ABSTRACT

SUN, LEI. Understanding the Performance and Topology of Multi-Hop Wireless Cognitive Radio Networks. (Under the direction of Dr. Wenye Wang.)

Cognitive Radio Networks (CRNs) are becoming an important supplementary technology to current communication systems. They offer dynamic spectrum opportunities for unlicensed users and greatly improve spectrum usage efficiency. However, CRNs confront many technical challenges that limit their full utilization. In such networks, the communication quality received by each unlicensed user depends highly on the cooperation of other unlicensed users and traffic of coexisting licensed users, which is constrained by many factors such as the opportunistic availability of radio spectrum, the heterogeneous communication capability of users, the mobility of users, the difficulty in user coordination, and the failure of user devices.

We intend to understand the performance and topology of wireless CRNs in this dissertation, which will help us to utilize CRNs effectively, efficiently and reliably. We identify four fundamental performance and topology aspects to investigate, namely, the communication capacity and spectrum sensing, the tempo-spatial limits, the node mobility, and the failure resilience. The study on the first two perspectives attempts to maximize the capacity and minimize the delay of CRNs, while the study on the last two perspectives evaluates and mitigates the impact of user mobility and failure on the network topology.

Specifically, we make the following contributions toward improving the utilization of CRNs. First, we have determined the maximum throughput capacity in large CRNs and designed a new sensing algorithm to achieve the maximum throughput in the order sense. Second, we have identified the sufficient and necessary conditions that a wireless CRN is connected and determined the fastest information dissemination for both connected and non-connected CRNs. Third, we have analyzed the distribution of information dissemination latency in finite CRNs, and its scaling law as the network size grows large, when the secondary users are mobile under a general mobility framework. Last, we have characterized the spread of user failures, identified
the formation and structure of Blackholes (components of failed nodes) and suggested strategies to maintain global communications in large CRNs, in the face of node failures. The work in this dissertation improves our understanding and enhances the potential applications of wireless CRNs.
Understanding the Performance and Topology of Multi-Hop Wireless Cognitive Radio Networks

by
Lei Sun

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Computer Engineering
Raleigh, North Carolina
2012

APPROVED BY:

Dr. Michael Devetsikiotis

Dr. George N. Rouskas

Dr. Huaiyu Dai

Dr. Wenye Wang
Chair of Advisory Committee
DEDICATION

To my parents.
BIOGRAPHY

Lei Sun was born in Jiangxi, China in 1981. He received the M.E. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2005 and the M.S. degree in Statistics from Stephen F. Austin State University, Nacogdoches, TX, USA, in 2007, respectively. Since 2007, he has been working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. His research interests include modeling and analysis of large-scale cognitive radio networks, network resilience to mobility and failure, and network optimization. He has been a member of IEEE since 2007.
ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Professor Wenye Wang for the continuous support of my Ph.D study and research. Her guidance helped me in all the time of research and writing of this thesis. I feel myself so fortunate to have the opportunity to learn from her research passion, vision, wisdom, and much more, the attitude toward life. I believe that the knowledge and skills that she has imparted to me are far beyond this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my advisory committee: Professor Michael Devetsikiotis, Professor George N. Rouskas and Professor Huaiyu Dai, for their encouragement and support. Their valuable and insightful suggestions and comments have helped me broaden my research view and improve my work quality. I am also grateful to all the professors that have taught me in and out of classes at NC State for imparting their knowledge, which allows me to be prepared for my future career.

I also thank my fellow labmates in Netwis Group: Ming Zhao, Fei Xing, Nurcan Tezcan, Avesh Algawal, Shawqi Kharbash, Xi Yu, Entong Shen, Chi Yi, Zhuo Lv, Mohit Khanna, Yujin Li, Haiyang Zheng, Xiang Lu and Jianhua Zhang. It has been my great pleasure to work closely with them. My research discussions with them have always been one of my most important sources of fresh ideas that I have benefited from significantly. The time spent with them will remain a great memory for me.

Last but not the least, I would like to thank my family: my parents Zhongping Sun and Huaiying Zeng, for giving birth to me at the first place and supporting me spiritually throughout my life.
# TABLE OF CONTENTS

List of Tables ................................................. ix

List of Figures ................................................. x

Chapter 1 Introduction ......................................... 1
  1.1 Motivation ................................................. 1
  1.2 Research Objectives ........................................ 9
    1.2.1 Capacity Limits and Spectrum Sensing .................. 9
    1.2.2 Tempo-spatial Limits of Information Dissemination ....... 10
    1.2.3 Latency Distribution and Scaling with Generic Node Mobility .... 12
    1.2.4 Network Resilience against Blackholes .................. 13
  1.3 Contributions ............................................. 14
  1.4 Organization .............................................. 15

Chapter 2 Capacity Limits and Spectrum Sensing Algorithms in Large Cognitive Radio Networks ............................................. 16
  2.1 Motivation and Related Work ................................ 17
  2.2 Constraints of Sensing Algorithms .......................... 19
    2.2.1 Spectrum Sensing ....................................... 19
    2.2.2 Common Channels Constraint ............................ 21
    2.2.3 Interference Constraint .................................. 22
    2.2.4 Relay Schemes and Connectivity Constraints ............... 23
    2.2.5 Discussions On Spectrum Sensing ........................ 24
  2.3 Characteristics of CRNs ..................................... 24
    2.3.1 System Models .......................................... 24
    2.3.2 Relay Schemes In CRNs ................................... 26
    2.3.3 Shared Common Channel Probability of Secondary Nodes ....... 28
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.4</td>
<td>Connectivity Constraints in CRNs</td>
<td>30</td>
</tr>
<tr>
<td>2.4</td>
<td>Throughput in Large-scale CRNs</td>
<td>31</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Definitions</td>
<td>32</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Throughput with Cooperative Relays</td>
<td>33</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Throughput with Hybrid Relay</td>
<td>37</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Discussions On Throughput in CRNs</td>
<td>42</td>
</tr>
<tr>
<td>2.5</td>
<td>Design of Sensing Algorithms To Achieve Maximum Throughput</td>
<td>42</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Properties of Sensing Algorithms to Achieve Maximum Throughput</td>
<td>43</td>
</tr>
<tr>
<td>2.5.2</td>
<td>An Example Sensing Algorithm CASA to Achieve Maximum Throughput under Cooperative Relay</td>
<td>44</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Performance Evaluation</td>
<td>47</td>
</tr>
<tr>
<td>2.6</td>
<td>Summary</td>
<td>52</td>
</tr>
</tbody>
</table>

Chapter 3 The Tempo-spatial Limits of Information Dissemination in Cognitive Radio Networks

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Motivation and Related Work</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>Preliminaries and Problem Definition</td>
<td>56</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Assumptions and Models</td>
<td>56</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Percolated Cognitive Radio Networks</td>
<td>58</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Limits of Information Dissemination</td>
<td>60</td>
</tr>
<tr>
<td>3.3</td>
<td>How Far Can Information Be Disseminated in A CR Network?</td>
<td>63</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Critical Phenomenon</td>
<td>63</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Conditions for Percolated Cognitive Radio Networks</td>
<td>64</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Analysis of Dissemination Radius</td>
<td>69</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Discussion and Applications</td>
<td>70</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Simulation Results of Dissemination Radius</td>
<td>72</td>
</tr>
<tr>
<td>3.4</td>
<td>How Fast Can Information Be Propagated in a Cognitive Radio Network?</td>
<td>75</td>
</tr>
</tbody>
</table>
Chapter 4 The Distribution and Scaling Law of Dissemination Latency in Mobile Cognitive Radio Networks

4.1 Motivation and Related Work ................................................. 89
4.2 System Models and Problem Formulation .................................. 91
  4.2.1 Network Models .......................................................... 91
  4.2.2 Problem Formulation ...................................................... 95
4.3 Main Results ........................................................................... 97
4.4 The Distribution of $T_d$ in Finite Networks ................................. 99
  4.4.1 Distribution of $T_d$ under EIHP Mobility $\mathcal{M}(\Phi_E, \Psi, \alpha)$ .... 100
  4.4.2 Distribution of $T_d$ under PIHP Mobility $\mathcal{M}(\Phi_P, \Psi, \alpha)$ .......... 104
  4.4.3 Distribution of $T_d$ under HHP Mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$ ............ 105
4.5 The Scalability of $T_d$ in Large CRNs ...................................... 110
  4.5.1 Scalability of $T_d$ under EIHP and PIHP Mobility ...................... 111
  4.5.2 Scalability of $T_d$ under HHP Mobility .................................. 114
4.6 Summary .............................................................................. 118
# Chapter 5 Understanding Blackholes in Large-Scale Cognitive Radio Networks under Generic Failures

5.1 Motivation and Related Work ................................. 120
5.2 System Models and Problem Formulation .................. 123
  5.2.1 Preliminary ......................................... 123
  5.2.2 Network and Failure Models ......................... 124
  5.2.3 Problem Formulation ................................. 127
5.3 Results and Applications .................................. 128
  5.3.1 Main Results ....................................... 128
  5.3.2 Applications ....................................... 130
5.4 How Large is a Blackhole? ................................. 131
  5.4.1 The Distribution of $|O_{v_1}|$ .................... 131
  5.4.2 The Expected Value of $|O_{v_1}|$ ................. 134
5.5 Is a Large CRN Blackhole Resilient? ...................... 137
  5.5.1 Challenges and Differences with Earlier Work .... 138
  5.5.2 Sufficient Condition for a Resilient CRN ......... 139
5.6 Simulations ............................................... 143
5.7 Appendix: Calculation of Critical Probability $p_c^\Omega$ ... 146
5.8 Summary ................................................. 148

# Chapter 6 Conclusion and Future Directions .................. 150
6.1 Conclusion ................................................. 150
6.2 Future Directions ....................................... 152

References .................................................. 154
LIST OF TABLES

Table 2.1 Connectivity-Agile Sensing Algorithm. .......................... 47
LIST OF FIGURES

Figure 2.1 Common channel constraint. ............................................ 22
Figure 2.2 Interference and connectivity constraints. .......................... 23
Figure 2.3 Node positions and relay schemes. The perimeter is an extended version of the grids. .......................................................... 27
Figure 2.4 The illustration of ALSR scheme. ......................................... 35
Figure 2.5 The illustration of MCDS-based routing scheme. ..................... 40
Figure 2.6 The Effect of Shannel Channel Constraint. ............................ 49
Figure 2.7 Comparison of Sensing Algorithms. ..................................... 50
Figure 2.8 Performance of Connectivity-Agile Sensing Algorithm (CASA). . . . 51

Figure 3.1 Cognitive Radio Network Model: secondary nodes and PR nodes . 57
Figure 3.2 Illustration of information dissemination in a percolated CR network. . 60
Figure 3.3 An illustration of dissemination area after long time $t$. ................ 61
Figure 3.4 The triangular lattice $\mathcal{D}$ with the “flower” $\mathcal{F}_s = ABCDEF$. . 65
Figure 3.5 The illustration of Theorem 3.1. ......................................... 68
Figure 3.6 The percolation region for an example emergency network. ............ 70
Figure 3.7 Percolation regions under different interference range $R_I$. ............... 72
Figure 3.8 An example of percolation conditions in terms of $\lambda$, $m$ and $\{\lambda_{pk}\}_{k=1}^m$. ......................................................... 72
Figure 3.9 Snapshots of packets disseminating gradually when $\lambda_{c,2} < \lambda < \lambda_{c,c}$. ....... 72
Figure 3.10 Dissemination radius and propagation speed analysis. .................... 80
Figure 3.11 Snapshots of packets disseminating locally within finite time when $\lambda < \lambda_{c,w}$. ... 82
Figure 3.12 Snapshots of packets disseminating globally within finite time when $\lambda > \lambda_{c,c}$. .... 82
Figure 3.13 Dissemination radius and propagation speed analysis. .................... 82
Figure 3.14 Link Correlation Coefficient $C_{LCC}$ Calculation. .................... 84
Figure 3.15 A triangular lattice and its dual, drawn with a dashed line. ................. 85

Figure 4.1 Communication links for secondary users. ................................ 92
Figure 5.13 A finite open cluster at the origin, surrounded by a closed circuit in the dual lattice.
Chapter 1

Introduction

1.1 Motivation

In recent years, wireless networks attracted increasing attention from researchers since they provide a convenient means of networking service, supplementary to the conventional wired (infrastructure-based) networks. In fact, various types of wireless networks have already become an integral part of our daily lives, changing the traditional ways for communication in our many work and social activities. For example, almost everyone in the world is benefiting from the communication convenience offered by the wireless cellular technology. Especially when the high-rate data transmission service is provided in the current 3G networks, the usage of and dependence on cellular networks are becoming unprecedentedly high. The WiMAX technology allows the users to access information on the Internet using portable mobile broadband devices, providing an alternative solution to substitute the traditional cable and DSL connections. The high-speed Internet access offered by WiMAX provides a number of real-time services to the users including, for example, Voice-over-IP and IPTV. The wireless LAN and bluetooth technologies are presenting a convenient data communication paradigm by removing the requirement of wired connections between different electronic devices.

The fast pace in the wireless network research and deployment is significantly changing our
world in many aspects and the use of wireless networks is expected to be even more ubiquitous and pervasive in the future. However, wireless networks consume a critical yet scarce resource, **spectrum**. Today’s wireless networks are regulated by a fixed spectrum assignment policy, i.e., the spectrum is regulated by governmental agencies like the Federal Communications Commission (FCC) and is assigned to license holders or services on a long term basis in large geographical regions [1]. A challenging question is that when the licensed spectrum is infeasible, whether we can still provide networking service? Cognitive Radio Networks (CRNs) present a positive answer.

A “Cognitive Radio” is a radio that can change its transmitter parameters (operating spectrum, modulation, transmission power and communication technology) based on interaction with the environment in which it operates [2]. In Cognitive Radio Networks (CRNs), there are two types of spectrum users: (i) **primary users** who have license from the regulator and thus have priority to utilize spectrum, and (ii) **secondary users** who are equipped with Cognitive Radios and opportunistically utilize spectrum without interfering with the coexisting primary users. A recent report from FCC reveals that under the traditional static spectrum assignment policy, merely 5% ∼ 15% of the spectrum is utilized on average [1]. By exploiting spectrum when its licensed users do not utilize it, not only do CRNs support more wireless applications, but they also greatly improve the utilization efficiency of spectrum.

One of the most interesting potential applications of an CRN is in a military radio environment. CRNs can enable the military radios choose arbitrary, intermediate frequency bandwidth, modulation schemes, and coding schemes, adapting to the variable radio environment of battlefield. Also military networks have a strong need for security and protection of the communication in hostile environment. CRNs could allow military personnel to perform spectrum band for themselves and their allies.

*Emergency network* is another possible application. In the case of natural disasters, which may temporarily disable or destroy existing communication infrastructure, emergency personnel working in the disaster areas need to establish *emergency networks*. Since emergency networks
deal with the critical information, reliable communication should be guaranteed with minimum latency. In addition, emergency communication requires a significant amount of radio spectrum for handling huge volume of traffic including voice, video and data. CRNs can enable the usage of the existing spectrum without need for an infrastructure and by maintaining communication priority and response time.

In addition, the primary network can provide a leased network by allowing opportunistic access to its licensed spectrum with the agreement with a third party without sacrificing the service quality of the primary user. For example, the primary network can lease its spectrum access right to a mobile virtual network operator. Also the primary network can provide its spectrum access rights to a regional community for the purpose of broadband access.

Although the potential advantageous usage of CRNs, however, there exist a large number of technical challenges that require satisfactory solutions in order for us to make the full utilization of the CRNs. It is worthy of noting that Cognitive Radio is much more powerful and flexible than the traditional hardware-based radio, which makes the performance analysis of CRNs different and more complicated than that of wireless networks. Thus many existing results on wireless networks in the literature may not be applicable to CRNs.

First of all, since Cognitive Radio is capable of reconfiguring radio frequency and switching to newly-selected channels, the number of channels that can be accessed by CRNs is much larger. In addition, due the the physical diversity and dynamic traffic of primary users, the subset of channels that can be used by each secondary users without interfering with the coexisting primary users may be different and vary over time. In wireless ad hoc networks, wireless nodes are assumed to communicate with each other using a single common channel. In contrast, secondary users in CRNs need to sense multiple channels and find such a common channel before communication. Spectrum sensing incurs overhead to the performance, e.g, capacity and delay, of CRNs. What is more, failure to find such common channel will cause some nodes disconnected from the network, which further complicates the problem.

Heterogeneous nodes impose another challenge. Note that there exist two types of heteroge-
neous users in CRNs, which diversifies routing and scheduling schemes and amplifies interference in CRNs. Specifically, secondary users may either use other secondary users, or primary users as relays. Since primary and secondary users have different bandwidth, transmission power and spatial density, hybrid relay modifies the transport capacity and information delivery speed among secondary users. And by allowing primary users to help relay data of secondary users, hybrid relay may also improves the connectivity of secondary users. Furthermore, since primary users have higher priority to utilize spectrum and secondary users cannot interfere with the co-existing primary users, the communication between secondary users is constrained by both primary users and other secondary users. To be specific, there are two types of interference, i.e., primary-to-secondary and secondary-to-secondary interference, for communication between between secondary users. Thus we need a new interference model to study the information delivery among secondary users.

Among all the factors that influence the performance of CRNs, node energy plays an important role. To save energy, or limit the interference to the primary users, secondary users need to adapt their transmission energy to the environment, such as the transmitting activities of neighboring primary users, from time to time. Thus the node transmission radius, which is closely coupled with the transmission power, changes over time, which leads to dynamic communication links in the network, due to the fact that two wireless nodes may communicate with each other directly only when they are within each others’ transmission radius. Furthermore, since the secondary users are powered by batteries and packet transmission and reception expend the limited battery energy gradually, a secondary user will finally drain out its battery. At the time of energy depletion, the node fails, causing link disruption to the other nodes that use it as a relay, which creates additional dynamics in network topology.

Node mobility [3–13] impacts the performance of CRNs too. In CRNs, secondary users are usually mobile, to avoid interference to the primary users, or to obtain better channel opportunities. When a node is mobile, its neighbors change over time. Since the moving node relies on its neighbors to relay its packets, node movement results in frequent network topology changes
and node communication interruptions. As the current communication links are broken and new links are established later, network overhead, packet loss and packet delay are introduced inevitably. The mobility impact is more pronounced in large-scale networks than small-scale ones as the average path length is longer in the former and any link disruption causes path failure.

Like all other complicated systems, wireless CRNs have to deal with failures. Such unavoidable faults can be brought out by malfunctions of electrical devices, energy depletion, natural disasters (fire, river overflow, earthquake, etc) or adversarial attacks (a bomb explosion for example). Communications may be disabled by jamming, traffic congestion or energy depletion. In addition, causal relations often exist among failures, i.e., some failures happen as a result of other earlier failures. One example of such correlated failures is traffic overloading and energy depletion [14], that is, when a node fails to deliver packets, the incoming and outgoing traffic is redistributed to the neighboring nodes. Some neighbors may work under heavy traffic loads, resulting in early energy depletion and node failures. Such correlation among failures and cascading effects lead to Blackholes (i.e., components of failed nodes) in the network, where information cannot be transmitted or forwarded.

Hence, though the Cognitive Radio Networks are a promising type of network technology, there are many problems that need to be solved before they can provide effective, efficient and reliable services. In this dissertation, we are dedicated to understand the performance and dynamic topology of CRNs, which lays the foundation for the further efforts in designing and operating CRNs in an efficient way. Specifically, motivated by the objective to optimize the performance of CRNs and thus best utilize them, we first place our emphasis on the theoretical understanding of network performance and then design practical solutions to achieve the limiting performance determined by our theoretical analysis. The solutions are practical in the sense that they do not make unrealistic assumptions and thus they are implementable. We target two fundamental performance metrics, throughput [15–29] and delay [30–39]. Ideally, we expect a network to be operated with large capacity and short delay. Additionally, when mobility and
failure occur, we also expect the network continue working normally, even the network topology changes frequently.

In wireless cognitive radio networks, spectrum sensing [40, 41] determines which portions of the spectrum is available (i.e., currently not used by licensed users), selects the best available channel to transmit over and detects the presence of licensed users. Spectrum sensing has been recognized as the most fundamental element in CRNs due to its crucial role of discovering “spectrum opportunity”, which refers to a time duration on a channel during which the channel is not used by primary users and thus can be utilized by secondary users. To maximize the discovery of spectrum opportunity, and eventually improve the spectrum utilization efficiency, many sensing algorithms have been proposed for CRNs [42–50]. However, as one of the most important benefits of cognitive radio networks, the achievable throughput remains an open question, which is supposed to be resulting from efficient spectrum usage. Therefore, we will take a new look at sensing algorithms. We first formally derive the throughput limit in large-scale CRNs and then propose a class of connectivity-agile sensing algorithms to achieve such a limit. Compared to the existing sensing algorithms, our work ensures the maximum throughput performance in the order sense in CRNs.

Though our study on spectrum sensing minimizes the impact of spectrum diversity and guarantees the network throughput, network performance is not measured solely by its throughput capacity. Packet delay is also an important performance metric. Packet delay has been a long-established topic in network research. It measures the time needed to transport a packet from its source node to its destination node. Connectivity [51–55], or coverage, is another equally important metric. It measures the largest area, or equivalently the longest distance, that a packet can be transmitted in the network. Delay characterizes the information delivery of networks in temporal dimension. Connectivity, on the other hand, studies the information delivery in spatial dimension. It is worthy of noting that the ultimate goal of CRNs is supposed to offer networking services with satisfactory performance when licensed spectrum is infeasible. Therefore, understanding how packets disseminate and their temporal and spatial limits, not only
provides knowledge of the *achievable* benefits of information dissemination in CRNS, but is also beneficial to the deployment, design and applications of such networks. Temporal and spatial limits of packet delivery have been studied for wireless ad hoc networks or sensor networks. For example, references [56–62] explore the conditions for connectivity or percolation in order to ensure the information can be *disseminated* to the whole network. In addition, information propagation speed or delay has been investigated in recent works [61,62], which categorize the delay into bandwidth-incurred propagation delay and topology-incurred delay. However, due to the unique challenges of CRNs discussed above, these existing results on wireless ad hoc networks may not be applicable to CRNs. Therefore, the lack of understanding of tempo-spatial limits of information dissemination motivates us to study the delay and connectivity of CRNs.

In addition, note that performance study must be based on a specific network topology. Therefore, as one of the most factors that will cause topology dynamics, node mobility plays a critical role on the network performance, e.g., packet delay, which has been evidenced by earlier results in wireless ad hoc networks. For instance, the seminal work [63] showed that mobility can improve the capacity in large wireless ad hoc networks at the price of large delay. This result is obtained by assuming that nodes move according to an ergodic process that is equally likely to visit any portion of the network area. That is, the nodes are *spatially homogeneous*. With the similar assumptions, capacity-delay tradeoffs have been extensively studied under various mobility models, such as the i.i.d model [64], the Brownian motion [65], the reshuffling model [66] and variants of random walk and random way-point models [67,68]. Later on, spatial inhomogeneity has been taken into account in [69,70] where the nodes are either restricted to move within an randomly chosen cell or the coverage of a home point. These studies motivate an interesting question about the latency under *general mobility*. Besides network topology, network size also impacts the network performance. A common assumption in most of the performance studies of wireless networks in the literature [57,58,61,62,71–74] is that the number of nodes is infinite or approaches to infinity. It is evident that the asymptotic results, though, provide good insights into network performance, may not explain the latency properties when
the number of nodes in real applications is finite. In other words, the stochastic properties of latency distribution in finite networks lead to understanding of real networks, rather than the large networks. Putting all together, we are interested in the latency distribution in finite networks, and the scaling law for large networks with infinite number of secondary nodes under general mobility where spatial inhomogeneity is considered in addition to common features of a variety of mobility models.

Failure resilience [75–77] is another important challenge in large CRNs. The current studies on the failures in wireless networks have mainly focused on either topology transitions in the face of isolated failures or prevention of a particular correlated failure spreading to the entire networks. Though the existing work in the literature helps improve the network resilience to random failures and understand failure spreading, there lacks a general understanding of the failure impact on the overall network topology and functions. In particular, failure correlation [78,79] has been recognized as one of the most important factors for the occurrence of Blackholes (i.e., components of failed nodes) and Xu et al. [79] further studied how an initial failure may incur a giant hole spanning over the entire network. Given its detrimental consequences, the occurrence of giant hole needs to be avoided in the initial network design [79] such that node failures can result in many finite holes in the network. Understanding the properties of Blackholes in the CRNs, or in particular, investigating structure and size of Blackholes, is of great importance in the design of basic networking operations. And the network resilience to Blackholes is hence necessary in order to prevent the network from potentially catastrophic failure impact.

Given the importance of these research problems, we are hence motivated to explore further the performance and topology properties of large-scale CRNs to gain a comprehensive understanding towards the maximal network utilization. We intend to determine the solutions for high traffic throughput, fast packet transportation, stable network structure and resilient Blackholes capability. Our work is helpful for the designers and operators of large-scale CRNs to evaluate, improve and optimize their network usage.
1.2 Research Objectives

In line with the research problems identified above, we will investigate four specific problems related to the performance and topology of CRNs, namely, the maximum throughput, the information dissemination latency, the topological dynamics under generic node mobility, and the network resilience against Blackholes. We next elaborate on these individual problems respectively.

1.2.1 Capacity Limits and Spectrum Sensing

Capacity has been a long-established topic in wireless networks, which measures the total volume of packets that can be served simultaneously. Many studies in the literature [56,65,72–74,80–82] have investigated the capacity of wireless ad hoc networks, in which the capacity bound has been derived and methods have been proposed to achieve at least a constant fraction of the capacity bound. However, unique features in CRNs, such as heterogeneous architecture and opportunistic spectrum access, make the existing results not applicable. For example, *spectrum sensing*, which is required before communication among secondary users, incurs sensing overhead over throughput. Moreover, two types of interference in CRNs make the collision-free transmission method used to obtain the theoretical bound on capacity in wireless ad hoc networks infeasible. In addition, a common channel is assumed to be used by all users in wireless ad hoc networks, which is used in constructing joint routing and scheduling schemes to achieve optimal network capacity, but each secondary user in CRNs may use different channels. Hence, we are interested in determining whether the same order optimal capacity is still achievable in CRNs, although the challenges. If the answer is positive, we will be able to provide the same quality of service to wireless users without spectrum license in terms of capacity. We attempt to achieve the following research objectives:

- The upper bound on throughput capacity will be determined. This bound is derived by assuming an ideal spectrum sensing without sensing overhead and thus it reflects the
maximum throughput capacity allowed in a wireless CRNs under any spectrum sensing algorithms.

- The guideline will be derived on the design of spectrum sensing algorithms to achieve the maximum throughput in the order sense.

- A Connectivity-Agile Sensing Algorithm (CASA) based on the guideline will be designed, whose performance will be evaluated by ns-2 simulation studies.

The theoretical analysis of the maximum throughput lays the foundation for designing sensing algorithms with guaranteed throughput performance, which is not seen in the existing sensing algorithms. We first examine the unique characteristics in CRNs and study their effect on achievable throughput. Through the mathematical reasoning, we then formally derive the theoretically achievable throughput limit of CRNs. The analytical results will reveal the conditions to achieve the maximum throughput and thus provide the guideline for spectrum sensing design that maximizes the network throughput. Finally, we propose a class of connectivity-agile sensing algorithms with guaranteed throughput performance. Our study on spectrum sensing and capacity limits will advance our understanding of the fundamental properties of spectrum sensing algorithms and optimize network performance from the throughput maximization perspective.

1.2.2 Tempo-spatial Limits of Information Dissemination

Besides network capacity, packet delay and connectivity are equally important metrics for network performance. Delay measures the time needed to transport a packet from its source node to its destination node, and connectivity evaluates the network capability of providing global services. The understanding of system capacity is not able to reveal *how fast* and *how far* a packet can be disseminated in a CRN, in temporal and spatial domains, respectively. The answers to these questions offer a straightforward interpretation of the potentials of CRNs for time-sensitive applications. For example, when a CRN is used for emergency rescue in the af-
termath of disasters or traffic accidents (e.g., vehicular networks), we need to ensure that help or warning messages can be disseminated to a sufficiently large area, and estimate how long it takes for such information to reach a chosen destination, which becomes more important than other performance metrics, such as the total network capacity in these circumstances.

Particularly, we address the following questions: (i) for a large multi-channel CRN, how far can a packet originated from an arbitrary node be disseminated? (ii) When a packet can be disseminated to a sufficiently large area, how long does it take this packet to reach a chosen destination? To tackle these problems, we define two new metrics, the disseminating radius \( \|\mathcal{L}(t)\| \) and the propagation speed \( S(d) \) to study the spatial and temporal limits, respectively. The former is the maximum Euclidean distance that a packet disseminates in time \( t \) and can be used to characterize the dissemination area; and the latter one is the speed that a packet transmits between a source and destination at distance \( d \) apart, which can be used to interpret the end-to-end delay. We intend to achieve the following research objectives regarding disseminating radius \( \|\mathcal{L}(t)\| \) and the propagation speed \( S(d) \):

- The sufficient and necessary conditions under which there exist theoretical limits of information dissemination will be determined, by identifying the correlation between spatial density of primary nodes, secondary nodes, and the number of accessible channels.
- When the packet cannot percolate to infinite area, the distribution of dissemination radius \( \|\mathcal{L}(\infty)\| \) is investigated.
- When the packets are able to reach infinite area, i.e., the dissemination radius approaches to infinity, the propagation speed \( S(d) \) is evaluated.

We first determine the constraint on network connectivity (percolation) imposed by the network factors including node transmission powers, transmission interference, and spectrum sensing. Through the mathematical analysis, we identify the sufficient and necessary conditions that enable us to maintain global communication and achieve the fastest packet transportation. We also consider the cases in which the desired conditions are not available. In these situations,
we determine the distribution of the maximum dissemination distance. Finally, we use the identified conditions to investigate the limiting propagation speed when information can and cannot be disseminated to the whole network, respectively. This study will optimize network performance from the delay and connectivity perspective.

1.2.3 Latency Distribution and Scaling with Generic Node Mobility

Node mobility creates a significant impact to the network topology and thus performance. Due to the limited transmission power, a node may communicate only to the other nodes that are close enough within the reach of its transmission. When a node moves, its neighbors change over time. Since the moving node relies on its neighbors to relay its packets and the moving node may also be actively forwarding the packets from others, node movement results in frequent network topology changes and communication interruptions. When the broken communication links and paths are repaired, network overhead, packet delay and packet loss are incurred inevitably. In addition, network size is another important factor. In most of the existing works on latency, it is commonly assumed that the network is large (i.e., consists of infinite number of wireless nodes.) It is evident that the asymptotic results, though, provide good insights into network performance, may not explain the latency properties when the number of nodes in real applications is finite. In other words, the stochastic properties of latency distribution in finite networks lead to understanding of real networks, rather than the large networks. We are hence motivated to study network topology in a mobile environment and attempt to achieve the following research objectives:

- A general mobility framework that captures most characteristics of the existing mobility models and takes spatial heterogeneity into account will be provided.

- The distribution of the dissemination latency in a finite CRN where secondary users are mobile under the general mobility framework will be determined.

- As the network size grows large, the latency scaling property will be identified.
The mathematical modeling and analysis lay the foundation for understanding topology and latency in mobile CRNs. To facilitate our study, we first define a general mobility framework depending on some fundamental parameters, and by adapting these parameters, this generic framework cover most of the existing mobility models. Then we analyze the network connectivity and derive the distribution of the dissemination latency when secondary users are mobile under this generic model and network size is finite. Finally, we investigate the scaling law of the latency as the network grows large.

1.2.4 Network Resilience against Blackholes

In wireless networks, node failures occur unavoidably, which impose a significant threat to the correct network functioning. Such unavoidable faults can be brought out by malfunctions of electrical devices, energy depletion, natural disasters (fire, river overflow, earthquake, etc) or adversarial attacks (a bomb explosion for example). Communications may be disabled by jamming, traffic congestion or energy depletion. When failures occur, not only the failed nodes suffer from the inability to communicate, but also the rest of the network may be impacted, as the failed nodes might serve as intermediate routers for the other portion of the network. Ideally, we want a wireless network to be robust such that it continues working when failures occur. In the scenario that the failures happen independently, we may ensure network resilience by deploying extra nodes in the network as backup. However, causal relations often exist among failures, i.e., some failures happen as a result of other earlier failures. Such correlation among failures and cascading effects lead to Blackholes (i.e., components of failed nodes) in the network. In order to design effective strategies to enhance the counter-failure capability of large-scale CRNs, we must understand the size and structure of Blackholes first. Therefore, we target the following research objectives:

- Generic models will be established to formulate the occurrence of random failures and their subsequent spreading, which incur Blackholes.
- Based on the failure models, we will determine the distribution and expected size of
Blackholes.

- We will identify sufficient condition when a CRN is resilient to Blackholes.

- The analytical results on Blackholes resilience can be used as a foundation for designing algorithms to locate and bypass Blackholes.

The first step in determining the scope of Blackholes is to understand the occurrence of random failures and the logical connections between these failures. We thus propose mathematical frameworks to model the appearance of Blackholes due to failure interdependence and spreading. Given the models, we then determine the size and structure of each Blackhole. Finally, we will provide resilient-Blackhole conditions in large-scale CRNs.

1.3 Contributions

Throughout this dissertation research, we contribute toward a comprehensive understanding of the fundamental performance aspects of Cognitive Radio Networks (CRNs). As the first step toward this objective, we have identified four aspects of CRNs, namely, capacity and sensing, information dissemination in tempo-spatial domains, latency distribution and scaling under generic mobility, and network resilience against Blackholes, to study. Because these performance and topology indicators reflect the fundamental expectations for wireless CRNs, we believe that our work provides useful insights for the network designers and operators to apply CRNs effectively, efficiently and reliably. In particular, our theoretical analysis on network capacity and sensing provides a practical and efficient solution to maximize the throughput in large-scale CRNs, our result on information dissemination in tempo-spatial domains determines how far and how fast information can be disseminated in large-scale CRNs, the result on latency distribution and scaling provide condition to maintain global communication and determines the fast message dissemination strategy in a mobile environment, and our result on the failure resilience provides a theoretical foundation to understand the formation of Blackholes and their
impact on network topology. The analytical results and proposed solutions in this dissertation can be used to evaluate, improve and optimize the utilization of CRNs.

1.4 Organization

The rest of this dissertation is organized as follows. Chapter 2 presents our research result on the spectrum sensing and capacity limits in large wireless CRNs and our solution to maximize network throughput. Chapter 3 investigates the dissemination latency and connectivity via studying the tempo-spatial limits of information dissemination in large CRNs. Chapter 4 explores the dissemination latency when the network is finite and secondary nodes are mobile under general mobility. Chapter 5 characterizes the size and structure of Blackholes due to the spreading properties of correlated failures, and provides network resilience conditions. Finally, Chapter 6 summarizes our research results and discusses the possible extension directions.
Chapter 2

Capacity Limits and Spectrum Sensing Algorithms in Large Cognitive Radio Networks

Spectrum sensing is essential to secondary nodes which may opportunistically use spectrum that are allocated to licensed or primary nodes without interrupting their communications. Prior research has focused on the design of sensing algorithms to improve spectrum efficiency. However, as one of the most important benefits of CR networks, the achievable throughput remains an open question, which is supposed to be resulting from efficient spectrum usage. Therefore, in this work, we first formally derive the throughput capacity limit in large-scale CRNs by examining unique characteristics of CRNs, such as connectivity constraints and hybrid relay schemes. The analytical result provides the fundamental understanding of the constraints on network capacity. Inspired by this understanding, we propose a class of connectivity-agile sensing algorithms to achieve such a limit. To validate our results, we design a Connectivity-Agile Sensing Algorithm (CASA) based on a cooperative relay scheme and show that CASA performs very well and yields higher throughput than the existing sensing algorithms.
2.1 Motivation and Related Work

Spectrum sensing has been recognized as the most fundamental element in CRNs due to its crucial role of discovering *spectrum opportunity*, which refers to the time duration during which the channel is not used by primary nodes. To maximize the discovery of spectrum opportunity, and eventually improve the throughput for secondary nodes, many sensing algorithms have been proposed for multi-channel CR networks [42–49], which in general target to optimize throughput either with the constraint of protecting primary nodes from interference or the trade-off between selecting the *best* channels and minimizing sensing overhead.

In the former category, the problem is to find optimal sensing periods such that primary nodes are protected against interference from secondary nodes, while making efficient reuse of legacy spectrum. Specifically, in [42, 43], a channel is modeled as an ON/OFF source, where an ON period represents the time duration during which primary nodes are actively using their channels. Secondary nodes are allowed to utilize the channel only during primary nodes’ OFF periods. With the knowledge of the distribution of the ON and OFF period, [42] proposes a sensing period optimization algorithm to maximize the utilization of OFF periods; [43] proposes a sensing period adaptation to minimize sensing overhead while satisfying the detectability requirements. In the latter category, throughput maximization is achieved by the design of sensing order and sensing stopping time, possibly along with prior information about the channel utilization. By sensing, secondary nodes can obtain information about channel quality and thus make better decisions regarding the channel for transmission, which is in turn at the cost of sensing overhead. In particular, considering multiple channels that may have different quality statistics and can be sequentially sensed with channel-dependent costs, this type of sensing algorithms focus on which channels to sense, in what order, when to stop and upon stopping which channels to use for transmission [45–49]. For example, optimal strategies for determining which channels to sense and which channel to use for transmission are derived in [49], with the assumption of independent but not identical channels for transmission. [48] analyzes a problem where channels can only be used immediately after sensing and that unsensed channels cannot
be used for transmission. Under these conditions the problem is formulated as an optimization problem of finding the stop sensing time for a given ordering of channels to be sensed.

However, as one of the most important benefits from CRNs, it is not clear what the achievable throughput is, and how to achieve the maximum throughput given the unique characteristics of CR networks, such as connectivity constraints and collaborative relay schemes. Furthermore, these unique characteristics, and thus the theoretical limit of achievable throughput may potentially impact our expectation of sensing algorithms significantly. For example, one commonly used assumption in existing sensing algorithms is that common channels are simultaneously available to transmitters and receivers. However, this assumption is rarely true for CRNs due to the dynamic traffic of primary nodes and the opportunistic nature of spectrum access and physical diversity of secondary users [83]. Therefore, instead of finding the best channel as in the existing algorithms, finding the best common channels among transmitter-receiver (Tx-Rx) pairs is the objective of sensing algorithms in secondary networks. Moreover, the existing sensing algorithms only focus on direct transmission between one (or multiple) transmitter-receiver (Tx-Rx) pair (pairs) and aim to maximize the throughput of the Tx-Rx pair (pairs). Nevertheless, in a large-scale CRN, the source and destination secondary nodes may be distributed far away and thus relay nodes must be involved for multi-hop communications. Therefore, we aim to achieve the throughput limit that can be transmitted by each node, instead of maximizing a particular Tx-Rx pair throughput that does not take the effects of other nodes’ transmission into account.

Our main contributions in this chapter are as follows. First, we examine the unique characteristics, including common channels constraint, multiple relays and connectivity constraint, in CRNs and study their effect on achievable throughput. Second, we formally derive the theoretically achievable throughput limit of CRNs and propose a class of connectivity-agile sensing algorithms to achieve such a limit. We find that although sensing algorithms may induce sensing overhead, which is usually considered as a negative factor for throughput and emphasized by the existing algorithms, nevertheless, sensing algorithms can also have a positive effect on
finding common channels among neighboring nodes, and thus improving throughput. Finally, as an example of using our results, we design a Connectivity-Agile Sensing Algorithm (CASA) based on cooperative relay. We examine its performance under a class of practical channel models through extensive simulations, which shows that CASA can achieve significant performance improvement.

2.2 Constraints of Sensing Algorithms

In this section, we will first formally define the spectrum sensing, and explain the objective of spectrum sensing for large-scale CRNs. Then we will discuss some challenges, including common channels, interference, relay schemes, and connectivity constraints, on spectrum sensing for a large-scale CRN in details.

2.2.1 Spectrum Sensing

Optimal sensing algorithms in multi-channel CR networks are for secondary nodes to maximize their rewards (e.g., achievable throughput) by searching and opportunistically transmitting over a subset of channels that currently are not used by primary nodes, among a potentially large number of channels. Particularly, the sequence of decisions on what sensing periods to use, which channels to sense, when to stop and which channel to transmit over on a sensing period will be called a sensing strategy. And an optimal sensing algorithm is a sequence of sensing strategies to maximize the benefits of secondary nodes without interrupting communications among primary nodes. Therefore, a spectrum sensing algorithm can be described as

**Definition 2.1.** (Spectrum Sensing) Given a set of channels and some priori knowledge, such as sensing cost and channel utilization, the optimal strategy in a particular sensing period is
defined as the strategy that maximize transmission gain, i.e., achieving the following maximum:

\[ G^* = \max_{\pi \in \Pi} g(\pi) \]

subjected to \( C(\pi) \leq C \),

where the objective function \( g(\cdot) \) denotes the transmission gain, and \( \pi \) denotes the strategy that decides the sensing period as \( \pi_s \) and sense channels in sequence \( \pi(1), \ldots, \pi(\tau - 1) \), and then transmits over channel \( \pi(\tau) \) at stopping time \( \tau \). \( \Pi \) denotes the set of all possible sensing strategies and \( C \) denotes the sensing requirements. Further, a sequence of optimal sensing strategies is called an optimal sensing algorithm.

**Constraints \( C \):** One of the major challenges in CR networks is to strike a balance between protection of primary nodes against interference from secondary nodes and efficient reuse of legacy spectrum. For maximal protection of primary nodes, FCC has set a strict guideline on sensing. For example, in IEEE 802.22, primary nodes should be detected within 2 seconds of their appearance with the probability of misdetection and probability of false detection less than 0.1, which is used as a mandatory requirement \( C \).

**Objective Function:** Given the sensing requirements, a number of objective functions \( g(\pi) \) can be considered for spectrum sensing. A commonly used objective function (see, e.g. [42, 43, 45–49]) is to maximize throughput. In these algorithms, the throughput is represented by the maximum reward of the sensed “available” channels times the transmission time (the sensing period minus sensing time). This design objective is focused on the transmitter of a particular Tx-Rx pair and thus these algorithms are characterized as “self-centric” algorithms. Also, such a definition of throughput may be not achievable because of “common channel constraint” in CRNs. It is important to be reminded that cognitive radio can be much more powerful and flexible than multi-channel multi-radio (MC-CR) technology that has been actively studied in recent years [72–74]. Note that MC-MR remains hardware-based radio technology: each radio can only operate on a single channel at a time and the number of concurrent channels is limited.
by the number of radio interfaces. In contrast, cognitive radio is capable of reconfiguring RF and switching to newly-selected channels. Thus, the number of channels that can be used by a cognitive radio is much larger. In addition, MC-MR based wireless networks typically assume that there is a set of “common channels” available for all nodes in the network. This assumption rarely holds for CRNs since each secondary node may have a different set of “available” channels based on its physical location. Channels available to one secondary node may not be available to its neighbors. Because a successful transmission must use a shared (or common) channel between a Tx-Rx pair, which may not exist or cannot be found during the sensing period, the achievable per-pair throughput using “self-centric” algorithms could potentially approach to zero. Furthermore, instead of the throughput of a particular Tx-Rx pair, maximization of the average per-node throughput is of importance for a large-scale secondary network. From a network’s standpoint, the per-node throughput cannot be simply characterized by the throughput of a single Tx-Rx pair when considering the effect of relay schemes and interference among simultaneous transmissions on the same channels. Therefore, taking common channel constraint into account, we need to study the average per-node throughput from a network’s standpoint, which will be used as the objective of the spectrum sensing in a large CR network. In the following subsections, we will first discuss the network constraints that impact the achievable throughput in details.

2.2.2 Common Channels Constraint

Note that secondary nodes can only opportunistically utilize the channels without interfering with communications among primary nodes. Therefore, it is likely that some channels may not be simultaneously available to the Tx-Rx pair in their vicinity. But only the shared or common channels (channels available for both Tx-Rx pair) can be used for successful transmission. The “self-centric” sensing algorithms which only focus on a single transmitter itself, however, may not find the “common” channels. For example, consider a secondary Tx-Rx pair $v_1$ and $v_2$ and a set of channels $\{ch_i : i = 1, 2, \ldots, N\}$, which are modeled as ON-OFF process [42, 43].
Suppose that channels $ch_1$ and $ch_2$ have the highest transmission rate and largest probability being available (not used by primary nodes) and their availability to $v_1$ and $v_2$ is shown in Fig. 2.1. Since one of $\{ch_1, ch_2\}$ is always available (see Fig. 2.1) and the transmission quality (characterized by transmission rate and availability) of $\{ch_1, ch_2\}$ is better than that of $\{ch_3, ch_4\}$, with self-centric sensing algorithms, the transmitter $v_1$ will only sense and transmit over $\{ch_1, ch_2\}$. However, from Fig. 2.1, it is evident that neither $ch_1$ nor $ch_2$ is available simultaneously for both $v_1$ and $v_2$ and thus the actual throughput for $v_1 - v_2$ would be zero. On the other hand, we find that at any time, there always exist common channels (either $ch_3$ or $ch_4$). Hence, a sensing algorithm that can identify a subset of channels $\{ch_1, ch_2, ch_3, ch_4\}$ will be able to find common channels, thus achieving non-zero throughput. Therefore, the common channel constraint cannot be ignored in the design of sensing algorithms for large-scale CRNs.

In the following subsections, we will show constraints that may affect the throughput from a network’s standpoint.

![Common channel constraint](image)

Figure 2.1: Common channel constraint.

### 2.2.3 Interference Constraint

Here we first show the effect of interference on the throughput. For example, consider two Tx-Rx pairs $v_1 - v_2$ and $v_3 - v_4$ and a set of channels $\{ch_i : i = 1, 2, \ldots, N\}$. Similarly, we assume that channels $ch_1$ and $ch_2$ have the highest transmission rate and largest probability being available for all these four nodes. Assume that the availability of the subset channels $\{ch_1, ch_2, ch_3, ch_4\}$
is the same for all four nodes as shown in Fig. 2.2 (b). As a result of self-centric sensing algorithms, two Tx-Rx pairs will both sense \( \{ch_1, ch_2\} \) and use either \( ch_1 \) or \( ch_2 \) simultaneously (see the upper left of Fig. 2.2 (a)). However, because \( v_2 \) is within the transmission range of \( v_3 \) and receiver \( v_4 \) is within the transmission range of \( v_1 \), there exists transmission collision which results in zero throughput. Meanwhile, if a sensing algorithm senses \( \{ch_1, ch_2, ch_3, ch_4\} \) and schedules different channels not interfering with each other, the network can achieve better throughput (see the lower left of Fig. 2.2 (a)). The optimal sensing algorithm must consider the interference constraint and sense more channels.

![Figure 2.2: Interference and connectivity constraints.](image)

### 2.2.4 Relay Schemes and Connectivity Constraints

In a large-scale CRN, the source and destination CR nodes may be distributed far away and thus the communication between them need to use relay nodes, which can be secondary nodes or PR nodes. There are two reasons that primary nodes may help secondary nodes as relays: (i) primary nodes can obtain some rewards or incentive in relaying packets for secondary nodes; and (ii) primary nodes can make contributions to spectrum efficiency under the condition that their own communications are not affected. We show later in Section 2.3.4 that the connectivity can be guaranteed by increasing either the transmitting power or the number of channels being sensed by each secondary node. For example, in Fig. 2.2 (c), the source node is \( v_s \) and destination node \( v_d \). Then nodes \( v_{r1} \) and \( v_{r2} \) can be used as relays to form a path \( v_s \rightarrow v_{r1} \rightarrow v_{r2} \rightarrow v_d \).
Assume that \(v_s\) and \(v_{r1}\) share common channel \(ch_1\) and \(v_{r2}\) and \(v_d\) share common channel \(ch_2\). However \(v_{r1}\) and \(v_{r2}\) share no common channels. Node \(v_{r3}\) may have shared channels with \(v_s\) and \(v_d\), but it lies outside their transmission range. Therefore, there is no path between \(v_s\) and \(v_d\). To enable communications between \(v_s\) and \(v_d\), we can either increase the transmission power of \(v_s\) or let secondary nodes sense more channels, or even make incentives to primary nodes to relay data. Therefore, the achievable throughput and thus the design objective of sensing algorithms can be also affected by relay schemes and connectivity constraint.

2.2.5 Discussions On Spectrum Sensing

We have shown that the throughput, which is commonly used as the objective function in sensing algorithms, can be affected by common channels, interference, relay schemes, and connectivity. Therefore, the design of spectrum sensing algorithms to achieve maximum throughput for large-scale CRNs confronts at least two challenges: (i) find potential relay schemes, and conditions for a connected network, and (ii) find the achievable throughput under such constraints in Section 2.3 and Section 2.4, respectively. Then we will be able to find guidelines on the design of sensing algorithms to achieve maximum throughput for large-scale CRNs.

2.3 Characteristics of CRNs

We investigate potential relay schemes that have impacts on the throughput in large-scale CRNs and the effect of common channels, positively and negatively. Also, the necessary condition of network connectivity will be presented. Based on these results, we will further find the achievable throughput in next section.

2.3.1 System Models

In this chapter, we study a cognitive radio network consisting of a set \(V = \{v_1, v_2, \ldots, v_n\}\) of \(n\) secondary nodes and a set \(U = \{u_1, u_2, \ldots, u_m\}\) of \(m\) primary nodes in a unit square area.
\( \Omega = [0,1]^2 \) and \( m = O(n) \). We assume that secondary nodes can use a set \( C = \{1,2,\ldots,\|C\|\} \) of channels, each with data rate \( \frac{W}{\|C\|} \), where \( \|C\| \) denotes the cardinality of \( C \) and the channels are numbered in the order of increasing frequency. Hence the total data rate is \( W \). The following assumptions are used throughout the paper.

**Node locations:** The secondary nodes are distributed according to Poisson point process with density \( n \) in \( \Omega \) and each secondary node is equipped with a cognitive radio. Primary nodes are distributed according to a Poisson point process with density \( m \) over \( \Omega \) and each primary node is equipped a hardware-based radio (or multiple hardware-based radio (MR)). Among the \( m \) primary nodes, there exist \( L \) high-power high-bandwidth nodes (PR BSs e.g), which are called PRR nodes for convenience. These nodes form a PRR network and their positions are known to secondary nodes. For example, the PRR nodes are placed regularly at positions \( \left( \frac{1}{\sqrt{L}} \cdot i \frac{1}{\sqrt{L}}, \frac{1}{\sqrt{L}} \cdot j \frac{1}{\sqrt{L}} \right) \) with \( 0 \leq i \leq \sqrt{L} - 1 \) and \( 0 \leq j \leq \sqrt{L} - 1 \). Clearly, these \( L \) regularly distributed PRR nodes divide the region \( \Omega \) into \( L \) subregions as Voronoi diagrams with side length \( \frac{1}{\sqrt{L}} \) (see Fig. 2.3). Here we generally assume that \( L \) is finite and a square of some integer for simplicity, and use \( S_j \) to denote the subregion centered at the primary node \( u_j \).

**Channel sets:** Different primary nodes may have license to different subset of channels and the primary nodes sharing the common subset of channels form a primary network. And we focus on the case where \( \|C\| = \Theta(logm) = O(logn) \) \(^1\). This is because in large scale deployments, the number of channels will typically be much smaller than the number of nodes. Moreover, as shown in [1], when \( \|C\| = O(logn) \), the capacity is independent of the number of interfaces. That is, we only need one interface to use the multiple channels, thus greatly simplifying our design and analysis.

**Interference model:** A protocol interference model [71] is used in our analysis because this model highlights the effects of transmission and interference range. Further, it has been shown that under some conditions, the more realistic Physical Model can be reduced to the Protocol

\(^1\)We use the following notation throughout: \( f(n) = O(g(n)) \) if \( \exists n_0 > 0 \) and constant \( c_0 \) s.t. \( f(n) \leq c_0 g(n) \) \( \forall n \geq n_0 \). \( f(n) = o(g(n)) \) if for every \( c > 0 \), \( \exists n_0 > 0 \) s.t. \( f(n) < c g(n) \) \( \forall n \geq n_0 \). We say \( f(n) = \Omega(g(n)) \) if \( g(n) = O(f(n)) \) and \( f(n) = \omega(g(n)) \) if \( g(n) = o(f(n)) \). And \( f(n) = \Theta(g(n)) \) if \( f(n) = O(g(n)) \) and \( f(n) = \Omega(g(n)) \).
Model [71]. Let \( r \) and \( R \) denote the transmission range of secondary radio and hardware-based radio, respectively. A secondary node \( v_i \) can successfully transmit to another secondary node \( v_j \) over channel \( k \) if \( \|v_i - v_j\| \geq (1 + \Delta_1)r \) for each secondary node \( v_i \) transmitting over \( k \) for some constant \( \Delta_1 \). Similarly, hardware-based radio \( u_i \) can successfully transmit to hardware-based radio \( u_j \) over channel \( k \) if \( \|u_i - u_j\| \geq (1 + \Delta_2)R \) for each hardware-based radio \( u_i \) transmitting over \( k \) for some constant \( \Delta_2 \).

### 2.3.2 Relay Schemes In CRNs

Most of the existing works on CRNs fall into one of the following two directions. The first assumes that there is no inter-operation between primary nodes and secondary nodes. Primary nodes either are oblivious of the presence of secondary nodes or have higher priority on accessing the channels and have the right to improve their own revenue by charging secondary nodes for using their licensed channels. The second considers some cooperation between primary nodes and secondary nodes. Primary nodes lease their spectrum to secondary nodes for a fraction of time and in exchange, they get the cooperative transmission power from secondary nodes. That is, primary nodes actively select some appropriate secondary nodes and involves them into the primary transmission as cooperative relays. In return, chances to access the wireless media are given to the selected secondary relays.

One important observation is that, in the existing works, secondary nodes do not involve primary nodes into their transmission as relays, maybe because a prerequisite in CRNs is that the secondary nodes can not interfere with the performance of PR nodes. On the other hand, reference [2] proposes an architecture where secondary nodes communicate with PRR nodes (primary base stations e.g.) directly on licensed channels. For these PRR nodes, the transmission power and transmission rate is high but the required traffic load demand is satisfied. Thus these PRR nodes have no interest to increase their transmission rate. Instead, they want to achieve certain benefit in other format, for example revenue, to fully utilize their resources. Based on this new observation, we propose a novel cooperative CR framework, where PRR nodes
are dominant in channel usage, while coordinating relays among secondary nodes. In return, secondary nodes make an extra payment, which is proportional to the cooperation time, to PRR nodes.

![Node positions and relay schemes. The perimeter is an extended version of the grids.](image)

Figure 2.3: Node positions and relay schemes. The perimeter is an extended version of the grids.

In particular, we consider two relay schemes: CR-Only Relay (cooperative) scheme and PRR nodes Assisted Relay (hybrid) scheme. In cooperative relay, we assume no cooperation between the secondary nodes and primary nodes. That is, secondary nodes communicate with each other in multi-hop only using other secondary nodes as relay nodes, e.g., \( v_1 \) to \( v_2 \) communication in Fig 2.3. In fact, this is similar to MC-MR networks, except these nodes do not share the common set of channels. In PRR nodes Assisted Relay (hybrid) scheme, PRR nodes participate in secondary transmission as relays. Instead of assuming all primary nodes can participate in relays, we focus on the scenario that only the PRR nodes coordinate the relays among secondary nodes for the following reasons. Firstly, it is not realistic to let all primary nodes participate in relays among secondary nodes since secondary nodes cannot interfere with primary nodes. Secondly, even all PR nodes can participate in relays, Gupta and Kumar [71] have shown that the throughput increase is bounded by \( \sqrt{\frac{mn}{n+1}} \)-fold. Thus the main benefit of hybrid relay is not throughput improvement but reliability and delay. Particularly, we assume that there exists
a PRR network and the secondary nodes can use PRR nodes as relays. And when the source secondary node and its destination secondary node fall into the same subregion, the source secondary node attempts to reach its receiver in multi-hop by using other secondary nodes in the same subregion as relays. Otherwise, the source node will reach the closest PPR node first, and then the latter relay the data to the PRR node which is closest to the receiver (may use some other PRR nodes as relay nodes). An example is shown as the communication between $v_3$ and $v_4$ in Fig 2.3. At last, the PRR node closest to the destination node relays the data to the destination by one- or multi-hop.

### 2.3.3 Shared Common Channel Probability of Secondary Nodes

As discussed earlier, there may not be a common set of channels among an arbitrary pair of secondary nodes due to primary nodes in the vicinity and imperfect sensing. While our observation does reflect the fact in CRNs, it induces another problem, that is, even for two secondary nodes within transmission range $r$, they can communicate with each other only when they share common channels. We show later that this requires larger transmission range of secondary nodes to make the secondary network connected. Here we derive the probability, denoted by $P_s$, that any pair of secondary nodes shares at least one common channel so that the communication between them is possible. The derived probability will be used in subsequent sections for deriving the throughput limits in CRNs.

We first show that at any point $z \in \Omega$, the usage efficiency of any channel is bounded by some constant $1 - p < 1$. Given $z \in \Omega$, the interfering area $A_z$ is a circle centered at $z$ with radius $(1 + \Delta_2)R$. That is, the channels used by the primary nodes in the surrounding area cannot be used by the secondary node at point $z$. Let $y_j$ be the indicator function such that $y_j = 1$ if the primary node $u_j$ is in the interfering area and $y_j = 0$ otherwise. Since the primary nodes are randomly and independently distributed, $P(y_j = 1) = \pi R^2 (1 + \Delta_2)^2$ for all $j$. Let $y$ denote the number of PR nodes in the interfering area. Thus $y = \sum_{j=1}^{m} y_j$ and $E(y) = m \cdot m \pi R^2 (1 + \Delta_2)^2$. Recall from [56], usually $R = \Theta(\sqrt{\log m})$ and thus $E(y) = \Theta(m \log m)$.
theorem, we can show that $|y - E(y)| \to 0$ as $m \to \infty$. Thus $y = \Theta(log m)$ with high probability. We assume that the number of PR nodes using any channel are of the same order because we do not consider the transmission quality difference among channels in our theoretical analysis. Thus, a randomly chosen primary node has the same probability to use any channel. That is, equivalently, for any channel $k \in C$, it is used by a primary node in $A_z$ with probability $\Theta(\frac{1}{\|C\|}) = \Theta(\frac{1}{log m})$. Thus $k$ is available at $z$ with probability $(1 - \Theta(\frac{1}{log m}))^y \approx e^{-\frac{y}{\Theta(\frac{1}{log m})}} = e^{-\epsilon}$ for some constant $\epsilon > 0$ when $m$ is sufficiently large. Let $p = e^{-\epsilon}$ and note that $p < 1$. This result corresponds to the FCC measurement which shows that 70% of the allocated spectrum in US is not used temporally and geographically.

For any secondary node $v_i$, let $A_i$ denote the subset of channels being sensed and $\|A_i\|$ be the cardinality of $A_i$. For simplicity, assume $\|A_i\| = \|A\|$ for all $i$. For any two secondary nodes $v_i$ and $v_j$, let $K_{i,j}$ denote the set of available channels’ shared by $v_i$ and $v_j$. That is $K_{i,j} = \{k: k \in A_i \cap A_j, \text{ and } k \text{ is available at } v_i \text{ and } v_j \}$, and $P_s(k)$ denote the probability that channel $k$ is shared by $v_i$ and $v_j$. Then we have

$$P_s(k) = P(k \in K_{i,j}) = \frac{\|A\|}{\|C\|}^2 \times p^2$$

$$P_s = P(K_{i,j} \neq \emptyset) = 1 - P(K_{i,j} = \emptyset) \approx 1 - \left[1 - \left(\frac{\|A\|}{\|C\|}\right)^2 \times p^2\right]^{\|C\|} \approx 1 - e^{-\frac{(\|A\|^2p^2)}{\|C\|}} \approx \frac{\|A\|^2p^2}{\|C\|}.$$ 

Note that in the above derivation, we assume that $k \in K_{i,j}$ is independent of other channels. This approximation is reasonable when $\|C\|$ is sufficiently large.
2.3.4 Connectivity Constraints in CRNs

The connectivity of CRNs is a prerequisite for communications since the throughput $\lambda(n) = 0$ if a CRN is disconnected.

Necessary Condition for Connected CRNs with Cooperative Scheme

We first derive the necessary condition under Cooperative scheme. Let $P_{\text{disc}}$ denote the probability that a CRN is disconnected and we have:

**Lemma 2.1.** Given a CRN with $n$ nodes uniformly deployed in a unit area $\Omega$, each with transmission range $r$, if $P_{\text{disc}} \approx \frac{\log n + \kappa(n)}{n}$ and the network is connected asymptotically almost surely, then $\lim_{n \to \infty} \kappa(n) = \infty$.

**Proof.** See Theorem 5 in [84] for details.

Lemma 2.1 gives a necessary condition of a multi-channel CR network to be connected under Cooperative scheme, which implies $r > \sqrt{\frac{\log n}{P_{\text{trans}}}}$. Furthermore, we will be showing later in Section 2.4.2, it is evident that $r = \sqrt{\frac{800 \log n}{P_{\text{trans}}}}$ suffices the network connectivity by a carefully designed routing scheme. Thus $r = \Theta\left(\sqrt{\frac{800 \log n}{P_{\text{trans}}}}\right)$ is a sufficient and necessary condition for connectivity.

Necessary Condition for Connected CRNs with hybrid Scheme

When a hybrid scheme is used, the entire network is connected if and only all secondary nodes and the primary node falling in each subregion $S_i$ form a connected subnetwork. In such a case, the transmission range of secondary nodes can be smaller than a fully connected network of secondary nodes. Let $N_i$ denote the number of nodes falling in subregion $S_i$, we have:

**Lemma 2.2.** For $1 \leq i \leq L$, $N_i = \frac{\rho n}{L}$ for some constant $\rho$ with high probability.

**Proof.** See Lemma 1 in [82] for details.
Since there are $N_i$ secondary nodes randomly and uniformly distributing on the subregion $S_i$ with area $\frac{1}{L_i}$, each with transmission range $r$, by Theorem 2.1, we have

**Lemma 2.3.** If a CRN with hybrid scheme is connected with high probability and $LP_s \pi r^2 = \frac{\log N_i + \kappa(N_i)}{N_i}$, then $\lim_{N_i \to \infty} \kappa(N_i) = \infty$.

Lemma 2.3 gives a necessary condition for a connected multi-channel CRN, which implies the transmission range $r$ must be at least $\Omega\left(\sqrt{\log(\frac{2\pi}{r^2})}\right)$ for some constant $\rho$.

So far, we have investigated some unique characteristics of CR networks, such as possible relay schemes and shared common channel problem. We also find that the connectivity of CRNs is dependent on relay schemes. Specifically, in Cooperative scheme, a necessary condition for connectivity is $r = \Theta\left(\sqrt{\frac{800 \log n}{P_n}}\right)$ and in hybrid scheme, a necessary condition for connectivity is $r = \Omega\left(\sqrt{\frac{\log(\frac{2\pi}{r^2})}{P_n}}\right)$, where $P_s \approx \frac{||A||^2 P^2}{||C||}$. The theoretical analysis is consistent with our intuition in in Section 2.2, that is, increasing the number of channels being sensed ($||A||$) can reduce the transmission range (or transmission power) required for connectivity. Next, we derive the achievable throughput in CRNs based on our analytical results. This indeed indicates that there is limited room for improving the path stability in high mobility environment.

### 2.4 Throughput in Large-scale CRNs

There have been many works on throughput of single-channel and multi-channel (MC) networks. In [71], Gupta and Kumar show that the achievable throughput in a random network with a common single channel is $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$, where $n$ is the number of nodes. They further show that using multiple channels cannot enhance the network throughput if there is one radio per channel at each node. Kyasanur and Vaidya studied the throughput of general multi-channel networks wherein the number of radios may be smaller than the number of channels in [72]. They show that the achievable throughput is still $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ when the ratio of the number of channels and the number of radio is $O(\log n)$. In [84], Bhandari and Vaidya study the throughput of networks in which nodes share heterogeneous channels and show that the achievable throughput is of the
same order as throughput of MC-MR networks, $\Theta\left(\frac{W}{\sqrt{n \log n}}\right)$ when the number of channels is $O(\log n)$ for a single radio.

These results may be not true for CRNs because of the unique characteristics of CRNs. For example, secondary nodes can only access the channels opportunistically without interfering with the primary nodes and the relay schemes in CR networks can be different from the MC-MR networks [71, 72, 84]. Therefore, the throughput derivation in CRNs must take spectrum sensing and relay schemes into account.

Here, we formally derive the achievable throughput of multi-channel CRNs with Cooperative scheme and hybrid scheme, respectively. We first analyze the throughput upper bounds with Cooperative and hybrid schemes. Then we derive the lower bounds of throughput, which is in the same order of the upper bounds, and further present joint routing and sensing schemes to achieve the throughput bounds.

### 2.4.1 Definitions

Throughout the paper, we consider throughput $\lambda(n)$ as the time average of the number of bits per second that can be transmitted by each secondary node to its destination CR node, similar to [71],

**Definition 2.2.** (Feasible throughput). The throughput of $\lambda(n)$ bits per second for each secondary node is feasible if there exists a spatial and temporal scheme for scheduling transmissions, such that by operating the network in a multi-hop fashion and buffering at intermediate nodes (secondary or primary) when awaiting transmission, every secondary node can send $\lambda(n)$ bits per second on average to its chosen secondary destination node.

**Definition 2.3.** (Throughput in CRNs) We say that the throughput of a class of Random Networks is of order $\Theta(f(n))$ bits per second if there are deterministic constants $c_1 > 0$ and $c_2 < \infty$ such that

$$\lim_{n \to \infty} P(\lambda(n) = c_1 f(n) \text{ is feasible}) = 1,$$
\[
\lim_{n \to \infty} P(\lambda(n) = c_2 f(n) \text{ is feasible}) < 1.
\]

### 2.4.2 Throughput with Cooperative Relays

#### Throughput Upper Bound with Cooperative Relay

Now we focus on the case when \(\|C\| = O(\log n)\). As mentioned in [72], when \(\|C\| = O(\log n)\), the throughput upper bound is dominated by the connectivity constraint. Specifically, given the number of hops taken by each packet \(H(b)\) for packet \(b\), and the number of simultaneous transmissions possible on each channel \(S\), if each source node generates data at rate \(\lambda(n)\), we have \(\lambda(n) n H(b) \leq SW\). In a random network model, \(H(b) = \Theta(\frac{1}{n})\) and \(S = O(\frac{1}{n})\). Thus we have \(\lambda(n) = O(\frac{W}{n^2})\). From previous part we have \(r = \Omega\left(\sqrt{\frac{\log n}{nP_s}}\right)\), by substituting \(P_s\) in (2.1), we have

\[
r = \Omega\left(\sqrt{\frac{\|C\| \log n}{\|A\|^2 n}}\right).
\]

Hence we obtain the following upper bound of the achievable throughput.

**Theorem 2.1.** For a CRN with \(n\) secondary nodes randomly and uniformly deploying into an unit area, there exists a deterministic constant \(c < \infty\), such that

\[
\lim_{n \to \infty} P\left(\lambda(n) = cW \sqrt{\frac{\|A\|^2}{\|C\| n \log n}} \text{ is feasible}\right) = 0.
\]

#### Throughput Lower Bound with Cooperative Relay

Here, we show that the throughput upper bound derived earlier is sharp by exhibiting a scenario where we can achieve a throughput with the same order of the upper bound throughput. Specifically, we provide a constructive joint sensing and routing scheme, namely Approximate Straight-Line Routing (ALSR) scheme, to demonstrate that one can spatially and temporally schedule transmissions so that when each randomly located secondary node has a randomly chosen destination, each source-destination pair can achieve throughput \(\Omega\left(W \sqrt{\frac{\|A\|^2}{\|C\| n \log n}}\right)\). Therefore, the throughput is lower bounded by \(\Omega\left(W \sqrt{\frac{\|A\|^2}{\|C\| n \log n}}\right)\). Since the lower bound has the same
order with the upper bound, the throughput is \( \Theta \left( W \sqrt{\frac{|A|^2}{|C|n \log n}} \right) \). This construction scheme shares similar ideas in \([65, 71, 72]\). The details of ALSR scheme are as follows:

(a). At the beginning of each time slot \( T \), each CR node \( v_i \) independently and randomly choose \( |A| \) channels to sense. We assume an ideal sensor and thus ignore sensing overhead here.

(b). Divide the unit area \( \Omega \) into squarelets, each of area \( \alpha(n) \) (see Fig. 2.4). The transmission range \( r \) is set to \( \sqrt{8\alpha(n)} \), thereby ensuring that two nodes from two squarelets (sharing a common side) be able to communicate with each other directly.

(c). A time division multiple access (TDMA) transmission scheme is used, in which each squarelet becomes active, i.e., its nodes can transmit successfully to nodes in the squarelet or in neighboring squarelets, at regularly scheduled time-slots.

(d). Let \( L_i, i = 1, 2, \ldots, n \) denote the straight line connecting a source node \( v_i \) to its destination \( v_{id} \). \( v_i \) transmits data to \( v_{id} \) in a sequence of hops. In each hop, the packet is transferred from one squarelet to another in the order in which they intersect \( L_i \) as shown in Fig. 2.4. When a squarelet becomes active, it transmits packets for all \( L_i \) passing through it.

(e). Since each secondary node independently scans a subset of channels \( A \), the source node \( v_i \) and its destination \( v_{id} \) may share no common channels. Therefore, channel transition from the source to the destination may be required. In our scheme, any relay node \( v_{ir} \) carrying the data (including \( v_i \)) can use any one of its channels to relay the data to the next squarelet (in the order of intersecting \( L_i \)). In other words, \( v_{ir} \) can choose any node sharing channels with it in the next squarelet as a relaying node until the data is transferred to the last relaying squarelet, which is just before the destination squarelet. The relay node \( v_{il} \) in the last relay squarelet may share no common channel with \( v_{id} \). If \( v_{ir} \) and \( v_{id} \) share no channels, \( v_{id} \) will relay the data to the channel transition node \( v_{it} \), i.e., the node shares a channel with both \( v_{il} \) and \( v_{id} \) in the destination squarelet. Then \( v_{it} \) relays the data to \( v_{id} \); otherwise, \( v_{il} \) relays the data to \( v_{id} \) directly.

In order to show that the above scheme is applicable, we need to prove i) each squarelet contains at least one node; ii) for each node in any squarelet, there exist nodes sharing at least
one common channel with it in any one of the neighboring squarelets; iii) there exist channel transition nodes in any destination squarelet; and vi) every squarelet can be scheduled to be active during bounded time period. We need the following lemmas to prove that our scheme satisfies these requirements.

**Lemma 2.4.** If \( \alpha(n) = \frac{100 \log n}{nP_s} \), where \( P_s \) is the probability that two secondary nodes share at least one channel which is obtained in (2.1), then the number of nodes \( N_D \) in any squarelet \( D \) is \( \frac{50 \log n}{P_s} \leq N_D \leq \frac{150 \log n}{P_s} \) with high probability.

**Proof.** see [71, 84] for details.

Lemma 2.4 shows that our scheme satisfies the first requirement and any squarelet contains at least \( N_D = \frac{50 \log n}{P_s} \) nodes with high probability. Given \( N_D \) nodes in each squarelet \( S_i \), the following lemma show that our scheme satisfies the second requirement, which require that for any node \( v \) in \( S_i \), there must exist some nodes sharing at least one common channel with \( v \) in the next squarelet in the order of intersecting with the straight line \( L_i \).

**Lemma 2.5.** If \( \alpha(n) = \frac{100 \log n}{nP_s} \), then given any relay node \( v \) carrying the data, there exist at least \( 25 \log n \) nodes sharing the channels with \( v \) in the next squarelet with high probability.

**Proof.** See Lemma 8 in [84] for details.
Lemma 2.6 ensures that our scheme satisfies the third requirement, that is, every destination squarelet contains channel transition nodes.

**Lemma 2.6.** Suppose the last relay node $v_{il}$ and the destination node $v_{id}$ can use available channels $k, k' \in C$ respectively. Then, with high probability there exists at least one channel transition node $v_{it}$, which can use both $k$ and $k'$, in the destination squarelet with high probability.

**Proof.** By Lemma 2.4, with high probability, there exist at least $N_D - 1 \geq \frac{50 \log n}{P_s} - 1$ nodes (except the destination node) in the destination squarelet. For any of the $\frac{50 \log n}{P_s} - 1$ node $w$, the probability that $w$ is channel transition node is $p_1 = \frac{\|A\|^2 P_s^2}{\|C\|^2}$. The probability that there exists no channel transition node is $P \leq \left(1 - p_1\right)^{\frac{50 \log n}{P_s} - 1} < e^{-\left(\frac{50 \log n}{P_s} - 1\right) p_1}$, which approaches zero as $n \to \infty$. Thus the result follows. \hfill \Box

The following Lemma 2.7 shows that every squarelet can be scheduled to be active at every bounded length of time, which ensures our scheme satisfies the fourth requirement.

**Lemma 2.7.** For some constant $c_1$, there exists a schedule for transmitting packets such that in every $(1 + c_1)$ slots, each cell can use one slot to transmit on each of the available channels.

**Proof.** See [65, 71] for details. \hfill \Box

Lemma 2.7 shows that each squarelet can be active for a guaranteed fraction of time on each of the available channel. Therefore, each squarelet can have a throughput $\frac{1}{1+c_1} WP$, where $p$ is the usage efficiency of each channel. Lemma 2.4, Lemma 2.5, Lemma 2.6 and Lemma 2.7 prove that our scheme is valid. The maximum number of $L_i$ passing through any squarelet is $O(n \sqrt{\alpha(n)})$ with high probability (see Lemma 3 in [65]).

If a squarelet contains the destination node of $L_i$, it may need to forward the packet from the transition node to the destination node. Otherwise, the traffic is relayed to the next squarelet. Hence the traffic handled by a squarelet is proportional to the number of lines passing through it. Since each line $L_i$ carries traffic of rate $\lambda(n)$, we obtained the following bound:
Theorem 2.2. With $\alpha(n) = \frac{100\log n}{nP_s}$, the achievable throughput for each source-destination flow is

$$\lambda(n) = \Omega\left(\frac{W}{n\sqrt{\alpha(n)}}\right) = \Omega\left(W\sqrt{\frac{\|A\|^2}{\|C\|n\log n}}\right).$$

(2.3)

Theorem 2.1 and Theorem 2.2 show that the upper throughput and lower throughput in multi-channel CRNs with Cooperative scheme are of the same order of $\Theta\left(W\sqrt{\frac{\|A\|^2}{\|C\|n\log n}}\right)$, which implies throughput $\Theta\left(W\sqrt{\frac{\|A\|^2}{\|C\|n\log n}}\right)$ is achievable in multi-channel CRNs with Cooperative scheme.

2.4.3 Throughput with Hybrid Relay

Here we study the achievable throughput of multi-channel CRNs with hybrid relay. Similarly, we first study upper bound of throughput limits. After that, we derive the lower bound and present a routing scheme to achieve such throughput limit.

Throughput Upper Bound with Hybrid Relay

Similar to the analysis for Cooperative relay, we first derive the upper bound due to connectivity constraint. Given the number of hops taken by each packet $H(b)$ for packet $b$ and the number of simultaneous transmissions on each channel $S$, the total number of transmissions required in time interval $T$ is at least $\lambda(n)TnH(b)$. Then we have $\lambda(n)TnH(b) \leq TSW$. Since $S = O\left(\frac{1}{r^2}\right)$, we further have $\lambda(n) = O\left(\frac{W}{r^2nH(b)}\right)$. Note that in hybrid, since PRR nodes participate in the relay, $H(b)$ will be much less than that in Cooperative. Next we derive the lower bound of $H(b)$ to achieve the upper bound of throughput limit. Let $Q$ denote the length transversed by bit $b$ among secondary nodes and $H(b) \geq \frac{Q}{r}$. By using Lemma 9 from [80] and Lemma 4 from [82], we derive $Q = \Theta\left(\frac{\sqrt{L}}{r}\right)$ with high probability, where $L$ is number of PRR nodes as defined in Section 2.4.1, for either the source and destination are in the same subregion, or the source and destination secondary nodes are in different subregions (see [80, 82] for details). Thus, $H(b) = \Theta\left(\frac{\sqrt{L}}{r}\right)$ and $\lambda(n) = O\left(\frac{\sqrt{LW}}{rn}\right)$. Hence we have the following result:
Theorem 2.3. For a multi-channel CRN with hybrid scheme, there is a deterministic constant $c < \infty$, such that

$$\lim_{n \to \infty} P\left( \lambda(n) = \frac{cW \sqrt{L \rho P_s}}{\sqrt{n^{\rho \frac{n_L}{L}}}} \text{ is feasible} \right) = 0.$$ (2.4)

Unfortunately, we show later that this upper bound is not achievable. After analysis, we find that in hybrid, secondary nodes in some subregion $S_i$ send data to the primary node $u_i$ when their destination nodes falling outside of $S_i$. Thus, if most of source nodes have destination outside the subregion, the primary nodes have much burden to relay data, thus become bottlenecks.

Lemma 2.8. In any subregion $S_i$, with high probability, there exist at least $\Theta\left( \frac{W}{\Gamma} (1 - \frac{1}{\Gamma}) \right)$ nodes that will send data to primary node $u_i$.

Proof. See [82] for more details.

Note that the data rate for any primary node to receive from adjacent secondary nodes is $\frac{W}{\|C\|}$. Therefore, by Lemma 2.8, we have the following tighter upper bound:

Theorem 2.4. For the multi-channel CRNs using hybrid scheme, the throughput $\lambda(n)$ is at most $\lambda(n) \leq \Theta\left( \frac{W}{\|C\| (1 - \frac{1}{\Gamma})} \right)$ with high probability.

Throughput Lower Bound with Hybrid Relay

Now, we derive the throughput lower bound of CRNs with hybrid relay. Specifically, we provide a constructive routing scheme and prove that the throughput achieved by our scheme matches the upper bounds asymptotically in the order of $\Theta\left( \frac{W}{\|C\|} \right)$. Our scheme is based on a good approximation of a minimum connected dominating set ($MCDS$, see [85] for details).

Before we construct an approximation of $MCDS$, we will first show the following lemma:

Lemma 2.9. In CRNs, given the $MCDS$, there exists a schedule solution by which all $MCDS$ nodes are able to transmit in every $\Gamma$ time slots for some constant $\Gamma$, which guarantees that each $MCDS$ node can transmit at rate of $\Theta\left( \frac{W}{\|C\|} \right)$.
Proof. From [80,85], we know that for a random network with a single channel, we can construct an approximation of \( MCDS \) and each \( MCDS \) node has at most a constant number of neighbors on the \( MCDS \); and thus an \( MCDS \) node can transmit at least once in every \( \Gamma \) time slots for some constant \( \Gamma \). In our case, for each subset \( V_k \subset V \) of all secondary nodes which can use the available channel \( k \in C \), we construct a connected domination set \( MCDS_k \) of \( V_k \) using similar scheme to [80, 85]. Each node \( v \in MCDS_k \) can only transmit over channel \( k \) even if it can also use other available channels. The union \( \bigcup_{k \in C} MCDS_k \) plus at most \( ||C|| \) channel transition nodes (which can communicate with nodes from two different channels, see Lemma 2.6), is a \( MCDS \) for our CRN. For the \( MCDS \), each node can transmit once in every \( \Gamma \) time slots with some channel \( k \in C \). Consequently, the per-node data transmission rate achieved by each node on the \( MCDS \) is at least \( \frac{W}{||C||} \). Thus the result follows. 

The details of Approximate MCDS-based Routing Scheme are as follows:

1. At the beginning of each time slot \( T \), each CR node \( v_i \) independently and randomly choose \( ||A|| \) channels to sense. We assume an ideal sensor and thus ignore sensing overhead here.

2. Partition the unit area into squarelets, each of area \( \alpha(n) \). The transmission range \( r \) is set to \( \sqrt{\alpha(n)} \), thereby ensuring that any two nodes from two squarelets (sharing a common side) can communicate with each other directly.

3. Let \( V_k \subset V \) denote the subset of all secondary nodes which can use channel \( k \in C \). At each squarelet \( S_j \), we choose a node \( v^j_k \in V_k \) for all \( k \in C \). Assume that \( v^j_k \) transmit only with channel \( k \). Thus \( \bigcup_{k \in C} \bigcup_{j} v^j_k \) plus at most \( ||C|| \) “connectors” (the channel transition node, whose existence is shown by Lemma 2.6), automatically form a CDS.

4. For each source node \( v_i \) in subregion \( S_j \) and its destination node \( v_{id} \) (if the destination of \( v_i \) falls outside \( S_j \), \( v_{id} \) denotes the primary node \( u_j \) in \( S_j \)), we find an MCDS node in each of these squarelets that are crossed by line \( v_i v_{id} \). We connect these MCDS nodes in sequence to form a path, denoted as \( P(v_i, v_{id}) \), and data are transmitted from \( v_i \) to \( v_{id} \).
through $P(v_i, v_{id})$ (see part (a) of Fig. 2.5). When $v_i$ and $v_{id}$ use different channels, $v_i$ first transmit data to proxy destination MCDS node $v_{id}''$, which use the same channel with $v_i$ and exist in the squarelet containing $v_{id}$, through path $P(v_i, v_{id}''$) constructed above. Then $v_{id}''$ relays data to $v_{id}$ at most two hops. (see part (b) of Fig. 2.5). When $v_i$ (or $v_{id}$ or both) is not in the CDS, we first connect $v_i$ (or $v_{id}$ or both) to its dominator (say $v_i'$ and $v_{id}'$) in the same squarelet (see part (c) of Fig. 2.5).

![Figure 2.5: The illustration of MCDS-based routing scheme.](image-url)

We show the validity of the Approximate MCDS-based routing scheme and the throughput achieved by this scheme in the following Lemmas. Let $\alpha(n) = \frac{100\log n}{nP_s}$ and we have:

**Lemma 2.10.** In Approximate MCDS-based Routing Scheme, there exist at least $\frac{25\log n}{P_s\|C\|}$ secondary nodes in any squarelet, which can use the available channel $k$ for each $k \in C$ with high probability.

**Proof.** See [82, 84] for details.

Lemma 2.10 ensures the correctness of Approximate MCDS-based Routing Scheme, which requires the existence of node being able to use channel $k$ for any $k \in C$ at each squarelet. We can later show that the total data rate of all routing requests on each MCDS node is at most $\Theta\left(\frac{W}{\|C\|}\right)$ almost surely based on our scheme, which guarantees that each secondary node can transmit once every $\Gamma$ time-slots by Lemma 2.9. This implies the total data rate achieved by all nodes is at least $n \cdot \frac{W}{\Gamma\|C\|}$. Therefore, if we can get the number $N_c$ of nodes that receive a
copy of the data packet during its transmission from the source node to its destination, we can derive the lower bound throughput \( \lambda(n) = \Omega\left(\frac{n^{\frac{W}{\|C\|}}}{N_c}\right) \).

Based on Lemma 4 and Lemma 5 in [82], we know that in hybrid scheme, the length between the source node \( v_i \) and \( v_{id} \), denoted by \( Q \) is \( \Theta(\frac{1}{\sqrt{L}}) \) with high probability, where \( L \) is number of PRR nodes as defined in Section 2.4.1. The number of squarelets crossed by \( uv \) is at most \( \frac{\sqrt{2Q}}{\sqrt{\alpha(n)}} \). Thus the length \( L_p \) of the resulting path \( P(v_i, v_{id}) \) satisfies

\[
L_p \leq \frac{\sqrt{2Q}}{\sqrt{\alpha(n)}} \times \sqrt{2\alpha(n)} = \Theta\left(\frac{1}{\sqrt{L}}\right).
\]

(2.5)

Let \( A(P) \) be the region covered by all transmitting disks of all transmitting nodes in the path \( P(v_i, v_{id}) \). Clearly, the area of \( A(P) \), denoted by \( |A(P)| \), is at most \( |A(P)| \leq 2rL_p + \pi r^2 < \psi \) for some constant \( \psi \) when \( n \) is large enough.

**Theorem 2.5.** (the upper bound of the data copies) With the Approximate MCDS-based Routing Scheme, the number of nodes \( N_c \) that get a data copy is, with high probability, at most \( N_c \leq \frac{3}{\psi} \psi n \), here \( \psi \) is a constant.

**Proof.** See Lemma 12 and Lemma 13 of [82] for details.

Based on Theorem 2.5, we have the following theorem.

**Theorem 2.6.** (the lower bound of the throughput) The throughput for multi-channel CRNs with hybrid scheme is at least \( \lambda(n) = \Omega\left(\frac{W}{\|C\|}\right) \).

Note that the lower bound matches the upper bound in Section 2.4.3 in the order of \( \Theta\left(\frac{W}{\|C\|}\right) \), which implies \( \lambda(n) = \Theta\left(\frac{W}{\|C\|}\right) \) is achievable in multi-channel CRNs with hybrid scheme. Also note that the correctness of Theorem 2.6 requires that the total traffic of all routing requests on each MCDS node is no more than \( \Theta\left(\frac{W}{\|C\|}\right) \). Since the achieved throughput is \( \lambda(n) = \Omega\left(\frac{W}{n}\right) \), the traffic load on any routing node (MCDS node) \( n\lambda(n) = \Omega(W) \) satisfies the requirement that the total traffic is no more than \( \Theta\left(\frac{W}{\|C\|}\right) \).
2.4.4 Discussions On Throughput in CRNs

We have considered two relay schemes, cooperative and hybrid relay in CR networks. Under cooperative relay, secondary nodes communicate with each other only using other secondary nodes as relays. This is similar to the relay in MC-MR networks, except that secondary nodes share no common channel sets. We find that the throughput under cooperative relay in CRNs is of the same order with the throughput in MC-MR networks, even though unlike MC-MR, secondary nodes can only use channels when primary nodes do not use them. Therefore, Cooperative relay can greatly improve the utilization efficiency of channels available in CRNs. Under hybrid relay in which PRR nodes assist in relaying packets by making the best use of their higher power and bandwidth in comparison to CR nodes, data transmission in hybrid relay can be much faster and more reliable. However, if there exist many secondary nodes whose destinations fall into different subregions and these secondary nodes may send routing requests to the primary nodes nearby, the primary nodes can potentially become the routing hot spots and hence network bottlenecks. Therefore, the throughput under hybrid relay in Section 2.4.3 is much lower than that under Cooperative scheme in Section 2.4.2. For application purposes, we can categorize data packets into two groups. For the real time data which need to be delivered to the destinations within a short time, hybrid relay is a good option; otherwise cooperative relay is preferred for data packets without delay constraint, such as delay tolerant networks.

2.5 Design of Sensing Algorithms To Achieve Maximum Throughput

In the previous sections, we showed that the achievable throughput in CRNs jointly depends on spectrum sensing, relay, and scheduling. Given optimal scheduling, we have derived the theoretically maximum achievable throughput, which can be used as the objective for spectrum sensing in large-scale CRNs. In this section, we first summarize general properties of sensing algorithms to achieve the maximum throughput in large-scale CR networks. Then given cooperative relay
as an example, we propose a *connectivity-agile sensing algorithm* (CASA) which aims to achieve optimal performance subject to connectivity constraint. Finally, we provide numerical results to evaluate the performance of the proposed sensing algorithm under a class of channel models.

2.5.1 Properties of Sensing Algorithms to Achieve Maximum Throughput

**Property 2.1.** Sensing algorithms need to be designed to ensure network connectivity in order to achieve the maximum throughput.

Connectivity is a precondition for nonzero throughput in large-scale networks. We have shown in Section 2.3.4 that connectivity jointly depends on sensing and transmission power. Particularly, given transmission power of secondary nodes, increasing the number of channels sensed can ensure the asymptotic connectivity. However, this observation is usually ignored and because *side effects* of sensing on throughput because of sensing overhead is overemphasized. Therefore, the advantage of sensing on throughput by increasing the probability of finding common channels among neighboring secondary nodes is neglected, which potentially underestimates the throughput and the benefits of sensing algorithms.

**Property 2.2.** Given optimal scheduling, sensing design depends on particular relay schemes.

In Section 2.4, we derived the theoretically maximum achievable throughput under cooperative relay and hybrid relay respectively. In particular, we find that without considering sensing time, the achievable throughput with cooperative relay obtained in (2.2) will almost increase linearly with the number of channels being sensed $\|A\|$ when the total number of channels $\|C\|$ is sufficiently large. The increase in the number of channels being sensed have double effects. First, it can increase the shared common channel probability $P_s = \frac{\|A\|^2p^2}{\|C\|}$, and thus reduce the critical transmission range for connectivity $r = \Theta(\sqrt{\frac{1}{P_s}})$, which has a *positive effect* on throughput. Second, it has a *negative effect* on the throughput since it induce sensing overhead time $T_I$ and thus reduce the fraction of transmission time $\frac{T_p-\|A\|T_I}{T_p}$. Combining these two effects, the achievable throughput will be the product of the result in (2.2) and the fraction of transmission
time $T_p - \|A\|T_2$, which increases with small $\|A\|$, even we consider the sensing overhead, given the sensing slot $T_p$ and sensing time of a single channel $T_I$.

Moreover, although the achievable throughput with hybrid is $\lambda(n) = \Theta\left(\frac{W}{n\|C\|}\right)$, which appears to be independent of $\|A\|$, we must note that the critical transmission range $r$ for connectivity, which is a prerequisite for non-zero throughput, still depends on $\|A\|$. Next we will provide a simple sensing algorithm as an example based on the above observations.

### 2.5.2 An Example Sensing Algorithm CASA to Achieve Maximum Throughput under Cooperative Relay

As an example, we consider a large-scale CRN using cooperative relay and consisting of a set of channels $C = \{1, 2, \ldots, \|C\|\}$. We assume that network connectivity is maintained by a well-designed scheduling algorithm, i.e., given the number of channels being sensed $\|A\|$, the transmission range will satisfy the conditions for connectivity in Section 2.3.4, which let us observe the performance of sensing under the same assumptions used in self-centric algorithms.

We further assume that each secondary node $v_i$ uses exactly one of the channels in transmission. Each channel $k$ is associated with a reward of transmission denoted by $R_{ik}$, which can be the data rate of using channel $k$ and it is in essence a random variable because of the randomness stemming from time-varying and uncertain nature of the wireless medium. As in [46, 48, 49], we assume the independence in channels and in time, i.e., the channel state within a single time slot is independent of the state during other slots and in a particular slot, $\{R_{ik}\}_{k \in C}$ are independent random variables.

For simplicity, we further assume that an optimal sensing period $T_p$ is found and it satisfies the sensing requirements (for example, given the sensing requirements, we can use some similar algorithms to [42, 43]). Then according to our discussion in Section 2.2, the connectivity-agile sensing algorithm is formulated as a problem of finding the subset of channels being sensed and the sensing order to maximize the achievable throughput. The problem can be described as follows.
**Definition 2.4.** Find sensing strategy that achieves the following maximum:

\[ G^* = \max_{\pi \in \Pi} E[R_{i\pi(\tau)} \cdot (T_p - T_I \cdot (\tau - 1)) \cdot (\tau - 1)] \]

where \( \pi \) denotes the strategy to sense channels in sequence \( \pi(1), \ldots, \pi(\tau - 1) \) and then transmit over channel \( \pi(\tau) \) at time \( \tau \). \( \Pi \) denotes the set of all possible strategies. Let \( \pi^* \) denote the optimal strategy to achieve \( G^* \).

Note that \( \tau \) is a random stopping time, which in general depends on the result of channel sensing. Based on our previous analysis, the maximum achievable throughput almost increase linearly with \( (T_p - T_I \cdot (\tau - 1)) \cdot (\tau - 1) \) and thus we use \( E[R_{i\pi(\tau)} \cdot (T_p - T_I \cdot (\tau - 1)) \cdot (\tau - 1)] \) as the objective function of sensing. We use a dynamic programming approach [86] to solve this optimization problem. Specifically, at any step, a sufficient information state [86] is given by the pair \((u, S)\), where \( S \subset C \) is the set of unsensed channels and \( u \) is the highest value among sensed channels \( C - S \). Let \( V(u, S) \) denotes the value function, i.e., maximum expected remaining reward given the system state \((u, S)\), then we have

\[ V(u, S) = \max \left\{ g(S)u, \max_{k \in S}\{ V(\max\{R_{ik}, u\}, S - k) \} \right\} \]

where \( g(S) = \|C - S\| \cdot (T_p - \|C - S\|T_I) \). The first term on the right hand side of (2.6) represents the expected reward of using the best-sensed channel, and the second one is the reward of sensing the unsensed channel.

Note that while the dynamic programs are readily available, computing the value function \( V(u, S) \) for every state is very difficult and practically impossible because the state space is potentially infinite and uncountable if \( R_{ik} \) is a continuous random variable. Therefore, instead of attempting to compute these values and the associated strategies, we use the above formulation to derive fundamental properties of optimal strategies and use them to find simpler ways of determining optimal strategies. Specifically, we have the following results:

**Theorem 2.7.** Given any state \((u, S)\), if \( \|A\| = \|C - S\| \leq \frac{T_p}{2T_I} \), then the optimal ac-
tion is to choose one channel from $S$ to sense, that is, $\pi^*(u, S) = \text{sense}(k^*)$, where $k^* = \arg \max_{k \in S} \{ E[R_{ik} | R_{ik} \geq u] \cdot P(R_{ik} \geq u) \}$.

Proof. When $\| C - S \| \leq \frac{T_p}{2IL}$, for any $k \in S$, $\max_{k \in S} \{ V(\max\{ R_{ik}, u \}, S - k) \} \geq V(\max\{ R_{ik}, u \}, S - k) \geq g(S - k) \cdot u \geq g(S) \cdot u$. Thus, $V(u, S) = \max_{k \in S} \{ V(\max\{ R_{ik}, u \}, S - k) \}$, which implies that the optimal action is to sense an unsensed channel. Now we need to determine which channel of $S$ should be sensed. Intuitively, the channels should be sensed in the decreasing order of the expected reward $E[R_{ik}]$. However, note that once a sender senses a channel with reward $u$, it can not increase its reward any further by sensing channels with reward $u$ or lower. Thus we sense the channels in the non-increasing order of $E[R_{ik} | R_{ik} \geq u] \cdot P(R_{ik} \geq u)$, which can be considered as the expected reward at state $u$. Actually, this can be seen as a generalization of Theorem 4.1 in [45], which constrains $R_{ik}$ to be a discrete random variables. \hfill \Box

**Theorem 2.8.** Given state $(u, S)$ with $\| A \| = \| C - S \| > \frac{T_p}{2IL}$, define $H_u = \{ k \in S : E(R_{ik}) > \frac{g(S)}{g(S - k)} u \}$ and $k^* = \arg \max_{k \in H_u} \{ E[R_{ik} | R_{ik} \geq u] \cdot P(R_{ik} \geq u) \}$. If $H_u \neq \emptyset$, then sense channel $k^*$; otherwise, choose $k \in C - S$ with $E[R_{ik}] = u$ to transmit.

Proof. Note that a node stops sensing and chooses one of the sensed channels with the maximum reward $u$ to transmit when $V(u, S) = g(S) \cdot u$. When $\| C - S \| > \frac{T_p}{2IL}$, $g(S)$ is an increasing function. If $u = \max_{k \in S} E[R_{ik}]$, then $V(u, S) = g(S) \cdot u$ and the node should stop sensing and choose one sensed channel to transmit. Otherwise, $V(u, S)$ may achieve higher value by sensing more channels. In addition, $\frac{g(S)}{g(S - k)}$ is a decreasing function of $S$ when $\| C - S \| > \frac{T_p}{2IL}$ and thus, define

$$H_u = \{ k \in S : E(R_{ik}) > \frac{g(S)}{g(S - k)} u \}.$$  

If $H_u = \emptyset$, $V(u, S) = g(S) \cdot u$, that is, the node stops sensing and chooses one sensed channel to transmit; otherwise, we have

$$\max_{k \in S} \{ V(\max\{ R_{ik}, u \}, S - k) \} \geq g(S - k) \cdot E[R_{ik}] > g(S) \cdot u$$

46
Table 2.1: Connectivity-Agile Sensing Algorithm.

A Connectivity-Agile Sensing Algorithm (CASA):

Given state \((u, S)\)

**Step 1:** Compute \(\frac{J_p}{2T_f}\) and \(\|A\| = \|C - S\|\).

**Step 2:** If \(\|C - S\| \leq \frac{J_p}{2T_f}\), \(\lambda^*(u, S) = \text{sense}(k^*)\), where \(k^* = \arg\max_{k \in S} \{E[R_{ik}|R_{ik} \geq u] \cdot P(R_{ik} \geq u)\}\).

Replace \(S\) with \(S - k^*\) and \(u = \max\{u, E(R_{ik})\}\), repeat 1.

**Step 3:** If \(\|C - S\| > \frac{J_p}{2T_f}\), find \(H_u = \{k \in S : E(R_{ik}) > \frac{g(S)}{g(S-k)} u\}\).

If \(H_u = \emptyset\), choose the channel \(k \in C - S\) with \(E[R_{ik}] = u\) to transmit. Otherwise, compute \(k^* = \arg\max_{k \in H_u} \{k \in S : E[R_{ik}|R_{ik} \geq u] \cdot P(R_{ik} \geq u)\}\) and \(\pi^*(u, S) = \text{sense}(k^*)\).

Replace \(S = S - k^*\) and \(u = \max\{u, E(R_{ik})\}\), repeat Step 1.

for some \(k \in H_u\) and the node should choose one channel from \(H_u\) to sense. Similarly, note that once a sender observes sensed channel with reward \(u\), it can not increase its gain. Thus we sense the channel in the non-increasing order of \(E[R_{ik}|R_{ik} \geq u] \cdot P(R_{ik} \geq u)\). Thus the result.

The above two Theorems show that at the beginning \((S = C)\), the shared common channel constraint dominates the sensing overhead. When \(\|A\| = \|C - S\| > \frac{J_p}{2T_f}\), the sensing overhead will gradually overshadow the constraint of shared channels. Therefore, the tradeoff between finding better shared channels and cost of sensing overhead needs to be balanced. The above two theorems provide a threshold structure of optimal sensing with respect to \(S\). The details are shown as Algorithm CASA (Fig. 2.1).

### 2.5.3 Performance Evaluation

From analysis and discussions, we know that shared common channels are very important in CRNs, especially large-scale networks in that each secondary node may need to sense a set of channels in order to identify a common channel for communications. Intuitively, the larger the set, the more likely that secondary nodes would be able to communicate; otherwise, they may not be able to communicate, yielding low or even zero throughput. Therefore, we use simulations...
to examine the impact of shared common channels, and then implement the proposed CASA algorithm with Network Simulator (NS2) for performance evaluation.

We find that in many previous works, sensing algorithms are studied analytically, without simulations. The main challenge is that we need to develop new extensions for multi-channel cognitive radio networks, and corresponding MAC-layer function, such as spectrum sensing, and channel assignment to avoid collisions because secondary nodes may use different channels. This is different from single/multiple radio networks where there exist common channels among nodes [87]. We configure an example network of 50 static secondary nodes using cooperative relay, which are randomly distributed in a $1000m \times 1000m$ square, and transmission range $r = 250m$. There are 26 sources (more than 50% of nodes in the network), sending 4 packets per second and the simulations are run for 1000 seconds. Therefore, we can find sufficient data traffic and evaluate network performance.

We show the impact of shared common channels on network performance in Fig. 2.6. Recall our discussions in Section 2.2.2, shared common channels are mandatory for secondary nodes to communicate. However, finding shared common channels depends on the number of channels and the set of channels to be sensed. Specifically, we assume that secondary nodes can use $\|C\| = 10$ iid channels and each channel is idle with probability 0.8. We find that even with ideal sensing without overhead, if the number of sensed channels is small (here for $\|A\| \leq 5$), the number of received packets is almost zero, which is omitted in Fig. 2.6 (a). This is because the secondary nodes cannot identify shared common channels when each node can use 10 channels in a medium-size network (50 nodes). In other words, the number of sensing channels must be sufficiently large to find shared common channels; otherwise, it is unlikely to have communications among secondary nodes, even though idle channels are present. We gradually increase the number of sensing channels, and find that when $\|A\| \geq 6$, there exist some communications among the sources and destinations. The more channels are sensed, the higher the packet delivery ratio is. In addition, Fig. 2.6 (b) shows that the average hop counts decrease from 4.8 to 3.0, which can be beneficial to routing performance as well because hop-
count may also be used as a routing metric. This is evidenced by a significant improvement of average delay in Fig. 2.6 (c), which decreases from 0.4 to 0.02 second and packet delivery ratio, which is defined as the number of received packets over the number of sent packets, increasing substantially from 28% to 56% in Fig. 2.6 (a). These results suggest that the number of sensed channels is very important to packet delivery ratio and average delay in CRNs.

![Graphs](image)

(a) average delivery ratio  
(b) average hops  
(c) average delay

Figure 2.6: The Effect of Channel Channel Constraint.

Next, we examine the performance of the proposed CASA algorithm under a practical class of channel models. We first take a look at the performance of CASA in a small network where a secondary source node $v_s$ transmits to a secondary destination node $v_d$ through a relay node $v_r$. Specifically, we use a two-state channel model where for each channel $k$ and secondary node $v_i$, $P(R_{ik} = r_{ik}) = p_{ik} = 1 - P(R_{ik} = 0)$ for some $r_{ik} > 0$, where $R_{ik}$ denotes the reward of channel $k$ on node $v_i$, which can be transmission rate (or normalized as the fraction of channel capacity) and availability. That means, channel $k$ is available for $v_i$ with probability $p_{ik}$ and when it is available, the reward is a random variable $r_{ik}$. Also, $r_{ik}$ and $p_{ik}$ are modeled as independent random variables, uniformly distributed in the interval $(0, 0.5, 1)$. For illustration purpose, we consider a sensing period $T_p = 1s$ and sensing time $T_I = 50ms$, though $T_I$ is much less in reality. This is because we only consider a small number of channels and increase $T_I$ to show the tradeoff between finding a better channel and sensing overhead.

We compare the average throughput and average idle time with four algorithms: Algorithm
CASA which abides by the constraints of CRNs; a self-centric algorithm, similar to [49] in which maximum throughput was used as the optimization target without consideration of the observations discussed earlier; and a ideal algorithm with perfect sensing without any overhead, which is used for observing the deficit due to sensing overhead; and no-CR technique to show the improvement of idle time. A total of $10^4$ random realizations are generated: for each realization of $r_{ik}$ and $p_{ik}$, the throughput and channel idle time are computed and then averaged values are obtained. It can be seen from Fig 2.7a that the average throughput with CASA is more than 2 times for a small number of channels, while it may be up to 10 times more for larger number of channels (e.g., $||C|| = 10$). However, in Fig. 2.7b, the idle time resulting from self-centric algorithm is about twice as much as the result of CASA, while the discrepancy becomes smaller as the number of channels increases because we consider a single secondary Tx-RX pair which do not generate sufficient traffic all the time, while both sensing algorithms have shown significant improvement of reduction in average idle time, i.e., better spectrum utilization. Thus we come to the conclusion that although CASA is designed for large-scale CRNs, it performs much better than self-centric algorithms even in small networks because it take common channel constraint into account.

Furthermore, we study the performance of CASA in our ns-2 simulation network. Note that our simulation results of 50-node network can be considered for medium-size networks in

![Figure 2.7: Comparison of Sensing Algorithms.](image-url)
comparison with similar studies with less than 10 nodes. We compare the number of received packets, average delivery ratio, average hop counts, and average delay with three algorithms: CASA, a self-centric algorithm, and an ideal algorithm with perfect sensing without any overhead. We assume sensing period $T_p = 1s$ and sensing time $T_I = 50ms$. Fig. 2.8a plots the total number of received packets. We find that using the self-centric algorithm, secondary nodes cannot receive any data because of the lack of common channels for communication and thus these performance metrics are not available. As shown in Fig. 2.8a, the number of received packets almost increase linearly with the number of usable channels, from 100 to more than 400. We also observe some sensing overhead between CASA and the ideal sensing algorithm. Fig. 2.8b and Fig. 2.8c plot the packet delivery ratio and average hop counts respectively and we find that CASA performs as good as the optimal sensing algorithm in terms of packet delivery ratio and hop counts. Particularly, we find that as the number of usable channels $\|C\|$ increases, the delivery ratio increases from 0.22 to 0.58 and the average hop count decreases from 3.3 to 3.04. This implies that increasing the number of channel accessible to secondary nodes can greatly improve the CRN performance with appropriate sensing algorithms. Fig. 2.8d plots the average delay which decreases almost linearly from 0.082 second to 0.025 second as $\|C\|$ increases from 3 to 10. It is also observed in Fig. 2.8d that there is a subtle delay overhead of CASA, compared with an ideal sensing algorithm without overhead. Therefore, we come to the suggestion that

Figure 2.8: Performance of Connectivity-Agile Sensing Algorithm (CASA).
unique features of CRNs do have evident impacts on the design of sensing algorithms toward the maximum throughput.

2.6 Summary

In this chapter, we studied the maximum achievable throughput in CRNs by considering unique characteristics and demonstrated that it may potentially impact the design objectives and expectations of spectrum sensing algorithms. In particular, we find that the lack of common channels among secondary nodes, relay schemes, and connectivity can have a mixed effect on the achievable throughput. For example, sensing algorithms may induce sensing overhead, thus decreasing the throughput. However, it also has a positive effect on finding common channels among neighboring nodes to improve throughput. In addition, while using primary nodes as relays can take the advantage of higher power and transmission rate than that of secondary nodes, the throughput may not be improved because they can potentially become the routing hot spots and hence network bottlenecks due to secondary nodes’ transmission collisions. By taking these networking constraints into account, we formally derived the throughput limit in CRNs and proposed a class of connectivity-agile sensing algorithms to achieve this limit. To validate our results, we designed a sensing algorithm CASA for a large-scale CRN using cooperative relay as an example and demonstrated very promising packet delivery ratio with negligible delay overhead.
Chapter 3

The Tempo-spatial Limits of Information Dissemination in Cognitive Radio Networks

In this chapter, we study the information dissemination in large multi-channel CRNs regarding their temporal and spatial limits, respectively, which is essential to understanding the fundamental properties and critical applications of CRNs. First, we define two metrics, dissemination radius $\|L(t)\|$ and propagation speed $S(d)$. The former is the maximum Euclidean distance that a packet can reach in time $t$, and the latter is the speed that a packet transmits between a source and destination at Euclidean distance $d$ apart, which can be used to measure the transmission delay. Further, we determine the sufficient and necessary conditions under which there exist spatial and temporal limits of information dissemination in a percolated CRN, by identifying the correlation between critical density of primary nodes, secondary nodes, and available channels. We find that when information cannot be disseminated to the entire network, the limiting dissemination radius $L(\infty)$ is statistically dominated by an exponential distribution, while the limiting information propagation speed approaches zero. Otherwise, the dissemination radius approaches infinity and the propagation speed $S(d)$ is no lower than some constant $\kappa$ for large
3.1 Motivation and Related Work

As a promising solution to the problem of limited frequency bandwidth and inefficiency in spectrum utilization, CRNs have become an important component of our communication infrastructure for a variety of application scenarios, such as military networks, emergency networks, cognitive mesh networks, and leased networks [2]. In CRNs, each secondary node is equipped with a cognitive radio and thus can opportunistically access multiple channels without interfering with the licensed nodes, which are also called primary nodes, by exploiting spectrum opportunistically [2]. Therefore, in recent years, there has been intensive research on understanding and optimizing performance limits, such as capacity, spectrum sensing, spectrum mobility, and spectrum sharing of CRNs [2,88–92], which provides insights on improving spectrum efficiency and traffic capacity in CRNs.

It is worthy of noting that the ultimate goal of CRNs is supposed to offer information delivery opportunities in circumstances otherwise hardly possible, which in turn, is the objective of exploiting inefficiently used frequency spectrum. Therefore, an interesting yet open question is that what are the achievable benefits of information dissemination in such networks, which is essential to fully explore the potentials and critical applications of CRNs. Understanding how packets disseminate and their temporal and spatial limits, such as the maximum information coverage, e.g., dissemination area, transmission speed and latency, can also be beneficial to the deployment, design and applications of CRNs. For example, when a CRN is used for emergency rescue in the aftermath of disasters or traffic accidents (e.g., vehicular networks), we need to ensure that help or warning messages can be disseminated to a sufficiently large area, and to estimate how long it takes for such information to reach a chosen destination, which become more important than other performance metrics, such as the total network capacity in these circumstances.

Similar problems, on the other hand, have been studied for wireless multihop networks or
sensor networks [56–62], which explore the conditions for connectivity or percolation in order to ensure the information can be disseminated to the whole network. In addition, information propagation speed or delay has been discussed in recent works [61,62], which categorized the delay into bandwidth-incurred propagation delay and topology-incurred delay. In particular, when the network topology remains unchanged or changes very slowly, the bandwidth-incurred propagation delay, which is the transmission time spent by a packet in all the links along its transportation path, is dominant, while topology-incurred delay is negligible.

However, these existing results on multihop wireless networks are not applicable to CRNs. For instance, the network topology in CR networks is dynamic not only because of factors such as user mobility and radio link quality, but it is more or less due to the opportunistic channels available over time. As a result, the network is more likely to be a percolated network, that is, a network is almost surely connected, instead of a fully connected one as assumed in earlier works [61,62,71,72,81,82]. Furthermore, in current studies on network topology and performance, the critical density is a key condition in identifying whether a network is percolated or not. When node density is higher than critical density, the network is considered percolated; otherwise, the network is not percolated. The challenge is that prior results are based on a common assumption of homogeneous networks in which all nodes are the same. Nonetheless, a CRN is intrinsically a heterogeneous network because primary nodes and secondary nodes are different from each other regarding their transmission ranges, usages of communication channels, locations, and routing functions. Therefore, how to determine the conditions for a percolated CRN is an unknown problem to be resolved.

Therefore, we focus on the following questions in this chapter: (i) for a large multi-channel CRN, how far can a packet originated from an arbitrary node be disseminated? (ii) When a packet can be disseminated to a sufficiently large area, how long does it take this packet to reach a chosen destination? To tackle these problems, we define two new metrics, the disseminating radius $\|L(t)\|$ and the propagation speed $S(d)$ to study the spatial and temporal limits, respectively. The former is the maximum Euclidean distance that a packet disseminates in time $t$
and can be used to characterize the dissemination area; and the latter one is the speed that a packet transmits between a source and destination at distance $d$ apart, which can be used to interpret the end-to-end delay. Here, we focus on the topology-incurred delay by ignoring bandwidth-incurred propagation delay.

Our main contributions in this chapter are as follows: (1) We determine the sufficient and necessary conditions under which there exist theoretical limits of information dissemination by identifying the correlation between critical density of primary nodes, secondary nodes, and accessible channels. (2) We find that when the packets cannot percolate to infinite area, the limiting dissemination radius $\|\mathcal{L}(\infty)\|$ is statistically dominated by an exponential distribution and the limiting information propagation speed approaches zero. When the packets are able to reach infinite area, i.e., the dissemination radius approaches infinity, the propagation speed $S(d)$ is no lower than some constant $\kappa$ for large $d$.

### 3.2 Preliminaries and Problem Definition

In this section, we first describe assumptions and network models used in the paper, along with a brief description of our approach. Then we define the dissemination radius and propagation speed in order to study the information dissemination in spatial and temporal domains.

#### 3.2.1 Assumptions and Models

We consider a large CRN consisting of $n$ secondary nodes independently and identically (i.i.d.) distributed in a region $\mathcal{A} = [0, \sqrt{\frac{\pi}{\lambda}}]^2$ for some constant $\lambda$. Denote $X_i (1 \leq i \leq n)$ as the random locations of secondary nodes. When $n \to \infty$, $\mathcal{H}_\lambda = \{X_1, \ldots, X_n\}$ can be presented by a homogeneous Poisson point process with density $\lambda$ [93]. In wireless multihop networks, each node is usually assumed to have a fixed transmission range. On the contrary, CR nodes may adapt their transmission ranges independently over time to save energy or limit interference with PR nodes. We assume that the transmission range $r_i$ of secondary node $v_i$ follows a common distribution $F_r$ with $\mathbb{P}(r_i < \gamma) = 1$ for some constant $\gamma$. And in this chapter, we only consider
\( \gamma = 1 \) but results derived here can be easily extended to the scenarios with any \( \gamma > 0 \). Let \( d = \|v_i - v_j\| \) denote the Euclidean distance between \( v_i \) and \( v_j \). Information may propagate between \( v_i \) and \( v_j \) in one hop only if \( d \leq \min(r_i, r_j) \) as shown in Fig. 3.1a.

\[ \frac{d}{\sqrt{r_i r_j}} \]

Figure 3.1: Cognitive Radio Network Model: secondary nodes and PR nodes

Moreover, we consider that the secondary nodes are overlaid with Poisson distributed PR networks using a sufficiently large independent channel set \( \mathcal{B} = \{ch_1, \ldots, ch_m\} \). Each \( ch_k \) is licensed to an overlaid primary network, where primary nodes are i.i.d distributed and the spatial density of the transmitting (active) primary nodes is \( \lambda_{pk} \). Denote \( R_I \) as the interference range of the primary nodes. Let \( d = \|v_i - v_j\| \) and \( \mathcal{S}(d, R_I) \) denote the region covered by the two circles with radius \( R_I \) centered at \( v_i \) and \( v_j \), as shown in Fig. 3.1b. We say \( ch_k \) is available to the edge \( v_i v_j \) if there is no transmitting PR node using \( ch_k \) within \( \mathcal{S}(d, R_I) \). Denote \( \mathbb{P}_s \) as the probability that there exist at least one channel available to \( v_i v_j \).

We need to emphasize here that since our model only assumes the spatially stationary distribution of transmitting primary nodes, this model is very general and accounts for a wide range of dynamics of primary nodes. For example, our model can be used to scenarios where primary nodes are mobile under the i.i.d mobility model [64]. And it can be also used to model the sensor network where sensor nodes are i.i.d distributed and each sensor alternates between active and inactive state independently [57].
Furthermore, since secondary nodes can only opportunistically use spectrum without interfering with the communications among primary nodes, spectrum sensing is essential to cognitive radio networks. In this chapter, we assume that an uncoordinated sensing \[57\] scheme is used for its simplicity of not inducing coordination overhead. Particularly, each secondary node alternates independently between the communicating (active) state and sensing (inactive) state with periods determined by the stationary i.i.d on/off process \(W(t)\). And \(\eta\) is denoted as the stationary probability of a secondary node being active.

The cognitive radio network is represented as \(G(\mathcal{H}_\lambda, F_r, W(t))\) in the rest of this paper.

3.2.2 Percolated Cognitive Radio Networks

*Full connectivity,* i.e., there exists a path between any two nodes, has been used as a prerequisite for applications and performance analysis of general purpose wireless ad hoc networks in the earlier studies \[61,62,82\]. However, considering the opportunistic channels available to secondary nodes over time, this *fully connectivity* criterion is overly restrictive and even impossible to achieve in cognitive radio networks.

Recent research has proposed a notion of connectivity based on *continuum percolation* \[94\]. The main result of percolation theory is that there exists a finite, positive value of the transmission range, or equivalently of the node spatial density, above which the network is percolated (super-critical) and below which it is not percolated (sub-critical). When the network is percolated, there exists a large connected component of nodes spanning almost the entire network (called *giant component* in \[94\]), and when the network is not percolated, the network consists only of small isolated connected components of nodes. This phase transition in the macroscopic behavior of large-scale wireless networks is defined as *critical phenomenon* in percolation theory \[94\].

In this chapter, instead of *full connectivity* which can ensure information disseminating to the entire network, we focus on *percolated connectivity*. Particularly, we are interested in how fast information can be disseminated in the *giant component* when the cognitive radio
network is percolated and how far information can reach when the cognitive radio network is not percolated. To address our problems, we introduce the following concepts first.

**Definition 3.1.** If $Z$ and $Z'$ are random variables such that $\mathbb{P}(Z > z) \leq \mathbb{P}(Z' > z)$, we say that $Z$ is stochastically dominated by $Z'$ and write $Z \prec \prec Z'$; and if $Z \prec \prec Z'$, there exists a random variable $\hat{Z}'$ which has the same distribution of $Z'$ such that $Z \leq \hat{Z}'$ ($\hat{Z}'$ is called a coupling of $Z'$, see p127 of [95] for more details.) Thus for two events $E_1$ and $E_2$ with $\mathbb{P}(E_1) > \mathbb{P}(E_2)$, if $E_2$ happens, we can assume $E_1$ also happens by coupling. Coupling is an elegant method used in probability theory and has been extensively used in continuum percolation (see p28 of [94]).

We have found that there are several challenges in using continuum percolation theory in solving problems in CRNs. For example, we find that

- **Poisson Boolean model is not valid:** Poisson Boolean model, where nodes are distributed as a Poisson process and each node has some fixed transmission range, has been extensively used to model and study wireless networks [14, 57–59, 71, 79, 96, 97]. One of the main features of secondary nodes is that they may have adaptive transmission range to the wireless radio environments in order not to interrupt the ongoing communications among PR nodes or reduce the interference with others. Let $d = \|v_i - v_j\|$ denote the Euclidean distance between secondary nodes $v_i$ and $v_j$ with random transmission range $r_i$ and $r_j$, instead of $d \leq (r_i + r_j)/2$ as required in Poisson Boolean model, the bond $v_i v_j$ is open (active) if $d \leq \min(r_i, r_j)$. Thus the Poisson boolean model cannot be used to study CR networks.

- **Bonds (edges) are not independent:** A common assumption used in the existing percolation models is that the bonds (edges) are independent of each other, which is necessary to the derivation of percolation results (see [94, 98, 99]). However, in CRNs, there exists some dependency among bonds. For example, given three nodes $v_i$, $v_j$ and $v_k$, if bonds $v_i v_k$ and $v_j v_k$ are open (active), which implies $v_i$ and $v_j$ are in the neighborhood of $v_k$ and there exist available channels to $v_i v_k$ and $v_j v_k$, then $v_i$ is in the neighborhood of $v_j$ and there
exist some channels available to $v_i v_j$ with higher probability. Equivalently speaking, given bonds $v_i v_k$ and $v_j v_k$ are open, the bond $v_i v_j$ has higher probability to be open. Therefore, bonds $v_i v_j$, $v_i v_k$ and $v_j v_k$ are not independent.

Hence the existing approaches and results from percolation theory cannot be directly applied to our network models, which is another challenge for this theoretical work. For example, to derive the upper bound of critical density, a wireless multihop network is usually mapped to an independent discrete percolation model, whose properties have been well studied [57, 58, 99]. However, a CRN can only be mapped to a $k$-dependent discrete percolation model, and further complicates the analysis.

### 3.2.3 Limits of Information Dissemination

![Illustration of information dissemination in a percolated CR network.](image)

Figure 3.2: Illustration of information dissemination in a percolated CR network.

We consider that information is disseminated through broadcasting in CRNs. Let us denote $\mathcal{V}(t)$ as the cluster of nodes that have received the packet by time $t$, given that the packet is sent at time $t = 0$. The dissemination area at $t$, $\mathcal{A}(t) \in \mathbb{R}^2$, that is, the total area covered by $\mathcal{V}(t)$, can be expressed by $\mathcal{A}(t) \triangleq \bigcup_{v_i \in \mathcal{V}(t)} B(v_i, 1)$, where $B(x, r)$ is a ball with radius $r$ centering at point $x \in \mathbb{R}^2$. An illustration of information dissemination in a CRN $G(H_\lambda, F_r, W(t))$ is shown in Fig.
3.2 in which a packet is originated by node $v_0$. When $v_0$ starts broadcasting this packet at time $0$, its neighbors receive the packet and they rebroadcast the packet. Here two nodes $v_i$ and $v_j$ are called “neighbors” at $t = t'$ if the link $G(\mathcal{H}_\lambda, F_r, W(t'))$. Without considering propagation delay, at $t = 0$, the packet has spread to cluster $\mathcal{V}(0) \subset G(\mathcal{H}_\lambda, F_r, W(t))$ containing $v_0$ (see Fig. 3.2a). We refer $\tau$ as topology-incurred delay for node $v \in \mathcal{V}(\tau) \setminus \mathcal{V}(\tau^-)$ to receive this packet from $v_0$. Thus at $t = \tau$, because some nodes change their states, a node $v_1 \notin \mathcal{V}(0)$ but inside $\mathcal{A}(0)$ receives the packet from its neighbors and then broadcasts the packet to the cluster containing $v_1$ at $\tau$.

![Diagram of dissemination area](image)

**Figure 3.3**: An illustration of dissemination area after long time $t$.

Without loss of generality, let us consider the node $v_0$ located at the origin $o \in \mathbb{R}^2$. Illustrations of dissemination area for sufficiently large $t$ are shown in Fig. 3.3. Since the network is not fully connected, any sufficiently large area $\mathcal{B}$ is only partially covered by $\mathcal{A}(t)$. The uncovered area, which is also called *vacancy* in this chapter, is shown in Fig. 3.3 as shaded areas. Letting $\mathcal{A}_\infty = \lim_{t \to \infty} \mathcal{A}(t)$, we illustrate $\mathcal{A}_\infty < \infty$ and $\mathcal{A}_\infty = \infty$ in Fig. 3.3a and Fig. 3.3b respectively. Denote $\mathcal{L}_\varphi$ as the line starting from the origin $o$ in the direction $\varphi \in [0, 2\pi)$ and $\mathcal{L}_\varphi(t) = oz$, where $z = \arg \max_{v \in \mathcal{L}_\varphi \cap \mathcal{A}(t)} \|v\|$ is the farthest intersection point between $\mathcal{L}_\varphi$ and $\mathcal{A}(t)$. For example, in Fig. 3.3b, $\mathcal{L}_{\varphi_1}(t)$ is the line segment $\overline{v_0 o}$. The length of $\mathcal{L}_\varphi(t)$, $\|\mathcal{L}_\varphi(t)\|$, is defined as the *transmitting distance* at $t$. 

61
**Definition 3.2** (Dissemination radius $\|L(t)\|$). The dissemination radius at time $t$ is defined as $\|L(t)\| = \max_{\phi \in [0,2\pi]} L_\phi(t)$; and the limiting dissemination radius is defined as $\|L(\infty)\| = \lim_{t \to \infty} \|L(t)\|$.

The dissemination radius indicates how far a packet can reach in spatial domain in a large network. Next, we move on to the temporal domain. Given $L_\phi(t)$, an intuitive definition of information propagation speed is $\frac{\|L_\phi(t)\|}{t}$, as in [62], which tells the maximum speed in direction $\phi$. However, due to the dynamics in CRNs, a node closer to $v_0$ may receive the packet later than a farther node. For example, in Fig. 3.2, $v_2$ receives the packet later than $v_1$. Instead of maximum speed, we are more interested in how long the packet can be disseminated to a chosen destination at $d$ apart. Thus, the Information Propagation Speed is defined as follows.

**Definition 3.3** (Information propagation speed $S_\phi(d)$). Let $r_\phi(d)$ be the point on $L_\phi$ with $\|r_\phi(d)\| = d$ (see Fig. 3.3). Denote $T(v_0,v) \triangleq \arg\min_{v \in V(t)} \{v \in V(t)\}$ as the topology-incurred delay of the node $v$. Denote $\tilde{v}_\phi(d)$ as the node closest to $r_\phi(d)$ which can receive the packet. That is, $\tilde{v}_\phi(d) = \arg\min_{v \in V_\infty} \|v - r_\phi(d)\|$, where $V_\infty = \lim_{t \to \infty} V(t)$. When $\|L(\infty)\| = \infty$, the Information Propagation Speed in direction $\phi$ is defined as $S_\phi(d) \triangleq \frac{d}{T(v_0,\tilde{v}_\phi(d))}$. And when $\|L(\infty)\| < \infty$, $S_\phi(d) \triangleq \frac{d}{T(v_0,\min_{\phi})}$ for $d \leq \|L(\infty)\|$, and $S_\phi(d) \triangleq 0$ for $d > \|L(\infty)\|$. The limiting propagation speed $S_\phi(\infty)$ is defined as $\lim_{d \to \infty} S_\phi(d)$.

The definition of $S(d)$ denotes the propagation speed in an arbitrary direction. For convenience of our analysis, we have the following definition, which is to differentiate infinite and finite limits of information dissemination in this paper.

**Definition 3.4.** For a packet $b$ originated by $v_0$ at $t = 0$, $b$ is said to be disseminated locally if $L(\infty) < \infty$ and globally otherwise. Particularly, $b$ is said to be disseminated globally “instantaneously” if $L(0) = \infty$, be disseminated globally “within finite time” if $\|L(0)\| < \infty$ but $\|L(\phi)\| = \infty$ for some bounded $\phi$, and be disseminated “gradually” if $L(t) < \infty$ for any $t$ but $L(\infty) = \infty$. 62
3.3 How Far Can Information Be Disseminated in A CR Network?

In this section, we identify the sufficient and necessary conditions under which information can be disseminated globally. We first prove the existence of the critical phenomenon in cognitive radio networks by coupling, then derive the sufficient condition for percolated cognitive radio networks by continuous-to-discrete mapping and the necessary condition by using a Link Correlation Coefficient (LCC) approach.

3.3.1 Critical Phenomenon

We first show that critical phenomenon in general purpose wireless networks (modeled by random geometric graph $G(H, 1)$ in [94]) also exists in the cognitive radio network $G(H, F, W(t'))$ at any time snapshot $t'$, through the coupling approach. Denote $\lambda_{c,w}$ as the critical density of $G(H, 1)$. Note that if all secondary nodes transmit with the maximum transmission range 1 and always stay active, and no primary nodes are transmitting, $G(H, F, W(t'))$ is equivalent to $G(H, 1)$, which implies that $G(H, F, W(t')) \subset G(H, 1)$. If $G(H, F, W(t'))$ percolates, $G(H, 1)$ also percolates by coupling. Therefore, when $\lambda < \lambda_{c,w}$, $G(H, 1)$ is not percolated, and thus $G(H, F, W(t'))$ is not percolated neither. On the other hand, as shown later, by mapping $G(H, F, W(t'))$ to a discrete lattice, $G(H, F, W(t'))$ percolates for some large enough $\lambda$. Let $\mathcal{E}$ denote the event that $G(H, F, W(t'))$ is percolated. Denote $P_\lambda(\mathcal{E})$ as the probability that $\mathcal{E}$ happens. Note that $\mathcal{E}$ is a tail event [100] and thus by Kolmogorov’s 0-1 theorem [100], $P_\lambda(\mathcal{E})$ is either 0 or 1. Since $P_\lambda(\mathcal{E})$ is nondecreasing with $\lambda$, and we have shown that when $\lambda$ is small, $P_\lambda(\mathcal{E}) = 0$ and when $\lambda$ is large, $P_\lambda(\mathcal{E}) = 1$, there exists a critical density $\lambda_{c,e}$ below which $P_\lambda(\mathcal{E}) = 0$ and above which $P_\lambda(\mathcal{E}) = 1$. This indicates the critical phenomenon in cognitive radios networks at any time point.

However, for practical applications, we are not only interested in a percolated network at any time snapshot, but more importantly, we need to know whether the network remains percolated...
all the time, despite of dynamics in network topology. By similar proof to Proposition 1 in [57], which proves that if a dynamic wireless network is percolated at any time point, then it is percolated all the time, we have the following lemma.

**Lemma 3.1.** Given \( G(H_\lambda, F_r, W(t)) \) and the critical density \( \lambda_{c,c} \), if \( \lambda > \lambda_{c,c} \), \( G(H_\lambda, F_r, W(t)) \) remains percolated for all \( t \geq 0 \) with probability 1. If \( \lambda < \lambda_{c,c} \), \( G(H_\lambda, F_r, W(t)) \) is not percolated for all \( t \geq 0 \) with probability 1.

### 3.3.2 Conditions for Percolated Cognitive Radio Networks

When a cognitive radio network is percolated, there exists a giant component, which includes the number of nodes of the order of network size \( n \). In this case, the network is *almost surely connected* such that a packet can be disseminated to most nodes in the network. Therefore, to find out the *dissemination radius* and *information propagation speed*, we need to know the *sufficient and necessary* conditions for a cognitive radio network to be percolated. As discussed in Section 3.3.1, critical density \( \lambda_{c,c} \) provides the condition for percolation. However, even for the simplified *random geometric graph* \( G(H_\lambda, 1) \), the exact value of critical density \( \lambda_{c,w} \) is unknown, although some numeric bounds were obtained from rigorous mathematical proofs with a wide range, e.g., \( 0.696 < \lambda_{c,w} < 3.372 \) [94]. Thus to determine the conditions for a cognitive radio network to be percolated is quite challenging. Particularly, we have the following necessary and sufficient conditions for a cognitive radio network to percolate.

**Theorem 3.1.** For a cognitive radio network \( G(H_\lambda, F_r, W(t)) \), given the number of channels \( m \) and the spatial density of primary nodes \( \lambda_p = \{\lambda_{pk}\}_{k=1}^m \), there exists a critical density \( \lambda_{c,c} \) on secondary nodes, above which \( G(H_\lambda, F_r, W(t)) \) remains percolated for all \( t \) and below which \( G(H_\lambda, F_r, W(t)) \) is not percolated for all \( t \). Specifically, we have

\[
\lambda_{c,c} < \frac{1.21 \log \left( \frac{1}{1 - \sqrt{\frac{e}{\eta}}} \right)}{\min_{0 < ||e|| < 0.5} \frac{||e||^2 \eta(1 - F_r(2||e||))}{\frac{1}{1 - \sqrt{\frac{e}{\eta}}}}},
\]

where \( \mathbb{P}_c \) is the bond open probability sufficient for percolation on a dependent triangular lattice.
Figure 3.4: The triangular lattice $\mathcal{D}$ with the “flower” $\mathcal{F}_{s_{i}} = ABCDEF$.

given in Appendix 3.6.2 and $P_{s} = 1 - \prod_{k=1}^{m} (1 - e^{-\lambda_{pk} \alpha})$ with $\alpha$ given in Appendix 3.6.1. Furthermore, almost surely,

$$\lambda_{c,c} > \frac{1}{\Gamma(1 - C_{LCC})},$$

(3.2)

where $\Gamma = 2\pi \eta \int_{0}^{1} \int_{0}^{r} x(1 - F_{r}(x))P_{s}dx dF_{r}$, and $C_{LCC}$ is the Link Correlation Coefficient given in Appendix 3.6.3.

We first derive the sufficient condition for percolation of $G(\mathcal{H}_{\lambda}, F_{r}, W(t))$ by using the technique of continuous-to-discrete percolation mapping. Note that a mapping of $G(\mathcal{H}_{\lambda}, 1)$ on a square lattice has been used in the existing work [58]. Although this square lattice mapping is easy for analysis, we find that the triangular lattice mapping can provide tighter results and thus is used in our study. Particularly, to obtain the sufficient condition, we compare $G(\mathcal{H}_{\lambda}, F_{r}, W(t))$ with a bond percolation model on the triangular lattice $\mathcal{D}$ as shown in Fig. 3.4 with each edge having length $\|e\| < \frac{1}{2}$. Each site (vertex) $s$ of the lattice is enclosed in an area $\mathcal{F}_{s}$, which is called “flower” in this paper. $\mathcal{F}_{s}$ is formed by the six arcs of circles, each of radius $\|e\|$ and centered at the midpoints of the six edges incident on $s$. This formation ensures that for any points $x \in \mathcal{F}_{s_{i}}$ and $y \in \mathcal{F}_{s_{j}}$, $\|x - y\| \leq 2\|e\| < 1$. Denote $\mathcal{B}(s, R_{I} + \|e\|)$ as a circle centering at $s$ with radius $R_{I} + \|e\|$. Thus, there exist available channels within $\mathcal{F}_{s_{i}} \cap \mathcal{F}_{s_{j}}$ if there are no primary nodes using $ch_{k}$ for some $k$ within $\mathcal{B}(s_{i}, R_{I} + \|e\|) \cap \mathcal{B}(s_{j}, R_{I} + \|e\|)$. We define a
site \( s \) to be open if there exist active secondary nodes with transmission range \( r \geq 2\|e\| \) within \( F_{s_i} \). A bond \( e = s_is_j \) is declared open if there exist available channels within \( F_{s_i} \cap F_{s_j} \) and both \( s_i \) and \( s_j \) are open. By this definition, for an open bond \( s_is_j \), there exist active secondary nodes \( v_i \) inside \( F_{s_i} \) and \( v_j \) inside \( F_{s_j} \) with \( \|X_i - X_j\| \leq 2\|e\| < \min(r_i,r_j) \) and some channels available for \( v_iv_j \) (\( X_i, X_j \) are the locations of \( v_i, v_j \)). That is, an open bond in \( D \) implies a link of \( G(H_{\lambda}, F_r, W(t)) \) around (see Fig. 3.4). Consequently, bond percolation of the triangular lattice \( D \) implies percolation in \( G(H_{\lambda}, F_r, W(t)) \).

According to our definition, for neighboring bonds \( s_is_j \) and \( s_js_k \), \( B(s_i, R_I + \|e\|) \cap B(s_j, R_I + \|e\|) \) and \( B(s_j, R_I + \|e\|) \cap B(s_k, R_I + \|e\|) \) are overlapping, which implies the above mapped bond percolation model on \( D \) is a dependent model. We then investigate the condition when the discrete bond percolation model on \( D \) percolates and we find that there exists a certain value \( P_c \) (see Appendix 3.6.2) such that if the bond open probability \( P_o > P_c \), the dependent bond percolation model percolates. A reverse mapping can be carried out back to the continuous plane. Then, we can finally obtain the percolation condition for \( G(H_{\lambda}, F_r, W(t)) \).

**Proof of Eq. (3.1) of Theorem 3.1.** By Thinning Theorem [94], the probability that there exist at least one active secondary nodes with transmission range \( r > 2\|e\| \) in \( F_{s_i} \) is \( 1 - e^{-\lambda[1 - F_r(2\|e\|)]A} \), where \( A = \|F_{s_i}\| \approx 0.8277\|e\|^2 \), \( \eta \) is the stationary probability of a secondary node being active and \( F_r \) is the common distribution of the transmission range \( r \). The probability that there is no primary node using the channel \( ch_k \) in \( B(s_i, R_I + \|e\|) \cap B(s_j, R_I + \|e\|) \) is \( e^{-\lambda pk\alpha} \), where \( \alpha \) denotes the area of \( B(s_i, R_I + \|e\|) \cap B(s_j, R_I + \|e\|) \) given in Appendix 3.6.1. Thus the probability that there exist available channels within \( B(s_i, R_I + \|e\|) \cap B(s_j, R_I + \|e\|) \) is \( P_s = 1 - \prod_{k=1}^{m} (1 - e^{-\lambda pk\alpha}) \). Therefore, \( P_o = \left(1 - e^{-\lambda[1 - F_r(2\|e\|)]A}\right)^2 P_s \). Thus the upper bound by \( P_o > P_c \).

Next, we study the necessary condition when \( G(H_{\lambda}, F_r, W(t)) \) percolates. There exist two typical methods in random geometric graph \( G(H_{\lambda}, 1) \) to derive the necessary condition for percolation: multi-type branching process method [94] and clustering coefficient method [101].
The former employs an argument based on comparing with a suitable branching process model to provide an upper bound on the expected number of nodes contained in a component of $G(H_\lambda, 1)$ (see p45 of [94] for more details). The latter incorporates clustering effect in the random geometric graph into the multi-type branching process method and thus can obtain tighter bounds. The clustering effect can be measured by clustering coefficient [101], which is defined as follows:

**Definition 3.5.** Given $G(H_\lambda, 1)$, distinct nodes $v_i$, $v_j$, and $v_k$, the clustering coefficient is the conditional probability that nodes $v_i$ and $v_j$ are adjacent given that $v_i$ and $v_j$ are both adjacent to node $v_k$.

Clustering coefficient can capture the location dependency among nodes in $G(H_\lambda, 1)$. However, besides this location dependency, there also exist channel dependency and transmission range dependency in cognitive radio networks (as shown in Appendix 3.6.3). By characterizing all these dependencies as Link Correlation Coefficient (LCC), the clustering coefficient method can be generalized to derive the necessary condition for percolation of $G(H_\lambda, F_r, W(t))$.

**Definition 3.6.** Given a cognitive radio network $G(H_\lambda, F_r, W(t))$ and distinct nodes $v_i$, $v_j$ and $v_k$, the Link Correlation Coefficient (LCC) $C_{LCC}$ is the conditional probability that nodes $v_i$ and $v_j$ are connected given that $v_i$ and $v_j$ are both connected to node $v_k$.

By using clustering coefficient method, the main result for percolation of $G(H_\lambda, 1)$ in [101] is the following lemma

**Lemma 3.2.** Given the mean node degree $\mu$ and clustering coefficient $C_c$ of $G(H_\lambda, 1)$. If $\mu < \frac{1}{1-C_c}$, $G(H_\lambda, 1)$ is not percolated.

Note that the proof for the above lemma in [101] does not require common channels or uniform transmission range. Hence the similar proof and result can be generalized to $G(H_\lambda, F_r, W(t))$ directly (see [101] for details). That is, given mean node degree $\mu$ and Link Correlation Coefficient (LCC) $C_{LCC}$ of $G(H_\lambda, F_r, W(t))$, $G(H_\lambda, F_r, W(t))$ is not percolated if $\mu < \frac{1}{1-C_{LCC}}$. In
fact, by *Thinning Theorem* [94], the active secondary nodes form a Poisson point process with density $\lambda \eta$. The mean degree of each secondary node is $\mu = \int_0^1 2\pi \lambda \eta \int_0^1 x (1 - F_r(x)) \mathbb{P}_s dx dF_r$.

With the *Link Correlation Coefficient (LCC)* $C_{\text{LCC}}$ in Appendix 3.6.3, we can obtain the lower bound on $\lambda_{c,c}$ given in Theorem 3.1. This completes the proof of Theorem 3.1.

**Remark 3.1.** As shown in Fig. 3.5, our results in Theorem 3.1 divide the spatial density of secondary users into three regions, if the parameters of coexisting primary networks are known. The upper region (given by Eq. (3.1)) determines the condition for percolation of cognitive radio networks and the bottom region identifies the condition for non-percolation. The middle region represents the gap between the upper and lower bounds of the critical density $\lambda_{c,c}$, whose exact value is unknown but is within this region. The perfect scenario is that there is no such gap and the mathematically rigorous value of critical density can be derived. However, due to the mathematical complexity and intractability in percolation analysis, there is no method to derive such exact value. And there are also some work, such as [79, 101], on reducing the gap. But to the best of my knowledge, our methods and results provide the tightest bounds for the critical density of cognitive radio networks.

**Remark 3.2.** Note that both the upper and lower bounds in Eq. (3.1) and Eq. (3.2) are increasing functions of $\{\lambda_{pk}\}_{k=1}^m$ and decreasing functions of $m$, which implies that $\lambda_{c,c}$ scales negatively with $m$ and positively with $\{\lambda_{pk}\}_{k=1}^m$. This agrees with our intuition: as $m$ increases or $\{\lambda_{pk}\}_{k=1}^m$ decreases, the channel accessible opportunities for secondary nodes increase and
thus there exist more links among neighboring nodes, which makes the necessary number of nodes for percolation less.

**Remark 3.3.** Eq. (3.1) and Eq. (3.2) indicate the correlation between $\lambda$, $\{\lambda_{pk}\}_{k=1}^m$, critical density of secondary nodes, and $m$. That is, for a fixed $\lambda$, these two results can be used to determine the sufficient and necessary conditions of $\{\lambda_{pk}\}_{k=1}^m$ or $m$ for a percolated cognitive radio network.

### 3.3.3 Analysis of Dissemination Radius

One of our main objectives of this study is to find how far information can be disseminated in a cognitive radio network. The problem is formulated to find the limits of dissemination radius, which is dependent on whether a cognitive radio network is percolated or not. Theorem 3.1 shows the critical phenomenon in $G(\mathcal{H}_\lambda, \mathcal{F}_r(r), \mathcal{W}(t))$ and provides upper and lower bounds on $\lambda_{c,c}$. Next, we analyze the dissemination radius and have the following results.

**Corollary 3.1.** For a cognitive radio network $G(\mathcal{H}_\lambda, \mathcal{F}_r, \mathcal{W}(t))$, given $m$ and $\lambda_p = \{\lambda_{pk}\}_{k=1}^m$, if $\lambda > \lambda_{c,w}$, a packet $b$ originated from $v_0$ can be disseminated globally with some positive probability, that is $P(\|\mathcal{L}(\infty)\| = \infty) > 0$. Further, if $\lambda > \lambda_{c,c}$, $b$ can be disseminated globally instantaneously with some positive probability.

**Proof.** When $\lambda > \lambda_{c,c}$, there exists a giant component $C_\infty(G(\mathcal{H}_\lambda, \mathcal{F}_r, \mathcal{W}(t)))$ in $G(\mathcal{H}_\lambda, \mathcal{F}_r(r), \mathcal{W}(t))$ and when a packet $b$ originated from $v_0 \in C_\infty(G(\mathcal{H}_\lambda, \mathcal{F}_r, \mathcal{W}(t)))$, all nodes in $C_\infty(G(\mathcal{H}_\lambda, \mathcal{F}_r, \mathcal{W}(t)))$ can receive this packet instantaneously. That is, $b$ can be disseminated globally instantaneously. When $\lambda_{c,w} < \lambda < \lambda_{c,c}$, there exists no giant component in $G(\mathcal{H}_\lambda, \mathcal{F}_r(r), \mathcal{W}(t))$ for any $t$ and thus $b$ cannot be disseminated globally instantaneously. However, there exists an infinite component $C_\infty(G(\mathcal{H}_\lambda, 1))$ in $G(\mathcal{H}_\lambda, 1)$. We will show in Section 3.4 that the packet originated from $v_0 \in C_\infty(G(\mathcal{H}_\lambda, 1))$ can be disseminated to $C_\infty(G(\mathcal{H}_\lambda, 1))$ at asymptotic constant speed. Thus the result. 

---

69
Corollary 3.2. When $\lambda < \lambda_{c,w}$, information can only be disseminated locally with probability 1 and

$$\mathbb{P}(L(\infty) > x) \leq \xi_1 e^{-\xi_2 x},$$

(3.3)

for some positive constants $\xi_1$ and $\xi_2$.

Proof. As discussed earlier, $G(H_\lambda, F_r, W(t))$ is a proper subset of $G(H_\lambda, 1)$. Therefore, considering $v_0$ and the component $C_{v_0} \in G(H_\lambda, 1)$ containing $v_0$, the secondary nodes which can receive the packet $b$ originated by $v_0$ are upper bounded by $C_{v_0}$. When $\lambda < \lambda_c$, $C_{v_0} \in G(H_\lambda, 1)$ is almost surely finite, that is $\mathbb{P}(\text{dia}(C_{v_0}) < \infty) = 1$. $\text{dia}(C)$ denotes the diameter of the component $C$. Thus, $\mathbb{P}(\|L(t)\| < \infty) = 1$ by $\|L(\infty)\| \leq \text{dia}(C_{v_0})$. This implies that the packet $b$ originated by $v_0$ can only be disseminated locally with probability 1. Further, as shown in [94] (see Lemma 3.3, page 68), when $G(H_\lambda, 1)$ is not percolated, for any component $C \in G(H_\lambda, 1)$, $\mathbb{P}(\text{dia}(C) > x) < \xi_1 e^{-\xi_2 x}$ for some positive constants $\xi_1$ and $\xi_2$. Thus Eq. (3.3) considering $\|L(\infty)\| < d(C_{v_0})$.

3.3.4 Discussion and Applications

In this section, we proved that there exist critical values on the spatial density of cognitive radio users, above which the limiting dissemination distance is infinite with some positive probability, i.e., $\mathbb{P}(L(\infty) = \infty) > 0$, and below which $L(\infty)$ is exponentially bounded. We further
mathematically derived the bounds of these critical values. Besides the theoretical importance of our findings, our results can be used practically not only in the initial deployment, but only dynamic reconfiguration of cognitive radio networks. Here are some examples.

- In the initial deployment, an appropriate value for the spatial density of cognitive radio nodes can be decided to guarantee that information from secondary users can be disseminated to the entire network, if the parameters of the overlaid primary networks are known. Particularly, Eq. (3.1) and Eq. (3.2) in Theorem 3.1 provide critical value required for global dissemination, in terms of primary networks' parameters. For example, consider a scenario to deploy an emergency network in the aftermath of traffic accidents, coexisting with a primary network with density 0.01 per $m^2$ and interference range $R_I = 1.2m$. If we assume that each secondary user transmit with fixed transmission range $r = 1$, Eq. (3.1) shows $\lambda_{c,c} \leq \frac{3.58}{\eta}$ (given $\left\| e \right\| = 0.2$). This implies that, by Corollary 3.1, if we set $\lambda > \frac{3.58}{\eta}$ (see Fig. 3.6), “emergency” or “help” information can be disseminated to the entire emergency network within finite time.

- In the real applications of secondary networks, such as emergency networks and military networks, secondary nodes are vulnerable to neighboring wireless radio environment (coexisting primary networks), which affect the communication connectivity and in turn impair the performance (coverage, e.g.) of secondary networks. Our results provide network designers a guideline on reconfiguring the secondary networks to ensure information being disseminated to the entire networks, based on the variance of the radio environment.

**Remark 3.4.** Our results provide theoretical insight into the dissemination distance of general cognitive radio networks. However, our findings can also be practically used in real applications of cognitive radio networks. Our theoretical results seem to be complicated, it can be tremendously simplified for practical applications (see the example emergency network mentioned above). Besides identifying correlation among network parameters as discussed in remark 3.2 and 3.3, our results can also be used to evaluate the exact values of network parameters required to ensure
Figure 3.7: Percolation regions under different interference range $R_I$.

Figure 3.8: An example of percolation conditions in terms of $\lambda$, $m$ and $\{\lambda_{pk}\}_{k=1}^m$.

**3.3.5 Simulation Results of Dissemination Radius**

We have implemented sever sets of simulations to further interpret our theoretical results on dissemination radius.

In the simulation, we distribute $n$ CR nodes independently and randomly with a uniform distribution to approximate a Poisson point process. Primary nodes using $ch_k$ are initially independently and uniformly distributed with a density $\lambda_{pk}^{int}$. For simplicity, we assume that PR nodes using different channels have the same initial density $\lambda_p^{int}$. That is, $\lambda_{pk}^{int}$ is independent of $k$. We further assume that at each time slot, each primary user independently uses the assigned channel with some constant $\varrho$. Let $\lambda_p = \lambda_p^{int} \times \varrho$. Although at any time snapshot, this
deployment is equivalent to deploying always-transmitting primary nodes with density $\lambda_p$ by thinning theorem [94], the on/off activities model the spatial and temporal dynamics of primary nodes. We consider scenarios where the interference range of PR nodes $R_I = 80m$, $120m$ and $200m$. The transmission range of each secondary node is randomly generated according to the same distribution $\text{Uniform}(60, 100)$. The communicating (active) and sensing (inactive) time of each CR node are independently generated from $\text{Uniform}(0, T_{on})$ and $\text{Uniform}(0, T_{off})$ respectively, for some constants $T_{on}$ and $T_{off}$.

Given $T_{on} = 0.9s$, $T_{off} = 0.1s$, and $\eta = \frac{T_{on}}{T_{on} + T_{off}} = 0.9$, the values of $\lambda$, $\lambda_p$, $m$ required for percolation under different $R_I$ are demonstrated in Fig. 3.7. Particularly, as shown in Fig. 3.7, there exists some bound on $\lambda_{pk}$ for any $m$ and $R_I$, above which $G(H_\lambda, F_r, W(t))$ cannot percolate no matter how large $\lambda$ is. This is because when the density of active primary nodes is beyond some certain value, there lack enough spectrum opportunities for secondary nodes to percolate. Furthermore, Fig. 3.7 shows that the percolation region will increase as $m$ increases and $R_I$ decreases. This corresponds to our intuition that as $m$ increases and $R_I$ decreases, the spectrum opportunity available for secondary users will be enlarged and thus the critical density $\lambda_{c,c}$ of secondary nodes required for percolation decreases. Fig. 3.7 also shows the critical value $\lambda_{c,c}$ is an increasing function of $\lambda_{pk}$ and decreasing function of $m$. The simulation results in Fig. 3.7 validate our analytical results in Theorem. 3.1.

Fig. 3.8 demonstrates the inverse impact of primary nodes activities on critical density $\lambda_{c,c}$.
of secondary nodes given $R_I = 120m$ and further validate our theoretical results in Theorem 3.1. Particularly, Fig. 3.8a and Fig. 3.8b simulate a network scenario with secondary node density $\lambda = 600$ (per $km^2$), $\eta = 0.9$ and $m = 2$. When no primary users are active ($\lambda_{pk} = 0$), the secondary network is percolated as shown in Fig. 3.8a. As more and more primary users transmit over their licensed channels, the secondary network will transit from percolation to non-percolation. Fig. 3.8b shows an example that when a enough large number of primary nodes ($\lambda_{pk} = 2$) are active, the secondary network is not percolated. Fig. 3.8a and Fig. 3.8c provide two examples to verify the analytical upper bound on $\lambda_{c,c}$ of Theorem 3.1. When $\lambda_{pk} = 0$, the dependent percolation model on triangular lattice in Appendix 3.6.2 reduces to 1-dependent model and thus $P_c = 3/4$. Thus the theoretical bound by Eq. (3.1) in Theorem 3.1 is $\lambda_{c,c} \leq 2107$ (per $km^2$). Fig. 3.8a shows that $G(\mathcal{H}_\lambda, F_r, W(t))$ percolates for $\lambda = 600$, which implies that our theoretical bound is correct. Given $m = 230$ and $\lambda_{pk} = 5$ (per $km^2$), Fig. 3.8c shows that $G(\mathcal{H}_\lambda, f_r, \mathbb{L}_{\lambda_{pk}, m}, W(t))$ percolates when $\lambda = 5000$, which is much smaller than the theoretical value 76846.

We also study how packets disseminates in a giant component. Fig. 3.9 simulates a scenario that when $\lambda_{c,w} < \lambda < \lambda_{c,c}$, how a packet $b$ originated from the origin node $v_0$ disseminates among the giant component, where the bigger dots connected by the solid line denote the cluster of nodes that have received $b$ at time $t$. Specifically, given $T_{on} = 0.8s$, $T_{off} = 0.1s$ and $\lambda = 230$ (per $km^2$), we find that $G(\mathcal{H}_\lambda, 1)$ is percolated but $G(\mathcal{H}_\lambda, F_r, W(t))$ is not percolated by simulations. As expected, in this case, only a very small set of secondary nodes will receive $b$, as shown in Fig. 3.9a. However, $b$ will be disseminated to more and more secondary nodes gradually, as shown in Fig. 3.9b and Fig. 3.9c. These results justify our theoretical analysis in Corollary 3.1 that when $G(\mathcal{H}_\lambda, 1)$ is percolated ($\lambda > \lambda_{c,w}$), $b$ can be disseminated globally. And simulation results in Fig. 3.9 also validate our theoretical analysis of propagation speed in Theorem 3.2 in the following section.
3.4 How Fast Can Information Be Propagated in a Cognitive Radio Network?

In this section, we investigate how fast information can disseminate in percolated and non-percolated cognitive radio networks. Particularly, we consider information propagation speed under the following three scenarios: first, \( \lambda_{c,w} < \lambda < \lambda_{c,c} \) (percolated) and \( v_0 \in C_\infty(G(H_\lambda,1)) \); second, \( \lambda > \lambda_{c,c} \) (percolated) and \( v_0 \in C_\infty(G(H_\lambda,1))\setminus C_\infty(G(H_\lambda,F_r,W(0))) \); and finally \( \lambda < \lambda_{c,w} \) (non-percolated). We have the following main results.

**Theorem 3.2.** (i) When \( \lambda_{c,w} < \lambda < \lambda_{c,c} \) and \( v_0 \in C_\infty(G(H_\lambda,1)) \), information \( b \) originated by \( v_0 \) can be disseminated gradually at some constant speed \( S_\varphi(d) = \kappa \), for sufficiently large \( d \). (ii) When \( \lambda > \lambda_{c,c} \) and \( v_0 \in C_\infty(G(H_\lambda,1))\setminus C_\infty(G(H_\lambda,F_r,W(0))) \), \( b \) can be disseminated globally within finite time with probability 1. (iii) When \( \lambda < \lambda_{c,w} \), \( b \) can only be disseminated locally with probability 1 and the limiting speed \( S_\varphi(\infty) = 0 \).

3.4.1 Differences from Earlier Studies

We first consider the information propagation speed when \( \lambda_{c,w} < \lambda < \lambda_{c,c} \) and \( v_0 \in C_\infty(G(H_\lambda,1)) \). This problem is similar to the first passage percolation problem [99]. Related problems have been studied in [57, 58]. In [57, 58], the authors consider the topology-incurred delay in wireless sensor networks where each sensor has independent or degree-dependent dynamic behavior, which can be modeled by a dynamic site percolation on random geometric graph \( G(H_\lambda,1) \). Particularly, by denoting \( T_{i,j} \) as the time until \( v_j \) receives \( b \) directly from \( v_i \) after \( v_i \) has received the packet \( b \) for each link \( v_iv_j \in G(H_\lambda,1) \), the authors coupled topology-incurred delay between any \( u,v \in C_\infty(G(H_\lambda,1)) \) to the first passage time in the weighted graph and expressed \( T(u,u) = \inf_{l(u,v)}\{\sum_{v_i,v_j \in l(u,v)} T_{i,j}\} \), where \( l(u,v) \) is an arbitrary path from the \( u \) to \( v \). By using subadditive ergodic theorem [102], the main result in [57, 58] is:

**Lemma 3.3.** \( T(u,u) \) asymptotically scales with the Euclidean distance \( d(u,v) \) between \( u,v \),
that is,
\[ \mathbb{P}\left( \lim_{d(u,v) \to \infty} \frac{d(u,v)}{T(u,u)} = \delta \right) = 1, \]
for some constant \( 0 < \delta < \infty \).

A fundamental assumption in [57, 58] is that when a node receives a packet, it stays active and broadcasts this packet until it is sure that the packet has reached all its neighbors. Thus \( T_{i,j} \) is assumed to be a constant or a light-tailed random variable with finite expected value. However, this assumption is not valid in cognitive radio networks. In cognitive radio networks, each secondary node must frequently alternates between communicating and sensing states. For example, in IEEE 802.22, each secondary node needs to sense the primary nodes once every 2 seconds. Furthermore, besides the Euclidean distance, communication links between neighboring secondary nodes also dynamically depend on opportunistic spectrum opportunity and random transmission ranges of secondary nodes. All these make information dissemination in cognitive radio networks much slower than that in wireless sensor networks. Thus the existing results on topology-incurred delay in wireless sensor networks may not be applied to our cognitive radio networks directly. However, we find that the proofs for the results in [57, 58] only require \( E(T_{i,j}) < \infty \). Therefore, if we can show that the information propagation delay \( T_{i,j} \) between neighboring nodes \( v_i \) and \( v_j \) has finite expected value in cognitive radio networks, the results about asymptotic propagation speed in [57, 58], i.e., Lemma 3.3, also hold in our cognitive radio networks.

**Lemma 3.4.** For each \( v_i,v_j \in G(\mathcal{H}_\lambda,1) \), let \( T_{i,j} \) be a random variable associated with \( v_i,v_j \) denoting the time needed by \( v_j \) to receive the packet from \( v_i \) directly after \( v_i \) has received the packet originated from the origin node \( v_o \). We have \( \mathbb{P}(T_{i,j} < \infty) = 1 \) and \( E(T_{i,j}) < \infty \).

**Proof.** Assume that \( v_i \) receives the packet at time 0 for simplicity. Thus, \( T_{i,j} = 0 \) when node \( v_j \) is active, \( \|X_i - X_j\| < \min\{r_i,r_j\} \) (recall that \( X_i,X_j \) are locations of \( v_i,v_j \) and there exist some channels available to \( v_i,v_j \). Denote \( Q = \min\{r_i,r_j\} \) and \( Y = \|X_i - X_j\| \). It is easy to show that their distributions \( \mathbb{P}(Q < q) = 1 - (1 - F_r(q))^2 \) for \( q \leq 1 \) and \( \mathbb{P}(Y \leq y) = \frac{\pi q^2}{\pi} = y^2 \) for
Thus $P(Y < Q) = \int_0^1 \int_0^Q dP(Y < y)dP(Q < q)$ and $P_d = P(T_{i,j} = 0) = \eta P(Y < X)\eta_s$, where $\eta$ and $\eta_s$ are defined in Section 3.2.1 as the probabilities of nodes being active and channels available for $v_i v_j$. Now consider $P(T_{i,j} > t)$. Since the sensing period is bounded by some constant $T_p$ (required by FCC to limit the interference with primary nodes), during time $t$, node $v_i$ will be active at least $t T_p$ times (ignoring the fraction factor here for simplicity).

Once $v_i$ turns active, the probability that the packet can be successfully delivered from $v_i$ to $v_j$ is at least $P_d$. Hence, $P(T_{i,j} > t) \leq (1 - P_d)^t T_p < e^{-t P_d T_p}$. Therefore, $P(T_{i,j} < \infty) = 1$ and $E(T_{i,j}) = \int_0^\infty P(T_{i,j} > t)dt < \infty$.

### 3.4.2 Asymptotic Information Propagation Speed

Based on the discussion and Lemma 3.4 in Section 3.4.1, we have seen that the proof techniques for asymptotic propagation speed in [57, 58] can also be used in this chapter. Particularly, by denoting $T_{i,j}$ as the topology-incurred between neighboring secondary nodes $v_i$ and $v_j$, the topology-incurred delay between any secondary nodes $v$ and $u$ defined in Section 3.2 can be coupled to the first passage time in the weighted graph and reexpressed as $T(v, u) \triangleq \inf_{l(v, u)} \{\sum_{(v, v_j) \in l(v, u)} T_{i,j}\}$, where $l(v, u)$ is an arbitrary path from the $v$ to $u$.

We consider the information propagation in a particular direction, say $x$-axis without loss of generality. For any point $r_0(d) = (d, 0) \in \mathbb{R}^2$, since the node distribution is continuous, there exists no node on $r_0(d)$ with probability 1. Thus as described in Section 3.2.3, we consider node $\tilde{v}_0(d) \in C_\infty(G(H_\lambda, 1))$ nearest to $r_0(d)$. In order to use subadditive ergodic theorem, we only consider discrete $d = nx$ for integers $n$ and some constant $x$. Since the discrete limit can be replaced by a continuous one by [100], the results derived here also applied for continuous $d$ when $d$ is sufficiently large. Let $\tilde{T}(mx, nx) \triangleq T(\tilde{v}_0(mx), \tilde{v}_0(nx))$ and define the collection of indexed variables $\{Z_{m,n} \triangleq \tilde{T}(mx, nx), m, n \in \mathbb{Z}^+\}$, for some constant $x > 0$. Using Liggett’s subadditive ergodic theorem, we can prove the part(i) of Theorem 3.2.

**Theorem 3.3.** [102], Liggett’s subadditive ergodic theorem] Let $\{Z_{m,n}\}$ be a collection of random variables indexed by integers satisfying $0 \leq m < n$. Suppose $\{Z_{m,n}\}$ has the following

\[\]
Proof. See Proposition 4 in [57].

Next, we show that \( \{Z_{m,n} : m, n \in \mathbb{Z}^+\} \) satisfy all the conditions in Theorem 3.3. By definition, \( Z_{0,n} \) is the first passage time from the origin \( v_0 \) to \( \tilde{v}_0(nx) \), which is clearly at most the first passage time from \( v_0 \) to \( \tilde{v}_0(mx) \) \( Z_{0,m} \) plus \( \tilde{v}_0(mx) \) to \( \tilde{v}_0(nx) \) \( Z_{m,n} \). Condition (i) is thus verified.

We then show that \( \mathbb{E}(Z_{m,n}) \) is bounded and nonnegative for any \( m, n \). As the first passage time cannot be negative, we have \( \mathbb{E}(Z_{m,n}) \geq 0 \). To show \( \mathbb{E}(Z_{m,n}) \) is finite, note that by Lemma 3.4, the average transmission delay at each hop \( \mathbb{E}(T_{i,j}) < \infty \), and thus we only need to find a path from \( \tilde{v}_0(mx) \) to \( \tilde{v}_0(nx) \) in \( C_{\infty}(G(H_\lambda, 1)) \) with finite hops.

**Lemma 3.5.** Let \( H_{m,n} \) be the number of hops in the shortest path between \( \tilde{v}_0(mx) \) and \( \tilde{v}_0(nx) \). Then \( \mathbb{E}(H_{m,n}) \) is finite.

*Proof.* See Proposition 4 in [57].

Since \( Z_{m,n} \) is defined in a stationary way, conditions (iii) and (iv) are clearly verified. Next, we show that \( Z_{m,n} \) is also ergodic, which is implied by mixing (see [100]).

**Lemma 3.6.** The sequence \( \{Z_{n,n+1}, n \geq 0\} \) is ergodic.

*Proof.* We only need to show \( \{Z_{n,n+1}, n \geq 0\} \) is mixing. Consider \( Z_{k,k+1} \) and \( Z_{3k,3k+1} \). We start by constructing two squares \( B_k \) and \( B'_k \) with edge \( kx \) and centering at \( (kx, 0) \) and \( (3kx, 0) \) respectively. From Theorem 3.5, we know that the paths from \( \tilde{v}_0(kx) \) to \( \tilde{v}_0(kx + x) \) and from \( \tilde{v}_0(3kx) \) to \( \tilde{v}_0(3kx + x) \) are included in \( B_k \) and \( B'_k \) respectively with high probability as \( k \to \infty \). Since \( B_k \) and \( B'_k \) do not overlap, the fastest path from \( \tilde{v}_0(kx) \) to \( \tilde{v}_0(kx + x) \) and from \( \tilde{v}_0(3kx) \) to \( \tilde{v}_0(3kx + x) \) are asymptotically independent, which means that \( \{Z_{n,n+1}, n \geq 0\} \) is mixing and thus ergodic. \( \square \)
Now we have seen that the topology-incurred delay $Z_{m,n}$ satisfies all the conditions of Theorem 3.3 and thus we have $\lim_{n \to \infty} \frac{Z_{0,n}}{n} = \zeta = \inf_{n \geq 1} \mathbb{E}(Z_{m,n})/n$. Next, we need to show that $0 < \zeta < \infty$. We have shown that $\forall m, n$, $\mathbb{E}(Z_{m,n}) < \infty$. Thus $\zeta = \inf_{n \geq 1} \mathbb{E}(Z_{m,n})/n \leq \mathbb{E}(Z_{0,1}) < \infty$. Note that $G(\mathcal{H}, f_r, \mathbb{L}_{\lambda_p,m}, W(t))$ is a subgraph of the dynamic site percolation model $G(\mathcal{H}, 1, W(t))$ defined in [57]. Thus $Z_{m,n}$ is lower bounded by the topology-incurred delay from $\tilde{v}_0(mx)$ to $\tilde{v}_0(nx)$ in $G(\mathcal{H}, 1, W(t))$, which has been proved positive in [57]. Thus $\zeta > 0$.

Proof of Part (i) of Theorem 3.2. By the definition of information propagation speed in Section 3.2.3, we have

$$S_0(\infty) = \lim_{d \to \infty} S_0(d) = \lim_{d \to \infty} \frac{d}{T(v_0, \tilde{v}_\varphi(d))} = \lim_{n \to \infty} \frac{nx}{T(v_0, \tilde{v}_0(nx))} = \lim_{n \to \infty} \frac{x}{\frac{Z_{0,n}}{n}} = \frac{x}{\zeta}.$$ 

Note that our proof in this section is independent of direction $\varphi$ and thus the result by letting $\kappa = \frac{x}{\zeta}$.

Next, consider the case when $\lambda > \lambda_{c,c}$ and $v_0 \in \mathcal{C}_\infty(G(\mathcal{H}, 1)) \setminus \mathcal{C}_\infty(G(\mathcal{H}, F_r, W(0)))$. Let $\psi$ be the first time that some node of $\mathcal{C}_\infty(G(\mathcal{H}, F_r, W(\psi)))$ receives this packet. Note that at $t = \psi$, the packet can be disseminated to all nodes of $\mathcal{C}_\infty(G(\mathcal{H}, F_r, W(\psi)))$ instantaneously. That is, $b$ can be disseminated globally within time $\psi$. Therefore, to prove part (ii) of Theorem 3.2, we only need to show that $\psi$ is finite with probability 1.

Proof of Part (ii) of Theorem 3.2. Let $\{Z_k\}_{k \geq 1}$ denote the i.i.d random variables which have the same distribution with $T_{i,j}$. Thus along the infinite path in $G(\mathcal{H}, 1)$, $b$ can be spread to at least $N = \sup_n \{\sum_{k=1}^n X_k < t\}$ nodes within time $t$. And at any time point $t'$, each $v \in \mathcal{C}_\infty(G(\mathcal{H}, 1))$ has the same probability $\theta > 0$ belonging to $\mathcal{C}_\infty(G(\mathcal{H}, F_r, W(t')))$. Thus we have $\mathbb{P}(\psi > t) < (1 - \theta)^N$. Considering $Z_k$ is bounded with probability 1, $N$ approaches to infinity as $t$ increases. Therefore, $\psi$ is finite with probability 1 and thus the result.
Finally we focus on $\lambda < \lambda_c$.

**Proof of Part (iii) of Theorem 3.2.** When $\lambda < \lambda_c$, since the packet cannot be disseminated to the node farther than $\|L(\infty)\|$ and $S_\varphi(d)$ is defined as 0 when $d > \|L(\infty)\|$, $S_\varphi(\infty) = \lim_{d \to \infty}(S_\varphi(d)) = 0$ almost surely considering $P(\|L(\infty)\| < \infty) = 1$. This completes the proof.

### 3.4.3 Discussions and Applications

![Diagram of dissemination radius and propagation speed analysis](image)

(a) Theoretical critical densities.  
(b) Mathematically achievable critical densities.

Figure 3.10: Dissemination radius and propagation speed analysis.

In this section, we provide theoretical insights into information dissemination speed in cognitive radio networks. Particularly, as shown in Fig. 3.10a, Theorem 3.2 identifies three regions, $[0, \lambda_{c,w}]$, $[\lambda_{c,w}, \lambda_{c,c}]$, and $[\lambda_{c,c}, \infty]$, on spatial density of secondary nodes $\lambda$. Theorem 3.2 proves that in the first region, the secondary networks are not percolated and the limiting propagation speed is zero; in the second region, although the secondary networks are still not percolated, information can be disseminated *globally* at some constant speed; and in the last region, information can be disseminated to the entire secondary networks within finite time. If
we know the exact value of $\lambda_{c,w}$ and $\lambda_{c,c}$, then we obtain perfect answers to the propagation speed problems in cognitive radio networks. However, due to the mathematical complexity in percolation theory, there are no such exact values.

Instead of exact values, we mathematically derive theoretical bounds on $\lambda_{c,c}$ in Theorem 3.1. Specifically, by using our results in Theorem 3.1, we mathematically determine there regions on $\lambda$ for the above mentioned three different network performances (see Fig. 3.10b), given primary network parameters. However, as shown in Fig. 3.10b, our results leave two gaps. And when $\lambda$ falls into these two gaps, the performance of secondary networks can be theoretically determined by our results. Actually, there are a lot of efforts on diminishing these gaps in percolation theory (see [79, 101]). But to the best of my knowledge, our approach provides the tightest results on critical density $\lambda_{c,c}$ for percolation in cognitive radio networks.

Besides the theoretical importance mentioned above, our results can also be used practically in real applications of cognitive radio networks.

- In emergency networks or military networks, the network designers need to ensure that emergency information can be disseminated to the entire network within finite time. As mentioned in the example in Section 3.3.3, although our theoretical results seem to be complicated, they can be much simplified in practical use. Therefore, given the information of the coexisting primary networks, our theoretical results can provide a quick estimate on $\lambda$ required for fast propagation (to ensure $\lambda$ to fall into the upper region in Fig. 3.10b)

### 3.4.4 Simulation Results on Information Propagation Speed

Similarly, we provide some simulation results to further explain and validate our theoretical results (Theorem 3.2) on propagation speed. The scenario that information is disseminated gradually among secondary nodes (part(i) of Theorem 3.2) when $\lambda_{c,w} < \lambda$ has been studied in Fig. 3.9 in Section 3.3. In this section, we study the rest two scenarios. Particularly, Fig. 3.11 studies a scenario $\lambda < \lambda_{c,w}$ and information can only be disseminated locally. As Fig. 3.11 demonstrates, in this scenario, even after enough long time ($t = 100s$), information is still be
disseminated around a small region. This justify Part (iii) of Theorem 3.2. Fig. 3.12 shows a scenario when $\lambda > \lambda_{c,c}$. In accordance with part (ii) in Theorem 3.2, information has been disseminated to the entire network within finite time ($t = 1s$, as shown in Fig. 3.12c).

To compare information dissemination in cognitive radio networks and in wireless sensor networks, we also consider an ideal dissemination strategy, which has been well studied in [57,58], by assuming that once a CR node receives $b$, it stays active and keeps sending $b$ using maximum transmission power, until all nodes within its maximum transmission range receive $b$. Comparing with Fig. 3.9c, Fig. 3.13a shows that, with ideal dissemination strategy, the information disseminates at least two times faster. This coincide with our analysis in Section 3.4.1.

The above simulation results validate our theoretical analysis that the spatial density $\lambda$ of
secondary nodes has been partitioned into three regions, and the macroscopic performance of cognitive radio networks are totally different in different regions. We next demonstrate that when $\lambda_{cw} < \lambda < \lambda_{cc}$, information disseminates gradually to the entire network at a constant speed. Particularly, the average dissemination radius based on 100 independent simulations with parameters $\lambda = 230$ (per km$^2$), $m = 10$ and $\lambda_{pk} = 10$ (per km$^2$) is shown in Fig. 3.13b. We find that the dissemination radius using ideal dissemination strategy almost scales linearly with time. This finding agrees with the results in [57, 58]. An interesting observation is that the dissemination radius without using ideal dissemination strategy still increases linearly with $t$ for any $\eta$, although the speed is much slower. We also note that the dissemination radius is an increasing function of $\eta$. This is natural since larger $\eta$ means more nodes in communicating state, which can help relay the packet farther. 5 independent simulations of propagation speed $S(d)$ with parameters $\lambda = 230$ (per km$^2$), $m = 10$ and $\lambda_{pk} = 10$ (per km$^2$) are shown in Fig. 3.13c. We observe that although the practical information dissemination in CR networks is much slower than that in wireless sensor networks with ideal dissemination strategy, the propagation speed $S(d)$ is still some constant for large transmission distance $d$. This validates our theoretical results in Theorem 3.2. And similar to Fig. 3.13b, we also observe that $S(d)$ can be increased by increasing $\eta$. 

![Figure 3.13: Dissemination radius and propagation speed analysis.](image)
In this Chapter we have studied the information dissemination radius and propagation speed to understand how far and how fast messages are transmitted in large multi-channel cognitive radio networks. We find that how far and how fast information can disseminate depend on the spatial density \( \lambda \) of secondary nodes. Particularly, we identify three intervals \([0, \lambda_{c,w}], [\lambda_{c,w}, \lambda_{c,c}]\) and \([\lambda_{c,c}, \infty]\) for \( \lambda \). In the first region, information can only be disseminated locally, information can be disseminated gradually to the entire network at some constant speed when \( \lambda \) falls into the second region, and information can be disseminated to the entire network within finite time when \( \lambda \) is in the last region. We further determine these intervals by identifying the bounds of \( \lambda_{c,w} \) and \( \lambda_{c,c} \), which depend on the environmental radio parameters, such as the densities of coexisting primary networks and spectrum opportunities. The work in this chapter provides fundamental understanding of how information propagation in large cognitive radio networks provides guideline on network deployment and maintenance to achieve desired performance.

It worth pointing out that there exist gaps between the upper bounds and lower bounds we achieved on \( \lambda_{c,w} \) and \( \lambda_{c,c} \) because of the mathematical complexity of percolation. Although there exist a lot of work on diminishing such gaps, to the best of my knowledge, our results provide the tightest bounds. Our results are validated through extensive simulations.
3.6 Appendix

3.6.1 The Area Of $S(d, R_I)$

From the elementary geometry, we have

$$\|S(d, R_I)\| = 2\pi R_I^2 - 2 \arccos\left(\frac{d}{2R_I}\right) R_I^2 + R_I^2 \sin(2 \arccos \frac{d}{2R_I}).$$

3.6.2 A Sufficient Bond Open Probability $\mathbb{P}_c$ for Dependent Percolation Models on The Triangular Lattice

Fig. 3.15 shows a triangular lattice $\mathcal{D}$ and its dual hexagonal lattice $\mathcal{D}_d$. We designate a dual edge $e_d \in \mathcal{D}_d$ open if and only if the corresponding edge $e$ of $\mathcal{D}$ is open. The derivation of the sufficient condition for percolation on $\mathcal{D}$ is based on the observation that the origin of $\mathcal{D}$ belongs to an infinite open cluster if and only if it lies in the interior of no closed circuit of $\mathcal{D}_d$ (see [99,103] for details). Let $\rho(\iota)$ be the number of circuits of length $\iota$ of $\mathcal{D}_d$ and $\rho(\iota) \leq 3 \cdot 2^{\iota - 2}$, which follows from the fact that any circuit of length $\iota$ $C_\iota$ surrounding the origin contains a path of length $\iota - 1$ starting from some point $(k \frac{\sqrt{3}}{3} + \frac{\sqrt{3}}{6}, 1/2)$ for $0 \leq k < \iota$. Let $q$ be the probability of edge being closed and $\Phi(\iota)$ be the probability that $C_\iota$ is closed. Then $\sum_{m} \mathbb{P}(C_i \text{is closed}) \leq \sum_{i=1}^{\infty} \rho(\iota) \Phi(\iota)$. 

85
For independent bond models, \( \Phi(\iota) = q' \).

For our model, the bonds within a circle with radius \( 2(R_I + \|e\|) \) are dependent by definition. Note that the area of each hexagon is \( \sqrt{3} \|e\|^2 \). Thus ignoring the edge effect, each circle contains at most \( \Upsilon = \left\lfloor \frac{4\pi(R_I + \|e\|)^2}{\sqrt{3} \|e\|^2} \right\rfloor \) hexagons and thus approximately \( 3\Upsilon \) bonds. Thus, if we partition \( C_i \) with such circles, there exist at least \( \theta = \left\lfloor \frac{\iota}{3\Upsilon} \right\rfloor \) independent bonds and thus \( \Phi(\iota) < q' \).

Thus \( \sum_m P(C_i \text{is closed}) \) converges if \( q < \frac{1}{2} \frac{1}{3\Upsilon} \), which implies that \( P_c = 1 - \frac{1}{2} \frac{1}{3\Upsilon} \) is enough for percolation.

### 3.6.3 \( \text{C}_{\text{LCC}} \) of \( G(\mathcal{H}_\lambda, f_r, \mathbb{L}_{\lambda_p, m}, W(t)) \)

To determine the Link Correlation Coefficient \( \text{C}_{\text{LCC}} \) of \( G(\mathcal{H}_\lambda, f_r, \mathbb{L}_{\lambda_p, m}, W(t)) \), which is the conditional probability that node \( v_j \) and \( v_k \) are connected, given they are both connected to a common node \( v_i \), we need to consider three possible cases: \( r_i \leq r_j \leq r_k \), \( r_j \leq r_i \leq r_k \) and \( r_j \leq r_k \leq r_i \), where \( r_i, r_j \) and \( r_k \) are the transmission ranges of \( v_i, v_j \) and \( v_k \) respectively (see Fig. 3.14). Denote \( h = \|v_i - v_j\| \) as the Euclidean distance between \( v_i \) and \( v_j \) and \( B_{m,n} \) as the circle centering at node \( v_m \) with radius \( r_n \) for \( m, n = i, j, k \) and \( \|B_{m,n}\| \) denote its area. Denote \( P_{sd} \) as the conditional probability that there exist channels available for \( v_jv_k \) given that there exist channels available for \( v_iv_j \) and \( v_iv_k \). Note that to calculate the exact value of \( P_{sd} \) is very complicated and make \( \text{C}_{\text{LCC}} \) intractable. Instead of the exact value of \( \text{C}_{\text{LCC}} \), we only need a lower bound of \( \text{C}_{\text{LCC}} \) to determine the lower bound of \( \lambda_c(\mathbb{L}_{\lambda_p, m}) \) in Eq. (3.2). Thus we use \( P_s \) to approximate \( P_{sd} \). Also note that when the spatial density of PR nodes is low, which is usually true in CR networks, \( P_s \) and \( P_{sd} \) are very close, which further verifies our approximation.

For Case I (see Fig. 3.14a), the conditional probability is \( P_s = \frac{\|B_{i,j}\|}{\|B_{i,i}\|} \). Thus if \( r_j > r_i + h \), the conditional probability is \( P_s \). If \( r_j < r_i + h \), this probability is \( b_1(h) = P_s \pi (r_i - h)^2 + \int_{r_i-h}^{r_j} 2\theta_1 x dx \pi r_i^2 \), where \( \theta_1 = \angle v_i v_j E = \cos^{-1}(\frac{x^2 + h^2 - r_i^2}{2xh}) \). Let \( f_{i,j,k} = f_r(r_i)f_r(r_j)f_r(r_k) \) be
the joint distribution of $r_i$, $r_j$ and $r_k$. Thus the conditional probability $\mathbb{P}_1$ for Case I is:

$$
\mathbb{P}_1 = \int_0^{1/2} \int_0^{1-h} \int_{r_i+h}^1 \int_{r_j}^1 2\mathbb{P}_s(h)f_{i,j,k}dr_kdr_jdr_i dh
+ \int_0^{1} \int_{h}^{1} \int_{r_i}^{1+h} \int_{r_j}^{1} 2\mathbb{P}_b(h)f_{i,j,k}dr_kdr_jdr_i dh.
$$

For Case II (see Fig. 3.14b), the conditional probability is $\mathbb{P}_s h_{ij}B_{i,i} \cap B_{j,j}$. Thus if $r_i > r_j + h$, this probability is $\mathbb{P}_s r_i^2$. If $r_i < r_j + h$, this probability is $b_2(h) = \mathbb{P}_s \frac{\pi(r_i-h)^2+\int_{r_j}^{r_i} \frac{2\theta_2 x dx}{\pi r_i^2}}{r_i^2}$, with $\theta_2 = \angle v_i v_j F = \cos^{-1}(\frac{x^2+h^2-r_i^2}{2xh})$. Hence for Case II, the conditional probability $\mathbb{P}_2$ is

$$
\mathbb{P}_2 = \int_0^{1/2} \int_0^{1} \int_{r_i+h}^{1} \int_{r_j}^{1} 2h\mathbb{P}_s \frac{r_i^2}{h}f_{i,j,k}dr_kdr_jdr_i dh
+ \int_0^{1} \int_{h}^{1} \int_{r_i}^{1+h} \int_{r_j}^{1} 2hb_2(h)f_{i,j,k}dr_kdr_jdr_i dh.
$$

For Case III, the conditional probability is $\mathbb{P}_s \frac{|B_{j,j} \cap B_{i,k}|}{|B_{j,j}||B_{i,k}|}$. Thus if $r_k > r_j + h$, this probability is $\mathbb{P}_s r_k^2$. If $r_k < r_j + h$, this probability is $b_3(h) = \mathbb{P}_s \frac{\pi(r_k-h)^2+\int_{r_j}^{r_k} \frac{2\theta_3 x dx}{\pi r_k^2}}{r_k^2}$, with $\theta_3 = \angle v_i v_j G = \cos^{-1}(\frac{x^2+h^2-r_k^2}{2xh})$. Hence for Case II, the conditional probability $\mathbb{P}_3$ is

$$
\mathbb{P}_3 = \int_0^{1/2} \int_0^{1} \int_{r_i+h}^{1} \int_{r_k}^{1} 2h\mathbb{P}_s \frac{r_k^2}{h}f_{i,j,k}dr_i dr_k dr_j dh
+ \int_0^{1} \int_{h}^{1} \int_{r_k}^{1+h} \int_{r_j}^{1} 2hb_3(h)f_{i,j,k}dr_i dr_k dr_j dh.
$$

Finally, by the symmetry of $v_j$ and $v_k$, we have $C_{LCC} \approx 2(\mathbb{P}_1 + \mathbb{P}_2 + \mathbb{P}_3)$. 

87
Chapter 4

The Distribution and Scaling Law of Dissemination Latency in Mobile Cognitive Radio Networks

A common assumption in most of the performance studies on dissemination latency of wireless networks in the literature [57, 58, 61, 62, 71–74] is that the number of nodes is infinite or approaches to infinity. It is evident that the asymptotic results, though, provide good insights into network performance, may not explain the latency properties when the number of nodes in real applications is finite. In other words, the stochastic properties of latency distribution in finite networks lead to understanding of real networks, rather than the large networks. In addition, although the latency of mobile wireless networks has been extensively studied under various mobility models [64–70], the fundamental impact of general mobility on latency is not clear. Thus in this Chapter, we are interested in the latency distribution in finite networks, and the scaling law for large networks with infinite number of secondary nodes under general mobility where spatial inhomogeneity is considered in addition to common features of a variety of mobility models.
4.1 Motivation and Related Work

Many efforts have been made recently to understand the characteristics of CRNs and thus to enable the deployment of such networks for realistic applications, including capacity limits, spectrum sensing, spectrum mobility, and spectrum sharing [2, 88, 89, 91, 92]. These works have presented a very good understanding of the potential of cognitive communications in optimizing spectrum utilization. However, the key question to the deployment of CRNs is not whether the spectrum efficiency is improved, but whether the CRNs are able to support applications. For example, spectrum can be overly used, with a very high throughput, but the latency may become extremely long, falling into the traditional problem of the tradeoff between network throughput and latency [63–66]. To this end, we aim to study a fundamental problem, i.e., what the stochastic properties of end-to-end latency in cognitive radio networks are.

Despite its importance, the latency is an under-explored problem and not well understood in multihop wireless networks. The pioneering work in [61, 62] studied the packet latency for the fully connected wireless ad-hoc networks and showed that there exist bounds on the latency which are tight when the number of nodes is large enough. Instead of full connectivity, some studies [57, 58] further showed that the latency scales asymptotically at least linearly with the transmission distance in wireless sensor networks when these networks are percolated. These results have greatly advanced our understanding of the nature of latency, and also laid a good foundation to approach the problem. Unfortunately, these results may not be applicable to CRNs because (i) asymptotic results were obtained by assuming that wireless nodes are static; and (ii) these results [57, 58, 61, 62] are only derived for large networks when the number of nodes approaches to infinity; (iii) these results are derived for homogeneous networks in which every node has the same capacity in propagation.

Particularly, node mobility plays a critical role on the latency, which has been evidenced by earlier results. For instance, the seminal work [63] showed that mobility can improve the capacity in large wireless ad hoc networks at the cost of the delay. This result is obtained by assuming that nodes move according to an ergodic process that is equally likely to visit any portion of
the network area. That is, the nodes are *spatially homogeneous*. With the similar assumptions, capacity-delay tradeoffs have been extensively studied under various mobility models, such as under the i.i.d model [64], the Brownian motion [65], the reshuffling model [66] and variants of random walk and random way-point models [67, 68]. Later on, spatial inhomogeneity has been taken into account in [69, 70] where the nodes are either restricted to move within a randomly chosen cell or the coverage of a home point. These studies motivate an interesting question about the latency under general mobility.

It is evident that the asymptotic results, though, provide good insights into network performance, may not explain the latency properties when the number of nodes in real applications is *finite*. In other words, the stochastic properties of latency distribution in finite networks lead to understanding of real networks, rather than the large networks. As the last point, CRNs feature *heterogeneity* in wireless nodes, since there are two types of nodes, primary nodes and secondary nodes [104], which is left open for study on the impact of latency distribution.

Putting all together, in this chapter we study the latency distribution in *finite* networks, and the *scaling law* for large networks with *infinite* number of secondary nodes under general mobility where spatial inhomogeneity is considered in addition to common features of a variety of mobility models. We find that in finite CRNs, the latency of information dissemination depends on the *mobility capacity* $\alpha$, which indicates how far a mobile node can reach in spatial domain. Also, there exists a *cutoff point* on $\alpha$, below which the latency has a *heavy-tailed* distribution; and above which its tail distribution is bounded by some Gamma distribution. In addition, as the network grows large, the latency asymptotically scales linearly with respect to the *distance* in terms of the number of hops or the Euclidean distance between the source and destination nodes if the network remains fully connected or percolated. It is interesting, though not surprising, that the density of primary nodes presents an adverse influence on the expected latency, but showing no obvious effect on the dichotomy of the latency tail in finite networks and linear scaling law of the asymptotic latency with respect to dissemination distance.
4.2 System Models and Problem Formulation

In this section, we first describe the network models and then collect basic assumptions, notations and definitions of the metric of interest that will be used throughout the paper.

4.2.1 Network Models

We consider a sequence of CRNs $\mathcal{F}_{m,n}$ indexed by the number of primary users $m$ and the number of secondary users $n$. Primary users $\{u_1, \ldots, u_m\}$ and secondary users $\{v_1, \ldots, v_n\}$ are supposed to move over a bi-dimensional Torus surface $\Omega_n$ of size $L_n \times L_n$. Let $U(t) = (u_1(t), \ldots, u_m(t))$ and $V(t) = (v_1(t), \ldots, v_n(t))$ denote the locations of primary users and secondary users at time $t$ respectively, and denote $V(0)$ and $U(0)$ as their initial positions. We further assume that $L_n$ scales with $n$ as $n^\delta$, where $0 \leq \delta \leq 1/2$. When $\delta = 0$, the network area remains constant and node density increases linearly with $n$. This scaling model has been widely used and known as the dense network [63, 66, 71, 80]. The scaling factor $\delta$ determines how the network area behaves as the number of nodes increases. When $\delta = 1/2$, for which the network area increases linearly with $n$ and the node density remains constant, the above model becomes another popular scaling model, the extended network. In this chapter, we only study the extended network model for simplicity (i.e., $L_n = \sqrt{n}$), as the results from this model can be extended by taking the scaling factor into consideration.

Interference Models

In CRNs, since secondary users cannot interfere with primary users and the latter are usually ignorant of the existence of the former, the primary to primary interference is similar to that in homogeneous wireless networks and thus can be characterized by the well-known protocol model [71], which has been widely adopted in wireless ad hoc and sensor networks [61–64, 71]. The challenge is how to model the interference to secondary users, which are subject to both secondary users, and active primary users who are transmitting or receiving. Let us denote $R_I$ and $r_I$, where $r_I < R_I$, as the interference ranges of the primary and secondary users
respectively, and \( r \) as the transmission range of secondary users. The secondary user \( v_i \) is permitted to transmit to \( v_j \) at time \( t \), only if: i) there are no primary users in the neighborhood, i.e., \( \|v_i(t) - u_k(t)\| > R_I \) and \( \|v_j(t) - u_k(t)\| > R_I \) for any \( k \), where \( \| \cdot \| \) denotes the Euclidean distance; and ii) the conditions from the protocol model are satisfied, that is, \( \|v_i(t) - v_j(t)\| < r \) and \( \|v_i(t) - v_l(t)\| \geq r_I \) for all \( l \neq j \).

In terms of the interference between secondary and primary users, we need to remark that, as illustrated in Fig. 4.1a, there exists a link from \( v_i \) to \( v_j \) when there are no primary receivers in the solid circle and no primary transmitters in the dash circle. When bidirectional links are used, then we consider the communication among secondary users is allowed only when there are no active primary users (either transmitter or receiver) in the neighborhood as shown in Fig. 4.1b, where no primary users in these two circles. This is because if there is an active primary user in the neighborhood of \( v_i \), the secondary user \( v_i \) is not supposed to send or receive information in order to avoid interference with primary users’ communication.

**Mobility Models**

We consider a general mobility model, \( \mathcal{M}(\Phi, \Psi, \alpha) \), which is characterized by three parameters \( \Phi, \Psi, \) and \( \alpha \) over the Torus surface \( \Omega_n \) defined in the network model. First, we consider that a node spends most of its time in a small region, and rarely visits the areas far away from it. We
model this behavior by assuming that each node $v_i$ has a home point [70], located at $v_i^h$. Nodes move “around” their home points according to independent stationary and ergodic processes.

We assume that each home point $v_i^h$ is associated with a fixed point $v_i^c$, which is called the center point of $v_i$. The center points are regularly placed in $\Omega_n$. For example, \{$v_1^c, \ldots, v_n^c$\} are placed regularly at positions \((\frac{1}{2\sqrt{\lambda}} + \frac{i}{\sqrt{\lambda}}, \frac{1}{2\sqrt{\lambda}} + \frac{j}{\sqrt{\lambda}})\) with $0 \leq i \leq \sqrt{n} - 1$ and $0 \leq j \leq \sqrt{n} - 1$ (we generally assume that $n$ is a square of some integer for simplicity, see Fig. 4.2). We describe the distribution of the home point $v_i^h$ around $v_i^c$ by a non-increasing probability density function $\Phi_i(x) = \Phi(x - v_i^c)$, which is assumed to be invariant in all directions and used as the first parameter in the mobility model. The second parameter, $\Psi_i(x) = \Psi(x - v_i^h)$ is used to describe the probability density of a node $v_i$ around $v_i^h$, which is again a non-increasing and direction-invariant function. We assume that $\Psi_i$ is non-zero in and only in a region characterized by a constant $\alpha$; that is, $\Psi_i(x) = \Psi(x - v_i^h) > 0$ when $\|x - v_i^h\| < \alpha$ and $\Psi_i(x) = \Psi(x - v_i^h) = 0$, otherwise. We refer $\alpha$ as mobility capacity since $2\alpha$ characterizes the moving diameter of nodes.

**Remark 4.1.** The idea of “home points” is not new [70] and it has been used to describe the spatial inhomogeneity incurred by the mobility of a particular wireless node. We introduce an additional concept, “center points” to model the heterogeneously spatial distribution of the home points, which characterizes the spatial inhomogeneity incurred by heterogeneous mobility of different wireless nodes. This two-level mobility model accounts for a wide range of mobility patterns. For example, if the probability density function $\Phi(x)$ is a constant function independent of $x$ (i.e., home points are uniformly distributed over $\Omega_n$), $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the Uniform Anisotropic model in [70]. Furthermore, if the probability density function $\Psi_i(x) = \Psi(x - v_i^h) = \delta(x - v_i^h)$, where $\delta(x)$ is the Dirac impulse function, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the static model in [71], where nodes are assumed to be static and uniformly distributed; if $\Psi(x)$ is also a constant function independent of $x$ and $\alpha$, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the homogeneous mobility model in [63]; and if $\Psi(x)$ is a threshold function whose value is zero when $x \geq \alpha$ and a nonzero constant when $x < \alpha$, $\mathcal{M}(\Phi, \Psi, \alpha)$ reduces to the constrained i.i.d model used in [58].
Mobility of Secondary Users: In this chapter, we assume that secondary users are mobile under the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$. To facilitate the study of the dissemination latency of secondary users, we consider three classes according to the spatial inhomogeneity of home points:

- Extremely Inhomogeneous Home Points (EIHP) mobility $\mathcal{M}(\Phi_E, \Psi, \alpha)$: Home points are fixed and regularly placed over $\Omega_n$. Here $\Phi_E(x) = \delta(x)$.

- Partial Inhomogeneous Home Points (PIHP) mobility $\mathcal{M}(\Phi_P, \Psi, \alpha)$: As shown in Fig. 4.2, center points $\{v_c^i\}_{i=1}^n$ partition the $\Omega_n$ into $n$ subregions $\{O_i\}_{i=1}^n$ as Voronoi diagrams. In this class, the home point $v_h^i$ is randomly distributed in $O_i$.

- Homogeneous Home Points (HHP) mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$: Home points $\{v_h^i\}_{i=1}^n$ are independently and uniformly distributed over $\Omega_n$. Here $\Phi_H(x)$ is a constant density function independent of $x$.

Mobility of Primary Users: Without losing generality, we assume that primary users are mobile under a homogeneous mobility $\mathcal{M}_H$ in [63]. Further, we consider that the location of a primary user $u_j$ at time $t$ $u_j(t)$ is ergodic and stationarily distributed uniformly over $\Omega_n$; and $\{u_j(t)\}_{j=1}^m$ are independent and identically distributed.
Discussions

In addition to the description of network models, we need to clarify that these models may not be able to cover every aspect in CRNs. For instance, whether primary users have active communications or not affects not only the available spectrum, but also the interference to secondary users, which results in spectrum dynamics. This factor is not considered in our interference model. However, by mapping a network of mobile nodes to a network of stationary nodes with dynamic links (see [58]), the mobility of primary users $M_H$ also characterizes the spectrum dynamics in large CRNs, where the primary users are static and uniformly distributed, and at any time $t$, primary users independently utilize the spectrum with some positive probability. Therefore, our analysis based on the mobility dynamics also imply the effect of spectrum dynamics, that is, the changing available spectrum to secondary users.

Another issue is that for simplicity of analysis, single-channel is considered for this study. However, in CRNs, multi-channel communication is feasible and preferable. As shown in [92], the main impact of multiple channels is to improve the spectrum opportunities for secondary users and thus increase the probability of the existence of links among secondary users. Since our proofs (see Theorem 4.3) do not rely on such a probability, our results are not limited to the single-channel requirements and can be extended to multi-channel scenarios by taking channel assignment into account, which is beyond the scope of this paper.

4.2.2 Problem Formulation

We denote $[F_{m,n}, M(\Phi, \Psi, \alpha), M_H]$ as a CRN $F_{m,n}$, where secondary and primary users are mobile under $M(\Phi, \Psi, \alpha)$ and $M_H$ respectively. Let $L(t)$ denote the set of communication links among secondary users and the interference model in Section 4.2.1 shows that $L(t)$ is dynamic as the primary and secondary users are mobile.

In this chapter, we focus on the dissemination latency, i.e., how fast information can be disseminated from the source to the destination secondary user. Therefore, rebroadcasting and “store-carry-and-forward” communication paradigm (also named mobility-assisted routing) are
considered. Specifically, by omitting the propagation delay, when source $v_s$ broadcasts a message at time 0, all the secondary users connected to $v_s$ in $\mathbb{L}(0)$ receive the message instantly. Denote $l_{i,j}$ as a communication link between secondary users $v_i$ and $v_j$ and $\mathcal{V}(t)$ as the set of secondary users that have received the message at time $t$.

**Definition 4.1.** The first hitting time between $v_i$ and $v_j$ is defined as $T_h(v_i, v_j) \triangleq \inf\{t \geq 0 : l_{i,j} \in \mathbb{L}(t)\}$.

**Definition 4.2.** The dissemination latency $T_d$ from the source $v_s$ and the destination $v_d$ is defined as:

$$T_d \triangleq \inf\{t \geq 0 : v_d \in \mathcal{V}(t)\}.$$

Based on the definitions and system models, we can formulate the problem as

1. In a finite $\mathcal{F}_{m,n}$, what the distribution of the dissemination latency $T_d$ is;

2. As the network grows large, say to infinity, whether the dissemination latency $T_d$ is scalable or not.

To further describe how fast information can be disseminated, we usually scale the dissemination latency $T_d$ with the distance between the source and destination secondary users. In $[\mathcal{F}_{m,n}, M(\Phi, \Psi, \alpha), M_H]$, three metrics can be used to characterize how far two nodes $v_i$ and $v_j$ are apart:

- $d(t)(v_i, v_j)$: the distance between $v_i$ and $v_j$ at time $t$.

- $d_h(v_i, v_j)$ and $d_c(v_i, v_j)$: the distance between home points and center points of $v_i$ and $v_j$ respectively.

Here the distance can be any $p$-norm metric function and we consider two of the most popular metrics transmission hops and Euclidean distance.

**Definition 4.3.** Denote $D$ as the distance, i.e., either the number of hops or the Euclidean distance between the source $v_s$ and destination $v_d$, which can be one of the three distances.
explained above. We define \( S_d \triangleq \frac{T_d}{D} \), which characterizes how fast information disseminates and is called dissemination speed in this chapter.

## 4.3 Main Results

The key question in this study is how fast information is disseminated in both finite and large CRNs under general mobility \( \mathcal{M}(\Phi, \Psi, \alpha) \). We first study the dissemination latency \( T_d \) in CRNs where secondary users are mobile under the three subclasses of models EIHP, PIHP and HHP, respectively. Then based on the generalization of these results, we obtain the fundamental properties of the dissemination latency \( T_d \) when secondary users are mobile under the general mobility \( \mathcal{M}(\Phi, \Psi, \alpha) \). We summarize our main results as follows.

**Theorem 4.1.** In a finite CRN \([F_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), \mathcal{M}_H]\), there exists a cutoff point on the mobility capacity \( \alpha \), above which the tail distribution of dissemination latency \( T_d \) is bounded by some Gamma distribution; below which \( T_d \) has a heavy-tailed distribution and \( \mathbb{P}(T_d = \infty) > 0 \).

**Remark 4.2.** \( \mathbb{P}(T_d = \infty) > 0 \) indicates a positive probability that the destination will not receive the message from the source. Thus the requirement \( \mathbb{P}(T_d < \infty) = 1 \) in mobile wireless networks is equivalent to the connectivity in the wired networks, which is used as a prerequisite to evaluate network functions. Moreover, a heavy tail of the dissemination latency \( T_d \) implies a significant probability that it takes long time to disseminate a message from the source to the destination. Thus in addition to a bounded dissemination latency (i.e., \( \mathbb{P}(T_d < \infty) = 1 \)), a light-tailed dissemination latency \( T_d \) (i.e., \( E(T_d) < \infty \)) is also crucial for time-critical applications in CRNs. Therefore, a light-tailed distribution of \( T_d \) is assumed or required in many deployments and performance studies of wireless networks in the literature. For example, the authors in [63] implicitly assume that the dissemination latency is exponentially bounded (light-tailed) so as to make their delay-capacity tradeoff analysis tractable.

Theorem 4.1 tells that to achieve a light-tailed dissemination latency (note that Gamma distribution is a type of light-tailed distribution), the mobility capacity of secondary users \( \alpha \)
need to be larger than some cutoff point, which is specifically identified in Proposition 4.1 for EIHP, Proposition 4.2 for PIHP and Proposition 4.3 for HHP, respectively. This result encourages the existing endeavor of deploying CRN for practical applications, including time-critical applications, such as emergency networks and military networks.

As the network size increases, we have the following theorem on the scalability of the dissemination latency $T_d$.

**Theorem 4.2.** Given a large connected\(^1\) cognitive radio network $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi, \Psi, \alpha), \mathcal{M}_H]$, there exists a finite constant $\kappa$ such that $\mathbb{P}(\lim_{D \to \infty} S_d = \lim_{D \to \infty} \frac{T_d}{D} = \kappa) = 1$.

**Remark 4.3.** Scalability has been one of the most fundamental problems that has discouraged the deployment of large wireless networks [64, 71]. Theorem 4.2 reveals that in large connected CRNs, the dissemination latency $T_d$ asymptotically scales linearly with the initial distance between the source and destination, i.e., the message sent by a source reaches its destination at a fixed asymptotic speed. This result enables the feasible deployment of CRNs for large applications.

It is worthy of noting that we aim to understand the fundamental properties of the dissemination latency $T_d$ in CRNs under general mobility. However, besides the theoretical importance of our findings, our results can be used practically not only in the initial deployment of a CRN, but also in evaluating the performance of network applications. For example, in a large deployment of CRNs as sensor-actuator networks, the result in Theorem 4.2 can be used to estimate the delay elapsed between the time at which an incoming event is sensed and the time that this event report is retrieved by the data collecting sink. In the next two sections, we present the proofs for Theorem 4.1 and Theorem 4.2, which studies the distribution and scalability of the dissemination latency $T_d$ in finite and large CRNs under EIHP, PIHP, and HHP mobility, respectively.

\(^1\)We consider two types of connectivity in large CRNs: full connectivity and percolation-based connectivity. The former is that there exists a communication path between any two nodes; and the latter is that there exists a large component well scattered over the entire network.
4.4 The Distribution of $T_d$ in Finite Networks

Here, we present the proof for Theorem 4.1 described in Section 4.3. This result tells that there exists a cutoff point on the *mobility capability* $\alpha$ in a finite cognitive radio network under general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, above which the tail distribution of the dissemination latency $T_d$ is bounded by some *Gamma* distribution and below which its distribution has a heavy tail. For the convenience of analysis, we first study the distribution of $T_d$ under these three subclasses of mobility models, i.e., EIHP $\mathcal{M}(\Phi_E, \Psi, \alpha)$, PIHP $\mathcal{M}(\Phi_P, \Psi, \alpha)$, and HHP mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$ based on the *spatial inhomogeneity* of home points defined in Section 4.2, respectively. Then we move on to identify the fundamental properties of $T_d$ under the general mobility. The scalability of the dissemination latency $T_d$ in large CRNs (Theorem 4.2) will be considered in the next section. To proceed, we find the following definitions useful toward the derivation of tail distribution of the latency $T_d$.

**Definition 4.4.** If $Z$ and $Z'$ are random variables such that $P(Z > z) \leq P(Z' > z)$ for all $z$, we say that $Z$ is stochastically dominated by $Z'$ and write $Z \overset{\mathcal{D}}{\lesssim} Z'$; and if $Z \overset{\mathcal{D}}{\lesssim} Z'$, there exists a random variable $\hat{Z}'$, which has the same distribution of $Z'$ such that $Z \leq \hat{Z}'$ ($Z'$ is called a coupling of $Z'$ [95].)

**Definition 4.5.** If $Z$ and $Z'$ are random variables such that $P(Z > z) \leq P(Z' > z)$ for large $z$, we say that the $Z$’s tail is stochastically dominated by $Z'$.

**Remark 4.4.** Coupling is a very important tool in probability theory which is used throughout the paper. To use the coupling method, stochastic domination is required (as shown in Definition 4.4). However, in finite CRNs, we are interested in the tail distribution of the dissemination latency $T_d$, which implies that only stochastic tail domination needs to be considered. Therefore, in order to use coupling, we need the following lemma, which bridges the gap between stochastic domination and stochastic tail domination.

**Lemma 4.1.** Given non-negative i.i.d random variables $\{X_i\}_{i=1}^{\infty}$ and $\{Y_i\}_{i=1}^{\infty}$ where $P(X_i > t) \leq P(Y_i > t)$ for large $t$, i.e., the tails of the former are stochastically dominated by the latter,
there exist i.i.d random variables \( \{\bar{X}_i\}_{i=1}^{\infty} \), which have the same tail distribution with \( \{X_i\}_{i=1}^{\infty} \) and are stochastically dominated by \( \{Y_i\}_{i=1}^{\infty} \). Furthermore, for any finite \( k \), \( \sum_{i=1}^{k} X_i \) has the same tail distribution with \( \sum_{i=1}^{k} \bar{X}_i \).

**Proof.** (Sketch.) Assume \( \mathbb{P}(X_i > t) \leq \mathbb{P}(Y_i > t) \) when \( t > t_c \) for some finite constant \( t_c \). We construct \( \{\bar{X}_i\}_{i=1}^{\infty} \) as \( \bar{X}_i = 0 \) when \( X_i \leq t_c \) and \( \bar{X}_i = X_i \) otherwise. This proves the first part.

For the second part, we only need to show \( \mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) = \mathbb{P}(X_1 + X_2 > t) \) for large \( t \):

\[
\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) = \mathbb{P}(\bar{X}_1 < t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t|\bar{X}_1 < t_c) + \\
\mathbb{P}(\bar{X}_1 > t - t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t|\bar{X}_1 > t - t_c) \\
+ \mathbb{P}(t_c < \bar{X}_1 < t - t_c)\mathbb{P}(\bar{X}_1 + \bar{X}_2 > t|t_c < \bar{X}_1 < t - t_c).
\]

Note that the third item on the right hand side is equal to its counterpart of \( \mathbb{P}(X_1 + X_2 > t) \) and the first two items are on the higher order of the third item as \( t \to \infty \). Thus \( \mathbb{P}(\bar{X}_1 + \bar{X}_2 > t) \to \mathbb{P}(X_1 + X_2 > t) \) for large \( t \) and this completes the proof. \( \square \)

### 4.4.1 Distribution of \( T_d \) under EIHP Mobility \( \mathcal{M}(\Phi_E, \Psi, \alpha) \)

Prior studies [57, 58] have shown that *propagation delay* in networks whose topologies change frequently (e.g., due to mobility) is negligible, in comparison with the latency incurred by the topology dynamics. Therefore, \( T_d \) can be coupled as the sum of a sequence of the first hitting time \( T_h \) between secondary users along a communication path from the source to the destination.
node. Hence we study $T_h$ first. In EIHP mobility, secondary users move around home points, which are overlaid with center points, with the Euclidean distance between neighboring home points being $\sqrt{\frac{1}{\lambda}}$ (see Fig. 4.2). The following lemma studies the property of the first hitting time $T_h(v_i,v_j)$ between $v_i$ and $v_j$ with neighboring home points.

**Theorem 4.3.** Given secondary users $v_i$ and $v_j$ in a finite CRN $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_E, \Psi, \alpha), \mathcal{M}_H]$ with $d_c(v_i, v_j) = \sqrt{\frac{1}{\lambda}}$, we have i) $P(T_h(v_i,v_j) = \infty) = 1$ if $\alpha < \frac{\sqrt{1-r}}{2}$; ii) otherwise, $E(T_h(v_i,v_j)) < \infty$ and $P(T_h(v_i,v_j) > t) \leq e^{-c_1t}$ for sufficiently large $t$ and some positive constant $c_1$.

**Proof.** At time $t$ if and only if $d(t)(v_i,v_j) < r$, nodes $v_i$ and $v_j$ may communicate directly. Also, $d(t)(v_i,v_j) > d_h(v_i,v_j) - 2\alpha$ for all $t$. Thus if $\alpha < \frac{\sqrt{1-r}}{2}$, $d(t)(v_i,v_j) > r$ for all $t$, which implies that $v_i$ and $v_j$ cannot communicate with each other. This completes the proof of part i).

For $\alpha > \frac{\sqrt{1-r}}{2}$, let $E_t$ denote the event that there exists no communication link between $v_i$ and $v_j$ at time $t$ and $\bar{E}_t$ as its complement. As shown in Fig. 4.3, a necessary condition for $\bar{E}_t$ is that there exist no primary users within the bigger circle centered at $o$, and a sufficient condition for $\bar{E}_t$ is that $v_i$ lies in the shaded region $S_1$, $v_j$ in $S_2$ and no primary users in the bigger circle. Therefore,

$$0 < 1 - (1 - \pi R_t^2/(n/\lambda))^m < P(E_t) = 1 - P(\bar{E}_t)$$

$$< 1 - (1 - \pi R_t^2/(n/\lambda))^m \Psi(S_1)\Psi(S_2) < 1, \quad (4.1)$$

where $(1 - \pi R_t^2/(n/\lambda))^m$ characterizes the probability of the necessary condition and $(1 - \pi R_t^2/(n/\lambda))^m \Psi(S_1)\Psi(S_2)$ characterizes the probability of the sufficient condition for $\bar{E}_t$, respectively. $\Psi(S) = \int_S \Psi dS$ and $n/\lambda$ is the area of $\Omega_n$. To proceed, we next find an index set $\mathcal{I}$ such that $\{E_t\}_{t \in \mathcal{I}}$ are independent and let $\varepsilon = P(E_t)$ for convenience.

Denote $\rho$ as a renewal interval for secondary users, i.e., for any $t > 0$, $\{v_i(t') : t' \leq t\}$ and
\{v_i(t'' + \rho) : t'' \geq t\} are independent. Denote \(\{\rho_i\}_{i=1}^{\infty}\) as a sequence of i.i.d random variables with the same distribution as \(\rho\). Now we consider the index set \(I_t = \{0, t_1, \ldots, t_{N(t)}\} \subset (0, t]\), where \(t_k = \sum_{i=1}^{k} \rho_i\) and \(N(t) = |I_t| = \max\{k : t_k \leq t\}\). Observe that

\[
P(T_h(v_i, v_j) > t) \leq P(\bigcap_{s \in I_t} E_s) = \prod_{s \in I_t} P(E_s) \] (4.2)

where the last equality is by the independency of \(\{E_s\}_{s \in I_t}\) and by conditioning on \(N(t)\), we have

\[
P(T_h(v_i, v_j) > t) \leq E(e^{-\beta N(t)}) = E(e^{-\beta N(t)}),
\]

where \(\beta = -\log \varepsilon > 0\). In addition, for any \(\tau > 0\),

\[
E(e^{-\beta N(t)}) = E(e^{-\beta N(t)} I_{N(t) \leq \tau t}) + E(e^{-\beta N(t)} I_{N(t) > \tau t}) \leq P(N(t) \leq \tau t) + e^{-\beta \tau N(t)}.
\]

Note that the finite sum of exponentially bounded random variables is still exponentially bounded [95, 105]. Thus, if we can show that \(P(N(t) \leq \tau t)\) is exponentially bounded, we will finish the proof. In order to proceed, we assume that the tails of renewals \(\{\rho_i\}_{i=1}^{\infty}\) are exponentially bounded. This assumption is reasonable considering the network is finite, which has been well-explained in many mobility models [58, 63, 67, 70, 105]. We next show that \(P(N(t) \leq \tau t)\) is exponentially bounded. By letting \(k = \tau t\), \(P(N(t) \leq \tau t) = P(\sum_{i=1}^{\tau t} \rho_i > t) = P(\sum_{i=1}^{k} r_i > \frac{k}{\varepsilon})\). The last item is obviously bounded by some exponential variable considering \(\{\rho_i\}_{i=1}^{\infty}\) are exponentially bounded. This completes the proof.

We next present our main result on the tail distribution of the dissemination latency \(T_d\) under EIHP mobility.

**Proposition 4.1.** Given \([F_{m,n}, M(\Phi_E, \Psi, \alpha), M_H]\) with finite users, if \(\alpha > \frac{\sqrt{\frac{\tau}{2} - r}}{2}\), the tail distribution of the dissemination latency \(T_d\) is stochastically dominated by a Gamma distribution, \(\Gamma(2\sqrt{n}, c_2)\), and the tail distribution of the dissemination speed \(S_d\) is stochastically dominated

102
by $\Gamma(\sqrt{\lambda}D, \frac{c_2}{2})$ for some positive constant $c_2$; otherwise, $T_d$ has a heavy-tailed distribution and $\mathbb{P}(T_d = \infty) > 0$.

Proof. (Sketch.) As the end to end latency, $T_d$ is clearly bounded by the transmission delay along any path from the source $v_s$ to destination $v_d$. Theorem 4.3 shows that, if $\alpha > \frac{\sqrt{r} - \tau}{2}$, a link exists between two neighboring secondary users with positive probability. Therefore, we can identify a Manhattan path through which $v_s$ first transmits the message vertically until the message reaches the secondary user whose center point has the same horizontal coordinate with $v_c$, and then transmits the message horizontally to $v_d$ as shown in Fig. 4.4. Denote $\{X_k\}_{k=1}^{\infty}$ as a sequence of random variables with identical distributions as the first hitting time between neighboring secondary users. Note that a Manhattan path consists of at most $2\sqrt{n}$ communication links and thus $T_d \leq 2\sqrt{n} \sum_{k=1}^{\infty} X_k$.

The next challenge is that the first hitting time of neighboring links, i.e., $X_i$ and $X_{i+1}$ are not independent. To tackle this challenge, we assume that after receiving the message, each secondary user will hold this message for a renewal time $\rho$ before it tries to relay the message. Let $\{\rho_i\}_{i=1}^{\infty}$ be a sequence of renewals and $Y_k = X_k + \rho_k$. It is clear that $T_d \leq 2\sqrt{n} \sum_{k=1}^{\infty} Y_k$. Note that $Y_k$ is bounded by $\text{exponential}(c_2)$ (since both $X_k$ and $\rho_k$ are both exponentially bounded) and $\{Y_k\}_{k=1}^{\infty}$ are clearly independent. Let $\{\hat{Y}_k\}_{k=1}^{\infty}$ be a sequence of independent random variables distributed as $\text{exponential}(c_2)$, we have

$$\mathbb{P}(T_d > t) \leq \mathbb{P}(\sum_{k=1}^{2\sqrt{n}} Y_k > t) \leq \mathbb{P}(\sum_{k=1}^{2\sqrt{n}} \hat{Y}_k > t),$$

(4.3)

where the last inequality is from Lemma 4.1 and coupling (Definition 4.4). By the moment generating function technique [106], we know that $Y$ follows a Gamma distribution, $\Gamma(2\sqrt{n}, c_2)$. This completes the proof for $T_d$.

To further describe how fast information can be disseminated, we study the dissemination speed $S_d = \frac{T_d}{D}$. We need first to specify the distance $D$ between $v_s$ and $v_d$. As analyzed above, since for any $v_i$ under EIHP mobility, the number of secondary users that can communicate
with $v_i$ directly is finite, and thus when $v_d$ is beyond the transmission range of $v_i$, hop by hop communication is necessary, in which case transmission hops can describe “how far” more accurately than the Euclidean distance. Therefore, $D$ here denotes the Manhattan distance between $v_s$ and $v_d$ by which the maximum number of transmission hops between $v_s$ and $v_d$ can be expressed as $\sqrt{\lambda D}$. Then it follows $S_d = \frac{T_d}{D} \leq \frac{\sum Y_k}{D}$. Similarly, let $Y' = \frac{\sum Y_k}{D}$ and $\mathbb{P}(S_d > t) \leq \mathbb{P}(Y' > t)$ by coupling. Then $Y'$ is distributed as $\Gamma(\sqrt{\lambda D}, \frac{c_2}{D})$ from [106], which obtains the result for $S_d$.

When $\alpha < \frac{\sqrt{\lambda - r}}{2}$, Theorem 4.3 says that the first hitting time between any two secondary users $T_h(v_i, v_j) = \infty$. Therefore, $T_d = \infty$, which completes the proof.

Remark 4.5. Proposition 4.1 shows that the tail of the dissemination speed $S_d$ is bounded by $Y' \sim \text{Gamma}(\sqrt{\lambda D}, \frac{c_2}{D})$. Note that the mean $E(Y') = \sqrt{\lambda c_2}$ and the variance, $\text{var}(Y') = \lambda c_2^2 / D$. As $D$ increases, $\text{var}(Y') \to 0$, which leads to $Y' = \sqrt{\lambda c_2}$ for large $D$. Intuitively, this implies that in large CRNs where $D$ is usually large, $S_d$ may be bounded by some constant and its tail disappears as $D$ increases. Actually, in Section 4.5, we will rigorously prove that in large CRNs, $S_d$ approaches to some constant, which agrees with our intuition here.

4.4.2 Distribution of $T_d$ under PIHP Mobility $M(\Phi_P, \Psi, \alpha)$

Note that the main difference between PIHP and EIHP mobility is that home points in the former are randomly located, and thus for neighboring secondary users $v_i$ and $v_j$ with $d_c(v_i, v_j) = \sqrt{\frac{1}{X}}$, $d_h(v_i, v_j) \neq \sqrt{\frac{1}{X}}$ in PIHP mobility. But under PIHP mobility, $d_h(v_i, v_j)$ is
still bounded and we have $\mathbb{P}(d_h(v_i, v_j) \leq \sqrt{\frac{\lambda}{\alpha}}) = 1$. Thus, by similar proof to Theorem 4.3, we are able to see that for any $v_i$ and $v_j$ with $d_c(v_i, v_j) = \sqrt{\frac{\lambda}{\alpha}}$, if $\alpha > \frac{\sqrt{\frac{\lambda}{\alpha}} - r}{2}$, the first hitting time $T_h(v_i, v_j)$ is exponentially bounded; and if $\alpha < \frac{\sqrt{\frac{\lambda}{\alpha}} - r}{2}$, $\mathbb{P}(T_h(v_i, v_j) = \infty) > 0$. Therefore, through the similar proof as that of Proposition 4.1, we have the following results about the dissemination latency $T_d$ and speed $S_d$ in finite CRNs under EIHP mobility:

**Proposition 4.2.** Given $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_P, \Psi, \alpha), \mathcal{M}_H]$ with finite users, if $\alpha > \frac{\sqrt{\frac{\lambda}{\alpha}} - r}{2}$, the tail distribution of $T_d$ is stochastically dominated by a Gamma distribution $\Gamma(2\sqrt{\frac{\lambda}{\alpha}}, c_3)$ and the tail distribution of $S_d$ is stochastically dominated by $\Gamma(\sqrt{\lambda D}, \frac{c_3}{D})$ for some positive constant $c_3$; otherwise, $T_d$ is heavy tail distributed and $\mathbb{P}(T_d = \infty) > 0$.

**Remark 4.6.** In Propositions 4.1 and 4.2, the distance $D$ between the source and destination nodes $v_s$ and $v_d$ has been considered as Manhattan distance between their center points $v^c_s$ and $v^c_d$. Note that under both EIHP and PIHP mobility, the mobile region of any secondary user $v_i$ is constrained to the coverage of home point $v^h_i$ and thus the number of nodes that can communicate with $v_i$ directly is finite. Hence information can only be delivered from the source to the destination hop by hop under these two mobility models. Thus Manhattan distance, which characterizes the number of transmission hops between the source and destination, is an appropriate metric to measure the distance between nodes $v_s$ and $v_d$.

### 4.4.3 Distribution of $T_d$ under HHP Mobility $\mathcal{M}(\Phi_H, \Psi, \alpha)$

When HHP mobility is considered, home points are homogeneously distributed in the whole network $\Omega_n$. Therefore, unlike EIHP and PIHP mobility, the distance between home points of secondary users $v_i$ and $v_j$ is homogeneous and may be any value in the interval $(0, \sqrt{\frac{\lambda}{2\alpha}})$ (see Fig. 4.5, where $\Omega_n$ is a torus surface without border effect). We next show that, to overcome the randomness of $d_h(v_i, v_j)$, secondary users need to move over the whole network (large mobility capability $\alpha$) to eliminate the heavy tail of the first hitting time.
Lemma 4.2. Given a CRN \([F_{m,n}, M(\Phi_H, \Psi, \alpha), M_H]\) with finite users, if \(\alpha > \frac{\sqrt{n^2 - r}}{2}\), the first hitting time \(T_h(v_i, v_j)\) is exponentially bounded; and if \(\alpha > \frac{\sqrt{n^2 - r}}{2}\), \(T_h(v_i, v_j)\) has a heavy tail and \(\mathbb{P}(T_d = \infty) > 0\).

Proof. As shown in Fig. 4.5, there may exist a communication link between \(v_i\) and \(v_j\) (i.e., \(\mathbb{P}(T_h(v_i, v_j) < \infty) = 1\)), if and only if \(v^h_i\) is located in the solid circle \(C\) centered at \(v^h_i\). Thus the probability that \(v^h_j\) is distributed outside \(C\) (i.e., \(\mathbb{P}(T_h(v_i, v_j) = \infty) > 0\)), will incur a heavy tail of \(T_h(v_i, v_j)\) (that is, \(E(T_h(v_i, v_j)) = \infty\)). Therefore, to eliminate the heavy tail, \(C\) must cover the whole network \(\Omega_n\), which requires \(2\alpha + r > \sqrt{\frac{n^2}{2}} \Rightarrow \alpha > \frac{\sqrt{n^2 - r}}{2}\). When \(\alpha > \frac{\sqrt{n^2 - r}}{2}\), \(v_i\) and \(v_j\) may communicate with each other with some positive probability at any time. Hence, with the similar proof of Theorem 4.3, we can show that \(T_h(v_i, v_j)\) is exponentially bounded. This completes the proof.

Different from the above two scenarios (i.e., EIHP and PIHP), under HHP mobility, any secondary user \(v_i\) may receive the message directly from the source \(v_s\), and any \(v_i\) that carries the message may in turn copy this message to all secondary users it encounters along its trajectory. Hence, we cannot apply the coupling method in calculating \(T_d\) hop by hop along the end to end path for HHP mobility. Instead, we use a stochastic model to analyze \(T_d\) and obtain the following result:

Proposition 4.3. Given \([F_{m,n}, M(\Phi_H, \Psi, \alpha), M_H]\) with finite users, if \(\alpha < \frac{\sqrt{n^2 - r}}{2}\), \(T_d\) has a heavy-tailed distribution and \(\mathbb{P}(T_d = \infty) > 0\); and if \(\alpha > \frac{\sqrt{n^2 - r}}{2}\), the tail of \(T_d\) is stochastically
dominated by a Gamma distribution.

**Proof.** When \( \alpha < \sqrt{\frac{2n - r}{\lambda}} \), there exists some positive probability that all the home points \( \{v_j^h, j \neq s\} \) are located outside the circle centered at \( v_s^h \) with radius \( 2\alpha + r \), which implies \( P(T_d = \infty) > 0 \) and thus a heavy tail.

When \( \alpha > \sqrt{\frac{2n - r}{\lambda}} \), Lemma 4.2 shows that the tail of the first hitting time \( T_h(v_i, v_j) \) between any \( v_i \) and \( v_j \) is stochastically dominated by exponential\((c_4)\) for some constant \( c_4 \). If we can show that when \( T_h(v_i, v_j) \) is distributed as exponential\((c_4)\), the tail of the resulting dissemination latency \( T_d' \) is stochastically dominated by a Gamma distribution, then by Lemma 4.1 and coupling, which shows that \( P(T_d > t) < P(T_d' > t) \) for large \( t \), we complete the proof.

Assume \( T_h(v_i, v_j) \sim \text{exponential}(c_4) \) for any \( v_i \) and \( v_j \). Denote by \( \zeta \) the number of secondary users, which carry the message sent by source \( v_s \) before this message is successfully delivered to the destination \( v_d \). The proof is based on modeling \( \zeta \) as an absorbing finite-state Markov chain. The Markov chain consists of states \( k = 0, 1, 2, \ldots, n - 1 \). The state \( k \) denotes \( \zeta = k \) and the state 0 denotes the absorbing state that \( v_d \) successfully receives the message as shown in Fig. 4.6.

When secondary users hit each other, messages will be forwarded to the ones without a copy of the message. Therefore, when there are \( k \) secondary users carrying the message, the message is sent to another secondary user at rate \( c_4 k(n - 1 - k) \) with the transition from \( k \) to \( k + 1 \), and to destination \( v_d \) at rate \( c_4 k \) with transition from \( k \) to 0, as shown in Fig. 4.6. The chain jumps from state \( k \) to \( k + 1 \) with probability \( \frac{n-1-k}{n-k} \) and transits from \( k \) to 0 with probability \( \frac{1}{n-k} \). The sojourn time \( S_k \) in state \( k \) is exponentially distributed with intensity \( c_4 k(n - k) \). Since \( S_1, S_2, \ldots, S_{n-1} \) are mutually independent. Thus \( P(\zeta = k) = \frac{1}{n-k} \prod_{j=1}^{k-1} \frac{n-1-j}{n-j} = \frac{1}{n-1} \).

Conditioning \( T_d' \) on \( \zeta \), we have \( T_d'|(\zeta = k) = \sum_{j=1}^{k} S_j \), which is clearly exponentially bounded and therefore, \( P(T_d' > t) = \sum_{k=1}^{n-1} P(\zeta = k)P(T_d'|(\zeta = k) > t) = \frac{1}{n} \sum_{k=1}^{n-1} P(T_d'(\zeta = k) > t) \) is exponentially bounded (note that the sum of exponential variables is still exponentially distributed). That is, the tail of \( T_d' \) is bounded by \( \text{exponential}(c_5) = \text{Gamma}(1, c_5) \) for some positive constant \( c_5 \). This completes the proof. □
Remark 4.7. Unlike EIHP and PIHP mobility, given $\alpha > \frac{\sqrt{\pi^2 - r^2}}{n}$, home points under HHP mobility are homogeneously distributed and secondary users move around the whole network according to some stationary process. Thus $T_d$ is homogeneous for any pair of secondary users and the distance between $v_s$ and $v_d$ $D$ has no obvious impact on $T_d$. Therefore here we only study the distribution of the dissemination latency $T_d$.

Remark 4.8. From Propositions 4.1, 4.2 and 4.3, it seems that the dissemination latency $T_d$ is independent of the density of primary users. We need to clarify here that this “independency” is related to the stochastic anatomy. However, the density of primary users may have a negative impact on the dissemination latency $T_d$ and thus the speed $S_d$ (see simulation results in Fig. 4.7). Particularly, as shown in the proof of Theorem 4.3, the expected first hitting time $E(T_h(v_i, v_j))$ and thus the expected dissemination latency $E(T_d)$ obviously increase as the number of primary users $m$ increases (see Figs. 4.7 and 4.8). On the other hand, in terms of the statistically macroscopic structure (i.e., light-tailed distribution or heavy-tailed distribution), the proofs for Propositions 4.1, 4.2 and 4.3 demonstrate that $T_d$ is independent of the primary users.

The results in Propositions 4.1, 4.2 and 4.3 collectively suggest that when secondary users are mobile under any subclass of the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, there exists a dichotomy structure on the mobility capacity $\alpha$. Above the cutoff point the tail distribution of $T_d$ is bounded by some Gamma distribution (light-tailed), and below which $T_d$ has a heavy-tailed distribution and $P(T_d = \infty) > 0$. The cutoff value also increases as the spatial homogeneity of home points increases (i.e., increases in the order EIHP, PIHP and HHP). This is equivalent to saying that the cutoff point exists in CRNs under the general mobility $\mathcal{M}(\Phi, \Psi, \alpha)$, which completes the proof of Theorem 4.1. We next further explain and validate the theoretical cutoff anatomy through simulations.

Our theoretical analysis on the distributions of the first hitting time $T_h$ and the dissemination latency $T_d$ has been verified by a series of simulation results. We here consider a finite CRN with $n = 16$ mobile secondary users within an area $2 \times 2$ square meters (i.e., $\lambda = 4$). The time is partitioned into unit slots. In each time slot, primary users are uniformly distributed.
Figure 4.6: Illustration of the Markov chain. Each state represents the number of secondary users carrying the message.

at random within the network area and secondary users are uniformly distributed around their home points (i.e., \( \Psi \) is uniform). Furthermore, home points are uniformly distributed around the center points under PIHP mobility (i.e., \( \Phi_P \) is uniform). The transmission range \( r \) of secondary users and the interference range \( R_I \) of primary users are set as \( r = 0.1 \) meter and \( R_I = 0.3 \) meter, respectively. The probability is calculated based on the average of 1000 independent simulations.

Fig. 4.7 shows the complementary distribution (CCDF) of the dissemination latency \( \mathbb{P}(T_d > t) \) on a log-log scale for EIHP, PIHP and HHP models with different values of mobility capacity \( \alpha \) and the number of mobile primary users \( m \). It is observed in Fig. 4.7 that as the number of primary users \( m \) increases, the curves move right-ward, which indicates the increasing expected dissemination latency. However, regardless of the value of \( m \), when \( \alpha = 0.4 \) meter, which is larger than the cutoff point under EIHP but smaller than those under PIHP and HHP, the dissemination latency \( T_d \) has a light tail under EIHP but heavy tails under PIHP. As \( \alpha \) increases to 0.8 meter, which is larger than the cutoff point in PIHP, but still less than that in HHP, the heavy tail of \( T_d \) in PIHP disappears, but \( T_d \) in HHP presents a heavy tail. Fig. 4.8 shows the complementary distribution (CCDF) of the first hitting time \( \mathbb{P}(T_h > t) \) on a log-log scale for EIHP, PIHP and HHP, which demonstrate a similar quantitative behavior as Fig. 4.7. These results are in good agreement with our theoretical analysis in Propositions 4.1, 4.2 and 4.3, and arguments in Remark 4.8.
The dissemination latency: \( \alpha = 0.4, m = 4 \)

![Figure 4.7: CCDF of dissemination latency \( T_d \) under general mobility.](image)

The dissemination latency: \( \alpha = 0.8, m = 4 \)

![Figure 4.7: CCDF of dissemination latency \( T_d \) under general mobility.](image)

The dissemination latency: \( \alpha = 0.4, m = 8 \)

![Figure 4.7: CCDF of dissemination latency \( T_d \) under general mobility.](image)

The first hitting time: \( \alpha = 0.4, m = 4 \)

![Figure 4.8: CCDF of the first hitting time \( T_h(v_i, v_j) \) between neighboring secondary users \( v_i \) and \( v_j \).](image)

The first hitting time: \( \alpha = 0.8, m = 4 \)

![Figure 4.8: CCDF of the first hitting time \( T_h(v_i, v_j) \) between neighboring secondary users \( v_i \) and \( v_j \).](image)

The first hitting time: \( \alpha = 0.4, m = 8 \)

![Figure 4.8: CCDF of the first hitting time \( T_h(v_i, v_j) \) between neighboring secondary users \( v_i \) and \( v_j \).](image)

4.5 The Scalability of \( T_d \) in Large CRNs

We next prove Theorem 4.2, which states that the dissemination latency \( T_d \) asymptotically scales linearly with the dissemination distance in large mobile CRNs. Our study on the distribution of \( T_d \) in Section 4.4 tells that \( E(T_d) \to \infty \) as the network size grows large, which implies that the distribution of \( T_d \) cannot be used to measure how fast information is disseminated in large CRNs. Therefore, in large CRNs, instead of probability distribution, we investigate the scalability of \( T_d \), i.e., the scaling behavior of \( T_d \) with respect to the distance \( D \) between the
source \( v_s \) and destination \( v_d \), which is characterized by \( S_d = \frac{T_d}{D} \) (called dissemination speed for convenience, see [57]).

The results in finite networks have presented an implication that the tail of \( S_d \) may disappear as the network size increases (see Remark 4.5 in Section 4.4), which will be validated in this section. It is worthy of noting that derivation in Section 4.4 is based on the assumption that the number of nodes is finite. For large CRNs, we use the large number theory to demonstrate that \( T_d \) asymptotically scales linearly with \( D \), i.e., \( \lim_{D \to \infty} S_d \) converges to some positive constant. First, we present the following theorem, which is helpful to our analysis to the limit of the dissemination speed \( S_d \):

**Theorem 4.4** (Liggett’s subadditive ergodic theorem, [102]). Let \( \{Z_{h,q}\} \) be a collection of random variables indexed by integers satisfying \( 0 \leq h < q \). Suppose \( \{Z_{h,q}\} \) has the following properties: (i) \( Z_{0,q} \leq Z_{0,h} + Z_{h,q} \); (ii) For each \( q \), \( \mathbb{E}(|Z_{0,q}|) < \infty \) and \( \mathbb{E}(Z_{0,q}) \geq cq \) for some constant \( c > -\infty \); (iii) The distribution of \( \{Z_{h,h+k}, k \geq 1\} \) does not depend on \( h \); (iv) For each \( k \geq 1 \), \( \{Z_{qk,(q+1)k}, q \geq 0\} \) is a stationary sequence; (v) If \( k \geq 1 \), \( \{Z_{qk,(q+1)k} : q \geq 0\} \) are ergodic. Then we have (a) \( \zeta = \lim_{q \to \infty} \mathbb{E}(Z_{0,q})/q = \inf_{q \geq 1} \mathbb{E}(Z_{0,q})/q \); (b) \( \mathcal{Z} = \lim_{q \to \infty} Z_{0,q}/q \) exists almost surely; (c) \( \mathbb{E}(\mathcal{Z}) = \zeta \); and (d) \( \mathcal{Z} = \zeta \) almost surely.

Liggett’s theorem provides a method to study the limiting behavior of a large random process. Our approach is to progressively increase the number of secondary and primary users \( n \) and \( m \) in \( \Omega_n \). Therefore homogeneously distributed primary users are asymptotically distributed as a two-dimensional Poisson point process with density \( \lambda_p = \lambda n/m \), which is a constant by increasing \( m \) and \( n \) proportionally.

### 4.5.1 Scalability of \( T_d \) under EIHP and PIHP Mobility

Since the study of the first hitting time \( T_h(v_i, v_j) \) between neighboring secondary users \( v_i \) and \( v_j \) under EIHP or PIHP models in Section 4.4 is independent of the network size, these results still hold for large CRNs. That is, when the mobility capacity \( \alpha > \sqrt{\frac{4-\rho}{2}} \) for EIHP or \( \alpha > \sqrt{\frac{4-\rho}{2}} \) for PIHP mobility, the first hitting time \( T_h(v_i, v_j) \) is exponentially bounded. Otherwise, \( T_h(v_i, v_j) \)
Figure 4.9: Illustration of the dissemination direction.

and thus $T_d$ have heavy tails. Therefore, we only need to study the scalability of $T_d$ with exponentially bounded $T_h(v_i, v_j)$. We next present our main result.

**Proposition 4.4.** Given $\alpha > \frac{\sqrt{1 - r}}{2}$ for a large network $[F_{m,n}, M(\Phi_E, \Psi, \alpha), M_H]$ (or $\alpha > \frac{\sqrt{1 - r}}{2}$ for $[F_{m,n}, M(\Phi_P, \Psi, \alpha), M_H]$), there exists some finite constant $\kappa$ such that $P(\lim_{D \to \infty} S_d = \lim_{D \to \infty} T_d = \kappa) = 1$.

To initiate the proof of Proposition 4.4, we first define the following notations. In Section 4.4, we argued that $D$ should be specified as Manhattan distance between the center points of the source and destination under EIHP and PIHP mobility models. Here we further denote by $d^{(1)}_c(v_i, v_j)$ the Manhattan distance between center points $v_i^c$ and $v_j^c$ for any $v_i$ and $v_j$. Let $N_h$ be the set of secondary users defined as,

$$N_h \triangleq \{v_i : d^{(1)}_c(v_s, v_i) = h\sqrt{\frac{1}{\lambda}}\},$$

where $\sqrt{\frac{1}{\lambda}}$ is the Manhattan distance between center points of neighboring secondary users (see Fig. 4.2). The information dissemination direction from the source $v_s$ to destination $v_d$ is denote by $g(v_s, v_d)$, which is the straight line joining $v_s^c$ and $v_d^c$ as shown in Fig. 4.9. We denote $v(h)$ as the secondary user whose distance to $v_s$ is $h\sqrt{\frac{1}{\lambda}}$ and which is in the information dissemination direction, i.e.,

$$v(h) \triangleq \{v_i : v_i \in N_h \text{ and } g(v_s, v_d) \cap O_i \neq \emptyset\},$$

where $O_i$ is the cell associated with $v_i^c$ as shown in Fig. 4.9. We next define the collection of indexed variables by $T_{h,q}$ as the dissemination latency from nodes $v(h)$ to $v(q)$ (thus $T_d =$
Therefore, Proposition 4.4 is equivalent to showing \( \Pr(\lim_{q \to \infty} T_{0,q} = \kappa \sqrt{\lambda}) = 1 \), which can be proved by Liggett’s theorem.

Particularly, if we can show that the sequence \( \{T_{h,q}, h \leq q\} \) satisfies the conditions \((i)- (v)\) of Liggett’s theorem, we can finish our proof. By definition, \( T_{0,q} \) is the shortest time that \( v(q) \) will receive the message from \( v(0) \), which is clearly at most \( T_{0,h} + T_{h,q} \). Condition \((i)\) is thus verified. As the latency cannot be negative, we have \( E(T_{0,q}) > 0 \). To compute an upper bound of \( E(T_{0,q}) \), we consider a Manhattan path between nodes \( v(0) \) and \( v(q) \) (see Fig.4.4) and thus have \( E(T_{0,q}) \leq qE(T_{h}(v_{i}, v_{j})) \). Theorem 4.3 tells that the first hitting time \( T_{h}(v_{i}, v_{j}) \) is exponentially bounded and thus \( E(T_{0,q}) < \infty \). This attests condition \((ii)\). Conditions \((iii)\) and \((iv)\) are clearly satisfied, as \( T_{h,q} \) is defined in a stationary way. The following lemma is to prove that the sequence \( T_{h,q} \) is ergodic, i.e., \( \{T_{h,q}, h \leq q\} \) satisfies condition \((v)\). Particularly, it shows that \( T_{h,q} \) is mixing (i.e., roughly speaking, asymptotically independent), which is a stronger property than ergodicity.

**Lemma 4.3.** The sequence \( \{T_{q,q+1}, q \geq 0\} \) is mixing.

**Proof.** We compute \( T_{q,q+1} \) by the following construction: Denote \( \mathcal{N}_{q,k} \) as the set of secondary users whose distance to \( v(q) \) is less than \( k \sqrt{\frac{1}{\lambda}} \), that is,

\[
\mathcal{N}_{q,k} \triangleq \{v_{i} : d_{c}^{(1)}(v(q), v_{i}) < k \sqrt{\frac{1}{\lambda}}\},
\]

and \( T_{q,q+1}^{(k)} \) as the transmission delay from \( v(q) \) to \( v(q+1) \) where nodes \( v \in \mathcal{N}_{q,k} \) are used as relays. Observe that

\[
\lim_{k \to \infty} \Pr(T_{q,q+1}^{(k)} < t) = \Pr(T_{q,q+1} < t)
\]

for all \( t \). Thus \( \{T_{q,q+1} \} \) is mixing by

\[
\lim_{k \to \infty} \Pr\left((T_{q,q+1} < t) \cap (T_{q+2k,q+2k+1} < t')\right) \\
= \lim_{k \to \infty} \Pr\left((T_{q,q+1}^{(k)} < t) \cap (T_{q+2k,q+2k+1}^{(k)} < t')\right) \\
= \lim_{k \to \infty} \Pr(T_{q,q+1} < t)\Pr(T_{q+2k,q+2k+1} < t') \quad \forall t, t'.
\]
The second equality follows that \( T^{(k)}_{q,q+1} \) and \( T^k_{q+2k, q+2k+1} \) are independent, as they depend on non-intersected node sets \( N_{q,k} \) and \( N_{q+2k,k} \).

Putting all together, We conclude that \( \{T_{h,q}, h \leq q\} \) satisfies all the conditions of Liggett’s theorem and thus prove Proposition 4.4. Since the proof for the scalability of \( T_d \) under PIHP mobility is similar, we omit the details. Next we study the scalability of the dissemination latency \( T_d \) under HHP mobility.

### 4.5.2 Scalability of \( T_d \) under HHP Mobility

In the previous analysis of information dissemination in large CRNs under EIHP or PIHP mobility, a fundamental assumption is that the mobility capacity \( \alpha \) is large enough, which is actually used to ensure that the networks are fully connected. That is, there exists a communication path (may be dynamic over time) between any two secondary users with high probability. Such a full connectivity requirement may be overly restrictive in large CRNs with HHP mobility. For example, it is well-known that to achieve the full connectivity in a large homogeneous networks with \( n \) static wireless nodes distributed as a Poisson process, the required transmission range is \( \Theta(\sqrt{\log n}) \) [71]. By ignoring the interference from primary users to secondary users, a large CRN under HHP with mobility capacity \( \alpha \) and transmission range \( r \) can be mapped as a static Poisson distributed homogeneous networks with transmission range \( 2\alpha + r \) [58]. This indicates that \( 2\alpha + r = \Theta(\sqrt{\log n}) \) is required for full connectivity, which is impractical to be satisfied when the number of nodes \( n \) is large.

Therefore, we study information dissemination in large percolated CRNs under HHP mobility rather than a fully connected network [57, 58, 104, 107]. The main result of percolation theory concerns a phase transition in the macroscopic behavior of large random networks [94]. Specifically, it tells that there exists a finite and positive value of the transmission range, or equivalently of the node spatial density, above which the network is percolated (supercritical) and under which the network is not percolated (subcritical). When the network is percolated,
there exists a \textit{large connected component} (usually called \textit{giant component}) of nodes spanning almost the entire network. When the network is not percolated, it consists of only small isolated components of nodes. Note that the ultimate goal of the \textit{full connectivity} requirement is to ensure that information can be disseminated throughout the entire network, which can also be achieved in percolated networks through the \textit{giant component}. But the conditions to achieve the latter are much less restrictive (e.g., finite transmission range $r$, or equivalently a finite $\alpha$). As a result, most of the networks in real world can be considered as \textit{percolated} other than \textit{fully connected}. To this end, we focus our study on how fast information is disseminated in large \textit{percolated} CRNs under HHP mobility.

The problem of percolation of a large static random CRN has been studied in [104]. The result is that given the density and interference range of primary users, a critical value can be identified on the transmission range of secondary users, above which the CRN is percolated. By mapping the CRN under HHP mobility to a large static random CRN (similar to [58]), we can derive a similar condition for percolated network easily (details omitted due to page limit). Now we focus on the dissemination latency $T_d$ and speed $S_d$ in large \textit{percolated} CRNs under HHP. By assuming that the mobility capacity $\alpha$ or transmission range $r$ is large enough to satisfy the percolation conditions, we obtain the following results on $T_d$ and $S_d$:
Proposition 4.5. Given a percolated CRN network $\mathcal{F}_{m,n}$, for any two nodes $v_s$ and $v_d$ in the giant component of $\mathcal{F}_{m,n}$, there exists a finite and positive constant $\kappa$ such that $\mathbb{P}(\lim_{D \to \infty} S_d = \lim_{D \to \infty} \frac{T_d}{D} = \kappa) = 1$.

Proof. (Sketch.) Instead of proving our results by using percolation theory from the scratch, we first map our networks to the existing models [57, 58, 104, 107] and thus make the best of use of existing findings. In particular, a similar setting is studied in [58], where a percolated wireless homogeneous network includes home points $\{v^h_i\}$ (called initial positions) are uniformly distributed and mobile nodes are independently and uniformly within the circular region centered at $v^h_i$ with radius $\alpha > 0$. It is shown that the limiting dissemination latency scales linearly with the dissemination distance $D$ (Euclidean distance between home points) in such a network. Our challenge is how to use this result in our study. We notice that there are two main differences between our model $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), \mathcal{M}_H]$ and the network used in [58]. First, in our model, a node $v_i$ is independently mobile within $A(v^h_i, \alpha)$ according to an arbitrary distribution $\Psi$ (not necessarily the uniform distribution). Second, primary users are present in our CRNs, which interfere with the communications among secondary users. However, we find that the proofs in [58] require neither uniform distribution of $v_i$ around $v^h_i$, nor non-interference from other nodes (e.g., primary users). Indeed, the fundamental requirement for proofs in [58] is that, given any two nodes $v_i$ and $v_j$ with $d^h(v_i, v_j) = \|v^h_i - v^h_j\| < 2\alpha + r$, the expected first hitting time $E(T_h(v_i, v_j)) < \infty$. Fortunately, our earlier study shown in Theorem 4.3 can satisfy this condition. Therefore, we are able to extend the results [58] to our network model $[\mathcal{F}_{m,n}, \mathcal{M}(\Phi_H, \Psi, \alpha), \mathcal{M}_H]$. This leads to the linear scaling law of information dissemination speed, which completes the proof.

Remark 4.9. Propositions 4.4 and 4.5 together complete the proof for Theorem 4.2, which demonstrates that information disseminates linearly in large connected (full connected or per-

\[2\] [57] defines two types of percolation for mobile networks: permanent percolation and cumulative percolation. The former is that the network is percolated at any time instant; the latter is that the network is not percolated at any time instant, but it is percolated over time. In this chapter, we use the latter definition for percolated mobile networks.
colated) CRNs under general mobility. More importantly, such a linear scaling property is invariant to the density of primary users, even though the specific values are negatively affected. For example, the asymptotic values of the dissemination speed are proportional to the expected first hitting time $E(T_h(v_i, v_j))$ between neighboring secondary users $v_i$ and $v_j$, which implies that the density of primary users (see Theorem 4.3) has an adverse impact on the dissemination speed. In other words, when primary users are dense, the communications among secondary users may be undermined, leading to slow dissemination speed. This is in agreement with our intuition. Furthermore, as shown in [104] and Fig. 4.11a, the required transmission range $r$ (or mobility capacity $\alpha$) for percolation increases as the density of the primary users increases. When the density of the primary users is larger than some critical value, the CRN is not percolated even with an infinitely large $r$ or $\alpha$. Therefore, our dissemination speed analysis in Proposition 4.5 is applicable to networks in which the density of primary users is not high such that the network is percolated.

![Percolation regions](image1)
(a) Percolation regions.

![An example of percolated CRN](image2)
(b) An example of percolated CRN.

![Dissemination speed](image3)
(c) Dissemination speed.

Figure 4.11: Percolation conditions and the dissemination speed in large percolated CRNs.

A series of simulations have been performed to further validate our theoretical results concerning the asymptotic latency. In these simulations, time is partitioned into unit slots. In each time slot, secondary users are independently and uniformly distributed around their home points. Fig. 4.10 shows the dissemination speed in CRNs under EIHP and PIHP models, re-
spectively, where the transmission range $r$ of secondary users and the interference range $R_I$ of primary users are set as $r = 0.1$ km and $R_I = 0.3$ km, and the spatial densities of secondary users and primary users are $\lambda = 4$ (per km$^2$) and $\lambda_p = 0.5$ (per km$^2$), respectively. As shown in Fig. 4.10a and 4.10b, no matter how large the mobility capacity $\alpha$ is, the dissemination latency $T_d$ scales linearly with the dissemination distance $D$ (Manhattan distance) as $D$ increases, which is in good accordance with Proposition 4.4. The dissemination speed in large percolated CRNs under HHP mobility is shown in Fig. 4.11. Particularly, Fig. 4.11a shows the simulated percolation conditions under different network parameters and Fig. 4.11b gives an example of such a percolated CRN. Fig. 4.11c shows that in percolated CRNs, the dissemination latency $T_d$ scales linearly with the dissemination distance $D$ (Euclidean distance) as $D$ increases. These observations provide a straightforward illustration of Proposition 4.5 and arguments in Remark 4.9 with regard to the information dissemination speed in large cognitive radio networks.

### 4.6 Summary

In this chapter, we have studied the distribution of the information dissemination latency $T_d$ in finite CRNs and the scalability of $T_d$ in large CRNs under general mobility. We found that in finite networks, there exists a cutoff point on the mobility capacity $\alpha$ of secondary users, above which the tail distribution of $T_d$ is bounded by some Gamma distribution and below which $T_d$ has a heavy-tailed distribution. When networks become large, the dissemination latency $T_d$ is (linearly) scalable with respect to the dissemination distance. Our results demonstrate that when secondary users can move in a large region, a Gamma distributed (light-tailed) latency in finite networks, or a scalable latency in large networks, is achievable, which encourages the deployment of CRNs for real-time and large applications.
Besides node mobility, node failure is another commonly observed threat to network topology. It has been demonstrated that in wireless networks, Blackholes, which are components of failed nodes, can easily result in devastating impact on network performance. Such Blackholes are typically generated by isolated node failures, and augmented by failure correlations due to traffic overloading, energy depletion, interrupted communication links, etc. In multihop wireless networks, many solutions, such as routing protocols and detection and restoration algorithms, are proposed to deal with Blackholes by locating and identifying alternative paths to bypass these holes such that the effect of Blackholes can be mitigated or even eliminated. These advancements are proved to be effective either in theory or simulations, which are based on an underlying premise that there exists at least one alternative path in the network. Even though it is widely adopted in the study of Blackholes in wireless networks, such a hypothesis remains an open question. In other words, we do not know whether the network is resilient to Blackholes or whether an alternative path exists. The answer to this question can complement our
understanding of designing routing protocols, as well as topology evolution in the presence of random failures. In order to address this issue, we focus on the topology of Cognitive Radio Networks (CRNs) because of their phenomenal benefits in improving spectrum efficiency through opportunistic communications. Particularly, we first define two metrics, namely the failure occurrence probability $p$ and failure connection function $g(\cdot)$, to characterize node failures and their spreading properties, respectively. Then we prove that each Blackhole is exponentially bounded based on percolation theory. By mapping failure spreading using a branching process, we further derive an upper bound on the expected size of Blackholes. With the observations from our analysis, we are able to find a sufficient condition for a resilient CRN in the presence of Blackholes through analysis and simulations.

5.1 Motivation and Related Work

Wireless communication has experienced an explosive growth in the past few decades, which imposes a significant demand for the already-crowded radio spectrum. However, a recent report by the Federal Communications Commission (FCC) indicated that over 90% of the licensed spectrum remains idle at a given time and location [1]. This observation immediately incurs considerable attentions to Cognitive Radio Networks (CRNs), which show great potential for improving spectrum usage efficiency [2] by permitting secondary networks to coexist with licensed primary networks. On one hand, many efforts have been devoted to understanding the performance limits of CRNs, including maximum capacity, minimum delay and connectivity [108–111]. These works have presented a very good understanding of the potential of CRNs for a variety of applications in theory. On the other hand, the properties and dynamics of global topology, which plays an important role in designing fundamental networking functionalities, such as point-to-point routing and scheduling algorithms, has never been well studied. The lack of knowledge about network topology greatly hinders the practical deployment of CRNs, which motivates the study on topological features of CRNs in this chapter.

Topology of wireless networks changes frequently due to different factors (e.g., node mobility,
failures) and in this chapter, we focus on topological transmutation by studying Blackholes due to node failures. Such unavoidable faults can be brought out by malfunctions of electrical devices, energy depletion, natural disasters (fire, river overflow, earthquake, etc) or adversarial attacks (a bomb explosion for example). Communications may be disabled by jamming, traffic congestion or energy depletion. In addition, causal relations often exist among failures, i.e., some failures happen as a result of other earlier failures. One example of such correlated failures is traffic overloading and energy depletion [14], that is, when a node fails to deliver packets, the incoming and outgoing traffic is redistributed to the neighboring nodes. Some neighbors may work under heavy traffic loads, resulting in early energy depletion and node failures. Such correlation among failures and cascading effects lead to Blackholes (i.e., components of failed nodes, see formal definition in Section 5.2) in the network, where information cannot be transmitted or forwarded.

Understanding the properties of Blackholes in the CRNs, or in particular, investigating structure and size of Blackholes, is of great importance in the design of basic networking operations. For example, a number of networking protocols exploit geometric intuitions for simple and scalable data delivery, such as geographical greedy forwarding [112,113]. These algorithms based on local greedy advances may not work properly in the presence of Blackholes, where routing messages will be lost. Backup and restoration methods, such as face routing on a planar subgraph, can help packets get out of Blackholes, but also create high traffic on hole boundaries and eventually undermine network lifetime [112,113]. In addition, a number of routing schemes address explicitly the importance of topological properties and propose routing with virtual coordinates that are adaptive to the intrinsic geometric features [114]. However, constructing these virtual coordinate systems requires the identification of topological features, especially Blackholes first in order to proceed routing.

Therefore, Blackholes have been extensively explored in wireless networks [78,115,116]. For example, Fang et al. [115] studied the difficulties imposed by Blackholes on geographic routing and proposed a distributed algorithm to build a path bypassing such holes. Wang et al. [116]
focused on topology discovery and presented an algorithm to identify Blackholes. The spatial features of the holes and their impact on data preservation have been investigated in [78]. These results significantly improve our understanding of the disadvantageous impact of Blackholes on network performance.

Meanwhile, it is evident that all of these studies presume a few interesting but more fundamental questions. First, what is the driving force in the formation of Blackholes and how large are these holes? Failure correlation [78, 79] has been recognized as one of the most important factors for the occurrence of Blackholes and Xu et al. [79] further studied how an initial failure may incur a giant hole spanning over the entire network. Given its detrimental consequences, the occurrence of giant hole needs to be avoided in the initial network design [79] such that node failures can result in many finite holes in the network. However, how to quantify the finite size of these holes has not been discussed. Moreover, existing works [78, 115, 116] are focused on locating and bypassing these holes in the network. But a fundamental question is whether the network is resilient to these Blackholes, i.e., whether we can always find alternative paths to bypass all holes. If such paths do not exist, routing protocols may not be a good solution, which is a fundamental issue in multihop networks.

In this chapter, we aim to provide insightful understanding of the above questions. In particular, we first study the process of how an initial failure “explodes” to a Blackhole and present theoretical analysis to quantify the scope of Blackholes. Using combinatorial arguments, we prove that the distribution of Blackhole size decays exponentially and we further provide an upper bound on the expected size of Blackholes by mapping failure spreading to a branching process. Then we investigate network resilience in the presence of Blackholes. A network is said to be resilient to node failures when there exists a large connected component of “surviving” (not failed) nodes spanning over the entire network. We have identified a sufficient condition for a resilient CRN against Blackholes by using techniques in percolation theory [94].

Our contributions to the understanding of topological resilience are as follows:

- We investigate the formation of Blackholes due to explosive spreading of random failures,
and prove that each Blackhole is exponentially bounded and provide an upper bound on its expected size.

- We identify a sufficient condition when a CRN is resilient to blackholes, which can be used as a prerequisite for the blackhole locating and bypassing algorithms in the existing works [78, 115, 116].

Although we only addressed topological features and resilience of CRNs, questions presented in this chapter are important yet remain unanswered in general multihop networks (e.g., wireless sensor networks and wireless ad hoc networks). Letting spatial density of primary users $\lambda_p = 0$, our results can be extended to other wireless multihop networks, which serves a timely complement to existing studies on restoration algorithms and protocols [78, 115, 116].

5.2 System Models and Problem Formulation

In this section, we first present a brief description of preliminaries, then describe the network models, basic assumptions and notations, and formulate the problem last.

5.2.1 Preliminary

Before introducing network models, we need a brief introduction of common models and tools used to study wireless networks for clarification. A continuum graph consisting of nodes $\mathcal{X}$ placed in space $\mathbb{R}^2$, with edges added to connect pairs of nodes which are close to each other, can be used to model wireless networks [79, 96, 108–110]. Rather than any specific positions, nodes $\mathcal{X}$ are usually assumed to be a Poisson point process for the following reasons. First, precise configuration of points may not be known. In addition, Poisson point process represents an average case. Some properties of graphs are unfeasible to compute for large graphs, and understanding their average behavior may be a useful alternative to exact computation. For example, given a Poisson point process $\mathcal{X} \subset \mathbb{R}^2$, the graph, denoted by $\mathcal{G}(\mathcal{X}, r)$, with vertex set $\mathcal{X}$ and edges connecting those pairs $\{x_1, x_2\} \in \mathcal{X}$ with $|x_1 - x_2| \leq r$, is called Boolean model.
and has been used in [96] to represent a large wireless network. If edge between $x_1$ and $x_2$ is added with probability $g(|x_1 - x_2|)$, the resulted graph $G(\mathcal{X}, r, g)$ is called *random connection model* and has been used in [79] to study failure spreading.

*Percolation theory*, which studies macroscopic structure of a large graph $G$, is a very useful tool to investigate topology of wireless networks [14, 79, 96, 108–110]. Particularly, a graph $G$ is composed of one or many disjointed components. Let $W \subset G$ denote the largest component of $G$ and $|W|$ be the cardinality of $W$. Percolation theory investigates the condition under which $W$ spans over the entire graph $G$, i.e., $|W| = \infty$ ($W$ is called *giant component* in this case). The main result of percolation theory is that there exists a finite, positive value of the node spatial density, above which the network is percolated ($P(|W| = \infty) = 1$), and below which it is not ($P(|W| = \infty) = 0$). This is called *percolation phenomenon*.

### 5.2.2 Network and Failure Models

In this chapter, we consider a large CRN consisting of $n$ secondary users $\{v_1, \ldots, v_n\}$, which are distributed independently and uniformly in a region $\Omega = [0, \sqrt{\frac{2}{\lambda}}]^2$ for some constant $\lambda$. Let $\mathcal{H}_\lambda = \{x_1, \ldots, x_n\}$ denote the random locations of secondary users and $\mathcal{H}_\lambda$ is a Poisson Point process with density $\lambda$ as $n \to \infty$ [93]. A set of $m$ channels $\{ch_1, \ldots, ch_m\}$ are assumed to be accessible by secondary users. For any $1 \leq k \leq m$, an overlay network of primary users with spatial density $\lambda_{pk}$ are transmitting with channel $ch_k$. We assume that $\lambda_{pk} = \lambda_p$ for any $k$ for simplicity. To model the dynamics of the primary traffic, we adopt a synchronized slotted structure, which has been used in [108] to study the connectivity of a large single-channel CRN. Particularly, time is slotted into units and at any time slot, primary users transmitting on any channel $ch_k$ are assumed to be uniformly and independently distributed in $\Omega_n$, and such distribution is i.i.d across slots.

124
Interference Models

In CRNs, there are two types of interference for information dissemination among secondary users: secondary-secondary and primary-secondary interference. The former interference can be characterized by the well-known protocol model [71]. Particularly, without interference from primary users, a successful transmission from a secondary user $v_i$ to $v_j$ is achievable if $\|x_i - x_j\| \leq r$ and for any other simultaneously transmitting node on the same channel $v_l$, $\|x_l - x_j\| \geq (1 + \Delta)r$, where $r$ is the transmission radius of secondary users, and $\Delta$ models the guard zone around $v_j$ in which any simultaneous transmission on the same channel causes collision at $v_j$. For the latter interference, denote $R_I$ as the interference range of primary users. And as Fig. 5.1 shows, two secondary users $v_i$ and $v_j$ are permitted to use channel $ch_k$ only when there are no primary users on $ch_k$ in the neighborhood, i.e., $\|x_i(t) - u(t)\| > R_I$ for any primary user $u$ transmitting with $ch_k$, where $u(t)$ is the position of $u$ at time $t$.

Failure Model and “Explosion”

In wireless networks, nodes fail unavoidably due to adversary attacks, natural hazards, resource depletion, etc. Node failures are often not independent and causal relations exist among these failures, i.e., some failures happen as a result of other earlier failures. Traffic overloading and energy depletion [14] is an example as a result of failures spreading. Because of failure correlation, each initial failure will “explode” and impact a component of nodes in the neighborhood.
An illustration of such process is shown in Fig. 5.2. In this example, random failures initially occur at nodes $v_1$, $v_5$, $v_8$, $v_{12}$, $v_{14}$ and $v_{15}$. As a result of the failure on $v_1$, node $v_2$ fails subsequently and spreads the failure further away to nodes $v_3$ and $v_4$. Similarly, nodes $v_6$, $v_7$, $v_9$, $v_{10}$, $v_{11}$, $v_{13}$, and $v_{16}$ fail subsequently due to random failures on $v_5$, $v_8$, and $v_{12}$. In Physics, a Blackhole is a region of spacetime from which nothing, not even light, can escape [117]. Likewise, in a wireless network, any information (e.g., routing packets) transmitting to components of failed nodes will be absorbed (lost). This similarity motivates us call a component of failed nodes incurred by a particular initial failure (see the shaded area in Fig. 5.2) a Blackhole for convenience.

The above example shows that the formation of Blackholes consists of the occurrence of initial failures and explosion of these failures. Thus we introduce the following models:

- Random failure model: each node is either surviving or failed independently and a node may fail with probability $p$ (failure occurrence probability). This model describes the initial occurrence of node failures.

- Failure explosion: We define failure connection function $g(\cdot)$ to model the likelihood of failure propagation from $v_i$ to $v_j$. If $|x_i - x_j| < r$, failure spreads from $v_i$ to $v_j$ with a probability $g(|x_i - x_j|)$ that depends on their distance but not their respective locations. If $v_j$ is beyond the transmission radius of $v_i$, failure cannot spread from $v_i$ to $v_j$ directly.

In this chapter, we assume that $g(\cdot) \equiv \tau$, which is called failure connection probability and
$r = 1$ by default, if there is no specific explanation. Thus failure spreading among secondary users can be represented as a random connection model $G(\mathcal{H}_\lambda, 1, \tau)$.

**Remark 5.1.** These two models are not new. Particularly, random failure model has been used in [96] to study topology transition of wireless networks because of independent node failures (without considering failure spreading) and failure connection function has been used in [79] to determine whether an initial failure will spread to the entire network. However, as discussed above, the occurrence of random failures and their subsequent explosion are inseparable, and we are interested in this chapter how these two processes together result in Blackholes in the network.

### 5.2.3 Problem Formulation

In order to understand the impact of Blackholes on network implementation (e.g., routing), we first focus on a particular hole, initiated by a failure on a node, say $v_1$, w.l.o.g, and denoted by $O_{v_1}$. Existing results on random connection model [94] shows that there exists some critical value $\zeta$ on node density $\lambda$, such that if $\lambda > \zeta$, $O_{v_1}$ may spread to the entire network with some positive probability; and if $\lambda < \zeta$, $O_{v_1}$ is finite. Given the devastating consequence of large-scale failure spreading, previous work in [79] provided bounds on $\zeta$, which helps network designers to operate network at $\lambda < \zeta$, making network be resilient to cascading failures. In this chapter, we are interested in when $\lambda < \zeta$, how large $O_{v_1}$ is.

**Definition 5.1.** *(BHG problem):* For a CRN which is resilient to large-scale failure spreading (i.e., $\lambda < \zeta$), how large does a Blackhole Grow?

**Definition 5.2.** *(Blackholes Resilient, BHR):* Given the existence of Blackholes, a CRN is said to be BHR if a giant component of surviving nodes, well spanning over the entire network, exists.

BHR property makes a CRN maintain global communication capability in the presence of Blackholes, i.e., information can bypass Blackholes and be disseminated to the entire network.
through the giant component. And BHR property provides theoretical foundation to the existing
studies on locating and bypassing Blackholes [78, 115, 116]. Next, we will formally define BHR
problem.

**Definition 5.3. (BHR problem).** For a large CRN which is assumed to be initially connected
(percolated), given node failures characterized by random failure model and failure explosion
model, determine the condition under which the network is BHR.

5.3 Results and Applications

In this section, we present our main results concerning BHG and BHR problems. We find that
the size of Blackholes is exponentially bounded and provide an upper bound on their expected
size. Based on the understanding of size of Blackholes, we further identify a sufficient condition
for a resilient CRN. In addition, we further discuss potential applications of our theoretical
analysis.

5.3.1 Main Results

We summarize our main results as follows. First, following theorems solve the BHG problem.

**Theorem 5.1. Exponential decay of $|O_{v_1}|$.** When Blackhole $O_{v_1}$ is not percolated, there
exists some $\epsilon > 0$ such that

$$\mathcal{P}(|O_{v_1}| \geq N) \leq e^{-N\epsilon} \text{ for all } N \text{ sufficiently large.} \quad (5.1)$$

**Remark 5.2.** Theorem 5.1 shows that when a failure cannot spread to the entire network,
the number of nodes that may be infected by this failure is exponentially bounded. Exponential
distribution is not enough to show how large Blackhole $O_{v_1}$ is, since the expected value $E(|O_{v_1}|)$
of $|O_{v_1}|$ is unidentified, i.e., the parameter $\epsilon$ in Eq. (5.1) is unknown. We provide $E(|O_{v_1}|)$ in
the next theorem.
**Theorem 5.2.** When Blackhole $O_{v_1}$ is not percolated, its expected size is upper bounded by

$$\beta = E(|O_{v_1}|) \leq \frac{1.43\pi\lambda^2\tau^2}{1 - 1.43\lambda\tau} + 1, \quad (5.2)$$

where $\lambda$ is spatial density of secondary users and $\tau$ is failure connection probability.

**Remark 5.3.** Theorem 5.2 indicates that the expected Blackhole size grows as failure connection probability $\tau$ increases, which corresponds to our intuition that the Blackhole is large when nodes are prone to be infected by their neighbors. Eq. (5.2) further implies that $1 - 1.43\lambda\tau > 0$ is necessary to guarantee that Blackhole $O_{v_1}$ is not percolated.

Theorems 5.1 and 5.2 study distribution and expected size of Blackhole $O_{v_1}$, and Eqs. (5.1) and (5.2) together answer the question that how large Blackhole is (BHG problem). Note that reasons incurring failure correlation (e.g., due to traffic overloading and energy depletion) are usually independent of primary users and thus our failure models (see Section 5.2.2) do not take primary users into account. That is, our failure models in CRNs are similar to those used in general wireless networks [79, 96], which implies that our results concerning Blackhole size can be applied to general wireless networks directly. In previous work [79], Xu et al. prove a value $\zeta$ such that when node density $\lambda > \zeta$, a failure is percolated, and it is not percolated otherwise. And our results further illustrate the size of nodes infected by a failure when $\lambda < \zeta$, which is an important and necessary complement to the existing work.

The next theorem answers the BHG problem, providing a sufficient condition for a resilient CRN in the presence of Blackholes.

**Theorem 5.3.** Given a CRN where each secondary node fails with probability $p$ according failure models defined in Section 5.2.2, it is BHR if $p < 1 - \frac{\Lambda e^{a_d}}{1 - e^{-p\alpha} + \Lambda}$, where

$$\Lambda = \sqrt{\frac{p_c^\Box}{(1 - e^{\lambda d_l^2})^2 (1 - (1 - e^{-\lambda p\alpha})^m)}}, \quad (5.3)$$

$d_l = \frac{r}{\sqrt{5}}$, $\alpha = (d_l + 2R_I)(2d_l + 2R_I)$ and $p_c^\Box$ is given in Section 5.7.
Remark 5.4. BHR problem is not only important in CRNs, it also remains unanswered in general wireless networks. Setting spatial density of primary users $\lambda_p = 0$, Theorem 5.3 also provides BHR condition for a general wireless network.

5.3.2 Applications

Besides the theoretical importance of our findings, our results can be used practically not only in the initial deployment, but also as a theoretical foundation in evaluating protocol designs. Here are some examples.

- In the initial deployment, an appropriate value for spatial density $\lambda$ of wireless nodes can be decided to guarantee that any random failure can only spread within a predefined area, if failure connection probability $\tau$ is known.

- There are many routing protocols [78, 115, 116] proposed to identify a path to bypass Blackholes through the entire network. However, when the network is not BHR, such path does not exist and thus these protocols will not work properly and waste network energy. Our result concerning BHR can be used as a prerequisite in determining whether adopt these routing protocols, or as a benchmark in evaluating the efficiency of these protocols.

- In wireless networks, node failures affect the communication connectivity and in turn impair network functionality. Thus as mentioned in [118], redeploying additional nodes is necessary to replace failed nodes so that a connected network topology can be maintained. Let $T_i$ ($1 \leq i \leq n$) denote the lifetime of node $v_i$ before it is failed. Given survival function $S(t) \triangleq \mathcal{P}(T_i > t)$ (thus failure occurrence probability $p = 1 - S(t)$), our results provide network designers a guideline on the optimal time that the redeployment of additional nodes should be carried out.
5.4 How Large is a Blackhole?

In this section, we demonstrate how to obtain the results concerning size of Blackhole $\mathcal{O}_{v_1}$ given in Section 5.3. Specifically, we investigate how many nodes will be infected by occurrence of failure on $v_1$. We first study the distribution of $|\mathcal{O}_{v_1}|$.

5.4.1 The Distribution of $|\mathcal{O}_{v_1}|$

Using percolation theory, Xu et al. [79] determine the condition under which $\mathcal{O}_{v_1}$ may be percolated to the entire network. However, when $\mathcal{O}_{v_1}$ is not percolated, how large $|\mathcal{O}_{v_1}|$ is, remains unknown. To study distribution of $|\mathcal{O}_{v_1}|$, our approach takes following procedures. We first map failure spreading process defined on continuous plane onto a discrete lattice, whose edges are declared open if certain properties are met (closed otherwise). In the discrete lattice, we then investigate the size of components consisting of open edges using combinatorial arguments. With a careful definition on the open edge in the lattice, a relation between the size of Blackholes and size of components of open edges can be derived. Finally, we obtain the distribution of Blackhole size $|\mathcal{O}_{v_1}|$ in Theorem 5.1. The detailed proof is presented as follows.

Proof of Theorem 5.1. When studying topology of continuum graph, an useful technique is the discretization of the graph on $\mathbb{R}^2$ into lattice on integer space $\mathbb{Z}^2$, since topological properties of the latter is easier to be analyzed [99]. One of the technical uses of such a discretization lies in the availability of combinatorial arguments for enumerating the sets in $\mathbb{Z}^2$. To proceed, we shall require a variety of notations. A set $A \subset \mathbb{Z}^2$ is said to be symmetric if $-x \in A$ for all $x \in A$. Vertices $x, y \in A$ are said to be $A$-adjacent ($x \sim_A y$) if and only if $y - x \in A$. A subset $S \subset \mathbb{Z}^2$ is $A$-connected if it induces a subgraph with adjacency relation $\sim_A$. The following lemma, which says that the number of $A$-connected subsets of $\mathbb{Z}^2$ of size $N$ containing the origin grows at most exponentially, is helpful.

Lemma 5.1. (Peierls argument, see page 178 in [93]) Let $A$ be a finite symmetric subset of $\mathbb{Z}^2$ with $|A|$ elements. The number of $A$-connected subsets of $\mathbb{Z}^2$ containing the origin, of
cardinality $N$, is at most $2^{|A|N}$.

In this chapter, we consider a discrete lattice $\mathcal{L} = d_l \times \mathbb{Z}^2$ with side length $d_l$. The coordinates of the vertices of $\mathcal{L}$ are $(d_l \times i, d_l \times j)$ for $(i, j) \in \mathbb{Z}^2$. Adjacency is defined by $A = \{ z \in \mathcal{L} : \|z\|_1 = d_l \}$ where $\| \cdot \|_1$ denotes 1-norm distance, i.e., an edge connects $x, y \in \mathcal{L}$ only when $\|x - y\|_1 = d_l$ (see solid lines in Fig. 5.3 and 5.4). For any $z \in \mathcal{L}$, we construct a box $B_z$ of size $d_l$ centered at $d_l \times z$ (see the dash lines in the Fig. 5.3 and 5.4). As Fig. 5.3 shows (for figures in this chapter, solid dots and circles denote failed and surviving nodes respectively), failure spreading, represented by random geometric graph $G(\mathcal{H}_\lambda, r, \tau)$ (i.e., graph consisting of failed nodes and edges connecting them), induces a realization of the bond percolation on $\mathcal{L}$ by setting an arbitrary bond $zz' \in \mathcal{L}$ to be open if there exists an edge $uv \in G(\mathcal{H}_\lambda, r, \tau)$ such that $u \in B_z$ and $v \in B_{z'}$. That is, given one or more failed nodes in $B_z$, at least one failed node connects to some some nodes in $B_{z'}$. And an example of open bond $zz'$ is shown in Fig. 5.3. Let $C(v_1)$ denote the cluster of open bonds and $|C(v_1)|$ denote its size. It is obviously true that if $|O_{v_1}| < \infty$, then $|C(v_1)| < \infty$, and vice versa. The mapping between the cluster of failed nodes and the cluster of open bonds allows us to find $|C(v_1)|$ and thus use it to study $|O_{v_1}|$.

Particularly, when $O_{v_1}$ is not percolated, $C(v_1)$ is not percolated. Bond percolation on discrete lattice (see Theorem 6.75 in [99]) shows that if $C(v_1)$ is not percolated, then there exist
constants $\mu > 0$, $n_0 > 0$ such that

$$\mathcal{P}(|C(v_1)| \geq N) \leq e^{-\mu N}, \quad N \geq n_0. \quad (5.4)$$

By Peirels argument (Lemma 5.1), there is a constant $\gamma$ such that, for all $N$, the number of open path of $L$ of cardinality $N$ containing the origin is at most $\gamma^N$. If $|C(v_1)| < N$ and $|O_{v_1}| > KN + 1$, then for at least one of these open paths, the union of associated boxes $B_z$ contains at least $KN$ nodes of $\mathcal{H}_\lambda$ (an example of such path and its associated boxes are shown in Fig. 5.4 as the bold line and shaded area). Therefore, we have

$$\mathcal{P}[\{|C(v_1)| < N\} \{ |O_{v_1}| > KN + 1\}] \leq \gamma^N \mathcal{P}[Po(N\lambda d^2_l) \geq KN], \quad (5.5)$$

where $Po(\cdot)$ denotes Poisson distribution. To continue, we need the following lemma (see (1.12) in [93]).

**Lemma 5.2.** Let $Po(\lambda)$ be a Poisson random variable with density $\lambda$. If $K > e^2\lambda$, then

$$\mathcal{P}[Po(\lambda) \geq K] \leq e^{-(\frac{K}{2})\log(\frac{K}{\lambda})}. \quad (5.6)$$

Letting $K \geq e^2d^2_l\lambda$ and putting Eq. (5.6) into Eq. (5.5), we have

$$\mathcal{P}[\{|C(v_1)| < N\} \{ |O_{v_1}| > KN + 1\}] \leq \gamma^N e^{-(\frac{KN}{2})\log(\frac{KN}{d^2_l\lambda})}. \quad (5.7)$$

If we take $K$ sufficiently large, we see from Eqs (5.4) and (5.7) that $\mathcal{P}(|O_{v_1}| > KN + 1)$ decays exponentially in $N$, so that Eq (5.1) follows.

After proving exponential distribution of Blackhole $O_{v_1}$, we next study its expected size.
5.4.2 The Expected Value of $|O_{v_1}|$

In this subsection, we investigate the expected number of nodes in Blackhole $O_{v_1}$ and prove the upper bound Eq. (5.2) given in Theorem 5.2. Specifically, we model failure spreading in CRNs as a branching process [119]. By studying the number of offspring in this branching process, we obtain our result. The detailed proof is given as follows.

Proof of Theorem 5.2. Denote our network with a graph $G(H_{\lambda}, 1, \tau)$. Let $x_1, x_2, \ldots$ be the points of the Poisson process $H_{\lambda}$ and assume that a failure initially occurs to $x_1$ (thus $x_1$ is initial member of the 0-th generation of the branching process, as shown in Fig. 5.5). The children of $x_1$ in this branching process are points which can be infected by $x_1$ directly. According to failure spreading model in Section 5.2.2, each point of $H_{\lambda}$ which lies in the ball $B(x_1, 1) = \{y \in \mathbb{R}^2 : |y - x_1| \leq 1\}$ (see the big circle in Fig. 5.5) may be a child of $x_1$ with probability $\tau$. If we take another Poisson process $X_1$ with density $\lambda \cdot \tau$, independent of $H_{\lambda}$ and let $x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}$ be all the points of $X_1$ which lie in the ball $B(x_1, 1)$, the children of $x_1$ in the branching process are equivalent to these points $x_{1,1}, \ldots, x_{1,n_1}$ by thinning theorem [94].

Let $x_{k,1}, x_{k,2}, \ldots, x_{k,n_k}$ be the members of the $k$-th generation of the branching process. To obtain the children of $x_{k,i}$, we consider a Poisson point process $X_{k+1,i}$ of density $\lambda \cdot \tau$ on $\mathbb{R}^2$, where $X_{k+1,i}$ is independent of all the processes described as yet. The children of $x_{k,i}$ are those points of the process $X_{k+1,i}$ which fall in the region $B(x_{k,i}, 1) \setminus B(x_{k-1,j}, 1)$ (see the shaded area...
in Fig. 5.5), where $x_{k-1,j}$ is the parent of $x_{k,i}$. The *type* of a child is defined as the distance between this child and its parent. For example, the type of $x_{k,i}$ is defined $|x_{k-1,j} - x_{k,i}| \in (0, 1)$ (e.g., the length of the solid line in Fig. 5.5). Clearly, the distribution of the number and types of children of $x_{k,i}$ depend only on $x_{k,i}$ and its type. Indeed, the distribution of the number of children of $x_{k,i}$ whose types lie in $(a, b)$, $0 \leq a < b \leq 1$ depends only on the area of the region $(B(x_{k,i}, 1) \setminus B(x_{k-1,j}, 1)) \cap \{ y : |y - x_{k,i}| \in (a, b) \}$, and this area depends on $x_{k-1,j}$ only through the distance $|x_{k-1,j} - x_{k,i}|$, which is precisely the type of $x_{k,i}$. Also, the distribution of the number and types of children of an individual $x_{k,i}$ does not depend on its generation $k$.

Given that $x_{k,i}$ is of type $h$, i.e., $|x_{k,i} - x_{k-1,j}| = h$, let $f(w|h)$ be the length of the curve given by $(B(x_{k,i}, 1) \setminus B(x_{k-1,j}, 1)) \cap \{ y : |y - x_{k,i}| = w \}$. A precise expression for $f(w|h)$ follows from an elementary trigonometric calculation, which yields

$$f(w|h) = \begin{cases} 2w \cos^{-1} \frac{1-h^2-w^2}{2hw} & \text{if } 1-h < w < 1 \\ 0 & \text{if } 0 < w \leq 1-h \end{cases}$$

Recalling our earlier discussion on the independence properties of the offspring distribution, we easily see that the expected number of children whose types lie in $(a, b)$ of an individual whose type is $h$ is given by $\int_a^b \lambda \tau f(w|h)dw$. Moreover, given that an individual is of type $h$, the
expected total number of grandchildren of this individual whose types lie in \((a,b)\) is given by

\[ \int_0^1 (\int_a^b \lambda^2 \tau^2 f(w|t)dw) f(t|h)dt. \]  \hfill (5.8)

In other words, if we let

\[ f_1(w|h) = \int_0^1 f(w|t)f(t|h)dt, \]

the integral in (5.8) reduces to

\[ \lambda^2 \tau^2 \int_a^b f_1(w|h)dw. \]

Thus defining recursively,

\[ f_i(w|h) = \int_0^1 f_{i-1}(w|t)f(t|h)dt, \]

we easily see that the expected number of members of the \(n\)-th generation having types in \((a,b)\) coming from a particular individual of type \(h\) as an ancestor \(n\) generations previously is given by

\[ \lambda^i \tau^i \int_a^b f_i(w|h)dw. \]

Hence the expected total number of individuals in the branching process if we start off with an individual of type \(h\) is

\[ \sum_{i=1}^{\infty} \lambda^i \tau^i \int_0^1 f_i(w|h)dw. \]  \hfill (5.9)

The node density \(\lambda\) is small enough to make Eq. (5.9) converge by the assumption that failure is not percolated. To estimate Eq. (5.9), we define

\[ T(h) = \int_0^1 f(w|h)dw. \]

It is easy to see that

\[ \int_0^1 f_i(w|h)dw = T^i(h). \]
Thus Eq. (5.9) reduces to
\[
\sum_{i=1}^{\infty} \lambda^i \tau^i T^i (h). \tag{5.10}
\]

By using Hilbert-Schmidt operator and standard numerical methods of calculating eigenvalues (see page 87 of [94]), we can show that \( T(h) < 1.43 \). Thus Eq. (5.9) reduces to
\[
\sum_{i=1}^{\infty} \lambda^i \tau^i T^i (h) \leq \sum_{i=1}^{\infty} \lambda^i \tau^i 1.43 = \frac{1.43\lambda\tau}{1 - 1.43\lambda\tau}. \tag{5.11}
\]

Come back to the 0-generation node \( x_1 \). By *thinning theorem* [94], the expected number of children of \( x_1 \) is \( \pi \lambda \tau \). Note that the expected total number of individuals starting of any child \( x_{1,j} \) of \( x_1 \) is upper bounded by Eq. (5.11), thus the expected number of nodes in each hole is upper bounded by Eq. (5.2). This completes the proof.

Now we have proved distribution and expected size of Blackhole \( O_{v1} \) given in Theorems 5.1 and 5.2, which solve the BHG problem. Next, we formally solve the BHR problem defined in Section 5.2.3, i.e., given exponentially distributed Blackholes, does there exist a giant component of surviving nodes spanning over the entire network?

### 5.5 Is a Large CRN Blackhole Resilient?

In this section, we study the macroscopic structure of a large CRN in the face of Blackholes and formally prove the sufficient condition for a BHR network addressed in Theorem 5.3. As mentioned in Section 5.2.1, *percolation theory* [94] is an useful tool to investigate topology of wireless networks. For example, by using *percolation theory*, Sun *et al.* [109] study the connectivity of a large CRN without failures, and identify a critical density \( \lambda_c \), above which (i.e., node density \( \lambda > \lambda_c \)) there exists a giant component of nodes. Xu *et al.* [79] prove a value \( \zeta \) such that when node density \( \lambda < \zeta \), a failure can only spread among a finite number of nodes. In this chapter, we are interested in the scenario that \( \lambda_c < \lambda < \zeta \). That is, the CRN is percolated initially. As time goes on, random failures may occur and each failure may infect a finite number
of neighboring nodes, i.e., a sequence of Blackholes may appear. An interesting question is that in the event of Blackholes, whether the network remains percolated, or the giant component breaks into many small components. Next, we target on answering this question.

5.5.1 Challenges and Differences with Earlier Work

As mentioned earlier, node failures and their impact on network topology have been studied in [79, 96]. Particularly, under a subtle assumption of no failure correlations, Xing et al. [96] provide a condition for a percolated network when random failures may occur independently. On the other hand, Xu et al. [79] focus on a particular failure and study the condition when this failure may spread to the entire network due to failure correlations. Note that existing results in percolation theory [93, 94] are based on the fundamental assumption that the nodes are distributed as a Poisson point process. Thinning theorem [93, 94] ensures that nodes after failures in [79, 96] are still distributed as Poisson point process and hence existing results can be applied directly. For example in [96], nodes are initially distributed as a Poisson point process with spatial density $\lambda$ and each node may fail independently with probability $p$. By Thinning theorem, the resulted network of surviving nodes is a Poisson point process with density $\lambda(1-p)$. Percolation theory [93, 94] states that a Poisson distributed network is percolated when node density is above some critical value $\phi$, and therefore a condition for network percolation in [96] is $\lambda(1-p) > \phi$.

In contrast to these two extreme scenarios investigated in [79, 96], we studied a generic scenario that initially, random failures may occur independently, and then each failure may explode and incur an exponentially bounded Blackhole. Therefore, Thinning theorem [93, 94], which requires independent failures, cannot be used here and obviously surviving nodes are no longer distributed as a Poisson point process. This indicates that existing results in percolation theory cannot be used to solve BHR problem directly. Fortunately, reference [94](Page 181) shows that percolation phenomenon (see Sec. 5.2.1) happens not only in the Poisson point process, but in any stationary point process. The occurrence of Blackholes does not change
the stationary property of the original Poisson point process, which motivates us to use some fundamental proof techniques in percolation theory to study the BHR problem.

5.5.2 Sufficient Condition for a Resilient CRN

In this section, we determine the BHR condition provided in Theorem 5.3 by using the technique of continuous-to-discrete percolation mapping. Specifically, we divide the network area into many small square cells and thus the graph consisting of surviving nodes and their connections now appear on the background of these cells, as illustrated in Fig. 5.6a. The size of components of surviving nodes is studied via bond percolation on a discrete grid, as shown in Fig. 5.6b. In particular, to obtain this discrete grid, we represent a cell in Fig. 5.6a by a site located at the center of this cell and two neighboring sites are connected by a bond, which represents
the neighborhood between the two corresponding cells. We choose the size of each cell small enough such that given two arbitrary locations in two neighboring cells, one in each, their distance is at most \( r \) (\( r \) is the transmission range of secondary users defined in Section 5.2.2). This small cell size guarantees that a secondary node is able to communicate with every node in the neighboring cell. Two nodes are separated farthest as shown in Fig. 5.7b, in which their distance is \( d_l\sqrt{5} \) and \( d_l \) is side length of the cell. Letting \( d_l\sqrt{5} = r \), we obtain \( d_l = \frac{r}{\sqrt{5}} \).

To proceed, we need the following notations for a given bond \( b = s_is_j \).

- Event \( A_{s_i} \): At least one surviving secondary node lies in the cell \( \Gamma_i \) associated with site \( s_i \) (see Fig. 5.7a).

- Event \( C_{s_is_j} \): The rectangle \( \text{Rec}_b \) associated with bond \( b \) is defined as the union of two squares associated with \( s_i \) and \( s_j \) respectively (e.g., see the solid rectangle in 5.7a). Particularly, denote \( d_l \) as the length of the square (see Fig. 5.7b), and \((X_{s_i}, Y_{s_i})\) and \((X_{s_j}, Y_{s_j})\) as coordinates of sites \( s_i \) and \( s_j \) respectively. Then \( \text{Rec}_b \triangleq [X_{s_i} - \frac{d_l}{2}, X_{s_i} + \frac{3d_l}{2}] \times [Y_{s_i} - \frac{d_l}{2}, Y_{s_j} + \frac{d_l}{2}] \). The extended rectangle is defined as \( \text{Rec}_{E_b} \triangleq [X_{s_i} - \frac{d_l}{2} - R_I, X_{s_i} + \frac{3d_l}{2} + R_I] \times [Y_{s_i} - \frac{d_l}{2} - R_I, Y_{s_j} + \frac{d_l}{2} + R_I] \) (see the dash rectangle in Fig. 5.7a), where \( R_I \) is the interference range of primary users. We define event \( C_{s_is_j} \) as the set of outcomes for which the following condition is satisfied: there exists at least one channel \( ch_k \) such that no primary users using \( ch_k \) lie in \( \text{Rec}_{E_b} \).

Note that event \( C_{s_is_j} \) guarantees that for some channel \( ch_k \), the distance between primary users using \( ch_k \) and any locations within \( \text{Rec}_b \) is larger than the interference range of primary users \( R_I \), which indicates that \( ch_k \) can be used by any secondary users in \( \text{Rec}_b \). We next define a bond \( s_is_j \) to be open when events \( A_{s_i}, A_{s_j} \) and \( C_{s_is_j} \) occur simultaneously. Particularly, let \( P_o \) be the probability that any bond is open. Then we have \( P_o = P(A_{s_i} \cap A_{s_j} \cap C_{s_is_j}) \). By this definition, an open bond \( s_is_j \) implies that surviving nodes exist in \( \Gamma_i \) and \( \Gamma_j \) respectively, and some channel can be used by these nodes. This is equivalent to saying that an open bond \( s_is_j \) implies an communication link across \( \Gamma_i \) and \( \Gamma_j \). By this mapping, bond percolation on the
discrete lattice ensures percolation of CRN. Therefore, we next investigate bond percolation condition for the discrete grid, which is sufficient for a BHR network.

In Section 5.4, we have shown that the number of failed nodes in each Blackhole $\mathcal{O}_{v_i}$ is upper bounded by $\Upsilon \sim \text{Exp}(-\beta)$, where the expected size $\beta$ of $\mathcal{O}_{v_i}$ is given in Eq. (5.2). That is, any node $v_i$ may be infected by at most $\Upsilon - 1$ nodes. Thus let $\mathcal{P}_l$ denote the probability that a node $v_i$ is surviving (not failed) and we have

$$\mathcal{P}_l = \sum_{\iota=1}^{\infty} \mathcal{P}(v_i \text{ is surviving} | \Upsilon = \iota) \mathcal{P}(\Upsilon = \iota)$$

$$= \sum_{\iota=1}^{\infty} (1 - p)^\iota (\mathcal{P}(\Upsilon \geq \iota) - \mathcal{P}(\Upsilon \geq \iota + 1))$$

$$= \sum_{\iota=1}^{\infty} (1 - p)^\iota (1 - e^{-\beta}) e^{-\beta \iota}$$

$$= \frac{(1 - e^{-\beta})(1 - p)e^{-\beta}}{1 - (1 - p)e^{-\beta}}.$$  

(5.12)

And

$$\mathcal{P}(A_{s_i}) \geq \mathcal{P}_l (1 - e^{\lambda d^2}).$$  

(5.13)

And by the assumption that primary users on any channel $ch_k$ are distributed as a Poisson point process with density $\lambda_p$, we have

$$\mathcal{P}(C_{s_i s_j}) = 1 - (1 - e^{-\lambda_p \alpha})^m,$$

(5.14)

where $\alpha = (d_l + 2R_I)(2d_l + 2R_I)$ denotes the area of $\text{RecE}_b$ (as illustrated in Fig. 5.7a).

To obtain $\mathcal{P}_o = \mathcal{P}(A_{s_i} \cap A_{s_j} \cap C_{s_i s_j})$, another challenge is that $A_{s_i}$ and $A_{s_j}$ are not independent. To continue, we need introduce the following concept and inequality.

**Definition 5.4.** For two geometric random graphs $\mathcal{G}$ and $\mathcal{G}'$, a partial ordering $\preceq$ is defined as $\mathcal{G} \preceq \mathcal{G}'$ if and only if $\mathcal{G}'$ can be induced from $\mathcal{G}$ by adding more (Poisson) points. Then an event $E$ is said to be increasing (decreasing) if $\forall \mathcal{G} \preceq \mathcal{G}', 1_E(\mathcal{G}) \leq 1_E(\mathcal{G}')$ ($1_E(\mathcal{G}) \geq 1_E(\mathcal{G}')$), where $1_E$
is the indicator function of event $E$.

**Lemma 5.3.** (KFG’s inequality [94]) If two events $E_1$ and $E_2$ are both increasing or decreasing, then

$$P(E_1 \cap E_2) \geq P(E_1)P(E_2).$$

By our definition, the more points in $\Gamma_i$, the more likely $A_{s_i}$ occurs. Thus $A_{s_i}$ is increasing. Therefore, we have

$$P_o = P(A_{s_i} \cap A_{s_j} \cap C_{s_is_j})$$

$$= P(A_{s_i} \cap A_{s_j})P(C_{s_is_j}) \geq P(A_{s_i})P(A_{s_j})P(C_{s_is_j})$$

$$\geq P_i^2(1 - e^{-\lambda d_i^2})^2\left(1 - (1 - e^{-\lambda p\alpha})^m\right). \quad (5.15)$$

Finally, by using the percolation condition in square lattice, we can achieve the sufficient BHR condition given in Theorem 5.3.

**Proof of Theorem 5.3.** As analyzed above, by careful definition of open bond in the square lattice, bond percolation on the mapped lattice guarantees the BHR property of CRN. In Section 5.7, we derived a probability $p^\square_c$ such that if bond is open with probability $P_o > p^\square_c$, the square lattice is percolated, which further indicates a BHR CRN. Plugging Eq. (5.15) and solving

$$P_i^2(1 - e^{-\lambda d_i^2})^2\left(1 - (1 - e^{-\lambda p\alpha})^m\right) > p^\square_c, \quad (5.16)$$

we arrive at

$$P_i > \Lambda, \quad (5.17)$$

where $\Lambda$ is given in Eq. (5.3). Then substituting Eq. (5.12) into Eq. (5.17), we have

$$1 - p > \frac{\Lambda e^\beta}{1 - e^{-\beta} + \Lambda}, \quad (5.18)$$

which indicates that $p < 1 - \frac{\Lambda e^\beta}{1 - e^{-\beta} + \Lambda}$ is sufficient for a BHR CRN. This completes the proof. \qed
5.6 Simulations

(a) Random failures
(b) A percolated CRN in the event of random failures

Figure 5.8: Network percolation in the event of random failures.

(a) An initial failure
(b) Failure percolation

Figure 5.9: An initial failure explodes to the entire network.

In this section, we have performed a series of simulations to explain and demonstrate the occurrence of Blackholes, and validate our theoretical analysis. In the simulation, secondary
users are distributed independently and uniformly with density $\lambda$. Time is slotted into units, and at each time slot, primary users on any channel are distributed as a Poisson point process with density $\lambda_p$. The transmission range $r$ of secondary users and interference range $R_I$ of primary users are set as $r = 50$ (meters) and $R_I = 80$ (meters) respectively.

We consider a CRN deployed within area $[0, 1000]^2$ (meters) with $m = 4$ channels, $\lambda = 0.0008 \text{ (per meter}^2\text{)}$ and $\lambda_p = 0.00001 \text{ (per meter}^2\text{)}$. To study Blackholes, we first investigate the occurrence of random failures (according to the random failure model in Section 5.2.2).
Assume that each secondary node fails independently with probability $p = 0.1$, as shown in Fig. 5.8a (in Fig. 5.8, 5.9, 5.10 and 5.11, solid dots and circles represent failed and surviving nodes respectively, a line connecting two failed (surviving) nodes denotes a failure connection (communication link), and the positions of primary users are not shown in the figures for simplicity). Ignoring failure correlation, the condition of whether a network is percolated in the event of such independent failures has been studied in [96]. An example of a percolated CRN in the face of independent failures has been shown in Fig. 5.8b. On the other hand, to understand the failure correlation, we simulate the scenario studied in [79]. Specifically, a particular failure occurs initially, as shown in Fig. 5.9a (see the solid dot in square area). This failure may infect its neighbors, according to the failure explosion model defined in Section 5.2.2, and similarly, infected neighbors may further impact more nodes. Xu et al. [79] determines when this failure will spread to the entire network and an example of such failure percolation has been shown in Fig. 5.9b.

In contrast, we will study the random failures and then their explosion subsequently. In particular, random failures may occur initially according to the random failure model, as shown in Fig. 5.8a. Each random failure then explodes according to the failure explosion model. By using the results in [79], we can set network parameters to ensure that each failure will not spread to the entire network. Therefore, each failure only infects a finite number of nodes and thus a sequence of Blackholes occur (see the components of solid dots in Fig. 5.10a and 5.11a, and two examples of Blackholes have been circled in Fig. 5.10a). The size of Blackholes depends on the failure connection probability $\tau$ in the failure explosion model (Blackhole size increases as $\tau$ increases, as shown in Eq. (5.2)). And Blackholes in Fig. 5.10 and 5.11 are incurred by setting $\tau = 0.2$ and $0.3$ respectively. When Blackholes are small, CRN is percolated (see the giant component of surviving nodes in Fig. 5.10b). As Blackholes grow, this giant component may disappear and CRN is not percolated, as shown in Fig. 5.11b. This motivates our study on the size of Blackholes (BHG problem) and determine when the network is percolated in the presence of Blackholes (BHR problem).
To study the size of Blackholes $O_{v_1}$, we run the simulation with $\lambda = 0.0008$ and $\lambda_p = 0.00001$ within $[0, 1000]^2$ 1000 times independently for variant failure connection probability $\tau$. The probability $P(|O_{v_1}| = N)$ is calculated by the frequency of the occurrence of Blackholes with size $N$. Using this method, the complementary distributions (CCDF) of $O_{v_1}$ under $\tau = 0.2, 0.25, 0.3$ have been calculated and shown in Fig. 5.12 on a semi-log scale. As illustrated in Fig. 5.12, CCDFs under different $\tau$ are approximately linearly under semi-log scale, which validates our analysis in Theorem 5.1 that the size of Blackholes $O_{v_1}$ decays exponentially. In addition, Fig. 5.12 further shows that the CCDF of $O_{v_1}$ decreases, which indicates the expected size of Blackholes $E(|O_{v_1}|)$ decreases, as failure connection probability $\tau$ decreases. This corresponds to our result about expected size of Blackholes in Theorem 5.2.

## 5.7 Appendix: Calculation of Critical Probability $p_c^\square$

Let $p_c^\square$ be the bond percolation probability of the square lattice mapped from CRN (see Fig. 5.6b). It was proved in [99] that bond percolation probability in square lattice is $\frac{1}{2}$. In discrete percolation theory, the open or closed state of every edge (or vertex) is independent from others. In our discrete lattice mapping, the state of an edge depends on, however, how primary and
secondary nodes are distributed around this edge, which implies that adjacent edges are not independent. Therefore, we cannot directly use the result in discrete percolation theory and we need to find out alternative percolation conditions for our mapping. Our method is based on the following observation.

Consider square lattice and its dual \( L \) and \( L' \) (see Fig. 5.13). The construction of \( L' \) is as follows: let each vertex of \( L' \) be located at the center of a square of \( L \). Let each edge of \( L' \) be open if and only if it crosses an open edge of \( L \), and closed otherwise. Now a key observation is that if the origin belongs to an infinite open edge cluster in \( L \), for which the event is denoted by \( E_L \), then there cannot exist a closed circuit (a circuit consisting of closed edges) surrounding the origin in \( L' \), for which the event is denoted by \( E_{L'} \), and vice versa (see page 17 in [99]). This is illustrated in Fig. 5.13. To proceed, we further need the following lemma.

**Lemma 5.4.** Given a lattice \( L \) containing the origin 0 and its dual \( L' \), let \( \sigma(z) \) be the number of paths with length \( z \) in \( L \) (i.e., comprising \( z \) edges) that start at 0, and \( \rho(z) \) be the number of circuits in \( L' \) with length \( z \) and containing 0 in their interiors, then \( \sigma(z) \leq 4 \cdot 3^{z-1} \) and \( \rho(z) \leq 2 \cdot (z-2) \cdot 3^{z-2} \).

**Proof.** See Lemma 2 in [96].

Let \( C_z \) be a circuit of the lattice \( L' \) with length \( z \) containing the origin in its interior,
then $\mathcal{P}(C_z \text{ is closed}) = \mathcal{P}(\text{all } z \text{ edges are closed})$. Based on the open edge definition described in Section 5.5.2, edges $a$ and $b$ are independent if their distance is larger than $\max\{R_I, d_i\}$ (the distance between two edges is defined as the minimum distance between any two points on edges $a$ and $b$). This implies that an independent subset of edges among $z$ edges of $C_z$ can be obtained by selecting an edge in every $\left\lceil \frac{z}{2R_I/d_i+1} \right\rceil$ edges. Thus at least $\kappa = \left\lfloor \frac{z}{2R_I/d_i+1} \right\rfloor$ have independent states among $z$ edges of $C_z$. Let $q$ be the probability of any edge being closed, i.e., $q = 1 - p_c^\Box$, then for any $C_z$, $\mathcal{P}(C_z \text{ is closed})$ is upper bounded by $q^\kappa$. Thus the probability that there exists a closed circuit surrounding the origin in $\mathcal{L}'$ is,

$$\sum_{C_z, \forall z} \mathcal{P}(C_z \text{ is closed}) \leq \sum_{z=4}^{\infty} q^\kappa \rho(z). \quad (5.19)$$

Therefore, $\sum_{z=4}^{\infty} q^\kappa \rho(z) < 1$ indicates that the probability of no closed circuit surrounding the origin in $\mathcal{L}'$ is strictly greater than 0, which provides a lower bound of $p_c^\Box$. For example, if $R_I < \frac{4}{7}$, $\kappa = \left\lfloor \frac{z}{7} \right\rfloor$ and thus

$$\sum_{C_z, \forall z} \mathcal{P}(C_z \text{ is closed}) \leq \sum_{z=4}^{\infty} q^{\left\lfloor \frac{z}{7} \right\rfloor} \rho(z) = \frac{4(9q)^2}{9(1-9q)^2}.$$ 

When $q < \frac{1}{15}$, $\frac{4(9q)^2}{9(1-9q)^2} < 1$ and thus the lattice is percolated, which implies $p_c^\Box > \frac{14}{15}$.

### 5.8 Summary

In this chapter we have studied the topology and resilience of large CRNs in the presence of node failures. When there exist causal relations, a single failure may initiate a component of related failures, and thus random failures may trigger a sequence of Blackholes in the network. In order to understand network topology in the face of Blackholes, two metrics, *failure occurrence probability* $p$ and *failure connection function* $g(\cdot)$ are defined to characterize the occurrence of random failures and their spreading to neighbors, based on which we prove that when a Blackhole cannot spread to the entire network, it is exponentially bounded. By mapping failure
spreading to a branching process, we derive an upper bound on the expected size of Blackholes. After studying Blackhole size, we then investigate network resilience. A network is said to be resilient to Blackholes if there exists a giant component of surviving nodes spanning through the entire network. By coupling with a continuum percolation process on the random geometric graph, we further obtain a sufficient condition for a resilient CRN to a sequence of Blackholes. We finally confirm correctness of our theoretical results by simulations. It is worthy of pointing out that although our results concerning Blackhole size and resilience are derived for CRNs, nevertheless, by setting spatial density of primary users $\lambda_p = 0$, these results can also be applied practically in general wireless networks. For instance, Fang et al. [115] described a distributed algorithm to build routes around Blackholes in wireless sensor networks, and our results can be used to determine the feasibility of such routes, and thus validate this algorithm.
Chapter 6

Conclusion and Future Directions

In this dissertation, we have presented our research results regarding the performance and topology of Cognitive Radio Networks in four fundamental aspects: the capacity and spectrum sensing, the information dissemination in spatial and temporal domains, the latency distribution and scaling under general mobility, and the Blackhole resilience. Next, we summarize our results and discuss the possible future extensions.

6.1 Conclusion

This dissertation has focused on the analysis and understanding of Cognitive Radio Networks on their fundamental performance and topology perspectives. In Chapter 2, we first studied the practical spectrum sensing algorithms in large-scale CRNs to achieve the same order optimal network capacity with that in wireless ad hoc networks. We then investigated how far and how fast information can be disseminated in large CRNs via two new metrics, the dissemination radius and information propagation speed, in Chapter 3. On the network topology aspects, we analyzed the impact of node mobility on network connectivity and performance in Chapter 4, and characterized the spread of correlated failures and studied resilience to Blackholes in Chapter 5. Next, we elaborate and highlight our major findings.

In Chapter 2, we explored the optimal sensing algorithms to achieve the maximum capacity
in the same order as that of wireless ad hoc networks. Due to unique features in CRNs, such as heterogeneous architecture and thus two types of interference, and opportunistic communication which requires spectrum sensing, the existing studies on capacity of wireless networks are not applicable here. For example, spectrum sensing will incur overhead on capacity. We first derive the theoretical bounds of capacity feasible in CRNs by taking these features into account and our analysis however shows that we can achieve comparable network capacity in CRNs. By following the analytical results as a theoretical guideline, we also designed a sensing algorithm to achieve the maximum capacity in the order sense. Our design maximizes the network capacity with minimum sensing overhead.

In Chapter 3, we have studied the information dissemination radius and propagation speed to understand how far and how fast messages are transmitted in large multi-channel cognitive radio networks. We have identified the sufficient and necessary conditions under which information can and cannot be disseminated to the entire network, depending on the spatial densities of primary and secondary users and the number of channels. When information can reach an infinite area, we find that the propagation speed is no lower than a constant $\kappa$. When information cannot percolate to the entire network, our analysis shows that the farthest information dissemination distance is statistically dominated by an exponential distribution and information cannot reach most destinations in the network. Our work advances our understanding of information dissemination in large CRNs.

In Chapter 4, we examined the network topology and performance in a mobile environment. When nodes are not static, the network topology experiences constant changes over time. The communications in the network suffer from the unstable topology as the traffic sessions are interrupted from time to time and the path establishment delays and overheads are increased. We have studied the distribution of the information dissemination latency $T_d$ in finite CRNs and the scalability of $T_d$ in large CRNs under general mobility. We found that in finite networks, there exists a cutoff point on the mobility radius $\alpha$ of secondary users, above which the tail distribution of $T_d$ is bounded by some Gamma distribution and below which $T_d$ has a heavy-tailed
distribution. When networks become large, the dissemination latency $T_d$ is (linearly) scalable with respect to the dissemination distance. Our results demonstrate that when secondary users can move in a large region, a Gamma distributed (light-tailed) latency in finite networks, or a scalable latency in large networks, is achievable, which encourages the deployment of CRNs for real-time and large applications.

In Chapter 5, we characterized the occurrence of random failures and their spread in large CRNs. As failures happen unavoidably in any communication systems and networks, it is imperative for us to gain an insightful understanding of the potential impact created by failures on the overall network performance. When failures propagate, many Blackholes appear in the network. Our work specifically focused on the structure and size of these holes. We first consider two metrics, failure occurrence probability $p$ and failure connection function $g(\cdot)$ to characterize the occurrence of random failures and their spreading to neighbors. By mapping failure spreading to a branching process, we determine the size of Blackholes. After studying Blackhole size, we then investigate network resilience. A network is said to be resilient to Blackholes if there exists a giant component of surviving nodes spanning through the entire network. By coupling with a continuum percolation process on the random geometric graph, we further obtain a sufficient condition for a resilient CRN to a sequence of Blackholes. Our result lays a theoretical foundation for designing efficient strategies to locate and bypass Blackholes.

6.2 Future Directions

The work presented in this dissertation is only part of the efforts in understanding and optimizing the performance of large-scale CRNs. In order to effectively, efficiently, and reliably to utilize this new technology, we may extend our study in the following directions.

Our current work studies and evaluates the network performance and topology from various stand-alone perspectives. A joint study on the interaction of different performance aspects will provide more comprehensive understanding of the large-scale wireless CRNs in general circumstances. For example, it has been evidenced that capacity can be improved at the price
of delay. To target this objective, we will introduce and discuss the concept of \textit{networking multi-modalities}, which brings capacity, delay, connectivity and networking resilience together. The \textit{network multi-modalities} will be determined under a generic network framework, which takes node failures and resilience into consideration, in addition to the network dynamics, such as node mobility and dynamic transmission power, considered above.

Moreover, \textit{heterogeneous networks} are a new trend. In the past ten years, wireless networks where homogeneous nodes are dynamically connected in a multi-hop manner, such as wireless ad hoc networks and wireless sensor networks, have been extensively studied because of their promising and indispensable usages in a wide scope of communication scenarios. However, such homogeneous wireless networks rarely exist because of various requirements of real applications and thus heterogeneous nodes, with different communication requirements, are usually necessitated to coexist and cooperate. For example, in the next-generation electric power systems (smart grid) \cite{120}, the protection information and commands exchanged between intelligent electrical devices (IEDs) will require much lower network latency than the messages exchanged between electrical sensors and control centers. Therefore, the IEDS should have priority to access the spectrum to the latter. In addition, as shown in \cite{1, 2}, the usable spectrum is limited but the spectrum usage efficiency in homogeneous networks is low, which motivates the network where heterogeneous nodes coexist and share the spectrum. These heterogeneous nodes can be classified as \{\textit{Type}_i\}_{i=1}^\gamma nodes in a hierarchical manner based on their priorities to spectrum access, i.e., \textit{Type}_j (j < i) nodes have higher priority than \textit{Type}_i nodes to access spectrum. And the recent development in software-defined radio and cognitive radio make such heterogeneous networks feasible. Understanding performance and topology of heterogeneous networks will be another interesting research problem.
REFERENCES


