ABSTRACT

LI, CHENGZHI. Analytical Study on Heterogeneous Systems. (Under the direction of Dr. Huaiyu Dai.)

A heterogenous system is composed of multiple different systems sharing the same space and frequency spectrum. In this dissertation we study two representative heterogenous systems: spectrum sharing systems and jamming-resistant systems. In the former various networks share the same spectrum with different priorities, in order to improve the spectrum efficiency; while in the latter jammers intend to destruct legitimate transmissions by producing strong interference.

The purpose of spectrum sharing systems is to relieve current conflict between increasing demand for radio spectrum and spectrum scarcity. It is typically composed of primary users with absolute right to utilize its licensed spectrum, and secondary users opportunistically accessing the same spectrum provided that their interference to primary receivers remains under an acceptable level. We start with our study on spectrum sensing algorithms in Chapter 2, one of the key components in spectrum sharing systems. A new spectrum sensing algorithm, adaptive CUSUM test, is proposed under the guidance of quickest detection theory, which can adaptively detect and track distribution changes with unknown and even time varying parameters. Our approach is efficient in computation and outperforms traditional parallel CUSUM test significantly.

The spectrum sharing system is attractive since it can potentially improve the spectrum efficiency. The question is to what extent the secondary network can gain in transmission of its own useful data, without harming the regular operation of the primary network. The throughput of secondary networks is studied in this thesis through three metrics: throughput scaling (Chapter 3), transmission throughput (Chapter 4) and single-hop transport throughput (STT) (Chapter 5). These studies provide insights from different perspectives. In particular our scaling law study shows that even under a weaker yet more practical scenario, where the densities of both networks are at the same order, the secondary network maintains the same throughput scaling as a stand-alone network, though subject to a certain outage. Our investigation on transmission throughput reveals a sufficient condition for the secondary network configuration such that the sum transmission throughput can be indeed boosted over that of the stand-alone primary network. The third metric, single-hop transport throughput, inherits the merits of transmission throughput and transport capacity, and incorporates transmission distance and outage probability into a uniform framework. The STT of the secondary network is explored with outage constraints imposed on both networks. Its limit, single hop transport capacity (STC), is derived, and the STT with secondary receivers randomly located in the field of interest is examined.

In spectrum sharing systems, we endeavor to regulate the operation of each individual system so that they can jointly achieve their maximal benefit under feasible constraints. In contrast, our effort in jamming-resistant systems is focused on enhancing jamming resistance of the legitimate systems. Particularly in Chapter 6 we study the communication efficiency of the Uncoordinated Frequency Hopping (UFH) technique, a recently proposed jamming countermeasure requiring no pre-shared secrets keys, in a large-scale broadcast network. We demonstrate that the average network broadcast delay incurred by UFH scales logarithmically with the network size, which substantiates its feasibility in a large scale network. With the help of relays, we further analytically show that its efficiency can be improved significantly through collaborative broadcast. Our study offers tradeoff and optimization on some key parameters in the UFH-based schemes.

Another topic we study for jamming-resistant networks concerns network connectivity and delay addressed in Chapter 7. We examine the connectivity issue in the framework of percolation theory. It is shown that the network experiences two phase transitions as the jammer density increases: from strong connection to weak connection then to disconnection. In addition there exists two cutting jammer densities beyond which the network is no longer strongly or weakly connected. It is also found that the multi-hop broadcast delay between two nodes on a weakly connected network scales linearly with their distance. This implies that in terms of scaling law, anti-jamming techniques such as UFH do not incur performance loss.

It is our hope that the work in this dissertation can help advance the understanding and deployment of current and emerging heterogeneous wireless networks.

Analytical Study on Heterogeneous Systems

by
Chengzhi Li

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

Electronical Engineering

Raleigh, North Carolina

2012

APPROVED BY:

Dr. Brian Hughes	Dr. Keith Townsend
Dr. Peng Ning	Dr. Huaiyu Dai Chair of Advisory Committee

DEDICATION

To my family.

BIOGRAPHY

Chengzhi Li was born on July 2nd, 1979, in Tianjin, China. He received his Bachelor of Science degree in Electrical Engineering in 2001 from Nankai University, Tianjin, China; and Master of Science degree in Signal Processing in 2004 from Peking University, Beijing, China. In August 2008, he started his work towards the PhD degree in Electrical Engineering at North Carolina State University, Raleigh, NC, under the guidance of Prof. Huaiyu Dai. His research interests include the general areas of wireless communications and signal processing, with focus on network analysis, distributed detection, and distributed computation.

ACKNOWLEDGEMENTS

I would like to acknowledge and extend my heartfelt gratitude to the following persons who have made the completion of this dissertation possible.

My deepest gratitude goes first and foremost to my advisor, Prof. Huaiyu Dai, for his continuous guidance, support and inspiration throughout this work. His wisdom, patience and encouragement have immensely influenced my academic growth. He walked me through every step of research and helped me tackle various problems. Without his guidance, this dissertation could not have reached its present form.

Secondly I am grateful to my committee members Prof. Brian Hughes, Prof. Keith Townsend, and Prof. Peng Ning for their illuminating instruction, encouragement, and their valuable suggestions and feedback on my work. I would also like to take the opportunity to thank all professors and teachers who have imparted to me all the knowledge in my student life.

I would like to thank my colleagues in my research group, Dr. Yanbing Zhang, Dr. Juan Liu, Mr. Huazi Zhang, Mr. Xiaofan He and Mr. Nakul Navarange. I have benefited greatly from enlightening discussions with them. And we had a great time in our personal life.

Finally, I would like to thank my parents for their unconditional love and consistent support for my studies. They set an excellent example for me to be an honest, warm-hearted and aspirant person. Every progress I made in my life owes to them. A special thanks goes to my wife, Dr. Hongxia Fang. She has always being there with me to share my pains and joys. Her love and support has encouraged me to overcome difficulties in both my life and research. This dissertation will not be completed so smoothly without the support and accompaniment she has provided

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

1.1 Motivation

In the past decade, wireless technology advances rapidly. Wireless communication becomes an indispensable way for people to exchange and share information worldwide. Satellites, Cellular systems, Wi-Fi, and bluetooth are some of the popular wireless systems. With the ever increasing demand for the scarce wireless spectrum, heterogeneous systems often coexist in the same space sharing or competing for the same spectrum resources. As the radio spectrum can not be created and can only become more precious, there is a pressing need for the study of heterogeneous wireless systems, as well as the investigation of sensible mechanisms that can protect the normal operation of legitimate systems and guarantee their healthy coexistence. In this dissertation, we are particularly interested in two heterogenous systems: spectrum sharing networks and jamming-resistant networks.

1.1.1 Spectrum sharing systems

Spectrum sharing systems are strongly propelled by the paradox between spectrum scarcity and underutilization. On the one hand the overly crowded frequency allocation chart drawn by Federal Communication Commissions (FCC) indicates that few radio frequency band are left for further use. On the other hand, actual spectrum usage measurements obtained by the FCC's Spectrum Policy Task Force [3] have reported vast temporal and geographic variations in the usage of allocated spectrum, which indicates a significant amount of spectrum white space. Spectrum sharing systems incorporate new technologies to exploit these spectrum white space or spectrum holes, admitting more intensive and efficient spectrum use. Such heterogeneous systems are typically composed of primary users, which are granted the licence to utilize the spectrum, and secondary users, which are equipped with cognitive radios to access the spectrum opportunistically. A cognitive radio is a software defined 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. Secondary operations are required transparent to primary nodes in order not to cause notable influence to the legacy (primary) system. The primary nodes in general do not cooperate with secondary nodes. It is the secondary nodes' responsibility to constrain their interference to primary nodes below an acceptable level (set according to the requirements of the primary nodes). Two spectrum sharing strategies have been proposed: spectrum underlay and spectrum overlay. In the former secondary transmissions occur concurrently with primary ones, however, the transmission power of secondary users have been strictly constrained to avoid incurring serious interference to primary receivers. In contrast, secondary transmissions in spectrum overlay occur only when there exists spatial and temporal spectrum white space around the secondary communication pairs. What matters to secondary nodes is the identification of spectrum holes rather than their transmission power.

One of the key components in a spectrum overlay system is spectrum sensing, the ability to locate spectrum holes by sensing and measuring the radio spectrum in the neighborhood of the cognitive radio transceiver. Basically spectrum sensing algorithms tell cognitive radios when, where and which spectrum is available for secondary users to utilize. Spectrum sensing is required not only before a secondary transmission to check for the the vacancy of spectrum, but also during the secondary transmissions to monitor the emergence of primary signals. Once the primary signals are detected in the frequency bands currently occupied by the secondary users, the latter have to vacate from the corresponding frequency band as quickly as possible. Spectrum sensing has been extensively investigated in the past few years. Some of the most common spectrum sensing techniques (see [4] and reference therein) include energy detection based sensing, cyclostationarity-based sensing and matched-filtering based sensing. A group of secondary nodes can also perform spectrum sensing collaboratively, to improve the reliability in environments with noise uncertainty, fading, and shadowing. Essentially, the spectrum sensing is to detect the change of spectrum activity. Therefore, besides the aforementioned classic approaches we can apply the theory of quickest detection [5, 6], which performs a statistical test to detect the change of distribution in observations as quickly as possible, in order to attain an agile and robust spectrum sensing. However, it is insufficient to apply the well-known cumulative sum (CUSUM) test [7] directly. There exist unknown parameters after the primary radio emerges, e.g. the amplitude of the received primary radio signal. Existing algorithms combatting the unknown parameters include generalized likelihood ratio (GLR) test [8], windowed GLR test [9] and parallel CUSUM test [10], all of which fail when the unknown parameters are time varying as well, which is inevitable in wireless fading channels. In Chapter 2, we propose an adaptive CUSUM test for quickest detection when there are unknown parameters in the post-change distribution and the parameters may be varying during

the detection process. Our new approach can narrow down the range of the unknown parameters to desired precision and track their changes adaptively, thus achieving significant performance improvement over the parallel CUSUM Test.

The spectrum sharing system is attractive since it can potentially improve the spectrum efficiency. The question is to what extent the secondary network can gain in transmission of its own useful data, without harming the regular operation of the primary network significantly. As a fundamental property of a network, the throughput of a spectrum sharing system deserves our investigation. In literature, relevant research on network throughput is dominated by the study of either the scaling law of transport capacity [11–13] or transmission capacity with outage consideration [14–17]. The former was proposed in the seminal work [18] and defined as the maximum bit·meters per second the network can achieve in aggregate; recent works in the capacity study of CR networks or ad-hoc overlaid networks [12, 13] show that there is no performance loss for the secondary network in terms of scaling law of transport capacity. However in these existing works the density of the secondary network is assumed to be higher than that of the primary one in the order sense. Under this assumption the influence of the primary network on the secondary network becomes increasingly negligible. In Chapter 3, we endeavor to study a weaker yet more practical scenario where the densities of both networks are at the same order. Our results show that the secondary network indeed performs as well as the primary one, in terms of scaling law, when the secondary-to-primary density ratio is larger than a threshold. In addition, secondary network inevitably suffers from a constant outage, defined as the percentage of the secondary S-D pairs which can not be served due to interference constraint.

Asymptotic analysis on the scaling law only characterizes the (rough) relationship between capacity and network size, neglecting the effect of many important system parameters. As an alternative, transmission capacity [19] quantifies the maximum spacial density under some outage probability constraint, and sparks an enormous interest recently (see [20] for its latest development). In spectrum sharing systems, it is more appropriate to study transmission throughput[1] than transmission capacity since both primary and secondary outage constraints can not be achieved simultaneously in general. In [16] it is shown that with small outage constraints on primary and secondary networks, the spectrum efficiency of the whole overlaid networks can be improved if extra outage is allowed for the primary network. However, the secondary network may not always improve the spectrum efficiency. With the additional interference incurred by the secondary transmitters, the primary throughput is inevitably degraded. In order to enhance the overall throughput of the spectrum sharing system the secondary network has to be appropriately configured. In chapter 4, we endeavor to quantitatively characterize the appropriate secondary setting and reveal a sufficient condition for the secondary network configuration

[1] defined as the product of the spatial density and the actual outage probability

such that the overall (transmission) throughput is boosted over that of the stand-alone primary network. Transmission capacity or throughput admits quantitative system analysis, but leaves out the consideration (and optimization) of transmission distance, a key parameter for wireless networks. To offer a comprehensive view of network throughput, in Chapter 5 we further study a new metric: single hop transport throughput (STT), which quantifies the total number of one-hop *reliable* transmissions in a unit area, weighted by corresponding transmission rates and distances. STT inherits the merits of both traditional transport capacity and transmission capacity, and incorporates transmission distance and outage probability into a uniform framework. It is worth investigating STT in the secondary network in that single hop transmissions may be preferred due to its inferior role in the spectrum access; it will also serve as a basis for extension to the multi-hop case.

1.1.2 Jamming-resistant networks

Secondary users in spectrum sharing systems have to guarantee all the primary users' right while sharing the spectrum with them. In contrast, jammers overlaid with legacy systems intend to damage the legitimate transmissions. Wireless networks are highly vulnerable to jamming attacks due to the openness of the communication medium. Lack of resistance to jamming, wireless signals can be easily blocked, modified or replaced, which may jeopardize personal safety and national security. As a result, jamming-resistant networks are crucial to security-critical applications; examples include emergency alert broadcast and navigation signal dissemination.

A conventional jamming countermeasure at the physical layer is the employment of Spread-Spectrum (SS) techniques, including Direct-Sequence Spread Spectrum (DSSS) and Frequency Hopping (FH). By deliberately spreading the signal's energy over a large bandwidth, such techniques prevent the legitimate transmission from being detected or intercepted, and exhibit resistance to narrow-band interference or jamming. However, these classic countermeasures rely on pre-shared secret keys, such as spreading code sequences and frequency hopping patterns, to achieve correct decoding at the receivers. Therefore, before the establishment or after the compromise of secrete keys, these approaches are inefficacious; and the schemes without secret keys are desired to further improve network jamming resistance. Uncoordinated SS (USS) attracts much attention recently due to its independency on secret keys, which includes Uncoordinated DSSS (UDSSS) [21] and Uncoordinated FH (UFH) [22, 23]. USS randomizes the key selection so that neither attackers nor malicious nodes are able to jam the communication. Due to their inherent difference in handling interference, UDSSS and UFH exhibit difference in anti-jamming performance [24]: generally speaking, the performance of UDSSS is mainly determined by the computing power of jammers, while that of UFH mainly decided by the hardware capacity (such

as sensing and switching frequency) of jammers. As it is anticipated that the improvement on the computing capability will continue its current trend in the foreseeable future, UFH assumes certain advantage in this aspect.

For a multi-channel network under jamming attacks, the communication efficiency of UFH is of our interest, particularly in a large-scale broadcast network with multiple cooperative relays. Such a large-system analysis is pertinent to emerging and future networked systems, and also helps us reveal the scaling behavior of the system performance with respect to the network size and other important system parameters. In Chapter 6 we reveal that the average broadcast delay in the UFH-based schemes scales only logarithmically with the network size, and there exists an optimal number of relays leading to the minimal broadcast delay. Network connectivity and message dissemination delay is our primary concern when a group of jammers with limited power are overlaid with the network. Subject to jamming, regular transmissions are no longer feasible and UFH can be exploited to enhance jamming resistance. However even UFH fails if all the channels are jammed by a group of cooperative jammers. As a consequence, the analysis on the network connectivity and information propagation delay in such a scenario is more complex than that of a jamming-free network.

In summary, our investigation on spectrum sharing systems, particularly on the spectrum sensing algorithm and throughput analysis, help expand our knowledge on the system design and the best attainable performance in spectrum sharing systems and advance our understanding of their potentials. In addition, our study on jamming-resistant networks analysis substantiates the feasibility and potentials of the UFH-based schemes in a large-scale network, and demonstrates the role of jammers and their countermeasures such as UFH in the network connectivity and message dissemination.With suitable modification, our analytical methodology can find applications in other heterogeneous systems.

1.2 Contributions

Our contribution in this dissertation are summarized below:

- Quickest detection techniques are applied into spectrum sensing, a key component of cognitive radio technology. In order to detect an abrupt distribution change with multiple not only unknown but also time varying parameters, we propose a novel adaptive approach, which is shown to outperform the traditional parallel CUSUM Test.

- The transport throughput of a spectrum sharing system is investigated in terms of scaling law, with a weaker and arguably more practical assumption that the densities of both primary and secondary networks are at the same order. Our analysis shows that with the ratio of the two network sizes larger than a threshold, this weaker condition does not

degrade either network throughput scaling. Our result further reveals the potentials of cognitive radio technology in real applications.

- We further examine the overall transmission throughput of decentralized overlaid networks. Our study provides a sufficient condition for the secondary network setting such that the overall throughput is improved over that of a stand-alone primary network. In addition both the maximal allowable secondary density and the optimal secondary density which maximizes the overall throughput are derived.

- A new metric, single hop transport throughput (STT), is proposed and studied in a spectrum sharing system. STT is a new metric that inherits the merits of both the traditional transport capacity and another popular metric, transmission capacity, incorporating transmission distance and outage probability into a uniform framework. The STT of the secondary network is explored with outage constraints imposed on both networks. We first derive the limit of STT, single hop transport capacity (STC), allowing optimal choice of the secondary transmitter density and the distribution of secondary receivers. Then we investigate STT with secondary receivers randomly located in the field of interest. Three models regarding the selection of receivers are considered: optimally selected, randomly selected, and the nearest neighbors. Our study provides a comprehensive view of achievable secondary network throughput, and offers insights into the configuration of secondary networks in spectrum sharing systems.

- We analytically evaluate the communication efficiency of UFH and cooperative UFH (CUFH) in large-scale networks, and investigate their scalability with the number of network nodes n and tradeoffs involved in other parameters such as the number of channels, relays, and packets that a message is divided into. In particular we show that the average network broadcast delay incurred by both protocols scales as $O(\ln n)$, which demonstrates the feasibility of UFH-based schemes in large-scale networks. In addition, the optimal number of relays in CUFH is derived, and the optimal cooperation gain achieved by CUFH over UFH is given, which increases with the number of channels. We also propose a suboptimal protocol CUFH-p to facilitate the implementation of CUFH, and reveal that CUFH-p achieves the optimal cooperation gain asymptotically.

- The connectivity of a multi-hop multi-channel network is investigated under jamming attacks, in the framework of percolation theory. A jamming region is specified for each jammer and the influence of multiple jammers are considered. We show that the network experiences two phase transitions as the jammer density increases: from strong connection where there exists an infinite component composed of regular links, to weak connection where there exists an infinite component composed of both regular and UFH links, then

to disconnection. Additionally there are two cutting jammer densities such that if the jammer density is beyond one of these densities, the network is either disconnected, or no longer strongly connected, for any network density. In addition the multi-hop delay incurred by UFH is examined on a percolated network. It is shown that the multi-hop broadcast delay between a pair of nodes scales linearly with their Euclidean distance, which implies that anti-jamming techniques such as UFH incurs no performance loss, in terms of scaling law.

1.3 Organizations

The remainder of this dissertation is organized as follows. The adaptive quickest detection is discussed in Chapter 2. And then the transport throughput in terms of scaling law, the transmission throughput and the single-hop transport throughput of secondary networks are investigated in Chapter 3, 4 and 5, respectively. Communication efficiency of UFH and CUFH in a large-scale jamming-resistant network is analyzed in Chapter 6, followed by study on network connectivity and multi-hop delay in Chapter 7. Finally the dissertation is summarized in Chapter 8, together with some directions for future work.

Chapter 2

Adaptive Quickest Change Detection with Unknown Parameters

2.1 Introduction

Quickest detection is a technique to detect distribution changes as quickly as possible based on sequential observations [5]. It admits a wide range of applications such as quality control, medical diagnosis and intrusion detection. Recently it is also applied in the study of Cognitive Radio [25]. There are two standard mathematical formulations for the quickest detection problem: Bayesian and minimax. The Bayesian formulation [26] assumes that the change point has a geometric distribution, while the minimax formulation [8], which is also our focus in this chapter, has no such an assumption. The goal of the minimax formulation is to minimize the worst case detection delay given some constraint on the false alarm. When both pre-change and post-change distributions are completely specified, one well-known procedure under the minimax formulation is the Cumulative Sum (CUSUM) test proposed by Page in [7]. Lorden showed that Page's CUSUM Test is optimal for independent observations [8], and Lai extended this study to dependent observations [27]. In many practical situations, however, the post-change distribution involves unknown parameters. The Generalized Likelihood Ratio (GLR) Test [8] is an optimal procedure to tackle such problems, but unbounded memory requirement makes it infeasible in practice. To improve efficiency in storage and computation, Nikiforov proposed the Parallel CUSUM Test [10], in which multiple Page's CUSUM Tests are carried out simultaneously on some specifically chosen values of the unknown parameters.

In this chapter, we propose an adaptive CUSUM algorithm for quickest detection when there are unknown parameters in the post-change distribution and the parameters may be varying during the detection process. Our new approach can narrow down the range of the unknown parameters to desired precision and track their changes adaptively, thus achieves significant

performance improvement over the Parallel CUSUM Test.

The remainder of this chapter is organized as follows. The system model is given in Section 2.2. After discussing existing approaches to the unknown-parameter problem in Section 2.3, we propose the adaptive CUSUM algorithm in Section 2.4, together with some performance analysis. The simulation results and conclusions are provided in Section 2.5 and Section 2.6, respectively.

2.2 System Model

Suppose a sensor is monitoring some property in the environment. Denote by $x(t)(t = 1, 2, ...)$ its (independent) observation at time slot t, whose probability density function belongs to $\{p_\theta\}_{\theta\in\Theta}$, where $\theta = (\theta_1, \theta_2, ..., \theta_d)^T$ is a d-dimensional vector. We assume the value of θ is changed from $\lambda = [\lambda_1, \lambda_2, ..., \lambda_d]^T$ to $\varphi = [\varphi_1, \varphi_2, ..., \varphi_d]^T$ at some unknown time instant v, i.e., the distribution of $x(t)$ is changed from p_λ to p_φ, where the vector φ is unknown but lies within a given set Φ, and λ is a known vector, outside of Φ. Therefore, two hypotheses of interest are:

$$\begin{cases} H_0 : \theta = \lambda \notin \Phi \\ H_1 : \theta = \varphi \in \Phi. \end{cases} \tag{2.1}$$

Correspondingly, the log likelihood ratio (LLR) is defined as:

$$l_\varphi(t) = \log\left(\frac{p_\varphi(x(t))}{p_\lambda(x(t))}\right). \tag{2.2}$$

For performance measurement, we use detection delay and mean time between false alarms [8], which are defined respectively as[1]

$$\bar{T}_1 = \sup_{v\geqslant 1}\{\text{esssup}E_v\left[(T^* - v + 1)^+ | x(1), ..., x(v-1)\right]\}, \tag{2.3}$$

$$\bar{T}_0 = E_\infty\left[T^*\right], \tag{2.4}$$

where T^* is the stopping time determined by the detection algorithms. E_v denotes the expectation, with respective to p_φ, under the assumption that the change happens at time slot v. $v = \infty$ means that the change never happens and E_∞ denoted the expectation with respective to p_λ.

[1] *esssup* refers to the worst case of all the pre-change distributions and $[x]^+ = \max(x, 0)$.

2.3 Existing Algorithms

In this section, we briefly review four existing algorithms of quickest detection: Page's CUSUM Test, GLR Test, Window-limited GLR Test and Parallel CUSUM Test. The first one applies to the case where complete knowledge of p_φ and p_λ is available, while the last three apply to the case where φ is unknown.

2.3.1 Page's CUSUM test

With full knowledge about the pre-change and post-change distributions, Page's CUSUM Test provides an optimal scheme minimizing the worst-case detection delay in (2.3) [8]. Specifically, the stopping time T^* in Page's CUSUM Test is given by

$$T^* = \inf\left\{ t \middle| \max_{1\leqslant k\leqslant t} \sum_{r=k}^{t} l_\varphi(r) \geq h \right\},$$

where h is a predetermined threshold and the metric $l_\varphi(r)$ is defined in (2.2). An alternative expression of T^* is:

$$T^* = \inf\left\{ t \middle| s(t) \geq h \right\}, \tag{2.5}$$

with $s(t) = \max\left(s(t-1) + l_\varphi(t), 0\right)$ (and $s(0) = 0$). It is equivalent to the above one but more efficient in computation and memory.

2.3.2 GLR test and window-limited GLR test

Unknown parameters in the post-change distribution make the detection more challenging. In such a scenario the optimal procedure is the GLR Test [8], in which the stopping time is given by:

$$T^* = \inf\left\{ t \middle| \max_{1\leqslant k\leqslant t} \sup_{\varphi\in\Phi} \sum_{r=k}^{t} l_\varphi(r) \geq h \right\}.$$

The sup operation implicates an optimal estimation for the parameters, which results in its asymptotical optimality [8] at the cost of large computation and memory requirement (since no recursive expression is available).

As a cost-effective alternative, Window-Limited GLR (WL-GLR) Test is proposed in [27,28], where the stopping time is given by:

$$T^*_{m\tilde{m}} = \inf\left\{ t > m \middle| \max_{max(0,t-m)+1\leqslant k\leqslant t-\tilde{m}} \sup_{\varphi\in\Phi} \sum_{r=k}^{t} l_\varphi(r) \geq h \right\},$$

where $0 \leqslant \tilde{m} \leqslant m$ and m is the size of a moving window. Since the number of samples are limited by a moving window, the computation cost of WL-GLR is considerably reduced. However, the window size still needs to be large enough (especially at low SNRs) to detect a change reliably due to uncertainty about the parameter. Therefore its computation cost may still be unacceptable in practice.

2.3.3 Parallel CUSUM test

To further reduce the computation burden, Nikiforov [10] proposed the Parallel CUSUM Test. Instead of online estimation of the unknown parameter, the Parallel CUSUM Test carries out a collection of Page's CUSUM Tests over L specially chosen values of φ, denoted by $c_1, c_2, ..., c_L$. The stopping time is given by:

$$T^* = \inf \{T_i, i = 1, 2, ..., L\}, \tag{2.6}$$

where $T_i = \inf \left\{ t \middle| s_{c_i}(t) \geq h \right\}$ and metric $s_{c_i}(t)$ is computed recursively by

$$s_{c_i}(t) = \max \left(s_{c_i}(t-1) + l_{c_i}(t), 0 \right).$$

The Parallel CUSUM Test reduces computation complexity but is suboptimal in nature. Its suboptimality lies in the fact that the candidates won't be changed during the detection process so that the inaccuracy of the parameters affects the performance throughout the process. L can be fairly large to achieve satisfactory detection delay $\bar{T}_1$, which, however, increases the computation burden and decreases the mean time between false alarms $\bar{T}_0$.

2.4 Adaptive CUSUM Algorithm

The exploration in the unknown-parameters problem is under development. Especially the study is still lacking on the practical scenarios where the unknown parameters vary over time, e.g., due to channel fading in the wireless environment. In such situations, all algorithms originally designed with the assumption of fixed parameters (such as those discussed in Section 2.3) will degrade.

A straightforward idea to solve the unknown parameters problem is separate operation, estimating the parameters first and then performing the detection. This idea is simple, however it hurts the performance of both estimation and CUSUM test. The estimation should not be stopped since more samples could lead to more accurate estimation. And the CUSUM test should start from the beginning no matter the estimates are accurate or not. Therefore joint operation, detecting the change while estimating the parameters, is more preferred in terms of

performance.

In this section, we propose a new quickest detection algorithm, the Adaptive CUSUM Test, which achieves a better tradeoff between performance and complexity and performs stably in changing environment. Our algorithm is recursive in nature, with each recursion comprising two main interleaved steps: parameter tracking and modified CUSUM test.

2.4.1 Parameter tracking

Our adaptive parameter tracking approach narrows down each unknown parameter's range to desired precision for a generic parameterized distribution satisfying the concavity property of $F(\hat{\varphi})$, where $\hat{\varphi}$ is an estimate of θ when $\theta = \varphi$ and $F(\hat{\varphi}) = E\left[l_{\hat{\varphi}}(t)\right]^2$, the mismatched Kullback-Leibler (KL) divergence between p_φ and p_λ. For rigorous presentation, we restrict our attention to the exponential family, i.e., p_θ is defined as:

$$p_\theta(x) = h(x)\exp(\sum_{i=1}^{s}\theta_i T_i(x) - A(\theta)), \tag{2.7}$$

where $\theta \in \Theta$, $T(x) = \{T_1(x), T_2(x), ..., T_s(x)\}$ is a sufficient statistic, $A(\theta)$ is a normalization factor and $s \geq d$.

We obtain the following result:

Proposition 1. *For distributions in the exponential family, $F(\hat{\varphi})$ is a concave function with global maximum achieved at $\hat{\varphi} = \varphi$.*

proof:

$$\begin{aligned} F(\hat{\varphi}) &= E\left[\log\frac{p_{\hat{\varphi}}(x(t))}{p_\lambda(x(t))}\right] \\ &= E\left[\log\frac{p_\varphi(x(t))}{p_\lambda(x(t))}\right] - E\left[\log\frac{p_\varphi(x(t))}{p_{\hat{\varphi}}(x(t))}\right] \\ &= E\left[l_\varphi(t)\right] - D(p_\varphi \parallel p_{\hat{\varphi}}), \end{aligned} \tag{2.8}$$

where $D(p_\varphi \parallel p_{\hat{\varphi}})$ is the KL divergence between p_φ and $p_{\hat{\varphi}}$. Since $D(p_\varphi \parallel p_{\hat{\varphi}})$ is nonnegative [29], $F(\hat{\varphi})$ achieves the global maximum when $\hat{\varphi} = \varphi$.

[2]Here the expectation is with respect to p_φ.

Substituting (2.7) in (2.8) and taking partial derivation with respect to $\hat{\varphi}_i$, we have

$$\begin{aligned}\frac{dF(\hat{\varphi})}{d\hat{\varphi}_i} &= -\frac{dD(p_\varphi \parallel p_{\hat{\varphi}})}{d\hat{\varphi}_i} \\ &= -E\left[\frac{d}{d\hat{\varphi}_i}\sum_i ((\varphi_i - \hat{\varphi}_i)T_i(x(t)) - (A(\varphi_i) - A(\hat{\varphi}_i)))\right] \\ &= E[T_i(x(t))] - \frac{dA(\hat{\varphi})}{d\hat{\varphi}_i}.\end{aligned}$$

Therefore when $\hat{\varphi}_i = \varphi_i$, $\frac{dF(\hat{\varphi})}{d\hat{\varphi}_i} = 0$ due to the fact that $E[T_i(x(t))] = \frac{dA(\varphi)}{d\varphi_i}$ [30].

For the second derivative $\frac{d^2F(\hat{\varphi})}{d\hat{\varphi}_i d\hat{\varphi}_j} = -\frac{d^2A(\hat{\varphi})}{d\hat{\varphi}_i d\hat{\varphi}_j}$. According to the differential identities of $A(\varphi)$ [30], $\frac{d^2A(\hat{\varphi})}{d\hat{\varphi}_i d\hat{\varphi}_j} = cov(T_i(x(t)), T_j(x(t)))$. Therefore the second derivative matrix M is given by:

$$M = -E\left[(T_1(x(t)), T_2(x(t)), ..., T_s(x(t)))^T (T_1(x(t)), T_2(x(t)), ..., T_s(x(t)))\right],$$

which is a non-positive definite matrix since the quadratic function $aMa^T \leq 0$, where a is any row vector with size s. Thus, $F(\hat{\varphi})$ is a concave function.

□

Corollary 1. *For a particular parameter θ_i and any estimates of other parameters, $F(\hat{\varphi}_i)$ is also a concave function with global maximum at $\hat{\varphi}_i = \varphi_i$.*

According to Corol. 1, for the parameter θ_i we can draw the following conclusions after distribution change:

- given a small value δ_i and estimations for other parameters, we can always find two estimates $\hat{\varphi}_{ia}$, $\hat{\varphi}_{ib}$ for θ_i such that $\hat{\varphi}_{ib} = \hat{\varphi}_{ia} + \delta_i$ and $F(\hat{\varphi}_{ia}) = F(\hat{\varphi}_{ib})$;
- θ_i's true value φ_i lies within the interval $(\hat{\varphi}_{ia}, \hat{\varphi}_{ib})$, denoted by $\hat{\Phi}_i$.

Intuitively, we can narrow down the range of φ_i from its original range Φ_i to $\hat{\Phi}_i$; this also allows parameter tracking in time-varying environments. An iterative procedure can be used to find $\hat{\varphi}_{ia}$ and $\hat{\varphi}_{ib}$. Given δ_i, one begins by arbitrarily choosing $\hat{\varphi}_{ia}$ and $\hat{\varphi}_{ib}$ within Φ_i, say $\hat{\varphi}^0_{ia}$ and $\hat{\varphi}^0_{ib}$. And then the succeeding values of $\hat{\varphi}_{ia}$ and $\hat{\varphi}_{ib}$ are obtained according to the recursion:

$$\hat{\varphi}^{k+1}_{ia} = \hat{\varphi}^k_{ia} + \xi\mathcal{D}^k_i, \tag{2.9}$$

$$\hat{\varphi}^{k+1}_{ib} = \hat{\varphi}^k_{ia} + \delta_i, \tag{2.10}$$

where $\hat{\varphi}^k_{ia}$ and $\hat{\varphi}^k_{ib}$ represent the values of $\hat{\varphi}_{ia}$ and $\hat{\varphi}_{ib}$ at the kth iteration, $\mathcal{D}^k_i$ is the difference

between $F(\hat{\varphi}_{ib}^k)$ and $F(\hat{\varphi}_{ia}^k)$ at the kth iteration, given by:

$$\mathcal{D}_i^k = F(\hat{\varphi}_{ib}^k) - F(\hat{\varphi}_{ia}^k) = E\left[\log \frac{p_{\hat{\varphi}_{ib}^k}}{p_{\hat{\varphi}_{ia}^k}}\right],$$

and ξ is a step size controlling the rate of adjustment.

In practice, it is not easy to calculate $\mathcal{D}_i^k$ due to the expectation operation. Two possible variations can be considered to overcome the difficulty.

- Continuous Updating

$$\hat{\varphi}_{ia}^{k+1} = \hat{\varphi}_{ia}^k + \xi \hat{\mathcal{D}}_i^k, \tag{2.11}$$

  where $\hat{\mathcal{D}}_i^k = l_{\hat{\varphi}_{ib}^k}(k) - l_{\hat{\varphi}_{ia}^k}(k) = \log \frac{p_{\hat{\varphi}_{ib}^k}(x(k))}{p_{\hat{\varphi}_{ia}^k}(x(k))}$. It is a common approach to replace the ensemble average by the time average, accomplished by means of the recursive first-order difference equation in (2.11).

- Block Updating

  Another variation is to average the log likelihood ratio of $p_{\hat{\varphi}_{ib}^k}$ and $p_{\hat{\varphi}_{ia}^k}$ over several iterations prior to making adjustment.

$$\hat{\varphi}_{ia}^{(k+1)N} = \hat{\varphi}_{ia}^{kN} + \xi \hat{L}_{kN}, \tag{2.12}$$

  where

$$\hat{L}_{kN} = \frac{1}{N} \sum_{n=0}^{N-1} \log \frac{p_{\hat{\varphi}_{ib}^{kN}}(x(kN+n))}{p_{\hat{\varphi}_{ia}^{kN}}(x(kN+n))}.$$

Convergence analysis is as follows. If $\mathcal{D}_i^k > 0$, φ_{ia}^{k+1} and φ_{ib}^{k+1} will grow according to (2.9) and (2.10), so that $\mathcal{D}_i^{k+1}$ will decrease due to the concavity of the function. Similarly, if $\mathcal{D}_i^k < 0$, φ_{ia}^{k+1} and φ_{ib}^{k+1} will become smaller so that $\mathcal{D}_i^{k+1}$ will increase. In both cases, $\mathcal{D}_i^k$ converges to zero surely. δ_i is a key factor in our algorithm. On the one hand larger δ_i leads to faster convergence. On the other hand, δ_i is the range for the unknown parameter, which is desired to be small. Therefore, choosing δ_i involves certain tradeoffs which depends on applications in practice.

2.4.2 Change detection

In the process of tracking the unknown parameters, change detection is conducted through an appropriate CUSUM test simultaneously. Basically, after the parameter tracking in time t, a

value is specified in each parameter's new range and substituted into the CUSUM test, i.e., Eq. (2.5) with $s(t) = \max(s(t-1) + l_{\hat{\varphi}}(t), 0)$, where $\hat{\varphi}$ is the parameter vector composed of the specified values of all the unknown parameters. As we discussed at the end of last section, we expect the range size δ_i as small as possible and the algorithm converges as fast as possible. However, they can not be achieved simultaneously. In the following, we'll discuss how to exploit the non-optimality coefficient to calculate δ_i to give a best tradeoff between convergence and performance. We first propose an approach for one unknown parameter case and then discuss the problems extending the approach to multiple unknown parameters.

Define a non-optimality coefficient ε_φ and the maximum non-optimality coefficient ε_m as:

$$\varepsilon_\varphi = 1 - \frac{\bar{T}_1^{opt}}{\bar{T}_1^{\varphi}}, \tag{2.13}$$

$$\varepsilon_m = \sup_{\varphi \in [\varphi_{min}, \varphi_{max}]} \varepsilon_\varphi, \tag{2.14}$$

where $\bar{T}_1^{opt}$ is the optimal detection delay given by the Page's CUSUM Test when φ is known and $\bar{T}_1^{\varphi}$ is the detection delay achieved by the Parallel CUSUM Test or our Adaptive CUSUM Test when the true value φ is unknown.

When there is one unknown parameter, given ϵ_m, we can predetermine L candidates a_j, according to the Parallel CUSUM Test [10], and confidence intervals $[\underline{a_j}, \bar{a_j}]$ associated with these candidates. [3] δ[4] can be determined as the minimal confidence interval, which is associated with the lth candidate. Specifically,

$$\delta = \min_{1 \le j \le L} (\bar{a_j} - \underline{a_j}); \quad l = \arg \min_{1 \le j \le L} (\bar{a_j} - \underline{a_j})$$

After (2.9) and (2.10), the value used in the CUSUM test is updated as:

$$\bar{\varphi}^k = \varphi_a^k + \alpha\delta,$$

where $\alpha = 1/2$ when $F(\hat{\varphi})$ is symmetric with respect to φ, or $\alpha = \frac{a_l - \underline{a_l}}{\delta}$ when $F(\hat{\varphi})$ is asymmetric. $\bar{\varphi}^k$ converges to some value $\bar{\varphi}$ when the parameter tracking procedure converges. If $\bar{\varphi}^k$ is out of the range $[\varphi_{min}, \varphi_{max}]$, it is set to φ_{min} or φ_{max}, whichever is closer. Our algorithm continues by substituting $\bar{\varphi}^k$ as the true value of φ in the CUSUM test.

Under H_1, $\bar{\varphi} = \varphi$ for the symmetric $F(\hat{\varphi})$ since when the algorithm converges, φ lies right in the middle of $\hat{\varphi}_a$ and $\hat{\varphi}_b$. In such a case this procedure provides an adaptive estimation for the unknown parameter. For the asymmetric $F(\hat{\varphi})$, $\bar{\varphi} \neq \varphi$ but $\varepsilon_{\bar{\varphi}} \le \varepsilon_m$ since the true value is within δ after convergence and δ is the confidence interval of $\bar{\varphi}$, which is equal to the minimal

[3] These candidates guarantee that the non-optimality coefficient ϵ_φ is less than ϵ_m for any true value φ. If the true value is within the confidence interval associated with a candidate, say a_1, $\varepsilon_{a_1} \le \varepsilon_m$.

[4] we omit the subscript i here since there is only one unknown parameter.

interval in the Parallel CUSUM Test. And $\varepsilon_{\bar{\varphi}}$ is actually much less than ε_m shown by the simulation result in the next section.

Extending the above approach to multiple parameters is theoretically feasible, however in practice it could be quite difficult to do so. First finding multiple candidates, each of which is a vector, and their associated "confidence space" is quite challenging since the Eq. (2.13) and (2.14) do not provide enough information for the candidates. Secondly, based on these confidence space, specifying δ_i for parameter i is nontrivial. Let's take two unknown parameters for example, in which "confidence space" becomes "confidence area". To find out δ_i we have to find out the largest square within each confidence area, which is time consuming. Therefore how to choose δ_i for multiple parameters needs further study.

The Adaptive CUSUM Test is summarized as follows:

1. Initialization: for parameter i, set an initial value φ_{ia}^0 and a new range size δ_i.
2. In time k,
   - update the estimation of parameter i φ_{ia}^k according to the parameter tracking approach;
   - specify a proper value within $\hat{\Phi}_i$;
   - substitute the specified value of each parameter into the CUSUM test.
3. Stop: The algorithm continues until the accumulated statistic $s(k)$ in the CUSUM test is larger than the threshold h.

This approach is suboptimal in nature. However the simulations in the next section show that its performance is close to the optimal one by choosing a fairly small range size for each parameter.

2.4.3 Performance analysis

The detection delay can be approximated by [8]

$$\bar{T}_1 \approx \frac{h}{E_1[l_{\bar{\varphi}}(t)]} \quad as \quad h \to \infty,$$

Under H_0, the mean time between false alarms admits [8]:

$$\bar{T}_0 \geq e^h,$$

which means there is no performance loss in term of $\bar{T}_0$, unlike the Parallel CUSUM Test in which $\bar{T}_0 \geq \frac{1}{L} e^h$.

2.5 Numerical Result

2.5.1 Single parameter

Symmetric $F(\hat{\varphi})$

In this section we take detecting a sinusoid wave with an unknown amplitude for an example to demonstrate the performance improvement of our Adaptive CUSUM Test over the Parallel CUSUM Test. The following hypotheses are assumed:

$$\begin{cases} H_0 : x(t) = n(t) \\ H_1 : x(t) = A\sin(wtT_s) + n(t), \end{cases}$$

where A is an unknown amplitude within the range $[2, 36]$; w is the carrier frequency and T_s is the sampling period, both of which are known. $n(t)$ is Gaussian noise with zero mean and unit variance.

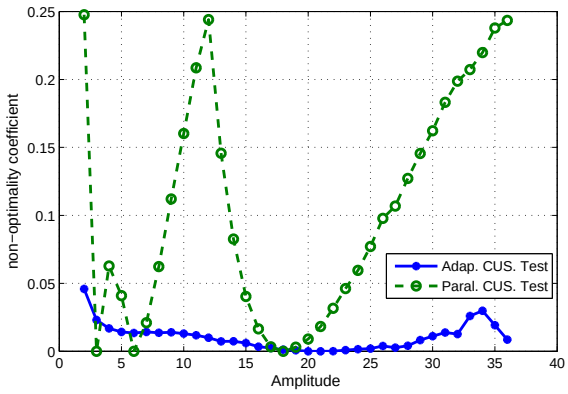

Figure 2.1: Performance comparison : fixed amplitude.

Denote by $\hat{A}$ an estimate of amplitude A. It is easy to check $F(\hat{A}) = E[l_{\hat{A}}(t)]$ is symmetric about A. We choose three candidates for the Parallel CUSUM Test so that its computation complexity is comparable with the Adaptive CUSUM Test. We set ε_m to 0.25 and the corresponding three candidates of the Parallel CUSUM Test are $3, 6, 18$. The thresholds for both tests are the same and set to 5000.

Fig. 2.1 compares non-optimality coefficients of our Adaptive CUSUM Test and the Parallel CUSUM Test for different amplitudes, staying fixed during the detection process. We can see non-optimality coefficients of the Parallel CUSUM Test are substantially larger than those of our Adaptive CUSUM Test at almost all possible amplitudes (except for the three candidates chosen

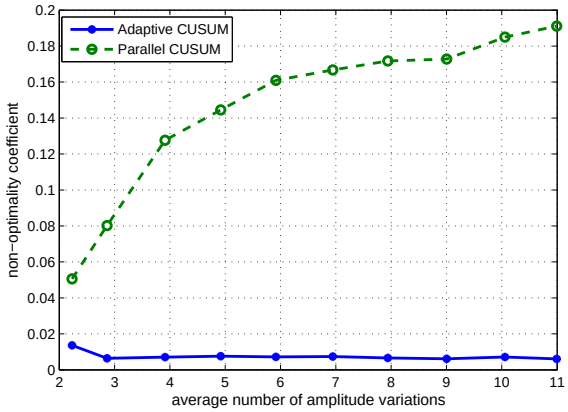

Figure 2.2: Performance comparison : time-varying amplitude

by the Parallel CUSUM Test). The average non-optimality coefficient of the Parallel CUSUM Test is 0.11, while for our Adaptive CUSUM Test, it is 0.01, indicating little degradation in optimality. In Fig. 2.2, it is assumed that the sinusoid wave goes though a block fading channel, where the amplitude changes every 600 time slots randomly within the range rather than fixed. We compare the average non-optimality coefficient of these two tests under different average number of variations, and for each point of Fig. 2.2, 250 independent experiments are conducted to guarantee its fidelity. We can see that the performance of our Adaptive CUSUM Test is stable and close to the optimal detection since it can track the amplitude change, while the Parallel CUSUM Test performs worse as more variations are involved.

Asymmetric $F(\hat{\varphi})$

To demonstrate the performance of our algorithm in an asymmetric case, we detect the change of the mean in a poisson distribution. The hypothesis is given by (2.1), where $\lambda = 1$ and $\Phi = [2, 40]$. The non-optimality coefficient $\epsilon_m = 0.25$ for the Parallel CUSUM Test, which results in three candidates 2.6, 8.1 and 56.9.

Fig. 2.3 compares non-optimality coefficients of these two tests. Surprisingly, the Adaptive CUSUM performs much better than the Parallel CUSUM Test even the true value of the parameter has not been achieved.

2.5.2 Multiple parameters

In this simulation, we assume a Gaussian variable changes its distribution from $\mathcal{N}(0, 1)$ to $\mathcal{N}(m, \sigma^2)$, where both m and σ are unknown to the detector and $m \in [3, 23], \sigma \in [2, 4]$. The simulation is composed of 50 rounds and in each round m and σ are randomly chosen from

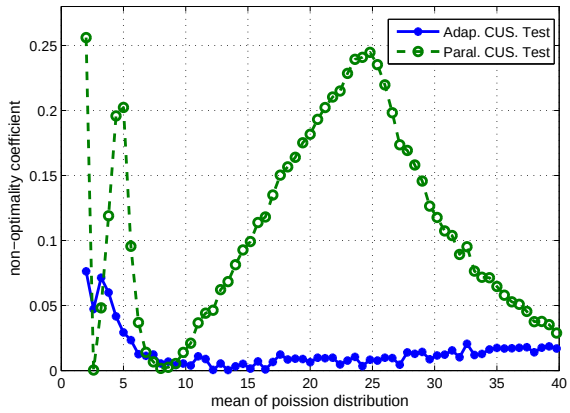

Figure 2.3: Performance comparison for the asymmetric case.

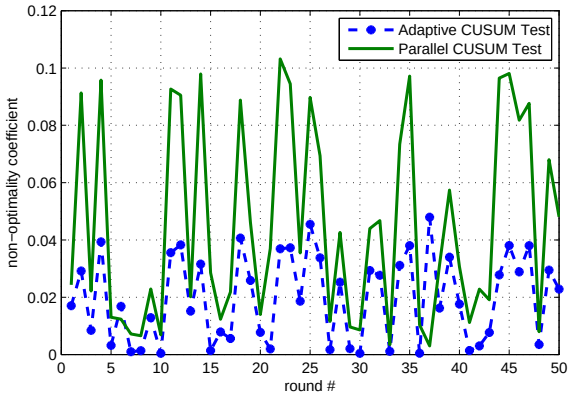

Figure 2.4: Performance of the Adaptive CUSUM test for Gaussian signal.

their ranges. Set $\delta_m = 6$ and $\delta_\sigma = 1$ in our adaptive CUSUM test and specify the values of m and σ in the middle of their ranges. As we discussed in the last section it's difficult to choose candidates for the Parallel CUSUM Test in the two parameters case. We choose them by experiments. For simplicity, we apply the true variance in the parallel CUSUM Test and choose the candidates 3, 6, 12 for the mean to guarantee that its non-optimality coefficient is less than 0.1. Figure 2.4 shows the non-optimality coefficient of both algorithms. We can see from the figure that the non-optimality coefficient of the Adaptive CUSUM Test is below 0.05, far outperforming the Parallel CUSUM Test. In addition compared to the single parameter cases in Fig.2.1 and 2.3, multiple unknown parameters do not degrade the performance of our approach notably.

2.6 Conclusions

In this chapter, we have studied quickest change detection with unknown parameters. An adaptive CUSUM algorithm is proposed to narrow down the ranges of the unknown parameters and track their possible change during the detection process. Analysis and numerical results show that the new algorithm achieves better performance than the Parallel CUSUM Test.

Chapter 3

On the Throughput Scaling of Cognitive Radio Ad Hoc Networks

3.1 Introduction

3.1.1 Motivation

The conflict between increasing demands for bandwidth and scarcity of spectrum in wireless communication strongly propels the study of Cognitive Radio (CR) technology in recent years [31]. This technology aims at providing a flexible way of spectrum management, permitting CR (secondary) users to temporally access spectrum that is not currently used by legacy (primary) users. In many cases, it is preferable that the operation of the secondary network is transparent to the primary network. Due to its secondary role, the CR network should prevent any unacceptable interference to the primary network, while tolerate the interference from the primary transmissions.

One primary concern for CR networks is network throughput. In [32], the authors showed that an n-node primary network and an m-node secondary network, while coexisting, can achieve the per-node throughput $\Theta(\sqrt{\frac{1}{n \log n}})$ and $\Theta(\sqrt{\frac{1}{m \log m}})$, respectively. The same results were achieved in [13] with a more practical assumption, requiring only the knowledge of the locations of primary transmitters, rather than the locations of all the primary nodes as in [32]. Both [32] and [13] focus on the dense network model, where the number of nodes grows with the node density in a fixed area, and assume that $m = n^{\beta}$ ($\beta > 1$), i.e., the density of the secondary network is higher than that of the primary one in the order sense. Under this assumption the influence of the primary network on the secondary network becomes increasingly negligible (so that full connectivity[1] of the secondary network is guaranteed); in addition, it is only valid in

[1] Given a transmission range r there is always a path connecting two arbitrary nodes in the network.

the dense network model and does not hold in the extended network model, where the number of nodes grows with the area with a fixed density. These facts motivate us to study a more practical scenario where $\frac{m}{n}$ = constant in this paper, i.e., the dimensions of the two overlaid networks are on the *same* order. We mainly consider an *extended* network[2] and show that even with this weaker condition the per-node throughput of both networks can be further boosted to $\Theta(\frac{1}{\sqrt{n}})$ and $\Theta(\frac{1}{\sqrt{m}})$, respectively, if m/n is larger than a threshold. Our results are derived with the help of intelligent design of "highway" systems originally proposed in [34]. It should be noted that 1) the weaker condition $\frac{m}{n}$ = constant no longer maintains full connectivity of the secondary network in the setting of [32] (see Appendix A.1 for more details) so that the multi-hop routing approach there may be infeasible, which advocates the necessity and effectiveness of our approach; 2) the states of the links in the secondary network are correlated (due to the influence of the primary network) instead of independent as in [34], which complicates the analysis.

3.1.2 Related works

The capacity of ad hoc networks has been extensively explored since the seminal work of Gupta and Kumar [18]. It was shown in [18] that in an arbitrary network where n nodes are arbitrarily located on the plane, per-node throughput of $\Theta(\frac{1}{\sqrt{n}})$ is achievable. In a random planar network, where n nodes are uniformly and independently distributed, per-node throughput only scales as $\Theta(\sqrt{\frac{1}{n \log n}})$. The gap was bridged by Franceschetti *et al.* in [34] through a percolation-based routing scheme. Both [18] and [34] simply treated simultaneous transmissions as interference. *Özgür et al.* proposed a hierarchical cooperation scheme incorporating distributed MIMO technique, and showed that the total throughput of the network could scale linearly with the number of nodes [35]. The same problem has also been studied through the information-theoretic approach in [36–38].

The capacity study of CR ad-hoc networks is more challenging in nature than that of a stand-alone network. Here is a brief overview of some important works in this area. The cognitive channel, an interference channel composed of one source-destination (S-D) pair from each of the two networks, was studied in [39,40] from an information-theoretic perspective. The throughput of a one-hop cognitive network was investigated in [11]. Recently the throughput in multi-hop CR networks was considered in [13, 32], which showed that there is no performance loss for CR networks, in terms of scaling law. Note that the same results could be trivially achieved if cooperations between two networks are permitted. Therefore, as [13,32] we preclude such an option in our study.

[2]In many scenarios, the (scaling law) results obtained in the dense and extended network model coincide, but it is felt that the extended network model is more realistic, where one is free from the concern of the near field effects of electromagnetic propagation [33].

The remainder of this chapter is organized as follows. The system model and problem formulation are given in Section 3.2. The primary and secondary protocols are presented in Section 3.3, together with our main results. Their throughput analysis is given in the following two sections. Finally Section 3.6 concludes this chapter.

3.2 System Model and Problem Formulation

3.2.1 Network model

Consider a primary network with n nodes randomly and uniformly deployed in a square $S_n \triangleq [0, \sqrt{n}] \times [0, \sqrt{n}]$. And m secondary nodes are distributed randomly and uniformly in the same area, sharing the same resources with the primary network. The ratio of their network sizes $\lambda_s = m/n$ is assumed a *constant*. Note that as $n \to \infty$, the positions of both primary and secondary nodes converge in distribution to homogeneous Poisson point processes (P.P.P.) with density 1 and λ_s, respectively. It is assumed that time is slotted and synchronized for both networks. Primary nodes are paired uniformly at random so that each node is the destination of exactly one source. The secondary source-destination pairs are randomly grouped similarly. The primary network performs as if it stands alone, while the secondary network accesses the spectrum *opportunistically* to prevent unacceptable interference to the primary network.

3.2.2 Propagation model

We assume identical transmission power P_p for all the primary nodes, and P_s for all the secondary nodes, respectively. For simplicity, we only consider path loss for the physical channels (as in the majority literature), i.e., the power attenuation function is given by:

$$l(d) = \min\{1, d^{-\alpha}\}, \tag{3.1}$$

where d is the Euclidean distance between a transmitter and a receiver, and $\alpha > 2$ is the path loss exponent.

3.2.3 Transmission rate model

We consider that the transmission rate R from a transmitter X_i to its corresponding receiver $X_{D(i)}$ is a continuous function of the Signal to Interference plus Noise Ratio (SINR) at $X_{D(i)}$, i.e.,

$$R(X_i, X_{D(i)}) = \log(1 + SINR). \tag{3.2}$$

Denote by $\{X_{p,k}; k \in \mathcal{T}_1\}$ the subset of concurrent primary transmitters, and $\{X_{s,k}; k \in \mathcal{T}_2\}$ the subset of concurrent secondary transmitters. For a primary transmitter $X_{p,i}$ the SINR at its receiver $X_{p,D(i)}$ is given by:

$$SINR_p = \frac{P_p l(\parallel X_{p,i} - X_{p,D(i)} \parallel)}{N_0 + I_p + I_{sp}}, \tag{3.3}$$

where N_0 is the noise power at the receiver; I_p is the interference power from all the other primary transmitters to the receiver $X_{p,D(i)}$, given by:

$$I_p = P_p \sum_{k \in \mathcal{T}_1, k \neq i} l(\parallel X_{p,k} - X_{p,D(i)} \parallel); \tag{3.4}$$

and I_{sp} is the interference power from all the secondary transmitters to the receiver $X_{p,D(i)}$, given by:

$$I_{sp} = P_s \sum_{k \in \mathcal{T}_2} l(\parallel X_{s,k} - X_{p,D(i)} \parallel). \tag{3.5}$$

The SINR at a secondary receiver is defined similarly as:

$$SINR_s = \frac{P_s l(\parallel X_{s,i} - X_{s,D(i)} \parallel)}{N_0 + I_s + I_{ps}}, \tag{3.6}$$

where I_s is the interference power from all the other secondary transmitters to the receiver $X_{s,D(i)}$, and I_{ps} is the interference power from all the primary transmitters to the receiver $X_{s,D(i)}$, defined similarly as Eq. (3.4) and (3.5).

3.2.4 Problem formulation

We study the throughput of both networks, which is the average number of bits per second that all source nodes can simultaneously transmit to their destinations w.h.p.. Formally the throughput of a network of size n is defined as:

Definition 1. *The throughput per S-D pair $T_\Pi(n)$ in a network of size n under some scheduling scheme Π is defined as the maximal quantity satisfying*

$$\Pr\left(\min_i \lim_{t\to\infty} \inf \frac{1}{t} B_{i,\Pi}(t) \geq T_\Pi(n)\right) \to 1, \tag{3.7}$$

as $n \to \infty$, where $B_{i,\Pi}(t)$ is the number of bits that S-D pair i can transfer in t time slots.

Note that the above definition of throughput is an asymptotical property, therefore we require the number of nodes in both networks go to infinity (with a fixed ratio) in our analysis.

We use the following probabilistic order notations throughput our paper. Given non-negative functions $f(n)$ and $g(n)$:

- $f(n) = \Omega(g(n))$ w.h.p. if there exists a positive constant c_1 such that $\lim_{n\to\infty} \Pr(f(n) \geq c_1 g(n)) = 1$.
- $f(n) = O(g(n))$ w.h.p. if there exists a positive constant c_2 such that $\lim_{n\to\infty} \Pr(f(n) \leq c_2 g(n)) = 1$.
- $f(n) = \Theta(g(n))$ w.h.p. if $f(n) = \Omega(g(n))$ and $f(n) = O(g(n))$.

3.3 Network Protocols and Main Results

In this section we introduce the protocols for both the primary and secondary network, followed by their achieved throughput.

3.3.1 Primary protocol

The primary network adopts the highway system proposed in [34]. Rooted from percolation theory, the highway system is composed of multiple horizontal and vertical paths, and every primary node in the plane can access at least one horizontal and vertical path through one hop (see Fig. 3.1). The primary protocol is summarized below :

- Tessellation: the square S_n is divided into $\frac{\sqrt{n}}{\sqrt{2}a_p k_p \log \frac{\sqrt{n}}{\sqrt{2}a_p}}$ horizontal corridors, each with dimension $\sqrt{n} \times \sqrt{2}a_p k_p \log \frac{\sqrt{n}}{\sqrt{2}a_p}$, where a_p and k_p are some constants, independent of n. The corridors are then tessellated into diamond cells with side length a_p (see Fig. 3.2); a highway within each corridor, composed of $\delta_p \log(\sqrt{n})$ horizontal paths for some constant δ_p, is built according to [34], using percolation theory (note that only one path is shown in Fig. 1 for simplicity). Similarly we divide the square S_n into $\frac{\sqrt{n}}{\sqrt{2}a_p k_p \log \frac{\sqrt{n}}{\sqrt{2}a_p}}$ vertical corridors, each with dimension $\sqrt{2}a_p k_p \log \frac{\sqrt{n}}{\sqrt{2}a_p} \times \sqrt{n}$ and $\delta_p \log(\sqrt{n})$ vertical paths.
- Routing: three phases are involved in the routing scheme:
  - access phase: the source drives its packet to one of the multiple horizontal paths in the horizontal corridor it is located, through one single hop;
  - express relay: the packet traverses horizontally and then vertically on the highway through multiple hops;
  - delivery phase: the packet is finally delivered to the destination by a node on a vertical path, through one single hop.

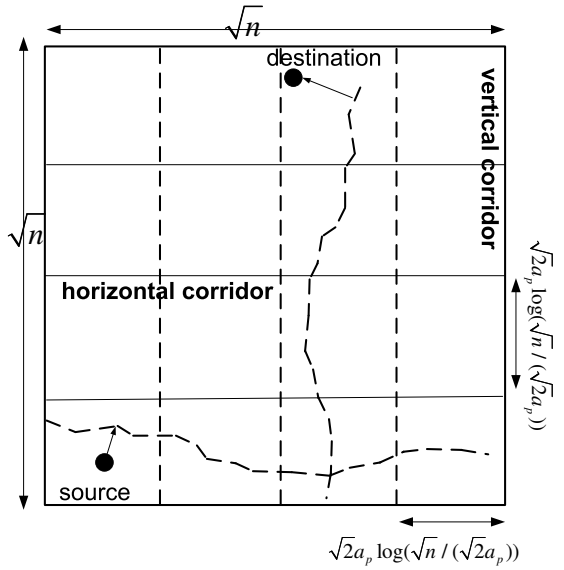

Figure 3.1: The highway system

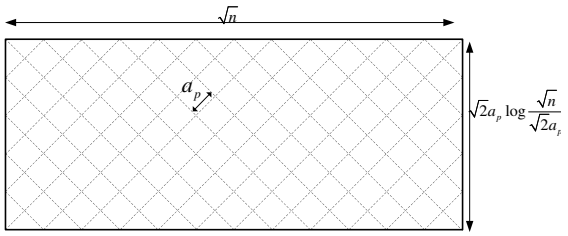

Figure 3.2: Tessellation of a primary horizontal corridor

- Transmission power: the transmission power of a primary transmitter is $P_p = Pa_p^\alpha$, where P is a constant [3], for all three routing stages.
- Scheduling: each time slot is further divided into three sub-slots. For the primary network, the three sub-slots correspond to the three routing phases, shown in Fig. 3.3. The $\Theta(\log n)$-TDMA scheme is used during the first and third routing phase and the 9-TDMA scheme is adopted during the highway transmission. The M-TDMA scheme is described in Appendix A.2.

3.3.2 Secondary protocol

The design of the secondary protocol is challenging. It is required that the secondary network keep its interference to the primary network at an acceptable level, which may limit its operation

[3]The same P is used in the secondary protocol to maintain the power ratio between two networks.

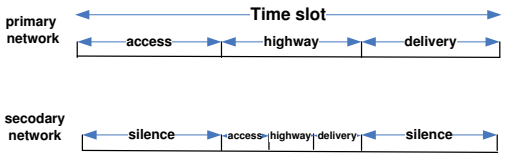

Figure 3.3: Time slot for both networks

and performance greatly. Inspired by the highway system design in [34], we propose a non-trivial secondary protocol, which performs as well as the highway system in a stand-alone network, while satisfies the above requirement.

The secondary protocol is summarized as follows:

- Tessellation: the square S_n is divided into $\frac{\sqrt{n}}{\sqrt{2}a_s k_s \log \frac{\sqrt{n}}{\sqrt{2}a_s}}$ horizontal corridors, each with dimension $\sqrt{n} \times \sqrt{2}a_s k_s \log \frac{\sqrt{n}}{\sqrt{2}a_s}$, where a_s and k_s are some constants, depending on $\lambda_s = m/n$ and independent of m. The corridors are then tessellated into diamond cells with side length a_s (see Fig. (D.2)). Correspondingly we divide the square into $\frac{\sqrt{n}}{\sqrt{2}a_s k_s \log \frac{\sqrt{n}}{\sqrt{2}a_s}}$ vertical corridors, each with dimension $\sqrt{2}a_s k_s \log \frac{\sqrt{n}}{\sqrt{2}a_s} \times \sqrt{n}$. A highway system is constructed in each corridor and it is shown in Theorem 3 that for large enough λ_s there are $\Theta(\log \sqrt{m})$ horizontal (vertical) disjoint pathes in each of the horizontal (vertical) corridors.

- A preservation region (see Fig.3.4) is set around each primary node, which is defined as a square composed of 9 secondary cells. The purpose of preservation regions is to limit the interference from the secondary transmitters to primary receivers. All the secondary nodes located in the preservation regions must keep "silent", i.e., they can not serve as transmitters or relays.

- Routing: the routing scheme of the secondary network is similar to that of the primary network, i.e., the routing is composed of three phases: access phase, express relay and deliver phase.

- Transmission power: the transmission power of a secondary transmitter is $P_s = Pa_s^\alpha$ for all three routing stages. Therefore the power ratio between the primary and secondary network is $(a_p/a_s)^\alpha$.

- Scheduling: the secondary network shares the same time frame structure with the primary network. The time slot is also divided into three sub-slots (Fig. 3.3). During the first and third sub-slot all the secondary nodes keep silent. The second sub-slot is further divided into three mini-slots, which correspond to three routing phases of the secondary network.

The $\Theta(\log m)$-TDMA scheme is used during the first and third routing phase and the 9-TMDA scheme is adopted during the highway transmission.

According to the primary and secondary protocol above, our main results are summarized below.

Theorem 1. *Under the primary and secondary protocol given in this section, the primary network achieves per-node throughput $\Theta(1/\sqrt{n})$, in the presence of the secondary network.*

Theorem 2. *Under the primary and secondary protocol given in this section and given the size ratio $\lambda_s > \min(5400, \lambda_s^*)$ and outage constraint[4] τ, the per-node throughput of the secondary network is $\Theta(\frac{1}{\sqrt{m}})$, where λ_s^* is the solution to the following equation*

$$\frac{\log(1+\lambda_s/9)}{\lambda_s} = \frac{\tau}{18}, \forall \epsilon > 0.$$

Remark 1. *As will be seen in our analysis below, the outage probability of the secondary network depends on the ratio m/n. The higher the ratio, the lower the outage is. The two theorems above imply that, at least in terms of the scaling law, there is no performance loss for either of the two coexisting networks; this is particularly interesting as to the CR network despite its secondary role.*

3.4 Throughput Analysis of the Secondary Network

In this section, we focus on the analysis of the per-node throughput of the *secondary* protocol, which is our primary concern. The corresponding analysis for the primary network will be given in the next section. We first show how to construct a highway with consideration of preservation regions. Then we evaluate the one-hop data rate achievable in each of the three routing phases, based on which we obtain the per-node throughput.

3.4.1 Highway construction

We describe the approach to build the highway in a horizontal corridor, which applies almost verbatim to a vertical corridor as well. As we mentioned, each (horizontal) corridor is partitioned into diamond cells of side length a_s. We call such a cell open (see Fig. (3.4)) if both of the following two events happen:

1. E_1 : there is at least one secondary node in the cell;

[4]The outage probability is defined as the percentage of the secondary S-D pairs which can not be served.

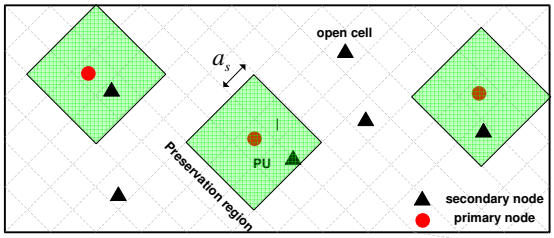

Figure 3.4: Preservation regions of the secondary network

2. E_2 : there is no primary node in either this cell or its 8 neighboring cells; this event differentiates the CR network with a stand-alone network studied in [3].

Otherwise we call it closed. According to the properties of P.P.P., the probabilities of the two events are given by:

$$p_{E_1} \triangleq \Pr(E_1) = 1 - (1 - \frac{a_s^2}{n})^m,$$
$$p_{E_2} \triangleq \Pr(E_2) = (1 - \frac{9a_s^2}{n})^n.$$

Due to the independence of E_1 and E_2, a cell is open with probability

$$\begin{aligned} p \triangleq p_{E_1} p_{E_2} =& [1 - (1 - \frac{a_s^2}{n})^m](1 - \frac{9a_s^2}{n})^n \\ \rightarrow& (1 - e^{-a_s^2 m/n}) e^{-9a_s^2}, \end{aligned} \tag{3.8}$$

where the convergence is made due to the fact that $(1 - \frac{c}{n})^m \rightarrow e^{-cn/m}$ for some constant c, as n and m goes to infinity.

This probability is desired to be as large as possible. To maximize p, a_s is set as

$$a_s = \sqrt{\frac{\log(1 + \lambda_s/9)}{\lambda_s}}, \tag{3.9}$$

after some calculations. p is then given by:

$$p = \frac{\lambda_s}{9 + \lambda_s} (\frac{9}{9 + \lambda_s})^{9/\lambda_s}. \tag{3.10}$$

Some interesting observations on p are in order:

1. The value of p increases with the size ratio $\lambda_s = m/n$. Therefore by choosing appropriate λ_s, p can be set as large as possible.

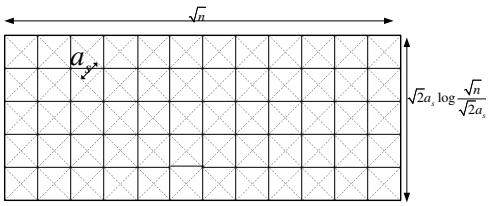

Figure 3.5: Tessellated secondary horizontal corridor and grid construction

2. The state of a secondary cell (open or closed) is *dependent* on the states of its neighboring cells, since one primary node can cause 9 secondary cells closed. In contrast, the cell states in a stand-alone network such as the one considered in [34] are *independent* from each other.

A key step in highway construction is to find (or ensure the existence of) sufficient paths across the network. For this purpose we map each of the tesselated secondary corridor into the bond percolation model on a $r \times k_s \log r$ grid G_r, where $r = \frac{\sqrt{n}}{\sqrt{2}a_s} = \frac{\sqrt{m}}{\sqrt{2\log(1+\lambda_s/9)}}$. Edges of the grid are composed of horizontal diagonals of some cells and vertical diagonals of the other cells. The grid G_r is given in Fig. D.2, where the dashed lines represent the tessellation of the corridor and the solid lines represent the grid. The states (open or closed) of the edges are the same as their corresponding cells, thus they are not independent either in G_r. For a horizontal corridor, an L(eft)R(ight) open path corresponds to a sequence of connected open edges on the grid, formally defined below.

Definition 2. *In a 2-D grid $G \triangleq [0, n] \times [0, \log n]$, let $C_v = (x_v, y_v)$ be the coordinates of a vertex v. If there exist a series of vertices v_1, v_2, ..., v_m, such that there is an open edge connecting two consecutive vertices and $\forall 1 \leq i < m$, $0 \leq x_{v_{i+1}} - x_{v_i} \leq 1$, $x_{v_1} = 1$ and $x_{v_m} = n$, then the path consisting of these edges is called an LR open path.*

For each open edges, there is at least one secondary node located in the corresponding cell which stays outside preservation regions. Therefore, an LR path can be mapped back to a routing path in the network. An LR path and its corresponding routing path are depicted in Fig. 3.6, where the dark solid line is an LR path and the dark dashed line is an actual routing path.

We show in the next theorem that with *dependent* edges, there are $\delta_s \log r$ paths in a horizontal corridor, for some constant δ_s. The same conclusion holds for a vertical corridor.

Theorem 3. *With $p > 80/81$, there are $\delta_s \log r$ LR paths in each horizontal corridor for some constant δ_s.*

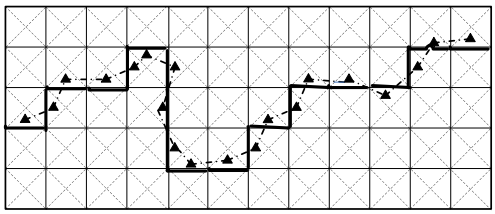

Figure 3.6: A routing path

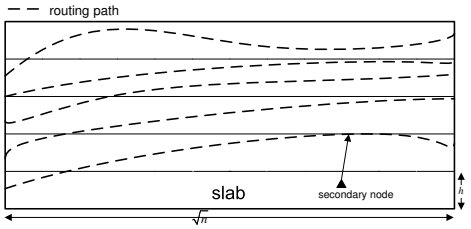

Figure 3.7: Slabs in a horizontal corridor

The proof is given in Appendix A.3. The key difference with [34] is that we need to deal with the case that the states of cells are dependent due to the introduction of preservation regions.

A horizontal (vertical) corridor is further evenly divided into multiple slabs with size $\sqrt{n} \times h$ ($h \times \sqrt{n}$), where $h = k_s/\delta_s$ is a constant. The value of h guarantees that the number of slabs are as many as the open paths in a corridor so that the nodes in each of the slabs share exactly one path. The slabs in a horizontal corridor are shown in Fig. 3.7, where there are 5 slabs corresponding to 5 routing paths.

All the nodes in a cell share one access point, which is the closest node located on their corresponding horizontal path. And in the access phase, source nodes load data to their access points through a single hop with length at most $\sqrt{2}a_s \log(\sqrt{n}/(\sqrt{2}a_s))$ (the width of the corridor). The operations in the delivery phase follows a similar way. One hop in a highway has to support the traffic of all the nodes in a slab since all the nodes in a slab drive their data to the same path.

3.4.2 Throughput analysis

We first calculate the number of nodes in a cell and a slab, and then evaluate the throughput of one hop during each phase.

Lemma 1. *The number of nodes in a cell is at most* $\log m$*, and the number of nodes in a slab is at most* $2h\sqrt{m/n}\sqrt{m}$*, both w.h.p..*

This lemma is direct application of the Vapnik-Chervonenkis Theorem given in Appendix A.4, and its proof is given in Appendix A.5.

The following two lemmas show that one hop on the highway can support a constant rate and the throughput for an access (delivery) link scales as $\Omega(1/(\log m)^{\alpha})$.

Lemma 2. *On the highway of the secondary network, each secondary cell can support traffic with a constant rate K_s, where $K_s > 0$ is independent of m.*

The proof is given by Appendix A.6.

Lemma 3. *The throughput for an access (delivery) link scales as $\Omega(1/(\log m)^{\alpha})$.*

The proof is given by Appendix A.7.

During the access and delivery phase, each link supports $\Omega(1/(\log m)^{\alpha})$ data rate (Lemma 3) and $\log m$ nodes (Lemma 1) share it. Therefore, the per-node throughput during these two phases scales as $\Omega(1/(\log m)^{\alpha+1})$. And on the highway one hop can support a constant data rate (Lemma 2) and there are at most $2h\sqrt{m/n}\sqrt{m}$ nodes (Lemma 1) sharing the same hop. Therefore the per-node throughput on the highway scales as $\Theta(1/\sqrt{m})$. Considering the throughput of all the phases we reach the conclusion of Theorem 2.

3.4.3 Outage analysis

According to the secondary protocol, outage occurs inevitably since any S-D pairs associated with the secondary nodes in the preservation regions can not be served. Since the total area of preservation regions is bounded by $n9a_s^2$, the probability that a secondary node stays in a preservation region is upper bounded by $\frac{n9a_s^2}{n} = 9a_s^2$, which is a constant. Denote by m_u the number of secondary nodes in all the preservation regions. Then according to Hoeffding inequality in Appendix A.8,

$$\Pr(m_u - m9a_s^2 \geq 9ma_s^2) \leq e^{-2m(9a_s^2)^2} \to 0 \text{ as } m \to \infty.$$

Thus,

$$m_u < 18ma_s^2 \text{ w.h.p. .}$$

The outage probability P_o can be calculated as follows:

$$P_o = \frac{m_u}{m} \leq \frac{18ma_s^2}{m} = 18a_s^2 = 18\frac{\log(1+\lambda_s/9)}{\lambda_s},$$

which is a constant depending on the network size ratio m/n. Therefore, higher size ratio leads to lower outage probability. Note that if the higher order condition $m = n^{\beta}(\beta > 1)$ is applied,

the outage probability is vanishing, w.h.p., which coincides with the conclusion in [32]. Solve $P_o(\lambda_s) = \tau$, we get λ_s^*.

The proof of Theorem 2 is then completed.

3.5 Throughput Analysis of the Primary Network

In this section we briefly analyze the throughput of the primary network, focusing on the key difference from [34]. Since the primary protocol is the same as the one in [34], so is the analysis on the access and delivery phase, when no secondary transmissions are allowed. This indicates that transmission rate $\Theta(1/\sqrt{n})$ is achievable during these two phases. Different from a real stand-alone network, the primary network suffers from additional interference from the secondary network during the express relay phase on the highway. However the next lemma shows that due to our design, the influence of the secondary network is negligible, at least up to scaling.

Lemma 4. *On the highway of the primary network, each primary cell can support traffic with a constant rate K_p, where $K_p > 0$ is independent of n.*

The proof is given by Appendix A.9.

According to [34], at most $O(\sqrt{n})$ nodes share one hop on the highway. Therefore, a transmission rate $\Omega(\frac{1}{\sqrt{n}})$ is achieved, the same as a stand-alone network. Theorem 1 is thus proved.

3.6 Conclusions

In this chapter, we studied the throughput of a type of heterogeneous networks consisting of a primary network of size n and a cognitive radio ad hoc network of size m under a more practical model. We show that the per-node throughput of the primary network scales as $\Theta(1/\sqrt{n})$ and there is indeed no performance loss, in terms of scaling law, with the coexistence of the CR network. The CR network can also achieve per-node throughput $\Theta(1/\sqrt{m})$, performing as well as a stand-alone network except suffering from a non-vanishing outage probability.

Chapter 4

Transmission Throughput of Decentralized Overlaid Networks with Outage Constraints

4.1 Introduction

It is of great interest to understand to what extent the secondary network in overlaid networks can gain in transmission of its own useful data, without harming the regular operation of the primary network significantly. Relevant research in literature mainly targets either the scaling law of transport capacity [11–13] or transmission capacity with outage consideration [14, 15, 17, 41]. The former was proposed in the seminal work [18] and defined as the maximum bit·meters per second the network can achieve in aggregate; pioneer work in the capacity study of overlaid or cognitive radio networks [11–13] showed that there is no performance loss for the secondary network in terms of *scaling law* of transport capacity. Nonetheless, asymptotic analysis on the scaling law only characterizes the (rough) relationship between capacity and network size, neglecting the effect of many important system parameters. As an alternative, transmission capacity [19] quantifies the maximum allowable spectrum efficiency per unit area under some outage probability constraint, and sparks an enormous interest recently (see [20] for an overview). The outage probability of the secondary network was studied in [14, 15]. It was pointed out in [17] that transmission capacity is not very applicable in the secondary network since the maximal outages for both networks generally can not be reached simultaneously. Instead, [17] proposed the concept of the *achievable* transmission capacity, defined as the product of the spatial density and the *actual* outage probability. The achievable transmission capacity is the metric of our interest in this work, and renamed as *transmission throughput* to avoid possible confusion.

Overlaid networks aim to improve the network spectrum efficiency by exploiting the "white space" in spectrum. Nevertheless, the secondary network should be configured appropriately to achieve its goal; otherwise the overall throughput of the overlaid networks, i.e., the sum throughput of the primary and secondary network, could potentially be hurt. In this work we endeavor to quantitatively characterize the "appropriate" secondary setting by investigating the *overall transmission throughput* of *decentralized* overlaid networks with outage constraints imposed on both primary and secondary networks. Note that the primary outage constraint, which should be satisfied by both networks, only guarantees the minimal acceptable performance for the primary network, and the actual primary throughput depends on its instantaneous outage probability and transmitter densities of both networks. Intuitively the actual primary throughput may decrease due to the additional interference incurred by the secondary network, and this loss may not be compensated by the throughput gain achieved in the secondary network. In [17] the *secondary* throughput was optimized with respect to the secondary transmitter density, while the primary one was ignored. The throughput ratio of the overlaid networks to the stand-alone primary one was derived in [41], with no analysis on whether this ratio is always larger than 1. In contrast, our study yields a sufficient condition for the secondary network configuration such that the overall (transmission) throughput is boosted over that of the stand-alone primary network. Additionally both the maximal allowable secondary density and the optimal secondary density maximizing the overall throughput are derived.

The remainder of this chapter is organized as follows. The system model is given in Section 4.2, followed by discussions on the allowable transmitter density region for the two overlaid networks in Section 4.3. The overall transmission throughput is investigated in Section 4.4. Finally the conclusion and future directions are provided in Section 4.5.

4.2 System Model

We assume the primary and secondary transmitters are distributed in the same two-dimensional plane, and their positions are modeled as two stationary Poisson Point Processes; the former is denoted by $\Pi_o^t = \{X_o(i)\} \subset \mathcal{R}^2$ with density λ_{ot} and the latter by $\Pi_s^t = \{X_s(i)\} \subset \mathcal{R}^2$ with density λ_{st}. Primary (secondary) transmitter i (j) is paired with a receiver located at $Y_o(i)$ ($Y_s(j)$). No cooperation between the primary and secondary network is allowed and the common assumptions about their intra-network communications are given below:

- All the primary (secondary) transmitters use the same transmission power P_o (P_s), which supports a fixed transmission range $l_o(l_s)$. And the power ratio between the primary and secondary network is denoted by $\theta = \frac{P_o}{P_s}$. Concurrent primary and secondary transmissions are simply treated as interference. Thermal noise is assumed negligible in this interference-limited scenario.

- For both networks large-scale path loss and small-scale Rayleigh fading are considered. Particularly the channel power gain for a communication link of length r is given by $g(r) = r^{-\alpha}u$, where $\alpha > 2$ is the path loss exponent, and u is exponentially distributed with *unit* mean. The Signal to Interference Ratio (SIR) at a primary receiver y_o, r_o-distance away from its transmitter x_o, is given by

$$SIR_o(r_o) = \frac{P_o g(r_o)}{I_o + I_{so}}, \tag{4.1}$$

  where $I_o = \sum_{X_o(k)\in\Pi_o^t\setminus\{x_o\}} P_o g(||X_o(k) - y_o||)$ ($I_{so} = \sum_{X_s(k)\in\Pi_s^t} P_s g(||X_s(k) - y_o||)$) is the sum of interference power from concurrent primary (secondary) transmitters, and $||\cdot||$ is Euclidean norm. The SIR at a secondary receiver, SIR_s, is defined similarly as

$$SIR_s(r_s) = P_s g(r_s)/(I_s + I_{os}), \tag{4.2}$$

  where I_s (I_{os}) is the sum of interference power from concurrent secondary (primary) transmitters to a secondary receiver, r_s-distance away from its corresponding transmitter.

- The primary (secondary) transmission is successful if the SIR_o (SIR_s) is no less than a threshold T_o (T_s), assumed fixed in our study. The transmission rate is a deterministic function of this threshold $R_o = f(T_o)$ ($R_s = f(T_s)$).

- There are outage constraints imposed on primary and secondary transmission links, given by

$$\Pr(SIR_o(l_o) < T_o) \leq \epsilon_o, \tag{4.3}$$
$$\Pr(SIR_s(l_s) < T_s) \leq \epsilon_s, \tag{4.4}$$

  respectively, where l_o and l_s are primary and secondary transmission range and $\epsilon_o(< 0.5)$ and $\epsilon_s(< 0.5)$ are predetermined small numbers.

- For convenience of analysis, it is typically assumed that a primary (secondary) receiver is located at the origin, which does not change the statistics of homogenous P.P.P. according to the Slivnyak's theory [42].

Remark 2. *The settings of the primary network such as its density and power are assumed fixed without considering the accommodation of the secondary network. In contrast, the secondary configuration can be tuned to improve the network spectrum efficiency, provided the primary outage constraint is satisfied.*

In this *decentralized* framework, the metric transmission capacity has received increasing interest recently, defined as the maximum density of successful transmissions subject to an out-

age constraint. In our study, we will mainly consider the scenario where the transmitter density of a network is given. To avoid possible confusion, we will call relevant metrics throughputs. In particular, for a Poisson network with transmitter density λ, a typical link length r, a predetermined SINR threshold T and transmission rate R, the transmission throughput is defined as:

$$\hat{C}(\lambda) = R\lambda(1 - \delta(\lambda, r)), \tag{4.5}$$

where the outage probability $\delta(\lambda, r) = \Pr(SINR(r) < T)$. The transmission capacity with outage constraint ϵ [19,20,43] is defined as $R\lambda_\epsilon(1-\epsilon)$, where λ_ϵ is the maximal density satisfying the outage constraint, i.e., $\delta(\lambda_\epsilon, r) = \epsilon$. It can be shown that when ϵ is small, the transmission capacity coincides with $\max_\lambda(\hat{C}(\lambda))$.

Remark 3. *In overlaid networks, transmission capacity may not be investigated directly since the maximal allowable outage of the two overlaid networks may not be reached simultaneously. In contrast, transmission throughput is well defined in either homogenous or heterogenous networks, while keeping the spirit of the transmission capacity.*

4.3 Density Region

Due to the primary and secondary outage constraints, the primary and secondary *transmitter* densities are constrained; the corresponding feasible density region is defined as $\{(\lambda_{ot}, \lambda_{st})\}$. Since the secondary network should remain transparent to the primary network, it can be easily derived from Lemma 3.1 in [44] that

$$\lambda_{ot} \leq \bar{\lambda}_{ot} = \frac{-\ln(1-\epsilon_o)}{K_\alpha T_o^{2/\alpha} l_o^2}, \tag{4.6}$$

where $K_\alpha = \frac{2\pi^2}{\alpha \sin(2\pi/\alpha)}$, regardless of the existence of the secondary network. The secondary network can be accommodated only when $\lambda_{ot} < \lambda_{ot,m}$; otherwise, there is no chance for a secondary network to survive. Given $\Delta\lambda_{ot} = \lambda_{ot,m} - \lambda_{ot} > 0$, it is the secondary network that is responsible for satisfying the primary outage constraint, as well as its own. To satisfy the primary outage constraint, the secondary transmitter density should be bounded above by:

$$\lambda_{st} \leq \bar{\lambda}_{st} = \theta^{2/\alpha} \Delta\lambda_{ot}. \tag{4.7}$$

The derivation is given in Appendix B.1. Note that $\bar{\lambda}_{st}$ is proportional to the power ratio θ and the "white space" $\Delta\lambda_{ot}$.

In order to satisfy the secondary outage constraint (c.f. Appendix B.1), the secondary

transmitter density should be further upper bounded by :

$$\lambda_{st} \leq \hat{\lambda}_{st} = \lambda_{st,m} - \lambda_{ot}\theta^{2/\alpha}, \tag{4.8}$$

where $\lambda_{st,m} = \frac{-\ln(1-\epsilon_s)}{K_\alpha T_s^{2/\alpha} l_s^2}$, represents the maximal allowable secondary transmitter density in a stand-alone setting (c.f. Eq. (6)), and the second term accounts for the additional constraints from the primary network. Intuitively, with increased interference from the primary network, the secondary transmitter density needs to be decreased to maintain the communication quality. From Eq. (4.8) an upper bound for θ can be deduced:

$$\theta < \theta_{\max} \triangleq \left(\frac{\lambda_{st,m}}{\lambda_{ot}}\right)^{\alpha/2}. \tag{4.9}$$

Considering both Eq. (5.7) and (4.8) a feasible secondary transmitter density should satisfy

$$\lambda_{st} \leq \lambda_{st,mm} \triangleq \min(\bar{\lambda}_{st}, \hat{\lambda}_{st}). \tag{4.10}$$

Proposition 2. *$\lambda_{st,mm}$ achieves the maximum $\frac{\lambda_{st,m}}{\lambda_{ot,m}}\Delta\lambda_{st}$ when $\theta = \theta_{opt} \triangleq (\frac{\lambda_{st,m}}{\lambda_{ot,m}})^{\alpha/2}$.*

Proof. Given λ_{ot}, $\bar{\lambda}_{st}$ increases with θ, while $\hat{\lambda}_{st}$ decreases with θ. Therefore $\lambda_{st,mm}(\lambda_{ot})$ is maximized when $\hat{\lambda}_{st}(\lambda_{ot}) = \bar{\lambda}_{st}(\lambda_{ot})$, which leads to $\theta_{opt} = (\frac{\lambda_{st,m}}{\lambda_{ot,m}})^{\alpha/2}$, independent of λ_{ot}. The corresponding maximum value is obtained after θ_{opt} is plugged in. □

An exemplary density region is shown in Fig. 4.1, where $T_o = 3, T_s = 2, r_o = 1, r_s = 0.8, \epsilon_o = \epsilon_s = 0.05, \alpha = 4$, and $\theta = 2$. We can see that the lines of $\bar{\lambda}_{st}$ and $\hat{\lambda}_{st}$ share the same slope $-\theta^{2/\alpha} = -\sqrt{2}$ with respect to λ_{ot}, and the density region is determined by Eq. (5.7) in this case. With the increase of θ, the curve of $\bar{\lambda}_{st}$ rotates around the point of $(\lambda_{ot,m}, 0)$ clockwise, while that of $\hat{\lambda}_{st}$ rotates around the point of $(0, \lambda_{st,m})$ clockwise. When $\theta = (\frac{\lambda_{st,m}}{\lambda_{ot,m}})^{\alpha/2} = 3.66$, two lines coincide (the dashed line in Fig. 4.1), and $\lambda_{st,mm}$ (as well as the density region) is maximized. Note that the maximal secondary transmission density can be interpreted as the secondary transmission capacity when the (constant) transmission rate and outage probability are left out.

In the following discussion, we assume all densities of interest $(\lambda_{ot}, \lambda_{st})$ are within the feasible density region.

4.4 Transmission Throughput of the Overlaid Networks

In this section we optimize the sum transmission throughput of two overlaid networks, with respect to the secondary transmitter density. Our investigation reveals two interesting results:

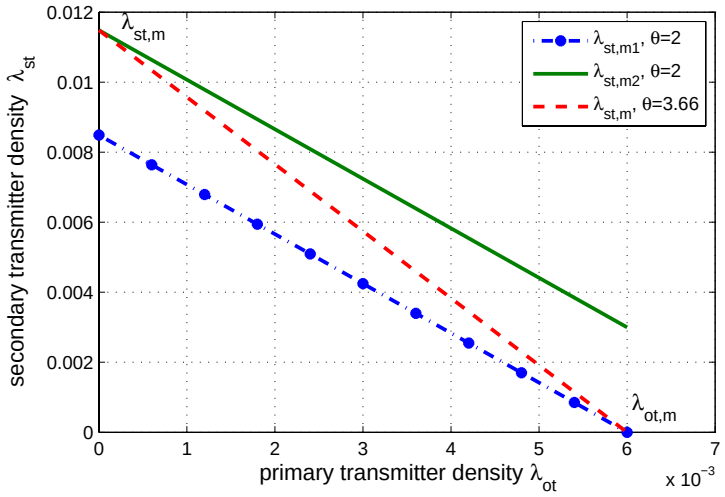

Figure 4.1: An exemplary density region $(\lambda_{ot}, \lambda_{st})$

1) a sufficient condition for the secondary transmitter density such that the overall throughput is larger than the throughput of a stand-alone primary network; 2) the optimal secondary transmitter density λ_{st}^{opt} such that the overall throughput is maximized.

According to Appendix B.1, the outage probabilities for the primary and secondary network are given by:

$$\delta_o(\lambda_{ot}, l_o) = 1 - \exp[-K_\alpha T_o^{2/\alpha} l_o^2 (\lambda_{st} \frac{1}{\theta^{2/\alpha}} + \lambda_{ot})],$$

$$\delta_s(\lambda_{st}, l_s) = 1 - \exp[-K_\alpha T_s^{2/\alpha} l_s^2 (\lambda_{ot} \theta^{2/\alpha} + \lambda_{st})].$$

Thus the overall transmission throughput is given by:

$$\begin{aligned} & \hat{C}_2 = \hat{C}_o(\lambda_{ot}) + \hat{C}_s(\lambda_{st}) \\ = \; & R_o \lambda_{ot}(1 - \delta_o(\lambda_{ot}, l_o) + R_s \lambda_{st}(1 - \delta_s(\lambda_{st}, l_s)) \\ = \; & R_o \lambda_{ot} \exp[-K_\alpha T_o^{2/\alpha} l_o^2 (\lambda_{st} \frac{1}{\theta^{2/\alpha}} + \lambda_{ot})] \\ + \; & R_s \lambda_{st} \exp[-K_\alpha T_s^{2/\alpha} l_s^2 (\lambda_{ot} \theta^{2/\alpha} + \lambda_{st})]. \end{aligned}$$

In the following we consider $\hat{C}_2$ as a function of λ_{st}, $\hat{C}_2(\lambda_{st})$, while fix all the other parameters. Note that $\hat{C}_2(0) > 0$ is the original throughput of the primary network before the secondary network is deployed, and $\hat{C}_2(\lambda_{st})$ $(\lambda_{st} > 0)$ is the sum throughput of the overlaid networks with $\hat{C}_2(\infty) = 0$. Our goal is to find out what conditions will guarantee that there exists a $\lambda_{st} > 0$ such that $\hat{C}_2(\lambda_{st}) > \hat{C}_2(0)$.

Define

$$d = \frac{\theta^{2/\alpha} R_s}{K_\alpha T_o^{2/\alpha} l_o^2 \lambda_{ot} R_o \exp(\theta^{2/\alpha} a \lambda_{ot})}, \tag{4.11}$$

where $a = -K_\alpha T_o^{2/\alpha} l_o^2 \frac{1}{\theta^{2/\alpha}} + K_\alpha T_s^{2/\alpha} l_s^2$. Our main result is summarized as follows.

Theorem 4. *For any feasible primary and secondary transmitter densities, if $d > 1$ the sum transmission throughput is augmented over that of the stand-alone primary network. And the overall throughput is maximized by $\lambda_{st}^{opt} = \min(\lambda_{st,mm}, \lambda_{st}^*)$, where λ_{st}^* is the solution of (4.12), and $\lambda_{st,mm}$ is given in (4.10).*

Proof. Given that $\bar{C}_2(\lambda_{st})$ is a continuous function of λ_{st}, and $\bar{C}_2(0) > 0$ while $\bar{C}_2(\infty) = 0$, if there is a single stable point in the range $\lambda_{st} \in (0, \infty)$, this point must be a local maximum. Taking derivative of $\hat{C}_2(\lambda_{st})$ on λ_{st} and setting it to zero, we get after some calculation:

$$\exp(a\lambda_{st}) = c\lambda_{st} + d, \tag{4.12}$$

where a is given in Eq. (4.11) and $c = \frac{-T_s^{2/\alpha} l_s^2 \theta^{2/\alpha} R_s}{T_o^{2/\alpha} l_o^2 \lambda_{ot} R_o \exp(\theta^{2/\alpha} a \lambda_{ot})} < 0$. It can be shown that Eq. (4.12) has one single positive solution when $d > 1$, as illustrated in Fig. 4.2. In contrast, there is no positive solution when $d \leq 1, c \leq a$, as illustrated in Fig. 4.3. The remaining case, $d \leq 1, a < c < 0$, is more complex, and so is not further discussed (see Fig. 4.4 for an illustration).

In summary, $d > 1$ is a sufficient condition for the improvement of the overall transmission throughput per unit area. Considering the outage probability setting for both primary and secondary network, the total throughput is maximized by setting $\lambda_{st} = \min(\lambda_{st,mm}, \lambda_{st}^*)$. □

Fig. 4.5 shows the overall throughput for different secondary transmitter densities. The common network setting for these two sub-figures is $\epsilon_o = 0.1, \epsilon_s = 0.1, T_o = 3, T_s = 3, l_o = 1, R_o = \log(1 + T_o), R_s = \log(1 + T_s), \lambda_{ot} = 2/3\lambda_{ot,m}$. $\theta = 2, l_s = 0.5$, and $\theta = 0.4, l_s = 2$ for the left and right sub-figure, respectively, corresponds to $d = 23$ and $d = 2.9$, both larger than 1. We can see that, in both cases, the overall throughput increases with the secondary transmitter density (note $\lambda_{st}^{opt} = \lambda_{st,mm}$ for both). Moreover the right sub-figure implies that it is feasible for the secondary network to accommodate long distance transmission ($l_s > l_o$) with higher power ($\theta < 1$) in practice, as long as the primary outage constraint is satisfied. In contrast, Fig. 4.6 demonstrates a 'bad' network setting where $\epsilon_o = 0.4, \epsilon_s = 0.4, T_o = 3, T_s = 3, l_o = 1, l_s = 2, R_o = \log(1 + T_o), R_s = \log(1 + T_s), \lambda_{ot} = 2/3\lambda_{ot,m}$ and $\theta = 0.03$. In this case $d < 1$, and it is shown in the figure that the throughput decreases with the secondary transmitter density. Although the network setting in this example may hardly be adopted in practice, it indicates that an inappropriate configuration could hurt the overall network throughput.

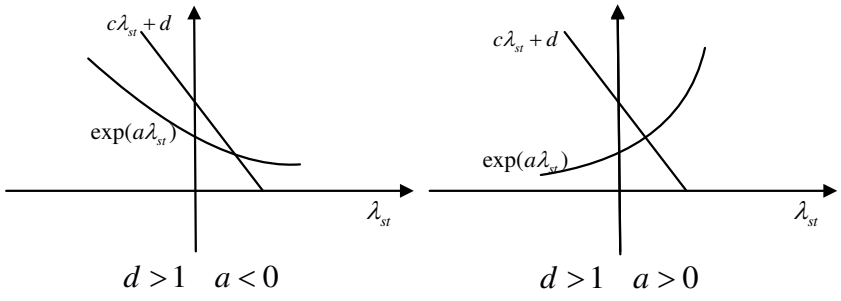

Figure 4.2: When $d > 1$, there is only one solution.

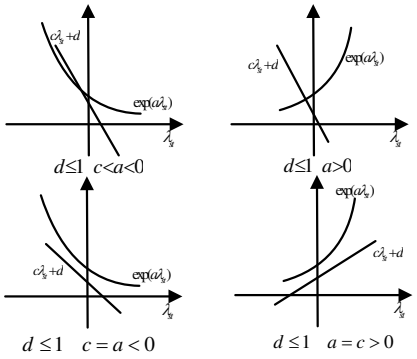

Figure 4.3: when $d \leq 1, c \leq a$, there is no solution.

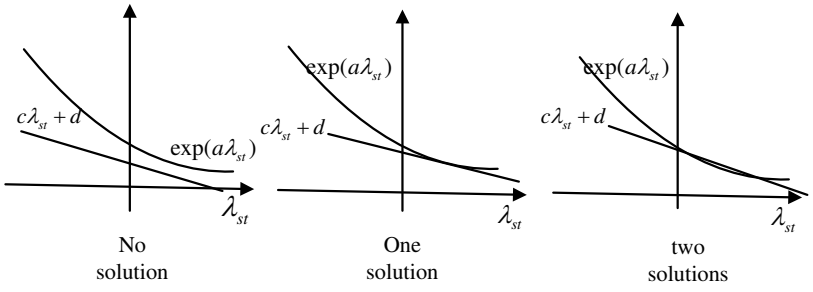

Figure 4.4: When $d \leq 1, a < c < 0$, there is zero, one and two solutions.

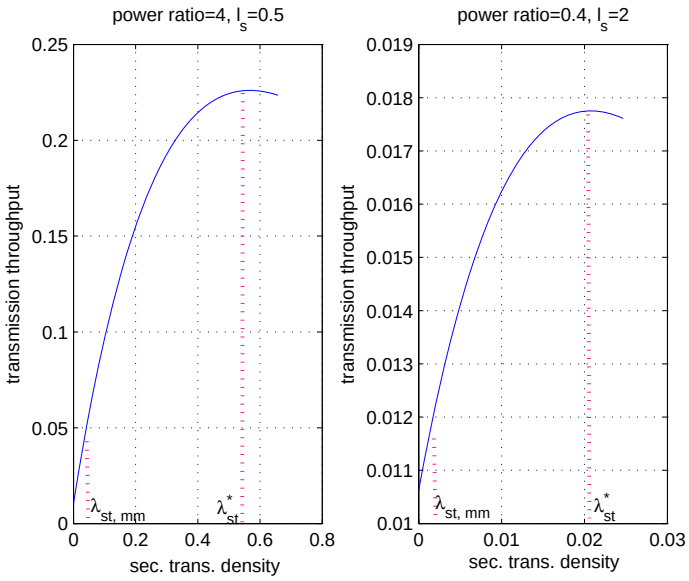

Figure 4.5: Overall throughput: an appropriate secondary setting

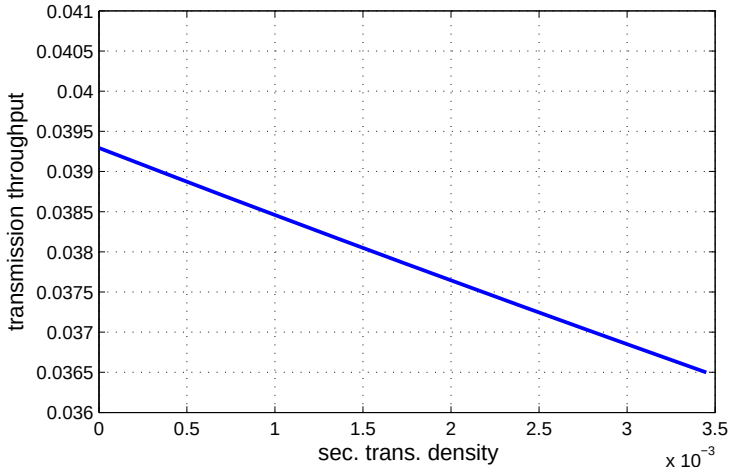

Figure 4.6: Overall throughput: an inappropriate secondary setting

Among secondary parameters, l_s, $T_s(R_s)$ and ϵ_s are usually application-specific. Thus, the power ratio θ becomes a vital parameter to determine the secondary or overall throughput. In the following, we endeavor to specify the power ratio range Θ such that $d > 1$. Set $x = \theta^{2/\alpha}$ and rearrange d as a function of x:

$$d(x) = \frac{R_s}{N\exp(-N)R_o}\frac{x}{\exp(Mx)}, \tag{4.13}$$

where $N = K_\alpha T_o^{2/\alpha} l_o^2 \lambda_{ot}$ and $M = K_\alpha T_s^{2/\alpha} l_s^2 \lambda_{ot}$. By checking the first and second differential of $d(x)$, we get its properties as follows: 1) $d(x)$ is maximized at $x^* = \frac{1}{M}$; 2) $d(x)$ is concave when $x < 2/M$ and convex, otherwise. Therefore if $d(x^*) > 1$ there exist x_1, x_2 ($x_1 < x^* < x_2$) such that $d(x_1) = d(x_2) = 1$ and $d(x) > 1$, $\forall x \in (x_1, x_2)$. Since x is a monotonically increasing function of θ for $\theta > 0$, we can conclude that if $d > 1$ at $\theta^* = \frac{1}{M^{\alpha/2}}$, there exist θ_1, θ_2 ($\theta_1 < \theta^* < \theta_2$) at which $d = 1$, and $d > 1$, $\forall \theta \in (\theta_1, \theta_2)$. Letting $\theta_{\min} = \theta_1$ and noting that $\theta_{\max} < \theta^*$ for $\epsilon_s < 0.5$ (see Eq. (4.9)), we have $\Theta = (\theta_{\min}, \theta_{\max})$ if $\theta_{\min} < \theta_{\max}$; otherwise $\Theta = \phi$.

The lower bound of Θ, $\theta_{\min}$, does not admit a closed-form expression. In the following we derive two analytical upper bounds $\theta'_{\min}$ and $\theta''_{\min}$, for $\theta_{\min}$.

Proposition 3. $\theta_{\min} < \theta'_{\min} < \theta''_{\min}$, *where*

$$\theta'_{\min} = \left(\frac{N\exp(-N)R_o}{(1-\epsilon_s)R_s}\right)^{\alpha/2},$$

and

$$\theta''_{\min} = \left(\frac{-\ln(1-\epsilon_o)(1-\epsilon_o)R_o}{R_s(1-\epsilon_s)}\right)^{\alpha/2}.$$

Proof. It is desired that

$$d > 1 \quad \Leftrightarrow \quad N\exp(-N)\exp(M\theta^{2/\alpha})R_o < R_s\theta^{2/\alpha},$$

where N and M are given after Eq. (4.13). According to Eq. (B.2) the secondary outage probability $\delta(0, l_s) < \delta(\lambda_{st}, l_s) \leq \epsilon_s$ for $\theta \leq \theta_{\max}$, which leads to $\exp(M\theta^{2/\alpha}) < \frac{1}{1-\epsilon_s}$. $\theta'_{\min}$ is thus obtained. Consider the function $y(N) = N\exp(-N)$. It is easy to check that $y(N)$ monotonically increases with N when $N \in (0, 1)$. According to Eq. (5.6) and its following discussion, $N < -\ln(1-\epsilon_o) < 1$ for $\epsilon_o < 0.5$, which results in $y(N) < y(-\ln(1-\epsilon_o))$. Then $\theta'_{\min} < \theta''_{\min}$ is proved. □

In summery, in order to design an optimal secondary network, in terms of overall throughput, the following procedures may be taken:

1. choose l_s, T_s and ϵ_s so that $d > 1$ at $\theta = \theta_{\max}$;

2. choose $\theta \in [\theta''_{\min}, \theta_{\max})$;

3. find the optimal transmitter density according to Theorem 4, and calculate the corresponding maximal throughput;

4. repeat 2) and 3) until the maximal throughput is obtained; an exhaustive search on power ratio may be used if applicable.

4.5 Conclusions

We have made some quantitative study on the transmission throughput of the overlaid networks subject to the outage constraints on both the legacy network and the secondary network. The feasible region of the primary and secondary transmitter densities is explored and the maximal allowable secondary transmitter density is derived. As our main contribution in this chapter, a sufficient condition is obtained for the secondary setting such that the sum throughput of the primary and secondary network is improved over that of the stand-alone primary network.

Chapter 5

Transport Throughput of Secondary Networks in Spectrum Sharing Systems

5.1 Introduction

As discussed in last chapter, asymptotic analysis on the scaling law only characterizes the (rough) relationship between capacity and network size, neglecting the effect of many important system parameters, while transmission capacity admits quantitative system analysis, but leaves out the consideration (and optimization) of transmission distance, a key parameter for wireless networks. To offer a comprehensive view of network throughput of decentralized overlaid networks, we study a new metric in this paper: single hop transport throughput (STT), which quantifies the total number of one-hop *reliable* transmissions in a unit area, weighted by corresponding transmission rates and distances. STT inherits the merits of both traditional transport capacity and transmission capacity, and incorporates transmission distance and outage probability into a uniform framework. STT deserves thorough investigation in the secondary network in that single hop transmissions may be preferred due to its inferior role in the spectrum access; it will also serve as a basis for extension to the multi-hop case. Note that our metric is similar in spirit to the *random access transport capacity* (RTC) recently proposed in [43]. Their difference is that in RTC the transmission distance is pre-determined while in STT it is dynamic across the network. On the one hand, we would like to determine the achievable network throughput when links are activated by specific network protocols (such as nearest neighbor routing). On the other hand, it is desired to explore the limiting performance of the network with transmission distance optimized.

In this work we study STT of the *secondary* network in a decentralized setting, subject

to outage constraints for both the primary and secondary network. We incorporate channel randomness and interference, two essential practical factors, into the performance analysis of the secondary networks, and quantitatively characterize the relationship and tradeoff among important network parameters. Our contributions are summarized below:

- We derive the limit of STT, single hop transport capacity (STC) of the secondary network, and show that it coincides with scaling law result in dense network.
- We further study the STT of the secondary network under the assumption that receivers are *randomly* distributed according to a Poisson Point Process (P.P.P) [45]. In such a setting achievable throughput may vary with the way how secondary transmitters select their corresponding receivers. Three important models are then considered: optimal receivers (OR), random receivers (RR) and nearest neighbors (NN). In the OR model each secondary transmitter chooses an optimal receiver so that the STT is maximized; in the RR model each secondary transmitter randomly chooses its receiver within a transmission radius; and each secondary transmitter chooses the receiver closest to it in the last model. It is shown that they share some similar properties.
- All of our results are in closed-form, which render more insights than scaling law study widely adopted in literature.

The remainder of this chapter is organized as follows. The system model is given in Section 5.2, followed by discussions on the feasible density region and the performance limit of the secondary STT in Section 5.3. Then the STT of the secondary network is studied in Section 5.4. And the simulation results are provided in Section 5.5. Finally the conclusion and future directions are provided in Section 5.6.

5.2 System Model

We follow the system model in the last chapter with some differences. In particular, we assume the primary and secondary transmitters are distributed in the same two-dimensional plane, and their positions are modeled as two stationary Poisson Point Processes; the former is denoted by $\Pi_o^t = \{X_o(i)\} \subset \mathcal{R}^2$ with density λ_{ot} and the latter by $\Pi_s^t = \{X_s(i)\} \subset \mathcal{R}^2$ with density λ_{st}. Primary transmitter $X_o(i)$ is paired with a receiver located at $Y_o(i)$. The locations of potential secondary receivers $Y_s(i)$ are randomly distributed according to a P.P.P Π_s^r with density λ_{sr}. According to the supposition theorem in [45] the density of secondary nodes (including both transmitters and receivers) is $\lambda_{st} + \lambda_{sr}$. No cooperation between the primary and the secondary network is allowed and the common assumptions about their intra-network communications are given below:

- All the primary (secondary) transmitters use the same transmission power P_o (P_s) and their power ratio is denoted by $\theta = \frac{P_o}{P_s}$. Concurrent primary and secondary transmissions are simply treated as interference.

- For both networks large-scale path loss and small-scale Rayleigh fading are considered. Particularly the channel power gain for a communication link of length r is given by

$$g(r) = r^{-\alpha}u,$$

where $\alpha > 2$ is the path loss exponent, and u is exponentially distributed with *unit* mean. The Signal to Interference and Noise Ratio (SINR) at a primary receiver y_o, r_o-distance away from its transmitter x_o, is given by

$$SINR_o(r_o) = \frac{P_o g(r_o)}{N + I_o + I_{so}}, \tag{5.1}$$

where N is the power of the thermal noise, and $I_o = \sum_{X_o(k)\in\Pi_o^t\setminus\{x_o\}} P_o g(||X_o(k) - y_o||)$ ($I_{so} = \sum_{X_s(k)\in\Pi_s^t} P_s g(||X_s(k) - y_o||)$) is the sum of interference power from concurrent primary (secondary) transmitters, and $||\cdot||$ is Euclidean norm. The SINR at a secondary receiver, $SINR_s$, is defined similarly as

$$SINR_s(r_s) = P_s g(r_s)/(N + I_s + I_{os}), \tag{5.2}$$

where I_s (I_{os}) is the sum of interference power from concurrent secondary (primary) transmitters to a secondary receiver, r_s-distance away from its corresponding transmitter.

- The primary (secondary) transmission is successful if the $SINR_o$ ($SINR_s$) is no less than a threshold T_o (T_s), assumed fixed in our study. The transmission rate is a deterministic function of this threshold $R_o = f(T_o)$ ($R_s = f(T_s)$).

- There are outage constraints imposed on primary and secondary transmission links. For the primary network, the constraint is given by:

$$\Pr(SINR_o(l_o) < T_o) \leq \epsilon_o, \tag{5.3}$$

where l_o is a typical primary transmission range and considered constant in our analysis, whose exact value stems from specification of the primary network and is immaterial for our analysis. $\epsilon_o (< 1)$ is a predetermined small number.

For the secondary network, we consider a similar constraint

$$\Pr(SINR_s < T_s) \leq \epsilon_s, \tag{5.4}$$

Under this constraint, we allow the secondary users to flexibly adjust their transmission distances, so long as the communication quality is satisfied for each active secondary link, i.e., $\Pr(SINR_s(r) < T_s) \leq \epsilon_s$ where the link length r is variable, and is a key parameter in our study.

In the study of transmission capacity (c.f. Eq. (4.5)) and its variants, the transmission distance is ignored. In this work, we explore a metric called single-hop transport throughput defined as follows.

Definition 3. *The* single-hop transport throughput *(STT) of a Poisson network with transmitter density λ, a pre-determined SINR threshold T and transmission rate R, is defined as:*

$$C(\lambda) = R\lambda E_r\left[r(1-\delta(\lambda, r))\right] \tag{5.5}$$

where the outage probability $\delta(\lambda, r) = \Pr(SINR(r) < T)$.

In STT the transmission distance is explicitly considered, and also allowed to change (over the space) according to some distribution. The *single-hop transport capacity* (STC)[1] with outage constraint ϵ is defined as the maximum of the single-hop transport throughput with respect to λ and the distribution of r, such that $\delta(\lambda, r) \leq \epsilon$. For a secondary network, λ and r should be chosen such that the outage constraint of the primary network is met as well. The notations and symbols in this work are summarized in Table **??**.

5.3 Single-hop Transport Capacity

The purpose of this paper is to explore the single-hop transport throughput of the secondary network given the outage constraints on both the legacy network and itself. We begin by setting the boundaries for this metric.

Due to the primary and secondary outage constraints, the primary and secondary *transmitter* densities have to be constrained; the corresponding feasible density region is defined as $\Phi = \{(\lambda_{ot}, \lambda_{st})\}$. According to section 4.3,

$$\lambda_{ot} \leq \bar{\lambda}_{ot} = \frac{-\ln(1-\epsilon_o) - n_o l_o^{\alpha}}{K_\alpha T_o^{2/\alpha} l_o^2}, \tag{5.6}$$

where $n_o = \frac{T_o N}{P_o}$, $K_\alpha = \frac{2\pi^2}{\alpha \sin(2\pi/\alpha)}$, regardless of the existence of the secondary network. And

$$\lambda_{st} \leq \bar{\lambda}_{st} = \theta^{2/\alpha} \Delta \lambda_{ot}. \tag{5.7}$$

[1]Some preliminary results on the STC of CR networks are provided in our recent work [46] without considering the secondary outage constraint.

This inequality intuitively reflects the inherent tradeoff of the overlaid networks. If the primary network is sparse, i.e., $\lambda_{ot} \ll \bar{\lambda}_{ot}$, there is ample "white space" left to secondary users. On the contrary, if the primary network is heavily loaded (i.e., λ_{ot} is close to $\bar{\lambda}_{ot}$), a secondary network can hardly "survive" without harming the performance of the primary network. However, an important observation is that the maximum secondary transmitter density is proportional to the power ratio θ, which implies that the secondary network may still operate with low transmission power given limited "white space". In such circumstances, some low power physical layer techniques such as spread spectrum may help improve the performance of secondary networks.

In the following discussion, we assume all densities of interest $(\lambda_{ot}, \lambda_{st})$ are within the feasible density region.

To explore the limiting performance of the secondary network, we assume the flexibility to choose an arbitrary distribution for the transmission distance r. It turns out that the optimal distribution is a degenerate one.

Lemma 5. *Given a secondary transmitter density λ_{st}, assume l and L satisfy the following functions respectively:*

$$\alpha n_s l^\alpha + 2B(\lambda_{st})l^2 - 1 = 0, \tag{5.8}$$

$$B(\lambda_{st})L^2 + n_s L^\alpha + \ln(1-\epsilon_s) = 0, \tag{5.9}$$

where

$$B(\lambda) \triangleq K_\alpha T_s^{2/\alpha}(\lambda_{ot}\theta^{2/\alpha} + \lambda).$$

The secondary single-hop transport throughput $C_s(\lambda_{st})$ achieves the maximum $\bar{C}_s(\lambda_{st})$ by choosing a common transmission distance r_s for each link,

$$r_s = \begin{cases} l & \epsilon_s \geq 1 - e^{-\frac{1}{2}+(\frac{\alpha}{2}-1)n_s l^\alpha} \\ L & otherwise, \end{cases} \tag{5.10}$$

In addition both l and L decreases with λ_{st}

The proof is given in Appendix C.1.

The single hop transport capacity is given by

$$\bar{C}_s = \max_{\lambda_{st}\in\Phi} \bar{C}_s(\lambda_{st}) = \max_{\lambda_{st}\in\Phi} \left[\lambda_{st} r_s \exp\left(-B(\lambda_{st})r_s^2 - n_s r_s^\alpha\right)\right].$$

Typically both l and L do not admit close-form expressions except for some particular α, such as $\alpha = 4$. However directly exploiting Eq. (5.8) and (5.9) we can still explore some properties of $\bar{C}_s$.

Theorem 5. *The single hop transport capacity is*

$$\bar{C}_s = \bar{C}_s(\bar{\lambda}_{st}), \tag{5.11}$$

where $\bar{\lambda}_{st}$ is given in Eq. (5.7).

Proof. It suffices to show that $\bar{C}_s(\lambda_{st})$ is increasing w.r.t. λ_{st}. Note that r_s is also a function of λ_{st}.

When $r_s = l$ the first derivative of $\bar{C}_s(\lambda_{st})$ is given by:

$$\begin{aligned}
&\frac{d\bar{C}_s(\lambda_{st})}{d\lambda_{st}} \exp\left(B(\lambda_{st})r_s^2 + n_s r_s^\alpha\right) = r_s + \lambda_{st}\frac{dr_s}{d\lambda_{st}} \\
&+ \lambda_{st}\left(-2B(\lambda_{st})r_s^2\frac{dr_s}{d\lambda_{st}} - K_\alpha T_s^{2/\alpha} r_s^3 - \alpha n_s r_s^\alpha \frac{dr_s}{d\lambda_{st}}\right) \\
&= r_s + \lambda_{st}\frac{dr_s}{d\lambda_{st}} + \lambda_{st}\left(-\frac{dr_s}{d\lambda_{st}} - K_\alpha T_s^{2/\alpha} r_s^3\right) \\
&= r_s - \lambda_{st} K_\alpha T_s^{2/\alpha} r_s^3,
\end{aligned}$$

where the second equality follows from Eq. (5.8). Also from Eq. (5.8), $r_s \leq \frac{1}{\sqrt{2B(\lambda_{st})}}$. Then,

$$\frac{d\bar{C}_s(\lambda_{st})}{d\lambda_{st}} = \frac{r_s\left(1 - \frac{\lambda_{st}K_\alpha T_s^{2/\alpha}}{2B(\lambda_{st})}\right)}{\exp(B(\lambda_{st})r_s^2 + n_s r_s^\alpha)} > 0.$$

When $r_s = L$ we first take derivative of both sides of Eq.(5.9) w.r.t λ_{st},

$$2B(\lambda_{st})r_s^2\frac{dr_s}{d\lambda_{st}} + K_\alpha T_s^{2/\alpha} r_s^3 + \alpha n_s r_s^\alpha \frac{dr_s}{d\lambda_{st}} = 0 \tag{5.12}$$

$$\Longrightarrow \frac{dr_s}{d\lambda_{st}} = \frac{-K_\alpha T_s^{2/\alpha} r_s^3}{2B(\lambda_{st})r_s^2 + \alpha n_s r_s^\alpha}. \tag{5.13}$$

Then,

$$\begin{aligned}
&\frac{d\bar{C}_s(\lambda_{st})}{d\lambda_{st}} \exp\left(B(\lambda_{st})r_s^2 + n_s r_s^\alpha\right) = r_s + \lambda_{st}\frac{dr_s}{d\lambda_{st}} \\
&= \frac{2B(\lambda_{st})r_s^3 + \alpha n_s r_s^{\alpha+1} - \lambda_{st}K_\alpha T_s^{2/\alpha} r_s^3}{2B(\lambda_{st})r_s^2 + \alpha n_s r_s^\alpha} \\
&= \frac{r_s^3(2B(\lambda_{st}) - K_\alpha \lambda_{st} T_s^{2/\alpha}) + \alpha n_s r_s^{\alpha+1}}{2B(\lambda_{st})r_s^2 + \alpha n_s r_s^\alpha}, \\
&> 0,
\end{aligned}$$

where the first equality is obtained according Eq. (5.12) and the second equality is achieved

after plugging in Eq. (5.13).

Hence the proof is completed. □

Remark 4. *Since $\bar{\lambda}_{st}$ increases with the power ratio θ, one can always improve the capacity $\bar{C}_s$ by increasing the power ratio. This implies that in order to achieve high throughput, secondary transmissions with low power and short transmission range are preferred.*

Corollary 2. $\bar{C}_s = \theta(\sqrt{\bar{\lambda}_{st}})$.

Proof. In the following we focus on $r_s = l$ case. The result for the other case can be shown similarly.

Since the secondary transmitter density $\bar{\lambda}_{st}$ increases with power ratio θ, according to Eq. (5.7), we have

$$\begin{aligned}
&\lim_{\bar{\lambda}_{st}\to\infty} \frac{\bar{C}_s(\bar{\lambda}_{st})}{\sqrt{\bar{\lambda}_{st}}} \\
&= \lim_{\theta\to\infty} \frac{\theta^{2/\alpha}\Delta\lambda_{ot} l \exp(-1/2 + (\alpha/2-1)n_s l^{\alpha})}{\sqrt{\theta^{2/\alpha}\Delta\lambda_{ot}}} \\
&= \frac{\sqrt{\Delta\lambda_{ot}}\exp(-1/2)}{2K_{\alpha}^{2/\alpha}\bar{\lambda}_{ot}}
\end{aligned}$$

where the last equality follows from the fact that

$$l \to \frac{1}{\sqrt{2B(\lambda_{st})}}, \text{ as } \lambda_{st} \to \infty. \tag{5.14}$$

Given the setting of the primary network this ratio is a positive constant. Thus the proof is completed. □

Remark 5. *Our result coincides with the existing scaling law result on capacity in a dense network model where the network size is allowed to grow with the network density in a fixed area. It is shown in [18] that, in a single n-node arbitrary dense network, the sum throughput scales with*[2] *$\Theta(\sqrt{n})$. Thus there is no performance loss in transport capacity of the secondary network, in terms of scaling law. Note that the critical constant term is usually unavailable in the scaling law study, while it is clearly presented in our expression. In addition, with larger spectrum opportunity $\Delta\lambda_{st}$, the secondary throughput increases faster.*

[2]Note that the upper bound in [18] is also obtained in a one-hop setting.

5.4 Transport Throughput of Secondary Network

In this section we investigate the single hop transport throughput of the secondary network defined in Section 7.2, with the secondary nodes modeled as a P.P.P. with density $\lambda_{st} + \lambda_{sr}$. The two overlaid networks operate under the primary and secondary outage constraints, i.e., their transmitter densities are within the density region. For the ease of calculation, we ignore the thermal noise in this section, which is reasonable in this interference-limited scenario when α is small. The optimal range in Lemma 5 is then given by

$$r_s = \begin{cases} l = 1/\sqrt{2B(\lambda_{st})} & \epsilon_s \geq 1 - e^{-1/2} \\ L = \sqrt{\frac{\ln(1-\epsilon_s)}{-B(\lambda_{st})}} & \text{otherwise,} \end{cases} \tag{5.15}$$

And the corresponding maximum single-hop transport throughput $\bar{C}_s(\lambda_{st})$ is given by

$$\bar{C}_s(\lambda_{st}) = \begin{cases} \frac{R_s \lambda_{st}}{\sqrt{2B(\lambda_{st})}} e^{-1/2}, & \epsilon_s \geq 1 - e^{-1/2} \\ \frac{R_s \lambda_{st}(1-\epsilon_s)}{\sqrt{B(\lambda_{st})}} \sqrt{-\ln(1-\epsilon_s)}, & \text{otherwise.} \end{cases} \tag{5.16}$$

Remark 6. *In practice, it is expected that ϵ_s is smaller than $1 - e^{-1/2} \approx 0.39$. Therefore in the rest of this paper, we focus on this scenario. It is also known from Appendix C.1 that $C_s(r_s, \lambda_{st})$ monotonically increases with $r_s \in (0, L)$ when $\epsilon_s < 1 - e^{-1/2}$, and L is the maximum allowable transmission distance satisfying the outage constraint ϵ_s.*

Communication links can be formed in different manners, depending on how transmitters select their corresponding receivers, which plays an essential role in determining the STT. To present a comprehensive view of throughput we are interested in the following settings:

1. Optimal receivers (OR): the secondary receivers are selected in order to achieve the best performance. Interestingly in the setting of our interest, the optimal receivers selected are the furthest ones from the transmitters.

2. Random receivers (RR): the secondary receivers are randomly selected.

3. Nearest neighbors (NN): namely, each secondary transmitter communicates with its nearest neighbor, which is a conservative but easy-to-implement approach.

The maximal secondary throughput in Eq. (5.16) serves as a benchmark for our study. We'll first derive the throughput in these settings and then discuss these results in the last subsection,

5.4.1 Optimal receivers

The optimal receiver case demonstrates the maximal achievable throughput when secondary nodes are randomly distributed (in contrast to the arbitrary distribution assumed in the STC study). According to remark 6, given the secondary transmitter density λ_{st}, the throughput monotonically increases over transmission distance[3] $r_s \in (0, L]$. Therefore, the throughput is maximized if each secondary transmitter communicates with the furthest receiver within the range of L. In particular, for a secondary transmitter X_s, we order the potential secondary receivers $\{Y_s(i)\}$ according to their Euclidean distance $r_i \triangleq ||Y_s(i) - X_s||$ to X_s such that $r_i \leq r_j$, $\forall i < j$. The receiver chosen by transmitter X_s is given by

$$k^* = \begin{cases} \max\{i : L - r_i \geq 0\} & r_1 \leq L, \\ NULL & otherwise. \end{cases} \tag{5.17}$$

Theorem 6. *Given a secondary transmitter density λ_{st} and a secondary receiver density λ_{sr}, the largest achievable single-hop transport throughput is given by:*

$$C_s^o(\lambda_{st}, \lambda_{sr}) = \begin{cases} A_1 \left[\frac{Le^{aL^2}}{2a} - \frac{\sqrt{\pi} erfi(\sqrt{a}L)}{4a^{3/2}} \right] & a > 0 \\ A_1 L^3/3 & a = 0 \\ A_1 \left[\frac{Le^{aL^2}}{2a} + \frac{\sqrt{\pi} erf(\sqrt{-a}L)}{4(-a)^{3/2}} \right] & a < 0 \end{cases}$$

where $A_1 = 2\lambda_{st} R_s \lambda_{sr} \pi e^{-\lambda_{sr}\pi L^2}$, $a = \lambda_{sr}\pi - B(\lambda_{st})$, $B(\lambda_{st})$ is given in Lemma 1, $erf(x)$ and $erfi(x)$ are the error function and the imaginary error function, defined as $2/\sqrt{\pi}\int_0^x e^{-t^2}dt$, and $2/\sqrt{\pi}\int_0^x e^{t^2}dt$, respectively.

Proof. Given that there is at least one receiver within the circle with radius L, the event $\{r_{k^*} > l\}$, where $l \leq L$, is equivalent to the event {there is at least one receiver in the annulus $C(l, L)$}. Therefore,

$$\Pr(r_{k^*} > l | r_1 \leq L) = \frac{1 - e^{-\lambda_{sr}\pi(L^2 - l^2)}}{1 - e^{-\lambda_{sr}\pi L^2}},$$

which leads to the conditional pdf of r_k,

$$f_{r_{k^*}|r_1 \leq L}(l) = \frac{e^{-\lambda_{sr}\pi L^2} 2\lambda_{sr}\pi l}{1 - e^{-\lambda_{sr}\pi L^2}} e^{\lambda_{sr}\pi l^2}.$$

[3]For ease of notation, we drop the argument λ_{st} in this expression and the following.

It follows that

$$\begin{aligned} C_s^o &= \lambda_{st} R_s E_{r_{k^*}}(r_{k^*}(1-\delta_s(r_{k^*}))) \\ &= \lambda_{st} R_s \Pr(r_1 \le L) E_{r_{k^*}|r_1 \le L}(r_{k^*}(1-\delta_s(r_{k^*}))) \\ &= \lambda_{st} R_s (1-e^{-\lambda_{sr}\pi L^2}) \int_0^L l e^{-B(\lambda_{st})l^2} f_{r_{k^*}|r_1 \le L}(l) dl, \end{aligned} \tag{5.18}$$

where the outage δ_s is given in Eq. (B.2). Eq. (5.18) is then derived after some calculation. □

5.4.2 Random receivers

The RR model demonstrates the STT in the average sense, where each secondary transmitter randomly chooses a receiver in a range at most L distance away.

Theorem 7. *Given a secondary transmitter density λ_{st} and a secondary receiver density λ_{sr}, the single hop transport throughput in the RR model is*

$$C_s^r(\lambda_{st}, \lambda_{sr}) = A_2 \frac{R_s \lambda_{st}}{\sqrt{B(\lambda_{st})}} \left(1 - (1-\epsilon_s)^{\lambda_{sr}\pi / B(\lambda_{st})}\right) \tag{5.19}$$

where $A_2 = \frac{\sqrt{\pi} erf\left(\sqrt{\ln \frac{1}{1-\epsilon_s}}\right) - 2(1-\epsilon_s)\sqrt{\ln \frac{1}{1-\epsilon_s}}}{-2\ln(1-\epsilon_s)}$ *and $B(\lambda_{st})$ is given in Lemma 1.*

Proof. Denote by E_k the event that there are k receivers in the circle centered at a secondary transmitter with radius L. Given the number of receivers in the circle, all the receivers are uniformly distributed with probability density function (pdf) $f(r) = \frac{2r}{L^2}$ according to the Poisson property. Then we have

$$\begin{aligned} & E\left[r(1-\delta_s(r))\right] = E[E(r(1-\delta_s(r))|E_k)] \\ &= \sum_{k=1}^{\infty} \frac{(\lambda_{sr}\pi L^2)^k e^{-\lambda_{sr}\pi L^2}}{k!} \int_0^L r \exp(-B(\lambda_{st})r^2) \frac{2r}{L^2} dr \\ &= \left(1 - \frac{1}{e^{\lambda_{sr}\pi L^2}}\right) \left[\frac{\sqrt{\pi} erf\left(L\sqrt{B(\lambda_{st})}\right)}{2L^2 B(\lambda_{st})^{3/2}} - \frac{e^{-B(\lambda_{st})L^2}}{B(\lambda_{st})L}\right]. \end{aligned}$$

The theorem follows after substituting L in Eq. (5.15) to the equation above.

□

5.4.3 Nearest neighbors

Nearest neighbors model is one of the popular models in analysis of wireless networks, especially in the study of network connectivity ([47] and reference there). In this subsection we

consider secondary transmitters communicate with their nearest receivers satisfying the outage constraints. For the ordered receivers introduced in Sec. 5.4.1, the receiver chosen by X_s in this model is $Y_s(1)$ if $r_1 \leq L$; otherwise no receiver is chosen since all the receivers are outside of the allowable transmission range.

Theorem 8. *Given a secondary transmitter density λ_{st} and a secondary receiver density λ_{sr}, the STT in the NN model is given by:*

$$C_s^n(\lambda_{st}, \lambda_{sr}) = A_1 \left(-\frac{Le^{-B(\lambda_{st})L^2}}{2b} + \frac{\sqrt{\pi} erf(\sqrt{b}L)}{4e^{-\lambda_{sr}\pi L^2} b^{3/2}} \right), \tag{5.20}$$

where $b = B(\lambda_{st}) + \lambda_{sr}\pi$, $B(\lambda_{st})$ is given in Lemma 1, and A_1 is given in Theorem 2.

Proof. The event $\{r_1 > l\}$ is equivalent to the event {there is no receiver in the circle with radius l centered at X_s}. Thus,

$$\Pr(r_1 > l) = e^{-\lambda_{sr}\pi l^2},$$

which leads to the conditional pdf of r_1,

$$f_{r_1}(l) = 2\lambda_{sr}\pi l e^{-\lambda_{sr}\pi l^2}.$$

Since we only consider the receivers located within L

$$\begin{aligned} C_s^n &= \lambda_{st} R_s E_{r_1 | r_1 \leq L}(r_1(1 - \delta_s(r_1))) \\ &= \lambda_{st} R_s \int_0^L l e^{-B(\lambda_{st})l^2} f_{r_1}(l) dl. \end{aligned} \tag{5.21}$$

The theorem then follows after some calculation. □

5.4.4 Discussions and insights

Before we explore the properties of secondary throughput, we introduce the average node degree defined as the average number of receivers which obtain a packet correctly from a broadcast caused by an arbitrary transmitter. Suppose an arbitrary secondary transmitter at the origin, then we have

Proposition 4. *The average node degree $k = \frac{\lambda_{sr}\pi}{B(\lambda_{st})}$.*

Proof. Consider a circle $c(o, a)$ of radius a centered at the origin and denote by $\mathcal{N}_a$ the number

of receivers in $c(o,a)$ connected to the origin, then we have

$$E(\mathcal{N}_a) = E\left[\sum_{i=1}^{N_a} I_i\right] = E\left[E\left(\sum_{i=1}^{k} I_i | N_a = k\right)\right] \tag{5.22}$$
$$= \sum_{k=1}^{\infty} \frac{(\lambda_{sr}\pi a^2)^k e^{-\lambda_{sr}\pi a^2}}{k!} E\left(\sum_{i=1}^{k} I_i | N_a = k\right)$$

where $N_a = \Pi_s^r(c(o,a))$ is the number of receivers in $c(o,a)$ and I_i is the indicator function such that $I_i = 1$ if receiver i is connected to the origin.

The probability density function that a secondary receiver is at a distance $r \leq a$ from the origin is $f_r(r) = \frac{2r}{a^2}$. Therefore,

$$E\left(\sum_{i=1}^{k} I_i | N_a = k\right) = kE(I_1) = k\int_0^a (1-\delta_s) f_r(r) dr. \tag{5.23}$$

where δ_s is given in Eq. (B.2).

Taking Eq. (5.23) into Eq. (5.22) and noticing the fact that $\sum_{k=1}^{\infty} \frac{(\lambda_{sr}\pi a^2)^{k-1} e^{-\lambda_{sr}\pi a^2}}{(k-1)!} = 1$, we get

$$E(\mathcal{N}_a) = \lambda_{sr}\pi a^2 \int_0^a (1-\delta_s) f_r(r) dr \tag{5.24}$$
$$= \lambda_{sr}\pi \frac{1-\exp(-B(\lambda_{st})a^2)}{B(\lambda_{st})}.$$

Thus,

$$k = \lim_{a\to\infty} E(\mathcal{N}_a) = \frac{\lambda_{sr}\pi}{B(\lambda_{st})}.$$

□

Define $\beta = \frac{\lambda_{sr}}{\lambda_{st}}$, the density ratio between the secondary receivers and transmitters. Given β

$$k = \frac{\beta\pi}{K_\alpha T_s^{2/\alpha}(\frac{\lambda_{ot}\theta^{2/\alpha}}{\lambda_{st}} + 1)} \tag{5.25}$$
$$\leq \frac{\beta\pi}{K_\alpha T_s^{2/\alpha}(\frac{\lambda_{ot}\theta^{2/\alpha}}{\bar{\lambda}_{st}} + 1)} = \frac{\pi\beta\Delta\lambda_{ot}}{K_\alpha T_s^{2/\alpha}\bar{\lambda}_{ot}},$$

where the inequality is due to the fact that $\lambda_{st} \leq \bar{\lambda}_{st}$. This indicates that in a p-Persistent Aloha system with probability $p = \frac{1}{1+\beta}$ such that a secondary node becomes a transmitter, the average node degree is upper bounded by a constant.

Substituting the average node degree $k = \frac{\lambda_{sr}\pi}{B(\lambda_{st})}$ into Eq. (5.18) ($a > 0$ case), (5.19) and (5.20), we have

$$C_s^o(\lambda_{st}, k) = Q_o(k)\frac{R_s\lambda_{st}}{\sqrt{B(\lambda_{st})}}(k > 1), \tag{5.26}$$

$$C_s^r(\lambda_{st}, k) = Q_r(k)\frac{R_s\lambda_{st}}{\sqrt{B(\lambda_{st})}}, \tag{5.27}$$

$$C_s^n(\lambda_{st}, k) = Q_n(k)\frac{R_s\lambda_{st}}{\sqrt{B(\lambda_{st})}}, \tag{5.28}$$

where
$Q_o(k) = \frac{k\sqrt{-\ln(1-\epsilon_s)}(1-\epsilon_s)}{k-1} - \frac{k\sqrt{\pi}erfi(\sqrt{(1-k)\ln(1-\epsilon_s)})}{2(k-1)^{3/2}(1-\epsilon_s)^{-k}}$,
$Q_r(k) = A_2(1-(1-\epsilon)^k)$, and
$Q_n(k) = \frac{k\sqrt{-\ln(1-\epsilon_s)}(1-\epsilon_s)^{1+k}}{-k-1} + \frac{k\sqrt{\pi}erf(\sqrt{-(1+k)\ln(1-\epsilon_s)})}{2(k+1)^{3/2}}$.

All the above results, as well as the maximal throughput derived in Eq (5.16), take the same form except for the coefficient $Q_i, i = o, r, n$. This similarity provides a common ground for our following discussion.

Corollary 3. *1) Given secondary transmitter density λ_{st}, C_s^o and C_s^r in the OR and RR model increase with the secondary receiver density λ_{sr}, or equivalently k, and*

$$\begin{aligned}
\lim_{k\to\infty} C_s^o(\lambda_{st}, k) &= C_{s,m}(\lambda_{st}) \\
&= (1-\epsilon_s)\sqrt{-\ln(1-\epsilon_s)}\frac{R_s\lambda_{st}}{\sqrt{B(\lambda_{st})}}, \\
\lim_{k\to\infty} C_s^r(\lambda_{st}, k) &= A_2\frac{R_s\lambda_{st}}{\sqrt{B(\lambda_{st})}},
\end{aligned}$$

where $C_{s,m}(\lambda_{st})$ is given in Eq. (5.16) and A_2 is given in Theorem 7.

2) Given λ_{st}, C_s^n in the NN model is maximized at $k \approx \frac{2}{\epsilon_s}$ for small ϵ_s.

3) Given β, C_s^o and C_s^r increases with λ_{st} and achieve the maximum at $\bar{\lambda}_{st,m}$

Proof. Property 1): It is straightforward to see that $Q_o(k)$ and $Q_r(k)$ increases with k and the convergence is achieved after some calculation.

Property 2): The existence of an optimal k such that C_s^n is maximized is due to the fact that $\lim_{k\to 0} C_s^n(\lambda_{st}, k) = \lim_{k\to\infty} C_s^n(\lambda_{st}, k) = 0$. To get the optimal k we solve the equation

$\frac{dQ_n(k)}{dk} = 0$.

$$\begin{aligned}\frac{dQ_n(k)}{dk} &\approx \sqrt{-\ln(1-\epsilon_s)}\frac{(1-(1-\epsilon_s)^{1+k})k}{1+k} \\ &\approx \sqrt{-\ln(1-\epsilon_s)}\frac{\epsilon_s^2}{2}(-k^2+\frac{2}{\epsilon_s}k),\end{aligned}$$

where the approximation is made due to the fact that

$$erf(\sqrt{-(1+k)\ln(1-\epsilon_s)}) \approx 2/\sqrt{\pi}\sqrt{-(1+k)\ln(1-\epsilon_s)}$$

and

$$(1-\epsilon_s)^{1+k} \approx 1-\epsilon_s(1+k)+\epsilon_s^2 k(1+k)/2$$

for small ϵ_s. The optimal $k \approx \frac{2}{\epsilon_s}$.

Property 3): Since $Q_o(k)$ and $Q_r(k)$ increase with k, which increases with λ_{st} given β (c.f. Eq. (5.25)), they also increase with λ_{st}, and achieves the maximum at $\lambda_{st} = \bar{\lambda}_{st}$. □

Intuitively, the property 1) and 2) indicate that the capacity discussed in the last section is achieved when the density of receivers in the OR model goes to infinity (so that a secondary receiver can be found at the optimal distance), and the RR and NN model suffer performance loss by a constant factor. In addition, the NN model achieves the optimal value at a finite k value, whereas the other two achieve the optimum when k goes to infinity. Property 3) reveals that in a p-Persistent Aloha system, the secondary throughput in the OR and RR model increases with the secondary density λ_{st}, as well as the total secondary density $\lambda_{st} + \lambda_{sr}$ (since β is fixed). Note that if the receiver density is fixed, the throughput is not necessarily increases with the transmitter density, which can be seen from the expressions in Eq. (5.18) and (5.19).

5.5 Numerical and Simulation Results

In this section the secondary transport throughput in the three models are compared by simulations. We randomly and uniformly distributed primary and secondary nodes in a square $[0, 1000]^2$. And their numbers are generated according to the Poisson distribution with their respective densities. The common settings for all the figures are $T_o = 3, T_s = 2, l_o = 6, \alpha = 4, \epsilon_o = 0.5$, and $\lambda_{ot} = 1/5\bar{\lambda}_{ot}$ for all the figures except the last one. We can see from all the figures the numerical results match perfectly with the simulation results.

The throughputs in the three models are depicted in Fig. 5.1, 5.2, 5.3, and 5.4, versus different parameters. We can see in Fig. 5.1 and 5.2 (where $\lambda_{st} = \bar{\lambda}_{st}$) that in the low

receiver density region, three models perform almost the same, but with the receiver density increased, the NN model performs the worst among the three models. The throughputs in both OR and RR model increase with λ_{sr}; while at $\lambda_{sr} \approx \frac{2B(\lambda_{st})}{\epsilon_s \pi}$ the throughput in the NN model is maximized. These observations coincide with our discussion above and admit an intuitive interpretation: with a low receiver density the receivers chosen by a transmitter in these models often coincide so that their performance is indiscernible; with more receivers the transmitters have more chance to select the receivers at the optimal positions in the OR and RR model; however, the distance between a transmitter and its closest receiver becomes smaller so that the throughput is decreased in the NN model. Another interesting observation is that with low outage probability imposed on the secondary network (Fig. 5.1), the throughput in the OR model approaches the upper bound rather slowly with the increase of the secondary receiver density; its converging speed is substantially faster with a high outage probability (Fig. 5.2). The reason behind this phenomenon is that with a high outage probability the optimal transmission distance is much larger (c.f. (5.15)), which can be easily satisfied even when the receiver density is low.

Fig. 5.3 describes the throughputs for different transmitter densities given the density ratio $\beta = 100$. The throughputs increases with the transmitter density and scales as $\sqrt{\lambda_{st}}$, which substantiates our analysis in section 5.4.4.

Fig. 5.4 shows the influence of the primary transmitter density on these three throughputs, where $\lambda_{st} = \bar{\lambda}_{st}$, given λ_{ot}. Note that the maximal secondary transmitter density is *inversely* proportional to the primary transmitter density (c.f. Eq. (5.7)). Therefore, the secondary throughput decreases with λ_{ot}, which is clearly presented in Fig. 5.4.

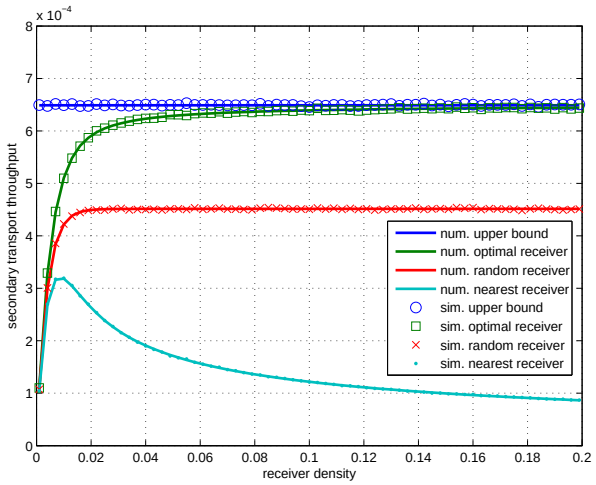

Figure 5.1: Secondary STT VS secondary receiver density $\epsilon_s = 0.1$

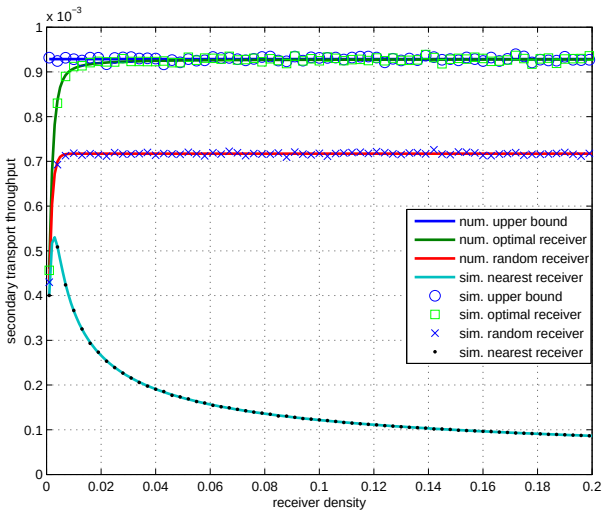

Figure 5.2: Secondary STT VS secondary receiver density $\epsilon_s = 0.3$

5.6 Conclusions

We have made some quantitative study on the throughput of the secondary network in spectrum sharing systems subject to the outage constraints for both the legacy network and the secondary network, aiming at revealing the relationship and tradeoff among key system parameters and providing insights into system design and optimization. A new metric, single hop transport throughput, was proposed, which incorporates transmission distance and outage constraint into a uniform framework. We showed that the secondary single hop transport capacity, the maximal secondary transport throughput given the settings of the primary network, coincides with existing scaling law result. This indicates that there is indeed no performance loss for the secondary network, in terms of scaling law. In addition, our analysis reveals that the optimal secondary STT is given when farthest receivers are chosen by their corresponding transmitters.

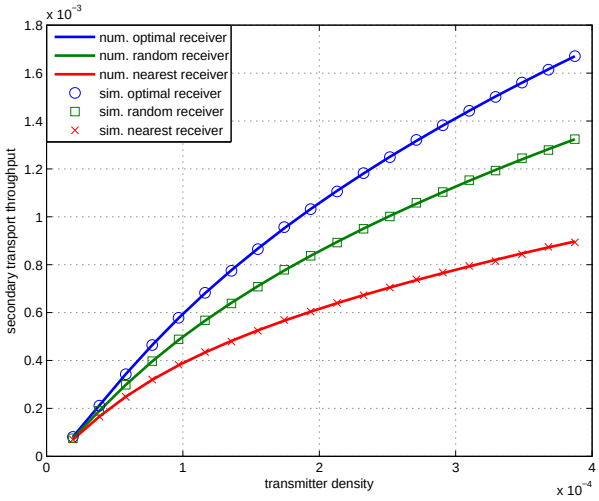

Figure 5.3: Secondary STT VS secondary transmitter density: receiver density $\lambda_{sr} = 100\bar{\lambda}_{st,om}$

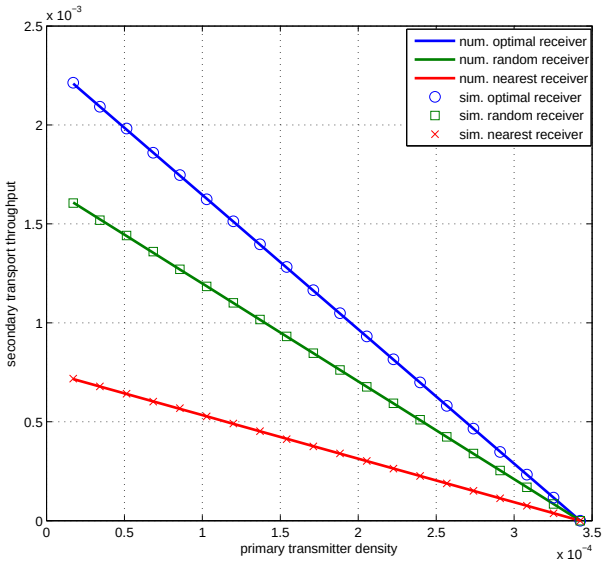

Figure 5.4: Secondary STT versus primary transmitter density

Chapter 6

Communication Efficiency of Anti-jamming Broadcast in Large-scale Multi-Channel Wireless Networks

6.1 Introduction

6.1.1 Motivation

Wireless networks are highly vulnerable to jamming attacks due to the openness of the communication medium. Lack of resistance to jamming, wireless signals can be easily blocked, modified or replaced, which may jeopardize personal safety and national security. As a result, jamming-resistant broadcast is crucial to security-critical applications; examples include emergency alert broadcast and navigation signal dissemination.

A prominent scheme against jamming threat at the physical layer is the employment of Spread-Spectrum (SS) techniques, including Direct-Sequence Spread Spectrum (DSSS) and Frequency Hopping (FH). By deliberately spreading the signal's energy over a large bandwidth, such techniques prevent the legitimate transmission from being detected or intercepted, and exhibit resistance to narrow-band interference or jamming. However, these classic countermeasures rely on pre-shared secret keys, such as spreading code sequences and frequency hopping patterns, to achieve correct decoding at the receivers. Therefore, before the establishment or after the compromise of secrete keys, these approaches are inefficacious.

The dependency on secret keys and the associated security issue are recognized recently, and one promising research direction, anti-jamming communication without pre-shared keys,

is initiated [22, 48, 49]. In this direction Uncoordinated SS (USS) is of particular interest, including Uncoordinated DSSS (UDSSS) [21] and Uncoordinated FH (UFH) [22, 23]. USS randomizes the key selection so that neither attackers nor malicious nodes are able to jam the communication. In particular transmitters in UDSSS randomly select a spreading code from a set of code sequences, and receivers attempt to decode the packets through a brute-force search, while in UFH a transmitter (receiver) randomly chooses a channel from a pool to transmit (listen), and successful transmission is made when the transmitter and receiver agree on the same channel. Due to their inherent difference in handling interference, UDSSS and UFH behave differently in the presence of jamming [24]: generally speaking, the performance of UDSSS is mainly determined by the computing power of jammers, while that of UFH is mainly decided by the hardware capacity (such as sensing and switching frequency) of jammers. As it is anticipated that the improvement on the computing capability will continue its current trend in the foreseeable future, UFH assumes certain advantage in this aspect. In this work, we will focus on the UFH framework for the concreteness of discussion.

One common issue with USS, including UFH, is low communication efficiency; this is the price paid for jamming resistance, due to lack of coordination between the sender and receiver. The efficiency of UFH was enhanced in [23] by incorporating erasure coding and one-way authentication based on bilinear maps. However the improved UFH only reduces the communication latency up to one half, still far less efficient than the conventional (coordinated) FH. In our recent work [24], Collaborative UFH (CUFH) was proposed to achieve significantly higher throughput and enhanced jamming resistance, where nodes having obtained the message serve as relays for the remaining nodes to expedite the broadcast process.

In existing literature on anti-jamming broadcast, relevant analysis on the communication efficiency is relatively underdeveloped, particularly for a large-scale network and when there are cooperative relays. In this work we analytically evaluate the communication efficiency of UFH and CUFH in large-scale networks, where the number of nodes n is large and may even exceed the number of channels C. Such a large-system analysis is pertinent to emerging and future networked systems, and also helps us reveal the scaling behavior of the system performance with respect to the network size and other important system parameters. Furthermore, such a study also brings about new research problems that were not considered in previous studies. In particular, when $n > C$, the number of relays actually needs to be carefully controlled for the best performance. The metric we are interested in is the broadcast delay for an arbitrary receiver and the whole network. Our contributions are summarized as follows:

- For UFH, the average node broadcast delay is derived in a succinct expression, and a tight upper bound of the average network broadcast delay is obtained. The former shows that the average node broadcast delay scales as $\theta(C \ln M)$, where M is the number of packets a message is divided into. The latter reveals that the network broadcast delay scales as

$O(C \ln n)$, i.e., *logarithmically* with the network size n, which demonstrates the feasibility of UFH in large-scale networks.

- In large-scale networks the performance of CUFH depends on the number of relays. The optimal number of relays is derived in terms of maximizing the packet reception rate and the optimal cooperation gain achieved by CUFH over UFH is quantified. In addition we propose a suboptimal protocol CUFH-p, which simplifies the procedure of constraining the relays in practice, and facilitates the derivation of an upper bound of the average network broadcast delay that CUFH incurs. It is shown that the network broadcast delay of CUFH scales as $O(\ln n)$ and outperforms UFH by a factor of C when $C = o(\ln n)$. As an intermediate step, we have studied an ideal cooperative network, whose results may be of independent interest.

- Our analysis is well supported by simulations, and the insights obtained from our analysis allow us to obtain a very good estimate of the average network broadcast delay of CUFH.

Note that our study assumes wider applicability: our analytical methodology applies largely to general anti-jamming broadcast, and many of our results will find applications for broadcast in multi-channel networks. For example in multi-channel cognitive radio networks, it's nontrivial to establish coordination between secondary transmitters and receivers since channel states (available or not) are dynamical. With the help of UFH they can build the initial connection.

6.1.2 Related works

As we mentioned above conventional jamming defense such as spread spectrum techniques requires pre-determined keys to build up jamming resistance, which suffers from poor scalability and severe threat once the keys are exposed. In order to remove such dependence, a common approach adopted by most recent works on this topic is to introduce randomness on the key selection. Besides Uncoordinated SS (USS), Randomized Differential DSSS (RD-DSSS) [50] also adopts this strategy, where each bit is encoded using the correlation of unpredictable spreading codes. Formulating the frequency hopping problem as a multi-armed bandit problem, an online adaptive UFH scheme was proposed in [51], where the receiver performs online learning and optimization in response to a potentially adaptive (responsive) jammer. Randomness enhances the jamming immunity, yet typically degrades the communication efficiency due to lack of coordination between the transmitters and receivers. CUFH is one effective way to improve the efficiency. Researchers have also considered the combination of uncoordinated SS with conventional SS for efficiency improvement, such as delayed-key UDSSS [21] and Uncoordinated Seed Disclosure-FH (USD-FH) [52]. In the former the transmitter sends a message spread by a random code sequence K and subsequently transmits K, both using UDSSS. Once the receiver

decodes K, the message can be easily extracted as in conventional DSSS. In the latter, a predetermined hopping patten is first broadcast through UFH by the transmitter; once it is successfully decoded by a receiver a conventional FH link is built.

Besides defenses at the physical layer, countermeasures implemented at the MAC layer have also been studied. Two such countermeasures, channel surfing and data exfiltration, were discussed in [53] and [54], respectively. A jamming-resistant MAC protocol was proposed in [55] for a single-hop network and is extended to a multi-hop network in [56] recently. A jamming detection scheme based on the Sequential Probability Ratio Test (SPRT) was proposed in [57]. The impact of placement and attacking ranges of jammers was experimentally studied in [58].

Another line of works [59–61] is worth mentioning, which addresses the same topic, jamming-resilient communications in single-hop multi-channel wireless networks without pre-shared secrets, from a computation-theoretic perspective. Relaying is considered in these works to facilitate message distribution. In [60] the authenticated pairwise message exchange problem is investigated, while in [59,61] the gossip problem (where all nodes like to get initial values of all others) is treated. The computation-theoretic protocols of these works are mainly evaluated by their running time (the number of synchronous runs) in the order sense.

Other related work includes anti-jamming study in 802.11-based networks [62,63] and Cognitive Radio networks [64,65].

The remainder of the chapter is organized as follows. The system model is introduced in Section 7.2. In Section 6.3 the performance of UFH is evaluated in terms of average node and network broadcast delay. In Section 6.4 the performance of CUFH with relay constraint is investigated. Our analysis is substantiated by simulation results provided in Section 6.5. Related works are discussed in Section 6.1.2. Finally we conclude this paper in Section 7.5.

6.2 System Model

Consider a large-scale multi-channel single-hop broadcast network with one source node, n identical destination nodes and C non-overlapping frequency channels, where both n and C are large, and $C < n$ is particularly considered in this study. Each channel can accommodate transmission at a constant data rate R. For simplicity we assume each node can only transmit or receive on a single channel at a time. Our following analysis can be extended to the multi-radio case straightforwardly. A message of L-bit length is divided into $M(\in [2, L])$ short packets, transmitted separately for improved jamming resistance. Each packet is expanded with O-bit overhead including the message ID, fragment number, hash index[1], etc. The time is slotted

[1]The hash links are used to prevent insertion attacks [22].

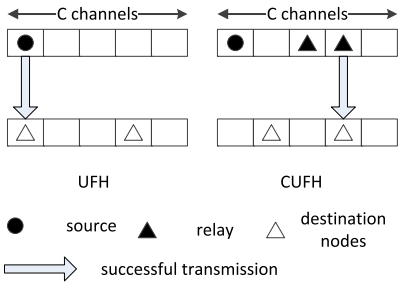

Figure 6.1: UFH and CUFH

and synchronized over the network[2], and the slot duration T_s is set for the transmission of one packet, i.e,

$$T_s = \frac{O + L/M}{R}. \tag{6.1}$$

In UFH the source hops at the start of every slot and transmits one packet in a randomly chosen channel. The packets are broadcast sequentially and repeatedly for a sufficiently long time or till the whole network successfully receives the message. Receivers may hop at the same rate as the source, or at a slower rate (which can relax the requirement on time synchronization between transmitters and receivers). At each hop receivers randomly and independently choose channels to listen. In CUFH, destination nodes serve as relays for the remaining receivers right after obtaining the message. Several channel selection strategies for relays and remaining receivers have been proposed in [24]. In this work we consider a representative one, where each relay randomly selects a channel for communication, and so does a receiver. This version of CUFH captures the essence of jamming resistant collaborative broadcast, and is amenable to analysis. No packet coordination is assumed among the transmitters[3], thus, collisions occur when they send *different* packets on the same channel. UFH and CUFH are illustrated in Fig. 6.1, where successful transmissions occur only when transmitters (source and relays) and receivers select the same channels.

As in [22–24], we consider omniscient jammers with bounded computation and transmission capacity. Jammers can be categorized into non-responsive and responsive ones, based on whether a jammer attempts to detect the ongoing transmission before jamming. It is assumed that a non-responsive jammer can block C_{nr} channels simultaneously, and change jamming channels every $t_{nr} \triangleq t_j + t_{sw}$ time, where t_j is the minimal time duration required for effective

[2]This is commonly assumed for a frequency hopping system, and our study does not require a high precision for this synchronization. There is also some discussion on the the asynchronous scenario; see Remark 8.

[3]The roles played by the source node and relays are identical for a receiver.

jamming of a channel and t_{sw} is the time cost for switching channels. A responsive jammer is more intelligent. It is assumed that she can sense C_r channels simultaneously, requiring t_r time to switch sensing channels, and jam the channels where transmissions are detected. There are two strategies for her to jam the channels: 1) jam the channels as soon as transmissions are detected; 2) jam the channels with (detected) ongoing transmissions altogether at the end of each slot. These two cases make no difference in UFH, while the former is typically more efficient than the latter in CUFH when there are many ongoing transmissions. In this work, we consider the first jamming strategy for a responsive jammer.

Without loss of generality we consider a single[4] jammer with full knowledge about the network protocol and the ability to acquire any pre-shared secrete keys (i.e., the jammer could be an insider). This jammer can block a channel with probability

$$P_{jam} = \frac{C_j T_s}{C} = \alpha_1 + \frac{\alpha_2}{M}, \tag{6.2}$$

where $\alpha_1 = \frac{C_j O}{CR}$, $\alpha_2 = \frac{C_j L}{CR}$, and C_j stands for the effective number of channels the jammer can block in a unit time. For a non-responsive jammer, $C_j = \frac{C_{nr}}{t_{nr}}$. For a responsive jammer, $C_j \in \left[\frac{C_r}{t_j+t_r}, \frac{C_r(T_s-t_j)}{T_s t_r}\right]$[5], where the upper bound is achieved when ongoing transmission occurs only in one channel, i.e., in UFH, and the lower bound is achieved when there are ongoing transmissions in most of channels as in CUFH (with a large number of relays). Note that in the latter case, sensing channels before jamming can barely benefit the jammer. Thus, a responsive jammer is less effective in CUFH than in UFH, which will be verified through simulation in Section V. However, our discussions in Sections 6.3 and 6.4 assume an abstract P_{jam}, which hold regardless of the jammer types. In addition, to possibly evade the attacks from the jammer it's reasonable to assume that $P_{jam} < 1$ (otherwise, all available channels can be blocked in a time slot). Under such an assumption, there exists an integer $M \in \left(\frac{\alpha_2}{1-\alpha_1}, L\right]$ subject to our design.

Jammers can launch other types of attacks including packets modification, replacement and insertion. We target on jamming attacks in this work as it is shown to be one of the most effective attacking strategies [22]. Message authenticity and confidentiality issues can generally be achieved on the application layer [21], and will not be addressed either. Solutions for other attacks in literature can be jointly considered with our approach without confliction.

The main metrics of our study are defined below.

Definition 4. Node broadcast delay *is defined as the time duration from the start of broadcast*

[4]In practice, multiple jammers may be deployed, whose influence can be modeled as one jammer with combined capability in a broadcast network.

[5]In practice $t_j \approx t_r << T_s$, thus the lower bound is smaller than the upper bound.

till the time when a particular node successfully receives the message; network broadcast delay *is defined as the time duration from the start of broadcast till the time when* all *the nodes in the network successfully receive the message.*

Definition 5. Packet reception rate *is defined as the probability that a destination node correctly decodes a packet in a slot.*

6.3 Performance of UFH

We first evaluate the performance of UFH in terms of network broadcast delay D^u. In particular,

$$D^u = \max\left(D_1^u, D_2^u, ..., D_n^u\right),$$

where $D_v^u, v = 1, 2, ..., n$ is the broadcast delay for node v. Note that $\{D_v^u\}$ are identically distributed but *not independent* in general since all the nodes staying at the same channel experience the same channel status (jammed or not). This dependency becomes prominent in a large-scale network as considered here. Since M packets are transmitted sequentially, we derive the network and node broadcast delay in terms of the number of *rounds* with each composed of M consecutive slots, for the convenience of analysis.

Let us first introduce a property of the Harmonic Number $\sum_{k=1}^{K} \frac{1}{k}$, which facilitates our analysis.

Lemma 6. *[66] The harmonic number* $\sum_{k=1}^{K} \frac{1}{k}$ *is bounded by*

$$\frac{1}{2(K+1)} < \sum_{k=1}^{K} \frac{1}{k} - \ln K - \gamma < \frac{1}{2K}, \tag{6.3}$$

where $\gamma = 0.572...$ *is the Euler-Mascheroni constant.*

Then, the average node broadcast delay $E\left(D_v^u\right)$ is derived below.

Proposition 5. *Given* $d \in \mathbb{N}$*, the Cumulative Distribution Function (CDF) of* D_v^u *in UFH is given by:*

$$\Pr\left(D_v^u \leq d\right) = \left(1 - (1 - p_1)^d\right)^M, \tag{6.4}$$

where

$$p_1 = \frac{1}{C}(1 - P_{jam}) \tag{6.5}$$

is the packet reception rate of UFH. Its average admits:

$$E\left(D_v^u\right) = \frac{-1}{\ln(1-p_1)} \sum_{k=1}^{M} \frac{1}{k} + o(1) = \theta(C \ln M). \tag{6.6}$$

The proof is given in Appendix D.1.

Some insights behind this proposition are given below:

- As long as $P_{jam} < 1$, a message can be delivered to an arbitrary node in a finite time with probability 1.
- Given the settings of the jammer there exist an optimal number of channels (C) and packets (M) such that the average node broadcast delay $E\left(D_v^u\right)$ is minimized. The optimal C can be obtained by maximizing p_1 in Eq. (6.5) (note that P_{jam} is a decreasing function of C), and the optimal M can be derived by minimizing $E(D_v^u)$ in Eq. (6.6) (note that p_1 decreases with M while the harmonic number increases with M).
- In practice either C or M might be set large in order to counteract jamming attacks. According to the scaling law result, a large M is more favorable since $E(D_v^u)$ scales *linearly* with C while only *logarithmically* with M.

Based on the marginal CDF of D_v^u, we bound the CDF of D^u as follows.

Theorem 9. *The CDF of D^u admits:*

$$\Pr(D^u \le d) \ge \left(1 - (1-p_1)^d\right)^{nM}.$$

Its average $E(D^u)$ is upper bounded by

$$E(D^u) \le \bar{D}^u \triangleq \frac{-1}{\ln(1-p_1)} \sum_{k=1}^{nM} \frac{1}{k} + o(1) = O(C \ln n), \tag{6.7}$$

where the equality holds when $P_{jam} = 0$ and receivers hop at the same rate as the source.

Proof. Since $D^u = \max(D_1^u, D_2^u, ..., D_n^u)$,

$$\begin{aligned}\Pr(D^u \le d) &= \Pr(D_1^u \le d, D_2^u \le d, ..., D_n^u \le d) \\ &= \prod_{v=1}^{n} \Pr(D_v^u \le d | D_{v-1}^u \le d, ..., D_1^u \le d) \\ &\ge \prod_{v=1}^{n} \Pr(D_v^u \le d) \\ &= \Pr(D_v^u \le d)^n = \left(1 - (1-p_1)^d\right)^{nM},\end{aligned} \tag{6.8}$$

where the inequality follows from the dependence of D_v^u, and becomes equality when $P_{jam} = 0$ and receivers hop at the same rate as the source (see Appendix D.2 for more details).

Then the average network broadcast delay is

$$\begin{aligned}E(D^u) &= \sum_{d=0}^{\infty} \Pr(D^u > d) \\ &\le \sum_{d=0}^{\infty} \left(1 - \left(1 - (1-p_1)^d\right)^{nM}\right) \triangleq z.\end{aligned}$$

Defining $\lambda = -\ln(1-p_1)$ and following the proof of Proposition 5 in Appendix D.1,

$$\frac{1}{\lambda} \sum_{k=1}^{nM} \frac{1}{k} < z < \frac{1}{\lambda} \sum_{k=1}^{nM} \frac{1}{k} + 1.$$

As a result $E(D^u) \le \frac{-1}{\ln(1-p_1)} \sum_{k=1}^{nM} \frac{1}{k} + o(1)$. The scaling law result follows from the property of the harmonic number in Lemma 6. The proof is hence completed. □

Some observations are in order:

- $E(D^u)$ scales logarithmically with the network size. This result was experimentally shown in [21] and hereby formally proved.
- According to Remark 14 in Appendix D.2, for a sufficiently large C, $\bar{D}^u$ is reasonably tight as long as the receivers hop at the same rate as the source.
- Eq. (6.8) implies that dependence of node broadcast delays $\{D_v^u\}$ is actually beneficial for the communication efficiency. For example, in the extreme case where receivers never hop, the network broadcast delay only depends on C, independent of the network size, since any nodes staying at the same channel can be treated as identical and there are at most C different nodes. Of course, this increases the vulnerability to jamming attacks, and there is a tradeoff between throughput and security.

- In practice the CDF of D^u can be exploited to set a time-out value for the transmission time of the source for energy conservation.

6.4 Performance of Collaborative UFH

6.4.1 Protocol

Without requiring pre-shared secret keys, UFH exhibits robustness to insider jamming attacks and good scalability with the network size, and yet poor communication efficiency due to the lack of coordination between the source and destination nodes. Collaborative UFH (CUFH) intends to accelerate the broadcast process by exploiting cooperative relaying. In order to apply CUFH in a large-scale network where $C < n$, the number of relays should be carefully controlled. Intuitively allowing more relays than the number of available channels in the network may harm, rather than benefit, the network performance, due to the fact that collisions, incurred by simultaneous transmissions, will congest most of the channels. Therefore, the original CUFH protocol in [24] is modified for the large-scale networks as follows. [6]

Protocol 1. *CUFH with Relay Constraint*

P0 Initialization: As in UFH, the source node randomly selects a channel from a pool of C frequency channels and broadcasts the packets of interest sequentially and repeatedly; similarly destination nodes randomly and independently choose channels to listen on;

P1 Relay accumulation: Destination nodes serve as relays right after obtaining the whole message. The relays randomly and independently select a channel for the transmission of each packet. This phase ends when there are n_r transmitters including the source node.

P2 Collaborative broadcasting: the n_r transmitters continue to broadcast their packets until all the remaining nodes receive them. No more new relays are admitted in this phase.

The number of cooperative transmitters, n_r, is a key parameter in this protocol. A small n_r would result in inadequate collaboration gain, while a large n_r , as indicated above, would incur severe collisions and deteriorate the network performance. An optimal value of n_r is given in Proposition 6 below.

6.4.2 Packet reception rate

Suppose there are n_r cooperative transmitters at some slot and an arbitrary destination node stays at channel c at the beginning of that slot. Denote by R_l the event that l relays select

[6] We will simply use CUFH to refer to this protocol in the following analysis.

channel c with the *same* packet to transmit. The packet reception rate of this node, in the absence of jammers, is derived as:

$$\begin{aligned}
&\tilde{p}_{n_r}(M) \\
&= \sum_{l=1}^{n_r} \Pr(R_l) \\
&= \sum_{l=1}^{n_r} \binom{n_r}{l} \left(\frac{1}{C}\right)^l \left(1-\frac{1}{C}\right)^{n_r-l} \left(\frac{1}{M}\right)^{l-1} \\
&= M\left[\sum_{l=0}^{n_r} \binom{n_r}{l} \left(\frac{1}{CM}\right)^l \left(1-\frac{1}{C}\right)^{n_r-l} - \left(1-\frac{1}{C}\right)^{n_r}\right] \\
&= M\left[a^{n_r} - b^{n_r}\right], \qquad (6.9)
\end{aligned}$$

where $a = 1-\left(1-\frac{1}{M}\right)\frac{1}{C}$, $b = 1-\frac{1}{C}$ and the last equality is due to the Binomial theorem (Eq. (1.111) in [67]).

Considering the influence of the jammer, the packet reception rate is given by:

$$p_{n_r}(M) = \tilde{p}_{n_r}(M)(1-P_{jam}), \qquad (6.10)$$

where P_{jam} is given in Eq. (6.2).

As a key metric, p_{n_r} offers many insights for performance evaluation and optimization.

Proposition 6. *The optimal number of transmitters, in terms of maximizing $p_{n_r}(M)$, is*

$$n_r^*(M) = \ln\left(\frac{\ln b}{\ln a}\right) / \ln\left(\frac{a}{b}\right) \approx -M\ln\left(1-\frac{1}{M}\right)C,$$

where a and b are given after Eq. (6.9), and the approximation is made for large C. We also have

$$n_r^*(M) \in [\underline{n}_r^*, \bar{n}_r^*], \qquad (6.11)$$

where $\underline{n}_r^* \approx C$ and $\bar{n}_r^* \approx 1.4C$ for large C.

The proof is given in Appendix D.3.

Remark 7. *As long as the number of nodes n is no larger than n_r^*, it is not necessary to control the number of relays. In addition, our simulation in Fig. 6.6 shows that the performance loss is not significant when the number of transmitters is within $[C, 1.4C]$, which facilitates the practical implementation of CUFH.*

Remark 8. *When timing between relays is significantly mismatched (e.g., due to malfunction of hardware), the packet reception rate is revised as (derivation omitted in the interest of space):*

$$\hat{p}_{n_r} = n_r \frac{1}{C}\left(1-\frac{1}{C}\right)^{2(n_r-1)}(1-P_{jam}), \tag{6.12}$$

due to the fact that one transmission of a particular transmitter can collide with two successive transmissions of others almost surely. Then following the derivation above, the optimal number of relay in this case is $n_r^* = \frac{-1}{2\ln(1-1/C)} \approx C/2$.

With the optimal number of transmitters we show next that $p_{n_r^*}$ is bounded at both ends by non-trivial values.

Proposition 7. *Given* $\alpha \triangleq \frac{\alpha_2}{1-\alpha_1} > 1.4$ *and a sufficiently large* C,

$$0.35\,(1-\alpha_1-\alpha_2/M) < p_{n_r^*} < 0.38(1-\alpha_1).$$

The proof is given in Appendix D.4. In practice the condition $\alpha > 1.4$ can be easily satisfied since the message length L (which α_2 depends on) is usually a large number (see Eq. (6.2)).

Finally we compare p_1 in UFH (Eq. (6.5)) with p_{n_r} in CUFH (Eq. (6.10)), through the cooperation gain defined as

$$g_c(n_r) \triangleq \frac{p_{n_r}}{p_1}.$$

Proposition 8. *The cooperation gain* $g_c(n_r)$ *is bounded by:*

$$n_r\left(1-\frac{1}{C}\right)^{n_r-1} < g_c(n_r) < 2C\left[\left(1-\frac{1}{2C}\right)^{n_r}-\left(1-\frac{1}{C}\right)^{n_r}\right],$$

and the optimal cooperation gain

$$g_c(n_r^*) \in (0.35C, 0.5C].$$

The proof is given in Appendix D.5.

Remark 9. *The left hand side of* $g_c(n_r)$ *indicates that when* n_r *is small (so that the item* $(1-1/C)^{n_r-1}$ *is close to* 1*) the cooperation gain almost increases linearly with* n_r*, the number of relays. This implies that the relay accumulation phase accelerates once the first relay appears. On the other hand the right hand side of* $g_c(n_r)$ *shows that when the number of relays exceeds certain value this gain drops exponentially.*

6.4.3 Ideal cooperative network

Ignoring the relay accumulating process in CUFH, we explore an ideal cooperative network in this subsection, which focuses on collaborative broadcast. The broadcast delay of this network will help us obtain relevant results for CUFH in the following (Theorem 12). Consider an ideal cooperative network (ICN) where there are $n_r(>> M)$ identical source nodes and n destination nodes. These source nodes transmit the packets sequentially and repeatedly but each starts transmission from a random packet (which implies there is no packet coordination among sources).

Denote by D_v^{icn} and D^{icn} the broadcast delay, measured by rounds, for an arbitrary destination node v and the whole network. We have the following results concerning these two quantities.

Theorem 10. *For the ideal cooperative network described above*

$$\begin{aligned}&\Pr(D_v^{icn} \le i) \triangleq \epsilon(n_r, i) \\ &= \begin{cases} \sum_{k=M}^{iM} \binom{iM}{k} p_{n_r}^k (1-p_{n_r})^{iM-k} P(M,k), & i \ge 1 \\ 0, & otherwise \end{cases}\end{aligned} \tag{6.13}$$

where

$$P(M,k) = \sum_{l=0}^{M} (-1)^l \binom{M}{l} (M-l)^k / M^k,$$

and p_{n_r} is given in Eq. (6.10). The mean of D_v^{icn} is given by

$$E(D_v^{icn}) = \frac{1}{p_{n_r}} \sum_{k=1}^{M} \frac{1}{k}.$$

The mean of D^{icn} is upper bounded by

$$E(D^{icn}) < \sum_{i=0}^{\infty} (1 - \epsilon(n_r, i)^n) \triangleq \bar{D}^{icn}.$$

Proof. Set $q = iM,\ i = 1, 2, ...$ The probability that k packets are successfully received within $q(\ge M)$ time slots at an arbitrary destination node v is $\binom{q}{k} p_{n_r}^k (1-p_{n_r})^{q-k}$, and the probability that these k successfully decoded packets include all the M different packets is $P(M,k)$ for $k \ge M$, derived in Appendix D.6. Thus, Eq. (6.13) is achieved. The empirical and analytical CDF of D_v^{icn} is given in Fig. 6.2, which indicates an excellent match.

Denote by Q_l the number of time slots before node v obtains l different packets. Set $t_l = Q_l - Q_{l-1}, \forall l \ge 1$ with $Q_0 = 0$. Then, the time needed to receive the whole message is

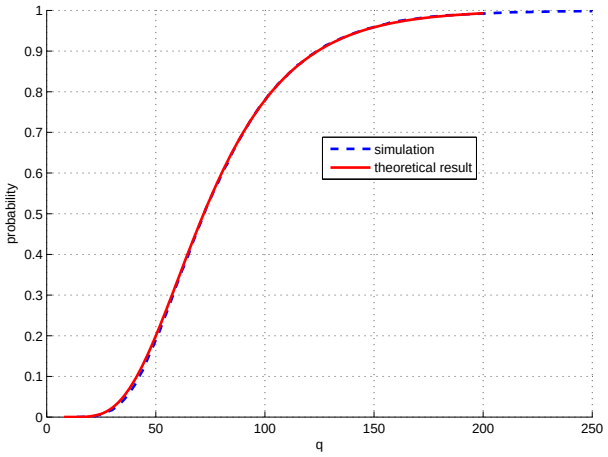

Figure 6.2: CDF of D_v^{icn}

given by (note that $D_v^{icn} = Q_M/M$)

$$Q_M = t_1 + t_2 ... + t_M, \tag{6.14}$$

where t_l, $l = 1, 2, ..., M$, is geometrically distributed with parameter[7] $\frac{M-l+1}{M} p_{n_r}$. Then,

$$E(Q_M) = E(t_1 + ... + t_M) = \frac{1}{p_{n_r}} \sum_{l=1}^{M} \frac{M}{M-l+1},$$

and $E(D_v^{icn})$ is verified.

Following the same line of the proof in Theorem 9, we get

$$E(D^{icn}) < \sum_{i=0}^{\infty} (1 - \epsilon(n_r, i)^n).$$

□

Corollary 4. *The average network broadcast delay of this ideal cooperative network can be approximated by:*

$$E(D^{icn}) \approx \frac{1}{p_{n_r}} \sum_{k=1}^{nM} \frac{1}{k}. \tag{6.15}$$

[7]The success probability decreases due to the increasingly stringent demand for a different packet from those already received.

The proof is given in Appendix D.7.

Remark 10. *Compared with UFH, it can be shown that when $n_r = n_r^*$,*

$$\frac{E(D^{icn})}{E(D^u)} \approx \frac{E(D_v^{icn})}{E(D_v^u)} \approx \frac{p_1}{p_{n_r^*}} = \frac{1}{g_c(n_r^*)},$$

where $E(D_v^u)$ and $E(D^u)$ are given in Eq. (6.6) and (6.7). Thus, this ideal cooperative network achieves the optimal cooperation gain characterized in Proposition 8.

6.4.4 Average network broadcast delay

Due to system dynamics in the relay accumulation process, the precise derivation of the average network broadcast delay $E(D^c)$ that CUFH incurs is intractable. Instead we turn to explore its lower and upper bound.

lower bound

As shown in Remark 9, the process is accelerated after the first relay appears. Before that no cooperation can be exploited and there is no difference between CUFH and UFH. Therefore, the time duration $D_{\min}^c$ from the beginning of the broadcast until the first relay appears contributes a substantial portion in the total network broadcast delay and $E(D_{\min}^c)$ serves as a lower bound for $E(D^c)$.

Theorem 11.

$$E(D^c) \geq E(D_{\min}^c) > \underline{D}^c \triangleq \sum_{i=0}^{\infty} \left(1 - (1 - (1 - p_1)^i)^M\right)^n. \tag{6.16}$$

Proof.

$$D^c = \max(D_1^c, D_2^c, ..., D_n^c) \geq D_{\min}^c = \min(D_1^c, D_2^c, ..., D_n^c),$$

where D_v^c is the broadcast delay for node v. Before the first relay appears there is only one transmitter in the network. Thus, $D_{\min}^c$ coincides with the minimal time delay among nodes in UFH. Then the probability that $D_{\min}^c$ is larger than i rounds is given by (c.f. proof of Theorem 9):

$$\begin{aligned}\Pr[D_{\min}^c > i] &= Pr(D_1^u > i, D_2^u > i, ..., D_n^u > i) \\ &> [\Pr(D_1^u) > i]^n = \left(1 - (1 - (1 - p_1)^i)^M\right)^n,\end{aligned}$$

where D_v^u is the broadcast delay of node v in UFH (with no relays) and the inequality is due to the dependence of D_v^u. Then,

$$E(D_{\min}^c) > \sum_{i=0}^{\infty} \left(1 - (1 - (1 - p_1)^i)^M\right)^n \triangleq \underline{D}^c.$$

□

upper bound

We examine the upper bound of the average network broadcast delay by studying a suboptimal protocol for CUFH, CUFH-p, which can be easily implemented in practice.

Protocol 2. *CUFH-p*

P0 The source node predetermines $n_r - 1$ destination nodes as potential relays and imbeds their IDs in the message broadcast. Then the source transmits the packets sequentially using UFH.

P1 These predetermined relay nodes serve as relays right after obtaining the whole message. The relays randomly and independently select channels for the transmission of each packet.

The only difference between CUFH-p and CUFH (Protocol 1) is the way to select relays. The relay nodes in CUFH-p are predetermined rather than dynamically selected among all the destination nodes as in CUFH. This incurs some performance degradation (since non-relay nodes cannot help even if they obtain the message before some relay nodes), while significantly simplifies the implementation of CUFH[8]. Our analysis and simulation show that CUFH-p still outperforms UFH dramatically, and asymptotically achieves the optimal cooperation gain.

Theorem 12. *The average network broadcast delay that CUFH-p incurs admits*

$$E(D^p) < \bar{T}_a + \bar{T}_b \triangleq \bar{D}^c,$$

where

$$\bar{T}_a = \frac{1}{p_1} \sum_{k=1}^{M(n_r-1)} \frac{1}{k} + 1,$$

and

$$\bar{T}_b = \sum_{i=0}^{\infty} \left(1 - \epsilon(n_r, i)^{n-n_r+1}\right),$$

with $\epsilon(n_r, i)$ given in Eq. (6.13).

[8]Dynamical constraint on the number of relays to a predetermined value usually requires feedbacks from destination nodes to the source node.

Proof. The broadcast process of CUFH-p consists of two phases. For the first phase, the time delay T_a for all the designated relay nodes to receive the message is bounded by the time delay in UFH with network size n_r. Then by Eq. (6.7),

$$T_a < \frac{1}{p_1} \sum_{k=1}^{M(n_r-1)} \frac{1}{k} + 1 \triangleq \bar{T}_a.$$

After all the relay nodes obtain the message there are n_r transmitters and at most $n - n_r + 1$ destination nodes which have not received the message. These destination nodes need to receive at most M packets. Then, according to Theorem 10, the time delay T_b for all the remaining nodes to obtain the message is upper bounded as

$$T_b < \sum_{i=0}^{\infty} \left(1 - \epsilon(n_r, i)^{n-n_r+1}\right) \triangleq \bar{T}_b \tag{6.17}$$

Hence the proof is completed. □

Corollary 5. *When $n_r = n_r^*$ and $C = o(\ln n)$, $\bar{D}^c = O(\ln n)$, and $\lim_{n\to\infty} \frac{\bar{D}^c}{\bar{D}^u} \approx \frac{1}{g_c(n_r^*)} \leq 3/C$, where $\bar{D}^u$ is given in Eq. (6.7).*

Proof. When $n_r = n_r^*$,

$$\begin{aligned}
\bar{T}_b &\approx \sum_{i=0}^{\infty} \left[1 - \left(1 - e^{-ip_{n_r}}\right)^{M(n-n_r^*+1)}\right] \\
&< \frac{1}{p_{n_r^*}} \sum_{k=1}^{M(n-n_r^*+1)} \frac{1}{k} + 1 \\
&< \frac{\frac{1}{2M(n-n_r^*+1)} + \ln(M(n-n_r^*+1)) + \gamma}{0.35(1-P_{jam})} + 1,
\end{aligned}$$

where the approximation follows from Corollary 4 and the last inequality follows from Proposition 7. According to Theorem 9, $\bar{T}_a = O(C \ln C)$. Thus, $\bar{D}^c = O(C \ln C + \ln n) = O(\ln n)$, and the ratio

$$\lim_{n\to\infty} \frac{\bar{D}^c}{\bar{D}^u} \approx \frac{p_1}{p_{n_r^*}} = \frac{1}{g_c(n_r^*)}.$$

□

Therefore CUFH-p asymptotically achieves the optimal cooperation gain, so as CUFH.

6.5 Simulation Results and Discussions

We substantiate our analysis by simulations in this section. The following simulation setting is adopted for a time-synchronized network unless otherwise noted:

- network configuration: the number of channels $C = 64$, message length $L = 550$ bits divided into $M = 5$ packets, each appended with $O = 10$ bits overhead, and data rate $R = 2$Mbps. Based on these parameters the time duration for one slot $T_s = 60\mu s$.
- jammer configuration: for a non-responsive jammer $t_j = 4\mu s$, $t_{sw} = 8\mu s$ and $C_{nr} = 7$; for a responsive jammer $t_r = 8\mu s$ and $C_r = 5$. Such a setting leads to the same jamming probability $P_{jam} = 0.547$ in UFH for both types of jammers.

As we will see in Fig. 6.7, the responsive jammer is less effective than the non-responsive jammer for CUFH, while they make no difference for UFH. To fairly compare the performance of UFH and CUFH, we consider a non-responsive jammer in Figs. 6.3-6.6.

The average node and network broadcast delay UFH incurs are depicted in Fig. 6.3 with respective to the number of nodes n. We can see that both node and network broadcast delay analyzed in Proposition 5 and Theorem 9 match the simulation very well.

Fig. 6.4 demonstrates the average network broadcast delay CUFH and CUFH-p incur, with $n_r = 71$, the optimal number of relays given in Proposition 6. An interesting observation is that the average network broadcast delay of CUFH decreases with n and that of CUFH-p increases logarithmically with n. This is due to the fact that the time delay for relay accumulation in CUFH decreases with network size n (reflecting the effect of multiuser diversity), which dominates in the overall network broadcast delay for the given range of n. This point is also substantiated by its lower bound (Theorem 11), which only accounts for the time duration before the appearance of the first relay. In contrast the delay for relay accumulation in CUFH-p is constant, independent of the network size, since relays are pre-determined. Therefore the average network broadcast delay of CUFH-p increases with n logarithmically as the delay in the collaborative broadcasting does (Corollary 5). However, compared with UFH in Figure 6.3, even CUFH-p significantly outperforms UFH, and their cooperation gain increases with the number of nodes. The upper bound is loose due to the fact that $\bar{T}_a$ in Theorem 12 is derived based on UFH. Nevertheless, with the insights obtained in our analysis, we may estimate the average broadcast delay of CUFH as $\underline{D}^c + D^{icn}$, where $\underline{D}^c$ and D^{icn} are given in Eq. (6.16) and (6.15), respectively. Basically the former accounts for the delay for accumulating relays and the latter accounts for the delay of cooperative broadcast. We can see in the figure that this estimation is quite accurate.

Fig. 6.5 compares the average network broadcast delay of UFH and CUFH with respective to the number of channels C, where $n = 800$, and the number of relays in CUFH is optimal (c.f.

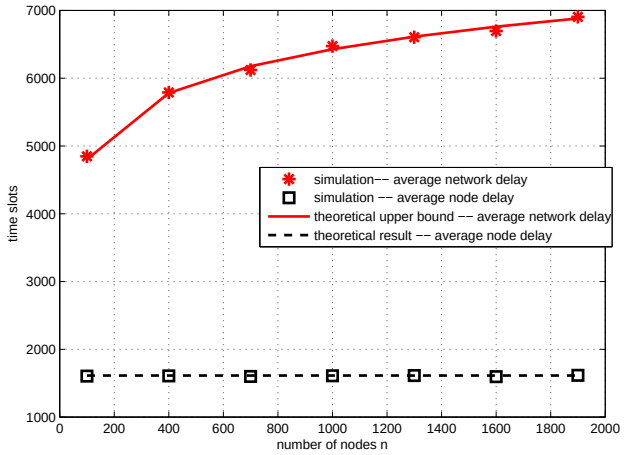

Figure 6.3: Average node and network broadcast delay of UFH

Proposition 6). We can observe that the average network delay of UFH increases significantly with C, while that of CUFH is much less sensitive with C. Therefore, in practice we may choose a large C (whenever applicable) for CUFH, to enhance its immunity to jamming attacks, without much concern on its degradation in communication efficiency.

The performance of CUFH with different number of relays is demonstrated in Fig. 6.6. It is clear from the figure that the best performance of CUFH is achieved when $n_r = 71$, regardless of network size, which coincides with our analysis in Proposition 6 (for $M = 5$, and $C = 64$). The reason why $n = 100$ incurs longer delay than $n = 400$ is similar to Fig. (6.4), i.e., more nodes accelerate the relay accumulation process, the delay of which dominates the overall broadcast delay. In addition, it is also observed that the performance loss is not significant if the number of relays is within the range of $C = 64$ and $1.4C = 90$, which gives us the flexibility in practice to control the number of relays.

Finally the performances of CUFH and UFH under attacks from both the non-responsive and the responsive jammers are compared in Fig. 6.7. According to our setting, the responsive jammer achieves the same jamming probability as the non-responsive jammer in UFH, which is clearly demonstrated in the figure (the upper two curves). In contrast, the jamming probability of the responsive jammer in CUFH is within the range $[0.391, 0.547]$ (see our discussion after Eq. (6.2) in Section 7.2). Intuitively, when there are effective transmissions on most channels, sensing before jamming loses its power. Thus, the responsive jammer is less effective in CUFH (including CUFH-p) than in UFH. This is confirmed in the figure (the bottom four curves), where the average network broadcast delay is smaller in the presence of the responsive jammer as opposed to the non-responsive jammer.

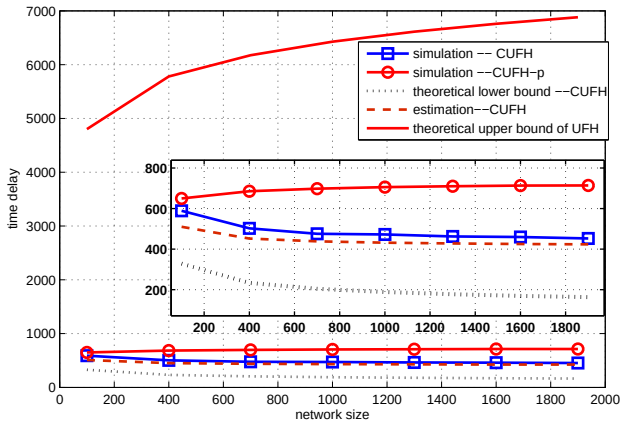

Figure 6.4: Average network broadcast delay of CUFH. The curves at the bottom are amplified in the square inside

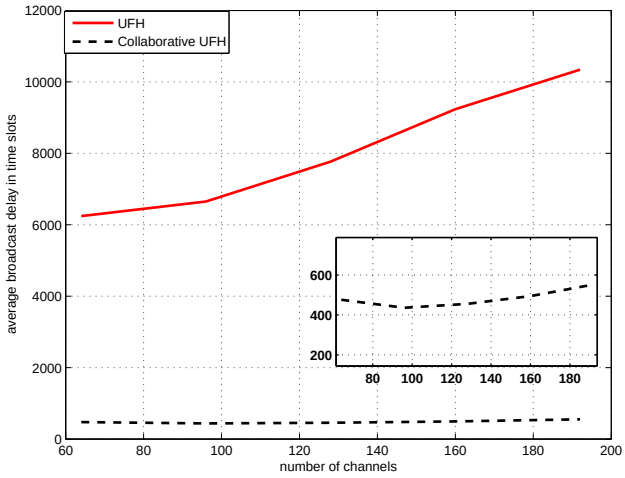

Figure 6.5: Comparison between CUFH and UFH on the number of channels C. The curve of CUFH is amplified in the square inside.

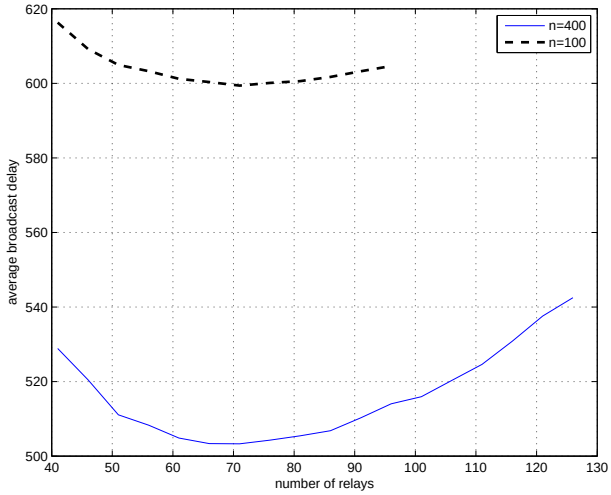

Figure 6.6: The average network broadcast delay of CUFH VS the number of relays in CUFH

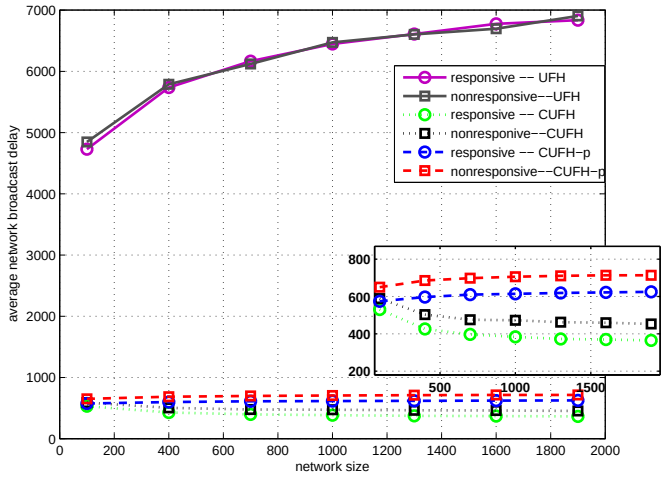

Figure 6.7: The average network broadcast delay with the responsive jammer. The four curves at the bottom are amplified in the square inside.

6.6 Conclusions

In this chapter we have examined the communication efficiency of UFH and Collaborative UFH in large-scale multi-channel networks, in terms of average node and network broadcast delay. Many insights related to protocol design and optimality have been provided for better understanding of anti-jamming broadcast in multichannel networks. We have also proposed a suboptimal protocol CUFH-p, which simplifies the implementation of CUFH and still outperforms UFH significantly. Our analytical results are well supported by simulation, and may find broader applications in emerging networks and systems.

Chapter 7

Connectivity and Multi-hop Delay of Multi-channel Wireless Networks Under Jamming Attacks

7.1 Introduction

7.1.1 Motivation

When network is under jamming attacks one of our primary concerns is network connectivity. This problem would be trivial if a single jammer was powerful enough to jam the whole space the network nodes occupy, which fortunately is not the general case in practice. Jammers are usually limited in power both due to hardware constraint and for the sake of reducing the risk of being detected. Usually each jammer is associated with a jamming region bounded in size. In a multi-hop multi-channel network the nodes staying in the jamming regions can switch to Uncoordinated Frequency Hopping (UFH) (see Chapter 6 for more details), a jamming countermeasure without pre-shared keys, to possibly resist jamming when jamming attacks are detected[1]. However for the nodes staying in the intersecting areas of several jamming regions, UFH may fail since they suffer from (much) stronger jamming than the nodes locating in a single jamming region. Any communication links involving these nodes are no longer available. As a consequence, the connectivity issue in such scenarios is more complex than that of a jamming-free network.

Our next concern is the multi-hop delay incurred by multi-channel networks under jamming attacks. Intuitively jamming countermeasures such as UFH help enhance network resistance to jamming, while typically degrade network performance such as throughput or transmission

[1]Jammer detection schemes are beyond the scope of this study.

delay. In particular UFH admits longer transmission delay than conventional communication protocol such as (coordinated) FH, due to lack of coordination for receivers and transmitters on channel selection. We'll show that actually UFH incurs no performance loss, in terms of scaling law.

Our contributions in this work are summarized below:

- We investigate the connectivity of a multi-hop multi-channel network with jamming attacks, in the framework of percolation theory. A jamming region is specified for each jammer and the influence of multiple jammers are considered. We show that the network experiences two phase transitions as the jammer density increases: from strong connection where there exists an infinite component composed of regular links, to weak connection where there exists an infinite component composed of both regular and UFH links, then to disconnection. Additionally there are two cutting jammer densities such that if the jammer density is beyond one of these densities, the network is either disconnected, or no longer strongly connected, for any network density.

- The multi-hop delay incurred by UFH on a percolated network is examined. It is shown that the multi-hop broadcast delay between a pair of nodes scales linearly with their Euclidean distance, which implies that anti-jamming techniques such as UFH incurs no performance loss, in terms of scaling law.

Network connectivity and multi-hop broadcast delay has been explored for homogenous networks ([33, 68–70] and references therein) in last few years. Recently these studies are extended to heterogenous networks such as Cognitive Radio networks [71, 72]. To the best of our knowledge there is no such study available in literature for multi-channel networks under jamming attacks.

The remainder of the chapter is organized as follows. The system model is introduced in Section 7.2. In Section 7.3 network connectivity is discussed with regular links replaced by the UFH links when jamming attacks are detected. The multi-hop delay of multi-channel networks under jamming attacks is then studied in Section 7.4. Finally we conclude this paper in Section 7.5.

7.2 System Model

7.2.1 Network model

Consider a multi-channel multi-hop network with n nodes and C non-overlapping frequency channels. All the nodes are uniformly distributed in a space $S_n = [0, \sqrt{\frac{n}{\lambda_n}}]^2$ with spatial density λ_n, and admit identical transmission range r_n. A group of m jammers are overlaid

with the legitimate network, and uniformly distributed in S_n as well. Due to limited power, each jammer jams a bounded area referred to as a jamming region, which is a circle centered at the jammer with radius r_m, and any nodes located in jamming regions are considered under jamming attacks if they communicate in the same channels where the jammers launch attacks. In practice r_m could be large, for instance, $r_m > r_n$, in order to cause severe destruction to the network, or small to reduce the risk of being detected. For simplicity, we assume that a jammer has no influence for nodes outside its jamming region. Let $\mathcal{X}_n \triangleq \{X_1, X_2, ..., X_n\}$ and $\mathcal{Y}_m \triangleq \{Y_1, Y_2, ..., Y_m\}$, where X_i and Y_j denote the random locations of node i and jammer j, respectively. As $n \to \infty$, $\mathcal{X}_n$ and $\mathcal{Y}_m$ converge in distribution to homogeneous Poisson point processes (P.P.P.) with density λ_n and $\lambda_m = \frac{m}{n}\lambda_n$, respectively. It is assumed that $m = \theta(n)$ such that λ_m is a constant; otherwise $\lambda_m \to 0$ for $m = o(n)$, or $\lambda_m \to \infty$ for $m = \omega(n)$, as $n \to \infty$, neither of which is of interest in the asymptotic study. In our analysis the jamming-free network is modeled as a random geometric graph $G(V_n, E_{\lambda_n})$, where V_n is the vertex set composed of all the nodes and E_{λ_n} is the edge set. An undirect edge exists between two nodes x_i and x_j if their Euclidean distance is no larger than r_n, i.e., $||x_i - x_j|| \leq r_n$. Under jamming attacks some links in E_{λ_n} may not be available, which leads to a random graph $G(V_n, E_{\lambda_n,\lambda_m})$ induced from $G(V_n, E_{\lambda_n})$, where $E_{\lambda_n,\lambda_m} \subseteq E_{\lambda_n}$.

Remark 11. *In general uniform distribution for jammers' positions is not necessarily optimal. However it's a reasonable model for research as 1) global knowledge on the network topology or distribution of network flows may not be available to jammers, 2) it's more efficient to uniformly distribute jammers than place them in predetermined positions, especially in a battle field and adverse environment; and 3) it is typically more difficult to detect uniformly distributed jammers.*

7.2.2 Uncoordinated Frequency Hopping

We consider jammers with powerful and yet bounded computation and transmission capability. They are aware of any network protocols including pre-shared secret keys (jammers may be insiders). Therefore traditional anti-jamming schemes such as the spread spectrum techniques are inefficacious[2]. In our work, the endpoints of the jammed links switch to Uncoordinated Frequency Hopping discussed in the last chapter, where a transmitter (receiver) randomly chooses a channel out of the pool to transmit (listen) and no channel synchronization between transmitters and receivers is required. The time is slotted and transmitters hop every time slot, while receivers hop much more slower so that they agree on the same channel within a finite time. In order to increase jamming resistance, a message is usually divided into multiple packets, which are transmitted sequentially and repeatedly. During one slot, we assume each

[2]They require pre-shared secrets such as spreading code sequences and frequency hopping patterns.

jammer can jam a particular channel with probability p_j. Multiple jammers admit stronger jamming capability than a single jammer, especially through collaboration. In our study, we assume a single jammer is not able to jam all the channels in its jamming region during one slot; while $J(\triangleq \lfloor \frac{1}{p_j} \rfloor)$ or more cooperative jammers can block any node in the intersection of their jamming regions. As a result in the random graph $G(V_n, E_{\lambda_n,\lambda_m})$,

$$E_{\lambda_n,\lambda_m} = \{e | e \in E_{\lambda_n} \text{ and } S_j(e) \text{ is satisfied}\},$$

where $S_j(e)$ denotes the condition that neither endpoint of e is covered by the jamming regions of J or more jammers.

7.3 Network Connectivity

Conventional communication schemes such as spread spectrum fail to offer jamming resistance once their secret keys are exposed. However as long as not all the channels accessible to receivers are blocked, communication links may still be active by switching to UFH, at the cost of longer (but finite) transmission delay. Links in E_{λ_n,λ_m} are divided into two types:

$L1$: jamming-free links (regular links): *neither* endpoint is in any jamming regions; communication over these links is operated as usual;

$L2$: jamming-resistant links (UFH links): links in E_{λ_n,λ_m} that are not of $L1$ type, i.e., at least one endpoint is subject to jamming but there still exists communication opportunities in between; regular communication over these links fails and UFH is exploited to avoid jamming attacks.

An $L2$ link is depicted in Fig. 7.1, with the assumption that $J \geq 4$. Note that 1) a jammer may jam both endpoints of a link; 2) for a successful reception of a packet on an $L2$ link, multiple attempts may be needed due to lack of channel coordination between the two endpoints and disruption from jammers. Thus $L2$ links incur (much) longer transmission delay than $L1$ links. For the ease of description a node is said *unjammed* if it belongs to an $L1$ link; a node is said *jammed* if it belongs to an $L2$ link; otherwise it is said *blocked*.

The connectivity of $G(V_n, E_{\lambda_n,\lambda_m})$ is investigated in the framework of percolation theory. In particular the network is connected or percolated if there exists a giant component almost surely (a.s.), which is a component in $G(V_n, E_{\lambda_n,\lambda_m})$ with infinite size as $n \to \infty$. Since $L1$ and $L2$ links perform differently, a percolated network is considered as *strongly* connected a.s. if the giant component is composed of only $L1$ links; otherwise *weakly* connected. Strong connection indicates that jammers are not powerful enough to have severe damage to the network.

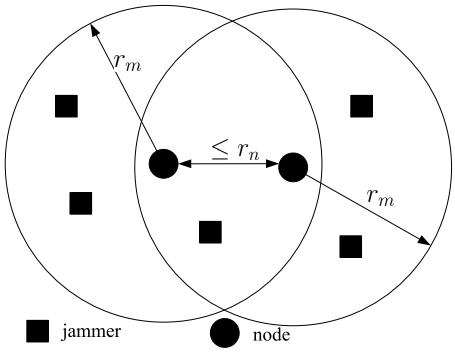

Figure 7.1: An $L2$ link

It is well known that for the random graph $G(V_n, E_{\lambda_n})$, there exists a critical density $\lambda_{n,c}$ such that when $\lambda_n > \lambda_{n,c}$ the network is connected; and if $\lambda_n < \lambda_{n,c}$, the network is disconnected. So the value of $\lambda_{n,c}$ has not been determined analytically, while its bounds were obtained in literature, e.g., $\lambda_{n,c} \in (0.696, 3.372)$ for $r_n = 1$ [73]. Our purpose in this work is to identify and analyze such phase transition phenomena in the graph $G(V_n, E_{\lambda_n,\lambda_m})$. Our results are summarized below and their proof is presented in the following subsections.

Theorem 13. *Given $\lambda_n > \lambda_{n,c}$, there exist two critical jammer densities $\lambda_{m,l}$ and $\lambda_{m,u}$ ($> \lambda_{m,l}$) depending on λ_n such that if $\lambda_m < \lambda_{m,l}$, $G(V_n, E_{\lambda_m,\lambda_n})$ is strongly connected; if $\lambda_m \in (\lambda_{m,l}, \lambda_{m,u})$, $G(V_n, E_{\lambda_m,\lambda_n})$ is weakly connected; and if $\lambda_m > \lambda_{m,u}$, it is disconnected. In addition, both functions $\lambda_{m,l}(\lambda_n)$ and $\lambda_{m,u}(\lambda_n)$ are non-decreasing.*

This theorem shows that the network experiences two phase transitions (Fig. 7.2) as the jammer density λ_m increases, given the capability of jammers. An instance of strongly and weakly connected networks is depicted in Fig. 7.3 and 7.4, respectively, where the solid red lines are $L1$ links and the dotted black lines are $L2$ links. And when $\lambda_m > 0.55$ the network is disconnected, the figure of which is omitted due to limited space. The network and jammer configuration in our simulation is $n = 2000, \lambda_n = 2$, $r_n = 1$, and $r_m = 1.3$.

Remark 12. *The weak connection is due to employment of UFH. Without such countermeasures as UFH, network may only experience one phase transition: if $\lambda_m < \lambda_{m,l}$ the network is connected, and if $\lambda_m > \lambda_{m,l}$ it is disconnected.*

In addition, the theorem above implies that increasing network density may help improve its jamming resistance since $\lambda_{m,l}(\lambda_n)$ and $\lambda_{m,u}(\lambda_n)$ are non-decreasing functions of λ_n. Unfortunately, the next theorem reveals that if the jammer density is high enough the network is no longer connected no matter how large the network density is.

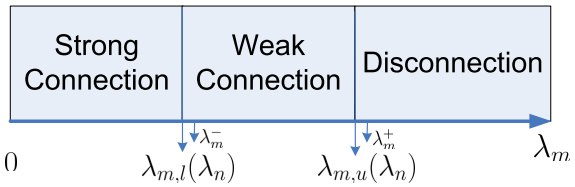

Figure 7.2: Phase transition of network

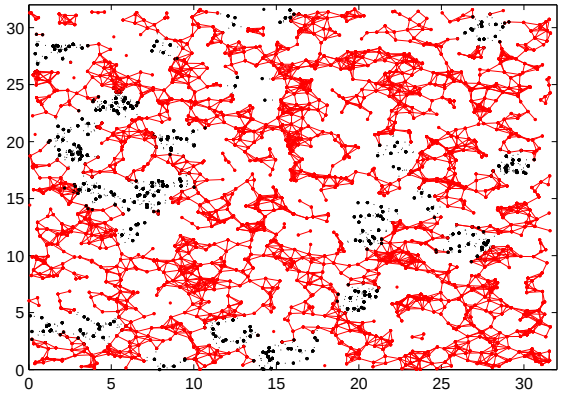

Figure 7.3: Strongly connected network:$\lambda_m = 0.03$

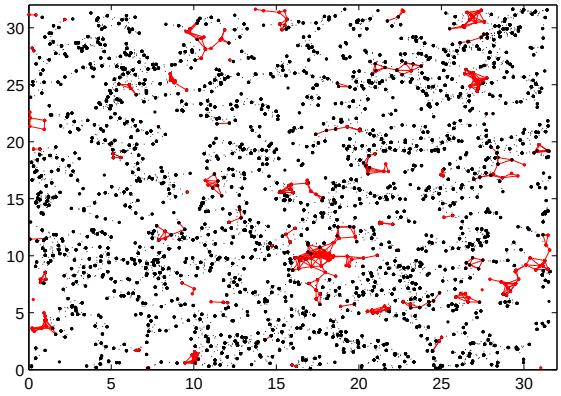

Figure 7.4: Weakly connected network:$\lambda_m = 0.35$

Theorem 14. *There exist two finite cutting jammer densities λ_m^+ and λ_m^- defined as:*

$$\lambda_m^+ = \sup\{\lambda_m : \exists \lambda_n \ s.\ t.\ G(V_n, E_{\lambda_n,\lambda_m}) \ \ is\ weakly\ connected\},$$

and

$$\lambda_m^- = \sup\{\lambda_m : \exists \lambda_n \ s.\ t.\ G(V_n, E_{\lambda_n,\lambda_m}) \ \ is\ strongly\ connected\},$$

respectively.

This theorem reveals that when $\lambda_m > \lambda_m^-$ the network is no longer strongly connected for any network density λ_n and the network is not even connected for any λ_n when $\lambda_m > \lambda_m^+$. The relationship among critical and cutting jamming densities is illustrated in Fig. 7.2.

7.3.1 Critical density $\lambda_{m,l}$ and $\lambda_{m,u}$

We focus our analysis on the critical density $\lambda_{m,u}$, which is the critical point separating connection from disconnection. The existence of $\lambda_{m,l}$ can be similarly proved with slight changes.

To begin with, we introduce some facts about the general Poisson Boolean Model $B(\mathcal{X}, r, \lambda)$ in percolation theory [73], where $\mathcal{X} = \{X_i\}$ is P.P.P. with density λ and each point X_i is associated with a ball with radius r. A region is occupied if it is covered by at least one ball. And vacant regions are the complement of occupied regions. A continuous occupied (vacant) region is named as an occupied (vacant) component and a continuous curve in the occupied region is called an occupied path.

Definition 6. *An L(left)-R(ight) occupied crossing in the rectangle $[0, l_1] \times [0, l_2]$ is an occupied path op, which intersects with both the left and the right boundaries of the rectangle, i.e., $op \cap 0 \times [0, l_2] \neq 0$, and $op \cap l_1 \times [0, l_2] \neq 0$, and the segment between the two intersecting points is fully contained in the rectangle. The T(op)-B(ottom) occupied crossing is defined similarly.*

Denote by $\lambda_c(r)$ the critical density for $B(X, r, \lambda)$, then we have the following facts:
Fact 1: when $\lambda < \lambda_c(r)$, there is no infinite occupied component a.s. and there is one infinite vacant component a.s.; and when $\lambda > \lambda_c(r)$, there is one infinite occupied component a.s. and there is no infinite vacant component a.s.
Fact 2: Define $LR(l_1, l_2, \lambda)$ ($TB(l_1, l_2, \lambda)$) as the event that there exists an occupied L-R (T-B) crossing in the rectangle $[0, l_1] \times [0, l_2]$. Then when $\lambda > \lambda_c(r)$, for every $k \geq 1$, $\lim_{n\to\infty} \Pr(LR(kn, n, \lambda)) = 1$; when $\lambda < \lambda_c(r)$, for every $k \geq 1$, $\lim_{n\to\infty} \Pr(LR(kn, n, \lambda)) = 0$. The same conclusion also holds for $TB(l_1, l_2, \lambda)$.

Based on the definition of the Poisson Boolean Model, the random graph $G(V_n, E_{\lambda_n})$ is equivalent to $B(\mathcal{X}_n, r_n/2, \lambda_n)$. Thus the critical density $\lambda_{n,c} = \lambda_c(r_n/2)$.

We prove the existence of $\lambda_{m,u}$ through the following two steps:

Lemma 7. *Given $\lambda_n > \lambda_c(r_n/2)$, $\exists \underline{\lambda}_{m,u}(\lambda_n) > 0$ such that there exists a giant component in $G(V_n, E_{\lambda_n,\lambda_m})$ a.s. for any $\lambda_m < \underline{\lambda}_{m,u}(\lambda_n)$.*

The proof is given in Appendix D.8.

Lemma 8. *Given $\lambda_n > \lambda_c(r_n/2)$, if $\lambda_m > \bar{\lambda}_{m,u}(\lambda_n)$, the network is disconnected, where $\bar{\lambda}_{m,u}(\lambda_n)$ is the solution of*

$$\sum_{i=0}^{J-1} e^{-\bar{\lambda}_{m,u}\pi r_m^2} \frac{(\bar{\lambda}_{m,u}\pi r_m^2)^i}{i!} = \frac{1}{\lambda_n \pi r_n^2}.$$

The proof is given in Appendix D.9.

Jointly considering the two lemmas above, we reach the conclusion that there exists a jammer density $\lambda_{m,u} \in (\underline{\lambda}_{m,u}, \bar{\lambda}_{m,u})$ such that the network is connected if $\lambda_m < \lambda_{m,u}$, and disconnected if $\lambda_m > \lambda_{m,u}$.

Proof of the non-increasing property of $\lambda_{m,u}(\lambda_n)$ is based on the coupling technique[3]. It is known that fixing the jammer density λ_m, increase in the network density at least does not hurt network connectivity. Given two networks $G(V_{\lambda_{n1}}, E_{\lambda_{n1},\lambda_m})$ and $G(V_{\lambda_{n2}}, E_{\lambda_{n2},\lambda_m})$, where $\lambda_{n1} < \lambda_{n2}$, if the former one is percolated, so is the latter one. Thus $\lambda_{m,u}(\lambda_{n1}) \leq \lambda_{m,u}(\lambda_{n2})$.

To prove the existence of the critical density $\lambda_{m,l}$, the edge set E_{λ_n,λ_m} is modified as

$$E_{\lambda_n,\lambda_m} = \{e | e \in E_{\lambda_n} \text{ and } S'_j(e) \text{ is satisfied}\},$$

where $S'_j(e)$ denotes the condition that there is no jammer within r_m distance from either endpoint of e. It's easy to see that E_{λ_n,λ_m} is only composed of $L1$ links.

Similarly, we have the following two lemmas.

Lemma 9. *Given $\lambda_n > \lambda_c(r_n/2)$, $\exists \underline{\lambda}_{m,l}(\lambda_n) > 0$ such that if $\lambda_m < \underline{\lambda}_{m,l}(\lambda_n)$, the network is strongly connected.*

The proof follows the same line as that of Lemma 7, with E_2 be defined as: there are no jammers within r_m distance from any node on the crossings involved in E_1.

Lemma 10. *Given $\lambda_n > \lambda_c(r_n/2)$, if $\lambda_m > \bar{\lambda}_{m,l}(\lambda_n)$, the network is not strongly connected, where $\bar{\lambda}_{m,l}(\lambda_n) = \frac{\ln(\lambda_n \pi r_n^2)}{\pi r_m^2}$.*

The proof follows the same line as that of Lemma 8, with the following modifications. The average degree u of an arbitrary node a is defined as

$$u = E(\deg(a) | \overline{Jam}(a)),$$

[3]Intuitively, coupling accounts for simultaneously constructing two realization of networks on the same probability space, with the condition that if some event occurs in one realization, it should also happen in the other.

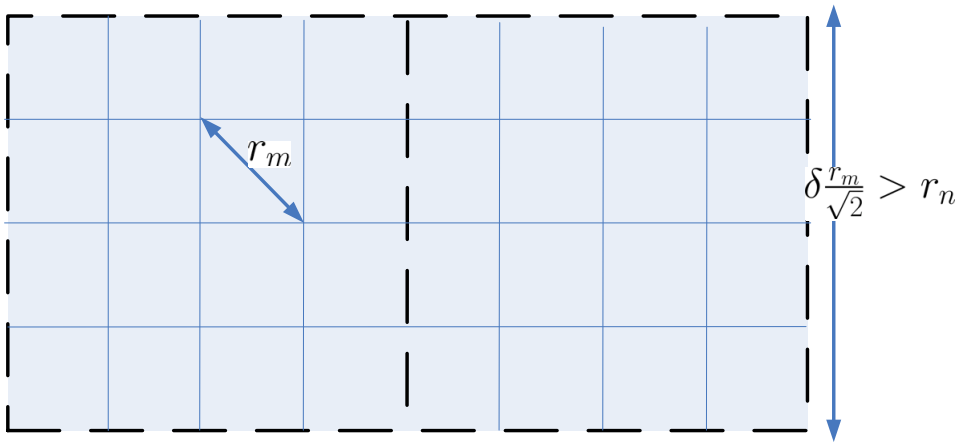

Figure 7.5: Grid $\mathcal{G}_\delta$ (dashed lines) and sub squares (solid lines)

where $\overline{Jam}(a)$ is the event that a is not jammed, i.e, there are no jammers within r_m distance from a. Then, following the same line as that of Lemma 8, we have

$$u \leq \lambda_n \pi r_n^2 \Pr[\overline{Jam}(b)] = \lambda_n \pi r_n^2 e^{-\lambda_m \pi r_m^2}.$$

$\bar{\lambda}_{m,l}$ follows by solving the equation $\lambda_n \pi r_n^2 e^{-\bar{\lambda}_{m,l}\pi r_m^2} = 1$.

As a result, the critical density $\lambda_{m,l} \in (\underline{\lambda}_{m,l}, \bar{\lambda}_{m,l})$.

7.3.2 Cutting density

Again we focus on the discussion of λ_m^+ and our arguments also hold for λ_m^- with slight modification. The intuition behind our proof lies in that if the jammer density is large enough the continuous unblocked region where no nodes are blocked is bounded in size and communication is not possible for nodes in separate unblocked regions. Thus, a giant component of $G(V_n, E_{\lambda_n,\lambda_m})$ no longer exists for any $\lambda_n > 0$.

Our argument is based on the site percolation model [47]. Let us construct a grid $\mathcal{G}_\delta$ (Fig. A.1) with edge length $\delta\frac{r_m}{\sqrt{2}} (\geq r_n)$, where $\delta = \lceil \sqrt{2} \lceil \frac{r_n}{r_m} \rceil \rceil$. Then, we partition each site (unit square in $\mathcal{G}_\delta$) into δ^2 sub-squares with side length $\frac{r_m}{\sqrt{2}}$. A site is called occupied if all the nodes located in it are blocked, otherwise it is called vacant. Let E_{ss} denote the event that there are at least J jammers in a sub-square. Then, $p \triangleq \Pr[\text{a site is occupied}]$ is bounded by:

$$p > \Pr[E_{ss}]^{\delta^2} = \left[\sum_{i=J}^{\infty} e^{-\lambda_m r_m/2} \frac{(\lambda_m r_m/2)^i}{i!} \right]^{\delta^2},$$

according to the fact that all the nodes in a sub-square are blocked if E_{ss} happens for that sub-square and the number of nodes in (disjoint) sub-squares are independent. According to the site percolation model there exists a critical probability p'' such that if $p > p''$ the grid $\mathcal{G}_\delta$ is percolated, i.e., there exists an occupied component, composed of occupied sites, with an infinite size. Any nodes located in the occupied component are totally blocked. Once the grid is percolated, each vacant component, corresponding to an unblocked region, is bounded in size. Since the side length of a site is larger than r_n, nodes located in different vacant components are out of the transmission range of each other. Thus the network is disconnected. $\bar{\lambda}_m^+$ is achieved by solving the function

$$\left[\sum_{i=J}^{\infty} e^{-\bar{\lambda}_m^+ r_m/2} \frac{(\bar{\lambda}_m^+ r_m/2)^i}{i!}\right]^{\delta^2} = p''.$$

7.4 Multi-hop Delay

Our connectivity study demonstrates the jamming resistance of UFH and offers insights on the maximal distance a message can traverse on the network. Our next concern is the multi-hop broadcast delay of a message, which reveals the performance of UFH. The metric of our interest is normalized message propagation delay, defined as

$$\lim_{d(u,v)\to\infty} \frac{t(u,v)}{d(u,v)},$$

where $d(u,v)$ is the Euclidean distance between an arbitrary source node u and a destination node v and $t(u,v)$ is the *minimum* message delay from u and v. Obviously when either u or v is located outside the giant component $\mathcal{C}(G(V_n, E_{\lambda_n,\lambda_m}))$, this metric is trivially infinite. Therefore we focus on the case $u, v \in \mathcal{C}(G(V_n, E_{\lambda_n}, \lambda_m))$ (when $\lambda_m < \lambda_{m,u}$).

In this work $L1$ can employ any communication protocol of interest, and is not the focus in our study. Our emphasis is on the performance of UFH. Thus, we study a simplified yet more powerful jamming model: the jamming region of each jammer covers the whole space S_n and the number of jammers $m = J - 1$ such that each channel over the space is jammed with probability $p_j' < (J-1)p_j < 1$. Note that the simplified jamming model does not change network connectivity. Thus, under the simplified jamming model and with $\lambda_n > \lambda_{n,c}$, the giant component in the network is the same as $\mathcal{C}(G(V_n, E_{\lambda_n}))$, which is denoted by $\mathcal{C}(G^s(V_n, E_{\lambda_n}))$ for distinction. In contrast with the giant component $\mathcal{C}(G(V_n, E_{\lambda_n,\lambda_m}))$, which is possibly composed of $L1$ and $L2$ links, the infinite component $\mathcal{C}(G^s(V_n, E_{\lambda_n}))$ is only composed of $L2$ links, i.e., UFH is applied to all the links belonging to E_{λ_n}. It is straightforward to see that the broadcast delay in $\mathcal{C}(G^s(V_n, E_{\lambda_n}))$ serves as an upper bound for that in $\mathcal{C}(G(V_n, E_{\lambda_n,\lambda_m}))$.

Theorem 15. *If* $u, v \in \mathcal{C}(G^s(V_n, E_{\lambda_n}))$, $\exists \delta \in (1/p_c, \infty)$ *such that*

$$\lim_{d(u,v)\to\infty} \frac{t(u,v)}{d(u,v)} = \delta,$$

where $p_c = 1 - p'_j$.

Note that higher δ implies worse performance.

Proof. The main technical tool we use to prove this theorem is Subadditive Ergodic Theorem given in Appendix D.11. First we show the existence of δ and its upper bound by studying an explicit broadcast protocol.

The simplest broadcast protocol is message flooding, where any node acts as a relay after receiving the message correctly. However, as discussed in Chapter 6, without constraining the number of relays for a receiver, the performance of UFH may degrade significantly. On the other hand, a small number of relays may not cover all the nodes in $\mathcal{C}(G^s(V_n, E_{\lambda_n}))$. Therefore to balance the interference and coverage caused by relays we consider a broadcast protocol based on a dominating set[4]. In particular, the network is divided into clusters with each containing one cluster head. The cluster heads are selected from the network to form a dominating set. It is well known that searching a minimal dominating set is NP-hard problem. Instead we use a suboptimal approach proposed in [74] to generate the clusters, which is distributed, easy to implement and guarantees that cluster heads are separated by distance at least r_n. Two clusters c_1 and c_2 are *neighbors* if there exist nodes x and y in c_1 and c_2 respectively with $||X_x - X_y|| \leq r_n$, and x, y can serve as *relay nodes* for both clusters. If multiple pairs of such nodes are available, one pair is randomly chosen. Only cluster heads and relays can help relay messages and all the others serve as receivers. To broadcast a message over the network, the source node first transmits the message to its cluster head if it's not, and ACK is required for notification of successful reception. And then the message is propagated by cluster heads and relays over the whole network. The protocol interference model is considered: node i receives packets correctly from node j if 1) i and j stay at the same channel; 2) $|X_i - X_j| \leq r_n$; 3) for any concurrent transmitter k in the same channel as i and j, $|X_k - X_i| > (1 + \Delta)r_n$, where constant $\Delta > 0$. Note that the transmitters are not necessarily time synchronized.

Lemma 11. *There are at most* $192(2.5 + \Delta)^2$ *relays within* $(1 + \Delta)r_n$ *distance from a node.*

The proof is given in Appendix D.12.

Lemma 12. *In our protocol one hop delay is finite.*

[4]A dominating set for a graph $G = (V, E)$ is a subset D of V such that every vertex not in D is joined to at least one member of D by some edge.

The proof is given in Appendix D.13. And this lemma also implies that the maximal average one hop delay $t_m < \infty$.

Next, we set $r_n = 1$ without loss of generality. Consider u as the original and the line connecting u and v as the X-axis. Define an auxiliary node w_i for every integer i:

$$w_i \triangleq \arg_{w \in \mathcal{C}(G^s(V_n, E_{\lambda_n}))} \min d(w, (i, 0)).$$

Let k be the closest integer to v and notice that $(k-1) \leq d(u,v) \leq (k+1)$, then,

$$\frac{t(w_0, w_k) - t(w_k, v)}{k+1} \leq \frac{t(u,v)}{d(u,v)} \leq \frac{t(w_0, w_k) + t(w_k, v)}{k-1}.$$

If $v = w_k$ $t(w_k, v) = 0$; otherwise since $d(w_k, v) \leq d(w_k, k) + d(k, v) \leq 1/2 + 1/2 = 1$, $t(w_k, v)$ is the one hop delay and is finite. Thus,

$$\lim_{d(u,v)\to\infty} \frac{t(u,v)}{d(u,v)} = \lim_{k\to\infty} \frac{T_{0,k}}{k},$$

where $T_{l,k} = t(w_l, w_k)$. To prove the theorem it suffices to show that

$$\lim_{k\to\infty} \frac{T_{0,k}}{k} = \delta.$$

In the following we will show that $T_{l,k} (\triangleq t(w_l, w_k))$ satisfies all the conditions required by the Subadditive Ergodic Theorem.

Condition (1), (2) and (3) are easily checked according to stationarity of P.P.P. $\mathcal{X}_n$. Condition 4) is confirmed through the following lemmas.

Lemma 13.

$$d_i = d(w_i, i) < \infty$$

The proof is given in Appendix D.14.

Let $L(w_l, w_k)$ be the shortest path from w_l to w_k and let $M(w_l, w_k)$ be the path from w_l to w_k with minimal delay. $L(w_l, w_k)$ and $M(w_l, w_k)$ do not necessarily coincide. Denote by $|L(w_l, w_k)|$ ($|M(w_l, w_k)|$) the number of hops of $L(w_l, w_k)$ ($M(w_l, w_k)$), and by $T^L_{l,k}$ the multi-hop delay along $L(w_l, w_k)$.

Lemma 14. $|L(w_l, w_k)| < \infty$

The proof follows the same line of that of Lemma 9 in [70].

Then, $\forall k \geq 1$

$$E(T_{0,k}) \leq E(T^L_{0,k}) \leq t_m E(|L(w_0, w_k)|) < \infty.$$

Condition (4) is then checked.

To check the condition (5) we show that $\{T_{(q-1)i,qi} : q \geq 1\}$ is mixing [5], which implies its ergodicity. Construct two circles C_q and C_{q+j} centered at $w_{(q-1)i}$ and $w_{(q-1+j)i}$ with radius $|M(w_{(q-1)i}, w_{qi})|$ and $|M(w_{(q-1+j)i}, w_{(q+j)i})|$. In the verification of condition (4) we know that $T_{(q-1)i,qi} \leq T^L_{(q-1)i,qi}$ and $|L(w_{(q-1)i}, w_{qi})| < \infty$, which implies that $|M(w_{(q-1)i}, w_{qi})| < \infty$ and $|M(w_{(q-1+j)i}, w_{(q+j)i})| < \infty$. Then, when $j \to \infty$, C_q and C_{q+j} do not overlap, which indicates that two paths $M(w_{(q-1)i}, w_{qi})$ and $M(w_{(q-1+j)i}, w_{(q+j)i})$ do not share common nodes. a.s. As a matter of fact, $T_{(q-1)i,qi}$ and $T_{(q-1+j)i,(q+j)i}$ are independent of each other as $j \to \infty$. The mixing property of $\{T_{(q-1)i,qi} : q \geq 1\}$ follows according to the fact that

$$\begin{aligned}&\lim_{j\to\infty} \Pr[T_{(q-1)i,qi} < t) \cap T_{(q-1+j)i,(q+j)i} < t'] \\ &= \Pr[T_{(q-1)i,qi} < t)] \Pr[T_{(q-1+j)i,(q+j)i} < t'].\end{aligned}$$

Since all the conditions in Subadditive Ergodic Theorem are satisfied

$$\exists \delta \geq 0, \lim_{k\to\infty} = \frac{T_{0,k}}{k} = \delta.$$

Then,

$$\delta = \inf_{k\geq 1} \frac{E(T_{0,k})}{k} \leq E(T_{0,1}) < \infty.$$

To derive the lower bound of δ, we jointly consider an ideal multi-hop transmission between u and v and nodes' collaboration[6]. First some relays are placed on the line connecting the source and destination with equal distance r_n. And the number of relays is $l_r = \lceil \frac{d(u,v)}{r_n} \rceil - 1$, which is also the minimal number of relays between u and v. In addition, we assume the time delay for each hop is geometrical distributed with parameter $p_c = 1 - p'_j$. Note that p_c is the upper bound of packet reception rate in UFH for any cooperation schemes[7]. Let $u = v_0$, $v = v_{l_r+1}$ and $v_i,\ i = 1, 2...l_r$ denote the relays between u and v. It's not difficult to check that the time delay $T_{i,j} = t(v_i, v_j)$ satisfies all the conditions in Subadditive Ergodic Theorem. Thus,

$$\lim_{m\to\infty} \frac{T_{0,m}}{m} = \inf_{m\geq 1} \frac{E(T_{0,m})}{m} = 1/p_c.$$

[5] A measure-preserving transformation T is said to be mixing on a probability space $(\Omega, \mathcal{F}, \mu)$ if $\forall A, B \in \mathcal{F}$, $\lim_{k\to\infty} \mu(T^k(A) \cap B) = \mu(A) \cdot \mu(B)$. A sequence $\{X_n\}$ is said to be mixing if the shift on sequence space is mixing. Mixing systems are ergodic [75].

[6] In single channel networks, concurrent transmissions always play a negative role in message broadcast. In contrast, in multichannel networks with UFH, simultaneous transmissions may accelerate message dissemination. Please refer to [76] for details.

[7] For a single UFH link the packet reception rate is $1/C(1 - p')$, where $1/C$ accounts for the probability that the transmitter and receiver agree on the same channel. The best case with nodes' cooperation is that there is a transmitter staying at each channel. Thus, the packet reception rate becomes $1 - p'$.

□

Remark 13. *Assume the minimal propagation delay in L1 link is τ. Following the same line as last theorem, we can show that for $u, v \in \mathcal{C}(G(V_n, E_{\lambda_n}))$ in a jamming-free network, $\exists \delta' \in (\tau, \infty)$ such that $\lim_{d(u,v)\to\infty} \frac{t(u,v)}{d(u,v)} = \delta'$. This indicates that there is no performance loss for UFH, in terms of scaling law.*

7.5 Conclusions

In this chapter we rendered study on the connectivity and multi-hop delay of multi-channel networks under jamming attacks. We showed that the network experiences two phase transitions as the jammer density increases. The existence of weak connection is due to the jamming resistance provided by UFH. In addition we investigated the multi-hop broadcast delay of multi-channel networks when weakly connected. It was shown that the broadcast delay on UFH links scales linearly with the distance between two nodes. This implies that in terms of scaling law, anti-jamming techniques such as UFH do not incur performance loss.

Chapter 8

Summary and Futher Work

8.1 Conclusions

In this dissertation, we have addressed some key issues in spectrum sharing systems as well as jamming-resistant networks, making contributions to both theory and practice of heterogeneous wireless systems. A central theme throughout this dissertation is the exploration of the inherent relationship and tradeoff between configurations of each coexistent system.

Chapter 2 to 5 are focused on spectrum sharing systems. In Chapter 2, we proposed a new spectrum sensing algorithm, an adaptive CUSUM test, in the framework of quickest detection. Through parameters tracking our algorithm can detect distribution change with unknown and time varying parameters, achieving significant performance improvement over the traditional parallel CUSUM test.

Our following efforts were put on the study of network throughput. We first studied the throughput scaling of cognitive radio networks in Chapter 3, with a weaker and yet more practical scenario where the densities of both primary and secondary networks are at the same order. Our results show that the secondary network indeed performs as well as the primary one, in terms of scaling law, when their density ratio is larger than a threshold. In addition, secondary network inevitably suffers from a non-varnishing service outage. Asymptotic analysis on the scaling law only characterizes the (rough) relationship between capacity and network size, neglecting the effect of many important system parameters. In order to discover the inherent tradeoffs among the key system parameters, we investigated the transmission throughput of a spectrum sharing system in Chapter 4. We quantitatively characterized appropriate secondary settings and revealed a sufficient condition for the secondary network configuration such that the overall (transmission) throughput is boosted over that of the stand-alone primary network. Transmission capacity or throughput admits quantitative system analysis, but leaves out the consideration and optimization of transmission distance, a key parameter for wireless networks.

To offer a comprehensive view of network throughput, in Chapter 5 we further studied a new metric: single hop transport throughput (STT), which quantifies the total number of one-hop *reliable* transmissions in a unit area, weighted by corresponding transmission rates and distances. The STT of the secondary network was explored with outage constraints imposed on both networks. The limit of STT, single hop transport capacity (STC), was derived, and STT with secondary receivers randomly located in the field of interest was investigated.

In the subsequent chapters, we studied some vital issues in networks under jamming attacks. In Chapter 6, the communication efficiency of Uncoordinated Frequency Hopping (UFH) and collaborative UFH was examined in a large-scale broadcast network. Many insights related to protocol design and optimality have been provided for better understanding of anti-jamming broadcast in multichannel networks. We showed that the UFH-based protocols scale logarithmically with network size and CUFH achieves the optimal cooperation gain asymptotically. We have also proposed a suboptimal protocol CUFH-p, which simplifies the implementation of CUFH and still outperforms UFH significantly. Our analytical results are well supported by simulation, and may find broader applications in emerging networks and systems. In Chapter 7, we explored the connectivity and multi-hop delay of jamming-resistant networks. It was shown that such a network experiences two phase transitions as the jammer density increases. In addition there exists two cutting jammer densities beyond which the network is no longer strongly or weakly connected. In addition we investigated the multi-hop broadcast delay of a weakly connected network. It was shown that the broadcast delay on the UFH links scales linearly with the distance between two nodes. This implies that in terms of scaling law, anti-jamming techniques such as UFH do not incur performance loss.

8.2 Achievements

We have tackled some open and interesting problems in heterogeneous systems, and made contributions to two main areas: distributed detection and cross-layer design. Part of this PhD work has been published or submitted for publication [77–87], which includes nine published conference papers, as well as three journal papers under review.

8.3 Future Works

The study in this dissertation can be improved upon in the following directions.

- In Chapter 2, the adaptive CUSUM test was proposed for a single node. It is of interest to study a cooperative scheme where a set of nodes jointly make a decision through collaborations. A preliminary work can be found in [88], where all the distribution parameters are assumed known for simplicity.

- In Chapter 5, the modified transport throughput was studied in a single-hop scenario. Extension to the multi-hop case is an interesting yet challenging problem.
- Cognitive radio networks and jamming-resistant networks admit some similarities in terms of network properties. For example, due to the existence of primary users or jammers, the secondary network or legitimate network may no longer be connected. It is of great interest to generalize both applications to a uniform theoretical framework.

REFERENCES

[1] V. N. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory Probab. its Applic.*, vol. 16, no. 2, 1971.

[2] T. Liggett, "An improved subadditive ergodic theorem," *Annals of Prob.*, vol. 13, 1985.

[3] Federal Communications Commission Spectrum Policy Task Force, "Report of the spectrum efficiency working group," tech. rep., Nov. 2002.

[4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE COMMUNICATIONS SURVEYS & TUTORIALS*, vol. 11, 2009.

[5] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Applications*. Englewood Cliffs: Prentice-Hall, NJ, 1993.

[6] T. Kailath and H. V. Poor, "Detection of stochastic processees," *IEEE Trans. on Inform. Theory*, vol. 44, p. 2230, 1998.

[7] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, p. 100, 1954.

[8] G. Lorden, "Procedures for reacting to a change in distributions," *Ann. Math. Statist.*, vol. 42, p. 1897, 1971.

[9] R. J. Stahl and P. K. Willett, "Quickest detection of a tonal burst," *IEEE Trans. on Signal Processing*, vol. 45, p. 2037, 1997.

[10] I. V. Nikiforov, "Suboptimal quadratic change detection scheme," *Trans. Inform. Theory*, vol. 46, p. 2095, 2000.

[11] M. Vu, N. Devroye, M. Sharif, and V. Tarokh, "Scaling laws of cognitive networks," in *Proc. Crowncom*, Aug. 2007.

[12] S. Jeon, N. Devroye, M. Vu, S. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," *IEEE Trans. Info. Theory*, 2009. Submitted.

[13] C. Yin, L. Gao, and S. Cui, "Scaling laws for overlaid wireless networks: A cognitive radio network vs. a primary network," *IEEE Trans. Networking*, vol. 18, Agu. 2010.

[14] K. Huang, V. K. N. Lau, and Y. Chen, "Spectrum sharing between celluar and mobile ad hoc networks: transmission capacity trade off," *IEEE J. Sel. areas in Commun., Special Issue on Stochastic Geometry and Random Graphs for Wireless Netorks*, Sep. 2009.

[15] C. Lee and M. haenggi, "Interference and outage in doubly poisson cognitive networks," in *ISIT*, 2010.

[16] C. Yin, L. Gao, T. Liu, and S. Cui, "Transmission capacities for overlaid wireless ad hoc networks with outage constraints," in *Proceedings of the IEEE ICC*, June 2009.

[17] J. Lee, S. Lim, J. G. Andrews, and D. Hong, "Achievable transmission capcity of secondary system in cognitive radio networks," in *ICC*, 2010.

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

[19] S. Weber, X. Yang, J. G. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraint," *IEEE Trans. Info. theory*, vol. 51, pp. 4091–4102, Dec. 2005.

[20] S. Weber, J. G. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Trans. Comm.*, 2010. To appear.

[21] C. Popper, M. Strasser, and S. Capkun, "Anti-jamming broadcast communication using uncoordinated spread spectrum techniques," *IEEE Journal on Selected Areas in Communications*, 2010.

[22] M. Strasser, S. Capkun, C. Popper, and M. Cagalj, "Jammming-resistant keys establishment using uncoordinated frequency hopping," in *Proc. IEEE symposium on security and privacy*, 2008.

[23] M. Strasser, C. Popper, and S. Capkun, "Efficient uncoordinated FHSS anti-jamming communication," in *Proc. Mobihoc*, 2009.

[24] L. Xiao, H. Dai, and P. Ning, "Jamming-resistant collaborative broadcast using uncoordinated frequency hopping," *IEEE Transactions on Information Forensics and Security*, 2011. to appear.

[25] H. Li, C. Li, and H. Dai, "Quickest spectrum sensing in cognitive radio," in *Proc. of CISS'08*, p. 203, Mar. 2008.

[26] A. N. Shiryaev, "The problem of quickest detection of a viola- tion of stationary behavior," *Dokl. Akad. Nauk SSSR*, vol. 138, p. 1039, 1961.

[27] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *Trans. Inform. Theory*, vol. 44, p. 2917, 1998.

[28] A. S.Willsky and H. L. Jones, "A generalized likelihood ratio approach to detection and estimation of jumps in linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, p. 108, 1976.

[29] T. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.

[30] http: //en.wikipedia.org/wiki/Exponential_family.

[31] I. F. Akyildiz, W. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Networks*, vol. 7, pp. 810–836, July 2009.

[32] S. Jeon, N. Devroye, M. Vu, S. Chung, and V. Tarokh, "Cognitive networks achieve throughput scaling of a homogeneous network," *IEEE Trans. Info. Theory*, 2009. Submitted.

[33] O. Dousse and P. Thiran, "Connectivity vs capacity in dense ad hoc networks," in *Proc. IEEE INFOCOM*, 2004.

[34] M. Franceschetti, O. Dousse, D. N. C. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Trans. Info. Theory*, pp. 1009–1018, Mar. 2007.

[35] A. Ozgur, O. leveque, and D. N. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. Inform. Theory*, vol. 53, Oct. 2007.

[36] L.-L. Xie and P. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation," *IEEE Trans. Infor. Theory*, vol. 50, pp. 748–767, May 2004.

[37] A. Jovicic, P. Viswanath, and S. R. Kulkarni, "Upper bunds to transport capacity of wireless networks," *IEEE Trans. Infor. Theory*, vol. 50, 2004.

[38] F. Xue, L.-L. Xie, and P. R. Kumar, "The transport capacity of wireless networks over fading channels," *IEEE Trans. Infor. Theory*, vol. 51, pp. 834–847, May 2005.

[39] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Infor. Theory*, vol. 52, pp. 1813–1827, May 2006.

[40] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Infor. Theory*, vol. 55, pp. 3945–3958, Sep. 2006.

[41] C. Yin, C. Chen, T. Liu, and S. Cui, "Generalized results of transmission capacities for overlaid wireless networks," in *IEEE International Symposium on Information Theory (ISIT)*, 2009.

[42] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Chichester: John Wiley & Sons, 1995.

[43] J. G. Andrews, S. Weber, M. Kountouris, and M. Haenggi, "Random access transport capacity," *IEEE Transactions on Wireless Communications*, vol. 9, Jun. 2010.

[44] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An aloha protocol for multiphop mobile wireless network," *IEEE Trans. Inform. Theory*, vol. 5, Feb. 2006.

[45] J. F. C. Kingman, *Poisson processes*. Oxford university press, 1996.

[46] C. Li and H. Dai, "Transport capacity and connectivity of cognitive networks with outage constraint," in *ISIT*, 2010.

[47] M. Franceschetti and R. Meester, *Random Networks for Communication*. Cambridege university press, 2007.

[48] C. Popper, M. Strasser, and S. Capkun, "Jamming-resistant broadcast communication without shared key," in *Proc. USENIX Security symposium*, 2009.

[49] L. Baird, W. Bahn, M. Collins, M. Carlisle, and S. Butler, "Keyless jam resistance," in *Proc. IEEE Information Assurance and Security Workshop*, 2007.

[50] Y. Liu, P. Ning, H. Dai, and A. Liu, "Randomized differential DSSS: jamming-resistant wireless boradcast commuincation," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2010.

[51] Q. Wang, P. Xu, K. Ren, and X. Li, "Delay-bounded apdaptive UFH-based anti-jamming wireless communication," in *IEEE INFOCOM*, Apr. 2011.

[52] A. Liu, P. Ning, H. Dai, Y. Liu, and C. Wang, "USD-FH: Jamming-resistant wireless communication using frequency hopping with uncoordinated seed disclosure," in *Proc. 7th IEEE international conference on mobile ad-hoc and sensor systems*, 2010.

[53] W.Xu, T. Wood, and Y. Zhang, "Channel surfing and spatial retreats," in *Proc. of workshop on Wireless Security*, 2004.

[54] G. Alnifie and R. Simon, "A multi-channel defense against jamming in wireless sensor networks," in *Proc. of Q2SWinet'07*, 2007.

[55] B. Awerbuch, A. Richa, and C. Scheideler, "A jamming-resistant MAC protocol for single-hop wireless netowrks," in *Proceding of PODC*, 2008.

[56] A. Richa, C. Scheideler, S. Schmid, and J. Zhang, "A jamming-resistant MAC protocol for multi-hop wireless networks," in *the 24th International Symposium on Distributed Computing (DISC)*, 2010.

[57] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attacks and network defense policies in wireless sensor networks," in *Proc. of IEEE INFOCOM*, 2007.

[58] K. Panyim, T. Hayajneh, P. Krishnamurthy, and D. Tipper, "On limited-range strategic/random jamming attacks in wireless ad hoc networks," in *the 5th LCN Workshop on Security in Communications Networks*, Oct. 2009.

[59] S. Dolev, S. Gilbert, R. Guerraoui, Calvin, and Newport, "Gossiping in a multi-channel radio network (An oblivious approach to coping with malicious interference)," in *the 21st International Symposium on Distributed Computing (DISC'07)*, pp. 208–222, 2007.

[60] S. Dolev, S. Gilbert, R. Guerraoui, and C. Newport, "Secure communication over radio channels," in *Proc. ACM PODC*, 2008.

[61] S. Gilbert, R. Guerraoui, D. Kowalski, and C. Newport, "Interferenceresilient information exchange," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2009.

[62] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using channel hopping to increase 802.11 resilience to jamming attacks," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2007.

[63] A. D. Wood, J. A. Stankovic, and G. Zhou, "Deejam: defeating energyefficient jamming in IEEE 802.15.4-based wireless networks," in *Proc. Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 2007.

[64] H. Li and Z. Han, "Dogfight in spectrum: jamming and anti-jamming in multichannel cognitive radio systems," in *IEEE Globecom 2009*, Dec. 2009.

[65] Y. Wu, B. Wang, and K. J. R. Liu, "Optimal defense against jamming attacks in cognitive radio networks using the Markov decision process approach," in *IEEE Globecom*, 2010.

[66] J. Havil and J. Gamma, *Exploring Euler's Constant.* Princeton, NJ: Princeton University Press, 2003.

[67] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series and products, seventh edition.* Elsevier Academic Press, 2007.

[68] P. Gupta and P. R. Kumar, "Critical power for asymptotic connecivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications: A Vlume in Honor of W. H. Fleming. Edited by W.M. McEneany, G. Yin, And Q. Zhang*, pp. 547–566, 1998.

[69] O. Dousse, F. Baccelli, and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," in *Proc. IEEE INFOCOM*, Apr. 2003.

[70] Z.N.Kong and E. M. Yeh, "connectivity, percolation, and information dissemination in large-scale wireless networks with dynamic links," *IEEE Trans. Infor. Theory.* submitted.

[71] W. Ren, Q. Zhao, and A. Swami, "Connectivity of heterogeneous wireless networks," *IEEE Trans. Info. theory*, 2009. accepted.

[72] W. Ren, Q. Zhao, and A. Swami, "On the connectivity and multihop delay of ad hoc cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, 2011.

[73] R. Meester and R. Roy, *Continuum Percolation.* Cambridge, U. K.: Cambridge Univ. Press, 1996.

[74] W. Li and H. Dai, "Cluster-based distributed consensus," *IEEE Trans. Wireless Communications*, vol. 8, no. 1, 2009.

[75] K. Petersen, *Ergodic Theory.* New York: Cambridge Univ. Press, 1989.

[76] L. Xiao, H. Dai, and P. Ning, "Jamming-resistant collaborative broadcast using frequency hopping," *IEEE Trans. Information Forensics and Security*, 2011. submitted.

[77] C. Li and H. Dai, "Efficient in-network computing with noisy wireless channels," *IEEE Trans. Mobile Computing.* submitted.

[78] C. Li and H. Dai, "Transport throughput of decentralized cognitive radio networks," *IEEE Trans. Mobile Computing.* submitted.

[79] C. Li and H. Dai, "Communication efficiency of anti-jamming broadcast in large-scale multi-channel wireless networks," *IEEE Trans. Signal Processing.* submitted.

[80] C. Li and H. Dai, "On the throughput scaling of cognitive radio ad hoc networks," in *2011 IEEE Conference on Computer Communications (INFOCOM), Shanghai, China*, Apr. 2011.

[81] C. Li and H. Dai, "Transport throughput of decentralized cognitive radio networks," in *2011 IEEE Conference on Computer Communications (INFOCOM), Shanghai, China*, Apr. 2011.

[82] C. Li, H. Dai, L. Xiao, and P. Ning, "Analysis and optimization on jamming-resistant collaborative broadcast in large-scale networks," in *2010 Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA*, Nov. 2010.

[83] C. Li and H. Dai, "Transport capacity and connectivity of cognitive radio networks with outage constraint," in *2010 IEEE International Symposium on Information Theory (ISIT), Austin, Texas*, Jun. 2010.

[84] C. Li and H. Dai, "Towards efficient designs for in-network computing with noisy wireless channels," in *2010 IEEE Conference on Computer Communications (INFOCOM), San Diego, CA*, Mar. 2010.

[85] C. Li, H. Dai, and H. Li, "Adaptive quickest change detection with unknown parameter," in *2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Taipei*, Apr. 2009.

[86] C. Li, H. Dai, and H. Li, "Finding the k largest metrics in a noisy broadcast network," in *2008 Allerton Conference on Communication, Control and Computing, Monticello, IL*, Sept. 2008.

[87] C. Li, H. Li, and H. Dai, "Collaborative quickest detection in ad hoc networks with delay constraint - part II: Multi-node network," in *2008 Conference on Information Sciences and Systems (CISS), Princeton, NJ*, 2008.

[88] H. Li, H. Dai, and C. Li, "Collaborative quickest spectrum sensing via random broadcast in cognitive radio systems," in *2009 IEEE Global Communications Conference (GLOBECOM), Honolulu, Hawaii*, Nov. 2009.

APPENDICES

Appendix A

On the Throughput Scaling of Cognitive Radio Ad Hoc Networks

A.1 Discussion on dense network model

In this section we revisit the dense network model used in [32] with our weaker condition $m/n =$ constant, and show that full connectivity of the secondary network is no longer maintained.

As in [32], we assume that the primary and secondary network are distributed in a *unit* square, according to a P.P.P with density n and m, respectively. The square is further divided into secondary cells with side length $d = r_m/\sqrt{5}$, where $r_m = k\sqrt{\log m/m}$ for some k is the transmission range of the secondary network, which guarantees full connectivity of the secondary network if it were stand-alone. To restrict the interference from the secondary network to the primary network, a preservation region is set around each primary node, which is defined as a square composed of 9 secondary cells (with the center cell containing the primary node). Any secondary nodes located in the preservation regions have to keep silent all the time. Note that full connectivity is still guaranteed a.s. for the secondary network with $m = n^\beta$ ($\beta > 1$) since

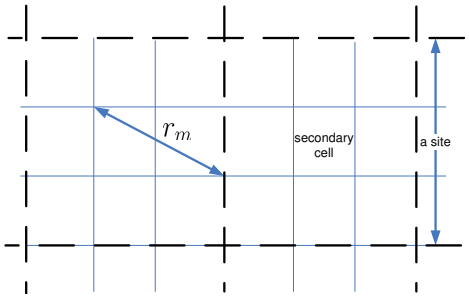

Figure A.1: A grid with each site containing 9 secondary cells

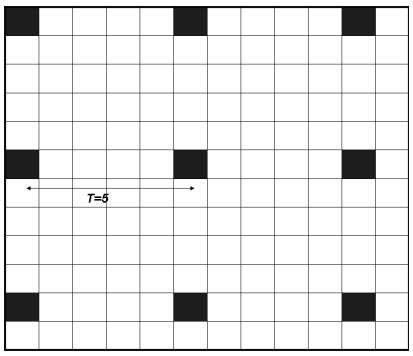

Figure A.2: M-TDMA scheduling scheme for $T = 5$.

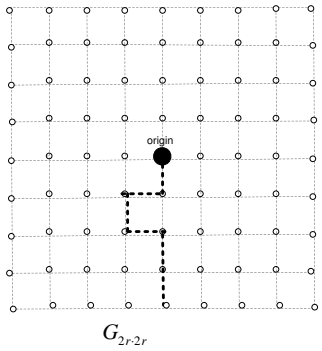

Figure A.3: Construction of $G_{2r\cdot 2r}$

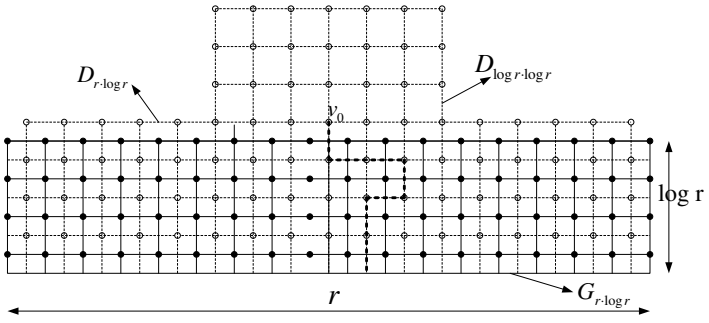

Figure A.4: Construction of $G_{r\cdot\log r}$, $D_{r\cdot\log r}$ and $D_{\log r\cdot\log r}$

clustering of preservation regions can be circumvented [32]. However it does not hold any more with $m/n =$ constant, which is the key assumption in our work.

Our arguments are based on the discrete site percolation model. We map the tesselated square into a grid, where each site is composed of 9 secondary cells (see Fig. A.1, where a site is indicated by dashed lines and a secondary cell is indicated by solid lines). A site is considered occupied if the secondary cell in its center contains at least one primary node. Otherwise it is vacant. Note that an occupied site corresponds to a preservation region, thus any secondary nodes in occupied sites are deactivated. The probability that a site is occupied is given by:

$$p_s = 1 - \exp(-n\pi k\frac{\log m}{5m}) = 1 - \exp(-\theta \log m),$$

where $\theta = \pi k\frac{n}{5m}$ is a constant. According to the discrete site percolation model [47], there exists a critical probability p_c such that if $p_s > p_c$, the grid is percolated, i.e., there exists an infinite occupied components a.s., and each vacant component is bounded in size. Since p_s increases with $\log m$, there exists an M_0 such that if $m > M_0$, $p_s > p_c$, which indicates that the grid is percolated a.s.. In addition, nodes in separate vacant components are beyond transmission range of each other since the side length of a site $\frac{3r_m}{\sqrt{5}} > r_m$. Our conclusion then follows from the fact that only nodes in the same vacant component can communicate with each other and vacant components are bounded in size.

A.2 M-TDMA scheme

To limit the interference caused by simultaneous transmissions, the $M(= T^2)$ TDMA cell scheduling scheme [18] is often adopted, where T is an integer. Each time slot is assigned to one of M cells in a $T \times T$ supper cell and the cells assigned in the same slot are at least T cells away from each other. Fig. A.2 depicts the 25-TDMA scheduling scheme, where nodes in the shaded cells transmit simultaneously. The protocol is executed in active cells and nodes in each active cell take turns to transmit their information.

A.3 Proof of theorem 3

Assume all the grids denoted by G below are constructed from the secondary network model. Note that the states of any neighboring nodes are correlated and two edges with their Manhattan distance [1] no smaller than 3 are independent. Denote by $G_{q\cdot s}$ a 2-D grid with dimension $q \times s$. Let $0 \leftrightarrow \partial G_{2r\cdot 2r}$ be the event that there is an open path connecting the center of the grid $G_{2r\cdot 2r}$

[1] the minimal number of adjacent edges that must be traversed to connect two edges

to its boundary (see Fig. A.3). Since the length of such paths is at least r Thus, we have

$$\begin{aligned} P(0 \leftrightarrow \partial G_{2r\cdot 2r}) &\leq P(N_r \geq 1) \\ &\leq p^{r/4}\delta(r) \\ &\leq p^{r/4} 4 \cdot 3^{r-1} \\ &= \frac{4}{3}(81p)^{r/4}, \end{aligned} \tag{A.1}$$

where N_r is the number of open paths of length r and $\delta(r)$ is the number of paths of length r, all starting at the center. According to the grid construction, any two edges along a path are independent if their Manhattan distance on the path is no smaller than 3. Therefore paths of length r are open with probability at most $p^{r/4}$, which results in the second inequality. The third inequality is due to the fact[2] that $\delta(r) \leq 4 \cdot 3^{r-1}$.

Now, let's consider a grid $G_{r\cdot k_s \log r}$ given in Fig. A.4, corresponding to a corridor in the secondary network. We construct a dual of the random grid by drawing an edge in its dual lattice [3] $D_{r\cdot k_s \log r}$, if it does not cross an edge of the original random grid, and deleting it otherwise. Obviously the edge in $D_{r\cdot k_s \log r}$ is open with probability $1-p$. Let v_0 be a vertex at the top side of $D_{r\cdot k_s \log r}$ which satisfies

$$P(E_{v_0}) \geq \frac{1}{r+1} P(D^{TB}_{r\cdot k_s \log r}), \tag{A.2}$$

where E_{v_0} denotes the event that there is a top to bottom crossing starting from v_0 in $D_{r\cdot k_s \log r}$ and $D^{TB}_{r\cdot k_s \log r}$ is the event that there is a top to bottom crossing in $D_{r\cdot k_s \log r}$. Then a grid $D_{2k_s \log r \cdot 2k_s \log r}$ centered at v_0 is constructed. Therefore,

$$\begin{aligned} P(D^{TB}_{r\cdot k_s \log r}) &\leq (r+1) P(E_{v_0}) \\ &\leq (r+1) P(0 \leftrightarrow \partial D_{2k_s \log r \cdot 2k_s \log r}) \\ &\leq \frac{4}{3}(r+1)(81(1-p))^{k_s \log r/2}. \end{aligned} \tag{A.3}$$

where the last inequality is due to Eq. (A.1).

Let $G^{LR}_{r\cdot k_s \log r}$ be the event that there is a left to right crossing in the grid $G_{r\cdot k_s \log r}$. According to the properties of a dual grid,

$$\begin{aligned} \Pr(G^{LR}_{r\cdot k_s \log r}) &= 1 - P(D^{TB}_{r\cdot k_s \log r}) \\ &\geq 1 - \frac{4}{3}(r+1)(81(1-p))^{k_s \log r/2}. \end{aligned}$$

[2]The origin has 4 edges to choose and then each consecutive vertex along a potential path has 3 edges to choose.

[3]The dual lattice is attained by shifting the original grid above by 1/2 side length.

The above equation indicates that there exists an open LR crossing in $G_{r\cdot k_s \log r}$ with positive probability. With the help of the following lemma (lemma 4.3.7 in [47] with paraphrase in our setting), a stronger conclusion can be reached that there exist $\delta_s \log r$ open LR paths.

Lemma 15. *Let a random grid G_1 constructed from a secondary network with size m_1. Denote by A an increasing event, and let $I_e(A)$ the event defined by the set of configurations in A for which A remains true even if we change the states of up to e arbitrary edges. Then construct another random grid G_2 from a secondary network with size $m_2 > m_1$, we have*

$$1 - P_{p_2}(I_e(A)) \leq (\frac{p_2}{p_2 - p_1})(1 - P_{p_1}(A)),$$

where p_1 and p_2 are the probabilities that an edge is open in G_1 and G_2, respectively.

Then denote by $N_c(i)$ the maximal number of paths in the i corridor. According to the lemma above,

$$\begin{aligned}\Pr(N_c(i) < \delta_s \log r) &\leq (\frac{p}{p-p'})^{\delta_s \log r}(1 - P_{p'}(G^{LR}_{r\cdot k_s \log r})) \\ &= (\frac{p}{p-p'})^{\delta_s \log r}\frac{4}{3}(r+1)(81(1-p'))^{k_s \log r/2}\end{aligned}$$

for any $80/81 < p' < p$.

Let $N_n = \min_i N_c(i)$.

$$\begin{aligned} & P(N_n < \delta_s \log r) \\ < \quad & \frac{r}{k_s \log r} P(N_c(i) < \delta_s \log r) \\ < \quad & \frac{r}{k_s \log r}\left((\frac{p}{p-p'})^{\delta_s \log r}\frac{4}{3}(r+1)(81(1-p'))^{k_s \log r/2}\right) \\ = \quad & \frac{4r(r+1)}{3k_s \log r} e^{\log r(\delta_s \log(p/(p-p'))+k_s/2 \log(81(1-p')))} \\ \to \quad & 0 \end{aligned}$$

when k_s and δ_s are chosen such that

$$\delta_s \log(p/(p-p')) + k_s/2 \log(81(1-p')) \leq -2.$$

A.4 The Vapnik-Chervonenkis Theorem [1]

If $\mathcal{C}$ is a set of finite VC-dimension[4] $d_{vc}(\mathcal{C})$, and $\{X_i | i = 1, 2, ..., N\}$ is a sequence of i.i.d. random variables with a common probability distribution P, then for every $\epsilon, \delta > 0$,

$$\Pr\left(\sup_{C\in\mathcal{C}}\left|\frac{\sum_{i=1}^{N} I(X_i \in C)}{N} - P(C)\right| \leq \epsilon\right) > 1 - \delta,$$

whenever $N > \max[\frac{8d_{vc}(\mathcal{C})}{\epsilon}\log\frac{16e}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}]$.

A.5 Proof of Lemma 1

Let $A = a_s^2$, $B = \sqrt{n}h$, m_c and m_s denote the number of nodes in a secondary cell and slab. Then, according to the Vapnik-Chervonenkis Theorem above

$$\Pr\left(\sup_{A\in S_n}\left|\frac{m_c}{m} - \frac{a_s^2}{n}\right| \leq \epsilon\right) > 1 - \delta,$$
$$\Pr\left(\sup_{B\in S_n}\left|\frac{m_s}{m} - \frac{\sqrt{n}h}{n}\right| \leq \epsilon\right) > 1 - \delta,$$

whenever $m > \max[\frac{24}{\epsilon}\log\frac{13}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}]$, where $d_{vc} = 3$ for the tesselated network, and $\epsilon = \delta = \frac{\log m - ma^2/n}{m}$.

Thus,

$$m_c \leq \log m, \text{ w.h.p. },$$

and

$$m_s \leq m\frac{\sqrt{n}h}{n} + \frac{\log m - ma^2/n}{m} < 2m\frac{h}{\sqrt{n}}, \text{ w.h.p. }.$$

A.6 Proof of Lemma 2

The secondary network schedules its transmissions on the highway according to the 9-TDMA scheme. For the ith active secondary TX-RX pair, the secondary receiver $X_{s,D(i)}$ is at most $2\sqrt{2}a_s$ distance away from its corresponding transmitter $X_{s,i}$, so the SINR at the secondary receiver is given by:

$$SINR_s = \frac{P_s l(\parallel X_{s,i} - X_{s,D(i)} \parallel)}{N_0 + I_s + I_{ps}}. \tag{A.4}$$

[4]Let $\mathcal{C}$ be a set of subsets. A finite set of points A is said to be shattered by $\mathcal{C}$ if for every subset B of A there is a set $C \in \mathcal{C}$ such that $A \cap C = B$. The VC-dimension of $\mathcal{C}$, denoted by $d_{vc}(\mathcal{C})$, is defined as the supremum of the sizes of all finite sets that can be shattered by $\mathcal{C}$.

There exist 8 first-tier secondary interfering cells at least a_s away from the receiver and 16 secondary-tier interfering cells at least $3a_s$ away. In general there are $8t$ tth-tier secondary interfering cells with a distance at least $(2t-1)a_s$. Thus, the interference power from all the other active secondary transmitters is bounded above by:

$$\begin{aligned} I_s &\leq \sum_{k=1,k\neq i}^{\infty} P_s l(\parallel X_{s,k} - X_{s,D(i)} \parallel) \\ &< \sum_{t=1}^{\infty} P_s 8t \min(1, ((2t-1)a_s)^{-\alpha}) \\ &< \sum_{t=1}^{q_s} P_s 8t + P \sum_{t=q_s+1}^{\infty} 8t(2t-1)^{-\alpha} \\ &= C_{I_s}, \end{aligned} \tag{A.5}$$

where q_s is the largest integer such that $(2t-1)a_s < 1$ and the sum $\sum_{t=q_s+1}^{\infty} 8t(2t-1)^{-\alpha}$ converges when $\alpha > 2$.

Since any secondary nodes out of preservation regions are at least a_s distance away from primary nodes, we have

$$\begin{aligned} I_{ps} &\leq \sum_{k=1}^{\infty} P_p l(\parallel X_{p,k} - X_{s,D(i)} \parallel) \\ &< \sum_{t=1}^{\infty} P_p 8t \min[1, ((2t-1)a_p)^{-\alpha}] + P_p \min(1, a_s^{-\alpha}) \\ &< \sum_{t=1}^{q_p} P_p 8t + P \sum_{t=q_p+1}^{\infty} 8t(2t-1)^{-\alpha} + P_p \min(1, a_s^{-\alpha}) \\ &= C_{I_{ps}}, \end{aligned} \tag{A.6}$$

where q_p is the largest integer such that $(2t-1)a_p < 1$. $C_{I_{ps}}$ is a constant due to the convergence of $\sum_{t=q_p+1}^{\infty} 8t(2t-1)^{-\alpha}$ for $\alpha > 2$. Thus,

$$\begin{aligned} SINR_s &= \frac{P_s l(\parallel X_{s,i} - X_{s,D(i)} \parallel)}{N_0 + I_s + I_{ps}} \\ &\geq \frac{P_s \min(1, (2\sqrt{2}a_s)^{-\alpha})}{N_0 + C_{I_s} + C_{I_{ps}}} \\ &= C_s, \end{aligned} \tag{A.7}$$

where C_s is a constant.

Therefore, the transmission rate $R_s = 1/9 \log(1 + SINR_s)$ is lower bounded by some con-

stant K_s.

A.7 Proof of Lemma 3

The secondary network schedules its transmissions according to the $\log m$-TDMA scheme during the access and delivery phase. Given the ith active secondary TX-RX pair, the secondary receiver $X_{s,D(i)}$ is at most $\sqrt{2}a_s \log(\sqrt{n}/(\sqrt{2}a_s))$ distance away from its corresponding transmitter $X_{s,i}$. Therefore, following the proof in lemma 2, I_{ps}, the interference power from primary transmitters to $X_{s,D(i)}$, is still bounded by a constant, while I_s, the interference power from concurrent secondary transmitters to $X_{s,D(i)}$, and the desired received power scale as $\Omega(1/(\log m)^\alpha)$. Therefore the $SINR_s$ at $X_{s,D(i)}$ scales as $\Omega(1/(\log m)^\alpha)$, so as the transmission rate.

A.8 Hoeffding inequality

Let X_i $(1 \leq i \leq n)$ be n i.i.d. random variables over the interval $[a, b]$. For the sum of these variables $S = \sum_{i=1}^{n} X_i$, we have the inequity

$$\Pr\left(S - E(S) \geq n\delta\right) \leq e^{-\frac{2n\delta^2}{b-a}}, \quad \delta > 0.$$

A.9 Proof of Lemma 4

Without loss of generality, let us focus on the second routing phase of the secondary network. During the second sub-slot (see Fig. 3.3) both primary and secondary network schedule their transmissions according to 9-TDMA scheme[5]. For the ith active TX-RX pair, the receiver $X_{p,D(i)}$ is at most $2\sqrt{2}a_p$ distance away from the transmitter $X_{p,i}$, i.e., $\parallel X_{p,i} - X_{p,D(i)} \parallel \leq 2\sqrt{2}a_p$, so the SINR at the primary receiver is given by:

$$SINR_p = \frac{P_p l(\parallel X_{p,i} - X_{p,D(i)} \parallel)}{N_0 + I_p + I_{sp}}. \tag{A.8}$$

Similar to the analysis in Appendix A.6, the interference power from all the other active

[5]In the first and third routing phase of the secondary network, the inferference to the primary network is even smaller.

primary transmitters is bounded above by:

$$\begin{aligned} I_p &\leq \sum_{k=1,k\neq i}^{\infty} P_p l(\parallel X_{p,k} - X_{p,D(i)} \parallel) \\ &< \sum_{t=1}^{\infty} P_p 8t \min(1, ((2t-1)a_p)^{-\alpha}) \\ &< \sum_{t=1}^{q_p} P_p 8t + P \sum_{t=q_p+1}^{\infty} 8t(2t-1)^{-\alpha} \\ &= C_{I_p}, \end{aligned} \tag{A.9}$$

where q_p is the largest integer such that $(2t-1)a_p < 1$ and C_{I_p} is a constant since $\sum_{t=q_p+1}^{\infty} 8t(2t-1)^{-\alpha}$ converges for $\alpha > 2$.

Preservation regions keep the secondary transmitters at least a_s away from the primary receiver (see Fig. 3.4). Thus, the interference from the secondary active transmitters is upper bounded by:

$$\begin{aligned} I_{sp} &\leq \sum_{k=1}^{\infty} P_s l(\parallel X_{s,k} - X_{p,D(i)}) \parallel) \\ &< \sum_{t=1}^{\infty} P_s 8t \min(1, ((2t-1)a_s)^{-\alpha}) \\ &< \sum_{t=1}^{q_s} P_s 8t + P \sum_{t=q_s+1}^{\infty} 8t(2t-1)^{-\alpha} \\ &= C_{I_{sp}}, \end{aligned} \tag{A.10}$$

where q_s is the largest integer such that $(2t-1)a_s < 1$ and $C_{I_{sp}}$ is a constant due to the convergence of $\sum_{t=q_s+1}^{\infty} 8t(2t-1)^{-\alpha}$ for $\alpha > 2$.

Thus

$$\begin{aligned} SINR_p &= \frac{P_p l(\parallel X_{p,i} - X_{p,D(i)} \parallel)}{N_0 + I_p + I_{sp}} \\ &\geq \frac{P_p \min(1, (2\sqrt{2}a_p)^{-\alpha})}{N_0 + C_{I_p} + C_{I_{sp}}} \\ &= C_p, \end{aligned} \tag{A.11}$$

where C_p is a constant.

Therefore, the transmission rate $R_p = 1/9 \log(1 + SINR_p)$ is lower bounded by some constant K_p.

Appendix B

Transmission Throughput of Decentralized Overlaid Networks with Outage Constraints

B.1 Derivation of Outage Probability

We are interested in a general result by considering the thermal noise N. Overlaid with the secondary network, the probability of a successful primary transmission, $\Pr(SINR_o(l_o) \geq T_o)$, is given by:

$$\begin{aligned}
&\Pr\left(\left(\frac{P_o l_o^{-\alpha} u}{N + I_o + I_{so}}\right) \geq T_o\right) \\
&= \Pr\left(u \geq \frac{T_o(I_o + I_{so} + N)}{P_o l_o^{-\alpha}}\right) \\
&= \exp(-n_o l_o^{\alpha}) \int_0^{\infty} \exp\left(-\frac{T_o}{P_o l_o^{-\alpha}} i\right) f_I(i) di \\
&= \exp(-n_o l_o^{\alpha}) \psi_I\left(\frac{T_o l_o^{\alpha}}{P_o}\right),
\end{aligned}$$

where $n_o = \frac{NT_o}{P_o}$, $I = I_o + I_{so}$, with pdf f_I, and $\psi_I(i)$ is the Laplace transform of f_I. Due to the independence of I_o and I_{so},

$$\begin{aligned}
\psi_I\left(\frac{T_o l_o^{\alpha}}{P_o}\right) &= \psi_{I_o}(\frac{T_o l_o^{\alpha}}{P_o}) \psi_{I_{so}}(\frac{T_o l_o^{\alpha}}{P_o}) \\
&= \exp\left[-K_\alpha T_o^{2/\alpha} l_o^2 \left(\lambda_{st} \frac{1}{\theta^{2/\alpha}} + \lambda_{ot}\right)\right],
\end{aligned} \tag{B.1}$$

where $\psi_{I_o}(i) = \exp\left[-K_\alpha \lambda_{ot}(P_o i)^{2/\alpha}\right]$, $\psi_{I_{so}}(i) = \exp\left[-K_\alpha \lambda_{st}(P_s i)^{2/\alpha}\right]$ (K_α is defined after (5.6)). Eq. (5.7) is obtained by letting $\Pr(SINR_o(l_o) \geq T_o) \geq 1 - \epsilon_o$.

Following the same line above, the outage probability of the secondary network can be calculated as follows:

$$\begin{aligned} \delta_s(r, \lambda_{st}) &= \Pr\left(\frac{P_s u r^{-\alpha}}{N + I_s + I_{os}} \leq T_s\right) \\ &= 1 - \exp\left[-K_\alpha T_s^{2/\alpha} r^2 \left(\lambda_{ot}\theta^{2/\alpha} + \lambda_{st}\right) - n_s r^\alpha\right], \end{aligned} \tag{B.2}$$

where $n_s = \frac{T_s N}{P_s}$, r is the transmission distance between a secondary transmitter and its corresponding receiver.

Appendix C

Transport Throughput of Secondary Networks in Spectrum Sharing Systems

C.1 Sketch of proof of Lemma 5

From the definition of transport throughput in Eq. (5.5),

$$\begin{aligned} C_s(\lambda_{st}) &= R_s \lambda_{st} E_r[r(1-\delta(r,\lambda_{st}))] \\ &\leq \max_{r_s} C_s(r_s,\lambda_{st}) \triangleq \bar{C}_s(\lambda_{st}), \end{aligned}$$

subject to the outage constraint $\delta_s(r_s,\lambda_{st}) \leq \epsilon_s$, where $C_s(r_s,\lambda_{st}) = R_s\lambda_{st}r_s(1-\delta_s(r_s,\lambda_{st}))$ and $\delta_s(r_s,\lambda_{st}) = 1-e^{-B(\lambda_{st})r_s^2-n_s r_s^\alpha}$ is given in (B.2).

Taking derivative of $C_s(r_s,\lambda_{st})$ w.r.t r_s we have

$$\frac{dC_s(r_s,\lambda_{st})}{dr_s} = \frac{R_s\lambda_{st}(1-2B(\lambda_{st})r_s^2-\alpha n_s r_s^\alpha)}{\exp(B(\lambda_s t)r_s^2+n_s r_s^\alpha)}.$$

Note that $f(r_s) = 1-2B(\lambda_{st})r_s^2-\alpha n_s r_s^\alpha$ is monotonically decreasing. Regardless of the outage constraint $C_s(r_s,\lambda_{st})$ achieves maximum when $f(r_s)=0$, i.e. $r_s = l$. The outage probability at $r_s = l$ is $1-e^{-\frac{1}{2}+(\frac{\alpha}{2}-1)n_s l^\alpha}$ and $C_s(r_s,\lambda_{st})$ (and outage) monotonically increases with r_s when $r_s \in (0,l)$. Thus, the throughput is maximized at $r_s = l$ when $\epsilon_s \geq 1-e^{-\frac{1}{2}+(\frac{\alpha}{2}-1)n_s l^\alpha}$, otherwise, when the outage constraint ϵ_s is achieved, i.e., $\delta(r_s,\lambda_{st}) = \epsilon_s \rightarrow r_s = L$.

Taking derivative of both sides of Eq. (5.8) w.r.t. λ_{st}, we get

$$\frac{dl}{d\lambda_{st}} = \frac{-2K_\alpha T_s^{2/\alpha} l^2}{\alpha^2 n_s l^{\alpha-1} + 4B(\lambda_{st})l} < 0.$$

Thus, l decreases with λ_{st}. Similarly L decreases with λ_{st} as well.

Appendix D

Communication Efficiency of Anti-jamming Broadcast in Large-scale Multi-Channel Wireless Networks

D.1 Proof of Proposition 5

For an arbitrary node v, denote by $D^u_{v,l}$ the delay for node v to receive packet l correctly. Then

$$D^u_v = \max(D^u_{v,1}, D^u_{v,2}, ..., D^u_{v,M}),$$

and the CDF of D^u_v is given by

$$\begin{aligned} \Pr(D^u_v \leq d) &= \Pr(\max(D^u_{v,1}, D^u_{v,2}, ..., D^u_{v,M}) \leq d) \\ &= \Pr(D^u_{v,1} \leq d)^M = \left(1 - (1-p_1)^d\right)^M, \end{aligned} \tag{D.1}$$

due to the fact that $D^u_{v,j}$ are i.i.d. random variables and geometrically distributed with parameter p_1.

Given the CDF of D^u_v, the mean of D^u_v is given by

$$E(D^u_v) = \sum_{d=0}^{\infty} \left[1 - \left(1 - (1-p_1)^d\right)^M\right] \triangleq z.$$

Let $1 - p_1 = e^{-\lambda}$. Noticing d is an integer and the above expression is a discrete summation,

we get

$$z > \int_0^\infty \left[1 - \left(1 - e^{-\lambda x}\right)^M\right] dx$$
$$= \sum_{k=1}^{M} \frac{1}{k}(-1)^{k+1}\binom{M}{k}\frac{1}{\lambda} = \frac{1}{\lambda}\sum_{k=1}^{M}\frac{1}{k}$$

where the last equality is from Eq. (0.155) in [67]. Similarly,

$$z - 1 = \sum_{d=1}^{\infty}\left[1 - \left(1 - (1-p_1)^d\right)^M\right]$$
$$< \int_0^\infty \left[1 - \left(1 - e^{-\lambda x}\right)^M\right] dx.$$

Therefore,

$$\frac{1}{\lambda}\sum_{k=1}^{M}\frac{1}{k} < z < \frac{1}{\lambda}\sum_{k=1}^{M}\frac{1}{k} + 1. \tag{D.2}$$

Thus

$$E(D_v^u) = \frac{1}{\lambda}\sum_{k=1}^{M}\frac{1}{k} + o(1) < \frac{1}{p_1}\sum_{k=1}^{M}\frac{1}{k} + o(1),$$

due to the fact that $\lambda = -\ln(1-p_1) > p_1$. The scaling law result follows straightforwardly from the property of the harmonic number in Lemma 6. The proof is hence completed.

D.2 Proof of dependence of $\{D_v^u\}$

We assume in the following that receivers hop at the same rate as the source, i.e., at the beginning of each slot, the receivers randomly select a channel to listen. Clearly, when the receivers hop at a slower rate, D_v^u would become more correlated[1].

Denote by $E_{i,d}$ the event that node i receives one specific packet within d rounds. Then for

[1]For example, in the extreme case when receivers never hop, the broadcast delays of nodes staying at the same channel are identical.

two nodes i and j

$$\begin{aligned}
&\Pr(D_i^u \le d, D_j^u \le d) \\
&= \Pr(E_{i,d} \cap E_{j,d})^M \\
&= \left\{ \sum_{k=1}^{d} \binom{d}{k} P_{jam}^{d-k} (1-P_{jam})^k [1-(1-1/C)^k]^2 \right\}^M \\
&= \left\{ 1 - 2(1-p_1)^d + [1-p_1(2-1/C)]^d \right\}^M,
\end{aligned} \tag{D.3}$$

where in the second equality $\binom{d}{k} P_{jam}^{d-k}(1-P_{jam})^k$ accounts for the probability that there are $k(\le d)$ rounds within which this specific packet is not jammed, and $[1-(1-1/C)^k]^2$ is the probability that both node i and j receive this specific packet within these k rounds.

In addition

$$\begin{aligned}
&\Pr(D_i^u \le d) \Pr(D_j^u \le d) \\
&= \left[1-(1-p_1)^d \right]^{2M} \\
&= \left\{ 1 - 2(1-p_1)^d + \left[1 - p_1 \left(2 - \frac{1-P_{jam}}{C} \right) \right]^d \right\}^M \\
&\le \Pr(D_i^u \le d, D_j^u \le d),
\end{aligned} \tag{D.4}$$

where in the last step the equality holds when $P_{jam} = 0$. Then,

$$\Pr(D_j^u \le d | D_i^u \le d) \ge \Pr(D_j^u \le d).$$

Intuitively successful packet reception at one node indicates that the packet is not jammed, so others can receive it at probability $1/C$, instead of $(1-P_{jam})/C$. That is, given $D_i^u \le d$ the event $D_j^u \le d$ occurs more possibly.

Remark 14. *Close observation shows that the difference between Eq. (D.3) and (D.4) is negligible for a sufficiently large C. Therefore, when receivers hop at the same rate as transmitters, D_v^u can be reasonably considered as i.i.d random variables.*

D.3 Proof of Proposition 6

It's easy to check that p_{n_r} is a concave function of n_r. The optimal $n_r^*(M)$ follows by solving

$$\frac{dp_{n_r}}{dn_r} = M(a^{n_r} \ln a - b^{n_r} \ln b)(1-P_{jam}) = 0.$$

It can be verified that n_r^* is a decreasing function with M, thus the maximum and infimum of n_r^* are given as follows:

$$\bar{n}_r^* = n_r^*(2) = \frac{\ln \frac{\ln(1-1/C)}{\ln(1-1/(2C))}}{\ln \frac{1-1/(2C)}{1-1/C}} \approx 2\ln 2C \approx 1.4C,$$
$$\underline{n}_r^* = \lim_{M\to\infty} n_r^*(M) = \frac{1}{-\ln(1-1/C)} \approx C.$$

The approximations in this proposition are made by the fact that $\ln(1+x) \approx x$ for small x.

D.4 Proof of Proposition 7

It is easy to check that

$$\tilde{p}_{n_r}(M) > \tilde{p}_{n_r}(\infty) = \frac{n_r}{C}\left(1-\frac{1}{C}\right)^{n_r-1} > \tilde{p}_{\bar{n}_r^*}(\infty),$$

where $\bar{n}_r^*$ is given in Eq. (6.11). Therefore

$$p_{n_r^*(M)}(M) > \tilde{p}_{\bar{n}_r^*}(\infty)(1-P_{jam}) \to 0.35(1-\alpha_1-\alpha_2/M),$$

where the convergence is made due to the fact that $\left(1-\frac{a}{C}\right)^{n_r} \to \exp\left(-\frac{an_r}{C}\right)$ for large C and n_r.

Next we show that $p_{n_r^*}(M)$ is an increasing function of M (Note that P_{jam} is also a function of M.). Fix $n_r \in [\underline{n}_r^*, \bar{n}_r^*]$ in Eq. (6.10), then, for large C and n_r, $p_{n_r}(M)$ can be well approximated by

$$\begin{aligned}&[(1-\alpha_1)M-\alpha_2]\left[\exp\left(-\frac{(1-1/M)n_r}{C}\right)-\exp\left(-\frac{n_r}{C}\right)\right]\\ &=[(1-\alpha_1)M-\alpha_2]\exp\left(-\frac{n_c}{C}\right)\left(\exp\left(\frac{n_r}{CM}\right)-1\right).\end{aligned}$$

The derivative of $p_{n_r}(M)$ with respective to M is

$$\frac{dp_{n_r}(M)}{dM} = (1-\alpha_1)\left[\exp(x_0)(1-x_0+\frac{\alpha}{M}x_0)-1\right]e^{-\frac{n_c}{C}},$$

where $x_0 = \frac{n_r}{CM}$. Due to the fact $M > \alpha$, function $f(x) = \exp(x)(1-x+\frac{\alpha}{M}x)$ monotonically increases when $x < \frac{\alpha}{M-\alpha}$. Given $\alpha > 1.4$, $x_0 < \frac{\alpha}{M-\alpha}$. Therefore, $f(x_0) > f(0) = 1$ and

$\frac{dp_{n_r}(M)}{dM} > 0$ for $n_r \in [\underline{n}_r^*, \bar{n}_r^*]$. Thus, $p_{n_r}(M)$ ($n_r \in [\underline{n}_r^*, \bar{n}_r^*]$) increases with M and converges to

$$y(n_r) \triangleq \lim_{M\to\infty} p_{n_r}(M) = \frac{n_r}{C}\left(1-\frac{1}{C}\right)^{n_r-1}(1-\alpha_1).$$

The upper bound of $p_{n_r^*(M)}(M)$ follows from the fact that

$$p_{n_r^*(M)}(M) \le p_{n_r^*(M)}(M+1) \le p_{n_r^*(M+1)}(M+1),$$

and $p_{n_r^*(\infty)}(\infty) = y(\underline{n}_r^*) \to 0.38(1-\alpha_1)$, where $\underline{n}_r^* = \lim_{M\to\infty} n_r^*(M)$ is given in the last proposition.

D.5 Proof of the Proposition 8

$$g_c(n_r) \;=\; \frac{p_{n_r}}{p_1} = \frac{M(a^{n_r}-b^{n_r})}{1/C} = C\tilde{p}_{n_r}(M).$$

It is easy to check that $\tilde{p}_{n_r}(M)$ decreases with M,

$$\tilde{p}_{n_r}(\infty) < g_c(n_r)/C < \tilde{p}_{n_r}(2).$$

The bounds of $g_c(n_r)$ are then established.

Noticing that function $y_1(n_r) = \tilde{p}_{n_r}(2)(y_2(n_r) = \tilde{p}_{n_r}(\infty))$ is maximized (minimized) at $n_r = \bar{n}_r^*$ for $n_r \in [\underline{n}_r^*, \bar{n}_r^*]$, the conclusion for the optimal cooperation gain follows after some calculation (plugging $\bar{n}_r^*$ into $y_1(n_r)$ and $y_2(n_r)$).

D.6 Derivation of $a(j,k)$

Our case is equivalent to the scenario where k eggs are independently put into one of j baskets. Let $a(j,k)$ be the number of possibilities that each basket at least has one egg, which can be calculated recursively as

$$a(j,k) = j^k - \sum_{i=1}^{j-1}\binom{j}{i}a(i,k),$$

and we can get $a(j,k) = \sum_{i=0}^{j}(-1)^i\binom{j}{i}(j-i)^k$ after some computation. Then $P(M,k) = a(M,k)/M^k$.

D.7 Proof of Corollary 4

$$
\begin{aligned}
&\epsilon(n_r, i) \\
&= \sum_{l=0}^{M} (-1)^l \binom{M}{l} \left[\sum_{k=M}^{iM} \binom{iM}{k} (p_{n_r} - \frac{p_{n_r} l}{M})^k (1 - p_{n_r})^{iM-k} \right] \\
&= \sum_{l=0}^{M} (-1)^l \binom{M}{l} \left[\left(1 - \frac{l p_{n_r}}{M}\right)^{iM} \right. \\
&\left. - \sum_{k=0}^{M-1} \binom{iM}{k} \left(p_{n_r} - \frac{l p_{n_r}}{M}\right)^k (1 - p_{n_r})^{iM-k} \right] \\
&\approx \sum_{l=0}^{M} (-1)^l \binom{M}{l} \exp(-il p_{n_r}) \\
&= (1 - \exp(-ip_{nr}))^M,
\end{aligned}
$$

where the second equality follows from the Binomial theorem and the approximation is made due to the fact that $(1 - \frac{lp_{nr}}{M})^{iM} \to \exp(-ilp_{n_r})$ and $\sum_{k=0}^{M-1} \binom{iM}{k} ((1 - \frac{l}{M})p_{n_r})^k (1 - p_{n_r})^{iM-k}$ is negligible for large iM. $\epsilon(n_r, i)$ and its estimation are numerically depicted in Fig. D.1. We can see an excellent match even when i and M are small.

Following the line of proof in Theorem 9, we can show that

$$
\begin{aligned}
E(D^{icn}) &= \sum_{i=0}^{\infty} (1 - \epsilon(n_r, i)^n) \\
&\approx \sum_{i=0}^{\infty} \left(1 - (1 - \exp(-ip_{nr}))^{Mn}\right) \\
&\approx \frac{1}{p_{n_r}} \sum_{k=1}^{nM} \frac{1}{k}.
\end{aligned}
$$

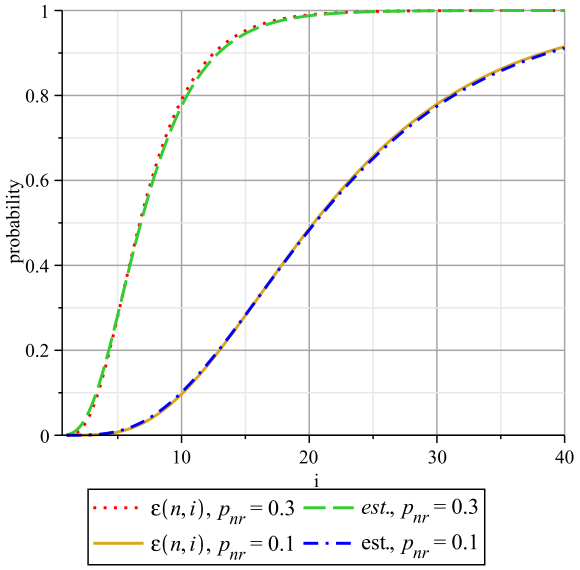

Figure D.1: $\epsilon(n_r, i)$ in Eq. (6.13) and its estimation

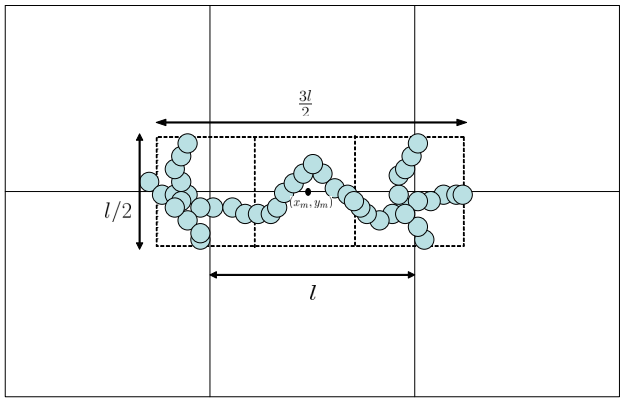

Figure D.2: Grid construction

D.8 Proof of Lemma 7

First we show the necessity of condition $\lambda_n > \lambda_{n,c}$ by coupling techniques[2]. According to Fact 1, network $G(V, E_{\lambda_n})$ is percolated if and only if $\lambda_n > \lambda_{n,c}$. Since jammers only have negative influence on the network connectivity $G(V, E_{\lambda_n,\lambda_m})$ is possibly connected if and only if $\lambda_n > \lambda_{n,c}$.

Next we map the Poisson boolean model $B(X, r_n/2, \lambda_n)$ to the discrete bond percolation model. Consider a random grid $\mathcal{G}_l$ with side length (inter-node distance) l (Fig.D.2). For each edge m denote by (x_m, y_m) its middle point. An edge is open, which is referred to as event E_o, if the following two events happen:

E_1 :There is an occupied L-R crossing of the rectangle $[x_m - 3l/4, x_m + 3l/4] \times [y_m - l/4, y_m + l/4]$ in $B(X_n, r_n/2, \lambda_n)$; AND there are two occupied T-B crossings of the square $[x_m - 3l/4, x_m - m/4] \times [y_m - l/4, y_m + l/4]$ and the square $[x_m + l/4, x_m + 3l/4] \times [y_m - l/4, y_m + l/4]$ in $B(X_n, r_n/2, \lambda_n)$;

E_2 : There are no more than J jammers within distance r_m of any node on the crossings involved in E_1.

Otherwise the edge is close. Given $\lambda_n > \lambda_c(r_n/2)$, our objective is to show that $\exists L > 0, \underline{\lambda}_{m,u} > 0$ such that for $l = L$ and an arbitrary original O, $\forall \lambda_m < \underline{\lambda}_{m,u}$, $\Pr(|C(O)| = \infty) > 0$, where $C(O)$ denotes the open connected component containing the original O and $|C(O)|$ is the number of edges in $C(O)$. This objective can be achieved as follows.

[2]Intuitively, coupling accounts for simultaneously constructing two realization of networks on the same probability space, with the condition that if some event occurs in one realization, it should also happen in the other.

For an arbitrary edge m in $\mathcal{G}_l$, let $p = \Pr(E_o)$, then we have

$$p = \Pr(E_1, E_2) = \Pr(E_2|E_1)P(E_1). \tag{D.5}$$

From Fact 2

$$\begin{aligned}\Pr(E_1) &\geq \Pr[LR(3l/2, l/2, \lambda_n)] \Pr[TB((l/2, l/2), \lambda_n)]^2 \\ &\to 1 \text{ as } l \to \infty,\end{aligned}$$

where the inequality is due to FKG's inequality[3]. As a result, given $\epsilon \in (0, 1)$, there exists an L such that $\Pr[E_1] > \sqrt{\epsilon}$ when $l = L$.

Let $l = L$, and e_J denote the event that there are at most J jammers in the rectangle $[x_m - 3l/4 - r_m, x_m + 3l/4 + r_m] \times [y_m - l/4 - r_m, y_m + l/4 + r_m]$

$$\begin{aligned}\Pr[E_2|E_1] &> \Pr[e_J|E_1] = \Pr[e_J] \\ &= \exp(-\hat{\lambda}_m) + \sum_{k=1}^{J} \frac{\hat{\lambda}_m^k}{k!} \exp(-\hat{\lambda}_m) \\ &\to 1 \text{ as } \lambda_m \to 0,\end{aligned}$$

where $\hat{\lambda}_m = \lambda_m(\frac{l}{2} + 2r_m)(\frac{3l}{2} + 2r_m)$, and the equality is due to the fact that e_J and E_1 are independent. Thus, there exists $\underline{\lambda}_{m,u}$ such that $\Pr[E_2|E_1] > \sqrt{\epsilon}, \forall \lambda_m < \underline{\lambda}_{m,u}$.

Then according to Eq. (D.5),

$$\Pr[E_o] > \epsilon \text{ for } l = L, \text{ and } \forall \lambda_m < \underline{\lambda}_{m,u}.$$

The state of an edge is correlated with its six adjacent edges and independent of other edges. According to the Peirels argument and dependent discrete percolation [47], we can show that there exists probability p' such that if $p > p'$, $\Pr[|C(O)| = \infty] > 0$, which implies that there exists a giant component in grid $\mathcal{G}_l$ due to network ergodicity. According to the construction of $\mathcal{G}_l$, a giant component in $\mathcal{G}_l$ implies a giant component in $G(V, E_{\lambda_n, \lambda_m})$. The proof is then completed.

[3]If two events A_1 and A_2 are increasing or decreasing, then $\Pr[A_1 \cap A_2] \geq \Pr[A_1] \Pr[A_2]$. An event A is increasing if adding an edge in any realization of the random network where A occurs, leads to a configuration which A still holds. An event is decreasing if its complement is increasing.

D.9 Proof of Lemma 8

Following the procedure on constructing a branching process in Theorem 2.5.1 in [47], it can be shown that the network is connected if the average node degree $u > 1$. In our case, the average degree u of an arbitrary node a is defined as

$$u = E(\deg(a)|\overline{Blk}(a)),$$

where $\overline{Blk}(a)$ is the event that a is not blocked, i.e., there are no more than J jammers within r_m distance from a.

Denote by $N_a(k)$ the event that there are k nodes within r_n distance from a and let b an arbitrary neighbor of a. Then we get,

$$\begin{aligned}
u &= \sum_{k=0}^{\infty} e^{-\lambda_n \pi r_n^2} \frac{(\lambda_n \pi r_n^2)^k}{k!} E(\deg(a))|\overline{Blk}(a) \cap N_a(k)) \\
&= \sum_{k=1}^{\infty} e^{-\lambda_n \pi r_n^2} \frac{(\lambda_n \pi r_n^2)^k}{k!} k \Pr(\overline{Blk}(b)|\overline{Blk}(a)) \\
&\leq \lambda_n \pi r_n^2 \Pr[\overline{Blk}(b)] \\
&= \lambda_n \pi r_n^2 \sum_{i=0}^{J} e^{-\lambda_m \pi r_m^2} \frac{(\lambda_m \pi r_m^2)^i}{i!},
\end{aligned}$$

where the inequality is due to the fact that the jammers around a may also have influence on b. The lemma follows by solving the equation

$$\lambda_n \pi r_n^2 \sum_{i=0}^{J} e^{-\lambda_m \pi r_m^2} \frac{(\lambda_m \pi r_m^2)^i}{i!} = 1.$$

D.10 Proof of Lemma 10

The average degree u of an arbitrary node a is defined as

$$u = E(\deg(a)|\overline{Jam}(a)),$$

where $\overline{Jam}(a)$ is the event that a is not jammed, i.e, there are no jammers within r_m distance from a. Then following the same line as Appendix D.9, we have

$$u \leq \lambda_n \pi r_n^2 \Pr[\overline{Jam}(b)] = \lambda_n \pi r_n^2 e^{-\lambda_m \pi r_m^2}.$$

$\bar{\lambda}_{m,l}$ follows by solving the equation

$$\lambda_n \pi r_n^2 e^{-\bar{\lambda}_{m,l}\pi r_m^2} = 1.$$

D.11 Subadditive Ergodic Theorem [2]

Let $T_{l,k}$ be a collection of random variables indexed by integers $0 \leq l < k$. Suppose $\{T_{l,k}\}$ has the following properties:

1. $T_{0,k} \leq T_{0,l} + T_{l,k}, 0 \leq l < k$;
2. $\{T_{(q-1)i,qi}, q \geq 1\}$ is a stationary process for each i.
3. the distribution of $\{T_{l,l+q} : q \geq 1\}$ does not depends on l;
4. $E[|T_{0,k}|] < \infty$ for each k.

Then

a) $\delta \triangleq \lim_{k\to\infty} \frac{E[T_{0,k}]}{k} = \inf_{k\geq} \frac{E[T_{0,k}]}{k}; T \triangleq \lim_{k\to\infty} \frac{T_{0,k}}{k}$ exists with probability 1 and $E[T] = \delta$;

Furthermore, if

5) the stationary process in (2) is ergodic,

then

b) $T = \delta$ with probability 1.

D.12 Proof of Lemma 11

An interfering cluster of an arbitrary node i is defined as a cluster which has relays within distance $(1 + \Delta)r_n$ of node i. Therefore the distance between node i and the cluster heads of its interfering clusters is at most $(2 + \Delta)r_n$. Since two cluster heads are at least r_n away from each other, the circles centered at cluster heads with radius $r_n/2$ do not overlap. Then the number of interfering clusters of node i is at most $\frac{\pi(2.5+\Delta)^2 r_n^2}{\pi 0.5^2 r_n^2}$. Similarly the distance between two neighboring cluster heads is at most $3r_n$, and a cluster has at most $\frac{3.5^2}{0.5^2} - 1 = 48$ neighboring clusters, which indicates that a cluster has at most 48 relays. The lemma follows after simple calculation.

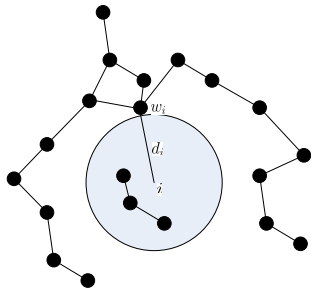

Figure D.3: Circle C_i in Lemma 13

D.13 Proof of Lemma 12

We consider the worst scenario where a receiver is interfered by $n_r = 192(2.5+\Delta)^2 - 1$ relays. Since time is not synchronized one slot of a particular transmitter can collide with two successive slots of other transmitters a.s.. Thus, the receiver receives a packet from its corresponding transmitter correctly at probability:

$$p_{n_r} = \frac{1}{C}(1-\frac{1}{C})^{2(n_r-1)}(1-p'_j). \tag{D.6}$$

Due to the fact that the one hop delay is geometrically distributed with parameter p_{n_r}, it is finite.

D.14 Proof of Lemma 13

Consider the circle C_i centered at i with radius d_i. According to the definition of w_i,

$$G^s(V, E_{\lambda_n}) \cap C_i \notin \mathcal{C}(G^s(V, E_{\lambda_n})).$$

In other words, C_i only contains nodes belonging to finite components of $G^s(V, E_{\lambda_n})$, or C_i does not contain any nodes (see Fig. D.3 for illustration). Therefore according to the fact 1 of last section, the area of C_i is bounded in size, which leads to $d_i < \infty$.