derived from multipath traditional routing based on results in [32].

For simplicity, in the four node network shown in Fig. 4.2(a), we assume each node transmits at the same rate R, and each link is associated with a PRR indicated in the pair on each link. Assume every node is in the carrier sensing range of any other nodes. We are going to find the maximum end-to-end throughput from node a



(a) Original graph

(b) Link conflict graph for traditional routing

(c) Transmitter conflict graph for OR

Figure 4.2: Conflict Graph

to d for traditional routing and OR.

For traditional routing, we first construct the link conflict graph as shown in Fig. 4.2(b). In the conflict graph, each vertex corresponds to each link in the original connectivity graph. There is an edge between two vertices when these two links conflict with each other. According to the protocol model, any two links in Fig. 4.2(a) cannot be scheduled simultaneously. So the link conflict graph for traditional routing is a complete graph (clique). There are four independent sets each containing one node in the conflict graph. Each independent set corresponds to one concurrent schedulable link set. By running the linear programming formulated in [32], we can find an optimal schedule on links to maximize the throughput. Assuming the whole communication period is τ, one feasible solution is assigning $\frac{3}{10}\tau$, $\frac{3}{10}\tau$, $\frac{2}{10}\tau$, $\frac{2}{10}\tau$ to l_{ab}, l_{ac}, l_{bd}, l_{cd}, respectively. So the maximum end-to-end throughput between a and d is $\frac{2(\frac{3}{10}R \cdot 0.5\tau)}{\tau} = \frac{3}{10}R$ for the traditional routing.

For OR, we construct the node conflict graph. Assume a chooses nodes b and c as its forwarding candidates, and b and c's forwarding candidate is just the destination d. According to the protocol model, the node conflict graph is constructed in Fig. 4.2(c), which only contains three vertices and is also a clique. So the three conservative transmission sets are $T_1 = \{a\}$, $T_2 = \{b\}$, and $T_3 = \{c\}$. Assume node b has higher

relay priority than node c, then we have $\widetilde{R}^1_{ab} = 0.5R$, $\widetilde{R}^1_{ac} = 0.25R$, and $\widetilde{R}^2_{bd} = \widetilde{R}^3_{cd} = 0.75R$. By running the linear programming formulated in Fig. 4.1, we get an optimal schedule that assigns $\frac{1}{2}\tau$, $\frac{1}{3}\tau$ and $\frac{1}{6}\tau$, to nodes a, b and c respectively. So the throughput of OR under this optimal schedule is $\frac{0.25R\tau+0.125R\tau}{\tau} = \frac{3}{8}R$, which is 25% higher than that of the traditional routing.

4.1.5 Maximum End-to-end Throughput of OR

The throughput bound we find based on the maximum conservative CTS's in section 5.3.6 is a lower bound of maximum end-to-end throughput. The CCTS's can be constructed based on either the protocol model or the physical model. However, the interference freedom at every intended receiver is a stricter requirement than necessary. It may be applicable under some protocol scenario but it fails to take full advantage of opportunistic nature of OR, because it excludes the situations where concurrent transmission is able to deliver throughput on some of the links even though some other links are having conflicts. In order to compute the exact capacity, we apply the same optimization technique to the greedy CTS's. Since greedy CTS's include all the possible concurrent transmission scenarios that generate non-zero throughput, the bound found by the optimization technique based on all greedy CTS's will be the maximum end-to-end throughput of OR.

Similar to the construction of CCTS's, GCTS's can be constructed based on either the protocol model or the physical model. Under the protocol model, the conflict between two links is binary, either conflict or no conflict. It is not difficult to construct the GCTS's under the protocol model with the proposed node conflict graph. On the other hand, it is well known that the physical model captures the interference property more accurately. However, it is more complicated to represent the interference when multiple transmitters are active at the same time. In this section, we discuss the construction of GCTS's based on the physical interference model.

Under the physical interference model, a link l_{ij}, from node n_i to n_j, is usable if and only if the signal to noise ratio at receiver n_j is no less than a certain threshold, e.g., $\frac{Pr_{ij}}{P_N} \geq SNR_{th}$, where Pr_{ij} is the average signal power received at n_j from n_i's transmission, P_N is the interference+noise power, and SNR_{th} is the SNR threshold, under which the packet can not be correctly received and above which the packet can be received at least with probability p_{td}. Note that, SNR_{th} is different for different data rates.

Under the physical model, the interference gradually increases as the number of concurrent transmitters increases, and becomes intolerable when the interference+noise level reaches a threshold. We define a weight function w_{ij_q}, to capture the impact of a transmitter n_i's transmission on a link l_{jj_q}'s reception. Link l_{jj_q} represents the data forwarding from node n_j to one of its forwarding candidate n_{j_q}.

$$w_{ij_q} = \frac{Pr_{ij_q}}{\frac{Pr_{jj_q}}{SNR_{th}} - P_{noise}} \tag{4.10}$$

where Pr_{ij_q} and Pr_{jj_q} are the received power at node n_{j_q} from the transmissions of nodes n_i and n_j, respectively, P_{noise} is the ambient noise power, and $\frac{Pr_{jj_q}}{SNR_{th}} - P_{noise}$ is the maximum allowable interference at node n_{j_q} for keeping link l_{jj_q} usable.

Then given a transmission set S and $n_j \in S$, a link l_{jj_q} is usable if and only if $\sum_{n_i \in S, i \neq j} w_{ij_q} < 1$. It means that link l_{jj_q} is usable even when all the transmitters in set S are simultaneously transmitting. For conservative mode, if this condition is true for every link associated with each transmitter in S, this set S is a CCTS. For greedy mode, if this condition is true for at least one link associated with each transmitter in S, the set S is a GCTS.

After finding all the GCTS's, we can apply the same optimization technique to the maximum flow problem based on all the GCTS's. The result is the exact bound of maximum end-to-end throughput.

When each node has only one forwarding candidate, OR degenerates to the tra-

ditional routing. Therefore, finding all the concurrent transmission sets is at least as hard as the NP-hard problem of finding the independent sets in [32, 74] for traditional routing. However, it may not be necessary to find all of them to maximize an end-to-end throughput. Some heuristic algorithm similar to that in [61], or column generation technique [76] can be applied to find a good subset of all the CTS's. In addition, complexity can be further reduced by taking into consideration that interferences/conflicts always happen for nodes within a certain range. How to efficiently find all the CTS's is out of the scope of this paper. We simply apply a greedy algorithm to find all the CTS's, say each time we add new transmitters into the existing CTS's to create new CTS's, until no any additional transmitter can be added into any of the existing CTS's.

4.1.6 Multi-flow Generalization

Our formulations in Fig. 4.1 can be extended from a single source-destination pair to multiple source-destination pairs using a multi-commodity flow formulation [17] augmented with OR transmission constraints. By assigning a unique connection identifier to each source-destination pair, we introduce the variable f_{ij}^k to denote the amount of flow for connection k on link l_{ij}. For each flow k, according to some OR routing strategy, the corresponding transmitters and their forwarding candidates can be decided. Then the CCTS or GCTS can be constructed over the union of all the OR modules. Referring to Fig. 4.1, the objective is now to maximize the summation of all the flows out of all the sources; the flow conservation constraints at each node apply on a per-connection basis (constraint (4.2)); the total incoming flow into a source node is zero only for the connection(s) originating at that node (constraint (4.3)); similarly, the total outgoing flow from a destination node is zero only for the connection(s) terminating at that node (constraint (4.4)); f_{ij}^k is non-negative (constraint (4.5)); f_{ij}^k is equal to zero if the flow k is not routed by any link (constraint (4.6)); and the sum

of all the flows traversing on a link is constrained by the total amount of flow that can be delivered in all activity periods of the OR modules which contain this link (constraint (4.9)).

4.2 Impact of Transmission Rate and Forwarding Strategy on Throughput

The impact of the transmission rate on the throughput of OR is twofold. On the one hand, different rates have different transmission ranges, which lead to different neighborhood diversity. High-rate usually has short transmission range. In one hop, there are few neighbors around the transmitter, which presents low neighborhood diversity. Low-rate is likely to have long transmission range, therefore achieves high neighborhood diversity. From the diversity point of view, low rate may be better. On the other hand, although low rate brings the benefit of larger one-hop distance which results in higher neighborhood diversity and fewer hop counts to reach the destination, it may still end up with a low effective end-to-end throughput because the low rate disadvantage may overwhelm all other benefits. It is nontrivial to decide which rate is indeed better.

We now use a simple example in Fig. 4.3 to illustrate transmitting at lower rate may achieve higher throughput than transmitting at higher rate for OR. In this example, we assume all the nodes operate on a common channel, but each node can transmit at two different rates R and $R/2$. We compare the throughput from source a to destination d when the source transmits the packets at the two different rates. Fig. 4.3(a) shows the case when all the nodes transmit at rate R, and the packet delivery ratio on each link is 0.5. So the effective data rate on each link is $0.5R$. There is no link from a to d because d is out of a's effective transmission range when a operates on rate R. Assume the four nodes are in the carrier sensing range of each

(a) Source broadcasts at rate R

(b) Source broadcasts at rate R/2

Figure 4.3: End-to-end throughput comparison at different transmission rates

other, so they can not transmit at the same time. Assuming b and c are the forwarding candidates of a, and b has higher relay priority than c. Then link l_{ac} has effective forwarding rate of $0.25R$. By using the formulations in Fig. 4.1, we obtain an optimal transmitter schedule such that a, b and c are scheduled to transmit for a fraction of time 0.4, 0.4 and 0.2, respectively. So the maximum end-to-end throughput from a to d is $0.3R$. While in Fig. 4.3(b), when a is transmitting at a lower rate $R/2$, it can reach d directly with packet delivery ratio of 0.6. Additionally, we get higher packet deliver ratio from a to b and c as 0.8. In this case, the lower rate achieves longer effective transmission range and brings more spacial diversity chances. Assume d, b, and c are forwarding candidates of a, and with priority $d > b > c$. Similarly, we calculate the maximum throughput from a to d as 0.36R, which is 20% higher than the scenario in Fig. 4.3(a) where the system operates on a single rate.

Besides the inherent rate-distance, rate-diversity and rate-hop tradeoffs which affect the throughput of OR, the forwarding strategy will also have an impact on the throughput. For example, different forwarding candidates may achieve very different throughput, and even for the same forwarding candidate set, different forwarding priority will also result in different throughput, etc.. We refer readers to [68] for detail analysis on the impact of forwarding strategy on the OR throughput.

4.3 Rate and Candidate Selection Schemes

How to efficiently select the transmission rates and forwarding strategy for each node such that the network capacity can be globally optimized is still an open research issue. We have shown the example in Fig. 4.3 that nodes transmitting at a lower rate may lead to a higher end-to-end throughput than when nodes are transmitting at a higher rate. Then, what criteria should node a follow to select transmission rate, forwarding candidates and candidate priority to approach the capacity? It is non-trivial to answer this question. Towards the development of distributed and localized OR protocol that maximize the capacity, in this section, we propose two rate and candidate selection schemes, one is enlightened by least-cost opportunistic routing (LCOR) proposed in [24], and the other is inspired by geographic opportunistic routing (GOR) [84, 28, 68, 67, 70].

4.3.1 Least Medium Time Opportunistic Routing

In traditional routing, the medium time metric (MTM) [8] and expected transmission time (ETT) [23] have shown to be good metrics to achieve high throughput. For OR, we define the opportunistic ETT (OETT) as the expected transmission time to send a packet from n_i to any node in its forwarding candidate set $\mathcal{F}_i$.

$$OETT_{n_i}^{\mathcal{F}_i} = \frac{L_{pkt}}{R_i P_{\mathcal{F}_i}} \tag{4.11}$$

where L_{pkt} is the packet length, R_i is the data transmission rate at node n_i, and $P_{\mathcal{F}_i}$ is the probability of at least one candidate in $\mathcal{F}_i$ correctly receiving the packet sent by n_i:

$$P_{\mathcal{F}_i} = 1 - \prod_{q=1}^{r} (1 - p_{ii_q}) \tag{4.12}$$

Note that this metric actually generalizes the unicast ETT, that is, for $|\mathcal{F}_i| = 1$, the OETT reduces to the unicast ETT.

Denote by D_i the expected medium time (EMT) to reach the destination n_d from a node n_i. Assume that n_i's forwarding candidates are prioritized according to their expected medium time D_{i_q}, such that $D_{i_1} < D_{i_2} ... < D_{i_r}$. Then we define the remaining EMT to the destination n_d when node n_i choose forwarding candidate set $\mathcal{F}_i$ as following:

$$EMT_{\mathcal{F}_i}^{n_d} = \frac{1}{P_{\mathcal{F}_i}} \sum_{q=1}^{r} D_{i_q} p_{ii_q} \prod_{k=0}^{q-1} (1 - p_{ii_k}) \tag{4.13}$$

where $p_{ii_0} := 0$.

$p_{ii_q} \prod_{k=0}^{q-1}(1-p_{ii_k})$ is the probability of candidate n_{i_q} receiving the packet correctly but all the higher-priority candidates do not. That is, it is the probability of n_{i_q} becoming the actual forwarder. So the summation $\sum_{q=1}^{r} D_{i_q} p_{ii_q} \prod_{k=0}^{q-1}(1 - p_{ii_k})$ is the expected remaining medium time needed for a packet to travel to the destination for one transmission from node n_i. $\frac{1}{P_{\mathcal{F}_i}}$ is the expected transmission count n_i needs to make in order to deliver the packet to one of its forwarding candidate.

Note that like the OETT, the EMT generalizes the single-path case: when $|\mathcal{F}_i| = 1$, it simply becomes the delay from the next-hop to the destination. We should also notice that for any two different transmitters, n_i and n_j, even if $\mathcal{F}_i = \mathcal{F}_j$, they may have different EMT, since this EMT is affected by the delivery probabilities from the transmitter to its each forwarding candidate. In other words, the remaining EMT from a forwarding candidate set to the destination depends not only on the candidate set itself, but also on the predecessor node of this set.

We now define the least EMT of node n_i to the destination n_d in a multi-rate

scenario:

$$D_i = \min_{\mathcal{F}_i^j \in 2^{\mathcal{C}_i^j}, 1 \le j \le J} (OETT_{n_i}^{\mathcal{F}_i^j} + EMT_{\mathcal{F}_i^j}^{n_d}) \tag{4.14}$$

where $\mathcal{C}_i^j$ is the neighboring node set of node n_i when n_i transmits at rate R^j, $\mathcal{F}_i^j$ is the corresponding forwarding candidate set.

We enumerate all the possible $\mathcal{F}_i^j$ to get the optimal one. This equation represents the steady-state of the least medium time OR (LMTOR), that selects the forwarding candidates and transmission rate for each node to achieve the minimum end-to-end EMT. A distributed algorithm running like Bellman-Ford can solve the LMTOR problem. That is, in one iteration, each node n_i updates its value D_i^k, where k is the iteration index. This D_i^k is the estimated EMT from n_i to the destination at the k^{th} iteration; it converges toward D_i. $D_d^k = 0, \forall\ k$. One iteration step consists of updating the estimated EMT to the destination from each node:

$$D_i^{k+1} = \min_{\mathcal{F}_i^j \in 2^{\mathcal{C}_i^j}, 1 \le j \le J} (OETT_{n_i}^{\mathcal{F}_i^j} + EMT_{\mathcal{F}_i^j}^{n_d}(k)) \ \forall\ n_i \neq n_d \tag{4.15}$$

where $EMT_{\mathcal{F}_i^j}^{n_d}(k)$ is the remaining EMT computed using the costs $D_{i_q}^k$ $(n_{i_q} \in \mathcal{F}_i^j)$ from the previous iteration.

The rate and candidates selected by n_i are determined as a byproduct of minimizing the Eq. (4.15). The algorithm terminates when: $D_i^{k+1} = D_i^k\ \forall\ n_i \neq n_d$. Similar to the proof in [24], this algorithm converges after at most N iterations, where $N = |V|$ is the number of nodes in the network. Although this algorithm needs to enumerate all the combinations of neighboring nodes of each node, which is in exponential complexity, it is feasible when the number of neighbors per node is not large. In a denser network, we propose another local rate and candidate selection scheme by leveraging on the node's location information as in GOR.

4.3.2 Per-hop greedy: Most Advancement per Unit Time

A local metric: Expected Advancement Rate The location information is available to the nodes in many applications of multihop wireless networks, such as sensor networks for monitoring and tracking purposes [84] and vehicular networks [28]. GOR has been proposed as an efficient routing scheme in such networks. In GOR, nodes are aware of the location of itself, its one-hop neighbors, and the destination. A packet is forwarded to neighbor nodes that are geographically closer to the destination. In Chapter 2, we proposed a local metric, *expected packet advancement (EPA)* for GOR to achieve efficient packet forwarding. EPA for GOR is a generalization of EPA for traditional geographic routing [56,42]. It represents the expected packet advancement achieved by opportunistic routing in one transmission without considering the transmission rate. In this paper, we extend it into a bandwidth adjusted metric, *expected advancement rate (EAR)*, by taking into consideration various transmission rates.

We define the EAR as follows.

$$EAR_{n_i}^{\mathcal{F}_i} = R_i \sum_{q=1}^{r} a_{ii_q} p_{ii_q} \prod_{k=0}^{q-1} (1 - p_{ii_k}) \tag{4.16}$$

The physical meaning of EAR is the *expected bit advancement per second* towards the destination when the packet is forwarded according to the opportunistic routing procedure introduced in section 7.1.

The definition of EAR is the rate R_i multiplying the EPA proposed in [67]. According to the proved **relay priority rule** for EPA in Section 2.2.2, we have the following theorem for EAR:

Theorem 4.3.1. ***(Relay priority rule)*** *For a given transmission rate at n_i and $\mathcal{F}_i$, the maximum EAR can only be achieved by giving the candidates closer to the destination higher relay priorities.*

This theorem indicates how to prioritize the forwarding candidates when a transmission rate and the forwarding candidate set are given. From the definition of

EAR, it is also not difficult to find that adding more neighboring nodes with positive advancement into the existing forwarding candidate set will lead to a larger EAR. Therefore, we conclude that *an OR strategy that includes all the neighboring nodes with positive advancement into the forwarding candidate set and gives candidates with larger advancement higher relay priorities will lead to the maximum EAR* for a given rate.

Then a straightforward way to find the best rate is: for node n_i, at each transmission rate R^m $(1 \leq m \leq J)$, we calculate the largest EAR according to the above conclusion, then we pick the rate that yields the maximum EAR. This would be the local optimal transmission rate and the corresponding forwarding candidate set. Note that for a node n_i, it is possible that no neighboring nodes are closer to the destination than itself. In this case we need some mechanism like face routing [36] to contour the packet around the void. However, solving the communication voids problem is out of the scope of this dissertation.

Note that the above discussion does not take into consideration protocol overhead. As we have shown in [68, 67, 70], including as many as possible nodes might not be the optimal strategy when overheads, such as the time used to coordinate the relay contention at MAC layer, are taken into consideration. To consider the protocol overhead, the EAR can be extended to the metric EOT (expected one-hop throughput) which we will study in Chapter 7. However, in this chapter, since our goal is on studying the end-to-end throughput bound of OR, we apply EAR as the local metric, which is the upper bound of the packet advancement rate that can be made by any GOR.

4.4 Performance Evaluation

In this section, we use Matlab to investigate the impact of different factors on the end-to-end throughput bound of opportunistic routing, such as source-destination distances, node densities, and number of forwarding candidates. Both line and square topologies are studied for each factor. We also compare the performance of single rate opportunistic routing and multi-rate ones, and the performance of OR with traditional routing (TR). We call a routing scheme "traditional" when there is only one forwarding candidate selected for each packet relay at each hop.

The OR schemes we investigate include single-rate ExOR [9], single/multi-rate GOR and single/multi-rate LMTOR introduced in Section 4.3.1. For ExOR [9], each transmitter selects the neighbors with lower ETX (Estimated Transmission count) to the destination than itself as the forwarding candidates, and neighbors with lower ETX have higher relay priorities. For GOR, the forwarding candidates of a transmitter are those neighbors that are closer to the destination, and candidates with larger advancement to the destination have higher relay priorities. The EAR metric proposed in Section 4.2 is used to select the transmission rate for each node in the multi-rate scenario. For multi-rate LMTOR, the algorithm and metric proposed in Section 4.3.1 is used to choose transmission rate and forwarding candidates at each node. All the evaluations are under protocol model.

4.4.1 Simulation Setup

The simulated network has 20 stationary nodes randomly uniformly distributed on a line with length L or in a $W \times Wm^2$ square region. The data rates 24, 12, and 6 Mbps (chosen from 802.11a) are studied. We use one of the most common models - log-normal shadowing fading model [52] to characterize the signal propagation. The

received signal power is:

$$P_r(d)_{dB} = P_r(d_0)_{dB} - 10\beta log(\frac{d}{d_0}) + X_{dB} \tag{4.17}$$

where $P_r(d)_{dB}$ is the received signal power at distance d from the transmitter, *beta* is the path loss exponent, and X_{dB} is a Gaussian random variable with zero mean and standard deviation σ_{dB}. $P_r(d_0)_{dB}$ is the receiving signal power at the reference distance d_0, which is calculated by Eq. (4.18):

$$P_r(d_0)_{dB} = 10log(\frac{P_t G_t G_r c^2}{(4\pi)^2 d_0^2 f^2 L}) \tag{4.18}$$

where P_t is the transmitted signal power, G_t and G_r are the antenna gains of the transmitter and the receiver respectively, c is velocity of light, f is the carrier frequency, and L is the system loss.

In our simulation, $d_0 = 1m$, $\beta = 3$, $\sigma_{dB} = 6$, G_t, G_r and L are all set to 1, $P_t = 15dbm$, $c = 3 \times 10^8 m/s$, and $f = 5GHz$.

We assume a packet is received successfully if the received signal power is greater than the receiving power threshold (P_{Th}). According to [66], for 802.11a, the P_{Th} for 24, 12, and 6Mbps is -74, -79, and -82dbm, respectively. Then according to Eq. (4.17) and (4.18), the packet reception ratio for each rate at a certain distance d can be derived. The PRR vs distance for each data rate is shown in Fig. 4.4. We set the PRR threshold p_{td} as 0.1, so the effective transmission radius for each rate (24, 12, and 6Mbps) is 47, 70 and 88m, respectively. As discussed in [75], 802.11 systems have very close interference ranges for different channel rates, so we use a single interference range 120m for all channel rates for simplicity.

4.4.2 Impact of source-destination distances

In this subsection, we evaluate the impact of the source-destination distance on the end-to-end throughput bound of OR and TR in line and square topologies. For line

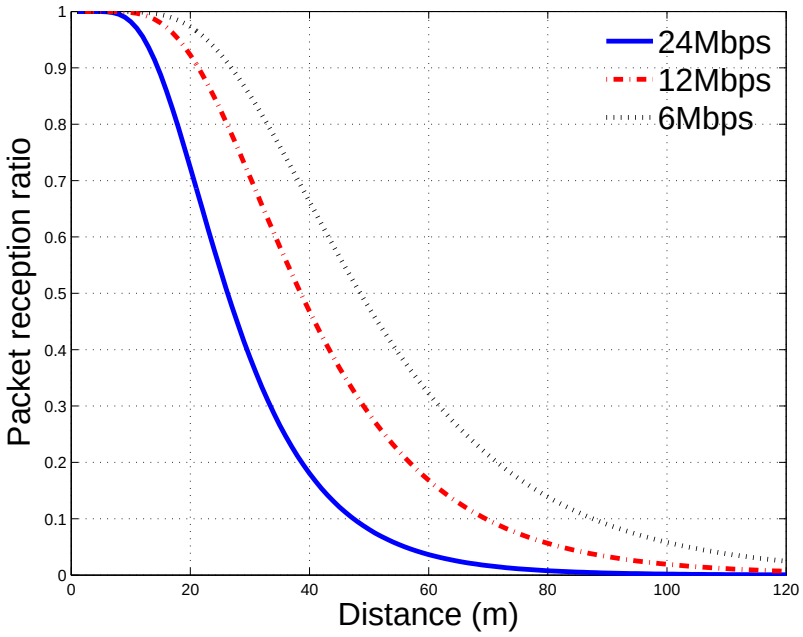


Figure 4.4: Packet reception ratio vs distance at different data rates

topology, the length L is set as 400m. We fix the left-end node as the destination, and calculate the throughput bounds from all other nodes to it under different OR and TR variants. For square topology, the side length is set as 150m. We fix the node nearest to the lower left corner as the destination, and calculate the throughput bounds from all other nodes to it. Therefore, there are 19 different source-destination pairs considered in the evaluation for each topology. We evaluate the performance under both single-rate and multi-rate scenarios. The average numbers of neighbors per node (indicated as ρ) under different topologies and data rates are summarized in Table 4.1.

In the single-rate scenario, for TR, we compute the exact end-to-end throughput bound between the source-destination pairs according to the LP formulations in [32], which normally result in multiple paths from the source to the destination. So we call it "Multipath TR". We also compute the end-to-end throughput of a single path that is found by minimizing the medium time (delay), and we call it "Single-path TR".

	ρ	
Rate (Mbps)	Line	Square
24	3.5	3.5
12	5.5	7.0
6	6.8	10.0

Table 4.1: Average number of neighbors per node under different topologies and data rates

The bound of single-path TR is calculated according to the formulations in [74]. For the three OR variants, we compute the throughput bounds under both conservative (indicated as 'c') and greedy (indicated as 'g') modes as we discussed in Section 4.1.2.

Fig. 4.5 shows the simulation results of LMTOR, ExOR, GOR and TR in a single rate (12Mbps) system under line topology. We have the following observations: 1) when the distance between the source and destination increases, the end-to-end throughput bound of each routing scheme decreases. 2) the OR achieves higher throughput bound than TR under different source-destination distances. 3) all the OR variants achieve the same performance under the same mode. 4) when source-destination distance is larger than 2 hops, OR in greedy mode results in higher end-to-end throughput than that in conservative mode, while when the source-destination distance is smaller than 2 hops, they represent the same performance. 5) the multi-path TR achieves almost the same throughput bound as single-path TR.

In the line topology, the throughput gain of OR over TR mainly comes from the opportunistic property. That is, for each packet transmission, multiple forwarding candidates help on forwarding the packet. The reliability of at least one forwarding candidate correctly receiving the packet is increased comparing to TR. The increased reliability reduces the retransmission overhead, and saves the medium time for each packet forwarding, thus improves the throughput.

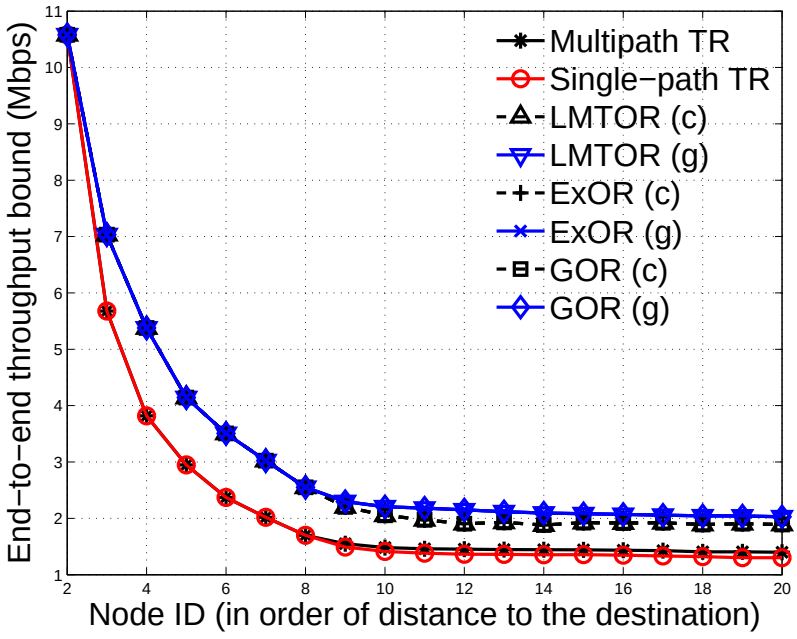


Figure 4.5: End-to-end throughput bound of OR and TR in a single rate (12Mbps) network under line topology

By tracing into the simulation, we find that the three OR variants result in the same forwarding candidate selection and prioritization at each forwarding node, although they follow different criteria to select the candidates and prioritize them. That's why we have the observation 3), which indicates that in the line topology the per-hop greedy behavior in GOR can approach the same end-to-end performance as that obtained by a distributed scheme like LMTOR.

For observation 4), when the source is near to the destination, all the nodes along the paths are in the interference range of each other, thus there is no concurrent transmission allowed in either greedy or conservative mode. Therefore, OR in both modes achieves the same performance when the source-destination distance is smaller than 2 hops. When the source-destination distance is lager than 2 hops, concurrent transmission in the network becomes possible. Since conservative mode requires interference free communication at all the forwarding candidates, foe each transmission, it consumes more space than greedy mode, which only needs interference free at least on forwarding candidate. That is, greedy mode achieves higher spatial reuse ratio than conservative mode and allows more concurrent transmissions in the network, thus results in higher throughput.

The observation 5) indicates that multipath TR does not really improve the wireless network throughput over the single-path TR in the line topology. The reason is that even when there are multiple paths between the source and destination, the links on different paths can not be scheduled at the same time due to interference. OR does make real use of multiple paths, in the sense that throughput can take place on any one of the outgoing links from the sender to its forwarding candidates.

Fig. 4.6 shows the simulation results of LMTOR, ExOR, GOR and TR in a single rate (12Mbps) system under square topology. One interesting observation is that the multipath TR achieves (up to 60%) higher throughput bound than single-path TR, and it can achieve comparable or even higher throughput than OR in conservative

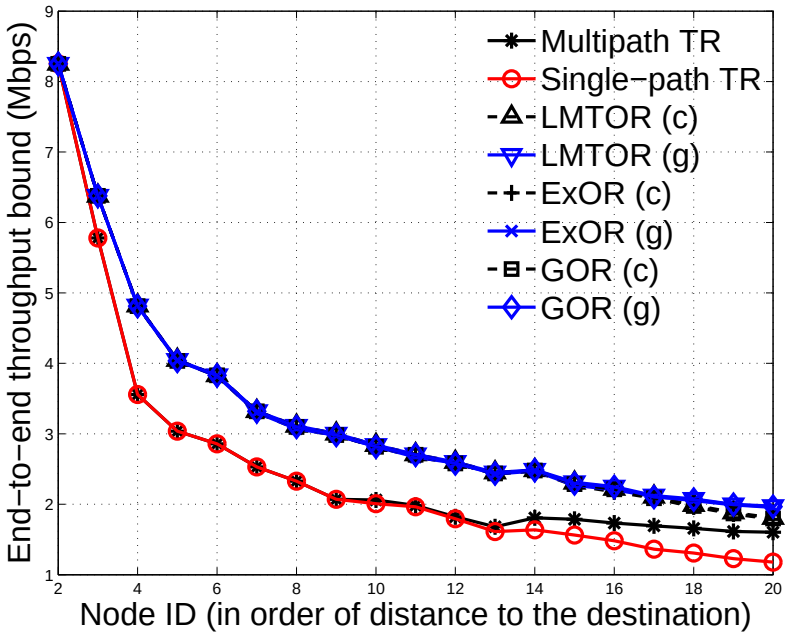


Figure 4.6: End-to-end throughput bound of OR and TR in a single rate (12Mbps) network under square topology

mode when the source-destination distance is larger than 2 hops. In the square topology, when the source and destination are far apart, real multipath routing becomes feasible. That is, different links on different paths can be activated at the same time and this improves the throughput. This observation also indicates that it is not a good idea to include as many as possible forwarding candidates into opportunistic routing when some protocol requires interference free at all the forwarding candidates. As we can see in Fig. 4.6 that OR in greedy mode still achieves higher throughput than OR in conservative mode and multipath TR. So the advantage of OR over TR is still validated.

Since OR in greedy mode always achieves higher throughput bound than that in conservative mode, in the following evaluation, the throughput bound of OR is only calculated under greedy mode. As the performance of ExOR is nearly the same as that of GOR, we will not show the simulation result of ExOR in the following figures. Now, we compare the throughput bounds of OR in multi-rate and single-rate systems.

Fig. 4.7 shows the simulation results of multi-rate LMTOR, multi-rate GOR, and single-rate GOR under line topology. We can see that generally multi-rate OR achieves better performance than any single-rate OR. When the distance between the source and destination is shorter than the interference range (corresponding to node ID 7), the system operating on 24Mbps achieves better performance than that on 12Mbps. However, the difference becomes smaller and smaller when the source-destination distance becomes larger, since more forwarding candidates are involved for 12Mbps and the spatial diversity is increased. When the source-destination distance is larger than the interference range, the performance of 24Mbps is as the same as that of 12Mbps. Fig. 4.8 shows the simulation results under square topology. An interesting difference from line topology is that the system operating at 24Mbps shows lower throughput bound than those operating at 12Mbps and 6Mbps for most of the source-destination pairs. The disadvantage of short transmission range and

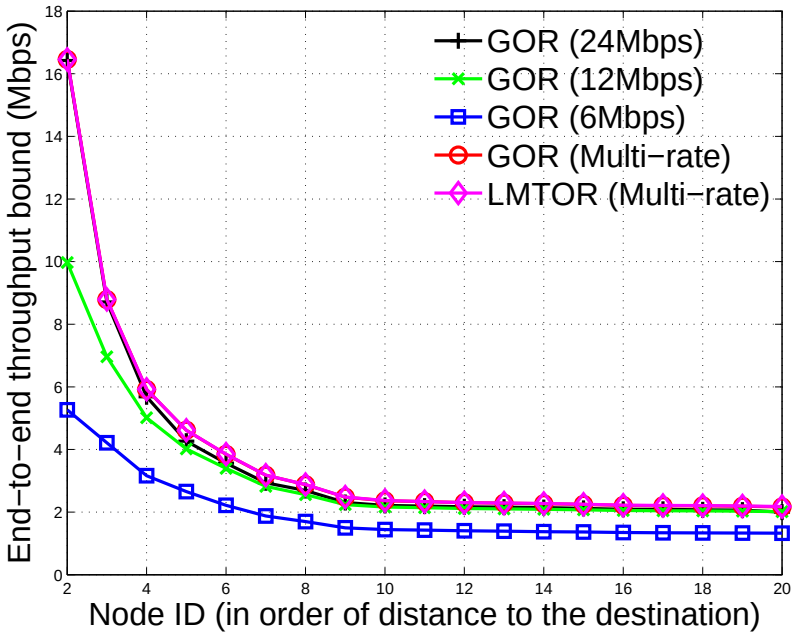


Figure 4.7: End-to-end throughput bound of OR in single-rate and multi-rate networks under line topology

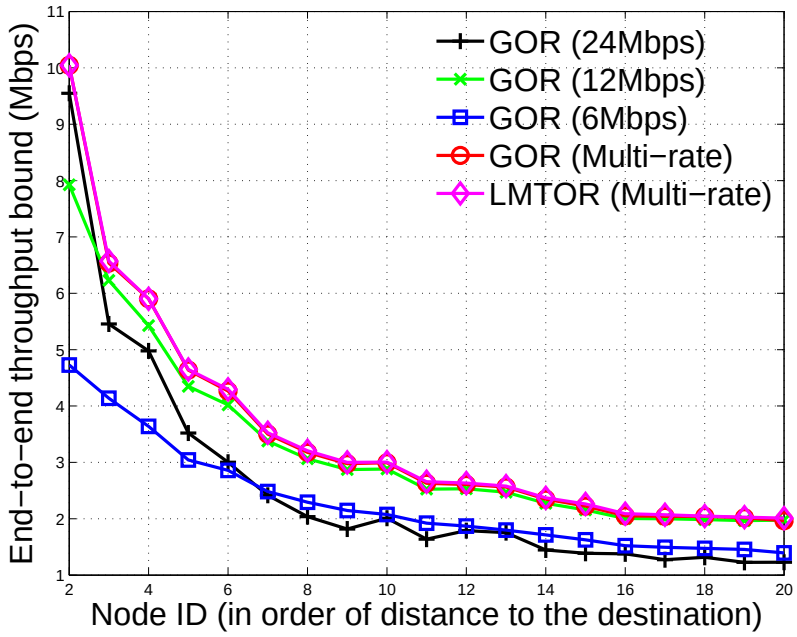


Figure 4.8: End-to-end throughput bound of OR in single-rate and multi-rate networks under square topology

lower spacial diversity of 24Mbps overwhelms its higher data rate advantage in the square topology.

4.4.3 Impact of forwarding candidate number

In this subsection, we study the impact of the number of forwarding candidates on the performance of OR. For line topology, we examine the bound between the two end nodes on the line. For square topology, we examine the throughput bound between the two end nodes on the diagonal. The topology sizes are set as the same as those in the previous simulation. For a transmitter, given a maximum number of forwarding candidates, the single-rate GOR selects the forwarding candidates as follows: first, it finds all the neighbors which are closer to the destination than the transmitter; second, if the number of the found neighbors is less than or equal to the maximum

number of forwarding candidates, GOR just involves all the found neighbors and gives the neighbors closer to the destination higher relay priorities. If the number of the found neighbors is greater than the maximum number of forwarding candidates, we apply the algorithm proposed in [67] to select the forwarding candidates which maximizes the EPA. For multi-rate GOR, we select the forwarding candidates for each single-rate GOR and calculate its corresponding EAR, then select the data rate with the highest EAR. For LMTOR, we apply the distributed algorithm proposed in Section 4.3.1. For the local search in Eq. (4.14) and (4.15), we only test a subset of all the neighbors with cardinality no larger than the maximum number of forwarding candidates.

Fig. 4.9 and 4.10 show the simulation results under line and square topologies, respectively. Generally, multi-rate OR achieves better performance than any single-rate OR, and multi-rate LMTOR achieves better performance than multi-rate GOR. In the square topology (Fig. 4.10), GOR on 12Mbps is always the best among all the single-rate GOR for all the different candidate sizes. The 24Mbps GOR performs even worse than 6Mbps GOR in square topology when the maximum forwarding candidate number is larger than 3. Since 24Mbps has the shortest transmission range, which results in the lowest node density as shown in Table 4.1, GOR on 24Mbps actually does not have 3 or more forwarding candidates to choose. Note that, the maximum number of forwarding candidates being equal to 1 corresponds to the TR. Although 6Mbps geographic TR (GTR) achieves lower throughput bound than 24Mbps GTR, it is not necessarily the truth for GOR. Since lower data rates have longer transmission ranges, this yields higher neighborhood diversities, which can help to increase the effective forwarding rate for each transmission when OR is used. In the line topology (Fig. 4.9), when the forwarding candidate number is greater than 3, GOR on 12Mbps achieves better performance than that on 24 Mbps which can be explained by the same reason. However, in the line topology, the disadvantage of low data rate of

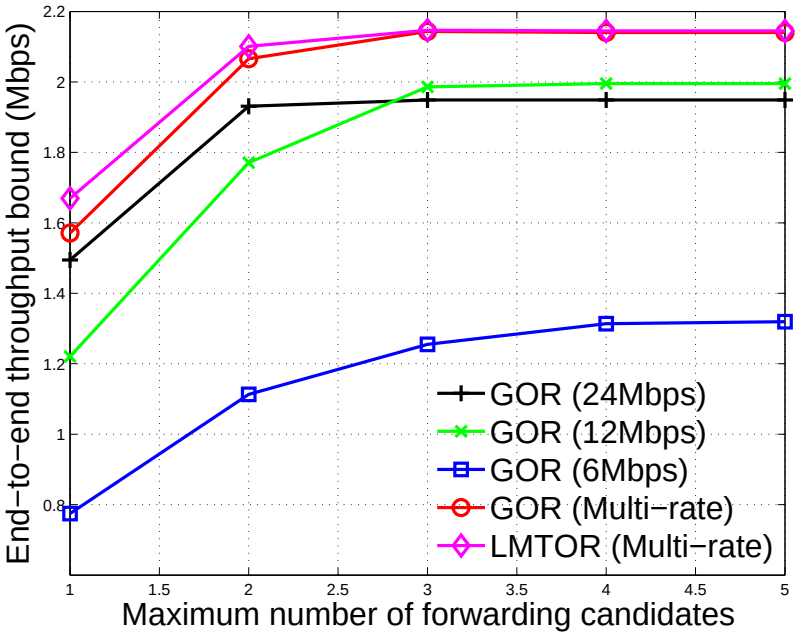


Figure 4.9: End-to-end throughput bound of OR with different number of forwarding candidates under line topology

6Mbps overwhelms its advantage on higher spatial diversity, GOR on 6Mbps shows the worst performance.

An interesting observation in both Fig. 4.9 and 4.10 is the concavity of each curve, which indicates that although involving more forwarding candidates improves the end-to-end throughput bound of OR, the capacity gained becomes marginal when we keep doing so. We can see that when the number of forwarding candidates is larger than 3, the end-to-end throughput bound remains almost unchanged. This end-to-end throughput observation is consistent with the local behavior found in Chapter 2. For a realistic MAC for OR, the coordination overhead is likely to increase when more forwarding candidates are involved. Since the throughput gain decreases when the number of forwarding candidates is increased, considering the MAC overhead, it may not be wise or necessary to involve as many as forwarding candidates in OR.

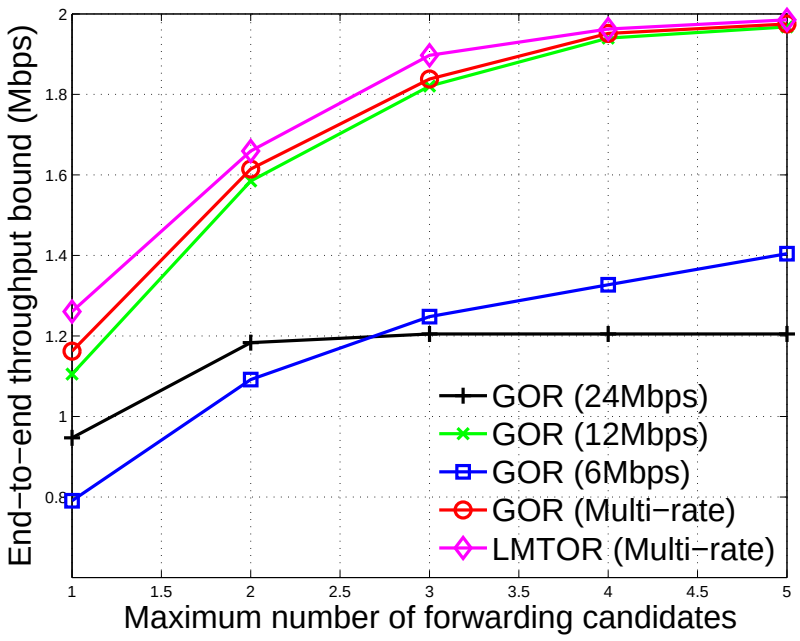


Figure 4.10: End-to-end throughput bound of OR with different number of forwarding candidates under square topology

	Line length			
Data rate (Mbps)	300	400	500	600
24	4.7	3.4	2.6	2.2
12	7.1	5.5	4.3	3.5
6	9.0	6.8	5.4	4.4

Table 4.2: Average number of neighbors per node at each rate under square topology with different side lengths

	Square side length			
Data rate (Mbps)	100	120	140	180
24	7.7	5.5	4.1	2.8
12	13.8	10.9	8.7	5.8
6	17	14.5	11.9	8.6

Table 4.3: Average number of neighbors per node at each rate under square topology with different side lengths

4.4.4 Impact of node density

The impact of the node density on the performance of OR is investigated in this subsection. Instead of single flow, we investigate the multi-flow case by randomly selecting four source-destination pairs in the network. The settings of the network terrain size and the corresponding number of neighbors per node under different data rates are summarized in Table 4.2 and 4.3.

Fig. 4.11 and 4.12 show the simulation results under line and square topologies, respectively. They show the same trend. There exists a threshold on the node density, higher than which, the GOR on 24Mbps performs better than that on 12Mbps, and lower than which, vice versa. The threshold is about 5.5 and 10.9 neighbors per node on 12Mbps for line and square topologies, respectively. Our proposed multi-rate

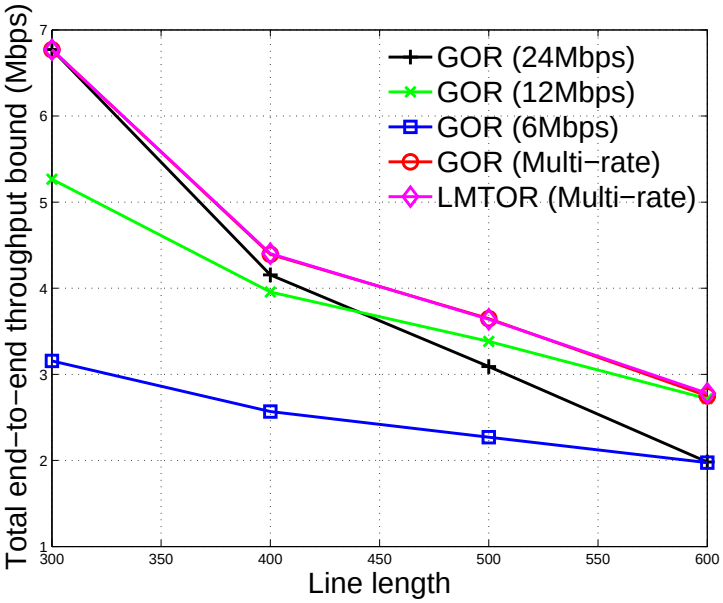


Figure 4.11: Total end-to-end throughput bound of OR under line topology with different lengths in multi-flow case

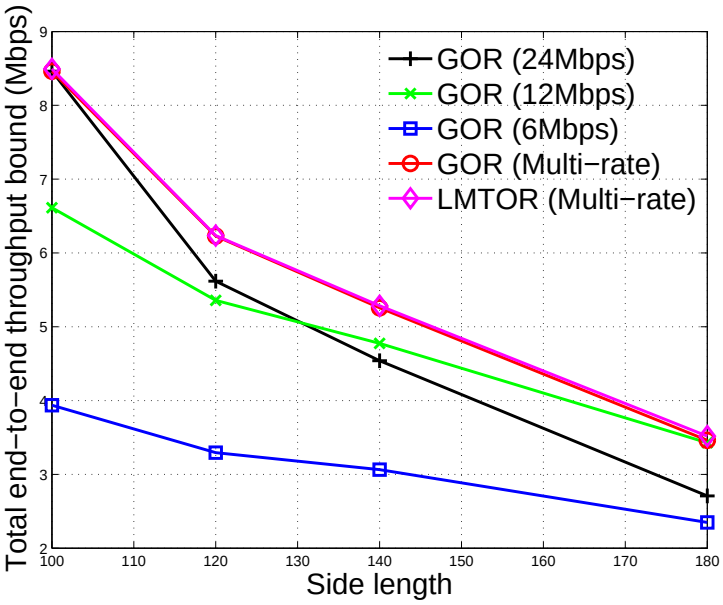


Figure 4.12: Total end-to-end throughput bound of OR under square topology with different side lengths in multi-flow case

GOR and LMTOR can adapt to the different node densities, and choose the proper transmission rate and forwarding candidate set to achieve the best performance than any single-rate GOR.

4.5 Conclusion

In this chapter, we studied the impact of multiple rates, interference, candidate selection and prioritization on the maximum end-to-end throughput of OR. Taking into consideration wireless interference, we proposed a new method of constructing transmission conflict graphs, and present a methodology for computing the end-to-end throughput bounds (capacity) of OR. We formulated the maximum end-to-end throughput problem of OR as a maximum-flow linear programming problem subject to the transmission conflict constraints and effective forwarding rate on each link. To the best of our knowledge, this is the first theoretical work on capacity problem of OR for multihop and multi-rate wireless networks.

We also proposed two metrics for OR under multi-rate scenario, one is *expected medium time* (EMT), and the other is *expected advancement rate* (EAR). Based on these metrics, we proposed the distributed and local rate and candidate selection schemes: LMTOR and MGOR, respectively. We validate the analysis results by simulation, and compare the throughput capacity of multi-rate OR with single-rate ones under different settings, such as different topologies, source-destination distances, number of forwarding candidates, and node densities. We showed that OR has great potential to improve the end-to-end throughput under different settings, and our proposed multi-rate OR schemes achieve higher throughput bound than any single-rate GOR. We observed some insights of OR: 1) The end-to-end capacity gained decreases when the number of forwarding candidates is increased. When the number of forwarding candidates is larger than 3, the end-to-end throughput bound remains

almost unchanged. 2) There exists a node density threshold, higher than which 24Mbps GOR performs better than 12Mbps GOR, and lower than which, vice versa. The threshold is about 5.5 and 10.9 neighbors per node on 12Mbps for line and square topologies, respectively.

Chapter 5

Theoretical End-to-end Throughput Bounds of Opportunistic Routing

Recent advantages in multi-radio multi-channel transmission technology allowing more concurrent transmissions in multi-hop wireless networks show the potential of substantially improving the system capacity [39, 7, 76]. In this chapter, we extend the framework in Chapter 4 to multi-radio multi-channel networks. Multi-radio/channel capability raises interesting issues on radio-channel assignment for OR. In single-radio single-channel system, OR naturally takes advantage of the redundant receptions on multiple neighboring nodes of a transmitter without consuming or sacrificing any extra channel resource. Because when a node is sending packets, other than its specified next-hop node, its one-hop neighbors usually can not send or receive other packets at the same time due to co-channel interference. So these one-hop neighbors have no other choice but to listen to the transmission of the sender. However, in multi-radio/channel systems, the one-hop neighbors have two choices: 1) they can operate on the same channel as the transmitter's to improve opportunistic diversity gain on

the receiver side, then more effective traffic can flow out of the sender and increase the system throughput; or 2) they can operate on other channels orthogonal to the sender's, thus increases chances to transmit/receive packets to/from other nodes, which may result in more concurrent effective traffic flow in the network and also increase the system throughput. Which choice the neighboring nodes should make is non-trivial. So the radio-channel assignment for optimizing OR throughput in multi-radio multi-channel systems deserves careful study. Furthermore, even if the channel assignment and scheduling are given, we still need to optimally (often dynamically) assign the relay priority among the forwarding candidates in order to maximize the system throughput. How to dynamically assign and schedule the forwarding priority among forwarding candidates has not been studied in the existing literature.

This chapter comprehensively studies the integrated radio-channel assignment, scheduling, candidate selection and prioritization problem for OR in multi-radio multi-channel systems. First, we propose a unified framework to compute the capacity of OR in single/multi-radio/channel systems, and formulate the capacity of OR as a linear programming (LP) problem subject to radio/channel constraints and effective forwarding rate constraints. Our model accurately captures the unique property of OR that instant throughput can take place from a transmitter to any one of its forwarding candidates. Second, we study the necessary and sufficient conditions on the schedulability of outgoing flows from a transmitter to its forwarding candidates. We also propose an LP approach and a heuristic algorithm to find a feasible scheduling of opportunistic forwarding strategies. Finally, leveraging our analytical model, we find that OR can achieve comparable or even better performance than TR by using less radio resource.

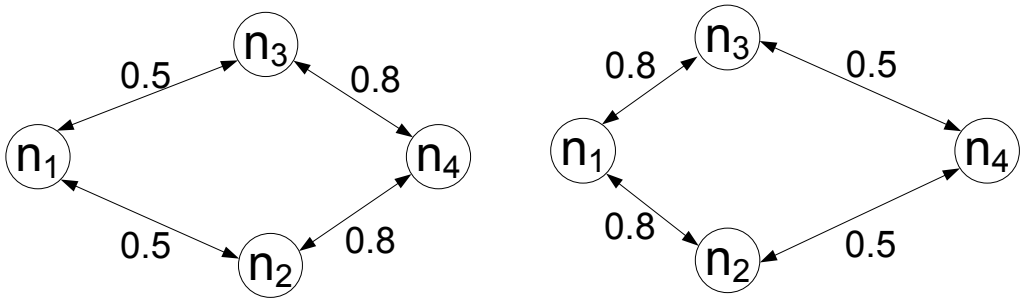
The rest of this chapter is organized as follows. Section 5.1 introduces the system model. Then we motivate the study of OR in multi-radio/channel systems in section 5.2. We propose the framework of computing the throughput bounds of OR in multi-

radio multi-channel systems in Section 5.3. Simulation results are presented and analyzed in Section 5.4. Conclusions are drawn in Section 5.5.

5.1 System Model

We consider a multi-hop wireless network with N nodes. Each node n_i $(1 \leq i \leq N)$ is equipped with one or more wireless interface cards, referred to as radios in this work. Denote the number of radios in each node n_i as t_i $(i = 1...N)$. Assume K orthogonal channels are available in the network without any inter-channel interference. We consider the system with channel switching capability, such that a radio can dynamically switch across different channels. However, we assume two radios on the same node are not far enough apart to each other to create spatial diversity, so there is no performance gain to assign the same channel to the different radios on the same node at any instant. For simplicity, we assume each node n_i transmits at the same data rate R_i among all its radios and channels. We also assume half-duplex on each radio, that is, a radio cannot transmit and receive packets at the same time. There is a unified transmission range R_T and interference range R_I for the whole network. Typically, $R_I > R_T$. Two nodes, n_i and n_j, can communicate with each other if the euclidian distance d_{ij} between them is less than R_T and they operate on the same channel. Due to the unreliability of wireless links, there is a packet reception ratio (PRR) associated with each transmission link.

We call the candidate selection and prioritization a **forwarding strategy**. Denote a forwarding strategy as $\mathcal{H} = (\mathbf{\Phi}, \mathcal{P})$, where $\mathbf{\Phi}$ is an indicator function on the potential forwarding candidates defined in Eq. (5.1), and $\mathcal{P}$ is a permutation function of the potential forwarding candidates. So $\mathbf{\Phi}$ represents selection of forwarding candidates and $\mathcal{P}$ represents prioritization of forwarding candidates. We denote $\mathcal{P}(j) < \mathcal{P}(k)$ if n_{i_j} has higher forwarding priority than n_{i_k}. Thus, a specified $\mathcal{H}$ can uniquely decide



(a) PRRs from the source (n_1) to relays (n_2 and n_3) are worse than that from the relays to the destination n_4

(b) PRRs from the source (n_1) to relays (n_2 and n_3) are better than that from the relays to the destination (n_4)

Figure 5.1: Four-node networks under different channel conditions (link PRRs).

a forwarding candidate sequence $\mathcal{F}_i$.

$$\mathbf{\Phi}(j) := \phi_j = \begin{cases} 1, & n_{i_j} \text{ is selected as a forwarding candidate;} \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

5.2 Problem Motivation

As discussed in Chapter 4, we know that candidate selection and prioritization affect the effective throughput on the egress links from a sender to its forwarding candidates, so they will affect the end-to-end throughput. Other than these two factors, multi-radio/channel capability, radio-channel assignment and scheduling also have impact on the end-to-end throughput. We present the following scenarios that motivate our study on OR in multi-radio/channel systems.

Consider two four-node network scenarios in Fig. 5.1 under different channel conditions (link PRRs), we want to find the maximum throughput from the source node n_1 to the destination n_4 under different radio/channel configurations. The PRR is indicated on each link. For ease of illustration, we assume the PRR is identical under different channels in each network. We assume each node is in the interference range

of each other. So there is only one transmitter can be active on the same channel at any instant in the network. The basic idea to achieve the optimal (maximum) end-to-end throughput is by activating transmitters/links periodically and in each period, we optimally decide which transmitters/links should be scheduled on what channel for how long. In the following discussion, we assume the scheduling period is one unit, and the network is saturated such that every node always has packets in its queue to be routed to its neighbors. All the optimal routing and scheduling results shown in the following discussion are obtained by applying the methodology that we will propose in Section 5.3. For each scenario, we discuss two radio/channel configuration cases: one radio one channel (1R1C) and one radio two channels (1R2C).

We first discuss the scenario in Fig. 5.1(a), where the PRRs from the source (n_1) to relay nodes (n_2 and n_3) are worse than that from the relays to the destination (n_4).

For 1R1C case, there is no need for channel assignment. The optimal transmission scheduling is as following. (1) In the first 0.258 time fraction, we activate n_1 to transmit packets to n_2 and n_3 by OR with n_2 having higher priority than n_3. According to the property of OR discussed in Section 4.1.3, at the end of the first time fraction, the effective traffic flows from n_1 to n_2 and n_3 are $1 \times 0.5 \times 0.258 = 0.129$ and $1 \times 0.5 \times (1 - 0.5) \times 0.258 = 0.0645$, respectively. (2) Similarly, in the second 0.258 time fraction, we activate n_1 to transmit packets to n_2 and n_3 by OR with n_3 having higher priority than n_2. So at the end of the second time fraction, the effective traffic flows from n_1 to n_2 and n_3 are 0.0645 and 0.129, respectively. Then at the end of the 0.516 time fraction, the effective traffic flows routed from n_1 to n_2 and from n_1 to n_3 are both $0.129 + 0.0645 = 0.194$. (3) We then activate link l_{24} for 0.242 time fraction. (4) Finally, we activate link l_{34} for 0.242 time fraction. So the maximum throughput from n_1 to n_4 is $\frac{0.242 \times 0.8 \times 2}{1} = 0.387$.

For 1R2C case, we should jointly consider channel assignment and forwarding scheduling. The optimal channel assignment and scheduling is summarized in Table

Link sets	$\langle l_{12}^1, l_{13}^1 \rangle$	$\langle l_{13}^1, l_{12}^1 \rangle$	$\{l_{12}^1, l_{34}^2\}$	$\{l_{13}^1, l_{24}^2\}$
Time fractions	0.14	0.14	0.36	0.36
Link throughput	$\langle 0.07, 0.035 \rangle$	$\langle 0.035, 0.07 \rangle$	{0.18, 0.29}	{0.18, 0.29}

Table 5.1: Channel assignment and scheduling of opportunistic forwarding strategies for Fig 5.1(a) in 1R2C case.

Link sets	$\{l_{24}^1, l_{13}^2\}$	$\{l_{34}^1, l_{12}^1\}$	l_{24}^1	l_{34}^1
Time fractions	{0.354, 0.3125}	{0.354, 0.3125}	0.146	0.146
Link throughput	$\{0.177, 0.25\}$	$\{0.177, 0.25\}$	0.073	0.073

Table 5.2: Channel assignment and scheduling of traditional routing strategies for Fig 5.1(b) in 1R2C case.

5.1, where we denote opportunistic routing by ordered set "$\langle\ \rangle$", and the traditional routing by general set "$\{\ \}$". l_{ab}^x means link l_{ab} operates on channel x. The scheduling is as following. (1) n_1, n_2, and n_3 all operate on channel 1 for 0.28 time fraction. In half of the fraction, n_1 forwards packets to the candidate set $\{n_2, n_3\}$ using OR strategy $\langle n_2, n_3 \rangle$, and for the remaining portion, changes the strategy as $\langle n_3, n_2 \rangle$. So at the end of the 0.28 time fraction, the effective flows from n_1 to n_2 and n_3 are both 0.11. (2) We then activate link l_{12} on channel 1 and l_{34} on channel 2 for 0.36 time fraction. At the end of this time fraction, the effective flows from n_1 to n_2 and from n_3 to n_4 are 0.18 and 0.29, respectively. (3) Finally, we activate link l_{13} on channel 1 and l_{24} on channel 2 for 0.36 time fraction. At the end of this time fraction, the effective flows from n_1 to n_3 and from n_2 to n_4 are 0.18 and 0.29, respectively. So the optimal throughput from n_1 to n_4 is 0.58. Comparing with the 1R1C case, we scale down the OR forwarding time fraction and increase the time fraction for concurrent transmissions by taking advantage of the multi-channel capability, and the throughput is increased by 50%.

Next, we discuss the scenario in Fig. 5.1(b), where the PRRs from the source (n_1)

to relay nodes (n_2 and n_3) are better than that from the relays to the destination (n_4).

For 1R1C case, the optimal routing and scheduling for this scenario are similar to that in Fig. 5.1(a). (1) In the first 0.171 time fraction, we activate n_1 to transmit packets to n_2 and n_3 by OR with n_2 having higher priority than n_3. (2) In the second 0.171 time fraction, we activate n_1 to transmit packets to n_2 and n_3 by OR with n_3 having higher priority than n_2. (3) We then activate link l_{24} for 0.329 time fraction to transmit packets from n_2 to n_4. (4) Finally, we activate link l_{34} for 0.329 time fraction to route the packets from n_3 to n_4. So the maximum throughput from n_1 to n_4 is $\frac{0.329\times0.5\times2}{1} = 0.329$. Comparing to the 1R1C case in Fig. 5.1(a), we can see that the time fraction allocated for opportunistic routing from the source to the relays is reduced since now the channel conditions from the source to the relays are better than that from the relays to the destinations. Therefore, we allocate more time for relays to forward packets to the destination.

However, for 1R2C case, the optimal routing and scheduling for this scenario, which are summarized in Table 5.2, are quite different from that in Table 5.1. We schedule the transmissions as follows. (1) We first activate link l_{24} on channel 1 for 0.354 time fraction, and at the same time activate link l_{13} on channel 2 for 0.3125 time fraction. (2) After the 0.354 time fraction, we then activate link l_{34} on channel 1 for 0.354 time fraction, and at the same time activate link l_{12} on channel 2 for 0.3125 time fraction. (3) At the end of 0.708 time fraction, we activate link l_{24} on channel 1 for 0.146 time fraction. (4) Finally, we activate link l_{34} on channel 1 for 0.146 time fraction. So the maximum end-to-end throughput from n_1 to n_4 is 0.5. In this case, it is not necessary to use opportunistic routing. Different from the 1R1C case, now the relay nodes (n_2 or n_3) can operate on a different channel from the sender's (n_1's), thus push more flow to the destination. Because the channel conditions from the source to the relays are better than that from the relays to the destination, the

maximum throughput now is constrained by the bottleneck links from the relays to the destination. So we should allow more concurrent transmissions to saturate the bottleneck links instead of making use of OR to push more flows out of the sender.

For 2R2C case, radio-channel assignment and OR forwarding scheduling are more complicated than the above two cases. We definitely need a unified framework to compute the throughput bound between two end nodes in single/multi-radio/channel systems. In the following section, we will propose this framework. Our framework can be used to compute the throughput bound between two end nodes as well as help us get insights into the OR behavior (candidate selection and prioritization) under different system configurations.

5.3 Problem Formulation

In this section we present our methodology to compute the throughput bound between two end nodes in a multi-radio multi-channel multihop wireless network. We assume that packet transmission at an individual node can be perfectly scheduled by an omniscient and omnipotent central entity. Thus, we do not consider issues such as MAC contention. We first introduce the **extended graph** to unify the multi-radio/channel and single-radio/channel cases. We further discuss the necessary and sufficient conditions of the schedulability of an egress flow demand vector associated with a transmitter in a concurrent transmission set (CTS). We also propose an LP approach and a heuristic algorithm to schedule opportunistic forwarding strategies to satisfy a flow demand vector. We finally present the LP formulation for computing throughput bounds between two end nodes based on CTS and the necessary condition.

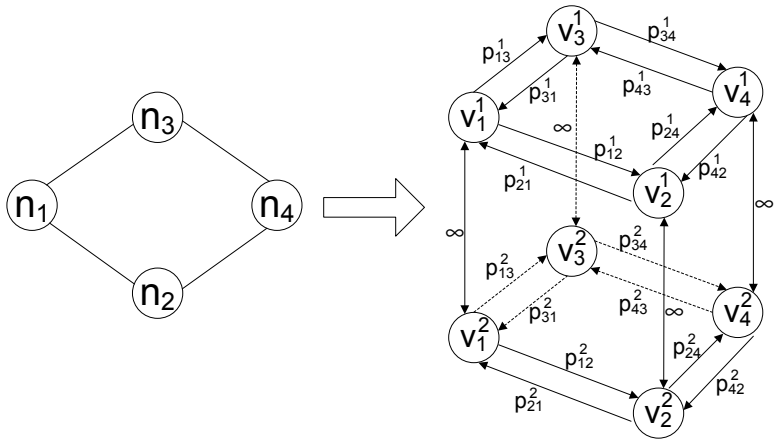


Figure 5.2: Transformation from a four-node network with two channels into an extended graph.

5.3.1 Extended Graph

We define an extended graph $G = (V, E + E')$, where $V = \{v_i^k | i = 1...N, k = 1...K\}$, which is the set of all possible **transceiver configurations** of each node n_i. v_i^k indicates node n_i operating on the channel k. E is the wireless link set. $E = \{l_{ij}^k | i = 1...N, j = 1...N, i \neq j, d_{ij} < R_T, k = 1...K\}$. Each wireless link l_{ij}^k is associated with a PRR p_{ij}^k. In this way, the original connected network is extended to K parallel connected subnetworks. $E' = \{l_i^{h,k} | i = 1...N, k = 1...K, h = 1...K, h \neq k\}$, which are wired links with infinite capacity between any vertices that share the same nodes in the original network. These links represent the in-node capacity among different radios and channels, because packets received by a radio/channel of one node, can be transmitted/forwarded by another radio/channel on the same node. Fig. 5.2 illustrates the transformation from a four-node network with two channels into an extended graph.

Each transceiver configuration in the extended graph can be in one of the three states: transmission, reception, and null. We call a vertex in transmission/reception state as transmitter/receiver, respectively. A transceiver configuration is in state null

means it does not send or receive any packets from any other transceiver configurations. We say a link l_{ij}^k is **active** if and only if v_i^k is in transmission state and v_j^k is in reception state, otherwise this link is **inactive**.

5.3.2 Concurrent Transmission Sets

In this subsection, we discuss which set of links in the extended graph can be active at the same time. We name a set of concurrent active links as a **concurrent transmission set** (CTS). The motivation of building concurrent transmission sets is similar to building independent sets in [32] and concurrent transmission patterns in [76]. That is, taking the benefit of time-sharing scheduling of different concurrent transmission sets, we could achieve a collection of capacity graphs, associated with capacity constraint on each link. OR can be performed on the underlying capacity graph to achieve the maximum throughput. The construction of CTS is more complicated than that in Chapter 4. Because, besides the co-channel interference, radio interface limits in the multi-radio system also impose constraint on concurrent transmissions in the extended graph.

For OR, we define the following co-channel interference model. Two wireless links, l_{ij}^k and l_{pq}^k, in the extended graph can be virtually active or active at the same time if they share the same transmitter ($i = p$), or v_j^k and v_q^k is out of the interference range of v_p^k and v_i^k, respectively. Otherwise, these two links interfere with each other. This co-channel interference model requires interference free only at receiver side as in [31, 32]. It can be extended to the case that requires interference free at both transmitter and receiver sides. We say a link l_{ij}^k is **usable** if it is not interfered by any other active links; otherwise, it is **unusable**. When a link l_{ij}^k is usable, we say receiver v_j^k is usable for the transmitter v_i^k. Note that the effective forwarding rate (capacity) on each active link is decided by the opportunistic routing strategy as discussed in Section 4.1.3.

Formally, a CTS α can be represented by an indicator vector on all links, written as $\alpha = \{\psi_{ij}^{k\alpha}, \forall\ l_{ij}^k \in E\}$.

$$\psi_{ij}^{k\alpha} = \begin{cases} 1, & l_{ij}^k \text{ is active in CTS } \alpha; \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

Denote the following indicator variable to represent the vertex state in CTS α:

$$\eta_i^{k\alpha} = \begin{cases} 1, & v_i^k \text{ is not null in CTS } \alpha; \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

Note that when a transceiver configuration v_i^k is in transmission state, it can transmit packets to multiple receivers. While when it is in reception state, it can only receive packets from one transmitter. This can be formally represented by:

$$\eta_i^{k\alpha} = \min(1, \sum_{l_{ij}^k \in E} \psi_{ij}^{k\alpha} + \sum_{l_{ji}^k \in E} \psi_{ji}^{k\alpha}), \forall\ i = 1...N, k = 1...K \tag{5.4}$$

Although any two active links operating on different channels do not interfere with each other, due to radio interface constraint, the number of channels being used on one node must be less than the number of radios installed on this node. To satisfy the constraint that the number of radios in use does not exceed the radio equipment at each node,

$$\sum_{k=1}^{K} \eta_i^{k\alpha} \leq t_i, \forall\ i = 1...N \tag{5.5}$$

If two wireless links are concurrently active on the same channel, they must not interfere with each other. This can be represented by

$$\psi_{ij}^{k\alpha} + \psi_{pq}^{k\alpha} \leq 1 + I(l_{ij}^k, l_{pq}^k), \forall\ k = 1...K \tag{5.6}$$

where

$$I(l_{ij}^k, l_{pq}^k) = \begin{cases} 1, & l_{ij}^k \text{ and } l_{pq}^k \text{ do not interfere;} \\ 0, & \text{otherwise.} \end{cases} \tag{5.7}$$

Note that the number of all the CTS's is exponential in the number of nodes, radios and channels. However, it may not be necessary to find all of them to maximize an end-to-end throughput. Some heuristic algorithm similar to that in [61], or column generation technique [76] can be applied to find a "good" subset of all the CTS's to approach the optimal solution. As it is not our main contribution, we will not go into detail of the technologies of finding CTS's. The CTS concept not only helps us calculate the capacity bound of OR, but also provides a way to study the behavior (such as candidate selection and prioritization) of OR in multi-radio multi-channel systems.

In the following subsection, we will first discuss the fundamental problem: given an opportunistic module in a CTS, what's the capacity region of the egress links from the transmitter to its forwarding candidates.

5.3.3 Capacity Region of Opportunistic Module

An **opportunistic module** in a CTS consists of a transmitter, all of its usable receivers and the corresponding wireless links from the transmitter to the receivers as illustrated in Fig. 5.3. Since the transmitter operates on the same channel as the receivers, for simplicity, we denote a transmitter, v_i^k as i, and the receivers as i_j's. The link $l_{ii_j}^k$ is simplified as l_j, and PRR $p_{ii_j}^k$ is simplified as p_j. In the active time of the opportunistic module in the CTS, we assume the time is further divided into sub-slots, in each sub-slot, a different forwarding strategy can be used. For example, if a transmitter i has two potential forwarding candidates, i_1 and i_2, four forwarding strategies can be used: 1) $\mathcal{H}_1$, only i_1 is used as the forwarder; 2) $\mathcal{H}_2$, only i_2 is used as the forwarder; 3) $\mathcal{H}_3$, i_1 and i_2 are both used, and i_1 has higher relay priority and than i_2, that is $\mathcal{F}_i = \langle i_1, i_2 \rangle$; 4) $\mathcal{H}_4$, $\mathcal{F}_i = \langle i_2, i_1 \rangle$. Strategy 1) and 2) correspond to the traditional routing, and strategy 3) and 4) correspond to the opportunistic routing. According to the definition of the effective forwarding rate in Eq. (4.1), as

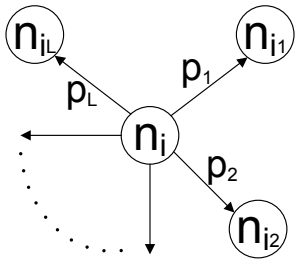


Figure 5.3: A transmitter n_i is transmitting a packet, and its potential forwarding candidate n_{i_j} $(1 \leq j \leq L)$ can correctly receive this packet with probability p_j.

	$\langle i_1 \rangle$	$\langle i_2 \rangle$	$\langle i_1, i_2 \rangle$	$\langle i_2, i_1 \rangle$
l_1	p_1	0	p_1	$p_1(1-p_2)$
l_2	0	p_2	$p_2(1-p_1)$	p_2

Table 5.3: Normalized effective forwarding rate on each link under different forwarding strategies with $L = 2$.

the transmission rate is normalized as one unit, we summarize the rate capacity on each link under each forwarding strategy as in Table 5.3.

Note that, no matter what forwarding strategy is used, the normalized forwarding rate (normalized by the transmission rate R) on each link can not exceed the link PRR, and the link rate summation can not exceed the **one-hop reliability** defined in Eq. (2.13). Recall that, for a forwarding candidate sequence $\mathcal{F}_i$, the one-hop reliability $P_{\mathcal{F}_i}$ is the probability of at least one candidate receiving the packet correctly. It is worthy pointing out that the forwarding reliability is only decided by the PRR from the transmitter to the receivers in the forwarding candidate set, but independent of the priority between the forwarding candidates.

The opportunistic module is equivalent to the single-server multi-user model in [62], which shows that any downlink with L independent ON/OFF channels has a capacity region that is given by a set of 2^L inequalities: Each inequality corresponds

to a subset of channels and indicates that the sum input rate into this subset is less than or equal to the probability that at least one channel within the subset is ON. Each link from a transmitter i to its forwarding candidate i_j in the opportunistic module can be seen as an ON/OFF channel with ON-probability of p_j. Then by applying the proved result in [62], we have the capacity region of the egress links from a transmitter i to its potential forwarding candidates as in Theorem 5.3.1.

Theorem 5.3.1. *For a transmitter i with L potential forwarding candidates, $i_1, ..., i_L$ ($L \geq 1$), in a CTS T_α, assume transmitter i broadcasts packets at rate R. In the active period τ_α of T_α, denote the traffic flow demand and PRR on link l_j ($j = 1...L$) as f_j and p_j, respectively. Then the traffic flow demand vector $\mathbf{f} = [f_1, ..., f_L]$ is in the capacity region if condition (5.8) holds.*

$$\sum_{i=1}^{L} f_i \cdot \phi_i \leq R \prod_{i=1}^{L} (1 - p_i \phi_i), \forall\ [\phi_1, ..., \phi_L] \in \{0, 1\}^L \tag{5.8}$$

Proof. The physical meaning of condition (5.8) is that any subset summation of traffic demand vector $\mathbf{f}$ must be bounded by the maximum achievable **candidate set forwarding rate** which is obtained by involving all the corresponding forwarders under the selection strategy $\mathbf{\Phi}$ defined in Eq. (5.1). □

Therefore, we say a traffic demand vector $\mathbf{f} = [f_1, ..., f_L]$ is *not schedulable* if condition (5.8) does not hold. Actually, Theorem 5.3.1 gives the necessary condition for the schedulability of a traffic demand vector $\mathbf{f}$. Now we discuss how to get a feasible schedule of opportunistic forwarding strategies to satisfy a traffic demand vector in the capacity region.

5.3.4 A Scheduling based on LP

A way to get a schedule of opportunistic forwarding strategies for traffic demand vector $\mathbf{f}$ is by solving a linear programming problem in Fig. 5.4. The basic idea of

$$Min \sum_{j=1}^{L!} \beta_j \tag{5.9}$$

$$s.t.$$

$$f_i \leq \sum_{j=1}^{L!} \beta_j \widetilde{R}_i^j, \ \forall \ i = 1...L \tag{5.10}$$

$$0 \leq \beta_j \leq 1, \ \forall \ j = 1...L! \tag{5.11}$$

Figure 5.4: LP formulations to test if a traffic demand vector is schedulable

this programming problem is to enumerate all possible $L!$ opportunistic forwarding strategies, and assign the j^{th} strategy a time fraction β_j. Then the effective rate $\widetilde{R}_i^j$ on link l_i in the j^{th} strategy can be calculated by Eq. (4.1). So the accumulated rate capacity on link l_i is $\sum_{j=1}^{L!} \beta_j \widetilde{R}_i^j$. If the solution of the objective function (5.9) is no greater than 1, then the traffic demand vector is schedulable, and β_j ($j = 1...L!$) is the byproduct of the linear programming; otherwise, the traffic demand vector is not schedulable. So the sufficient condition for the schedulability of a traffic demand vector $\mathbf{f}$ can be stated as in Theorem 5.3.2

Theorem 5.3.2. *Any traffic demand vector* $\mathbf{f} = [f_1, ..., f_L]$ *is schedulable if the solution of the objective function (5.9) in the linear programming problem formulated in Fig. 5.4 is no greater than 1.*

The linear programming problem in Fig. 5.4 provides a way to judge the schedulability of a traffic demand vector for an opportunistic module. Typically, L is at most the number of all the one-hop neighbors of a transmitter, so it tends to be a relatively small number. Thus, the linear programming can be an efficient approach to find a scheduling of forwarding strategies to satisfy a schedulable traffic demand vector. However, it is not necessary to enumerate all the possible forwarding strategies to find a feasible scheduling. In the following subsection, we propose a heuristic algorithm to

find a feasible scheduling of forwarding strategies that satisfies a flow demand vector.

5.3.5 A Heuristic Scheduling for an Opportunistic Module

Table 5.4 describes the heuristic recursive algorithm that finds a scheduling of opportunistic forwarding strategies satisfying a traffic demand vector **f**. The basic idea of this algorithm is to satisfy each flow one-by-one by two priority settings: assigning the corresponding candidate the highest and lowest priority in the existing subset of the candidates. One key property of OR we take advantage of is that from the lower-priority candidate point of view, the impact of higher-priority candidates on its effective forwarding rate is not dependent on the priority relationships among the higher-priority candidates, but only relative to their PRRs. Then we can consider a group of forwarding candidates $\mathcal{F}$ as a virtual candidate, whose PRR is the candidate set forwarding reliability defined in Eq. (2.13), and the flow demand to this virtual candidate is the accumulated flow to all the candidates in $\mathcal{F}$. Here we only need to take into account non-zero flow demand.

In Table 5.4, the input of the prioritizing and scheduling algorithm PS includes: **F**, the vector of flows; **P**, the corresponding PRR vector; **I**, the corresponding forwarding candidate index vector; r, the number of candidates/flows in **I**/**F**; β, the active time fraction of the links corresponding to candidates in **I**; ω, a scalar on the PRR which is used to calculate time fraction β_1 and β_2 in line 8. Initially, $\beta = \omega = 1$. The output of this algorithm is a set of opportunistic forwarding strategies (forwarding candidate sequences), S, and the corresponding time fraction vector, Γ. Lines 1 and 2 indicate the basic case where there is only one flow in F, then the flow index and the corresponding time fraction β are returned. When the flow number is larger than 1, we first pre-process the traffic demand vector (in lines 4 and 5) so that if there is a flow equal to its corresponding scaled PRR or smaller than the scaled effective forwarding rate when the corresponding candidate is assigned the lowest

(S,Γ) = PS($\mathbf{F}$,$\mathbf{P}$,$\mathbf{I}$,r,β,ω)

1 **if** r==1

2 **return**($\langle \mathbf{I}[1] \rangle$,$\beta$)

3 **else**

4 **if** $\exists\ \mathbf{F}[i] == \omega \mathbf{P}[i]$ || $\mathbf{F}[i] \leq \omega \mathbf{P}[i] \prod_{j \neq i}(1 - \mathbf{P}[j])$

5 swap($\mathbf{F}[1], \mathbf{F}[i]$); swap($\mathbf{P}[1], \mathbf{P}[i]$); swap($\mathbf{I}[1], \mathbf{I}[i]$);

6 $f_1 = \mathbf{F}[1]$; $p_1 = \mathbf{P}[1]$;

7 $f_2 = \sum_{i=2}^{r} \mathbf{F}[i]$; $p_2 = 1 - \prod_{i=2}^{r}(1 - \mathbf{P}[i])$;

8 β_2 =min($\frac{p_1 \cdot \omega - f_1}{p_2 p_1 \cdot \omega}$, 1); $\beta_1 = 1 - \beta_2$; $\omega' = \omega(1 - p_1 \beta_1)$;

9 (S_{11}, Γ_{11}) = PS($\mathbf{F}[1], \mathbf{P}[1], \mathbf{I}[1], 1, \beta\beta_1, \omega'$);

10 (S_{12}, Γ_{12}) = PS($\mathbf{F}[1], \mathbf{P}[1], \mathbf{I}[1], 1, \beta\beta_2, \omega'$);

11 (S_{21}, Γ_{21}) = PS($\mathbf{F}$[2$\sim$r],P[2$\sim$ r],I[2$\sim$ r],r-1,$\beta\beta_2, \omega'$);

12 (S_{22}, Γ_{22}) = PS($\mathbf{F}$[2$\sim$r],P[2$\sim$ r],I[2$\sim$ r],r-1,$\beta\beta_1, \omega'$);

13 (S_1, Γ_1)=Merge(S_{11},S_{22},Γ_{11},Γ_{22});

14 (S_2, Γ_2)=Merge(S_{21},S_{12},Γ_{21},Γ_{12});

15 **return**($S_1 \bigcup S_2, \Gamma_1 \bigcup \Gamma_2$);

Table 5.4: Pseudocode of a heuristic recursive algorithm for finding a scheduling of opportunistic forwarding strategies

priority, we put this flow at the first place of the traffic demand vector. We then split the candidates and the corresponding flows into two parts, part 1: $\mathbf{I}[1]/\mathbf{F}[1]$ and part 2: $\mathbf{I}[2{\sim}\mathrm{r}]/\mathbf{F}[2{\sim}\mathrm{r}]$. Next, we calculate the accumulated flows f_1 and f_2 in the two parts, and the corresponding forwarding candidate set reliability p_1 and p_2 (in lines 6 and 7). In line 8, we calculate the time fractions β_1 and β_2 corresponding to prioritization $\langle \mathbf{I}[1],\mathbf{I}[2{\sim}\mathrm{r}] \rangle$ and $\langle \mathbf{I}[2{\sim}\mathrm{r}],\mathbf{I}[1] \rangle$, respectively. Note that $\langle \mathbf{I}[1],\mathbf{I}[2{\sim}\mathrm{r}] \rangle$ implies the candidate $\mathbf{I}[1]$ has higher relaying priority than the group of candidates $\mathbf{I}[2{\sim}\mathrm{r}]$, and vice versa. Then we recursively call the function PS on $\mathbf{I}[1]$ and $\mathbf{I}[2{\sim}\mathrm{r}]$ (in lines 9 and 12). The returned S_{ij} is the set of forwarding strategies when part i is in the j^{th} place ($j = 1, 2$ indicates higher and lower priority, respectively). Then we combine the sequences in S_{11} and S_{22} to get S_1 which are sequences of candidates with $\mathbf{I}[1]$ having higher priority than $\mathbf{I}[2{\sim}\mathrm{r}]$ (in line 13). Similarly, we combine S_{21} and S_{12} with group of candidates $\mathbf{I}[2{\sim}\mathrm{r}]$ having higher priority than $\mathbf{I}[1]$ (in line 11). Finally, we return the whole series of prioritization by taking the union of S_1 and S_2.

Now we describe the Merge algorithm in Table 5.5. Assume both input (S_1, S_2, Γ_1 and Γ_2) and output (S and Γ) are stored in stacks. The basic idea of this Merge algorithm is to concatenate the sequence (corresponding to a prioritization) in the top of S_1 with that in the top of S_2 (in line 3) to create a new sequence (prioritization). The time fraction of this new sequence is the minimum of the time fractions of these two subsequences. After creating a new sequence, we pop the sequence with smaller time fraction, and update the time fraction of the other sequence by subtracting the used time fraction (in lines 5, 7, and 9). When all the sequences in S_1 and S_2 are popped out, a series of new sequences S and the corresponding time fraction vector Γ are returned (in line 11).

The computation complexity of Merge algorithm is $\Theta(|S_1| + |S_2|)$, where $|S_i|$ ($i = 1, 2$) is the number of sequences in S_i. For S_i with N receivers, we have at most $O(2^N)$ and at least $\Omega(1)$ sequences in it. So the complexity of the algorithm

```
(S, Γ)=Merge(S_1,S_2,Γ_1,Γ_2);
1  S = ∅; Γ = ∅;
2  while (S_1 ≠ ∅ || S_2 ≠ ∅)
3   push(S,top(S_1)|top(S_2));
4   if (top(Γ_1)>top(Γ_2))
5    push(Γ,top(Γ_2)); pop(Γ_2); pop(S_2); top(Γ_1)=top(Γ_1)−top(Γ_2);
6   else if (top(Γ_2)>top(Γ_1))
7    push(Γ,top(Γ_1)); pop(Γ_1); pop(S_1); top(Γ_2)=top(Γ_2)−top(Γ_1);
8   else
9    push(Γ,top(Γ_1)); pop(Γ_1); pop(S_1); pop(Γ_2); pop(S_2);
10 end while
11 return(S, Γ);
```

Table 5.5: Pseudocode of merging two prioritized sub-sets of candidates

PS is $O(2^{L-1})$ in the worst case and $\Omega(L)$ in the best case, where L is the number of forwarding candidates. We want to argue that although the worst-case complexity of this algorithm is exponential, in our simulations this algorithm runs much faster than $O(2^{L-1})$, and L is also a small number (less than the node degree in the network).

5.3.5.1 Correctness of the Heuristic Algorithm

This heuristic algorithm does not guarantee to return a feasible schedule of opportunistic forwarding strategies even when the traffic demand vector is schedulable. When this happens, we need to run the LP in Fig. 5.4 to get a feasible schedule. However, we will prove that this heuristic algorithm does return a feasible schedule for a schedulable traffic demand vector $\mathbf{f}$ when $|\mathbf{f}| \leq 2$. We have the following Proposition.

Proposition 5.3.3. *When the potential forwarding candidate number L is no greater*

than 2, any traffic demand vector $\mathbf{f} = [f_1, ..., f_L]$ *in the capacity region defined in Theorem 5.3.1 can be satisfied by the schedule obtained by the heuristic algorithm in Table 5.4.*

Proof. First, when $L = 1$, it's obvious that any f_1, s.t. $f_1 \leq R \cdot p_1$, is schedulable. Lines 1 and 2 in Table 5.4 deal with this case.

Second, when $L = 2$, as discussed previously, there are four forwarding strategies illustrated in Table 5.3. As opportunistic routing $\langle i_1, i_2 \rangle$ achieves the same effective rate on link l_1 as traditional routing $\langle i_1 \rangle$, from capacity point of view, we only need to think of the opportunistic routing case. Similarly, we only take into account $\langle i_2, i_1 \rangle$. So, there are two priority settings: $\mathcal{F}_1 = \langle i_1, i_2 \rangle$ and $\mathcal{F}_2 = \langle i_2, i_1 \rangle$. Then we can divide the active period τ_α into two subperiod τ_{α_1} and τ_{α_2}, s.t. $\tau_\alpha = \tau_{\alpha_1} + \tau_{\alpha_2}$, $\tau_{\alpha_1} \geq 0$ and $\tau_{\alpha_2} \geq 0$. Without loss of generality, assume in subperiod τ_{α_j}, we use priority $\mathcal{F}_j$, and let $\beta_j := \frac{\tau_{\alpha_j}}{\tau_\alpha}$, $1 \leq j \leq 2$. Then according to the effective forwarding rate defined in Eq. (4.1), we have the effective achievable rates on these two links as

$$R_1 = \beta_1 \cdot R \cdot p_1 + \beta_2 \cdot R \cdot p_1(1 - p_2) \tag{5.12}$$

$$R_2 = \beta_1 \cdot R \cdot p_2(1 - p_1) + \beta_2 \cdot R \cdot p_2 \tag{5.13}$$

Then we only need to prove, for any f_1 and f_2, s.t. $0 \leq f_1 \leq R \cdot p_1$, $0 \leq f_2 \leq R \cdot p_2$, and $f_1 + f_2 \leq R(1 - (1 - p_1)(1 - p_2))$, $\exists$ β_1 and β_2, s.t. $0 \leq \beta_1 \leq 1$, $0 \leq \beta_2 \leq 1$, and $\beta_1 + \beta_2 = 1$, to make $f_1 \leq R_1$ and $f_2 \leq R_2$.

With $f_2 \leq R_2$, $f_2 \leq R \cdot p_2$, Eq. (5.13) and $\beta_1 = 1 - \beta_2$, we have

$$0 \leq \beta_1 \leq \frac{R \cdot p_2 - f_2}{R \cdot p_1 p_2} \tag{5.14}$$

With $f_1 \leq R_1$, $f_1 \leq R \cdot p_1$, Eq. (5.12) and $\beta_2 = 1 - \beta_1$, we have

$$0 \leq \beta_2 \leq \frac{R \cdot p_1 - f_1}{R \cdot p_1 p_2} \tag{5.15}$$

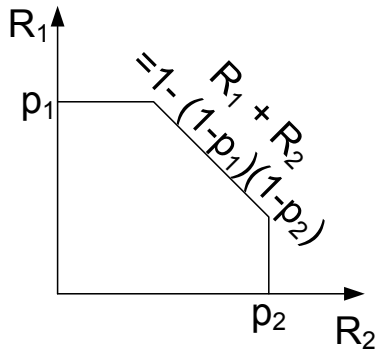


Figure 5.5: Capacity region for two forwarding candidates assuming broadcast rate $R = 1$.

By satisfying f_1, we set

$$\beta_2 = \min(\frac{R \cdot p_1 - f_1}{R \cdot p_1 p_2}, 1),\ \beta_1 = 1 - \beta_2 \tag{5.16}$$

By substituting Eq. (5.16) into Eq. (5.12) and (5.13), we can verify that $f_1 \leq R_1$ and $f_2 \leq R_2$. Note that, the setting of β_1 and β_2 makes inequality (5.14) and (5.15) hold. Eq. (5.16) exactly corresponds to line 8 in Table 5.4. So we proved the correctness of the heuristic algorithm PS for $L = 2$. □

The proof of the correctness of the heuristic algorithm also indicates that any normalized traffic demand vector in the capacity region shown in Fig. 5.5 is schedulable when $L = 2$.

5.3.5.2 An Example

Now we show an example to illustrate how the PS algorithm works. Assume in a CTS T_α a transmitter i has three forwarding candidates $\{i_1, i_2, i_3\}$, the corresponding normalized flow (normalized by λ_α) on each link l_j (j=1,2,3) is 0.2, 0.3, and 0.46, and the corresponding PRR on these links are 0.5, 0.6, and 0.8, respectively. Fig. 5.6 shows the running result of algorithm PS. In the first stage, f_1 is satisfied, and in the

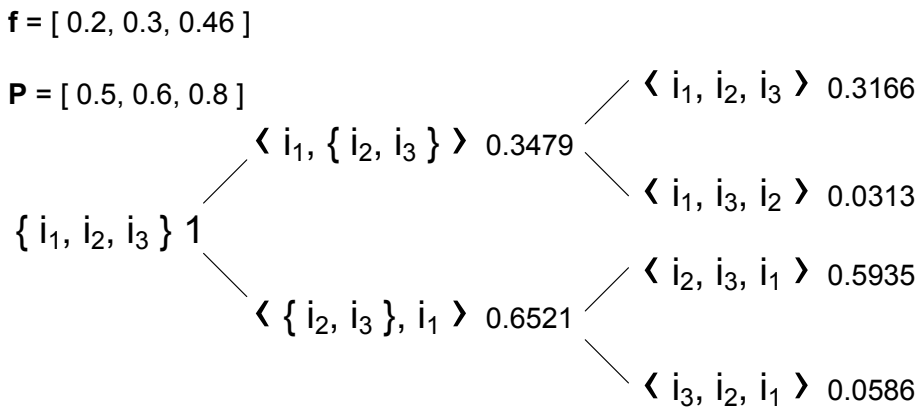


Figure 5.6: An example of opportunistic forwarding strategy scheduling for three forwarding candidates.

second stage f_2 is satisfied, then f_3. The time fraction β of each forwarding strategy is listed at the right of the strategy.

5.3.6 Maximum End-to-end Throughput of OR

Assume we have found all the CTS's $\{T_1, T_2...T_M\}$ in the network. At any time, at most one CTS can be scheduled to transmit. When one CTS is scheduled to transmit, all the nodes in that set can transmit simultaneously. Let λ_α denote the time fraction scheduled to CTS T_α ($\alpha = 1...M$). Then the maximum throughput problem can be converted to an optimal scheduling problem that schedules the activation of the CTS's to maximize the end-to-end throughout. Therefore, considering communication between a single source, n_s, and a single destination, n_d, with opportunistic routing, we formulate the throughput capacity problem between the source and the destination as a linear programming problem corresponding to a maximum-flow problem under additional constraints in Fig. 5.7.

In Fig. 5.7, $f_{ij}^{k\alpha}$ denotes the amount of flow on link l_{ij}^k in the CTS T_α. Recall that $\mathbf{E}$ is a set of all the wireless links in the extended graph G, and $\mathbf{V}$ is the set of all

$$Max \sum_{k=1}^{K} \sum_{l_{si}^k \in \mathbf{E}} \sum_{\alpha=1}^{M} f_{si}^{k\alpha} \tag{5.17}$$

$$s.t.$$

$$\sum_{k=1}^{K} \sum_{l_{ij}^k \in \mathbf{E}} \sum_{\alpha=1}^{M} f_{ij}^{k\alpha} = \sum_{k=1}^{K} \sum_{l_{ji}^k \in \mathbf{E}} \sum_{\alpha=1}^{M} f_{ji}^{k\alpha},$$

$$\forall\ i = 1...N, i \neq s, i \neq d \tag{5.18}$$

$$\sum_{k=1}^{K} \sum_{l_{is}^k \in \mathbf{E}} \sum_{\alpha=1}^{M} f_{is}^{k\alpha} = 0 \tag{5.19}$$

$$\sum_{k=1}^{K} \sum_{l_{di}^k \in \mathbf{E}} \sum_{\alpha=1}^{M} f_{di}^{k\alpha} = 0 \tag{5.20}$$

$$f_{ij}^{k\alpha} \geq 0, \quad \forall\ k = 1...K, l_{ij}^k \in \mathbf{E} \tag{5.21}$$

$$\sum_{\alpha=1}^{M} \lambda_\alpha \leq 1 \tag{5.22}$$

$$\lambda_\alpha \geq 0,\ \forall\ \alpha = 1...M \tag{5.23}$$

$$\sum_{\mathcal{C}} f_{ij}^{k\alpha} \cdot \phi_j \leq \lambda_\alpha R_i (1 - \prod_{\mathcal{C}} (1 - p_{ij}^k \cdot \phi_j)),$$

$$\mathcal{C} = \{j | l_{ij}^k \in \mathbf{E}, \psi_{ij}^{k\alpha} == 1\},$$

$$\forall\ v_i^k \in \mathbf{V},\ \alpha = 1...M, \forall\ \Phi(\mathcal{C}) \in \{0,1\}^{|\mathcal{C}|} \tag{5.24}$$

Figure 5.7: LP formulations to compute the capacity of OR in multi-radio/channel systems

the transceiver configurations. The maximization states that we wish to maximize the sum of flow out of the source, which is the accumulated flow on all outgoing links on all channels from the source in all CTS's. The constraint (5.18) represents flow-conservation, i.e., at each node, except the source and the destination, the amount of incoming accumulated flow is equal to the amount of outgoing accumulated flow. The constraint (5.19) states that the incoming accumulated flow to the source node is 0. The constraint (5.20) indicates that the outgoing accumulated flow from the destination node is 0. The constraint (5.21) restricts the amount of flow on each link to be non-negative. The constraint (5.22) represents that at any time, at most one CTS will be scheduled to be active. The constraint (5.23) indicates that the scheduled time fraction should be non-negative. In the constraint (5.24), $\Phi(\mathcal{C})$ is a vector of ϕ_j's with length $|\mathcal{C}|$. The constraint (5.24) states that the flows out of a transmitter in a CTS must be in the capacity region discussed in Section 5.3.3, that is, in any CTS, any sub-summation of the flow on usable outgoing links from the same transmitter is bounded by the corresponding forwarding set reliability timing the transmission rate.

The key difference of our maximum flow formulations from the formulations for traditional routing in [32,74,76] lies in the methodology we use to schedule concurrent transmissions. With OR, we virtually schedule the links from the same transmitter at the same time. While for traditional routing, any two links sharing the same sender can not be scheduled simultaneously. When a packet is not correctly received by the intended receiver but opportunistically received by some other neighboring nodes, traditional routing will retransmit that packet instead of making use of the correct receptions on the other receivers. However, OR takes advantage of the correct receptions. That's why OR achieves higher throughput than traditional routing. This LP formulation is also different from that in Chapter 4, where an opportunistic forwarding strategy is given and fixed at each node. We assume each transmitter can dynamically change its forwarding strategy in any CTS. So the solution of the

objective function (5.17) is the upper bound of the capacity between two nodes. The byproduct of the LP in Fig. 5.7 is a channel assignment and scheduling for OR in multi-radio/channel system. By applying the heuristic algorithm PS or LP in Fig. 5.4 on the normalized flow $\frac{f_{ij}^{k\alpha}}{\lambda_\alpha R_i}$ sharing the same transmitter v_i^k, we can further obtain the scheduling of the opportunistic forwarding strategies for the forwarding candidates of v_i^k in each CTS.

5.4 Performance Evaluation

In this section, we investigate the throughput bound of OR and TR in multi-radio multi-channel systems and compare the results with that in single-radio single-channel systems. The simulations are implemented in Matlab. We examine linear topology as well as rectangle topology. In each topology, we randomly deploy 12 nodes in the network. We select node n_1 at the left end or corner of the network as the destination, then calculate the throughput bound from other nodes to the destination using the LP formulations in Fig. 5.7. Therefore, there are 11 different source-destination pairs considered in the evaluation. In all the simulations, we assume the PRR is inversely proportional to the distance, and the interference range $R_I = 2R_T$. The transmission range R_T is set as 100 units. The performance metric is the normalized end-to-end throughput bound (by assuming the transmission rate is unit one).

Fig. 5.8 shows the throughput bound of OR and TR under different number of radios, channels and potential forwarding candidates in a linear topology with length of 300 units. In the legend, "TR" represents traditional routing, "OR" represents opportunist routing, "xRyC-z" represents x radios and y channels, with z maximal number of potential forwarding candidates. We can see that with the number of radios and channels increasing, the throughput of TR and OR are both increased. Generally OR achieves higher throughput than TR, and the multi-radio/channel capability has

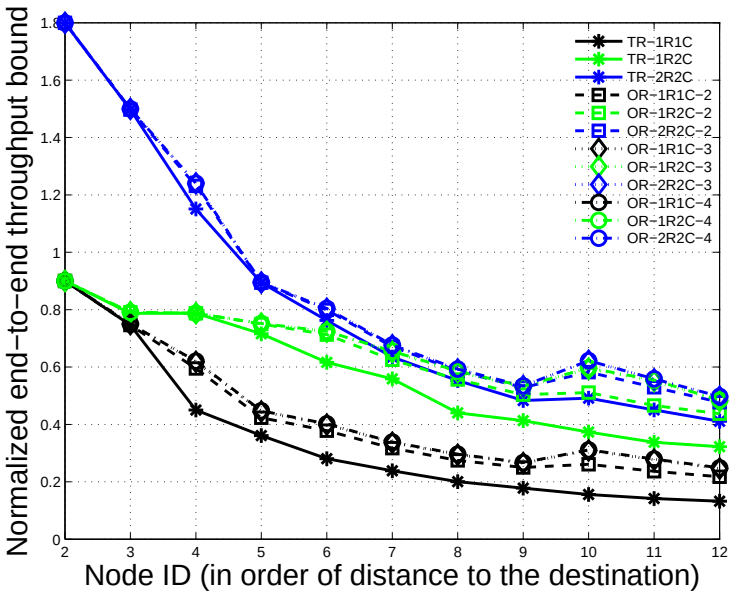


Figure 5.8: Normalized end-to-end throughput bound under different number of radios, channels and potential forwarding candidates in linear topology.

greater impact on the throughput of TR than OR. When the source is farther away from the destination, the OR presents more advantage than TR. The opportunistic forwarding by using multiple forwarding candidates do help increase the throughput. An interesting result is that, for node 8 to 12, the throughput of 1R2C case for OR is even greater than that of 2R2C case for TR. This result indicates that OR can achieve comparable or even better performance as TR by using less radio resource.

Another interesting observation is that the throughput gained decreases as the number of potential forwarding candidates increases. When this number is larger than 2, the gained throughput becomes marginal. This result is consistent with that found in [67, 68]. So it is not necessary to involve all the usable receivers of the transmitter into the opportunistic forwarding, and selecting a few "good" forwarding candidates is enough to approach optimal throughput.

Fig. 5.9 shows the throughput bound of OR under different number of radios,

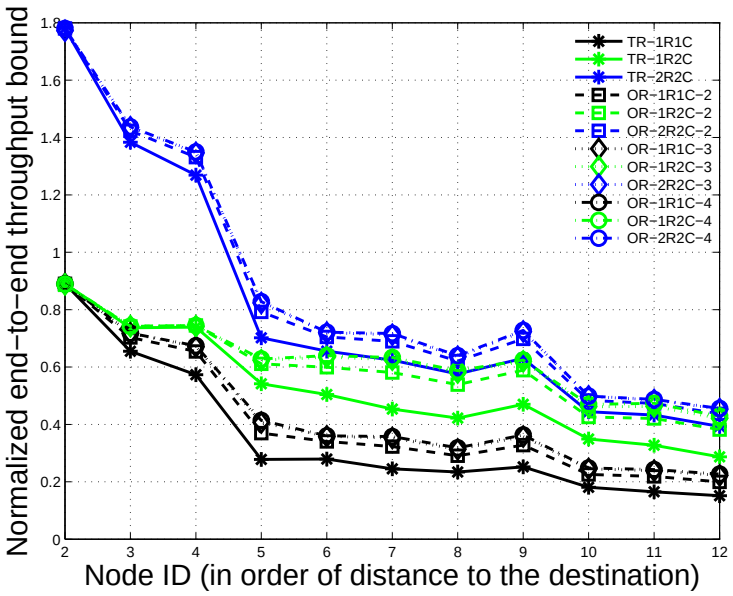


Figure 5.9: Normalized end-to-end throughput bound under different number of radios, channels and potential forwarding candidates in rectangle topology.

channels and potential forwarding candidates in a rectangle topology of 200 units × 300 units. We can see that the OR performance in the rectangle topology represents the same trend as that in the linear topology.

5.5 Conclusions

In this chapter, we proposed a unified framework to compute the throughput bound of opportunistic routing between two end nodes in single/multi-radio/channel multihop wireless networks. Our model accurately captures the unique property of OR that multiple outgoing links sharing the same transmitter can be virtually scheduled at the same time under particular rate constraints. We also studied the necessary and sufficient conditions for the schedulability of a flow demand vector associated with a transmitter in a concurrent transmission set. We further proposed an LP approach and a heuristic algorithm to obtain an opportunistic forwarding strategy schedul-

ing that satisfies a flow demand vector. Our methodology provides a framework to calculate the end-to-end throughput bound of OR and TR in multi-radio/channel multihop wireless networks, and can be used to study the OR behaviors (such as candidate selection and prioritization). Leveraging our analytical model, we found that OR can achieve comparable or even better performance than TR by using less radio resource.

Chapter 6

Medium Access Control for Opportunistic Routing - Candidate Coordination

One important and challenging issue in OR is candidate coordination. That is, in order to avoid duplication, we should ensure that only the "best" receiver of each packet forwards it. However, it is non-trivial to achieve this goal in an efficient way. The existing candidate coordination schemes have some inherent inefficiency such as high time delay at each one-hop transmission, potential duplicate forwarding, etc. Improperly designed coordination schemes will aggravate these problems and even overwhelm the potential gain provided by OR.

In this chapter, we carry out a comprehensive study on the candidate coordination in OR and propose a new scheme "fast slotted acknowledgment" (FSA) to further improve the efficiency of OR, which adopts a single ACK to confirm the successful reception and suppresses other candidates' attempts to forward the data packet with the help of channel sensing technique. We also confirm the benefit of our scheme by simulation. The result shows that FSA can decreases the average end-to-end delay up

to 50% when the traffic is relatively light that all the coordination schemes can still handle and can improves the throughput up to 20% under heavy traffic load where the other coordination schemes are already unable to delivery all the data packets.

The rest of this chapter is organized as follows. Section 6.1 describes the state-of-the-art coordination schemes in detail. Section 6.2 presents FSA's design and analysis, followed by Section 6.3, where evaluation and analysis of FSA's performance are presented. Section 6.4 concludes.

6.1 Existing Candidate Coordination Schemes

In this section, we review two state-of-the-art candidate coordination schemes: slotted acknowledgement and compressed slotted acknowledgement, and point out their potential vulnerability and inefficiency.

6.1.1 Slotted Acknowledgment (SA)

SA is proposed by Biswas and Morris in [10]. It applies a similar acknowledgment scheme as the one used in traditional 802.11, however, requires each candidate who has received the data packet to broadcast an ACK in different time slots according to their priorities. Instead of only indicating the success of reception, each ACK contains the ID of the highest-priority successful recipient known to the ACK's sender. All the candidates listen to all ACKs before deciding whether to forward the data packet, in case a lower prioritized candidate's ACK reports a higher prioritized candidate's ID. In order to protect all the ACKs from being interrupted by other transmissions, SA extends the Network Allocation Vector (NAV) in the MAC header of the data packet to reserve the channel for longer time. Thus the total coordination time for SA with n candidates is $n \times (T_{SIFS} + T_{ACK})$, where SIFS is short inter frame space [1]. This scheme has a serious vulnerability which makes it fail to work properly

in some scenarios. Taking one transmission with 3 candidates as an example in Fig. 6.1. Suppose that when the sender is transmitting, another node within the sender's transmission range, which is willing to transmit, does not hear the data packet clearly (for example, received corrupted one that cannot get the NAV value from the MAC header). At the same time, the highest-priority candidate also failed to receive the data packet. In this case, the first ACK is missing and the potential transmitter that does not update its NAV accordingly senses the channel to be clear for $2 \times T_{SIFS} + T_{ACK}$ which is obviously greater than DIFS (Distributed Inter Frame Space), the idle time needed before sending packet in 802.11 protocols [1]. Thus it

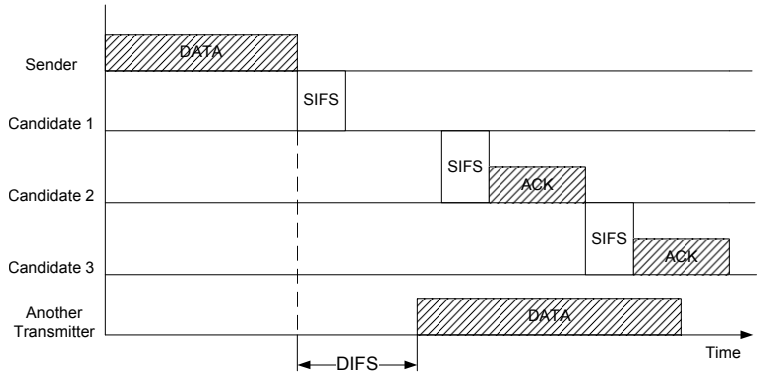


Figure 6.1: SA with first ACK missing

sends its own packet which will collide with the subsequent ACKs from $candidate_2$ and $candidate_3$ as illustrated in Fig. 6.1. Since no one hears a clear ACK, the consequence is that both $candidate_2$ and $candidate_3$ will forward the packet, which results in duplication, and the sender will unnecessarily retransmit the packet. The scenario shown above is not rare, especially in networks under heavy traffic loads.

6.1.2 Compressed Slotted Acknowledgment (CSA)

A.Zubow et. al. [86] try to alleviate the potential collision in SA by introducing the channel assessment technique and refine SA to a "compressed slotted acknowledg-

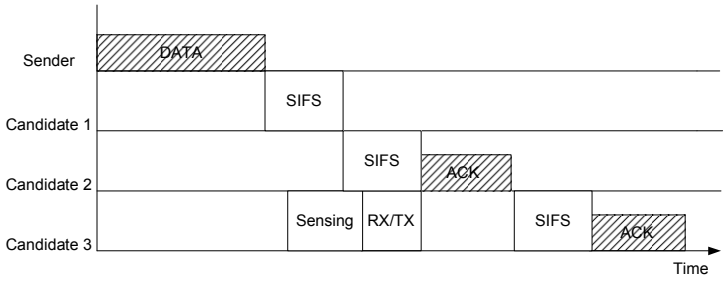


Figure 6.2: CSA with the first ACK missing where RX/TX is the turnaround time for radio to change from receive state to transmit state

ment". The general idea is as following: With a delay of SIFS after receiving the data packet, the highest-priority candidate sends its ACK out. From this time point, all other candidates who also successfully received the data packet sense the channel by received signal strength indicator (RSSI), a parameter in PHY layer. If the RSSI value increases significantly within the predefined detecting period determined by the priority, the ACK is considered as sent and they will continue to wait for their corresponding ACK slots before sending ACKs. Otherwise, if no such increase in signal strength is observed, the other candidates conclude that the highest priority candidate did miss the data packet. In that case, the second highest-priority candidate prematurely sends its ACK to compress the channel's idle space to be smaller than DIFS. All the other candidates behave in the same way as before except that all subsequent events happen earlier. Figure 6.2 depicts a case with 3 candidates. the use of channel assessment technique makes the SA's fixed ACK slots mechanism more flexible and gives CSA better performance on alleviating the potential collision. However, this detection-based scheme still requires multiple ACKs thus suffers from the same high coordination delay as SA.

6.2 DESIGN AND ANALYSIS OF FSA

6.2.1 Design of FSA

The main objective of FSA is to achieve an agreement among multiple candidates with lower coordination delay than SA and CSA. At the same time, FSA must be robust enough to deal with potential collision and unnecessary retransmission. Since all the inefficiencies in SA and CSA are mainly due to the use of multiple ACKs, we adopt a single ACK in FSA, which will be sent by the highest priority candidate in the set of successful receivers. This single ACK plays two roles. On one hand, it informs the sender of the successful reception, which is the same as SA and CSA; On the other hand, it suppresses all the other lower priority candidates' attempts to forward the data packet. This is different from the ACKs in SA and CSA schemes which are to help candidates share the information about the reception status. Accordingly, we also choose to use channel assessment technique to detect the appearance of ACK.

The FSA works as follows: Each candidate waits for $T_{SIFS} + (n-1) \times T_{Sensing\_Slot}$ before deciding whether it should broadcast ACK, where n is its priority order in the candidate set. So with a time delay of SIFS after the data packet was received, the highest-priority candidate sends out an ACK. From that point in time, all the other candidates detect the channel for a $T_{Sensing\_Slot}$ time to tell whether they detect this ACK. If the answer is positive, they stop detecting the channel and simultaneously suppress their own attempt of sending ACK and forwarding the data packet. Otherwise, if they did not detect any signal within this period, they think the highest priority candidate missed the data packet. In that case, the second highest priority candidate takes the responsibility of sending ACK in the beginning of the second $T_{Sensing\_Slot}$ and all the remaining lower priority candidates continue to monitor this ACK within this time. The coordination process goes on like that until some successful receiver finally sends an ACK in the channel.

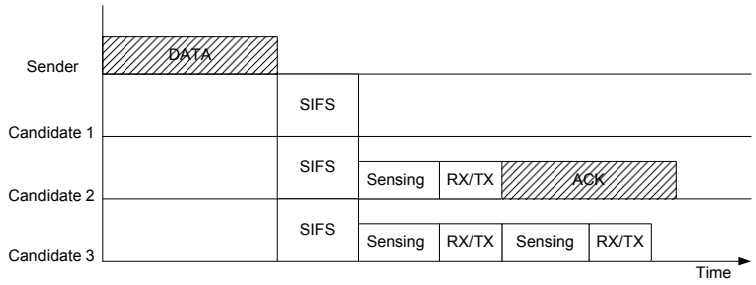


Figure 6.3: FSA with the first ACK missing

An example of a transmission with 3 candidate nodes is illustrated in Figure 6.3. During this one-hop transmission, $candidate_1$ failed to receive the packet. Thus the $candidate_2$ and $candidate_3$ detect nothing in the first $T_{Sensing\_Slot}$ time. Then $candidate_2$ thinks itself is the highest successful receiver and sends ACK at the beginning of the second $T_{Sensing\_Slot}$. $candidate_3$ detects this ACK and immediately suppresses itself. The total coordination time for FSA with n candidates is: $T_{SIFS} + T_{ACK} + (n-1) \times T_{Sensing\_Slot}$ compared with SA and CSA's $T_{SIFS} + n \times T_{ACK}$. Since the $T_{Sensing\_Slot}$ is far less than T_{ACK} (for 802.11b, the former is 15 micro-seconds while the latter is more than 200 micro-seconds), FSA can significantly reduce the time cost for candidate coordination.

6.2.2 Analysis

The key difference of FSA from SA and CSA is that it only uses single ACK to suppress other potential forwarders and acknowledge the sender. FSA uses channel assessment technique to infer some raw information such as whether some packet is transmitting rather than more detailed information like the content of the packet. At first sight, it seems to be less reliable than SA and CSA. However, it is not true. On the opposite, just because the information required by FSA is raw, it can be obtained more easily and reliably which makes the whole scheme works well in a wireless environment, where the most distinct property is unreliability. Suppose one

transmission with two candidates A and B where A possesses higher priority and both candidates got the data packet. If the link between A and B is not good when A is transmitting its ACK, then the ACK received by B may be corrupted. In SA and CSA coordination schemes, node B needs to get the ID of the highest priority successful recipient known to the ACK's sender (in the case it is A itself) from the received ACK. However, with this corrupted ACK, B will fail to do that and consider itself as forwarder which leads to duplicate forwarding and unnecessary retransmissions. However, in FSA, B just needs to know the happening of a transmission instead of the detailed content within the received packet. Even the ACK is corrupted, B may still be able to infer that there is an ACK transmission from higher-priority node A. Thus FSA is more robust in this case.

Another seemingly weakness of FSA is that a single ACK would not be reliable enough to ensure the sender to receive it correctly. Because once this single ACK is lost, the sender needs to retransmit the data packet unnecessarily. With multiple ACKs like SA and CSA, the sender would hear at least one clear ACK with high probability. However, it is also not true because of the following reasons. (1) If the data packet (which is generally longer than ACK and sent in higher rate) has already been received by the corresponding candidate successfully, the subsequent ACK sent along the reverse direction in a lower rate (1Mbps for 802.11b) will be received by the sender successfully with very high probability [54]. (2) The other ACKs except the first one in SA and CSA are sent in relatively long intervals (several ACK slots) after receiving the data packet. The link states may have already changed at that time. Then those following ACKs may not be able to be received correctly by the sender. So the added reliability by those extra ACKs is quite limited. In another word, single ACK is already strong enough and multiple ACKs are not indispensable.

The real potential vulnerability of FSA is its dependence on the precision of channel assessment technique. For example, if the detecting node considered some other

interferences to be ACK sent by some higher priority candidate, it will falsely suppress itself from forwarding the packet. It is also possible that in some situation the detecting node fails to sense the transmission of ACK from higher priority candidate and then sends its own ACK which will collide with the transmitting one. However, this dependence problem can be greatly alleviated through careful design. For the first case, we make the whole coordination process highly synchronized and also introduce more precise channel assessment technique (see details in following subsection). Thus such probability will be constrained in a rather low level. Even if this scenario indeed happened, the consequence is just that those lower priority candidates suppress themselves "over cautiously" and cause the sender to retransmit the packets unnecessarily. This will make opportunistic routing behave like the traditional routing. For the second case, just as we described in the beginning of this section, this possibility will be very low because the forwarding candidates are usually in each other's carrier sensing ranges. If this case really happens, the consequence will be duplicate forwarding and unnecessary retransmission, which has serious impact on the performance of OR protocols. However, we should notice that this false detection results in the same consequence in CSA scheme, which means FSA will not introduce extra chance for duplicate forwarding in the worst case.

6.2.3 More on Channel Assessment Techniques

Carrier Sensing Multiple Access with Collision Avoidance (CSMA-CA) is de facto medium access control protocol for 802.11 WLAN. It follows the LBT (Listen Before Talk) principle and works in time-slots manner, which requires the sender to sense the channel status within one time slot before sending packets. Such sensing mechanism is called clear channel assessment (CCA) [1]. Generally speaking, CCA performance could be characterized by a pair of detection and false alarm probabilities (P_d and P_{fa}) in which P_d refers to the probability of detecting the channel to be busy when

the channel is indeed busy and P_{fa} refers to the probability of detecting the channel to be busy when the channel is actually idle. There is an inherent trade-off between P_d and P_{fa} with the constraint of limited detection time [79]. CCA module can be implemented in two ways:

1. Energy detection (ED)

   ED-based CCA is a simple non-coherent detection approach. It integrates the square of the incoming signal from the radio front end during the CCA window to get an average signal strength, then compares it with a predefined threshold level which represent the normal background noise and make a judgment. The main advantage of ED-based CCA is simplicity and the main disadvantage is relatively poor detection reliability (especially in 802.11b/g, which works on 2.4 GHZ, coexisting with other technologies such as microwave ovens, Bluetooth devices, etc.)

2. Preamble detection (PD)

   PD-based CCA tries to use the correlation of well-knowned preambles with the received signal to detect the presence of a packet. It can be implemented by a cross-correlation based matched filter which is more complex. Since it can fully take advantage of the processing gain, thus has a more enhanced reliability. The main disadvantage of PD-based CCA is that it needs to run continuously thus brings relatively high energy cost.

Both detection methods are supported by most of the current wireless cards and can promise a P_d no less than 99% with the CCA window specified by IEEE 802.11 standards [1]. However, the PD-based CCA outperforms ED-based CCA in P_{fa}, especially in noisy scenarios. Thus in FSA we choose PD-based CCA technique. Another reason that we prefer PD-based CCA is that its main disadvantage can be avoided in the scenario of coordination in OR protocols. Because what we need to

know is that if any ACK being sent during the CCA window, and do not care the channel state that is out of this period. Thus we don't require the PD module running continuously and it can be turned on only when the MAC layer requires a CCA from the PHY layer, just as what ED module does.

6.3 Simulation Results and Evaluation

In this section, we evaluate and compare the performance of FSA with SA and CSA in GloMoSim [73]. We also introduce a perfect detection-based scheme, IDEAL, as the baseline for comparison. The IDEAL scheme is the same as FSA, except that all the ACKs in IDEAL are 100% reliable and the detection judgement is 100% precise. Since our focus is on the efficiency of different coordination schemes given the same candidate set and the corresponding forwarding priorities, we use an existing candidate selection algorithm based on node's geographic locations and adopt the local metric expected one-hop throughput (EOT) [70] to select candidates. We will elaborate this local metric in Chapter 7.

Because the existing popular network simulators, such as NS-2 [46], OPNET [47], GloMoSim [73], have not implemented the PHY layer's function like energy integration or matched filter module currently, we have done some modification to the PHY layer in GloMoSim and make the detection judgment based on the following probability model. We define the CCA error floor [79] at the optimal threshold, which can be achieved by equating $1 - P_d$ and P_{fa} where P_d is detection possibility and P_{fa} is the false alarm possibility. Then the CCA error floor for ED-based CCA and PD-based CCA can be expressed in terms of the Q function [37]:

$$P_{CCA\_ef\_ed} = Q(\sqrt{N}\frac{SNR}{1+\sqrt{1+2SNR}})$$
$$P_{CCA\_ef\_pd} = Q(\sqrt{\frac{N}{2}SNR})$$

We define the following performance metrics.

- Throughput: the ratio of the number of received bits to the whole session time.
- Delay: the per packet end-to-end time delay from the packet being sent out until it reaches the destination.
- Packet deliver ratio: the number of successfully received packets over the number of sent packets.
- Number of transmissions: the total number of data transmissions happened during the whole simulation time.
- Duplicate deliver ratio: the number of duplicate packets received at all the destinations over the total number of received packets.
- Retransmission ratio: the transmission number needed for a successful one-hop forwarding.

The simulation results of all metrics except for the number of transmissions are averaged over 25 flows under 5 simulation runs with different seeds.

6.3.1 Simulation Setup

We developed a simulation environment with Glomosim. The MAC protocol is based on 802.11b, however, with some modifications. Since the source code of SA and CSA schemes are not publicly available, we implemented our own version. Table 6.1 lists all the related simulation parameters.

Table 6.1: simulation parameters

Simulation Parameter	Value
number of nodes	50
stationary or dynamic	stationary
data transmission rate	11Mbps
ACK transmission rate	1Mbps
Retry limit	5
Collision window	31..1023
Radio sensing threshold for data	-100dbm
Radio receiving threshold for data	-83dbm
Radio sensing threshold for ACK	-100dbm
Radio receiving threshold for ACK	-91dbm
pathloss model	two-ray
fading mode	rician
rician k factor	4
radio reception SNR	10
Hello packet intervial	1s
Size of candidate set	3
CCA window	15μs
SIFS	10μs
Radio receive/transmit turnaround time	5μs

Table 6.2: average number of neighbors per node and average hops per packet under different network densities

Terrain Size	Neighbors	Hops
1400	12.42	2.17
1500	10.90	2.50
1600	9.65	2.67
1700	8.60	2.96
1800	7.79	3.17

All these 50 nodes are randomly uniformly distributed in a $d \times d$ m^2 square region where d=1400,1500,...,1800. The corresponding average number of neighbors per node and average hop counts per flow are listed in Table 6.2.

We randomly choose 25 communication pairs in the network. The sources are CBR (constant bit rate) and each packet being 512 bytes long. UDP is used at the transport layer. Each communication session lasts 120 seconds. Before all the transmissions start, the simulation environment will go through a 30 seconds' warm-up phase, during which each node sends out "Hello" packet periodically to learn the neighbors information and this learning process lasts through the simulation.

6.3.2 Simulation Results and Evaluation

6.3.2.1 Delay

Figure 6.4 shows the average per packet end-to-end time delay of SA, CSA, FSA and IDEAL. In order to make a fair comparison, we set the packet interval of all the data flows to be 120 milliseconds which promises all the schemes can handle the traffic demand (in this case, all the protocols achieve 100% delivery ratio and almost the same average per flow throughput of 34k bps, thus we will not show the performance comparison of these two metrics in this setting). We see that SA has the

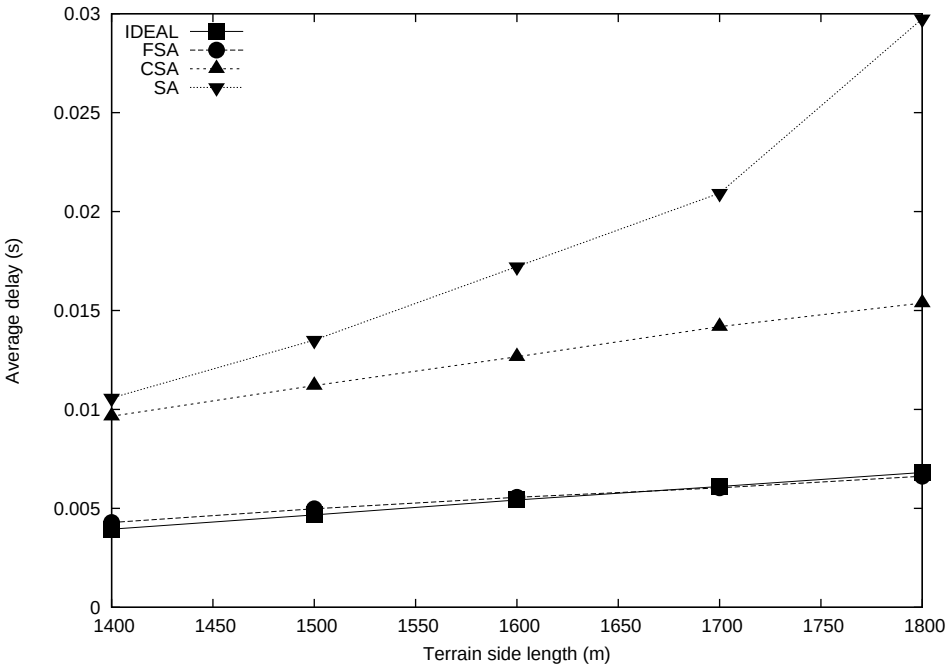


Figure 6.4: Average per packet end-to-end time delay

highest delay value under all terrain side lengths and CSA performs slightly better. FSA achieves far lower time delay than these two schemes and very close to the performance of IDEAL, which has the lowest delay value. From this result, firstly, we can get the conclusion that the use of channel assessment technique indeed can alleviate the potential collision problem caused by the ACK's unexpected missing. Secondly, we also notice that the time delays for CSA, FSA and IDEAL grow very slow as the increasing of terrain side length. This demonstrates that applying the channel assessment technique can also make the delay more stable under different network densities. Finally, we observe that FSA achieves less than half time delay of CSA under all the terrain side lengths. This reduction in time delay can be mainly attributes to the design of single ACK.

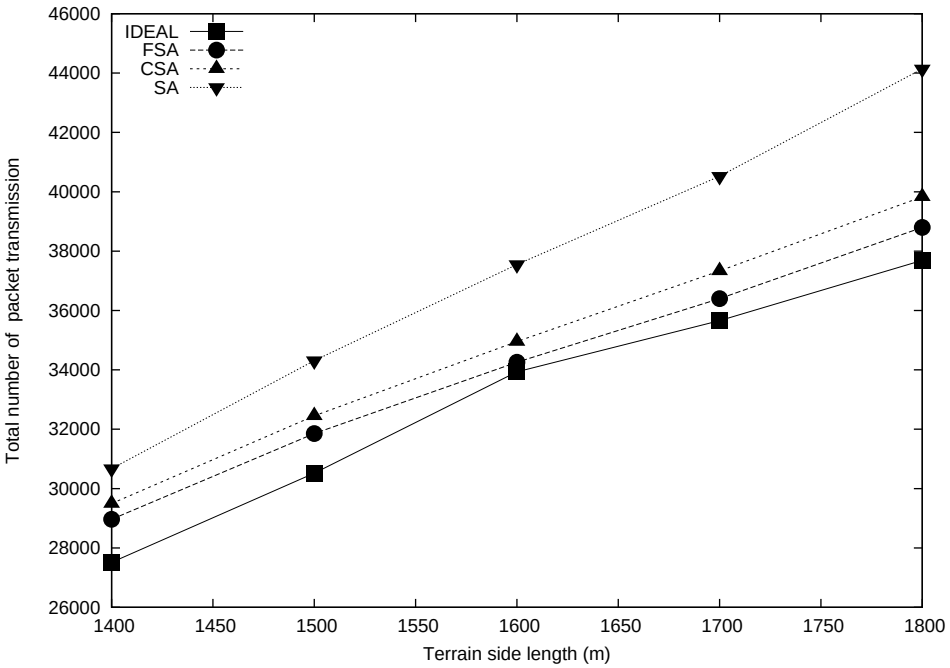


Figure 6.5: Total number of transmissions needed for delivering all the data flows

6.3.2.2 Number of transmissions

Figure 6.5 shows the total number of data transmissions during the simulation time. We see that as the increase of terrain side length, all schemes need more number of transmissions to deliver these data flows. This can be explained by the simultaneous increment of average hop count shown in Table 6.2. We also can observe that FSA need less number of transmissions than CSA and SA. This proves that FSA not only can greatly reduce the time cost for coordination process, but also can achieve better coordination reliability, which contributes to the reduced number of transmissions.

6.3.2.3 Duplicate deliver ratio and average retransmission ratio

The per packet duplicate ratio shown in Figure 6.6 and average retransmission ratio shown in Figure 6.7 can further demonstrate that FSA is more reliable. In Figure 6.6, we see that of all the data packets received successfully by the destinations of these

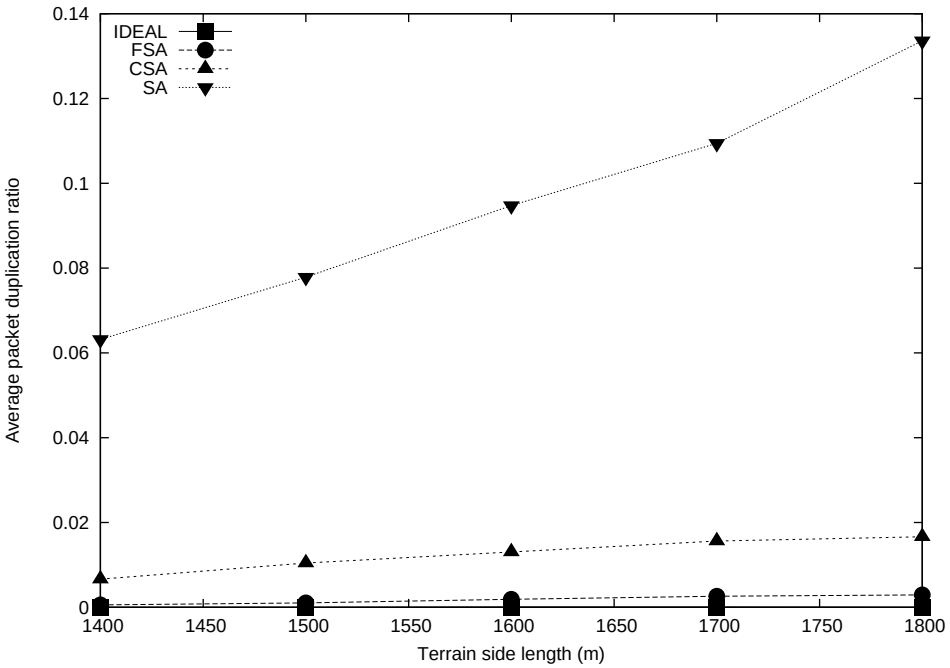


Figure 6.6: Average per packet duplicate ratio counted in the final receivers

data flows, there are approximately 6%-14% duplicated ones for SA and 0.4%-2% for CSA under different terrain side lengths. However, the duplicate ratios for FSA are almost zero under all terrain side lengths, which are very close to the performance of IDEAL. This confirms our analysis in section 6.2.2 which concludes that the probability for candidates in FSA to miss the presence of a higher priority candidate'ACK and result in duplicate forwarding is very low, but the probability for candidates in CSA to receive corrupted ACKs from other candidates and leads to duplicate forwarding is not negligible. In Figure 6.7, we see that IDEAL achieves an average retransmission ratio of approximately 1.01 under all terrain side lengths. Since the ACKs in IDEAL scheme are exempt from fading or interference, the only reason for retransmission in IDEAL is because all the candidates fail to receive the data packet. Such low retransmission ratio shows that OR schemes with multiple candidates indeed can greatly increase the forwarding reliability. We also notice that FSA's performance is close to the IDEAL, with an average higher ratio of 0.5%. This shows that the use of a single

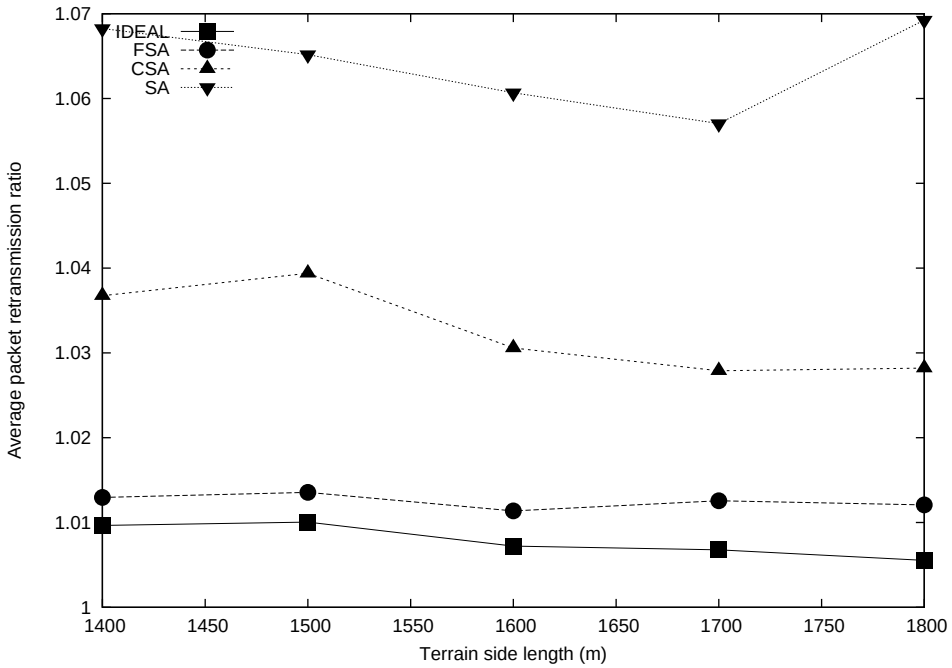


Figure 6.7: Average one-hop retransmission ratio

ACK in FSA is sufficiently reliable to acknowledge the sender.

6.3.2.4 Packet delivery ratio and throughput

In order to evaluate the throughput performance of all the schemes, we set the packet interval of all data flows to be 70 milliseconds, which makes a relatively heavy traffic load. From Figure 6.13 we see that SA and CSA are unable to handle the traffic demand and can only achieve packet delivery ratio of 87% - 81% and 81% - 74% respectively under different terrain side lengths. However, FSA still performs well and achieves 100% delivery ratio under different terrain sizes, just like IDEAL. Since our throughput metric is the ratio of the number of received bits to the whole session time and all the schemes have the same simulation time, thus the throughput is proportional to the packet delivery ratio. From Figure 6.8 we can see that FSA can achieve a throughput of approximately 54k bps, with an average gain of 12.5%-20% compared with CSA's throughput under different terrain side lengths. The

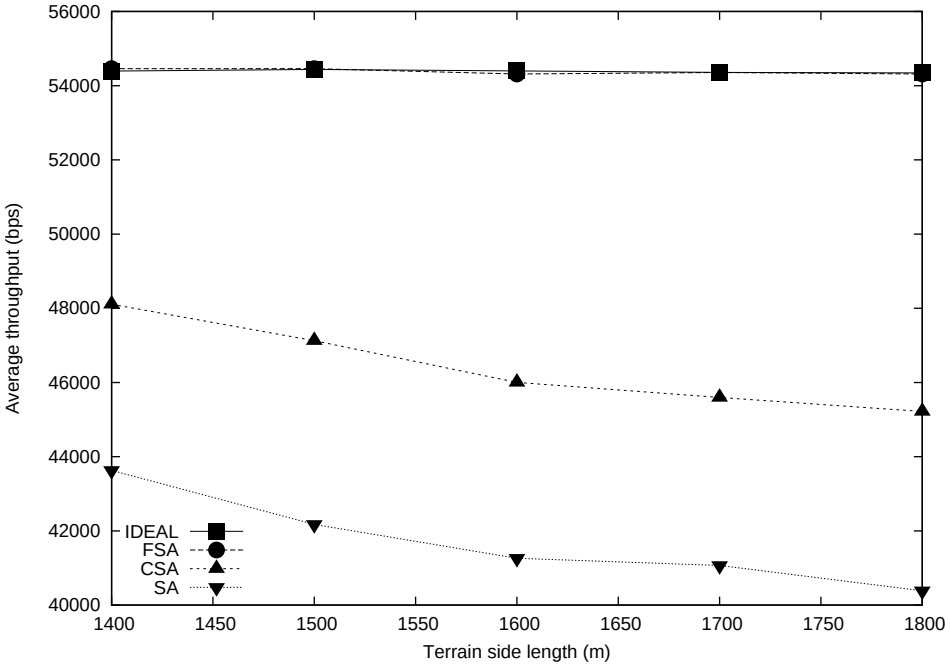


Figure 6.8: Average per flow throughput

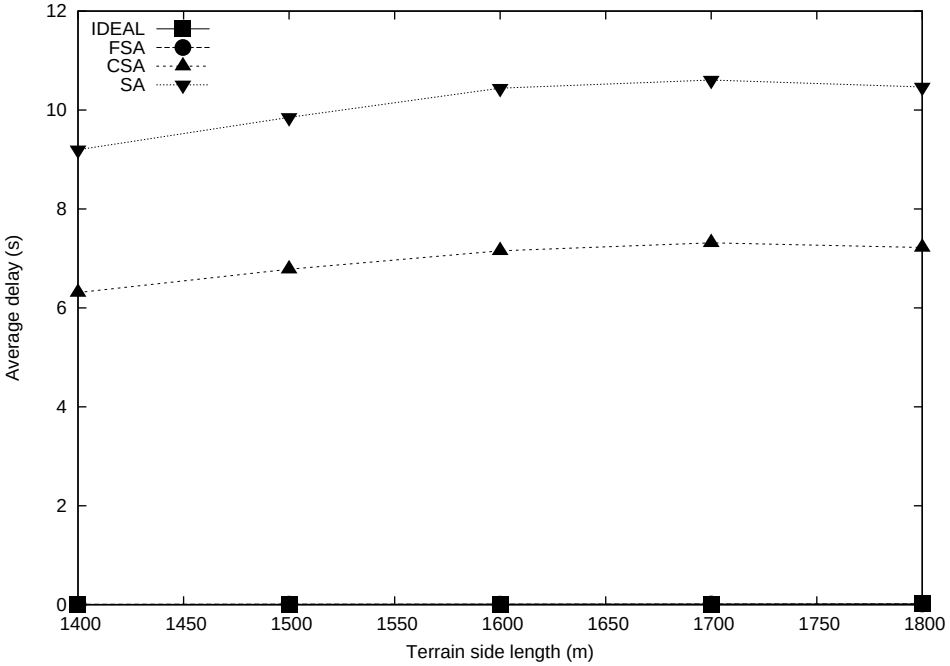


Figure 6.9: Average per packet end-to-end time delay

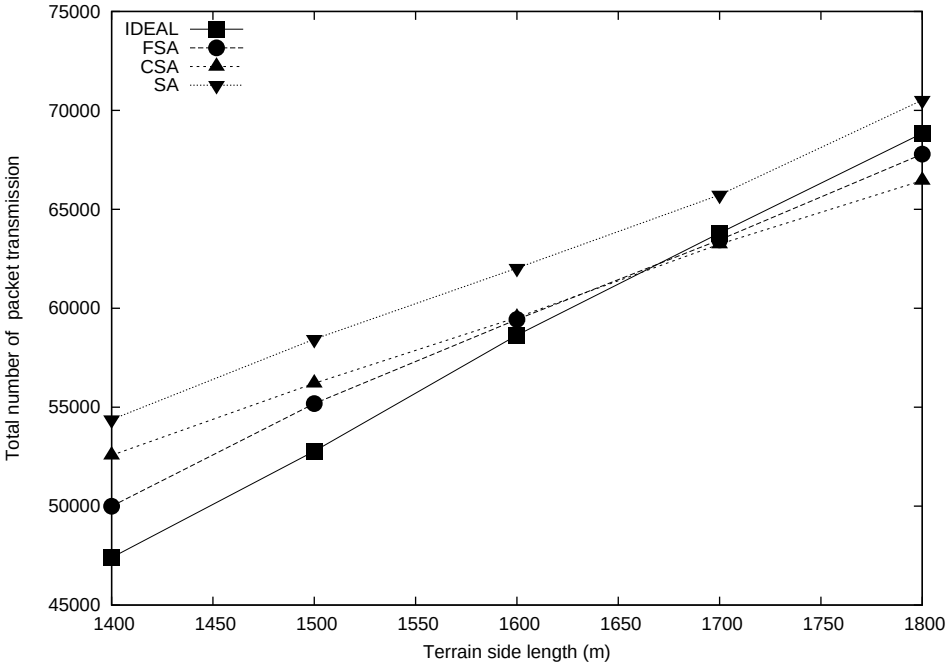


Figure 6.10: Total number of transmissions needed for delivering all the data flows

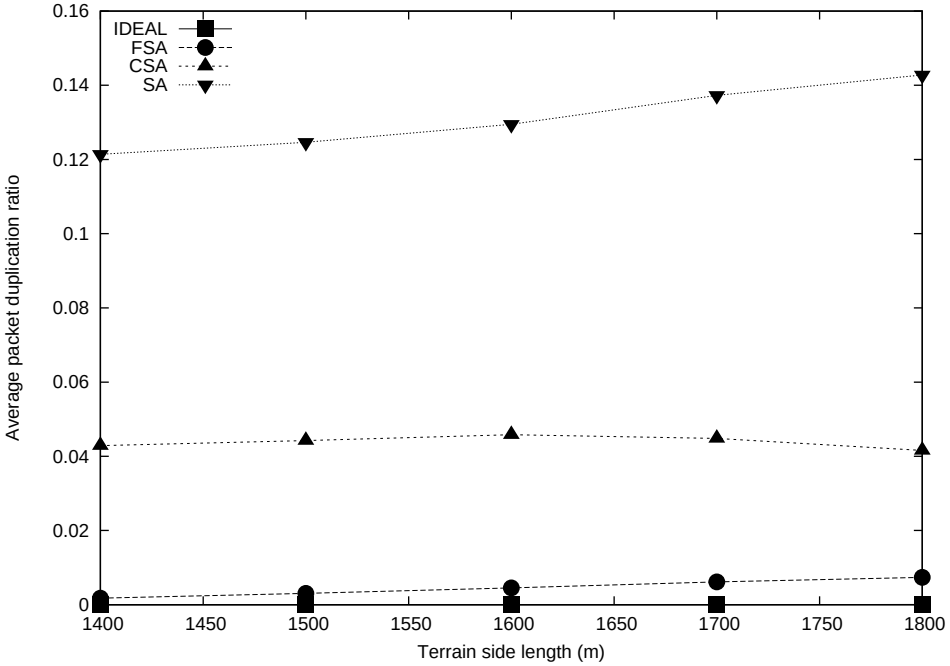


Figure 6.11: Average per packet duplicate ratio counted in the final receivers

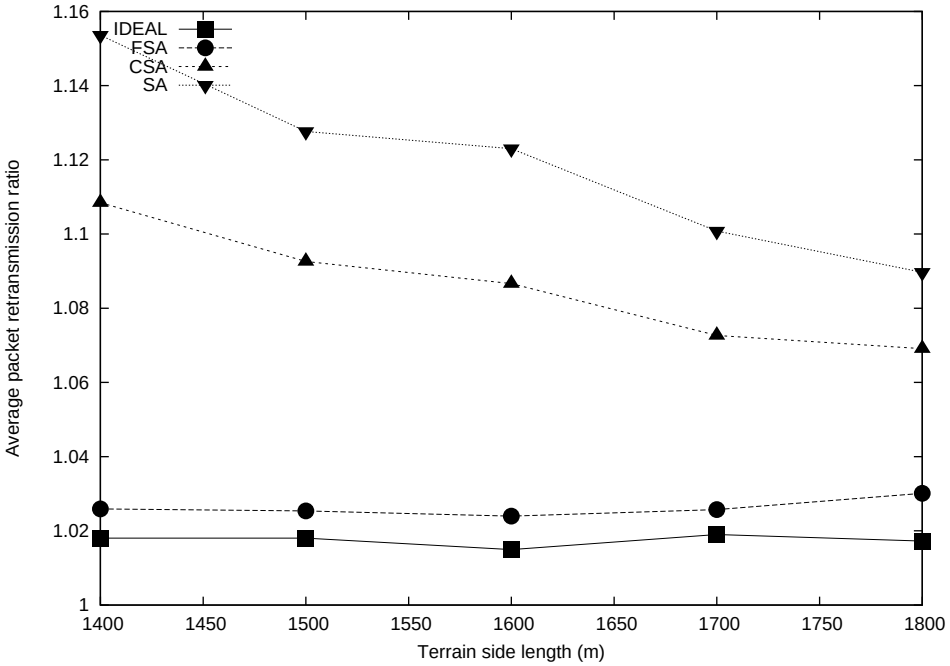


Figure 6.12: Average one-hop retransmission ratio

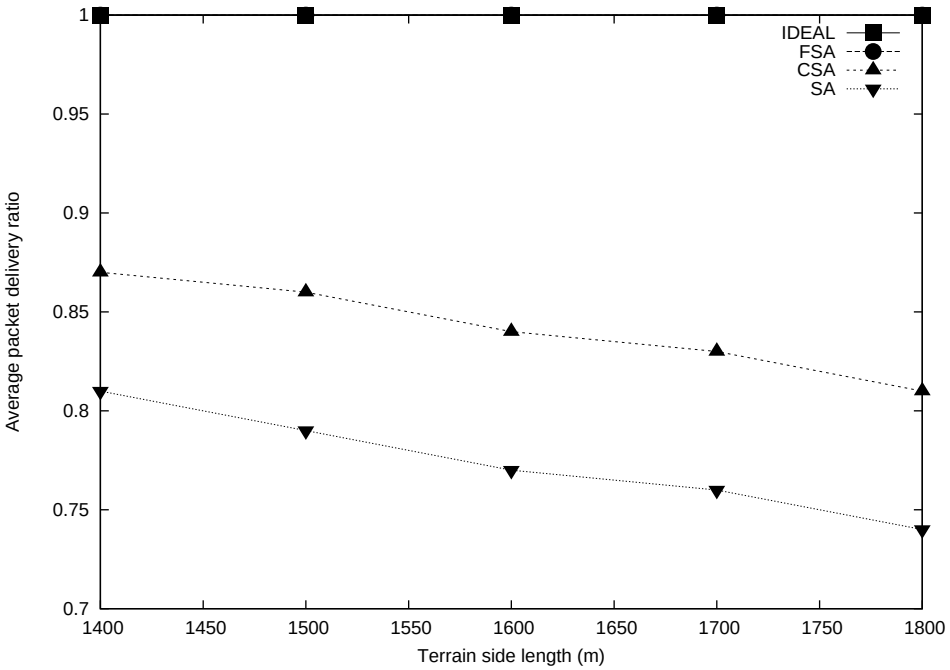


Figure 6.13: Average per flow delivery ratio

reason is that FSA takes less medium time in each hop's transmission and thus possesses a higher throughput capacity and the traffic demand in this setting is still within its capacity range. The significant higher time delay of SA and CSA shown in Figure 6.9 is also due to the fact that the traffic demand under this setting is beyond the throughput capacity of these two schemes, which cause each packet to suffer a long waiting time in the packet queue of every intermediate relay node. This long queueing delay further aggravates the duplication and retransmission problems, which can be observed in Figure 6.11 and Figure 6.12. From Figure 6.10 we observe that SA still has the highest number of transmissions during the simulation, but it can achieve the lowest throughput. This is not strange because the potential collision problem is exacerbated under heavy traffic load. However, we also observe that the total transmission number of FSA and IDEAL is higher than CSA under terrain side lengths of 1700m,1800m. This "abnormal" case also proves that nodes in FSA and IDEAL are more positive in transmitting rather than waiting in queue or backing off.

6.4 Conclusions

In this chapter, we analyzed the coordination problem in opportunistic routing. Based on these analysis, we proposed a new coordination scheme"fast slotted acknowledgment" (FSA) which fully takes advantage of the channel detection approach to meet an agreement among multiple candidates. We compared FSA with those state-of-the-art schemes and simulation results show that it achieves better performance in all the metrics, especially in time delay. The simulation also validated that FSA can achieve similar performance as ideal coordination where relay priority can be ensured and duplicate packet forwarding is avoided.

Chapter 7

Geographic Opportunistic Routing Protocol Design

In this chapter, we endeavor to study the impact of multiple rates, candidate selection, prioritization and coordination on the throughput of GOR by considering the protocol overhead. We introduce a local metric, *expected one-hop throughput* (EOT), to balance these factors. We further propose a rate adaptation and candidate selection algorithm to approach the local optimum of this metric. Simulation results show that MGOR incorporating the proposed algorithm achieves better throughput and delay performance than the corresponding opportunistic routing and geographic routing at any single rate.

The rest of this chapter is organized as follows. We discuss the impacts of multirate capability, forwarding strategy and candidate coordination delay on the throughput of opportunistic routing in Section 7.3. The local metric is introduced in Section 7.4. We propose the heuristic algorithm in Section 7.5. A multirate link quality measurement mechanism is proposed in Section 7.6. Simulation results are presented and analyzed in Section 7.7. We draw the conclusions in Section 7.8.

7.1 System Model

In this chapter, we consider the local MGOR scenario as the example in Fig. 7.1. Assume node S, i.e., the sender, is forwarding a packet to a remote destination D. S can transmit the packet at J different rates $R_1, R_2, ..., R_J$. Each rate corresponds to a **effective communication range**, within which the nodes can receive the packet sent by S with some non-negligible probability which is larger than a threshold, e.g., 0.1. The **available next-hop node set** $\mathcal{C}_j$ $(1 \leq j \leq J)$ of node S under a particular transmission rate R_j is defined as all the nodes in the communication range of S that are closer to D than S. We denote the nodes in $\mathcal{C}_j$ as $s_{j_1}, s_{j_2}, ..., s_{j_{M_j}}$, where $M_j = |\mathcal{C}_j|$. Define the **packet advancement** as d_{j_m} $1 \leq m \leq N_j$ in equation (7.1), which is the Euclidian distance between the sender and destination ($dist(S, D)$) minus the Euclidian distance between the neighbor s_{j_m} and destination ($dist(s_{j_m}, D)$).

$$d_{j_m} = dist(S, D) - dist(s_{j_m}, D) \tag{7.1}$$

Then at each rate R_j, each node in $\mathcal{C}_j$ is associated with one pair, (d_{j_m}, p_{j_m}), where p_{j_m} is the data packet reception ratio (PRR) from node S to s_{j_m}. Note that for different data rates, the PRR from node S to the same neighbor may be different. Let $\mathcal{F}_j$ denote the **forwarding candidate set** of node S at rate R_j, which contains the nodes that participate in the local opportunistic forwarding. Note that, here $\mathcal{F}_j$ is a subset of $\mathcal{C}_j$, while in the existing pure opportunistic routing schemes [84, 85, 9], $\mathcal{F}_j = \mathcal{C}_j$.

7.2 Candidate Coordination Mechanism

In this chapter, we use the "fast slotted acknowledgement" (FSA) mechanism, which is proposed in Chapter 6, to ensure the relay priority among the candidates. We briefly reiterate FSA as follows. When the channel is idle for a DIFS (distributed

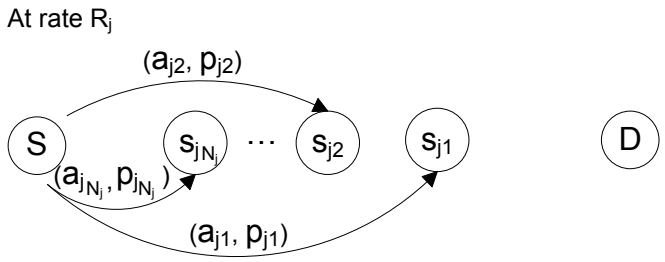


Figure 7.1: Node S is forwarding a packet to a remote destination D with transmission rate R_j.

inter-frame space), the sender broadcasts the data packet at the selected rate. In the header of the packet, the intended MAC addresses of the forwarding candidates and the corresponding relay priorities are identified. If the first-priority candidate receives the packet correctly, it broadcasts an ACK (acknowledgement packet) with a delay of SIFS (short inter-frame space) after the successful data reception. The ACK is used for informing the sender of the data packet reception as well as suppressing lower-priority candidates from forwarding duplicated copies. If the first-priority candidate does not receive the packet correctly, it just remains silent. For the second-priority candidate, it sets a waiting period from T_{SIFS} to $2T_{SIFS} - T_{rx/tx}$ after it received the data packet correctly, where T_{SIFS} and $T_{rx/tx}$ is the time duration of SIFS and radio receive/transmit status turnaround delay, respectively. If within the waiting period, it senses there is a significant signal strength increase in the channel, the ACK packet is considered as sent (It is not necessary that the first-priority candidate successfully receives the packet.) Then it just drops the received packet. On the other hand, if no such increase in signal strength is observed, the second-priority candidate conclude that the highest prioritized candidate did miss the data packet. So the second-priority candidate will turn around its radio from receiving status to transmitting status, and send out the ACK with $2T_{SIFS}$ delay after it received the packet. Generally, the i^{th}-priority ($i > 1$) candidate which receives the data packet correctly will set a

waiting period as $i \times T_{SIFS} - T_{rx/tx}$ after the data packet reception. If it detects a signal strength increase in this period, it will suppress itself from forwarding the packet; otherwise, it will send out an ACK at $i \times T_{SIFS}$ to claim its reception. We emphasize that although the throughput will be analyzed based on this specific MAC mechanism in this chapter, the analysis methodology and framework apply to other MAC schemes.

7.3 Impact of Transmission Rate and Forwarding Strategy on Throughput

Both transmission rate and forwarding strategy (including candidate selection, prioritization and coordination) will affect the throughput of MGOR.

The impacts of transmission rate on the throughput of opportunistic routing are twofold. On the one hand, different rates achieve different transmission ranges, which lead to different neighborhood diversity. Explicitly, high-rate causes short transmission range, then in one hop, there are few neighbors around the sender, which presents low neighborhood diversity. Low-rate is likely to have long transmission range, therefore achieves high neighborhood diversity. So from the diversity point of view, low rate may be better. On the other hand, although low rate brings the benefit of larger one-hop distance which results in higher neighborhood diversity and fewer hop counts to reach the destination, it is still possible to achieve a low effective end-to-end throughput when using low-rate communication links, because the low rate disadvantage may overwhelm this benefit. So it is nontrivial to decide which rate is indeed better.

Besides the inherent rate-distance, rate-diversity and rate-hop trade-offs which affect the throughput performance of opportunistic routing, the forwarding strategy will also have an impact on the throughput. That is, for a given transmission rate,

different candidate forwarding sets, relay priority assignments, and candidate coordinations will all affect the throughput.

In the following subsections, we will examine the impact of transmission rate and forwarding strategy on the one-hop performance of opportunistic routing, which leads us to the design of efficient local rate adaptation and candidate selection scheme. First we will analyze the one-hop packet forwarding time introduced by opportunistic routing.

7.3.1 One-hop Packet Forwarding Time of Opportunistic Routing

We define the one-hop packet forwarding time cost by the i^{th} candidate as the period from the time when the sender is going to transmit the packet to the time when the i^{th} candidate becomes the actual forwarder. Although the one-hop packet forwarding time varies for different MAC protocols, for any protocol, it can be divided into two parts. One part is introduced from the sender and the other part is introduced from the candidate coordination, which are defined as follows:

- T_s: the sender delay which can be further divided into three parts: channel contention delay (T_c), data transmission time (T_d) and propagation delay (T_p):

$$T_s = T_c + T_d + T_p \tag{7.2}$$

  For a contention-based MAC protocol (like FSA), T_c is the time needed for the sender to acquire the channel before it transmits the data packet, which includes the back-off time and Distributed Interframe Space (DIFS). T_d is equal to protocol header transmission time (T_h) plus data payload transmission time (T_{pl}), which is

$$T_d = T_h + T_{pl} \tag{7.3}$$

where T_h is determined by certain protocols at all layers, and T_{pl} is decided by the data payload length L_{pl} and the data transmission rate. The payload may be transmitted at different rates.

T_p is the time for the signal propagating from the sender to the candidates, which can be ignored when electromagnetic wave is transmitted in the air.

- $T_f(i)$: the i^{th} forwarding candidate coordination delay which is the time needed for the i^{th} candidate to acknowledge the sender and suppress other potential forwarders. Note that $T_f(i)$ is an increasing function of i, since the lower-priority forwarding candidates always need to wait and confirm that no higher-priority candidates have relayed the packet before it takes its turn to relay the packet. For the protocol we introduced in Section 7.1, $T_f(i) = i \times T_{SIFS} + T_{ACK}$, where T_{ACK} is the ACK transmission time.

Thus, the total medium time needed for a packet forwarding from the sender to the i^{th} forwarding candidate is

$$t_i = T_s + T_f(i) \tag{7.4}$$

7.3.2 Impact of Transmission Rate on Throughput

We examine the impact of transmission rate on throughput by using two examples. In one example, transmission at higher rate is better; while in the other example, lower rate achieves higher throughput. The throughput definition we use is the same as that proposed in [31] which is the bit-meters successfully delivered per second with unit bmps.

Assume the data payload $L_{pl} = 1000$ bytes. For simplicity, we assume the sender delay only includes the data transmission time (T_d), and T_h in Eq. (7.3) is fixed at $200\mu s$. So $T_s = \frac{1000 \cdot 8}{R_j} + 200\mu s$. Recall that R_j is the data transmission rate. According to the MAC protocol we discussed in Section 7.1, assuming $T_{SIFS} = 10\mu s$

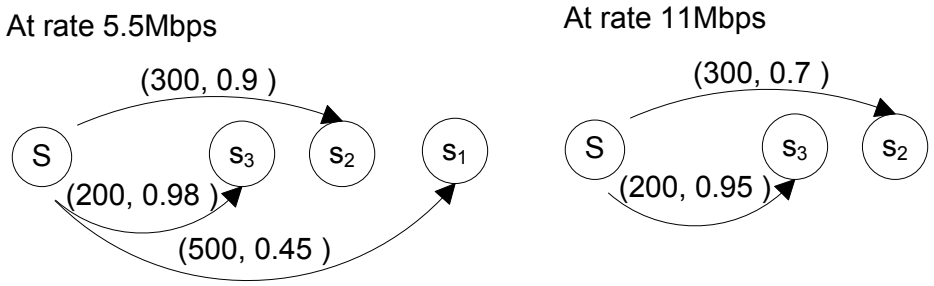


Figure 7.2: Different transmission rates result in different next-hop neighbor sets

and $T_{ACK} = 192\mu s$, $T_f(i) = 10i + 192\mu s$. Then $t_i = \frac{8000}{R_j} + 10i + 392\mu s$. In Fig. 7.2, the sender S transmits the data at 5.5Mbps and 11Mbps respectively. The next-hop neighbor set at each transmission rate and the corresponding (advancement, PRR) pairs associated with each neighbor are indicated in the figure. Assume at each rate, the neighbor closer to the destination is assigned higher relay priority. For a long term, S sends out a sufficient large number of packets, say N. Then when $R_j = 11Mbps$, there are $L_{pl}(300 \cdot 0.7N + 200 \cdot 0.95 \cdot 0.3N) = 2.136NM$ bit-meters are delivered, and the corresponding total packet forwarding time is $(t_1 \cdot 0.7N + t_2 \cdot 0.3N) = 1132.27N\mu s$. So there are $1.886G$ bit-meters successfully transmitted per second. We name it as the one-hop throughput. Similarly, the one-hop throughput at 5.5Mbps is $1.651G$ bmps, which is smaller than the throughput at 11Mbps. That is, in this example, although lower rate introduces more spatial diversity (more neighbors), this benefit does not make up the cost on the longer medium time. Now let's assume the neighbor s_3 is removed from Fig. 7.2 for each rate. Then when S is transmitting at 5.5Mbps, the one-hop throughput is $1.60G$ bmps. While when S is transmitting at 11Mbps, it achieves $1.49G$ bmps, which is smaller than that using rate 5.5Mbps. So transmitting at lower rate is better than higher rate in this case, because the extra spatial diversity brought by lower rate does help to improve the packet advancement but only introduce moderate extra packet forwarding time.

7.3.3 Impact of Forwarding Strategy on Throughput

We have seen that multi-rate capability has an impact on throughput. Other than this factor, for any given rate, different candidate prioritization also results in different throughput in opportunistic routing. Returning to the example in Fig. 7.2 at rate 5.5Mbps. If we assign s_2 the highest priority, then s_1, then s_3. The one-hop throughput is 1.306G bmps, which is lower than that achieved by assigning higher priority to the candidate closer to the destination. Actually, it has been proved in [67] that giving candidates closer to the destination higher priorities achieves maximum expected packet advancement (EPA).

7.3.4 Impact of Candidate Coordination on Throughput

The coordination delay is another key factor affecting the one-hop throughput. We use two extreme cases to illustrate the impact of this factor. First, we assume this delay is much larger than the sender delay, then it would be better to retransmit the packet instead of waiting for other forwarding candidates to relay the packet in order to save the packet forwarding time. In this case, one candidate may be optimal. On the other hand, we assume this delay is negligible, that is, the lower-priority candidates can relay the packet immediately when higher-priority candidates failed to do so. In this case, it is not difficult to imagine that we should involve all the available next-hop neighbors into opportunistic forwarding, because any extra candidates would help to improve the relay reliability but without introducing any extra delay. We should also give candidates closer to the destination higher relay priorities, since larger-advancement candidates should always try first in order to maximize the EPA. If they failed to relay the packet, the lower-priority candidates could instantaneously relay the correctly received packet without needing to wait. Therefore, the coordination delay has a great impact on throughput. Since we use the compressed slotted acknowledgement, which introduces small coordination delay

among candidates, it would be better to give candidates closer to the destination higher relay priorities.

ACK reliability from the candidates to the sender as well as among candidates are also important factors affecting throughput. On the one hand, if the sender does not receive the ACK sent by the candidate, it will retransmit the packet, which wastes bandwidth. However, recent study [54] has shown that since ACK is transmitted in a synchronous way, in the sense that it is sent out with a short delay (T_{SIFS}) after data packet reception, its reliability is considerably higher than that of asynchronous data messages. For the coordination mechanism we use, the candidates send ACKs in a synchronous way such that the i^{th}-priority candidate broadcasts the ACK at $i \times T_{SIFS}$ after successful packet reception if it does not detect any higher-priority candidate's transmission. Furthermore, since the ACK is transmitted at the basic rate (1Mbps), it has higher reliability than the data packet which is transmitted at higher rates. So the ACK can usually be correctly received by the sender with high probability. On the other hand, if the lower-priority candidates do not sense the ACK transmission of the higher-priority candidates, it will send out an ACK, which may result in ACK collision at the sender side and data packet duplication at the candidate side. However, this scenario could never happen. Since all the forwarding candidates are in the data transmission range of the sender, the longest distance between any two candidates are twice of the data transmission range. Typically, carrier sensing range is around double of the data transmission range. So other candidates should be able to detect a signal strength increase if a candidate does send out an ACK. False positive could happen when a lower-priority candidate sense a signal but it is from other transmission from a remote node. In this case, lower-priority candidate would drop its received packet. If all the lower-priority candidates who received the packet correctly believe there is a higher-priority candidate receives the packet but actually there is not, no ACK would be sent back to the sender, then the sender

would retransmit the packet. We have shown in Chapter 6 that FSA takes advantage of physical layer information and is able to suppress lower-priority candidates with very high probability.

7.4 Expected One-hop Throughput (EOT)

According to the analysis above, for a given next-hop neighbor set $\mathcal{C}_j$, we now introduce the local metric, *Expected One-hop Throughput* (EOT) (in Eq. (7.5)), to characterize the local behavior of GOR in terms of bit-meter advancement per second.

$$EOT(\mathcal{F}_j) = L_{pl} \cdot \frac{\sum_{i=1}^{r} d_{j_i} p_{j_i} \cdot \prod_{w=0}^{i-1} \overline{p}_{j_w}}{t_r \overline{P}_{\mathcal{F}_j} + \sum_{i=1}^{r} t_i p_{j_i} \cdot \prod_{w=0}^{i-1} \overline{p}_{j_w}} \tag{7.5}$$

where $\mathcal{F}_j = \langle s_{j_1}, ..., s_{j_r} \rangle$, which is an ordered subset of $\mathcal{C}_j$ with priority $s_{j_1} > ... > s_{j_r}$; $r = |\mathcal{F}_j|$; $p_{j_0} := 0$; $\overline{p}_{j_w} = 1 - p_{j_w}$; and

$$\overline{P}_{\mathcal{F}_j} = \prod_{i=1}^{r} (1 - p_{j_i}) \tag{7.6}$$

which is the probability of none of the forwarding candidates in $\mathcal{F}_j$ successfully receiving the packet in one physical transmission from the sender.

The physical meaning of the EOT defined in Eq. (7.5) is the expected bit advancement per second for a local GOR procedure when the sender S transmits the packet at rate R_j. EOT integrates the factors of packet advancement, relay reliability, and one-hop packet forwarding time. Now for multi-rate GOR, our goal is to select an R_j and the corresponding $\mathcal{F}_j$ to locally maximize this metric. The intuitions to locally maximize the EOT are as the following: 1) as the end-to-end achievable throughput is smaller than per-hop throughput on each link, to maximize the local EOT is likely to increase the path throughput; 2) the path delay is the summation of per-hop delay, which is actually relative to the delay introduced by transmitting the packet and coordinating the candidates. As the per-hop delay factors (T_s and $T_f(i)$) are integrated

in the denominators of EOT, to maximize EOT is also implicitly to decrease per-hop delay, which may further decrease the path delay. 3) as the transmission reliability $\mathcal{F}_j$ is also implicitly embedded in EOT, maximizing EOT also tends to improve the reliability. Reliability is a key factor affecting throughout and delay for the following reason. If a packet is transmitted on a low reliable link, several retransmissions are needed to make a successful packet forwarding at one hop. These retransmissions not only harm the throughput and delay performance of the flow which the packet belongs to, but also introduce huge medium contentions to other flows, thus further decrease the whole system performance. However, only to maximize the one-hop reliability is not enough to achieve a good end-to-end throughput. Because reliable links likely have short hop distance, this short hop distance may result in taking many hops to deliver a packet from the source to the destination, which may also introduce large delay or more medium contention to other flows. Our EOT metric jointly takes into account the hop advancement, reliability and packet forwarding time.

7.5 Heuristic Candidate Selection Algorithm

A straightforward way to get the optimal R_j and $\mathcal{F}_j$ to maximize the EOT is to try all the ordered subset of $\mathcal{C}_j$ for each R_j, which runs in $O(keM!)$ time, where k is the number of different rates, e is the base of natural logarithm, and M is the largest number of neighbors at all rates. It is, however, not feasible when N is large. In this section, we propose a heuristic algorithm to get a solution approaching the optimum.

As there are a finite number of transmission rates, a natural approach is to decompose the optimization problem into two parts. First, we find the optimal solution for each R_j; then, we pick the maximum one among them. So we only need to discuss how to find the solution approaching the optimum for a given rate, R_j, and the corresponding available next-hop neighbor set, $\mathcal{C}_j$. The following Lemma guides us

to design the heuristic algorithm.

Lemma 7.5.1. *For given R_j and $\mathcal{C}_j$, define $\mathcal{F}_j^r$ as one feasible candidate set that achieves the maximum EOT by selecting r nodes, then $\forall\ r$ $(1 \leq r \leq |\mathcal{C}_j|)$, $\exists\ \mathcal{F}_j^r$, s.t. $\mathcal{F}_j^1 \subseteq \mathcal{F}_j^r$.*

Proof. We prove this Lemma by contradiction. Assume $\forall\ r\ (1 \leq r \leq |\mathcal{C}_j|)$, we could find a feasible $\mathcal{F}_j^r$, s.t. $\mathcal{F}_j^1 \nsubseteq \mathcal{F}_j^r$. Then from that $\mathcal{F}_j^r$, we can obtain a new ordered set by substituting the lowest-priority candidate in $\mathcal{F}_j^r$ as the node in $\mathcal{F}_j^1$. According to Eq. (7.5) and the fact that $\mathcal{F}_j^1$ achieves the maximum EOT by selecting 1 node, we can derive that the EOT of the new set is larger than that of the $\mathcal{F}_j^r$. It is a contradiction, so the assumption is false, then the Lemma is true. □

Lemma 7.5.1 basically indicates that for given R_j and $\mathcal{C}_j$, the candidate achieving the maximum EOT by selecting 1 node from $\mathcal{C}_j$ is contained in the candidate set achieving the maximum EOT by selecting more number of nodes from $\mathcal{C}_j$.

Actually, the numerator of EOT is the EPA defined in Chapter 2. The EPA has three nice properties: **relay priority rule**, **containing property** and **concavity**. We recall these properties as follows without proof. These properties will help us design the rate and candidate selection algorithm.

Property 7.5.2. Relay Priority Rule*: Given a forwarding candidate set $\mathcal{F}$, the maximum EPA can only be achieved by giving candidates closer to the destination higher relay priorities.*

The ***Relay Priority Rule*** guides us to prioritize forwarding candidates by only examining their advancement to the destination. Next, we present the relationship among the optimal forwarding candidate sets (in the sense of maximizing EPA) with different number of candidates selected from a given candidate set $\mathcal{C}$.

Property 7.5.3. Candidate Set Containing Property: *Given an available next-hop node set $\mathcal{C}$ ($M = |\mathcal{C}|$), let $\mathcal{F}_r^*$ be a feasible ordered candidate set that achieves the maximum EPA by selecting r candidates from $\mathcal{C}$, $\forall\ \mathcal{F}_{r-1}^*$, $\exists\ \mathcal{F}_r^*$, s.t.*

$$\mathcal{F}_{r-1}^* \subset \mathcal{F}_r^* \quad \forall\ 1 \leq r \leq M \tag{7.7}$$

Property 7.5.3 indicates that an $r-1$-candidate set that achieves the maximum EPA is a subset of at least one of the feasible r-candidate sets that achieve the maximum EPA.

We also has the following concave property of the maximum EPA.

Property 7.5.4. Maximum EPA Concavity: *The maximum EPA is an increasing and concave function of the number of forwarding candidates.*

This property indicates that involving more forwarding candidates will increase EPA, but the gained EPA becomes marginal when we keep doing so. It has shown in Chapter 3 that the maximum EPA nearly does not increase when the number of forwarding candidates is larger than 3. Furthermore, involving more forwarding candidates may increase the probability of false positive, that is, lower-priority candidates are more likely to be falsely suppressed by other transmissions in the network. So in our algorithm design, we set a maximum allowable forwarding candidate number, r_{max}.

Now we examine the denominator of the EOT in Eq. (7.5). For the compressed slotted ACK mechanism, the denominator can be further simplified as $T_s(j)+T_{ACK}+T_{SIFS}(\sum_{i=1}^{r} i \cdot p_{j_i} \prod_{w=0}^{i-1} \overline{p}_{j_w} + r \cdot \overline{P}_{\mathcal{F}_j})$, where $T_s(j)$ is the delay at the sender side when the data packet is transmitted at rate R_j. The third part of this summation is the expected time introduced by candidate coordination, which is upper bounded by $r \cdot T_{SIFS}$. Since $T_{SIFS} \ll T_s(j) + T_{ACK}$ and r is a small number, the denominator can be seen as a constant at a fixed rate R_j. So maximizing the EOT is equivalent to maximizing its numerator, EPA.

FindMEOT($\mathcal{C}_j$'s, R_j's, r_{max})

$R^* \leftarrow 0$; $\mathcal{F}^* \leftarrow \emptyset$; $EOT^* \leftarrow 0$;
for each $\mathcal{C}_j$
  $\mathcal{F}_m \leftarrow \emptyset$; $EOT_m \leftarrow 0$; $\mathcal{A} \leftarrow \mathcal{C}_j - \mathcal{F}_m$;
  while ($\mathcal{A} \neq \emptyset$ && $|\mathcal{F}_m| < r_{max}$) **do**
    for each node $s_n \in \mathcal{A}$
      $\mathcal{F}_t \leftarrow$ Insert s_n into $\mathcal{F}_m$ according to ***Relay Priority Rule***;
      Get EOT on $\mathcal{F}_t$ according to Eq. (7.5);
      if ($EOT > EOT_m$)
        $EOT_m \leftarrow EOT$; $\mathcal{F}_m \leftarrow \mathcal{F}_t$
    end for
    $\mathcal{A} \leftarrow \mathcal{C}_j - \mathcal{F}_m$;
  end while
  if ($EOT_m > EOT^*$)
    $R^* \leftarrow R_j$; $\mathcal{F}^* \leftarrow \mathcal{F}_m$; $EOT^* \leftarrow EOT_m$;
end for
return(R^*, $\mathcal{F}^*$);

Table 7.1: Pseudocode of finding an transmission rate R^* and forwarding candidate set $\mathcal{F}^*$ approaching the maximum EOT

Therefore, according to Properties 7.5.2, 7.5.3, 7.5.4 and the analysis above, we propose a heuristic greedy algorithm which finds the transmission rate and the corresponding forwarding candidates approaching the maximum EOT. This heuristic algorithm FindMEOT is described in Table 7.1, where the input is the multi-rates R_j's, the corresponding $\mathcal{C}_j$'s and the maximum allowable forwarding candidate number r_{max}, and the output is the selected rate R^* and forwarding candidate set $\mathcal{F}^*$. For each rate R_j, this algorithm first finds the set $\mathcal{F}_m$ with one candidate that maximizes the EOT, then it incrementally adds more candidates into the existing $\mathcal{F}_m$ (line 6). Whenever adding a new candidate, it calculates the EOT (line 7), then updates the $\mathcal{F}_m$ when finding a new set achieving higher EOT than the existing one. Note that, according to Lemma 7.5.1, when the final returned set contains no more than 2 nodes, it is indeed the global optimum. Otherwise, it is an approximate optimal solution. An interesting finding is that this algorithm almost surely returns the global optimal solution even when the returned set contains more than 2 candidates.

7.6 Multirate Link Quality Measurement

To make our MGOR protocol work, we need to estimate the link quality (PRR) at different data rates. We propose a broadcast-based multirate link quality measurement scheme in this section. This link quality measurement scheme also serves for multirate neighborhood management.

Recall that there are k different data rates. Each node maintains k neighbor tables corresponding to the k data rates. The j^{th} table stores the bidirectional PRR information about its neighbors at rate R_j. For every τ second, each node broadcasts k "Hello" messages with each transmitted at a different data rate, e.g. 11Mbps, 5.5Mbps, and 2Mbps. Whenever a node n receives a "Hello" message sent from a node m at rate R_j, it will include node m into the corresponding neighbor table. Two

events drive the updating of PRR_{mn} at R_j on node n: one is the periodical updating event set by node n, for example, every t_u seconds node n will update PRR_{mn}. We denote this event as T; the other is the event that node n receives a "Hello" packet sent from m at rate R_j. We denote this event as H.

The Exponentially Weighted Moving Average (EWMA) method [64] is used to update PRR information. Since at each rate, the PRR is updated according to the same EWMA mechanism, we only describe the EWMA at a particular rate as follows. Let PRR_{mn} be the current estimation made by node n, $lastHello$ be the time stamp of the last event H, N_m be the number of known missed "Hello" packets between the current event H and last event H based on "Hello" message sequence number difference, and N_g be a guess on the number of missed packets based on "Hello" message broadcast frequency $\frac{1}{\tau}$ over a time window between the current T event and last H or T event. N_l and N_g are initialized to be 0, and FDR_{mn} is initialized to be 1.

This technique allows node n to measure PRR_{mn} and m to measure PRR_{nm}. Each "Hello" message sent at rate R_j by a node n contains PRR measured by n from each of its neighbors N_n at that rate during the last period of time. Then each neighbor of n, N_n, gets the PRR to n whenever it receives a "Hello" message from n.

The pseudocode of EWMA algorithm for node n to estimate PRR_{mn} at rate R_j is described in table 7.2, where $current_{Seq}$ and $last_{Seq}$ denote the sequence numbers of the current received "Hello" message and the last received "Hello" message, respectively, and $0 < \gamma < 1$ be the tunable parameter.

7.7 Performance Evaluation

In this section, we evaluate the performance of MGOR by simulation, and compare the performance of MGOR with multirate geographic routing (MGR), single-rate geo-

For node n:

When H event happens

$N_l = current_{Seq} - last_{Seq} - 1$

$last_{Seq} = current_{Seq}$

$lastHello =$ current time

$l = Max(N_l - N_g, 0)$

$N_g = 0$

$PRR_{mn} = PRR_{mn} \cdot \gamma^{l+1} + (1 - \gamma)$

When T event happens

$N_g = (current\ time - lastHello) \times \frac{1}{\tau}$

$l = N_g$

$PRR_{mn} = PRR_{mn} \cdot \gamma^{l}$

Table 7.2: Pseudocode of EWMA for a particular data rate

graphic routing (GR), and single-rate opportunistic routing. Our MGOR degenerates into MGR, when we choose only one forwarding candidate, and further degenerates into GR, when we also fix the transmission rate. For all the OR protocols, candidates closer to the destination are assigned higher relay priorities. The performance metrics we evaluate include: throughput, delay, and packet delivery ratio. In order to get insight into our rate and candidate selection algorithm, for MGOR, we show the number of packets transmitted at each rate in the whole network, and the average number of forwarding candidates used at each node on each data rate.

7.7.1 Simulation Setup

We implement the multirate link quality measurement mechanism and MGOR protocol with FSA in GlomoSim [73]. The FindMEOT algorithm proposed in Section 7.5 is used to select transmission rate and forwarding candidates for MGOR. This algorithm

is also used to select forwarding candidates for single-rate GOR by fixing the transmission rate. According to the analysis in Section 7.5 and considering the candidate coordination overhead, the maximum allowable forwarding candidate number (r_{max}) is set as 3. Other than the candidate coordination scheme, our OR protocol follows the same CSMA/CA medium access mechanism as that in 802.11 [1]. The simulated network has 50 stationary nodes randomly uniformly distributed in a $d \times d$ m^2 square region, with nodes having identical fixed transmission power of 15dbm. Each node can transmit data packets at three different rates: 11Mbps, 5.5Mbps, and 2Mbps. The ACK is transmitted at basic rate 1Mbps. According to the finding in [54] and discussion in Section 7.3.4, we assume the candidate coordination can be ensured by the compressed slotted ACK mechanism. As discussed in [75], 802.11 systems have very close interference ranges and the optimum carrier sensing ranges for different data rates, we set a single carrier sensing threshold as -100dbm for all rates. The receiving thresholds for 11Mbps, 5.5Mbps, 2Mbps and 1Mbps are -83dbm, -87dbm, -91dbm, and -94dbm, respectively. The packet reception decision is based on the SNR threshold and receiving threshold. When the SNR is larger than a defined threshold and the signal receiving power is above the corresponding threshold, the packet is received without error. Otherwise the packet is dropped. To simulate a randomly lossy channel, we assume Ground Reflection (Two-Ray) path loss model and Ricean fading model with $K = 4$ [52] for signal propagation. The multirate link quality measurement mechanism proposed in Section 7.6 is used to probe the link quality at each data rate, and the γ in EWMA method is chosen to be 0.9. We examine the impact of node density on the performance by setting $d = 1500, 1800, 2100, 2400$. The corresponding network density in terms of average number of neighbors per node at each rate is summarized in Table 7.3. We randomly choose 25 communication pairs in the network. The sources are CBR (constant bit rate) with packet interval of 75 mini-seconds (which makes the network saturated) and each packet being 512 bytes

Data rate (Mbps)	Terrain side length			
	1500	1800	2100	2400
2	19.7	14.4	11.3	8.8
5.5	16.3	11.9	8.8	6.8
11	11.1	7.9	5.8	4.3

Table 7.3: Average number of neighbors per node at each rate under different network densities

long. UDP (User Datagram Protocol) is used as transportation layer protocol. Each communication session continues 30 seconds. All the simulation results are averaged over 25 flows under 5 simulation runs with different seeds.

7.7.2 Simulation Results and Analysis

7.7.2.1 Throughput

Figure 7.3 shows the throughput of MGOR, single-rate GOR, MGR, and single-rate GR. First, MGOR achieves the highest throughput among all the protocols under all the network densities. Second, generally, single-rate GOR achieves higher throughput than the corresponding single-rate GR under each data rate and density. The spacial diversity gain introduced by involving multiple forwarding candidates in GOR dose increase the probability of a successful transmission at each hop, thus avoids the retransmission overhead, which results in higher throughput. Third, although GOR at 2Mbps can not support the traffic demand in the network, it achieves much higher throughput than the corresponding GR.

7.7.2.2 Packet Delivery Ratio

The packet delivery ratio is shown in Figure 7.4. We can see that MGOR achieves the highest packet delivery ratio among all the protocols. It delivers all the packets

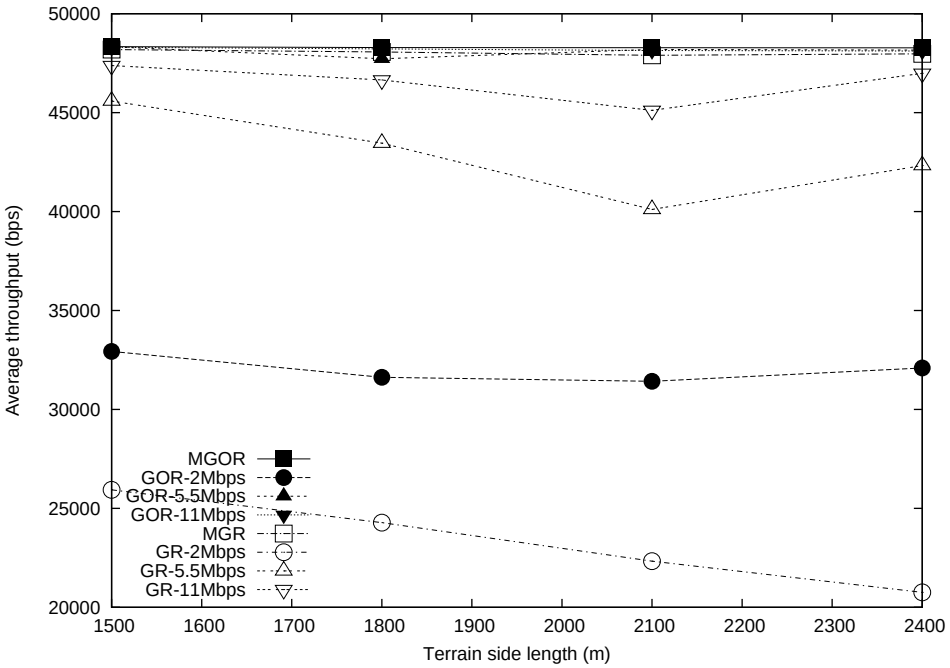


Figure 7.3: Average throughput of MGOR, single-rate GOR, MGR, and single-rate GR under different network densities

under all the network densities. Generally, GOR at each rate achieves higher packet delivery ratio than that of the corresponding GR. This figure is consistent with Figure 7.3.

7.7.2.3 Delay

The delay performance of these protocols is shown in Figure 7.5. We can see that GOR achieves much lower delay than the corresponding GR at any single rate. The reduction on retransmission does help reduce the end-to-end delay. Figure 7.6 enlarges the display of the delay performance of MGOR and GOR at 11Mbps. We can see that MGOR achieves lower delay than GOR at any single rate, especially when the network density is low. In some situation (e.g. the number of neighbor at 11Mbps is small) MGOR transmit packets at 5.5Mbps in order to achieve higher transmission advancement and reliability than 11Mbps; in some other situation, if transmission at

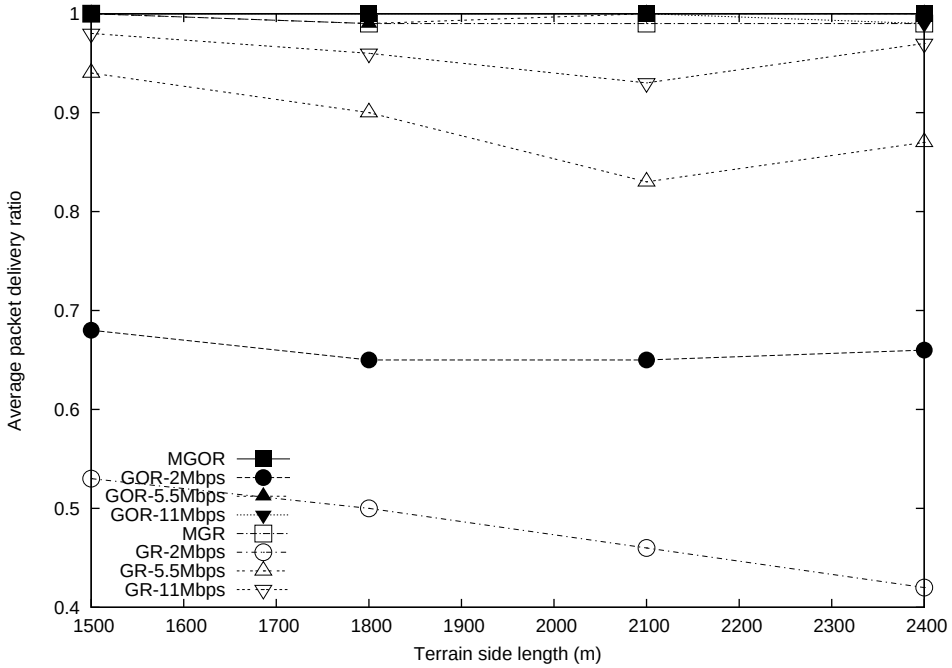


Figure 7.4: Average packet delivery ratio of MGOR, single-rate GOR, MGR, and single-rate GR under different network densities

11Mbps already introduces sufficient spacial diversity, MGOR chooses to transmit at higher rate (11Mbps). That's why MGOR has better performance than any single-rate GOR. It's obvious that the delay of each protocol increases when the network area is expanded, since more hops are needed for delivering packets. The hop count performance is shown in Figure 7.7.

7.7.2.4 Hop count

From Figure 7.7, we can see that GOR has slightly larger hop count than GR at each rate. Although GOR allows packets to be forwarded on long-distance links, some forwarding candidates with smaller advancement may also be chosen as the actual forwarder. The hop count of MGOR is between those of GOR at 11Mbps and 5.5Mbps, but closer to that at 5.5Mbps. The rate-distance trade-off is explicitly shown in the figure for both GR and GOR, that is, the hop count of lower rate

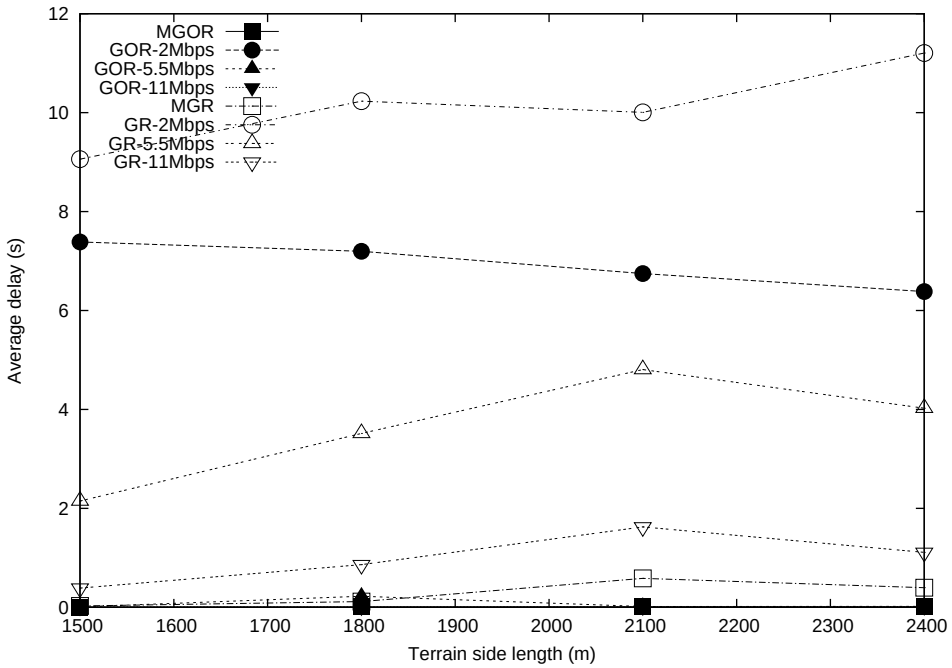


Figure 7.5: Average delay of MGOR, single-rate GOR, MGR, and single-rate GR under different network densities

is smaller than that of higher rate, since lower rates results in longer transmission ranges.

7.7.2.5 Average number of forwarding candidates

Figure 7.8 shows that for MGOR the number of forwarding candidates at each rate decreases when the network density is decreased. Furthermore, transmission at lower rate (5.5Mbps) results in more forwarding candidates than that at higher rate (11Mbps). In our MGOR, we do not choose 2Mbps transmission rate, since the traffic demand is already larger than the supportable rate of 2Mbps.

7.7.2.6 Portion of packets transmitted per node at each rate

Figure 7.9 shows that when the network becomes sparser, more packets are selected to transmit at 5.5Mbps in our MGOR protocol. Lower transmission rate results in

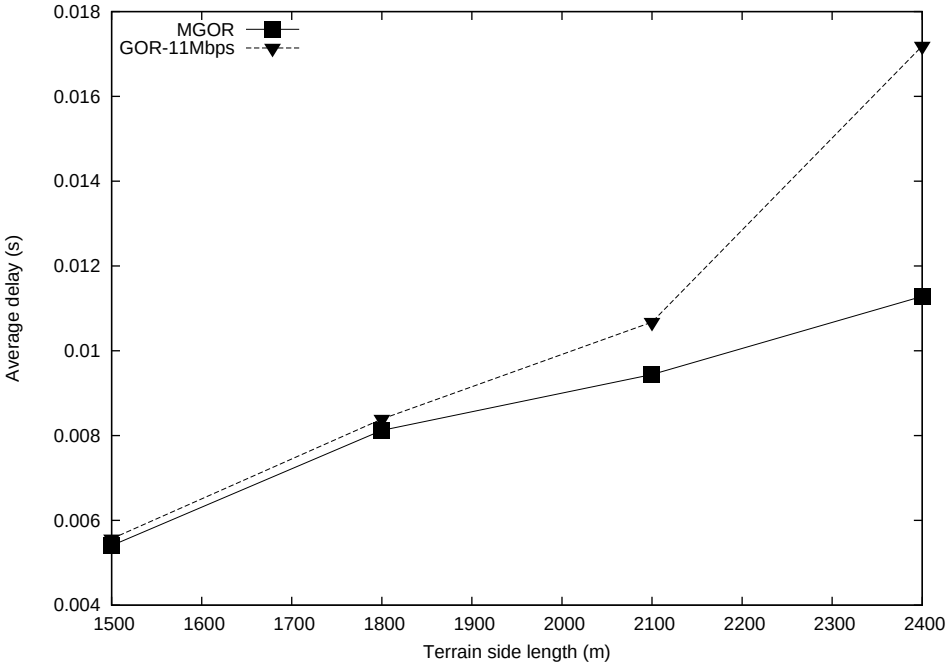


Figure 7.6: Average delay of MGOR and 11Mbps GOR under different network densities

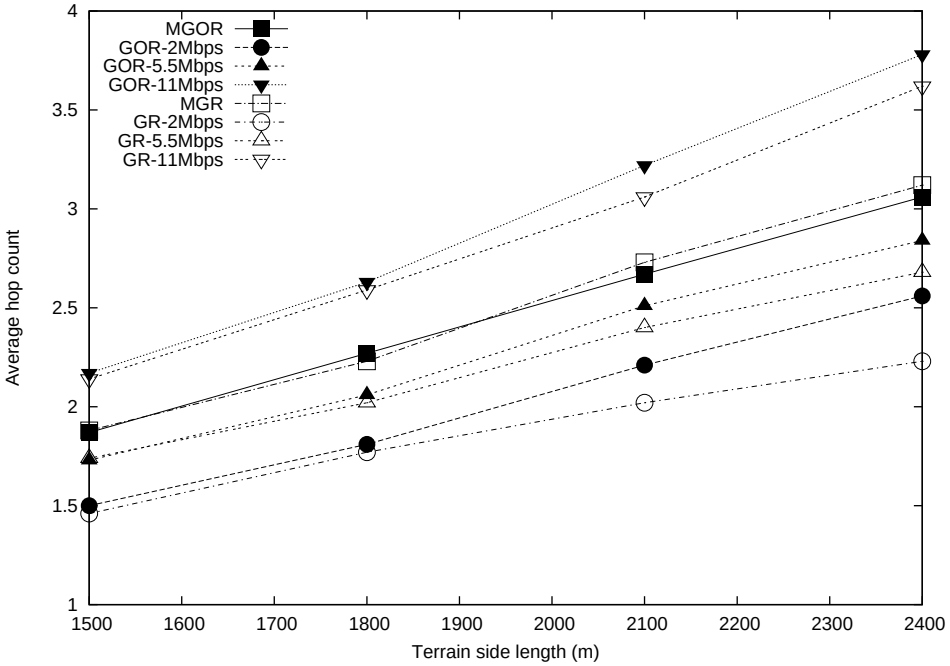


Figure 7.7: Hop count of each protocol

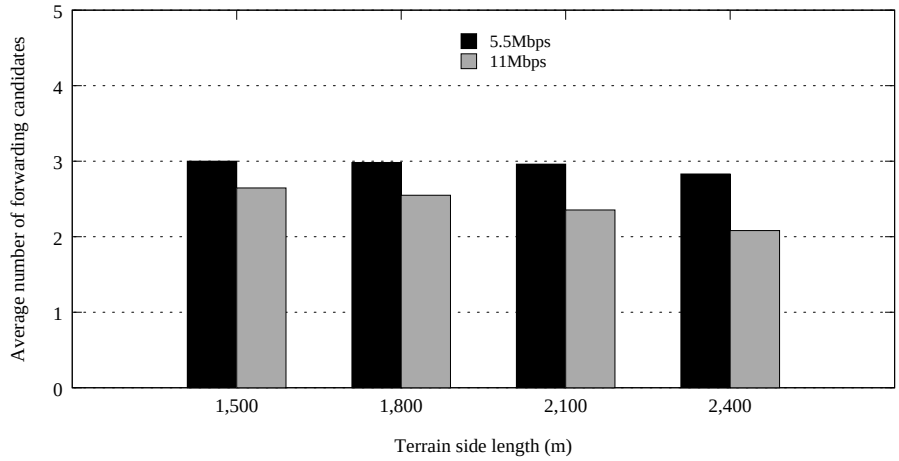


Figure 7.8: Average number of forwarding candidates of MGOR at each rate under different network densities

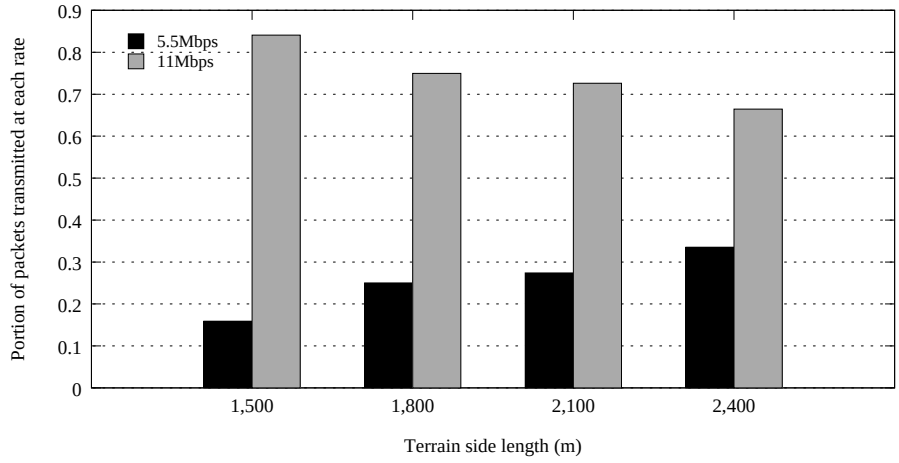


Figure 7.9: Average portion of packets transmitted per node at each rate of MGOR under different network densities

longer transmission range, which leads to more number of neighbors, thus increases spatial diversity. Therefore, in MGOR, transmission at lower rate does introduce spatial diversity gain and increase the probability of a successful transmission, then decrease the end-to-end delay.

7.8 Conclusion

In this Chapter, we studied multi-rate geographic opportunistic routing (MGOR) in a contention-based scenario, and examined the factors that affect its throughput, which includes multi-rate capability, candidate selection, prioritization, and coordination. Based on our analysis, we proposed the local metric, the *Expected one-hop throughput* (EOT), to characterize the trade-off between the packet advancement and medium time cost under different data rates. We further proposed a rate and candidate selection algorithm to approach the local optimum of this metric. We also presented a multirate link quality measurement mechanism. Simulation results show that MGOR incorporating our algorithm achieves better throughput and delay performance than the corresponding opportunistic routing and geographic routing operating at any single rate, which indicates that EOT is a good local metric to achieve high end-to-end throughput and low delay for MGOR.

Chapter 8

Secure Link Quality Measurement

The packet reception ratio (PRR) has been widely used as an indicator of the link reliability in multihop wireless networks. It has been shown that routing performance is significantly improved by considering the link PRR information. For example, *expected transmission count* (ETX) based routing achieves much higher throughput than traditional minimum-hop routing protocols in wireless mesh networks [21]. The ETX is defined as $\frac{1}{p_f \cdot p_r}$, where p_f and p_r is the forward and reverse link PRR, respectively. Recent work in sensor networks [54] suggests a link metric (ETF), *expected number of transmissions over forward links*, which only considers forward link PRR. State-of-the-art geographic routing protocols [56, 72] and most opportunistic routing protocols [9, 68] rely on link quality information to make routing decision.

Providing accurate link quality measurement (LQM) [1] is essential to ensure right operation of the above protocols/schemes. Furthermore, LQM is also important to supporting QoS guarantee in multihop wireless networks. Lastly, accurate long-term statistics of link-quality information is necessary to diagnose a network to identify the source of network failures, and reduce the management overhead.

[1] In this Chapter, we mainly focus on PRR measurement. Without specifying, the link quality indicates PRR.

The existing LQM mechanisms proposed in the literature [21, 38, 54] can be generally classified into three types: active, passive, and cooperative [38]. For broadcast-based active probing [21], each node periodically broadcasts hello/probing packets, and its neighbors record the number of received packets to calculate the PRRs from the node to themselves. In passive probing [38], the real traffic generated in the network is used as probing packets without introducing extra overhead. For cooperative probing [38], a node overhears the transmissions of its neighbor to estimate the link quality from the neighbor to itself.

However, for any of the existing LQM mechanisms, the inherent common fact is that a node's knowledge about the forward PRR from itself to its neighbor is informed by the neighbor. Since multihop wireless networks are generally deployed in an ad hoc style or in untrusty environments, nodes may be compromised and act maliciously. This receiver-dependent measurement opens up a door for malicious attackers to report a false measurement result, thus disturb the routing decision for all the PRR-based protocols. For example, in Fig. 8.1, suppose A is the source and D is the destination, and the actual PRR is indicated above each link in Fig. 8.1(a). The ETF-based shortest path routing would select the path $A \rightarrow B \rightarrow D$, since it has the lowest ETF path cost. However, if C is a malicious node, and reports to A that the PRR from A to itself is 0.9 (indicated below the link in Fig. 8.1(b)), then A would select path $A \rightarrow C \rightarrow D$. In such a way, a suboptimal path is selected between A and D, thus degrades routing performance. More severely, C attracts all the traffic from A, then with the control of the traffic, it can further maliciously drop the packets.

To the best of our knowledge, none of the existing work addresses security vulnerabilities in the existing LQM mechanisms. As LQM is becoming an indispensable component in multihop wireless networks, it is necessary to make this component work securely and provide actual and accurate PRR information.

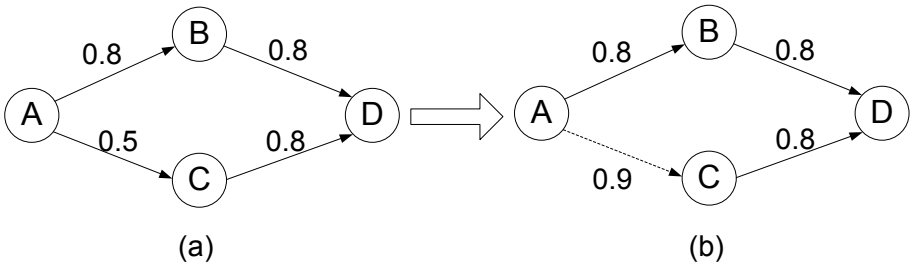


Figure 8.1: A 4-node example. (a) The actual PRR on each link is indicated, and the ETF-based routing selects the optimal path $A \rightarrow B \rightarrow D$. (b) The malicious node C bluffs A into believing that the PRR from A to C is 0.9, then the ETF-based routing would select the suboptimal path $A \rightarrow C \rightarrow D$

In this Chapter, we analyze the security vulnerabilities in the existing LQM mechanisms. We then propose a broadcast-based secure LQM mechanism, which prevents the malicious attacker from reporting a higher PRR than the actual one. This framework can be easily applied to unicast-based and cooperative LQM mechanisms.

The rest of this Chapter is organized as follows. Section 8.1 introduces the existing link quality measurement mechanisms and point out their security pitfalls. We propose a broadcast-based secure LQM (SLQM) mechanism and analyze its security strength and overhead in Section 8.2. Conclusions are drawn in Section 8.3.

8.1 Existing Link Quality Measurement Mechanisms and Vulnerabilities

This section gives an overview of the existing LQM mechanisms and analyzes their security vulnerabilities. According to the type of probing packets, LQM can be classified into broadcast-based and unicast-based probing. While based on the generation source of probing packets, LQM can also be categorized into active, passive, and cooperative probing [38].

8.1.1 Broadcast-based Active Probing

For broadcast-based active probing [21], each node broadcasts link probes of a fixed size, at an average period τ (e.g. 1 second). Every node remembers the probes it receives during the last w seconds (e.g. 10 seconds), allowing it to calculate the PRR from the measuring node at any time t as: $r(t) = \frac{count(t-w,t)}{w/\tau}$, where $count(t-w,t)$ is the number of probes received during the window w, and w/τ is the number of probes that should have been received. In the case of two neighboring nodes A and B, this technique allows A to measure the PRR from B to A, and B to measure the PRR from A to B. Each probe sent by a node A contains the number of probing packets received by A from each of its neighbors during the last w seconds. This allows each neighbor of A to calculate the forward link PRR to A whenever it receives a probe from A.

The security vulnerability in the broadcast-based active probing is that a malicious node can easily report a false measurement result. For example, if node B is an attacker, it can bluff A into believing that the PRR from A to itself is 1 by claiming that it received w/τ packets in the last probing window w.

8.1.2 Unicast-based Passive Probing

Unicast-based passive probing [38] makes use of the real unicast traffic as the "natural" probing packets without incurring extra overhead. It is applicable when there is enough unicast traffic on a measured unidirectional link. It runs as follows: for instance, suppose node A has enough traffic to node B. Then, A gets the information about the number of successful transmissions (N_s) and the total number of transmissions (N_t) from its MAC's MIB (Management Information Base) for the traffic. At the end of an update period, the PRR is derived as $\frac{N_s}{N_t}$, and is further smoothed by moving average [38].

For unicast-based passive probing, it is hard but not impossible for an attacker to

cheat on the link quality. In 802.11 [1], the Distributed Coordination Function (DCF) defines two access mechanisms for packet transmissions: basic access mechanism, and RTS/CTS access mechanism. We analyze the security vulnerability of the unicast-based passive probing under these two access mechanisms as following.

In the basic access mechanism, a sender starts the transmission of a DATA frame after it senses the channel is idle for a while. Upon successful decoding the whole DATA frame, the receiver sends an ACK frame back to the sender, indicating successful reception of the DATA frame. In this case, even when it can not decode the whole data frame, a receiver may decode some parts of it [33]. So it is possible for a malicious receiver to figure out the sender's address and send back an ACK to claim a correct reception even when it receives a corrupted data frame.

The RTS/CTS access mechanism uses a four-way handshake in order to reduce bandwidth loss due to the hidden terminal problem. Different from the basic access mechanism, a sender will send a RTS frame to the receiver before it sends out the DATA frame. Upon successful reception of the RTS frame, the receiver then sends a CTS frame back to the sender. The sender can start sending the DATA frame after the reception of the CTS frame. As in the basic access mechanism, upon successful reception of the DATA frame, the receiver sends an ACK frame back to the sender. In this case, by receiving the RTS, a malicious receiver can figure out the sender's address, so even it receives a corrupted data frame, it can still claim a successful reception by sending back an ACK.

In summary, although a sender estimates the link quality based on its own MIB information in the unicast-based passive probing, this information is still dependent on the feedback (ACK) from the receiver. A malicious receiver may still be able to make use of the ACK to bluff the sender into believing that there exists a high quality link from the sender to the receiver.

8.1.3 Cooperative Probing

Cooperative probing [38] is used when there is not enough unicast traffic from a measuring node to its neighbor, but to others. For example, a measuring node A has two one-hop neighbors, B and C. A has no egress traffic to C, but to B. The neighbor node (C) with no traffic to it from the measuring node (A) is called a "cooperative" node. Due to the broadcast nature of wireless media, the node C can overhear the traffic from the measuring node A to B. This traffic is called *cross traffic*. The overhearing result is then used for the measuring node to derive the quality of link $A \rightarrow C$. [38] assumes the node C cannot receive duplicate frames from its MAC layer even in the promiscuous mode, the retransmitted packets are not used for measurements. So node A counts first-time successful transmissions (C_c) within the cross traffic. In the update period, a report of overheard results (C_a) from C is sent to A, and then the PRR in this period is calculated as $\frac{C_a}{C_c}$.

To attack cooperative probing, similar to the unicast-based passive probing, a malicious "cooperative" node does not need to decode the whole data frame correctly. As long as it can figure out the sender's address and the status (0/1) of the "retry" bit in the data frame, it can increase its count of C_a.

8.1.4 Unicast-based Active Probing

When there is no egress/cross traffic, unicast-based active probing can be applied [38]. For example, if node A has no traffic to B or C, A initiates a unicast-based active probing on link $A \rightarrow B$ by generating unicast probing packets. Then, the link quality from A to B is measured by the same way as passive probing. At the same time, the quality of link $A \rightarrow C$ can be measured by cooperative probing. In this way, unicast-based active probing acts similarly as the broadcast-based active probing, with difference in that in unicast-based probing the receiver need send back an ACK to the sender when it receives the data frame correctly and the sender will retransmit

data frames when no ACK receives, while in broadcast-based active probing, any node does not need to send ACK.

For unicast-based active probing, the security vulnerabilities in measuring the link quality from the measuring node (e.g. A) to the intended receiver (e.g. B) and to the "cooperative" node (e.g. C) are the same as the that in unicast-based passive probing and "cooperative" probing, respectively.

To sum up, all the existing LQM mechanisms can not prevent a receiver cheating on the PRR. The inherent fact is that the receiver can claim a correct data frame reception without showing any evidence. To fix this vulnerability, we propose a broadcast-based secure LQM (SLQM) mechanism based on the challenge-response mechanism in the following section. We will show that this broadcast-based mechanism can be easily applied to unicast-based and cooperative SLQM mechanisms.

8.2 Broadcast-based Secure Link Quality Measurement

In this section, we propose a broadcast-based secure LQM mechanism, and then analyze its security strength and its computation, storage, and communication overhead. In this paper, we assume that a malicious node always wants to report a higher PRR than the actual measured one, thus disturb PRR-based routing performance. We also assume that a unique pair-wise key has been established between each pair of neighbors. The neighborhood pair-wise key establishment mechanisms have been extensively studied in multihop wireless networks [25].

8.2.1 Broadcast-based SLQM Framework

Assume a node A has N one-hop neighbors $A_1, A_2, ..., A_N$, and needs to measure the link PRR (p_i) to each of its neighbors (A_i). Similar to [38], the measurement is done

periodically. Each measurement period consists of three consecutive phases: probing, reporting, and updating phases, which are described as follows.

Probing phase: In this phase, A broadcasts N_s packets to its neighbors. In the j^{th} packet r_j, it embeds a random number. It keeps the broadcasted packets in its buffer within this measurement period. Receiver A_i only stores the XOR-ed result (R_i) of all the correctly received packets, and the corresponding indicator vector V_i defined in Eq. (8.1) that indicates the index of the received packet. Note that A_i can compute the XOR-ed result on the fly whenever it receives a new probing packet.

$$V_i(j) = \begin{cases} 1, & A_i \text{ received the } j^{th} \text{ packet correctly;} \\ 0, & \text{otherwise.} \end{cases} \tag{8.1}$$

where $V_i(j)$ is the j^{th} bit from the higher (left) end of the vector V_i.

Reporting phase: When the probing phase is ended, each neighbor A_i sends A a report $Rep_i := \{H_i, V_i\}$, where $H_i = h_{\mathcal{K}_i}(R_i)$ is a keyed hash of R_i with the pairwise key $\mathcal{K}_i$ shared between A and A_i. The hash function can be any of the existing cryptographic hash functions, such as MD5 [53].

Updating phase: On receiving A_i's report, A figures out how many and which packets A_i receives by examining the positions of bit '1's in vector V_i. Since A keeps all the packets that it broadcasted, it computes $R_i^{'}$ by doing XOR of the packets that A_i claims it received. A then computes $H_i^{'} = h_{\mathcal{K}_i}(R_i^{'})$. If $H_i^{'} = H_i$, A accepts this report; otherwise, it rejects the report. Suppose A counts there are N_{r_i} bit '1's in V_i, after A accepts the report, A calculates the PRR $p_i = \frac{N_{r_i}}{N_s}$ in this measurement period. A moving average method is further used to smooth the measured result. Denote the measured result in the k^{th} measurement period as $p_i[k]$, the smoothed PRR, $\widetilde{p}_i(k)$, at the end of the k^{th} period is calculated as

$$\widetilde{p}_i(k) = (1-\alpha)\widetilde{p}_i[k-1] + \alpha p_i[k] \tag{8.2}$$

where α is a smoothing constant in the range of (0,1).

Figure 8.2 shows an example of the broadcast-based SLQM mechanism in a measurement period. Suppose in the measuring phase, A broadcasts 5 probing packets (r_1,...,r_5), and A_i receives the packets r_1, r_3, and r_5. In the reporting phase, A_i calculates $H_i = h_{\mathcal{K}_i}(r_1 \oplus r_3 \oplus r_5)$, then sends H_i and a 5-bit vector $V_i = 10101$ back to A. When it receives the H_i and V_i, A examines V_i and get the indices $(u_1, ..., u_c)$ of the packets A_i claims it receives, then calculates $H_i' = hash_{\mathcal{K}_i}(r_{u_1} \oplus ... \oplus r_{u_c})$. If $H_i = H_i'$, A accepts A_i's report; otherwise, rejects it.

8.2.2 Security Strength

We now analyze the security strength of our broadcast-based SLQM mechanism. This mechanism achieves the security goal that prevents a malicious attacker from reporting a higher PRR than the actual one. We assume A_i is malicious in the following discussion.

First, it's computationally impossible for A_i to guess the packets which it does not receive, even when A_i overhears other's report. For example, in Figure 8.2, if A_i wants to claim it receives r_1, r_3, r_4, r_5, it needs to create a hash value $H_i = h_{\mathcal{K}_i}(r_1 \oplus r_3 \oplus r_4 \oplus r_5)$. Since it has no idea what r_4 is, the only thing it can do is to make a guess on r_4. However, it's hard to make a correct guess according to the weak collision resistance property of the hash function that given $x = (r_1 \oplus r_3 \oplus r_4 \oplus r_5)$, it's hard to find a $y = r_1 \oplus r_3 \oplus r_4^{'} \oplus r_5$, such that $h_{\mathcal{K}_i}(x) = h_{\mathcal{K}_i}(y)$. Even A_i overhears A_j's report indicating that A_j receives r_4, A_i still can not get any information about r_4 because of the one-way property of the hash function.

Second, our mechanism prevents A_i from replaying its own or other neighbor's report. According to the randomness embedded in each probing packet, even A_i receives all the probing packets in some measurement period, it can not replay this report in the following measurement period. Furthermore, if A_i replays A_j's report, this report can not pass the verification by A, because A uses $\mathcal{K}_i$ instead of $\mathcal{K}_j$ to

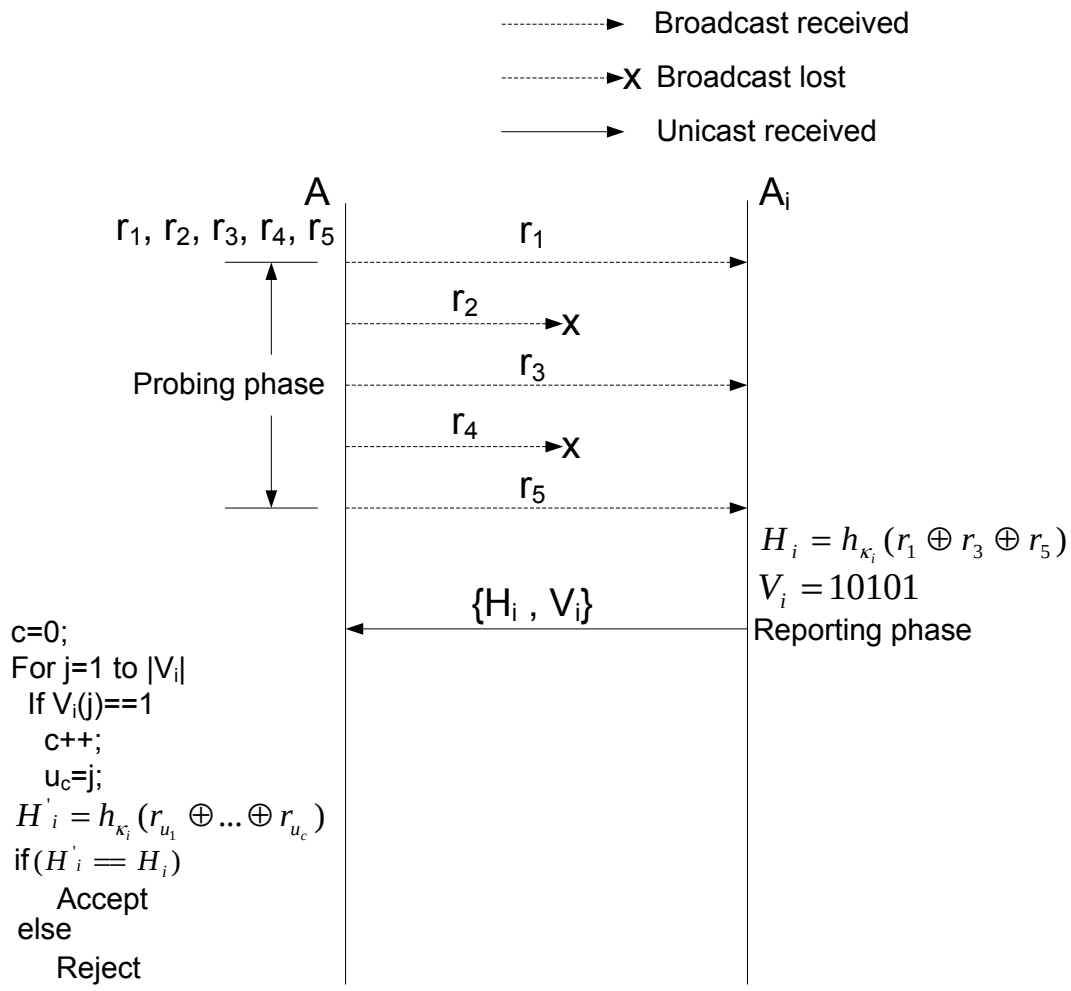


Figure 8.2: Probing and reporting phases of secure link quality measurement between A and A_i in a measurement period

verify A_i's report.

8.2.3 Computation, Storage and Communication Overhead

Computation overhead: On the sender side, A needs to generate a random number sequence. According to its computation and storage capability, A can generate a large random number sequence to be used for several measurement periods, and refresh this sequence when it is used up. Any of the existing efficient pseudorandom number generators, such as linear congruential generator [29], can serve this purpose. To do verification, A only needs to do XOR and hash operations, which are computationally efficient. On the receiver side, to create the report digest, each neighbor only needs to do a hash computation.

Storage overhead: On the sender side, A only needs to store the generated random numbers. Suppose the length of each random number is L_r bytes, the probing packet broadcast rate is B packet/second, and the probing phase is P seconds. Then in a measurement period, A needs $S = L_r \cdot B \cdot P$ bytes storage space. For example, if $L_r = 16$, $B = 1$, and $P = 10$, $S = 160 bytes$, which is supportable even on sensor nodes.

Communication overhead: The communication overhead of our SLQM mechanism is comparable to any existing broadcast-based probing mechanism, such as that in [21]. As the probing packet broadcast rate is usually low, e.g. $B = 1$, SLQM introduces very light local traffic into the network.

8.2.4 Applicability

As discussed above, our SLQM mechanism has very low computation, storage and communication overhead, so it's applicable to resource-constraint networks, such as wireless sensor networks, as well as more powerful networks, such as wireless mesh networks. Basically, broadcast-based SLQM can be implemented at any of applica-

tion, networking and MAC layers. Our SLQM framework can also be easily applied to unicast-based and cooperative LQM with a slight modification such that we embed a random number in each unicast packet (including retransmitted packets at MAC layer). For unicast-based SLQM, we can ask receiver to attach a hash value of the received packet in the corresponding ACK. For cooperative probing, the cooperative receiver does the same thing as the broadcast-based SLQM.

8.3 Conclusion

In this Chapter, we investigated the existing link quality measurement mechanisms, and analyzed the security vulnerabilities in them. A common inherent fact in all the existing LQM mechanisms are receiver-dependent measuring, that is, a node's knowledge about the forward PRR from itself to its neighbors is informed by its neighbors. We then proposed a broadcast-based secure LQM mechanism that prevents a neighboring node from maliciously claiming a higher measurement result. Our mechanism has very low computation, storage, and communication overhead, thus can be implemented in resource-constraint sensor networks as well as mesh networks. Our SLQM mechanism can be easily applied to unicast-based and cooperative LQM with slight modifications.

Chapter 9

Conclusions and Future Research

9.1 Summary

The essential idea of opportunistic routing is to exploit the broadcast nature and space diversity provided by the wireless medium. By having multiple forwarding candidates, the successful rate of each transmission can be much improved. However, a good OR protocol is to decide which set of nodes (in contrast to which single node) are good to form the forwarding candidate set and how they should be prioritized. Although we are taking opportunities, we want the packets to be routed to the destination through a set of paths that are statistically optimal. In this dissertation, we presented principles of the local behavior of OR, we analyzed the capacity, throughput and energy efficiency of OR, we developed new candidate coordination scheme for OR, and we designed secure link quality measurement mechanism.

In Chapter 2, we found and proved properties of the local behavior of OR and the associated candidate selection and prioritization issues. The contributions of Chapter 4 and 5 present analytical model to compute the end-to-end throughput bound and capacity of OR in multi-radio multi-channel and multi-rate wireless networks. In Chapter 3 we proposed an energy-efficient geographic opportunistic routing frame-

work and the corresponding local candidate selection and prioritization algorithms. In Chapter 6, we presented a new efficient candidate coordination scheme which takes advantage of the physical layer information. We studied the performance of multirate GOR under a contention-based medium access scenario in Chapter 7. In Chapter 8, we presented a secure link quality measurement mechanism which is able to prevent a malicious attacker reporting a fault link quality. We summarize our results by Chapter below.

- Chapter 2. In this Chapter, we generalized the definition of EPA for arbitrary number of forwarding candidates in GOR. Through theoretical analysis, we showed that the maximum EPA can only be achieved by giving the forwarding candidates closer to the destination higher relay priorities when a forwarding candidate set is given. We give the analytical result of the upper bound of the EPA that any GOR can achieve. We also showed that giving an available next-hop neighbor set with M nodes, the maximum EPA achieved by selecting r ($1 \leq r \leq M$) nodes is a strictly increasing and concave function of r. We proved that a feasible subset of the available next-hop neighbor set that achieves that maximum EPA is contained in at least one feasible subset with more nodes. We also showed that the increasing of the maximum EPA is consistent with the increasing of the one-hop reliability.

- Chapter 3. In this Chapter, we studied the geographic opportunistic routing strategy with both routing and energy efficiencies as the major concerns. We proposed a new routing metric which evaluates EPA per unit of energy consumption so that the energy efficiency can be taken into consideration in routing. By leveraging the proved findings in Chapter 2, we proposed two localized candidate selection algorithms with $\mathbf{O}(M^3)$ and $\mathbf{O}(M^2)$ running time in the worst case, respectively, and $\Omega(M)$ in the best case, where M is the number of available next-hop neighbors. The algorithms efficiently determine the forwarding

candidate set that maximizes the proposed new metric for energy efficiency, namely, the EPA per unit of energy consumption. We further proposed an EGOR framework applying the node selection algorithms to achieve the energy efficiency. Simulation results show that EGOR achieves better energy efficiency than geographic routing and blind opportunistic protocols in all the cases while maintaining very good routing performance. Our simulation results also show that the number of forwarding candidates necessary to achieve the maximum energy efficiency is mainly affected by the reception to transmission energy ratio but not by the node density under a uniform node distribution. Although the EPA can be maximized by involving the most number of nodes in GOR, in terms of energy efficiency, only a very small number of forwarding candidates (around 2) are needed on average. This is true even when the energy consumption of reception is far less than that of transmission.

- Chapter 4. Taking into consideration of wireless interference and the unique property of OR, we proposed a new method of constructing transmission conflict graphs, and presented a methodology for computing the end-to-end throughput bounds (capacity) of OR giving forwarding strategies. We formulate the maximum end-to-end throughput problem of OR as a maximum-flow linear programming subject to the transmission conflict constraints and effective forwarding rate constraints on each link in different concurrent transmission sets. We also proposed two metrics for OR under multirate scenario, one is *expected medium time* (EMT), and the other is *expected advancement rate* (EAR). Based on these metrics, we proposed the distributed and local rate and candidate selection schemes: LMTOR and MGOR, respectively. We validate the analysis results by simulation, and compare the throughput capacity of multi-rate OR with single-rate ones under different settings, such as different topologies, source-destination distances, number of forwarding candidates, and node densi-

ties. We show that OR has great potential to improve the end-to-end throughput under different settings, and our proposed multi-rate OR schemes achieve higher throughput bound than any single-rate GOR. We observe some insights of OR: 1) the end-to-end capacity gained decreases when the number of forwarding candidates is increased. When the number of forwarding candidates is larger than 3, the throughput almost remains unchanged. 2) there exists a node density threshold, higher than which 24Mbps GOR performs better than 12Mbps GOR, and lower than which, vice versa. The threshold is about 5.5 and 10.9 neighbors per node on 12Mbps for line and square topologies, respectively.

- Chapter 5. We proposed a unified framework to compute the capacity of opportunistic routing between two end nodes in single/multi-radio/channel multihop wireless networks by allowing dynamic forwarding strategies. Our model accurately captures the unique property of OR that multiple outgoing links sharing the same transmitter can be virtually scheduled at the same time under particular rate constraints. We also studied the necessary and sufficient conditions for the schedulability of a flow demand vector associated with a transmitter to its forwarding candidates in a concurrent transmission set. We further proposed an LP approach and a heuristic algorithm to obtain an opportunistic forwarding strategy scheduling that satisfies a flow demand vector. Our methodology can not only be used to calculate the end-to-end throughput bound of OR and TR in multi-radio/channel multihop wireless networks, but also be used to study the OR behaviors (such as candidate selection and prioritization) in multi-radio multi-channel systems. Leveraging our analytical model, we found that OR can achieve comparable or even better performance than TR by using less radio resource.

- Chapter 6. We analyzed the coordination problem in opportunistic routing, and based on these analysis, we proposed a new coordination scheme “fast slotted

acknowledgment" (FSA) which fully takes advantage of the channel detection approach to meet an agreement among multiple candidates. We compared FSA with those state-of-the-art schemes and simulation results show that it achieves better performance in all the metrics, especially in time delay. The simulation also validated that FSA can achieve similar performance as ideal coordination where relay priority can be ensured and duplicate packet forwarding is avoided.

- Chapter 7. We studied multi-rate geographic opportunistic routing (MGOR) in a contention-based scenario, and examined the factors that affect its throughput, which includes multi-rate capability, candidate selection, prioritization, and coordination. Based on our analysis, we proposed the local metric, the *Expected one-hop throughput* (EOT), to characterize the trade-off between the packet advancement and medium time cost under different data rates. We further proposed a rate and candidate selection algorithm to approach the local optimum of this metric. We also presented a multirate link quality measurement mechanism. Simulation results show that MGOR incorporating our algorithm achieves better throughput and delay performance than the corresponding opportunistic routing and geographic routing operating at any single rate. It indicates that EOT is a good local metric to achieve high end-to-end throughput and low delay for MGOR.

- Chapter 8. We investigated the existing link quality measurement mechanisms, and analyzed the security vulnerabilities in them. A common inherent fact in all the existing LQM mechanisms are receiver-dependent measuring, that is, a node's knowledge about the forward PRR from itself to its neighbors is informed by its neighbors. We then proposed a broadcast-based secure LQM mechanism that prevents a neighboring node from maliciously claiming a higher measurement result. Our mechanism has very low computation, storage, and communication overhead, thus can be implemented in resource-constraint sensor

networks as well as mesh networks. Our SLQM mechanism can be easily applied to unicast-based and cooperative LQM with slight modifications.

9.2 Future Research Directions

The frameworks proposed in Chapter 4 and 5 which compute the throughput bound and capacity of OR need to find all the feasible concurrent transmitter sets, which is a NP-complete problem. How to efficiently find a good subset of all the CTS's to approach the optimal solution within a controllable gap could be an interesting topic. Some heuristic algorithms similar to that in [61], or column generation technique [76] may be adopted to serve this purpose.

The distributed algorithm LMTOR proposed in Chapter 4 needs to enumerate all the combinations of forwarding candidates, which may not be feasible when the network is dense. To design more efficient algorithms with smaller searching space is a valuable direction.

Routing metrics with various performance objectives, such as maximizing throughput, minimizing delay, and maximizing energy efficiency, can be studied and tradeoff between conflicting goals can be analyzed and considered for OR.

Another direction of effort is to investigate further the error of link quality (PRR) estimation and its impact on the OR performance in different types of networks, and design protocols accordingly that are robust to estimation error. We plan to break down this task into three subtasks.

First, geographical routing has been well studied in the literature in networks where location information is available to the nodes, which is true in many applications of multihop wireless networks. GOR has been proposed as an efficient routing scheme in such networks. In GOR, the Euclidean distance between nodes is known and can be used as the cost function in routing. We can start with GOR in wire-

less sensor networks where the distance is a fixed value and not affected by the link estimation error. We can design cost functions which are less affected by PRRs but represent the space diversity along the path.

Second, a local OR decision depends on $OETT$ and EMT in Chapter 4. $OETT$ only depends on the local PRRs while EMT depends on remote PRRs through D_is. Thus the impact of link estimation error is propagated through the network by D_is. For very dynamic networks, such as mobile ad hoc network and vehicular networks, the link condition may change very fast that may diminish the benefit from our optimization based on link error estimation. In our second step of understanding the impact of estimation error on routing performance, we will study D_is that are less sensitive to such changes. One option is to use the cost based on the traditional routing, which is less affected by the link estimation error than OR. The goal is to mitigate such impact from remote nodes and to focus on the impact on local estimations. Another option is to develop on-demand protocols [34, 50]. Similar to multipath on demand routing [65, 45, 71], multiple replies can be enabled from the destination and nodes learn its local spatial diversity opportunity and report it in the reply messages. Spacial diversity along the paths can then be taken into consideration in routing decisions.

Third, after understanding the impact on each local decision, we will then extend the investigation to the whole paths. This study will be in a relatively stable setting such as sensor networks and mesh networks. We may adopt ideas similar to the fisheye state routing (FSR) [48] which allow multi-level routing information exchange depending on the distance to the destination. The focus is to control the routing overhead while trying to take advantage of OR and path diversity in a larger scale. This study will help us to gain deeper understanding of the OR and the capability of gaining performance benefits in the face of inaccurate link quality estimation. We believe the theoretical results and insights from this research will be valuable to

research community and crucial to the design of practical and efficient OR protocols approaching optimal performance.

Combining OR with network coding [15] is a promising research direction.

Other than performance, security is another major concern in multihop wireless networks. OR, by its indeterministic nature, is more robust to many attacks aiming to disrupt routing and data forwarding functions. It is valuable to investigate the security application of OR. We will propose secure OR protocols and integrate it into existing security framework to provide more robust and more secure information delivery service.

Bibliography

[1] IEEE Std 802.11b 1999. *http://standards.ieee.org/*.

[2] Daniel Aguayo, John Bicket, Sanjit Biswas, Glenn Judd, and Robert Morris. Link-level measurements from an 802.11b mesh network. *SIGCOMM Comput. Commun. Rev.*, 34(4):121–132, 2004.

[3] Jing Ai, Alhussein A. Abouzeid, and Zhenzhen Ye. Cross-layer optimal decision policies for spatial diversity forwarding in wireless ad hoc networks. In *Mobile Adhoc and Sensor Sysetems (MASS)*, 2006.

[4] I. F. Akyildiz and I. H. Kasimoglu. Wireless sensor and actor networks: Research challenges. *Ad Hoc Networks Journal (Elsevier)*, 2(4):351–367, Oct. 2004.

[5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Commun. Mag.*, 40(8):102–116, Aug. 2002.

[6] I. F. Akyildiz, X. Wang, and W. Wang. Wireless mesh networks: A survey. *Computer Networks*, Mar. 2005.

[7] Mansoor Alicherry, Randeep Bhatia, and Li (Erran) Li. Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks. In *MobiCom '05*, pages 58–72, New York, NY, 2005. ACM.

[8] Baruch Awerbuch, David Holmer, and Herbert Rubens. The medium time metric: High throughput route selection in multi-rate ad hoc wireless networks. *MONET*, 11(2):253–266, 2006.

[9] S. Biswas and R. Morris. Exor: Opportunistic multi-hop routing for wireless networks. In *SIGCOMM'05*, Philadelphia, Pennsylvania, Aug. 2005.

[10] Sanjit Biswas and Robert Morris. Opportunistic routing in multihop wireless networks. In *HotNets-II*, Cambridge,MA, 2003.

[11] Aggelos Bletsas, Ashish Khisti, David P. Reed, and Andrew Lippman. A simple cooperative diversity method based on network path selection. *IEEE Journal on Selected Areas in Communications*, 24(3):659–672, March 2006.

[12] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia. Routing with guaranteed delivery in ad hoc wireless networks. In *3rd International Workshop on Discrete Algorithms and methods for mobile computing and communications*, Seattle, WA, August 1999.

[13] N. Bulusu, J. Heidemann, and D. Estrin. Gps-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5):28–34, Oct. 2000.

[14] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, and J. Zhao. Habitat monitoring: Application driver for wireless communications technology. In *ACM SIGCOMM Workshop Data Comm. Latin America and the Caribbean*, Costa Rica, Apr. 2001.

[15] Szymon Chachulski, Michael Jennings, Sachin Katti, and Dina Katabi. Trading structure for randomness in wireless opportunistic routing. In *ACM SIGCOMM*, Kyoto, Japan, 2007.

[16] J. Chang and L. Tassiulas. Energy conserving routing in wireless ad-hoc networks. In *IEEE INFOCOM'00*, Tel Aviv, Israel, March 2000.

[17] V. Chavtal. *Linear Programming*. W.H. Freeman, New York, 1983.

[18] C.-Y. Chong and S. P. Kumar. Sensor networks: Evolution, opportunities, and challenges. *Proc. of IEEE*, Aug 2003.

[19] R.R. Choudhury and N. H. Vaidya. Mac layer anycasting in ad hoc networks. *SIGMOBILE Mobile Communication Review*, 34(1):75–80, 2004.

[20] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Second Edition*. MIT Press and McGraw-Hill, 2001.

[21] D.S.J. De Couto, D. Aguayo, J. Bicket, and R. Morris. A high-throughput path metic for multi-hop wireless routing. In *ACM MobiCom'03*, San Diego, California, Sept. 2003.

[22] MICA2 Mote Datasheet. http://www.xbow.com. 2004.

[23] Richard Draves, Jitendra Padhye, and Brian Zill. Routing in multi-radio, multi-hop wireless mesh networks. In *MobiCom '04*, 2004.

[24] Henri Dubois-Ferriere, Matthias Grossglauser, and Martin Vetterli. Least-cost opportunistic routing. Technical Report LCAV-REPORT-2007-001, School of Computer and Communication Sciences, EPFL, 2007.

[25] Laurent Eschenauer and Virgil D. Gligor. A key-management scheme for distributed sensor networks. In *CCS '02: Proceedings of the 9th ACM conference on Computer and communications security*, pages 41–47, New York, NY, USA, 2002. ACM.

[26] D. Estrin, D. Culler, and K. Pister. Connecting the physical world with pervasive networks. *IEEE Percasive Computing*, Jan-Mar 2002.

[27] Gregory G. Finn. Routing and addressing problems in large metropolitan-scale internetworks. Technical Report ISI/RR-87-180, USC/ISI, March 1987.

[28] H. Fussler, J. Widmer, M. Kasemann, M. Mauve, and H. Hartenstein. Contention-based forwarding for mobile ad-hoc networks. *Elsevier's Ad Hoc Networks*, 1(4):351–369, Nov. 2003.

[29] Martin Greenberger. Notes on a new pseudo-random number generator. *J. ACM*, 8(2):163–167, 1961.

[30] M. Grossglauser and D. Tse. Mobility increases the capacity of adhoc wireless networks. *IEEE/ACM Transactions on Networking*, 10(4):477–486, August 2002.

[31] P. Gupta and P. R. Kumar. The capacity of wireless networks. *Trans. Inform. Theory*, 46(2):388–404, Mar. 2000.

[32] Kamal Jain, Jitendra Padhye, Venkata N. Padmanabhan, and Lili Qiu. Impact of interference on multi-hop wireless network performance. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 66–80, New York, NY, USA, 2003. ACM Press.

[33] Kyle Jamieson and Hari Balakrishnan. Ppr: Partial packet recovery for wireless networks. In *ACM SIGCOMM*, Kyoto, Japan, August 27C31 2007.

[34] David B. Johnson, David A. Maltz, and Josh Broch. Dsr: The dynamic source routing protocol for multi-hop wireless ad hoc networks. In Charles E. Perkins, editor, *Ad Hoc Networking*, chapter 5, pages 139–172. Addison-Wesley, 2001.

[35] K. Kar, M. Kodialam, T. V. Lakshman, and L. Tassiulas. Routing for network capacity maximization in energy-constrained ad-hoc networks. In *IEEE INFOCOM*, Sanfrancisco, March 2003.

[36] B. Karp and H.T. Kung. Gpsr: Greedy perimeter stateless routing for wireless networks. In *ACM MOBICOM*, Boston, August 2000.

[37] Steven M. Kay. *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall Signal Processing Series, New Jersey, 1998.

[38] Kyu-Han Kim and Kang G. Shin. On accurate measurement of link quality in multi-hop wireless mesh networks. In *MobiCom '06: Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 38–49, New York, NY, USA, 2006. ACM Press.

[39] Murali Kodialam and Thyaga Nandagopal. Characterizing the capacity region in multi-radio multi-channel wireless mesh networks. In *MobiCom '05*, pages 73–87, New York, NY, 2005. ACM.

[40] Fabian Kuhn, Roger Wattenhofer, Yan Zhang, and Aaron Zollinger. Geometric ad-hoc routing: Of theory and practice. In *22nd ACM Symposium on the Principles of Distributed Computing (PODC)*, Boston, July 2003.

[41] P. Larsson. Selection diversity forwarding in a multihop packet radio network with fading channel and capture. *SIGMOBILE Mobile Communication Review*, 5(4):47–54, 2001.

[42] S. Lee, B. Bhattacharjee, and S. Banerjee. Efficient geographic routing in multihop wireless networks. In *MobiHoc*, 2005.

[43] Q. Li, J. A. Aslam, and D. Rus. Online power-aware routing in wireless ad-hoc networks. In *Mobicom'01*, Rome, Italy, July 2001.

[44] K. Lorincz, D. Malan, T. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, S. Moulton, and M. Welsh. Sensor networks for emergency response: Challenges and opportunities. *IEEE Pervasive Computing*, Oct-Dec 2004.

[45] M. K. Marina and S. R. Das. On-demand multipath distance vector routing in ad hoc networks. In *9th International Conference on Network Protocols*, Riverside, CA, November 2001.

[46] S. McCanne and S. Floyd. *The LBNL network simulator*. Lawrence Berkeley Laboratory, 1997.

[47] OPNET. *http://www.opnet.com*.

[48] G. Pei, M. Gerla, and Tsu-Wei Chen. Fisheye state routing: A routing scheme for ad hocwireless networks. In *International Conference on Communications (ICC)*, June 2000.

[49] Charles E. Perkins and Pravin Bhagwat. Dsdv: Routing over a multihop wireless network of mobile computers. In Charles E. Perkins, editor, *Ad Hoc Networking*, chapter 3, pages 53–74. Addison-Wesley, 2001.

[50] Charles E. Perkins and Elizabeth M. Royer. The ad hoc on-demand distance-vector protocol. In Charles E. Perkins, editor, *Ad Hoc Networking*, chapter 6, pages 173–219. Addison-Wesley, 2001.

[51] G. J. Pottie and W. J. Kaiser. Wireless integrated network sensors. *Communications of the ACM*, 43(5):51 – 58, 2000.

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

[53] R. Rivest. *The MD5 Message-Digest Algorithm*.

[54] Lifeng Sang, Anish Arora, and Hongwei Zhang. On exploiting asymmetric wireless links via one-way estimation. In *MobiHoc '07: Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, pages 11–21, New York, NY, USA, 2007. ACM.

[55] A. Savvides, C. Han, and M. B. Strivastava. Dynamic finegrained localization in ad-hoc networks of sensors. In *IEEE/ACM MobiCom*, July 2001.

[56] Karim Seada, Marco Zuniga, Ahmed Helmy, and Bhaskar Krishnamachari. Energy efficient forwarding strategies for geographic routing in wireless sensor networks. In *ACM Sensys'04*, Baltimore, MD, Nov. 2004.

[57] R. C. Shah, A. Bonivento, D. Petrovic, E. Lin, J. van Greunen, and J. Rabaey. Joint optimization of a protocol stack for sensor networks. In *IEEE Milcom*, Nov. 2004.

[58] R. C. Shah, S. Wietholter, A. Wolisz, and J. M. Rabaey. When does opportunistic routing make sense? In *IEEE PerSens*, Mar. 2005.

[59] S. Singh, M. Woo, and C. S. Raghavendra. Power-aware routing in mobile ad hoc networks. In *ACM/IEEE MOBICOM'98*, Dallas, Texas, Oct. 1998.

[60] Michael R. Souryal and Nader Moayeri. Channel-adaptive relaying in mobile ad hoc networks with fading. In *IEEE SECON*, 2005.

[61] Jian Tang, Guoliang Xue, and Weiyi Zhang. Cross-layer design for end-to-end throughput and fairness enhancement in multi-channel wireless mesh networks. *IEEE Transactions on Wireless Communications*, 6(10), 2007.

[62] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Transactions on Information Theory*, 39(2):466–478, 1993.

[63] Nitin H. Vaidya and Matthew J. Miller. A mac protocol to reduce sensor network energy consumption using a wakeup radio. *IEEE Transactions on Mobile Computing*, 4(3):228–242, 2005.

[64] A. Woo and D. Culler. Evaluation of efficient link reliability estimators for low-power wireless networks. Technical report, University of California, Berkeley, Apr. 2003.

[65] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi. A framework for reliable routing in mobile ad hoc networks. In *IEEE INFOCOM*, Sanfrancisco CA, March 2003.

[66] J. Yee and H. Pezeshki-Esfahani. Understanding wireless lan performance trade-offs. *CommsDesign.com*, Nov. 2002.

[67] Kai Zeng, Wenjing Lou, Jie Yang, and D. Richard Brown. On geographic collaborative forwarding in wireless ad hoc and sensor networks. In *WASA'07*, Chicago, IL, August 2007.

[68] Kai Zeng, Wenjing Lou, Jie Yang, and D. Richard Brown. On throughput efficiency of geographic opportunistic routing in multihop wireless networks. In *QShine'07*, Vancouver, British Columbia, Canada, August 2007.

[69] Kai Zeng, Wenjing Lou, and Hongqiang Zhai. On end-to-end throughput of opportunistic routing in multirate and multihop wireless networks. In *Infocom*, Phoenix, AZ, April 2008.

[70] Kai Zeng, Wenjing Lou, and Yanchao Zhang. Multi-rate geographic opportunistic routing in wireless ad hoc networks. In *IEEE Milcom*, Orlando, FL, Oct. 2007.

[71] Kai Zeng, Kui Ren, and Wenjing Lou. Geographic on-demand disjoint multipath routing in wireless ad hoc networks. In *IEEE MILCOM*, Atlantic City, NJ, October 2005.

[72] Kai Zeng, Kui Ren, Wenjing Lou, and Patrick J. Moran. Energy aware efficient geographic routing in lossy wireless sensor networks with environmental energy supply. *Wireless Networks (WINET)*, pages 477–486, 2007.

[73] X. Zeng, R. Bagrodia, and M. Gerla. Glomosim: a library for parallel simulation of large-scale wireless networks. In *Proceedings of PADS'98*, Banff, Canada, May 1998.

[74] Hongqiang Zhai and Yuguang Fang. Impact of routing metrics on path capacity in multirate and multihop wireless ad hoc networks. In *IEEE ICNP*, 2006.

[75] Hongqiang Zhai and Yuguang Fang. Physical carrier sensing and spatial reuse in multirate and multihop wireless ad hoc networks. In *IEEE Infocom*, 2006.

[76] Jihui Zhang, Haitao Wu, Qian Zhang, and Bo Li. Joint routing and scheduling in multi-radio multi-channel multi-hop wireless networks. In *IEEE Broadnets*, 2005.

[77] B. Zhao and M.C. Valenti. Practical relay networks: A generalization of hybrid-arq. *IEEE Journal of Selected Areas in Communications*, 23, 2005.

[78] J. Zhao and R. Govindan. Understanding packet delivery performance in dense wireless sensor networks. In *ACM Sensys'03*, LA,CA, Nov. 2003.

[79] Bin Zhen, Huan-Bang Li, Shinsuke Hara, and Ryuji Kohno. Clear channel assessment in integrated medical environments. *EURASIP Journal on Wireless Communications and Networking*, 2008.

[80] Z. Zhong, J. Wang, G.-H. Lu, and S. Nelakuditi. On selection of candidates for opportunistic anypath forwardning. In *ACM MOBICOM (Poster Session)*, August 2005.

[81] Zifei Zhong and Srihari Nelakuditi. On the efficacy of opportunistic routing. In *IEEE SECON*, 2007.

[82] Zifei Zhong, Juling Wang, and Srihari Nelakuditi. Opportunistic any-path forwarding in multi-hop wireless mesh networks. Technical Report TR-2006-015, USC-CSE, 2006.

[83] M. Zorzi and A. Armaroli. Advancement optimization in multihop wireless networks. *Proc. of VTC*, Oct. 2003.

[84] M. Zorzi and R. R. Rao. Geographic random forwarding (geraf) for ad hoc and sensor networks: energy and latency performance. *IEEE Transactions on Mobile Computing*, 2(4), 2003.

[85] M. Zorzi and R. R. Rao. Geographic random forwarding (geraf) for ad hoc and sensor networks: multihop performance. *IEEE Transactions on Mobile Computing*, 2(4), 2003.

[86] Anatolij Zubow, Mathias Kurth, and Jens-Peter Redlich. Multi-channel opportunistic routing. In *IEEE European Wireless Conference*, Paris, France, April 2007.

[87] M. Zuniga and B. Krishnamachari. Analyzing the transitional region in low power wireless links. In *IEEE Secon'04*, 2004.