On the Scalability of Ad Hoc Dynamic Spectrum Access Networks

Umair Ahsan

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Electrical Engineering

Claudio R. C. M. da Silva, Chair
R. Michael Buehrer
Allen B. MacKenzie

September 30, 2010
Blacksburg, Virginia

Keywords: Scalability, Dynamic Spectrum Access, DSA, Ad Hoc, Network Capacity, Cognitive Radio

On the Scalability of Ad Hoc Dynamic Spectrum Access Networks

Umair Ahsan

ABSTRACT

Dynamic Spectrum Access allows wireless users to access a wide range of spectrum which increases a node's ability to communicate with its neighbors, and spectral efficiency through opportunistic access to licensed bands. Our study focuses on the scalability of network performance, which we define in terms of network transport capacity and end-to-end throughput per node, as the network density increases. We develop an analytical procedure for performance evaluation of ad hoc DSA networks using Markov models, and analyze the performance of a DSA network with one transceiver per node and a dedicated control channel. We also develop and integrate a detailed model for energy detection in Poisson networks with sensing. We observe that the network capacity scales sub-linearly with the number of DSA users and the end-to-end throughput diminishes, when the number of data channels is fixed. Nevertheless, we show that DSA can improve network performance by allowing nodes to access more spectrum bands while providing a mechanism for spectrum sharing and maintaining network wide connectivity. We also observe that the percentage of relative overhead at the medium access layer does not scale with the number of users. Lastly, we examine the performance impact of primary user density, detection accuracy, and the number of available data channels. The results help to answer the fundamental question of the scaling behavior of network capacity, end-to-end throughput, and network overhead in ad hoc DSA networks.

ACKNOWLEDGEMENTS

I would like to thank my parents for their love and support, my brothers for always being there for me, Dr. Claudio da Silva for his support and advice, Dr. Buehrer for his continued mentorship, Dr. Mackenzie for serving on my committee, and all of the inhabitants of the town of Blacksburg, Virginia for making it a wonderful place to live and study. The work presented in this thesis was partially supported by a research contract from the Defense Information Systems Agency (DISA).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Since wireless spectrum is a limited resource, it must be shared between different wireless networks that desire to co-exist in a given area. Typically, most wireless networks are allocated orthogonal blocks of spectrum through a complicated network and frequency planning procedure regulated by government agencies such as the FCC. This procedure is cumbersome, costly, and leads to inefficiencies such as low spectrum utility, which measures the amount of licensed spectrum that is actually occupied in a given area. According to studies by the FCC and others [1-3], spectrum utility averages only 15% to 85% at any given time—even in urban areas with heavy wireless use.

A new paradigm for spectrum access commonly referred to as Dynamic Spectrum Access (DSA) or Opportunistic Spectrum Access (OSA) is envisioned to increase spectrum utility and ease the burden on spectrum planners as wireless networks grow in size and number. DSA is proposed to allow next-generation wireless networks to self-organize and share spectrum efficiently. Since most of the usable part of the radio spectrum is already licensed to primary users (PU) with privileged access to the spectrum, the first DSA networks are designed to co-exist as secondary users (SU) and opportunistically access spectrum. Knowledge of the radio environment allows DSA users to access unused spectrum at a given time and location while avoiding interference to primary users.

Managing spectrum access is particularly challenging for a class of wireless networks known as *ad hoc networks*. Ad hoc networks—defined in Section 1.1—may have special spectrum access needs, such as requiring access for a short period of time and only in particular locations. As with certain centralized networks, the amount and duration of spectrum required may vary temporally and spatially in an irregular and unpredictable fashion. Large ad hoc networks with mobile nodes, dynamic traffic patterns, and individualized QoS requirements require spectrum assignments that may be too complex for human planning. Therefore, ad hoc networks are an ideal candidate for DSA since licensing whole blocks of spectrum may be especially challenging and inefficient. DSA may also help the *scalability* issue that ad hoc networks have long been

known to suffer from [4]. Scalability refers the behavior of the network performance as the network density increases.

DSA introduces a platform for flexible network architectures and the freedom for each node to re-configure its communications with the changing environment (including RF and networking conditions). DSA allows users to access a wide range of spectrum bands which significantly increases a node's ability to communicate with its neighbors, make better use of available spectrum, and may potentially enable ubiquitous network wide connectivity. However, DSA is a new paradigm and comes bundled with new challenges that need to be characterized. For instance, DSA requires an even greater level of network cooperation as nodes must coordinate their spectrum use to enable communication and mitigate interference. Such cooperation requires overhead and may further limit the scalability of ad hoc DSA networks.

Our scalability study focuses on the scaling behavior of network performance as the network density increases. This work studies the trends in scalability seen with conventional and DSA ad hoc networks. Along the way, we identify key areas of focus and develop simulation and analytical tools to evaluate the network performance. We start, in Chapter 2, with an in-depth examination of previous work in the scalability of conventional ad hoc networks. Then in Chapter 3, we show simulation results on the scalability of single-channel ad hoc networks with random access MAC schemes. Then, in Chapter 4, we develop an analytical model which allows us to study the performance and scaling behavior of general ad hoc DSA networks. We use our analytical method to show the scaling properties of a class of ad hoc DSA networks. Lastly, in Chapter 5, we summarize our results and discuss important conclusions.

1.1 Background

1.1.1 Ad Hoc Networks

Ad hoc wireless networks—also known as a multi-hop or infrastructure-less networks—contain no centralized controller or hierarchy. Consequently, ad hoc networks usually operate in conjunction with a traffic model where each node also serves as a router for other nodes in the network [4-6]. Relays in an ad hoc network usually follow a *decode-and-forward* model in which each relaying node decodes the message, re-encodes it, and forwards it to the next relay [7]. There are a number of routing strategies available to ad hoc networks that can be classified as flat, hierarchal, and location assisted [8]. Each route formation method impacts the performance and scalability of ad hoc networks, nevertheless each strategy attempts to select a relay among neighboring nodes which minimizes the distance, time, or number of hops to the final destination. We do not examine routing methods in this study and direct the reader towards the study by Xiaoyan et. al. [8]

Mobile ad hoc networks remove the need for complex network planning and can be deployed instantaneously which allows for many interesting applications. For example, ad hoc networks can be set up in minutes for crisis management scenarios in which the existing communication infrastructure fails [5]. Also, ad hoc networks are ideal for certain military applications, since military operations are often spontaneous and in remote locations, where the existing communication infrastructure might not be available or usable due to security and reliability

concerns. Additionally, since communication is limited to nearby neighbors and occurs at low powers, it is easier to escape detection and interception [4]. Several existing types of ad hoc networks are commercially available such as Wireless Personal Area Networks (WPANs), Wireless Local Area Networks (WLANs), and sensor networks [6]. Ad hoc wireless networks can also be used for data sharing on the fly in a business environment [6].

The characteristics of ad hoc networks that make them attractive also lead to major design challenges, in addition to those faced by traditional wireless networks. In the absence of a central controller, each node must take on the responsibility of interference management (power control and medium access control) and routing. Alternatively, nodes may form groups at the expense of additional overhead, hierarchy and complexity. Distributed algorithms may be better functional alternatives in ad hoc networks, however distributed algorithms are much harder to implement optimally, and also carry additional overhead. Many applications of ad hoc networks require mobility, thus ad hoc networks must be capable of self-organization and self-configuration [4], [6]. Mobility and dynamic routes contribute to the lack of fixed references which do not tolerate a fixed hierarchal structure. Due to these factors, it is well known that scalability is one of the most critical challenges facing ad hoc networks [4]. Additionally, and as shown later, adopting a multi-hop transport model—where traffic is repeated from one node to the next in a series of "hops"—directly limits the scalability of ad hoc networks [9].

1.1.2 Ad hoc Network MAC Schemes

As mentioned previously, due to the distributed nature of ad hoc networks, scalability issues can arise in the physical, medium access control (MAC), and routing layers. We focus our study on the fundamental scalability limits of ad hoc networks and the scalability impact at the MAC layer. The general problem of spectrum sharing in conventional (non-DSA) ad hoc networks is also inherited by DSA ad hoc networks, which are discussed in Section 1.2. Therefore, we review spectrum sharing techniques in multi-hop networks in this section, and carefully evaluate their scalability impact in Chapter 3.

Traditional infrastructure-based networks rely on MAC schemes such as FDMA, TDMA, and CDMA where simultaneous transmissions occur in orthogonal domains of frequency, time, or codes to avoid mutual interference. However, these schemes usually rely on a centralized network entity to schedule transmissions for all users. Ad hoc networks, in the absence of a centralized controller, cannot rely on such methods without excessive network overhead due to mobility and dynamic links between nodes. Consequently, random access MAC schemes such as Aloha and CSMA, which allow for users to access the channel in an uncoordinated or semi-coordinated manner, are preferred for ad hoc networks.

The Aloha protocol was originally developed in 1970 by Norman Abramson at the University of Hawaii as a multiple access solution for a university wide, single-hop ad hoc network [10]. The basic Aloha protocol does not restrict transmissions; rather each node transmits when it has traffic and attempts to resolve collisions through a retransmission scheme. Under the protocol, users send data as it arrives and wait for an acknowledgement (ACK) from the receiver, if the ACK is not received the transmitter assumes there was a collision and retransmits the data after a random delay. The random delay before retransmitting attempts to avoid repeated collisions on

the channel. The Aloha protocol can also be implemented in a slotted form, where nodes are synchronized and may only begin transmission at discrete time slots. If the slot length is equal to or greater than the packet length, the slotted form reduces the probability of collision seen by each user and in turn increases efficiency [11]. Due to the characteristics of ad hoc networks, traditional clock synchronization algorithms are not applicable and maintaining synchronism in an ad hoc network can be challenging [12]. A central time server like GPS can be utilized to maintain synchronism but GPS also has limitations especially in indoor or dense urban environments [12].

Carrier Sense Multiple Access (CSMA), first introduced by Tobagi and Kleinrock in 1975 in [13], improves on Aloha introduces an opportunistic element in medium access. An ideal medium access scheme requires knowledge of the state and interference level at the receiver, and the state of nearby nodes. CSMA allows the transmitter to estimate the level of interference at the receiver, and nearby nodes, by sensing its own interference level. If the sensed interference exceeds a predefined threshold, the transmitter withholds transmission until the channel is clear. The benefit of sensing acts as a multiplier since it prevents new transmissions that are likely to fail and reduces interference seen by ongoing transmissions.

Existing CSMA protocols can be classified as p-*persistent* or *non-persistent* [13]. Upon sensing a busy channel, a node following a p-persistent CSMA protocol continuously senses the channel and attempts a transmission with probability *p* once the channel is clear. Continuously sensing the channel requires resources and nearly similar performance can be achieved with non-persistent CSMA. In non-persistent CSMA, once a node has traffic to transmit it senses the channel and transmits if it is idle. If the channel is busy, the node periodically senses after random delays until the channel is clear, and then transmits. There are also slotted forms of CSMA which require synchronization and transmission at the beginning of discrete time slots [13].

In single-hop networks, where all nodes can hear each other, CSMA provides a significant gain in performance over Aloha and mitigates the problem of performance dependence on the network load [13]. The benefits of CSMA start to deteriorate in multi-hop networks due to *hidden nodes*, which cause interference to the receiver but are out of the sensing region of the transmitter[14]. Hidden nodes often occur in a multi-hop ad hoc network, where transmissions are localized between neighboring nodes.

1.1.3 Dynamic Spectrum Access

Cognitive radio ad hoc networks [15], also referred to as DSA networks [16], are envisioned to provide mobile ad hoc users with access to more spectrum. The biggest advantage of DSA networks is allowing higher usage of spectrum by opportunistically working in licensed bands in the presence of primary users [15], [16]. The key enabling technology of DSA networks is the cognitive radio platform, originally introduced in 1999 by Mitola III and Maguire [17]. A cognitive radio can change its transmission parameters based on interaction with the environment in which it operates [17]. Two main characteristics of cognitive radios are their *cognitive capability* and *reconfigurability* [15] which lead to performance gains on the physical layer.

Cognitive capability refers to a radio's ability to learn from its radio environment, and adapt to the network state to meet user requirements. Cognitive radios are also intelligent to transform observations to information and evaluate the best course of action to reach certain goals. Additionally, cognitive radios are reconfigurable and able to modify operational parameters such as transmission power, modulation, coding, operating frequency, bandwidth, etc. adaptively. Cognitive radios optimize performance through a combination of sensing, cognition, and reconfigurability [15], [17]. Finally, cognitive radio networks can also increase performance through the idea of *cognitive routing.* Cognitive radios should be able to form routes, and re-route as necessary, based on spectrum sensing, network conditions and changing traffic patterns. This functionality is necessary to avoid cases where an end-to-end route may result in a high level of interference to the primary user or a high level of latency [15].

Cognitive radios can be used to realize key enabling requirements for DSA, which we group under the umbrella term *spectrum management*. Spectrum management in DSA networks must perform all of the following functions [15], [16]:

- Spectrum sensing: Detecting unused spectrum and avoiding interference to PUs.
- Spectrum analysis: Estimating characteristics such as channel gain and path loss of available spectrum.
- Spectrum decision: Accessing the best available spectrum to meet QoS requirements while also considering route and topology formation.
- Spectrum mobility: Maintaining seamless communication while transitioning between spectrum.
- Spectrum sharing: Ensure fair spectrum scheduling among coexisting DSA users.

Traditional network protocols lack support for functions required for operation of DSA networks. Specifically, the DSA medium access layer, in addition to the traditional duty of spectrum sharing, must take on an expanded role with additional duties of spectrum management. Route formation is also coupled with spectrum access in DSA networks since two nodes must share common spectrum to form a link. Additionally, the physical layer must account for the differences in propagation characteristics and interference levels between spectrum bands. Traditional approaches to routing require a common broadcast channel for route formation which may not be available in ad hoc networks. The spectrum access and routing functions of DSA networks must be considered in tandem since nodes do not necessarily share a common channel. Consequently, the existing layered networking model may not be a suitable approach to protocol design for DSA networks [16]. In fact, cross-layer route selection and spectrum allocation approaches that take advantage of knowledge of the availability and quality of spectrum between links have been shown to outperform traditional approaches where routes and spectrum are selected independently [18], [19].

For the purpose of our scalability study, we focus on the spectrum management features of DSA as these features form the core of DSA. In the next sections, we discuss the details of these features, give examples of some proposed DSA MAC protocols that contain these features, and discuss the additional challenges associated with ad hoc DSA networks.

1.1.4 Spectrum Sensing

In DSA networks, avoiding interference to the primary users is a high priority, therefore radios must continually sense for the appearance of primary users. A radio may adopt a periodic sensing policy or use on-demand sensing if the network lacks synchronization. Alternatively, secondary users may be able to access an internet database containing spectrum bands which are accessible in a given area. If a user is unable to sense the spectrum it would require a constant connection to the database, along with its own position information which can be costly and require modifications to legacy licensed systems [20]. Spectrum sensing allows secondary users to access spectrum opportunistically, without increasing infrastructure costs and allows for operation in remote areas.

Spectrum sensing maybe implemented using an energy detector, matched filter, or more sophisticated techniques like cyclostationary feature detection. A matched filter is an optimal detector if the primary user signal's structure is known, but it requires dedicated circuitry for synchronization. If a SU can access a wide variety of bands, matched filter detection increases the cost and complexity of the receiver [21]. In that case, or if the primary user signal structure is unknown, energy detection is an optimal solution. Energy detection compares the measured received energy level with a predetermined threshold. Energy detection is unable to determine the source or structure of the received signal, but is the simplest to implement since it does not require synchronization with the source. If more information about the received signal is available, cyclostationary feature detection can be used, but at the expense of longer sensing time and complexity [20].

Accurate sensing may require an extremely sensitive detector for worst case scenarios, such as high shadowing, fading, or noise uncertainly. Distributed sensing, where two or more users cooperatively sense the spectrum, has been shown to produce accurate results with less sensitive detectors [20], [22]. Distributed sensing leverages location diversity and also reduces the hidden node problem. For example, if one user misses a transmitter due to a strong fade, another user may sense the missed transmitter.

1.1.5 Spectrum Decision

Spectrum decision refers to the ability of cognitive radios to access the best available spectrum band based on the radio environment and the QoS requirements. The process involves *spectrum characterization* based on local observations and the behavior of primary users, and *spectrum selection* where nodes select the best available channel [15], [16]. In ad hoc networks, spectrum decision must also consider the additional constraints of the end-to-end route over multiple hops. Network connectivity, route formation and spectrum availability are inter-dependent and change over time and location.

Cognitive radios are envisioned to access a wide range of spectrum bands, spectrum characterization is important. The radio must observe the RF environment at various bands of interest for the level of interference, path loss, and link errors. Based on the RF conditions, the radio must adaptively determine the transmission power, modulation scheme, and link layer protocols used. Monitoring and characterizing PU activity also aids in spectrum selection, for

example nodes can learn to avoid those bands which are heavily used by primary users to minimize interference. Spectrum selection rules can then be developed based on the nodes' local observations, the applications QoS requirements and end-to-end routing requirements.

1.1.6 Spectrum Sharing

DSA users operating in licensed bands must not block primary users from access and only opportunistically use the spectrum. DSA operation and spectrum sharing is intimately tied with spectrum sensing since each node must assure that the channel is free of primary users before transmission. A node must allocate resources to sensing to satisfy the dual goals of minimizing interference and meeting QoS requirements. Sensing is integral to DSA operation and forms the core of many DSA MAC protocols [18]. In order to differentiate between primary and secondary users, some current proposals DSA MAC proposals [15], [18], [23] include periodic *quiet periods* where all secondary users in a region stop transmission and sense for the PU. Section 1.1.7 surveys currently proposed DSA MAC protocols.

Additionally, DSA users must share the available spectrum with other secondary users in a way that ensures fairness and manages self-interference. Traditionally, the MAC layer schedules transmissions between users to manage multiple access interference. In addition to inheriting conventional ad hoc MAC layer problems, DSA networks introduce new challenges to spectrum sharing. Since DSA nodes can access a wide range of spectrum, there must be a handshake mechanism between transmitter and receiver to ensure that both nodes *rendezvous* on the correct channel. DSA nodes must also be able to switch to a different channel if a primary user comes on during a data transmission. Most current DSA approaches include a common control channel, either regional or network wide, which is used for handshakes and network authentication. Current DSA MAC protocols realize the control channel in two different ways:

> A. Common Control Channel (CCC): A separate frequency channel is devoted to the transport of control messages. The channel maybe out of band and free of primary users or shared with the primary users. If the control channel is in a separate band which is licensed to the DSA network and free of primary users, the design is called a dedicated common control channel (DCCC).
>
> B. Split phase CC (SPCC): This approach allows for exchange of data information over the control channel by dividing time into control and data transmission phases. SPCC allows for data transmissions over the control channel which may increase performance, especially when nodes have short data packets.

A common control channel solves the rendezvous problem, and also offers a performance advantage by ensuring connectivity between nodes. However, a common control channel between all nodes is characteristic of a fixed spectrum system and does not fully utilize DSA. DSA nodes are supposed to be spectrum agile and not rely on a fixed band anywhere. Additionally the control channel may serve as a single point of failure because if the control channel is saturated or unavailable no communication can take place on any other channel. This makes the control channel architecture vulnerable for jamming in military networks and constraints the performance since nodes must access the control channel before data transmissions.

1.1.7 Survey of DSA MAC Schemes

Protocol	Arch.	Access Scheme	Radios	Control Channel
DOSS-MAC [25]	Ad hoc	Random	3	Dedicated CC
SRAC-MAC [26]	Ad hoc	Random	1	Split CC
HC-MAC [27]	Ad hoc	Random	1	Dedicated CC
SYN-MAC [28]	Ad Hoc	Hybrid	Multi.	Dedicated CC
SCA-MAC [29]	Ad hoc	Hybrid	2	Dedicated CC
DCA-MAC [30]	Infra.	Random	Multi.	Dedicated CC
C-MAC [23]	Infra.	Time Slotted	1	Dedicated CC
OS-MAC [32]	Infra.	Hybrid	1	Dedicated CC
POMDP [31]	Infra.	Hybrid	Multi.	Dedicated CC
Opportunistic MAC [33]	Infra.	Hybrid	2	Dedicated CC
OSA-MAC [34]	Infra.	Hybrid	1	Dedicated CC

Table 1.1 Classification of DSA MAC protocols.

Several DSA MAC protocols have been proposed in literature [18], [24] for ad hoc [25-29] and centralized networks [23], [30-34] and can be further classified based on their spectrum access mechanism as random access, time slotted, and hybrid. Random access protocols [25-27], [30] allow nodes to opportunistically send control and data information based on spectrum sensing. Time slotted protocols [23] designate slots for each user for control and data packets and may not be suitable for ad hoc networks due to strict synchronization requirements. Hybrid protocols [28], [29], [32-34] may use time slots for control information and employ random access and contention for data transmissions. Another, and more common, hybrid approach is to divide the time into super frames, or beacon intervals, with distinct phases for sensing, control and data transmissions while using random access over each phase. DSA MAC protocols can also be classified as single-radio [23], [26], [27], [32], [34] or multi-radio [25], [28-31], [33] depending on the number of transceivers required at each node. For protocols employing a common control channel, having multiple radios is advantageous since it allows a node to continuously monitor the control channel. Table 1.1 summarizes and classifies some of the DSA MAC protocols currently proposed in literature.

1.2 Motivation

With the accelerated development of cognitive radios and DSA techniques for achieving better spectral efficiency, it is desirable to examine the impact of using DSA techniques to realize ad hoc wireless networks. While, the scaling behavior of conventional ad hoc networks is well known, the analysis does not immediately translate to DSA networks. Therefore, the primary objective of our study is to evaluate performance and scalability trends to be expected from ad hoc DSA networks.

We ground our study of the scalability of ad hoc DSA networks in the study of conventional ad hoc networks, to establish a baseline of expected performance. There have been significant developments in the scalability and performance analysis of ad hoc networks in the past [6], [9], [11], [13], [14], [35-40] which we review in Chapter 2. In addition to establishing theoretical performance limits and scaling relationships that are inherent in all ad hoc networks, we also

review some of the practical limitations at the network, MAC, and physical layers. A thorough understanding of the issues related to medium access control is sought, since the biggest change of implementing DSA can be seen at the MAC layer. DSA provides each node with the ability to dynamically utilize the available spectrum and maximize throughput to neighboring nodes by re-using available spectrum. For the scope of this work, we define ad hoc DSA networks as opportunistic, multi-channel, multi-hop, ad hoc networks. Our work differs from multi-channel ad hoc CSMA networks because ad hoc DSA networks that operate in the presence of primary users with strict interference avoidance constraints, and a varying amount of accessible spectrum. We also use detailed sensing models, while most CSMA work assumes perfect sensing within a radius.

DSA networks inherit the design challenges associated with traditional wireless networks such as interference management and route management. DSA networks also bring many new design challenges that cannot be addressed by traditional wireless protocols. DSA networks must be able to perform spectrum sensing, spectrum management, spectrum decision, and spectrum sharing. These additional features of DSA networks are envisioned to allow autonomous operation by networks to increase performance. However, the additional functions of DSA also compete for network resources. In centralized networks, spectrum management is usually performed by the base station which has privileged hardware and software resources. The base station senses the spectrum, collects and processes sensing data from the mobile units to ensure optimal performance. However, in ad hoc networks—which lack a centralized controller—each node must perform spectrum sensing and spectrum management individually and in a distributed fashion. Therefore, the amount of network overhead—defined as the amount of time and resources that are not used for data transfer—associated with the increased amount of cooperation between nodes could become a bottleneck and prevent DSA networks from achieving optimal operation. Therefore, an important motivation of this research effort is also to characterize the additional overhead and signaling costs of employing DSA.

Since DSA is a new paradigm, various forms of DSA implementations can be found in literature as seen in Section 1.1.7. For example, DSA sensing protocols may perform individual or cooperative spectrum sensing, or they may connect to a database to download spectrum availability data based on their geo-location. DSA spectrum management protocols proposed in literature may use a dedicated control channel to aid spectrum management and sharing between secondary users, or use a switchable control channel shared with the primary user. Other DSA networks may not use a control channel at all, since it may suffer as a single point of failure. Some DSA protocols require strict synchronization while others use a hybrid or random access structure. Generally speaking, DSA may be implemented with varying amounts of cognition. There has not been a comprehensive study into the performance and scalability of DSA in ad hoc networks. There is also a need for an evaluation method which can model the expected behavior of such protocols in large networks. In Chapter 4, we develop a flexible analytical framework that may be adapted to varying DSA schemes. We show scalability and performance trends for a DSA scheme well suited for ad hoc networks. Future study using adaptations of our analytical procedure could be used to identify performance trends seen by various implementations of DSA—requiring various levels of cooperation, hardware resources, and cognitive abilities—to offer insight in to the value of the different strategies of deploying DSA.

1.3 Contributions of Thesis

We make the following set of contributions in this thesis:

- Develop a model for energy detection in heterogeneous Poisson DSA networks.
- Develop an analytical procedure for the performance evaluation of ad hoc DSA networks using Markov models.
- Present results and analysis on the performance and scalability of an ad hoc DSA network with one transceiver per node and a dedicated control channel.
- Characterize network performance as a function of the number of DSA nodes, the number of primary users, the number of data channels, and detector accuracy.

2 SCALABILITY OF AD HOC NETWORKS

Ad hoc wireless networks have no centralized entity and a traffic model where each node also serves as a router for other nodes in the network [4-6]. The first phase of this research effort is devoted to analysis and understanding of the scaling behavior of *conventional*, that is non-DSA, ad hoc networks, where communication takes place over a single channel using a single radio for each node. Through a detailed literature study and independent simulations, network models and key performance metrics are evaluated for conventional ad hoc networks. The theoretical results presented in this chapter are heavily based on the seminal paper on the capacity of wireless networks by Gupta and Kumar [9]. Three performance metrics of interest are identified as:

- Network Transport Capacity (C_T)
- End-to-End Throughput per Node ($\lambda(n)$)
- End-to-End Delay (D)

Network transport capacity (C_T)—given in bit·meters/second—is a *measure of the amount of information transported towards its final destination by the overall network at a give instance of time [9]*. Network transport capacity, which may also be referred to as *network capacity*, *transport capacity*, or *capacity*, is a powerful indicator of performance since it is a measure of not only rate but distance as well. The *end-to-end throughput per node* $\lambda(n)$, which may also be referred to as *feasible throughput or throughput capacity*, is the average rate in bits per second at which every node can communicate with its chosen destination [9]. This implies the existence of a scheduling and routing scheme which operates in a multi-hop fashion and may require buffering at intermediate nodes when awaiting transmission. The *end-to-end delay* (D) is a measure of the average delay experienced by one packet after it leaves the source node and before it is received by the chosen destination [35], [36].

This chapter reviews the scaling behavior of the above mentioned metrics, showing the main results from past groundbreaking papers on the scalability limits of ad hoc networks. An extensive review of [9] reveals scaling laws for the capacity of conventional ad hoc wireless networks as the number of nodes n increases in a fixed area. We also examine scalability of

limits of single channel ad hoc networks through a network simulation in Matlab. While unachievable by real networks, an examination of theoretical capacity bounds yields key performance metrics and optimal scaling laws as a function of network density. The effort also illustrates key relationships and network characteristics that directly affect the capacity as the network density increases. These key relationships may serve as guidelines in the design of implementable physical, MAC, and network layer algorithms.

2.1 The Capacity of Wireless Networks

Numerous studies have been conducted in examining the scaling behavior and capacity bounds of ad hoc networks, see [9], [35-40] among others. This section presents a framework for analysis of capacity from the widely referenced paper *The Capacity of Wireless Networks* by Gupta and Kumar [9]. The paper summarizes many important results regarding the scalability of transport capacity and end-to-end throughput per node as a function of increasing network density. End-to-end delay is not explicitly considered in the analysis, but is studied in related work as discussed in Section 2.2.1. Gupta and Kumar also propose a global scheduling and routing scheme that achieves the optimal scaling laws. In this section, we review the system model, present scaling laws, and show intermediate findings as proposed by Gupta and Kumar. Later, we use some of the intermediate results from [9] to arrive at the capacity bounds. The analysis is not a rigorous mathematical proof but an intuitive way to arrive at the scaling laws, the reader is referred to [9] for a full proof.

2.1.1 System Model

Assume a region of area 1 m^2 with n nodes sharing a common wireless channel which allows each node to transmit at *W* bits per second. Two types of networks are considered: *arbitrary networks* and *random networks.* In arbitrary networks, nodes are arbitrarily located in a disk of unit area in the plane, traffic patterns are arbitrarily assigned, and each transmission's range is arbitrarily chosen in order to optimize capacity. In random networks, nodes are randomly distributed and homogenous— therefore each node generates an equal amount of data and all transmissions occur at a common transmission range (or power). Nodes are independently and uniformly distributed either on a disk of area 1 m^2 in the plane or on the surface of a three-dimensional sphere of area 1 m^2 which eliminates edge effects. Additionally, each node chooses a destination node which is the closest node to a uniformly selected point in the region. In both networks, traffic is routed in a multi-hop, decode-and-forward [7] fashion along the source to destination path.

Gupta and Kumar construct the following routing and scheduling scheme in their proof of the lower bound on the end-to-end throughput per node for random networks [9]. The network is tessellated using Voronoi cells with a set of generators which yield the following properties [41]:

- Every Voronoi cell contains a disk of area $100\ \frac{logn}{n}$ with radius $\rho(n)$ and is contained in a disk of radius $\rho(n)$.

- Every Voronoi cell has no more than c_1 interfering neighbors where c_1 depends only on the constant Δ and grows no faster than linearly in $(1+\Delta)^2$. Δ models a guard zone in the protocol model.
- Every node in a cell is within a distance of $r(n) = 8\rho(n)$ from every node in the same cell and adjacent cells.

These properties allow for the creation of a scheduling scheme in which a node in each cell is allotted one slot out of $(1 + c_1)$ slots to transmit to a node in the same cell or any of the adjacent cells. The scheduling scheme assures that each node transmits a fixed number of slots, as the number of nodes in a cell is shown to be bounded by Gupta and Kumar. The route of each packet approximates the shortest path along a straight line from the source node to destination. The packet is relayed in a series of hops from cell to cell in the order in which they intersect a line drawn from source to destination.

2.1.2 Interference Model

Simultaneous transmissions are treated as interference and may cause collisions, which prevent a receiving node from correctly decoding its intended transmission. The successful reception of a transmission over one hop is given by two interference models—the *protocol model* and *physical model*. The physical model requires a minimum signal-to-interference ratio at the receiver, while the protocol model only requires a minimum spatial separation between the intended receiver and all other simultaneously transmitting nodes. Let $d(i,j)$ denote the distance between nodes *i* and *j*, the constant $\Delta > 0$ model a guard zone, N_T be the set of all nodes simultaneously transmitting in the network, and $r(n)$ be the common transmission range for random networks. According to the *protocol model for arbitrary networks*, if node *i* is transmitting to node *j*, then *j* successfully receives the transmission if

$$d(k,j) \geq (1+\Delta)d(i,j), \tag{2.1}$$

for all other nodes $k \in N_T$ and $k \neq i$. Similarly, the *protocol model for random networks* determines a successful transmission if

$$d(k,j) \geq (1+\Delta)r(n), \tag{2.2}$$

for all other nodes $k \in N_T$ and $k \neq i$.

A successful transmission by the physical model for interference requires

$$\frac{\frac{P_i}{d(i,j)^\alpha}}{N + \sum_{k \in N_T} \frac{P_k}{d(k,j)^\alpha}} \geq \beta, \tag{2.3}$$

where P_k is the transmit power of node k, N is the ambient noise power, α is the path loss exponent, and β is the minimum SINR threshold. For random networks, note that $P_k = P$ for all k since all transmissions occur at equal power.

2.1.3 Main Results

Gupta and Kumar's analysis leads to the following scaling laws, which use the notation $f(n) = \Theta(g(n))$ signifying that there are deterministic constants c and c' such that $f(n) = cg(n)$ is feasible with probability one and $f(n) = c'g(n)$ is feasible with probability less than one. The scaling relationships shown below bound the performance of ad hoc wireless networks as the number of nodes increases under the protocol model. Gupta and Kumar also derive similar bounds under the physical model for certain values of α and β and the interested reader is directed to [9]. The following conclusions summarize the main results of Gupta and Kumar's paper:

- "The transport capacity of an arbitrary network under the protocol model is $\Theta(W\sqrt{n})$ bit-meters per second if the nodes are optimally placed, the traffic pattern is optimally chosen, and if the range of each transmission is chosen optimally." [9]
- For random networks, "the order of the throughput capacity is $\Theta\left(\frac{W}{\sqrt{n\,log n}}\right)$ bits per second for the protocol model." [9]

Increasing the network density increases *spatial reuse* and increases the transport capacity of the overall network – but at a much slower rate than the number of nodes. There is an inverse square root relationship between end-to-end throughput and the network size, directly limiting the feasibility of ad hoc networks to small networks that can transport only a fraction of the total link throughput. For example, an ad hoc network of 100 nodes can only guarantee an end-to-end throughput of one tenth of the available link rate. This result implies that conventional ad hoc networks are not scalable.

Gupta and Kumar also show that, if the network domain is scaled from 1 m^2 to A m^2, the bounds on transport capacity scale by $\sqrt{A}$ [9]. Consequently, if the network domain increase proportionally with n, such that the network density remains constant, than *the network transport capacity scales linearly with n but the end-to-end throughput remains constant because the average path length also increases with $\sqrt{A}$*. It is possible to side-step the Gupta and Kumar bounds by limiting communication to nodes within a local area, or connect clusters of nodes through a wire-line network [9]. However, these solutions do not address the root cause of the throughput limitations which we discuss in the next section.

The throughput and capacity show a linear dependence on the maximum link throughput—or equivalently, channel bandwidth—*W*. The scaling laws hold even if the single channel of bandwidth *W* is broken up into multiple sub-channels, whose sum equals *W* [9]. Intuitively speaking, the scaling laws must hold with multiple channels because communication over each one of those sub-channels must follow the $\sqrt{n}$ trend. Nevertheless, the presence of multiple channels may be advantageous in the presence of realistic network and MAC protocols [42]. The end-to-end delay associated with such transactions is not directly examined by Gupta and Kumar, but is expected to be increase with n. Since the transmission ranges reduce with the network density, larger networks contain more relays between source and destination, so the delay is expected to increase as the number of hops between the source and destination increases.

Section 2.2.1 discusses some previous studies examining the delay characteristics of such networks.

2.1.4 Insights into Scalability

In this section, we discuss intermediate results obtained in [9] and [43] that allow for intuitive understanding of the scaling behavior of random networks. Gupta and Kumar make the following observations regarding the average distance from one node to the remaining nodes and the maximum average distance between a node and its nearest neighbor:

- *If source-destination pairs are picked randomly, the mean end-to-end path length* $\bar{L}$*, remains constant as the network density increases*. For a planar network of unit area, $\bar{L}$ is equal to $1/\sqrt{\pi}$ meters [9].

- *As the number of nodes increases, the average separation between the nodes decreases.*
- The *minimum radius for asymptotic connectivity* [43] is

$$r(n) = \Theta\left(\sqrt{\frac{logn}{\pi n}}\right). \tag{2.4}$$

The minimum radius for asymptotic connectivity ensures the absence of isolated nodes with high probability, a condition which Gupta and Kumar consider sufficient to ensure connectivity in the entire network [9]. In a related paper [43], Gupta and Kumar show the transmission range to be lower bounded by

$$r(n) > \sqrt{\frac{logn}{\pi n}}, \tag{2.5}$$

as a graph with edges constrained by $r(n) = \sqrt{\frac{logn+\kappa_n}{\pi n}}$, where κ_n is arbitrary, is connected with probability approaching 1 as $n \to \infty$ if and only if $\lim_{n\to\infty} \kappa_n = \infty$. The scaling behavior of $r(n)$ greatly influences the scaling laws governing the capacity and throughput of ad hoc networks under the Gupta and Kumar model. The mean path length remains constant while r decreases with n, therefore the mean number of hops taken by each flow from source to destination increases on the order of $\sqrt{n/logn}$. Consequently, total load imposed on the network by n nodes, which must be at least $\frac{n\bar{L}\,\lambda(n)}{r(n)}$, also increases. If the load is distributed evenly across n nodes through a fair routing scheme, the mean relaying load per node, which is at least $\frac{\bar{L}\,\lambda(n)}{r(n)}$, increases on the order of $\sqrt{n/logn}$.

The scaling behavior of the nodes' transmission range, reflective of the average distance between each node, directly impacts the capacity of ad hoc networks in two opposing ways. A small transmission range increases transport capacity through higher spatial reuse, yet the connectivity constraint limits how slowly the transmission radius decreases which also limits how quickly

transport capacity can increase. Additionally, since the average path length is fixed, the transmission range determines the minimum number of hops per flow, and in turn the *multi-hop burden*—the relaying load imposed on the network by each flow. As the network scales, the network load due to multi-hop transmission increases faster than the number of nodes, causing the relaying load per node to increase as the network grows. *The relaying load per node lowers the average throughput available per node since each node must dedicate a larger part of its fixed transmission time for relaying information for other nodes.* This is precisely the root cause of the diminishing end-to-end throughput phenomenon seen with the Gupta and Kumar bounds.

2.1.5 Transport Capacity

Recall that transport capacity is a measure of the amount of bits being transmitted and the forward progress they make in a given duration of time towards their final destination. The existence of a globally optimum routing scheme, which ensures that transmissions are forwarded appropriately to ensure information travels towards its ultimate destination, allows for a simple upper bound on transport capacity. The transport capacity is bounded by the product of the number of active transmissions, the link rate, and the average distance from transmitter to receiver $\bar{d}(i,j)$, as proposed in [9]:

$$C_T \leq W * N_T * \bar{d}(i,j). \tag{2.6}$$

Recall that the protocol model excludes any nodes—other than the original transmitter—from transmitting inside an *interference region* of radius $(1+\Delta)r(n)$ centered at the intended receiver. Consequently, there must be a minimum spacing of $(2+\Delta)r(n)$ between each active transmitter. Therefore, disks of radius $\left(1+\frac{\Delta}{2}\right)r(n)$ centered at each transmitter are disjoint [9]. The area consumed by each disk is $\pi\left(1+\frac{\Delta}{2}\right)^2 r^2(n)$, so for a planar network of area 1 m^2 and ignoring edge effects, N_T is bounded by [9]

$$N_T \leq \frac{1}{\pi\left(1+\frac{\Delta}{2}\right)^2 r^2(n)}. \tag{2.7}$$

Since all transmissions occur at a common range $r(n)$ and $\bar{d}(i,j) \leq r(n)$, we arrive at the following upper bound on transport capacity under the protocol model [9]:

$$C_T \leq \frac{W}{\pi\left(1+\frac{\Delta}{2}\right)^2 r(n)}. \tag{2.8}$$

There is an *inverse relationship between the transmission range and transport capacity, illustrating the need to minimize the transmission ranges to maximize spatial concurrency and capacity through a multi-hop transport model in ad hoc networks.* Maintaining network connectivity provides a lower bound for the minimum transmission range, since the transmission range must be large enough to ensure a multi-hop communication link between any two nodes in the network. The minimum radius given by (2.5), as shown in [43], ensures connectivity, therefore the above relation can be written in terms of n as

$$C_T < \frac{W}{\left(1+\frac{\Delta}{2}\right)^2}\sqrt{\frac{n}{\pi \log n}}. \quad (2.9)$$

The above analysis demonstrates the strengths and limitations of multi-hop networks in maximizing capacity through the idea of maximizing spatial reuse. The opposing constraint on achieving infinite spatial reuse is the minimum range of each transmission required to maintain connectivity. Thus, even from this simplistic approach we can derive the upper bounds on transport capacity of a multi-hop network and its relationship to the network density. We can also observe that the order of the transport capacity is proportional to the length of each transmission divided by the area covered by each transmission. The simple yet powerful protocol model maps transport capacity as a function of space, where analysis using only geometric principles can be used to find tight bounds on transport capacity, as shown in [9].

2.1.6 End-to-End Throughput

When considering the end-to-end throughput per node, one may assume that $\lambda(n)$ diminishes because increased interference from adding more nodes reduces the transmission time of each node. On the contrary, if the transmission range for each node is scaled appropriately, Gupta and Kumar—in their derivation of a lower bound on end-to-end throughput per node of random networks—show that the number of interfering neighbors for each node remains steady and bounded by a constant c_1—which depends only on Δ and grows no faster than linearly with $(1+\Delta)^2$. This allows for the creation of a scheduling scheme where in every $(1+c_1)$ slots each node[1] gets one slot to transmit, therefore the transmission rate of each node remains fixed at $\frac{W}{1+c_1}$ as the network grows. As illustrated next, *the end-to-end throughput diminishes because the average relaying load that each node must carry in relation to its own traffic increases, while its transmission rate remains fixed.*

The total traffic carried by each node can be calculated as the product of the end-to-end throughput and the number of routes that a node relays. Gupta and Kumar show that the number of routes intersecting a node is bounded by $c_5\sqrt{nlogn}$, where the constant c_5 does not depend on $n, \Delta \; or \; W$, so the rate at which each node needs to transmit is less than $c_5\lambda(n)\sqrt{nlogn}$ [9]. Since the scheduling scheme allows each node to transmit at a fixed rate, Gupta and Kumar arrive at the following lower bound for end-to-end throughput per node in random networks, where the deterministic constant $c > 0$ does not depend on $n, \Delta \; or \; W$:

$$\lambda(n) = \frac{cW}{(1+\Delta)^2\sqrt{n \log n}}. \quad (2.10)$$

[1] Note that Gupta and Kumar tessellate the network using 2 dimensional cells in their analysis while we simply opt for nodes (points) to simplify the discussion. Our simulations and system model also use nodes instead of cells when creating the scheduling and routing scheme.

The derivation for the upper bound on end-to-end throughput per node relies on the fact that the number of simultaneous transmissions N_T is upper bounded as follows. As shown in [9], since no other receiver can simultaneously receive a separate transmission on the same channel within a distance $\Delta r(n)$ of a receiving node, disks of radius $\frac{\Delta r(n)}{2}$ centered at each receiver are disjoint. The area of each disk is $\frac{\pi\Delta^2 r^2(n)}{4}$, therefore the maximum number of simultaneous transmissions a network of 1 m^2 may support is

$$N_T \leq \frac{4W}{\pi\Delta^2 r^2(n)}. \tag{2.11}$$

If $\bar{L}$ is the mean path length and $r(n)$ is the maximum hop distance, than the mean number of hops taken by a packet is at least $\frac{\bar{L}}{r(n)}$ [9]. For a network of n nodes communicating at $\lambda(n)$ the total traffic served by N_T nodes is at least $\frac{\bar{L}n\lambda(n)}{r(n)}$. Therefore, an upper bound on the end-to-end throughput per node of random networks, as derived in [9], is

$$\lambda(n) \leq \frac{c'W}{\Delta^2\sqrt{nlogn}}, \tag{2.12}$$

where the constant $c' < +\infty$ does not depend on $n, \Delta, or\ W$.

2.2 Related Results

2.2.1 End-to-End Network Delay

El Gamal *et al*. [35], [36] demonstrate the throughput-delay trade-off that exists in ad hoc networks. El Gamal *et al.* observe that the delay is dependent on: 1) number of hops; 2) the transmission range; and 3) node mobility and velocity. Intuitively, we can see that the scaling order of the end-to-end delay must be equal to or greater than the mean number of hops $\left(\sqrt{n\ /\ logn}\right)$ since each hop introduces a delay in the chain. By increasing the transmission range per node, the delay can be reduced since each flow requires a lower number of hops. However, the end-to-end throughput is also reduced since larger transmissions induce more interference in the network. El Gamal's derivation of the scaling relationship is similar to Gupta and Kumar's analysis and relies on a *relaxed protocol model* for collisions that allows for differentiated transmission ranges. The relaxed protocol model is identical to the protocol model for arbitrary networks given by (2.1).

El Gamal *et al*. [35] characterize the optimal throughput-delay trade-off in fixed ad hoc networks as $D(n) = \Theta(n\lambda(n))$, where $D(n)$ is the average delay per packet from source to destination. Note that $D(n) = \Theta\left(\sqrt{\frac{n}{\log n}}\right)$ since $\lambda(n) = \Theta\left(\frac{1}{\sqrt{n\log n}}\right)$. Additionally, it is shown in [36], that using a constant packet size requires buffering at intermediate nodes as the network grows. Diminishing throughput combined with increasing delay greatly impacts the scalability of ad hoc networks. Network delay is very important for any kind of real-time communication. An

increasing delay limits the feasible use of medium or large ad hoc networks to merely sensor networks or other non-critical applications that require low throughputs and can tolerate large delays.

2.2.2 Multiple Channels

Kyasanur and Vaidya [37] analyze capacity limits under the Gupta and Kumar network model with multiple channels and radio interfaces. Most practical multi-channel systems incorporate multiple channels with fewer interfaces than channels, which can further restrict the scalability of ad hoc networks. As shown in [37], *there is always a capacity loss in arbitrary networks whenever there are fewer interfaces per node than channels*. For random networks, if the ratio of channels to interfaces is $\Theta(logn)$ or greater, the optimal single-channel scaling laws can be achieved with multiple channels [37]. In the absence of switching delays, even a single radio interface per node suffices, while a few additional interfaces per node are required in the presence of switching delay.

2.2.3 Other Results

Another important result established in a research study by Grossglauser and Tse [38] is that mobility increases the capacity of ad hoc networks. The previous works cited in this chapter consider only stationary nodes. Allowing for mobility with smart protocols increases capacity and lowers the number of hops necessary for transport by taking advantage of the motion of the nodes. As a trivial example, if the application can tolerate arbitrarily large delays, transmission from source to a random destination can occur in a single hop at the maximum link rate if the nodes simply just queue the transmission until they are near each other.

The studies cited so far also do not consider fading. Kumar and Xue [40] show that the presence of fading may also increase the transport capacity of ad hoc networks. In fact, the results in [40] show that under certain network conditions, the transport capacity scales linearly with respect to the number of nodes, even with independent fading, with or without the presence of channel state information at the nodes. The increase in transport capacity is explained by the fact that ad hoc networks are interference limited.

2.3 Simulations

Our ultimate goal is to examine the scaling behavior of DSA networks. Nevertheless, we find it beneficial to characterize and reproduce the ultimate theoretical performance limits of single channel non-DSA ad hoc networks. The work in this chapter does not consider the impact of overhead required for multiple access control, routing, and power control schemes. We characterize the performance loss associated with realistic MAC schemes in Chapter 3. In order to reproduce the Gupta and Kumar bounds, we developed a time-driven MATLAB simulation that allows for the analysis of an idealized network in which all nodes have complete knowledge of the network topography and state of other nodes.

Recall that Gupta and Kumar assume that all nodes employ a common transmission range in the case of random networks. This assumption is made mainly for ease of analysis and, as discussed

in [39], bidirectional wireless links and silencing protocols such as RTS/CTS in IEEE 802.11 work best with a common transmission range. Since we are interested in simulating theoretical performance limits and aren't using any silencing protocols, we build our simulation model with differentiated transmission ranges. Each node changes its transmission range, per transmission, based on the proximity of the chosen receiver. This change leads to a modified version of the protocol model for successful transmission, as used in [35], given by (2.1). For simulation purposes we set $\Delta = 1$ so that transmissions cause interference to receivers as far as twice the transmission range. We only program our simulations under the protocol model and cite Gupta and Kumar's analysis for proof that the resulting performance should be matched by the physical interference model given by (2.3).

2.3.1 System Model

We assume a planar region of area 1 m^2 with n nodes sharing a single common wireless channel which allows each node to transmit at W bits per second. Node locations are independent and identically distributed in a uniform fashion. Each node randomly chooses a destination to which it wishes to send data at $\lambda(n)$ bits per second. While Gupta and Kumar chose the node nearest to a randomly located point, our simulator selects the destination for each node uniformly from the remaining $n - 1$ nodes.

We simulate a slotted system with perfect synchronization between nodes and zero propagation delay. We assume that one slot is equal to the duration of transmitting one packet at W bits per second. Each node modifies its transmission range before each transmission to be equal to the distance to the intended receiver. Successful transmissions are defined by the relaxed protocol model with the guard zone $\Delta = 1$ for simulation. Transmissions are spatially and temporally scheduled through a global multiple access scheme which randomly schedules non-interfering nodes for transmission in every slot and ensures collisions do not occur.

We define a *flow*, F_s, as the set of nodes which serve as relays for the transfer of data from the source node s to final destination node d: $\{F_s \epsilon\, s, s_1, s_2, \dots, s_i\}$. A routing table is generated for each flow through the use of *nearest neighbor sets.* The nearest neighbor set for a node is determined as nodes which are connected by lines on a Delaunay triangulation of the network. This method ensures that the entire network is connected since the Euclidean minimum spanning tree of a set of points is a subset of the Delaunay triangulation of those points [41]. Note that in a real network, constricting immediate transmissions inside nearest neighbor sets may not be an optimal use of resources; however, this approach is sufficient for our purposes to identify scaling behavior.

For routing, we implement a scheme similar to Gupta and Kumar's where routes approximate near straight lines drawn from source to destination. Since we do not tessellate the network into cells, we follow the simple routing strategy proposed in [44], in which each node is able to determine the nearest neighbor which is closest to the final destination and forward the traffic accordingly. Packets hop from one node to the nearest neighbor which is closest to the final destination, ensuring that the shortest and straightest possible route is selected. Consequently, forward progress is made with each hop and the total distance traveled by each packet is minimized.

The simulator generates $\mathcal{M}$ network instantiations of n nodes uniformly distributed over the network domain. Each simulated network instantiation m is run for T time slots. We assume that each node has the same amount of information to send and that there is an infinite backlog of packets at each node—for its own flow and those that it is relaying. For each time slot τ, a set of nodes are selected to transmit by an *ideal* global multiple access scheme. The ideal scheme is designed to 1) allow all nodes a statistically fair chance to transmit, 2) ensure there are no collisions, and 3) ensure high spatial reuse by maximizing the number of active transmissions. All nodes desire to transmit at every time slot, but only a subset of nodes is selected to transmit such that there are no collisions. The set of transmitters is chosen uniformly from non-interfering nodes, and maximized so turning any inactive node on would cause a collision. This random selection scheme ensures that each node gets an equal chance to transmit on average. Each node then selects a packet out of its parallel queues with equal probability, which could be its own data or data that it is forwarding for another flow. This process happens in an iterative manner, but in zero time, at the beginning of each time slot. Appendix A further discusses the scheduling scheme.

For each time slot, the range of each active transmission is recorded and summed to bound the transport capacity $C(m,\tau)$ achieved by the network. The sum is an upper bound since the true transport capacity only reflects the forward progress made from source to destination with each hop [9], [39]. Let $N_T(m,\tau)$ be the number of active transmissions in time slot τ and network instantiation m, and r_i is the receiving node chosen by node i, then the transport capacity achieved in one time slot for one instantiation of the simulation is bounded by the sum

$$\mathrm{C}(m,\tau) \leq W \sum_{k}^{N_T(m,\tau)} d(i,r_i). \tag{2.13}$$

After running the simulation for T time slots per $\mathcal{M}$ network instantiations, the final network transport capacity $C(n)$, bounded by average of the transport capacity achieved during each time slot and instantiation, is

$$C(n) \leq \frac{1}{T}\sum_{m}^{\mathcal{M}}\sum_{\tau}^{T} C(m,\tau). \tag{2.14}$$

The flow and the relay number which was activated are also stored to calculate the average end-to-end throughput $\lambda(m)$ achieved per flow for the given network instantiation. Gupta and Kumar define the feasible throughput $\lambda(n)$ such that in every time interval of $\mu < \infty$ every node can send $\mu\lambda(n)$ bits to its corresponding destination [9]. This implies that every node in the chain from source to destination must transmit at least $\mu\lambda(n)$ bits during the time interval μ. Since we opt for a random scheduling strategy instead of a fixed one, and only attempt to ensure equivalent transmission time for each node on average, we calculate the feasible throughput for a flow $\lambda(s)$ as equivalent to the *average* time that each relay node in the chain dedicates to a flow. Let $\delta(F_s)$ represent the average number of slots that each relay in the flow F_s was active over one network instantiation of T slots. The end-to-end throughput for a given network instantiation is the average number of slots that were utilized for transmission by each of the nodes relaying for each flow, divided by T:

$$\lambda(m) = \frac{1}{T} \sum_{s}^{n} \delta(F_s). \tag{2.15}$$

The final end-to-end throughput achieved over $\mathcal{M}$ network instantiations of n nodes is

$$\lambda(n) = \frac{1}{\mathcal{M}} \sum_{m}^{\mathcal{M}} \lambda(m). \tag{2.16}$$

2.3.2 Analytical Approximations

In this section, we derive approximations to the results expected from simulations. We base our analysis on Gupta and Kumar's work and make changes where necessary to adapt to our simulation model. Recall that the critical radius for asymptotic connectivity, given by (2.5), was identified by Gupta and Kumar as a bound for the common transmission range. We allow for differentiated transmission ranges in our simulations, but the concept of the critical radius helps in approximating the average transmission range. Our system model limits communication links in between nodes in the nearest neighbor sets, which are determined as nodes connected by edges of a Delaunay triangulation of the network. Figure 2.1 illustrates the simulated scaling behavior of the distance between neighboring nodes as n increases from 10 to 500. For $n > 50$, the average distance to a nearest neighbor closely follows the minimum value of Gupta and Kumar's definition of the critical radius for connectivity. The average distance to the furthest nearest neighbor is seen to be approaching, but consistently higher than the critical radius. The average distance to the closest nearest neighbor and the average nearest neighbor are seen to fall slightly below the critical radius, thus validating Gupta and Kumar's choice for a common transmission range. Therefore, we take the expected *average hop distance* for a network of n nodes to be

$$\bar{d}_H(n) = \sqrt{log\, n\,/(\pi n)}. \tag{2.17}$$

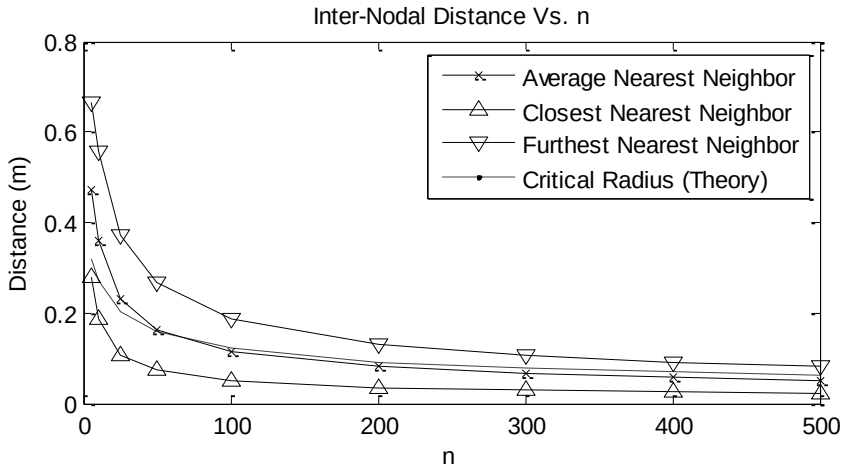

Figure 2.1 Nearest neighbor dynamics.

The distribution of the hop distance is also necessary to estimate the expected number of simultaneous transmissions. The distribution of the hop distance seen in simulation, $f(d_H; n)$, is approximated by a Rayleigh distribution with Raleigh parameter $b(n) = \bar{d}_H(n)\sqrt{2/\pi}$ conditioned on n. The approximation matches simulation well, as shown by the histogram of simulated values and the Rayleigh probability density function for $n = 100$ nodes in Figure 2.2.

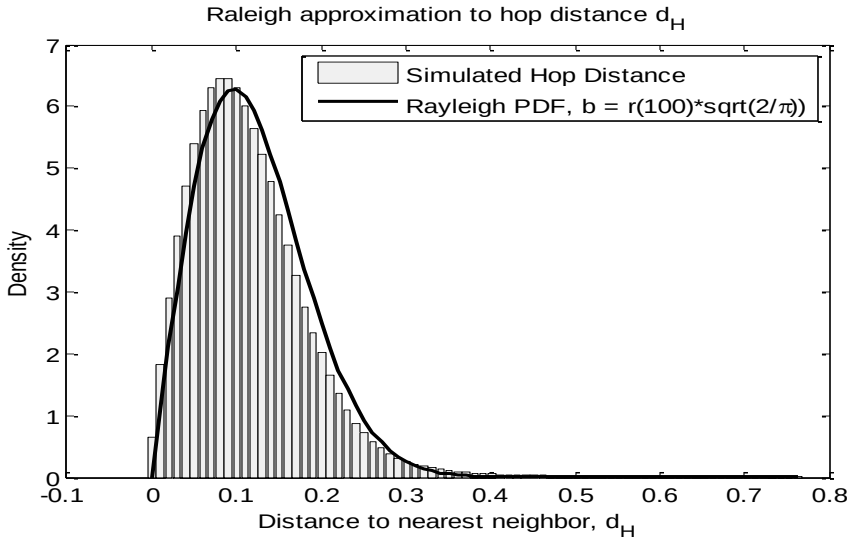

Figure 2.2 Simulated and approximated distribution of hop distance, *n* = 100.

The average number of simultaneous transmissions expected to be seen in simulations is approximated from the distribution of d_H and the protocol model. If two nodes forming a link are a distance d_H apart, there can be no other transmitter inside a radius of $(1 + \Delta)d_H$ of the receiver, and no other receiver inside a radius of $(1 + \Delta)d_H$ of the transmitter. Therefore, the minimum area consumed by a successful transmission between two nodes a distance of d_H apart is the *area of intersection* between two circles of radius $(1 + \Delta)d_H$ with centers d_H meters away. The area of this region is [45]

$$A(d_H) = 2(1 + \Delta)^2 {d_H}^2 q\left(\frac{1}{2(1 + \Delta)}\right), \tag{2.18}$$

where $q(t) = \arccos(t) - t\sqrt{1 - t^2}$. Consequently, the maximum number of simultaneous transmissions of length d_H in a network area of 1 m^2 is bounded by

$$N_T(d_H) \leq \frac{1}{A(d_H)}. \tag{2.19}$$

The effect of differentiated transmission ranges is taken into account by integrating $N_T(d_H)$ over the distribution of d_H:

$$\bar{N}_T(n) \leq \int_0^\infty \frac{1}{A(d_H)} f(d_H; n) dd_H. \tag{2.20}$$

In the simulation, the bound for transport capacity of a single network instantiation is found as the sum of each transmission occurring in the network over a single slot. This value is averaged over multiple time slots, traffic iterations, and network instantiations. The transport capacity is approximated in a similar manner as the number of simultaneous transmissions, noting that for a fixed distance between nodes of d_H the transport capacity is $C_T(d_H) \leq W * d_H * N_T(d_H) \leq \frac{W*d_H}{A(d_H)}$. Therefore, the average transport capacity expected is upper bounded by

$$\bar{C}_T(n) \leq \int_0^{\infty} \frac{W d_H}{A(d_H)} f(d_H; n) dd_H.^2 \tag{2.21}$$

The end-to-end throughput for a single flow is calculated in the simulation by averaging the number of slots each relaying node is able to utilize to forward information for the given flow. It is assumed that each node monitors the flows it is relaying for—including its own—and randomly chooses one of them at each slot where it is active, determined by the probability of transmission. The theoretical end-to-end throughput can be derived from transport capacity as shown in [9]. The transport capacity must be shared by n nodes trying to transmit at $\lambda(n)$ over a distance $\bar{L}$, where the mean path length $\bar{L}$ only depends on the network domain. Therefore, the average end-to-end throughput per node is approximated as

$$\bar{\lambda}(n) = \frac{\bar{C}_T(n)}{n\bar{L}}. \tag{2.22}$$

For comparison, if a common transmission range of $r(n) = \sqrt{logn/\pi n}$ is used each active transmitter must be separated by a distance of ($r_T(2 + \Delta)2 + \Delta)r(n)$. Therefore, disks of area $\pi(1 + \frac{\Delta}{2})^2 r(n)^2$ centered at each transmitter must be disjoint [9]. Then, the maximum number of simultaneous transmissions is bounded by

$$\bar{N}_T(n) \leq \frac{1}{\pi\left(1 + \frac{\Delta}{2}\right)^2 r(n)^2}. \tag{2.23}$$

Assuming that the PDF of transmissions chosen by the scheduling scheme follows the PDF of the hop distance d_H, the transport capacity achieved is bounded by

$$C_T(n) \leq \int_0^{r(n)} \frac{W * d_H}{A\left(d_H \pi \left(1 + \frac{\Delta}{2}\right)^2 r(n)^2\right)} f(d_H; n) dd_H. \tag{2.24}$$

In Appendix A, we compare the average hop distance against the average transmission length, as chosen by the scheduling scheme. The bound for end-to-end throughput per node is calculated from (2.24) in the same manner as (2.22).

---

[2] When numerically calculating the integral, the lower limit is changed to $r(n)/10$ since allowing for vanishingly small transmissions leads to infinitely many simultaneous transmissions.

2.3.3 Simulation Results

The simulation results confirm the trends in node spacing, the number of hops per flow, the number of simultaneous transmissions, transport capacity, and end-to-end throughput presented in Section 2.1. First, we confirm the validity of the routing scheme and the behavior of the multi-hop burden as the network scales. Figure 2.3 illustrates the average and maximum number of hops required by each simulated flow from source to destination. The mean number of hops per flow matches closely with the expected number of hops per flow:

$$\frac{\bar{L}}{\bar{d}_H(n)} = \sqrt{\frac{n}{log\ n}}. \tag{2.25}$$

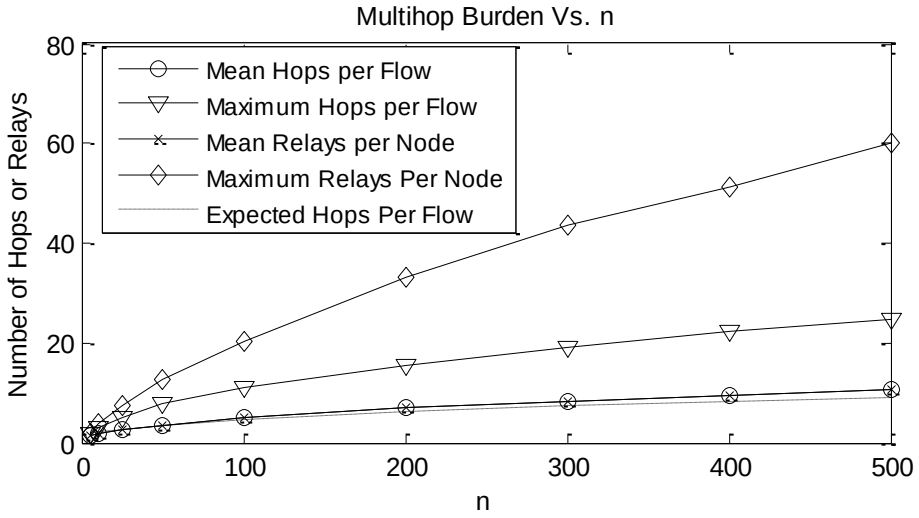

Figure 2.3 Multi-hop burden vs. *n*.

Additionally, the average number of flows that a single node relays (mean relays per node), a good indicator for the multi-hop burden, is found to be identical to the average number of hops per flow. This observation clearly illustrates that a multi-hop transmission strategy directly determines the relaying burden on each node. When compared to the maximum number of hops per flow, a much higher value for the maximum number of relays per node is seen and may indicate the presence of hot spots since our simulation uses a planar network. As discussed by Gupta and Kumar [9]—in planar networks, hot spots can build up near the center of the domain since several source-destination pairs may route traffic through the center. However, the scaling behavior of capacity and throughput of such networks remains unchanged as it is still dominated by interference (spatial reuse) and the multi-hop burden.

As discussed in Section 2.3.2, the number of active transmissions is bounded by equations (2.20) and (2.23) for differentiated and common ranges, respectively. Figure 2.4 compares the bounds with the average number of simultaneous transmissions seen in simulations. The number of simultaneous transmissions seen in simulations exceeds the theoretical bound for common

ranges, while coming under the bound for differentiated the form of differentiated transmission ranges allows twice as many nodes to transmit simultaneously than the maximum possible value for common transmission ranges.

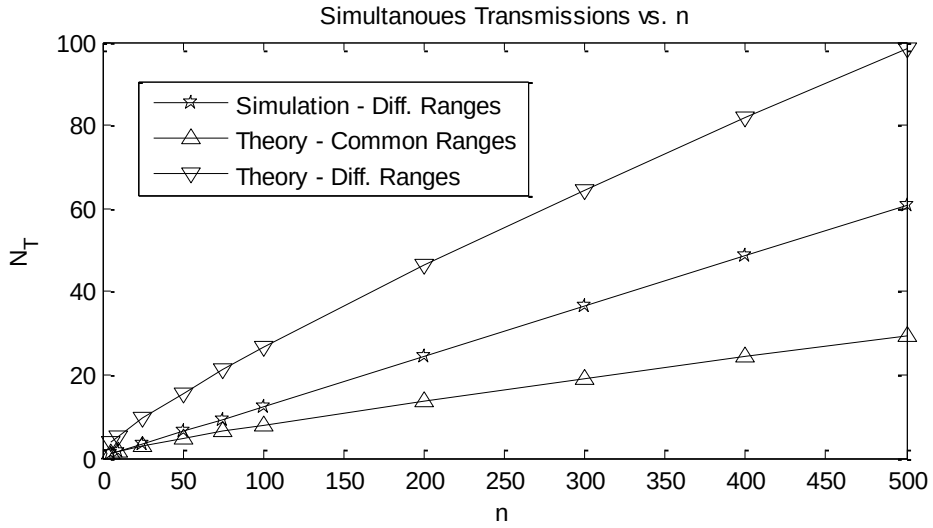

Figure 2.4 Simultaneous transmissions.

Figure 2.5 illustrates the capacity and throughput results obtained from our Matlab simulations. Since we assume that each slot is equal to one packet, the transport capacity is presented in normalized terms of packet-meters per second and end-to-end throughput is normalized in terms of packets per second. The transport capacity achieved is close to but less than the theoretical bound for differentiated ranges given by (2.21) and (2.22) , as it is ideal and assumes 100% spatial reuse. We do observe the same scaling behavior for simulation and the approximation.

As seen in Figure 2.5, going from five to five hundred nodes only increases the transport capacity by a factor of five, and limits the end-to-end throughput to near zero. The Gupta and Kumar bounds are shown to hold and reached in spite of the modifications made to the system model. The transport capacity achieved is well approximated by the curve $\frac{1}{\pi r(n)(1+\frac{\Delta}{2})^2}$, illustrating the Gupta and Kumar bound on transport capacity, $C_T(n) = \Theta\left(\sqrt{\frac{n}{log\, n}}\right)$ was reached by the simulations. The end-to-end throughput per node is also seen to diminish as the number of users increases. The simulated throughput and the theoretical approximation match closely though the simulation seems to perform better than the approximation for large n. This may be because the average path length seen in simulation is slightly less than that predicted by Gupta and Kumar: $\bar{L}_{sim} \approx 0.53 < \bar{L} = \sqrt{1/\pi}$. The results also confirm that our form of medium access control and routing methods were able to closely meet the approximated theoretical bounds. The results show the validity of our simulation code and provide a baseline for optimal scaling that can be achieved by multi-hop networks. The simulation code will be modified in Chapter 3 to characterize the scalability impact of realistic MAC schemes for ad hoc networks.

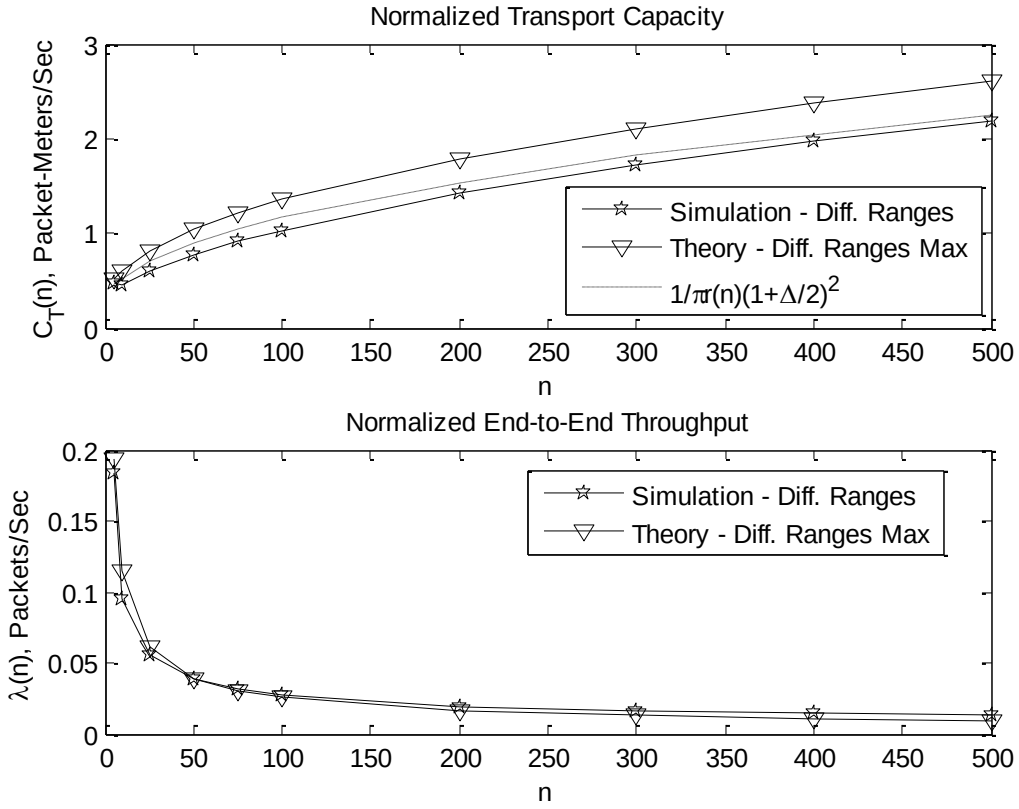

Figure 2.5 Simulated and theoretical scaling laws.

2.4 Conclusion

We have studied theoretical bounds on the capacity and throughput of conventional ad hoc networks. We have examined literature on capacity bounds and scaling behavior under the current wireless network paradigm—which treats unwanted transmissions as interference. It was shown that the currently accepted form of cooperation in an ad hoc network—a multi-hop, decode-and-forward, transport model—suffers from major scalability issues. Network transport capacity is not order optimal with the number of nodes and the end-to-end throughput diminishes to zero as the number of users increases. The scalability issues facing such wireless networks have important implications as wireless devices continue to proliferate in higher densities, and question the viability of visions of ubiquitous wireless networks and pervasive computing. The work presented in this chapter meets the goal of characterizing the ideal scaling behavior of conventional ad hoc networks. Our focus now shifts towards the scaling behavior of realistic ad hoc networks. We identify scalability issues in protocols for medium access in conventional ad hoc networks in Chapter 3.

Our ultimate research goals, and motivation for this study, are to characterize the scalability and performance issues of ad hoc DSA networks. To that end, all indications are that DSA alone is not enough to resolve the theoretical scalability issues discussed in this chapter. The scaling behavior of network transport capacity witnessed is interference limited, and the diminishing

behavior of the end-to-end throughput results from interference and the multi-hop burden due to the decode and relay form of cooperation. Nevertheless, advances in radio technology such as cognitive and software defined radios and shifts in spectrum policy towards unlicensed bands and dynamic spectrum access are steps in the right direction. DSA introduces a new paradigm for spectrum access and spectrum sharing among nodes, and addresses many of the challenges facing current wireless networks, such as inefficient spectrum utilization, and complicated network and spectrum planning [16]. DSA also provides a mechanism for interference avoidance but DSA alone does not address the root causes of the scaling issues facing ad hoc wireless networks—which are limitations in the multi-hop transport model.

3 SCALABILITY OF RANDOM ACCESS MAC

In Chapter 2, we developed system and simulation models for scalability analysis of single-channel ad hoc networks under ideal power control and multiple access schemes. Following [9] for random networks, it was shown that the network transport capacity scales no faster than $\sqrt{n/logn}$, and the end-to-end throughput diminishes with n. As Gupta and Kumar observed in [9], "we have not considered in this paper the *additional burden in coordinating access to the wireless channel*, and the additional burden caused by mobility and link failures and the consequent need to route traffic in a distributed and adaptive way. These can only further throttle capacity." The "additional burden" referred here is due in part to overhead and sub-optimal behavior of realistic routing, MAC, and power control protocols. For example, achievable capacity may be reduced due to sub-optimal power control which would cause retransmissions and unnecessary interference. Likewise, sub-optimal routing schemes may lead to unfair load balancing, longer path lengths, and longer delays.

Sub-optimal medium access behavior may manifest itself in the form of *collisions* and unnecessary prevention of transmissions. A collision occurs at a receiver when the level of interference (resulting from unwanted transmissions) prevents the receiver from successfully decoding a transmitted packet. A transmission may also result in a collision if the intended receiver is not ready for the transmission, as is the case if the node is transmitting to or receiving from another source. A medium access protocol may also behave sub-optimally by unnecessarily preventing transmissions that would otherwise be successfully received without causing interference to other simultaneous transmissions. This chapter considers the scalability and performance impact of random MAC schemes on ad hoc networks.

We examine random access MAC schemes which allow users to access the channel in an uncoordinated or semi-coordinated manner suitable for general ad hoc networks. Three basic forms of random access—Aloha, slotted Aloha and CSMA—are analyzed for scalability and performance in this chapter. There are more sophisticated MAC schemes for ad hoc networks,

such as the silencing RTS/CTS protocol for 802.11 networks [11]. However, we observe that more sophisticated distributed MAC schemes also rely on random elements to coordinate transmissions. Therefore, we study random access techniques as a worst case scenario to find an upper bound on the scalability impact of realistic MAC protocols in ad hoc networks.

Furthermore, a study of the scalability issues at the MAC layer may provide insight to the expected behavior of DSA, since the two share much in common. Medium access control manages users' transmissions and channel access on a micro—per packet—level, while DSA manages the users' access to the spectrum on a macro level. In Section 3.2, we characterize the performance impact of the Aloha and slotted Aloha protocols, in which collisions are resolved through retransmissions. In Section 3.3, we characterize the performance impact of a more realistic form of random access: Carrier Sense Multiple Access (CSMA). CSMA relies on the node's ability to sense energy on the channel and withhold transmissions that may cause collisions. It is important to note similarities in the operating principles behind CSMA and DSA. CSMA attempts to opportunistically coordinate access for a single channel while DSA attempts to opportunistically coordinate access to the overall spectrum. Therefore, evaluating and characterizing the performance of CSMA is an important motivation for this chapter.

The simulation model developed in Chapter 2 is expanded here to simulate random access behavior. The traffic model in the previous chapter assumed a backlog at every node and an *ideal MAC* scheme scheduled transmissions in every slot such that spatial reuse was maximized without causing any collisions. It is well known that random access MAC protocols are sensitive to the network load, a measure of the number of nodes desiring to access the channel at any given time [11]. To simulate performance under a varying network load, we assume that each node desires to access the channel at any given slot with a common *medium access probability* p_T between 0 and 1. Since we previously assumed that each node desires to transmit with probability one in each slot, in Section 3.1, we repeat the ideal MAC simulations while varying p_T to provide a baseline for the random access schemes. Four variations to the system model described in Section 2.3 are presented in this chapter:

- Ideal MAC with varying probability of transmission
- Slotted Aloha
- Aloha
- CSMA

3.1 Ideal MAC with Probability of Transmission

3.1.1 System Model

As discussed in Chapter 2, the network and interference model used in simulations are derived from Gupta and Kumar's analysis of the capacity of wireless networks [9]. The simulation models a network of unit area with *n* nodes. The node locations are randomly generated and communication is limited between *nearest neighbors*. A simple and efficient straight line routing scheme is developed where each node relays information to the node out of its nearest neighbor sets which is closest to the final destination. We use a relaxed protocol model to model interference with differentiated transmission ranges which reflect ideal power control. A backlog of packets is assumed at every node and transmissions are scheduled to avoid collisions and maximize spatial reuse.

The original scheduling scheme presented in Section 2.3.1 is kept intact in this section except for the following modification: at each time slot, a Poisson traffic model with probability p_T narrows the set of nodes that desire to access the channel. Each node generates a uniform number μ between 0 and 1, if and only if $\mu \leq p_T$ then the node desires to transmit in the given slot. Collisions can still occur if all of the nodes with traffic are allowed to transmit as in Aloha and slotted Aloha. Out of the subset of nodes with traffic, the ideal MAC scheme uniformly selects non-interfering nodes for transmission as described in Section 2.3.1. Those nodes which would collide with its transmission are removed from the subset of potential transmitters, and another node is selected at random from those that remain. The process is repeated in zero time, at the beginning of each slot, until all nodes with traffic are either transmitting or silenced so other nodes can transmit successfully.

3.1.2 Simulation Results

Figure 3.1 shows the simulated network transport capacity and end-to-end throughput per node for an ideal MAC with medium access probability from 0 to 1, parameterized by the number of users (*n*). Increasing the medium access probability increases the simulated performance because starting with a larger set of nodes desiring to access the channel allows the ideal MAC scheme to schedule more parallel transmissions. The increase is rapid for lower p_T and becomes marginal for $p_T > 0.5$ which implies that the network is nearly fully loaded at $p_T = 0.5$. Once the network is operating near full capacity, more transmissions could decrease network capacity by causing additional interference.

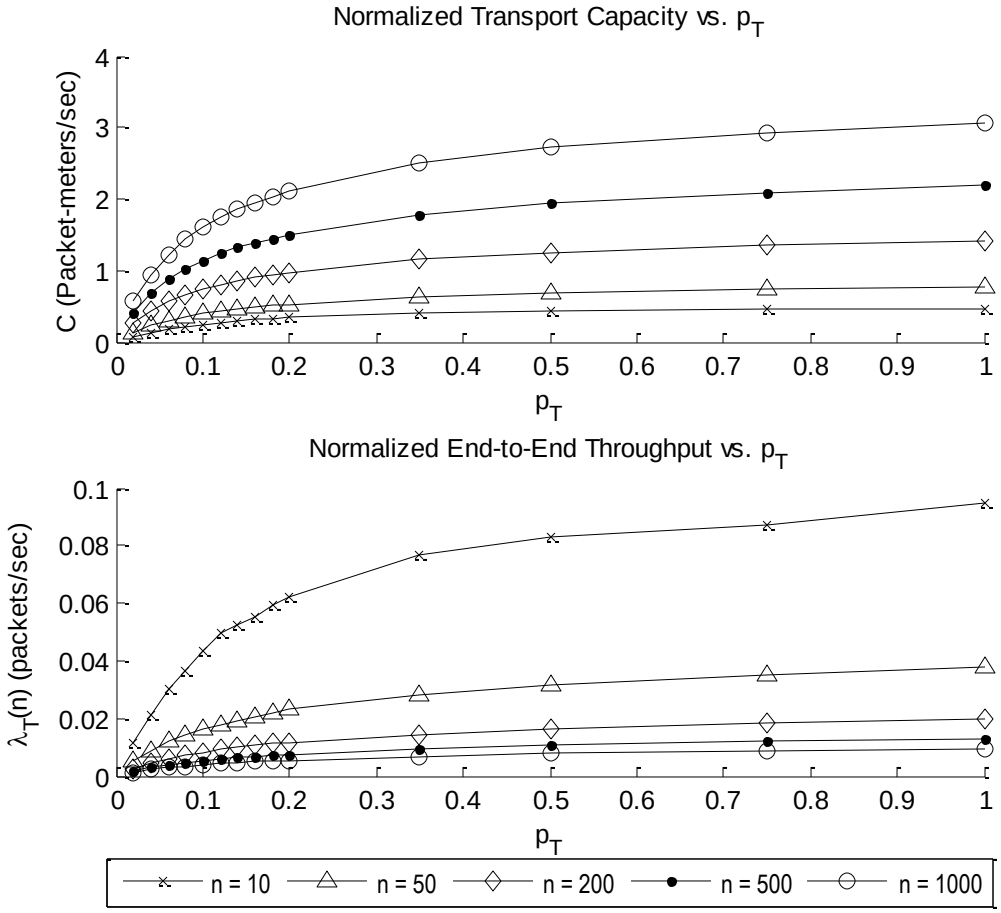

Figure 3.1 Transport capacity and end-to-end throughput vs. p_T, Ideal MAC.

Figure 3.2 shows transport capacity and end-to-end throughput as a function of n parameterized by selected values of p_T. The scaling behavior with respect to n is consistent for all values of medium access probability. The analytical upper bounds for the average transport capacity and end-to-end throughput, given by (2.21) and (2.22), respectively, are also plotted for comparison. The transport capacity scales no faster than $\sqrt{n/logn}$ and the end-to-end throughput diminishes as the network density increases.

Figure 3.3 shows the average number of simultaneous transmissions seen per slot $\bar{N}_T$ and the average length of each transmission for changing values of p_T. As expected, the number of simultaneous transmissions increases with p_T for a given number of nodes. Figure 3.3 also shows the mean hop distance $\bar{d}_H$ to be largely independent of p_T. This is because the routing scheme generates routes and selects relays based on the network topology and nearest neighbor sets, while the scheduling scheme uniformly chooses nodes and relays for transmission. The average hop distance decreases slightly as the network becomes loaded, since smaller transmissions lead to more spatial reuse.

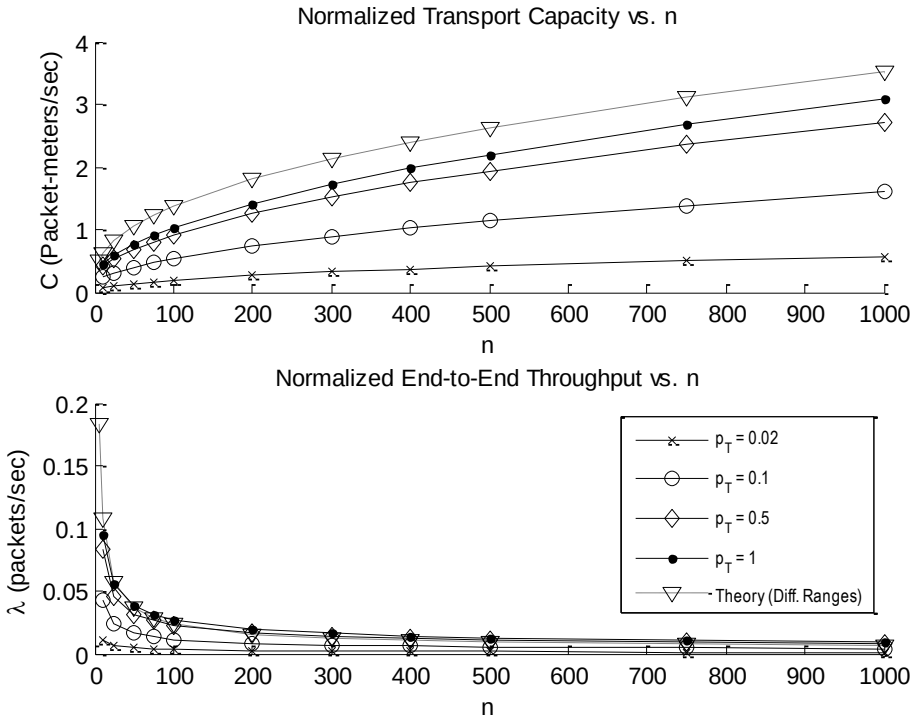

Figure 3.2 Transport capacity and end-to-end throughput vs. *n*, Ideal MAC.

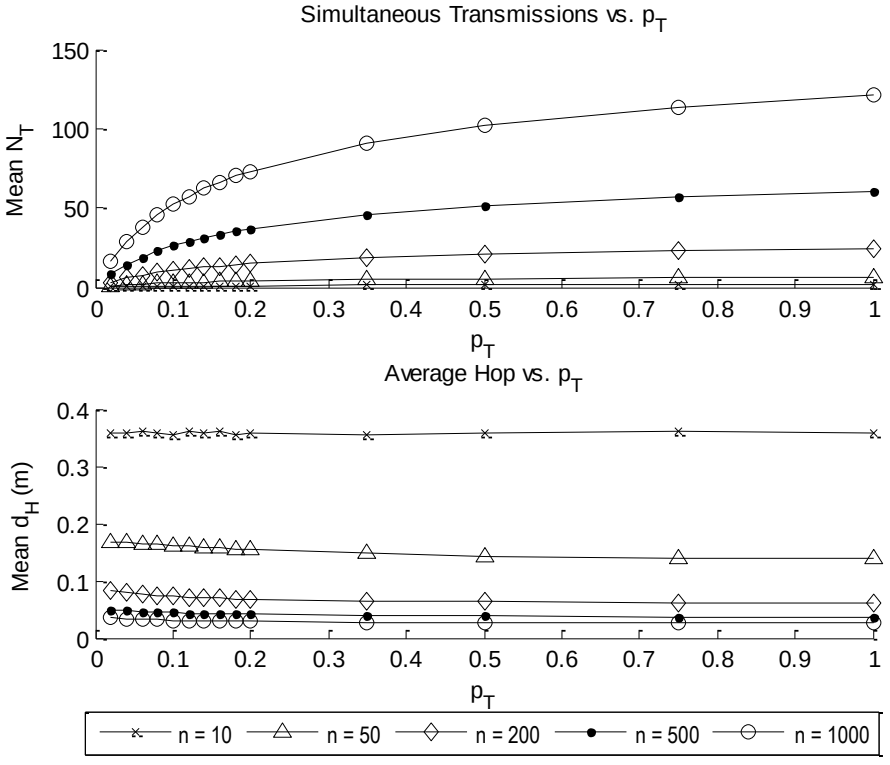

Figure 3.3 Simultaneous transmissions and transmission length vs. p_T, Ideal MAC.

Figure 3.4 presents $\bar{N}_T$ and $\bar{d}_H$ as a function of the number of nodes n parameterized by select values of p_T. As seen in Section 2.3.3, the average number of simultaneous transmissions increase linearly with *n,* and exceed the theoretical bound for common transmission ranges, given in (2.23), for high medium access probability. As observed in the lower plot, the average hop distance decreases with the number of users and follows the same general trend of $r(n) = \Theta\left(\sqrt{logn/\pi n}\right)$. The average hop distance seen in simulation is smaller than the critical radius because of the use of differentiated transmission ranges, as discussed in Appendix A.

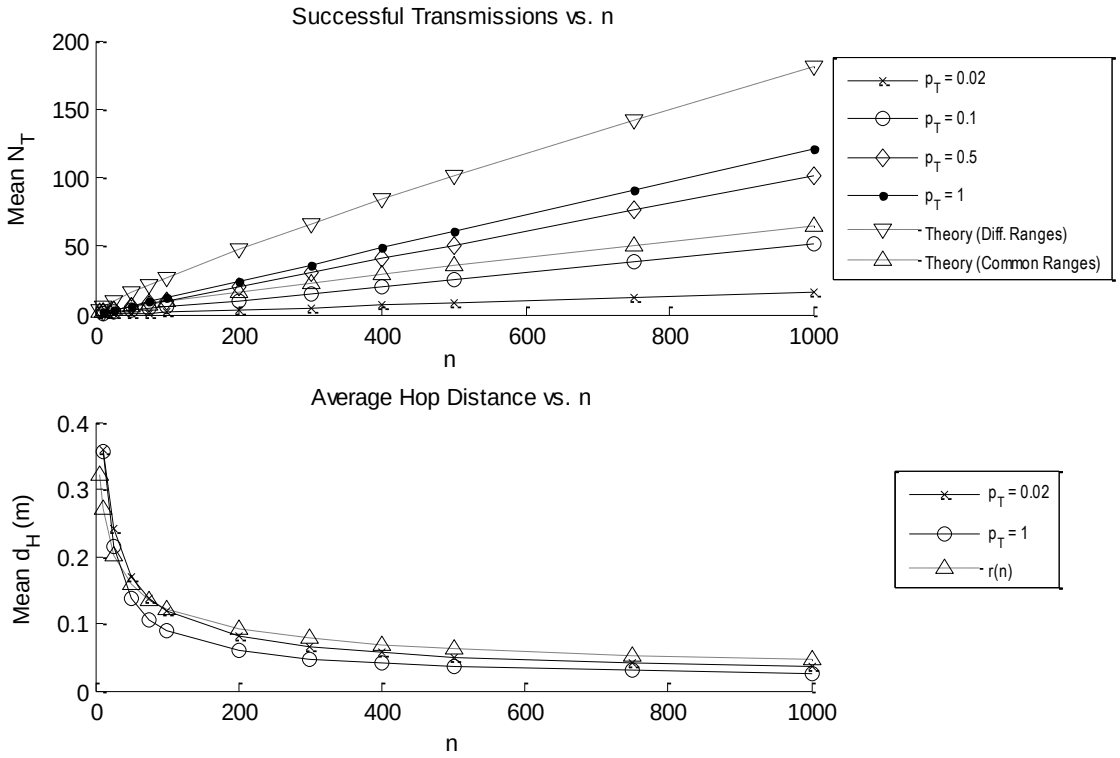

Figure 3.4 Simultaneous transmissions and transmission length vs. *n*, Ideal MAC.

To gain insight into the network behavior, we derive a metric for spatial use to indicate the total network area which is occupied by transmissions. The total "transmission area" is calculated from the average number of simultaneous transmissions and the average hop length, given by

$$A_{Tx} = \bar{N}_T \pi \bar{d}_H{}^2. \tag{3.1}$$

Recall that each transmission requires a guard zone of $(1 + \Delta)r_T$, where r_T is the transmission range and Δ models a guard zone. Therefore, each transmission is surrounded by an "interference area" where no other receiver may be present. The interference area is

$$A_\Delta = N_T(1 + \Delta)^2 \pi \bar{d}_H{}^2 - N_T \pi \bar{d}_H{}^2. \tag{3.2}$$

The total network area consumed is the sum of the transmission area and the interference area and gives an indication of the network area that is "spent" on transmissions. Figure 3.5

illustrates the spatial use seen in simulations for select values of p_T and n. Because of edge effects, the total area exceeds 1 m^2 for simulations with low n, where longer transmissions occur frequently. The bottom plot shows that the simulations maximize spatial reuse as the medium access probability increases. Also, over 90% of the total network area is consumed at $p_T = 0.5$, justifying the earlier observations that there is only marginal improvement in capacity for $p_T > 0.5$.

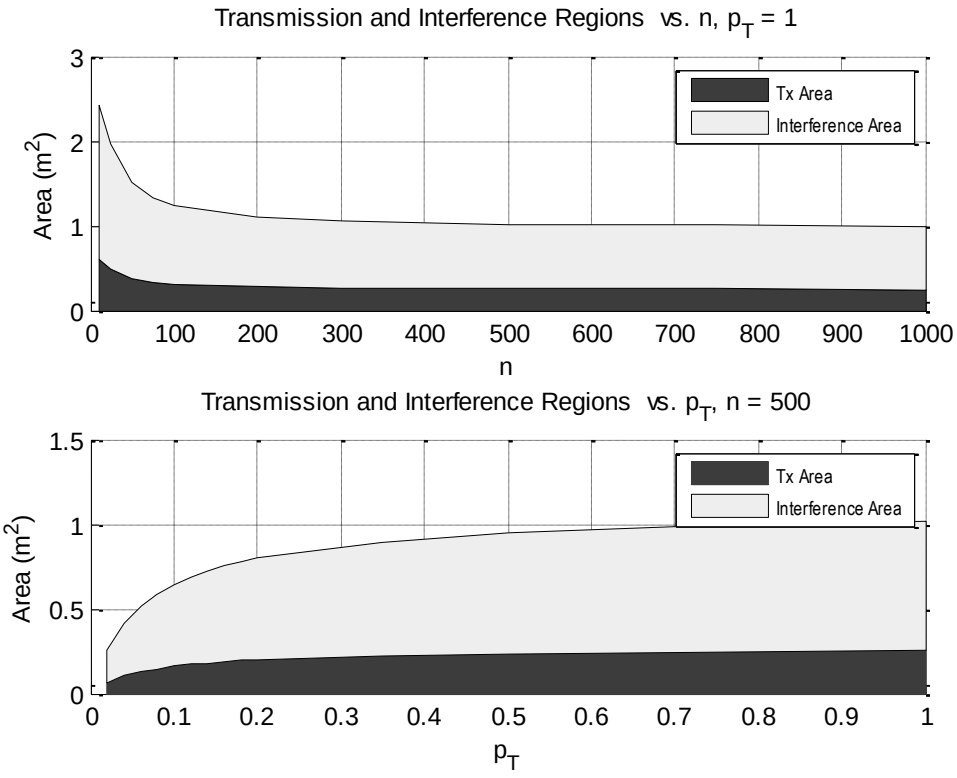

Figure 3.5 Spatial use, Ideal MAC.

Figure 3.6 plots the average number of packets sent per node, in 1000 slots, with respect to the network density (top) and the medium access probability (bottom). As discussed in Section 2.1.6, Gupta and Kumar create a scheduling scheme where each node is guaranteed a fixed transmission rate [9]. Observe that the ideal MAC scheme also behaves the same for $n \geq 25$, as the average number of slots allotted to a node stays constant as the network density increases. The simulations show that out of 1000 slots, each node was able to utilize an average of 120 slots for a successful transmission under the ideal MAC scheme and $p_T = 1$. Also, the number of slots does not increase linearly with the medium access probability and instead plateaus after interference builds in the network. Figure 3.6 clearly illustrates that with ideal power control and MAC the transmission rate of each node stays fixed as the number of users increases. The resulting *decrease in end-to-end throughput can only be explained by the fact that the multi-hop burden carried by each node increases as n increases.*

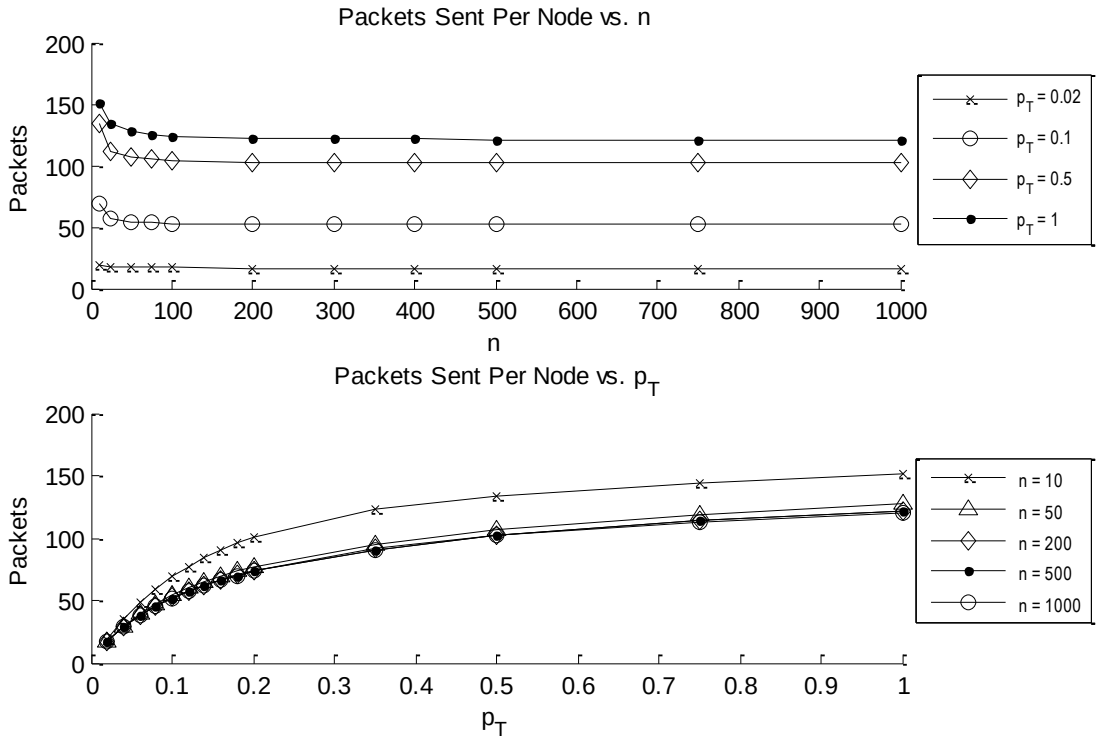

Figure 3.6 Packets sent per node, Ideal MAC.

3.1.3 Discussion

This section modified the ideal MAC scheduling scheme to simulate a varying network load. We introduced the medium access probability p_T as a variable to limit the number of nodes which desire to access the channel in a given slot. The ideal MAC scheme completely avoids collisions; therefore, it performed best at high p_T values. We also observed that the ideal MAC scheme maximizes spatial reuse and allows for a steady transmission rate per node as the network scales.

3.2 Slotted Aloha and Aloha

The first random access MAC schemes to be studied are Aloha and slotted Aloha. As discussed in Chapter 1, the Aloha protocol allows unrestricted transmission and resolves collisions through retransmissions. If nodes are synchronized, slotted Aloha can be used for better performance. Throughput analysis of the Aloha protocol is straightforward in a *single-hop network* where all nodes can hear each other. The performance for single-hop networks under a Poisson traffic model is highly dependent on the traffic load and the network density [11]. The throughput S of an Aloha network with n nodes with medium access probability p_T is $S = np_Te^{-2np_T}$, where S is defined as the average number of packets successfully transmitted during the time it takes to send one packet [11]. The throughput S of an equivalent slotted Aloha system is defined by a similar equation except the probability of a collision is reduced by ½ since the nodes are

synchronized. The throughput of a slotted Aloha system is $S = np_T e^{-np_T}$ [11]. The maximum throughput achieved by a node in the Aloha system in a single-hop network is 18.4% of the link rate, and doubles to 36.8% for the slotted case [11]. There is an inverse exponential relationship between throughput and the product of the number of users n and the medium access probability p_T. The Aloha and slotted Aloha protocols are vulnerable to become overloaded and inefficient under a heavy network load in a single-hop network.

In multi-hop networks, communication is limited to nearest neighbors which leads to multiple collision domains where only a subgroup of nodes can hear each other. Consequently, the analysis for the operation of Aloha in multi-hop networks is much more complicated as seen in [44]. In [44], a multi-hop spatial reuse Aloha scheme that achieves the Gupta and Kumar bounds is proposed. We show similar results for slotted Aloha and Aloha using modified simulation models discussed in Sections 3.2.1 and 3.2.2, respectively.

3.2.1 System Model (Slotted Aloha)

In order to simulate slotted Aloha we keep the assumption that one slot is equal to one packet length. We do not model return ACKs from the receiver and assume it is sent on an orthogonal channel. As in the previous section, we assume that each node desires to access the channel with a common medium access probability $p_T \leq 1$ in each slot. Collisions are determined according to the protocol model for differentiated ranges. A transmission from node i to node j results in a collision if any of the following conditions are met:

1) The intended receiver j also tries to transmit in the same slot with probability p_T.
2) Two transmitters, i and k, pick the same intended receiver.
3) Another node k transmits in the same slot such that $d(j,k) \leq (1+\Delta)d(i,j)$.

A transmission resulting in a collision is not counted towards the final transport and throughput capacity. Retransmissions are not explicitly modeled; it is assumed that retransmissions are scheduled in such a way that the medium access probability is sufficient to model both new and repeated transmissions.

3.2.2 System Model (Aloha)

The Matlab simulation inherently utilizes discrete time slots. Therefore, we equate 1 packet with 10 slots, and allow nodes to transmit at any given slot to simulate an unsynchronized, un-slotted system. In order to maintain consistency in the traffic load between slotted and un-slotted Aloha, packets arrive at idle nodes according to a Poisson process with medium access probability $p'_T = \frac{p_T}{10}$ per slot. The total probability of an idle node transmitting a packet over the packet duration of 10 slots is simply $10 * \frac{p_T}{10}$, which is equivalent to the previous cases. A transmission results in a collision if any of following conditions are met:

1) The intended receiver j is busy receiving or transmitting a previous packet.
2) The intended receiver j also tries to transmit in the same slot with probability $p_T/10$.
3) Two transmitters, i and k, pick the same intended receiver.

4) Another node k transmits in any of the 10 slots such that $d(j,k) \leq (1+\Delta)d(i,j)$.

3.2.3 Simulation Results (Slotted Aloha and Aloha)

The un-slotted Aloha simulation results look similar to the slotted Aloha simulations, therefore only major results are presented while slotted Aloha is discussed in detail. The simulated transport capacity and end-to-end throughput for slotted Aloha as a function of p_T parameterized by n are shown in Figure 3.7. The simulation shows that the performance of Aloha is heavily dependent on the medium access probability and the network density, as it is in the single-hop network case. Initially, the network performance increases with the medium access probability until the interference in the network and the high levels of collisions dominate the behavior and the performance begins decreasing. Compared to slotted Aloha, the Aloha protocol shows a similar trend in performance, but peaks at a lower network load and at about half of the performance.

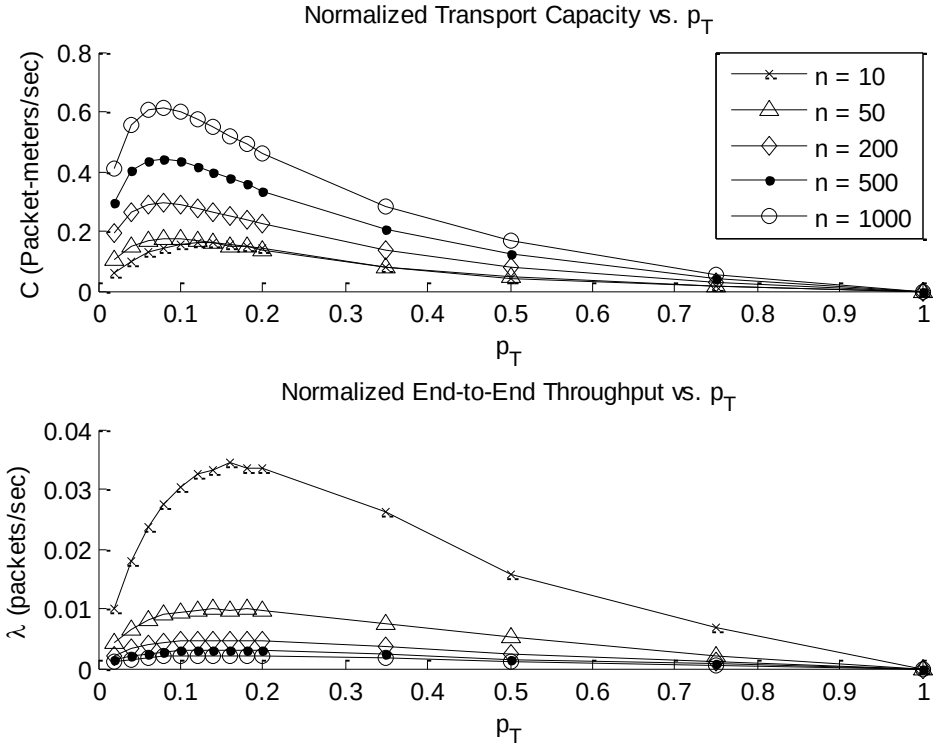

Figure 3.7 Transport capacity and end-to-end throughput vs. p_T, slotted Aloha.

Figure 3.8 shows the transport capacity and end-to-end throughput with respect to network density for select values of p_T. The top plot shows that the transport capacity increases with the network density under the slotted Aloha protocol, however there is a severe performance degradation compared to the ideal MAC case. The bottom plot shows that the end-to-end throughput still diminishes as the network density increases.

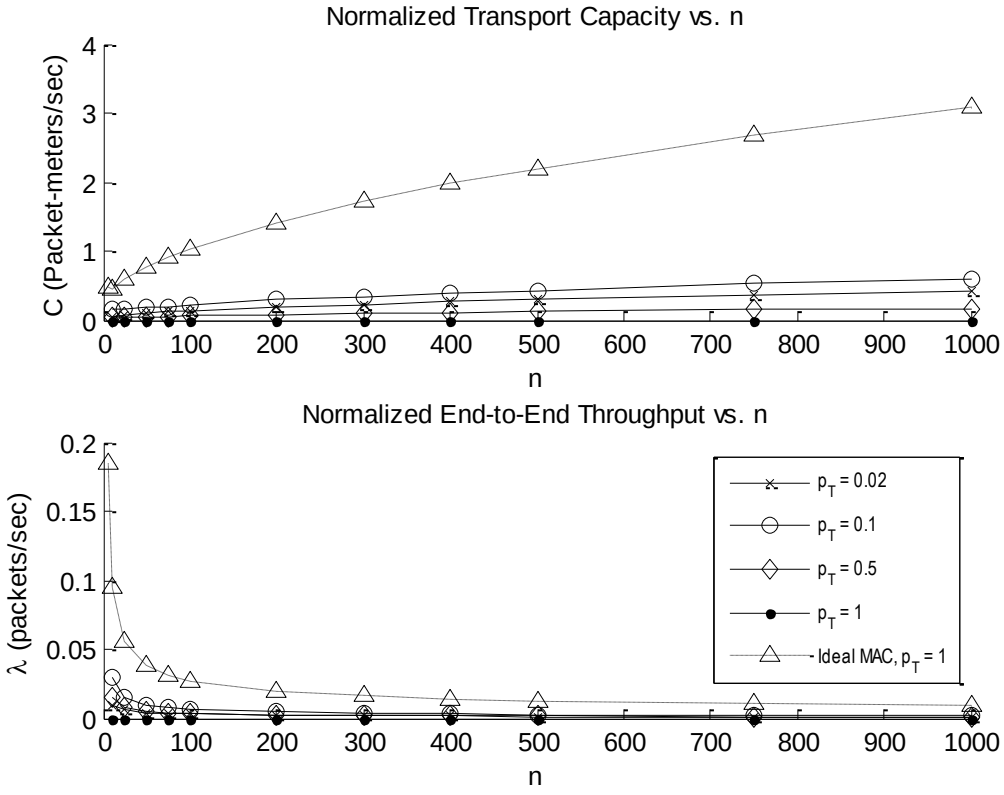

Figure 3.8 Transport capacity and end-to-end throughput versus *n*, slotted Aloha.

The transport capacity is dependent on the number of successful simultaneous transmissions and the average length of each transmission. Figure 3.9 presents the average number of attempted (top left), and successful (top right) simultaneous transmissions and the average attempted (bottom left), and successful (bottom right) transmission length with respect to the medium access probability, parameterized by n. The number of attempted transmissions increases linearly with the medium access probability, while the number of successful transmissions increases initially, and then falls as interference builds up in the network. The attempted hop distance stays constant with the medium access probability but the average successful transmission length decreases as the number of transmissions increases. The observation implies that shorter transmissions have a statistically better chance of success, and the collision probability associated with a transmission is dependent on the distance to the receiver.

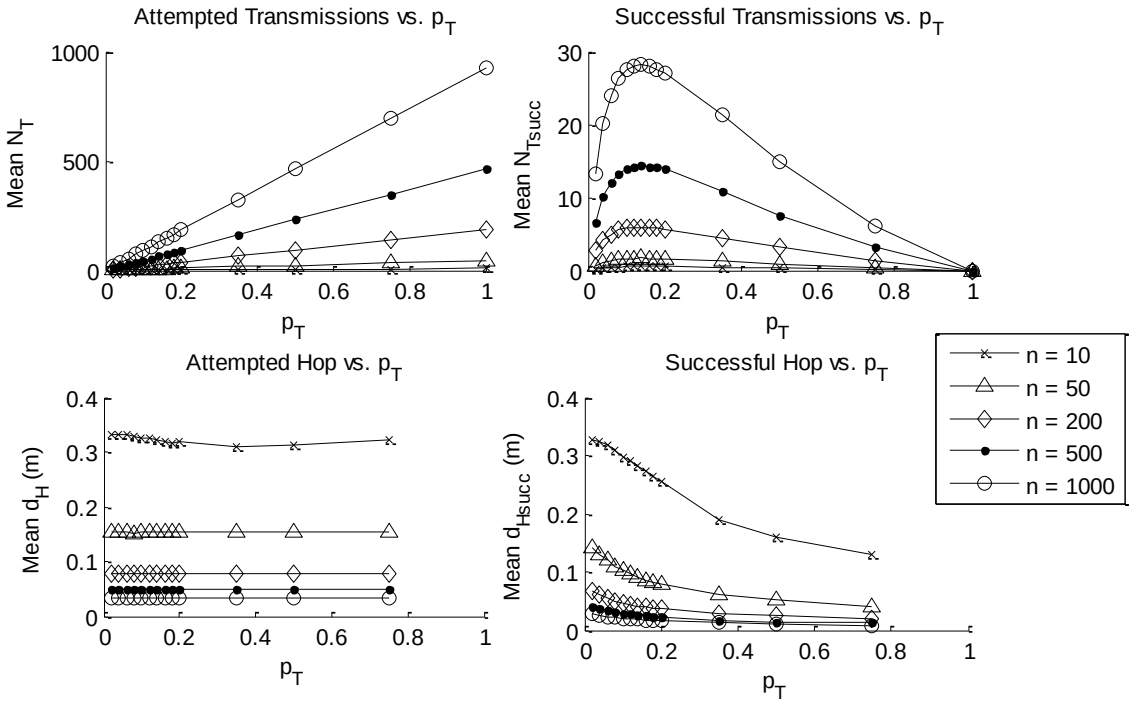

Figure 3.9 Simultaneous transmissions and transmission length vs. p_T, slotted Aloha.

Figure 3.10 plots the average number of successful transmissions (top) and the average successful transmission length (bottom) with respect to the network density for slotted Aloha. The values from the ideal MAC simulation are also plotted to illustrate the linear trend in the number of transmissions is maintained, along with the $\sqrt{logn/n}$ decreasing trend in the mean transmission length, as the network scales.

Clearly, Aloha performs poorly compared to the ideal case. To understand the performance loss, we combine the number of attempted and successful transmissions into a single performance metric, *probability of success.* We define the probability of success as

$$P_{Succ} = \frac{N_{Tsucc}}{N_T}. \tag{3.3}$$

Since the ideal MAC scheme avoids collisions, it always has unit probability of success. As shown in Figure 3.11 for slotted Aloha, the probability of success diminishes as the medium access probability increases, illustrating that uncoordinated transmissions lead to congestion and inefficiency (collisions) in the network. For a given p_T, the probability of success remains steady as the number of nodes increases for $n > 25$. From this observation, we can conclude that the local interference seen by a node remains steady since the average transmission range scales down as the network density increases.

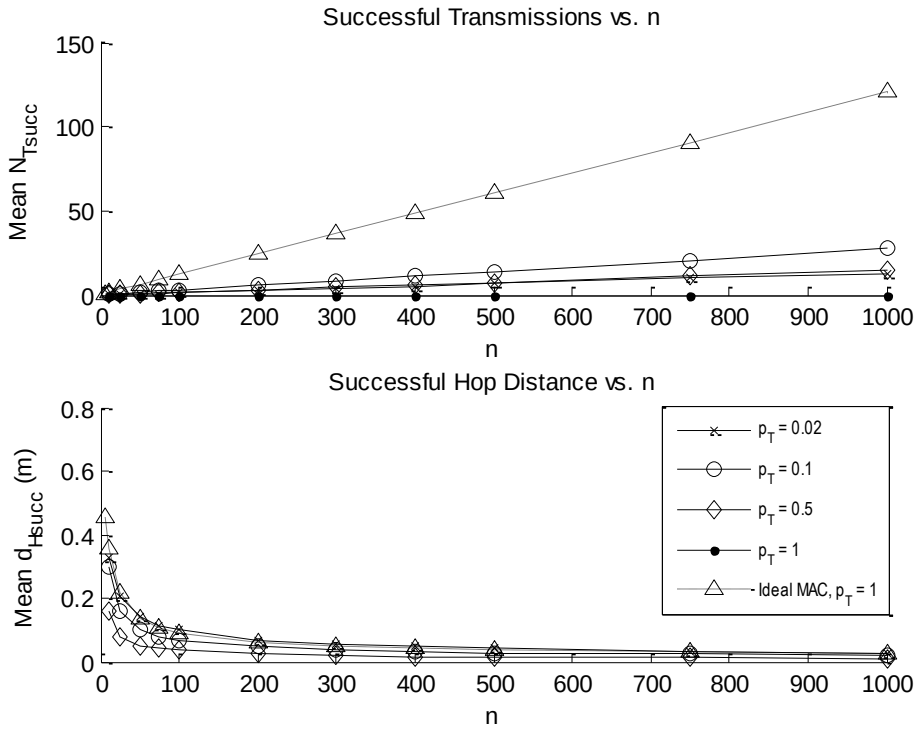

Figure 3.10 Simultaneous transmissions and transmission length vs. *n*, slotted Aloha.

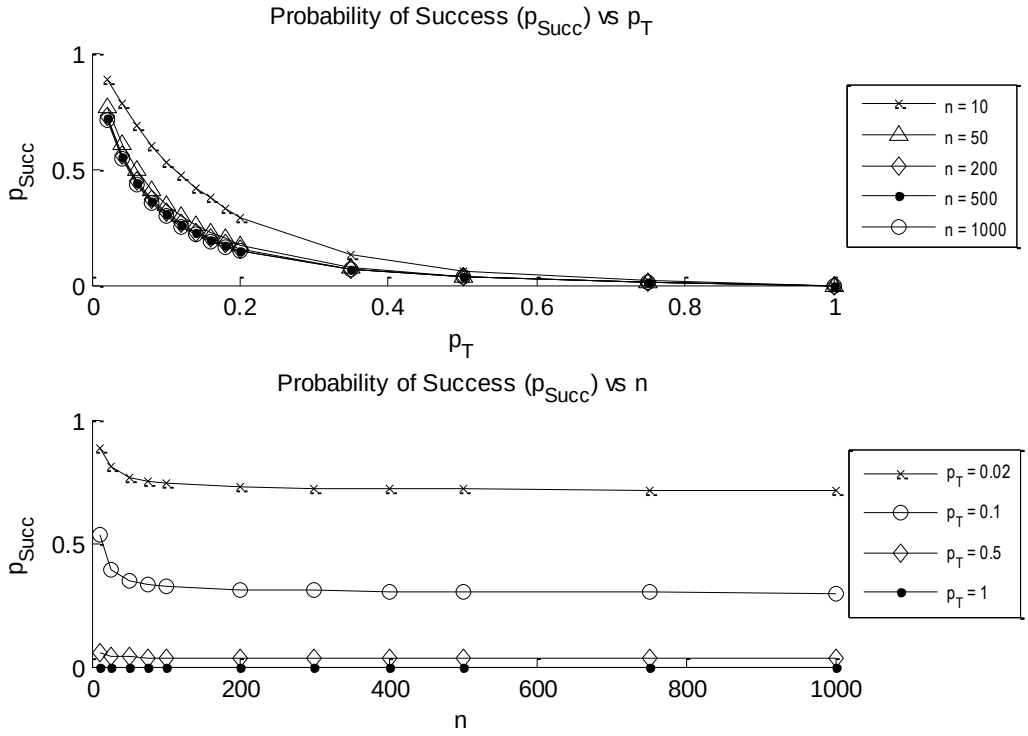

Figure 3.11 Probability of success, slotted Aloha.

All of the results show that the performance of the Aloha and slotted Aloha protocols is highly dependent on the chosen medium access probability, and to a lesser extent the network density. A single-hop network employing the Aloha protocol must adjust the medium access probability per node in accordance with the network density to maintain peak performance. In a multi-hop network with dynamic power control and communication limited between nearest neighbors, the number of interfering neighbors stays constant, as shown in Figure 3.11 for $n > 25$. Therefore, the optimum medium access probability may be determined apriori.

Figure 3.12 plots the medium access probability that optimizes the number of successful transmissions, the end-to-end throughput, and the transport capacity for slotted Aloha (left) and Aloha (right). Interestingly, in both cases, the probability which optimizes the throughput (0.14 and 0.08) is slightly higher than the probability which optimizes the transport capacity (0.08 and 0.04). The difference between the two may be explained by the earlier observation that the average successful transmission distance decreases with p_T, which negatively impacts transport capacity which is a measure of rate and distance. This is confirmed by the observation that the optimum value for the number of successful simultaneous transmissions also correlates closely with that for the end-to-end throughput, as both measures do not consider the transmission length. Comparing the two plots we can see that the un-slotted protocol gives peak performance at approximately half p_T. This observation is consistent with the behavior of the Aloha protocol in a single hop network, since the lack of synchronism between nodes doubles the collision probability [11].

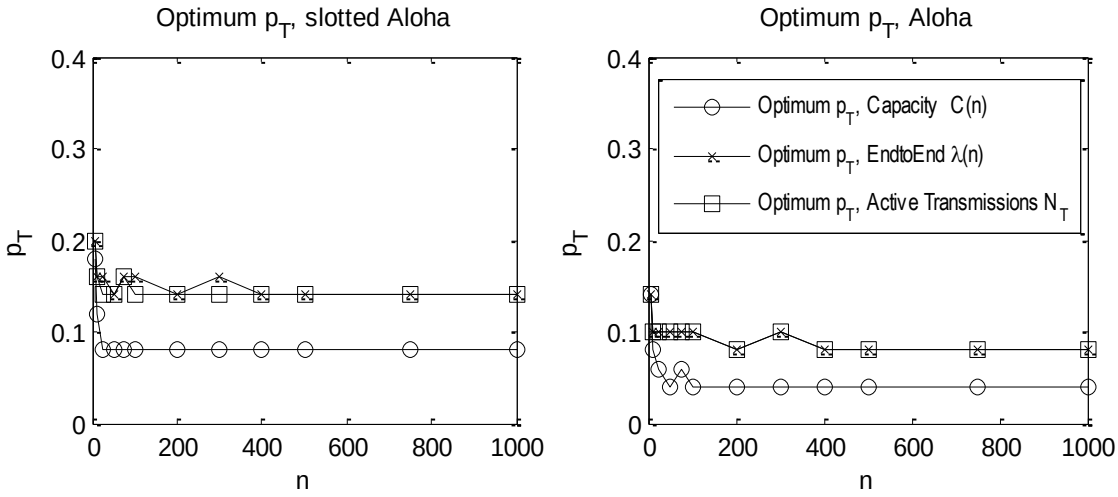

Figure 3.12 Choosing optimum p_T for slotted Aloha (left) and Aloha (right).

Figure 3.13 directly compares the peak performance of un-slotted Aloha and slotted Aloha. Slotted Aloha outperforms un-slotted Aloha for all three performance metrics of transport capacity, end-to-end throughput, and number of simultaneous transmissions. We observe that the peak transport capacity shows a 74-80% gain with slotted Aloha than with un-slotted Aloha, while peak end-to-end throughput shows a 70-76% gain. The result is similar to the behavior in a single-hop network, although slotted Aloha does not quite provide a 100% gain in performance. The observation suggests that there are other factors in play in a multi-hop network. For example, in multi-hop networks, a collision can also occur if the intended receiver is busy with another transmission.

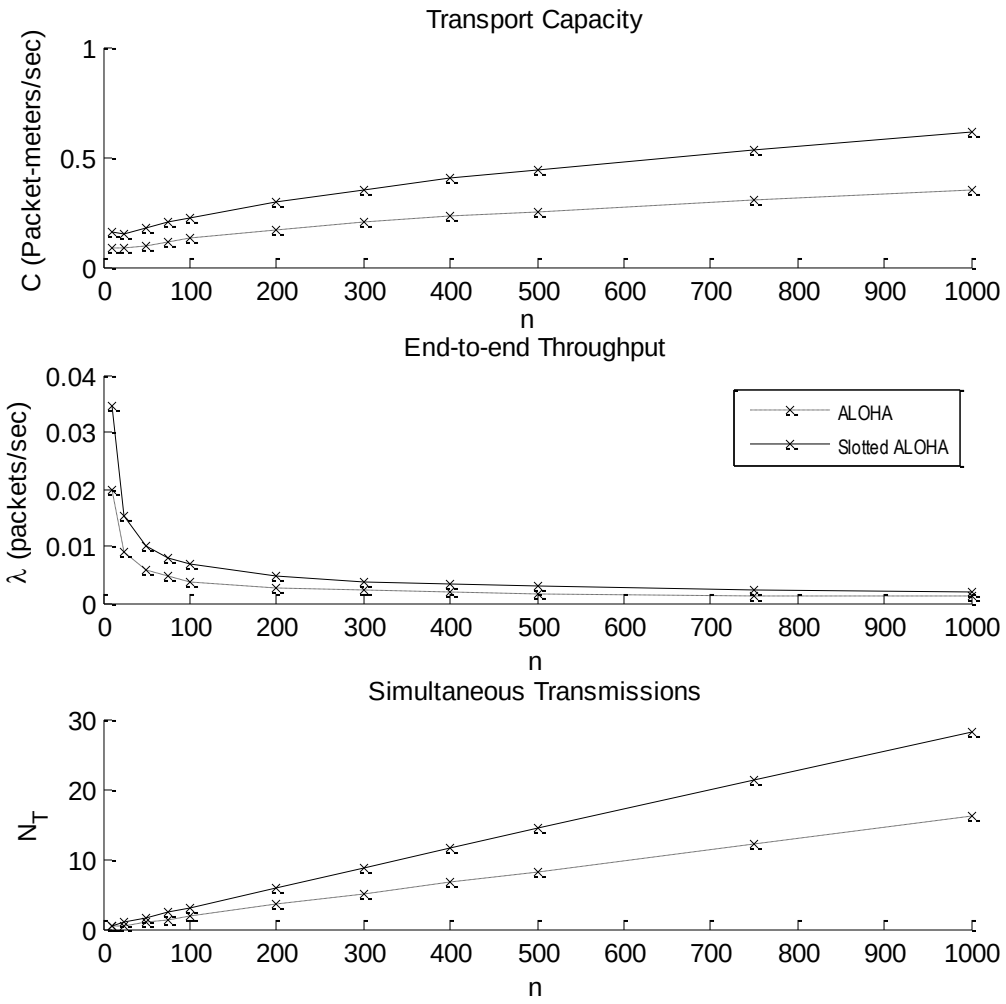

Figure 3.13 Slotted Aloha & Aloha.

Figure 3.14 compares the maximum performance of an ad hoc network with slotted Aloha and an optimally chosen medium access probability with the ideal MAC case discussed in Section 3.1. The transport capacity achieved by slotted Aloha in the optimum case pales in comparison to the maximum case for ideal MAC, achieving approximately 22% for large values of n. In a direct comparison with a smaller medium access probability of 0.08—which gives peak capacity for slotted Aloha—slotted Aloha achieves nearly 50% of the transport capacity of ideal MAC. A similar loss in end-to-end throughput can also be seen in the lower plot at $p_T = 0.14$ (which optimizes slotted Aloha end-to-end throughput). The throughput and capacity degradation seen illustrates the consequences of uncoordinated access to the channel and reinforces the importance of a good MAC scheme. For DSA, *the results show the impact on performance that can be expected without adequate spectrum sensing.*

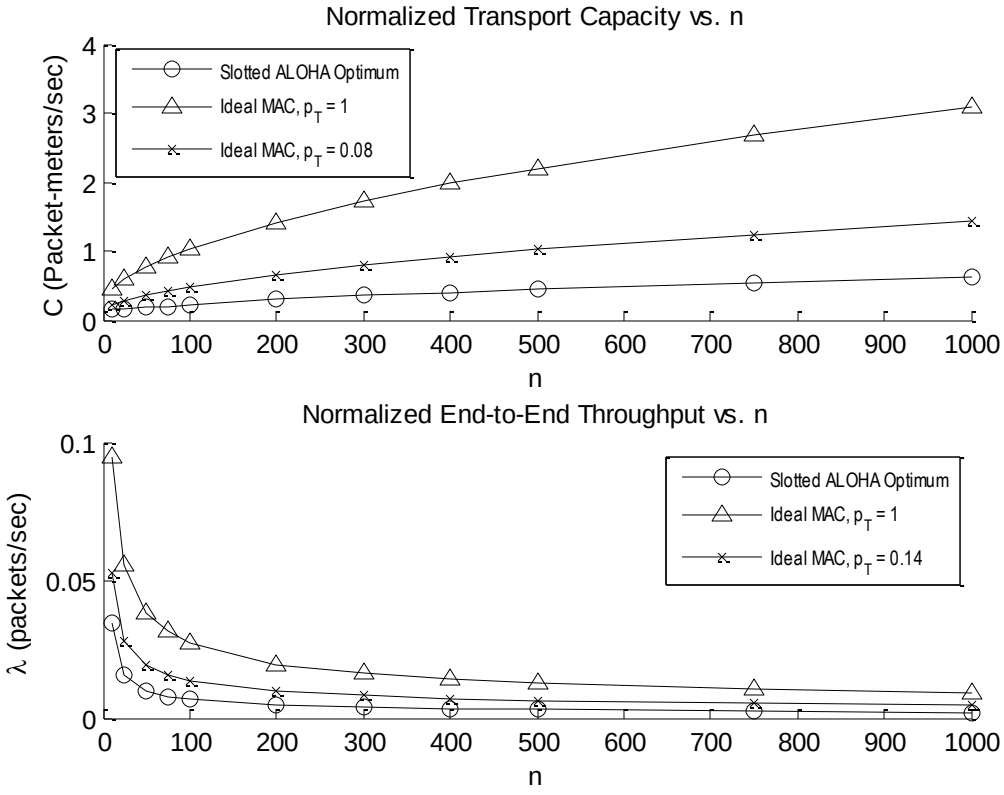

Figure 3.14 Slotted Aloha & Ideal MAC.

3.2.4 Discussion

In this section, we presented simulation results comparing the performance of slotted Aloha, a basic random access MAC scheme, with an ideal MAC scheme for random networks. We observed a drastic reduction in performance with slotted Aloha, and the performance was highly dependent on the medium access probability. Curiously, the optimal medium access probability values for transport capacity and end-to-end throughput stabilized as the network density increased. It is well known that, in a single-hop network of N nodes optimal throughput is achieved when $p_T = \frac{1}{N}$ [11]. The optimum p_T value remains constant in our simulations, which implies that the local interference seen at a node remains constant if nodes scale down transmission power according to the critical radius. In other words, as observed by Gupta and Kumar in [9], the average number of interfering neighbors of a node in a multi-hop network remains steady as the network grows.

We also observed that for low p_T slotted Aloha achieves nearly 50% of the performance of the ideal MAC scheme with the same p_T. Therefore slotted Aloha may be a viable design choice in certain ad hoc network applications, such as sensor networks, with slow and intermittent traffic. Especially since slotted Aloha requires very little overhead, other than synchronization, and there can be optimizations made to the basic form of slotted Aloha. For example, Yao and Li in [46] show that allowing for differentiated medium access probabilities, based on the relaying load of

a node, increases the performance of slotted Aloha in multi-hop networks. Lee and Kim in [47] show that network coding can be used at a congested node to increase the performance of Aloha in multi-hop networks. Xie and Haenggi [48] have studied the delay performance associated with slotted Aloha in multi-hop networks and compared it with the performance of a TDMA MAC scheme.

3.3 Carrier Sense Multiple Access (CSMA)

Carrier Sense Multiple Access (CSMA), first introduced in [13], allows for opportunistic medium access in single channel networks. CSMA allows nodes to opportunistically access the channel by estimating the interference at the receiver through channel sensing. *P-persistent* CSMA protocols continuously sense the channel and attempt to transmit with probability *p* once the channel is clear and *non-persistent* CSMA protocols periodically sense the channel after a random delay [11]. We simulate a non-persistent un-slotted CSMA scheme to evaluate transport capacity and end-to-end throughput. In single-hop networks, CSMA solves the congestion problem seen in Aloha networks and provides a significant gain in performance [11]. In multi-hop networks, carrier sensing may not detect all interferers at the intended receiver—an issue widely known as *the hidden node problem* [14].

Figure 3.15 illustrates a transmission scenario between a transmitter i and receiver j at a distance d_H where the sensing radius R_{CS} is assumed to be equivalent to the interference radius as defined by the relaxed protocol model. As seen in Figure 3.15, a *hidden node* is inside the interference region of the receiver but outside of the sensing region of the transmitter. Likewise, an *exposed node* is inside the sensing region of the transmitter but does not cause any interference to the receiver.

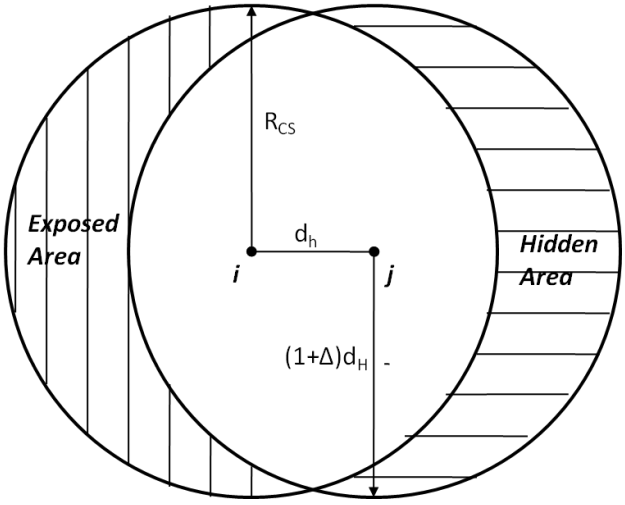

Figure 3.15 Hidden and exposed nodes.

Pre-existing hidden nodes occur in simulation when a transmitter senses the channel as clear and begins to transmit, but the transmitter fails to see another ongoing transmission which is inside the interference region of the receiver. This hidden transmitter causes interference at the intended receiver which leads to a collision. Hidden nodes may also turn on and cause a

collision during the transmission if the transmitter is out of the sensing range of the hidden node. Exposed nodes lower performance by unnecessarily preventing transmissions when the channel is busy at the transmitter but free at the receiver.

3.3.1 System Model

The network model from previous simulations is kept intact but the access scheme is modified to model a non-persistent CSMA network. As discussed in previous sections, outgoing traffic at each node is modeled according to a common medium access probability p_T. We assume Poisson arrival of packets, of 10 slots in length, at idle nodes at a common rate of $p_T/10$. Once a node has traffic it senses the channel and begins transmission if the channel is sensed idle. We assume proximity based channel sensing similar to the protocol model used for interference. If the channel is sensed busy, it waits and senses the channel again after a random delay uniformly distributed between 1 and 20 slots.

Under the protocol model, a successful transmission between nodes i and j occurs if $d(k,j) \geq (1+\Delta)d(i,j)$ for all other transmitters k in the entire duration of the transmission. Since we assume differentiated ranges for transmission, we also assume the sensing radius R_{CS} varies for each transmission depending on the proximity of the intended receiver. The sensing radius for node i intending to transmit to node j is

$$R_{CS} = (1+\Delta)d(i,j)^{1}. \tag{3.4}$$

If $d(k,i) \geq R_{CS}$, for all existing transmitters k, then node i assumes the channel is clear and begins transmission. Otherwise, the node declares the channel to be busy and queues the transmission until a random delay counter expires. At the end of the counter, the node senses the channel again and transmits or begins another counter if the channel is sensed busy. In order to characterize the performance limits of CSMA compared to an ideal scenario, it is assumed that sensing occurs in zero time and the propagation delay is zero. Additionally, we assume perfect sensing inside radius R_{CS}, so no collisions occur due to sensing errors.

If the protocol model for interference is violated for any duration of a transmission, a collision is declared and the packet is not counted towards the transport capacity or the end-to-end throughput. Neither returning ACKs from the receiver nor retransmissions due to collisions are modeled. A transmission from node i to node j results in a collision if any of the following conditions are met:

1) The intended receiver j also tries to transmit in the same slot as i.
2) Two transmitters, i and k, pick the same intended receiver and transmit at the same slot.
3) Another node k inside the *interference region* of j such that $d(j,k) \leq (1+\Delta)d(i,j)$ transmits in the same slot as i.
4) The beginning of a transmission of a *hidden node* during any of the 10 slots that node j is receiving.

---

[1] The value of $R_{CS} = (1+\Delta)d(i,j)$, where $\Delta = 1$ was found in simulations to come close to optimal.

3.3.2 Simulation Results

The transport capacity and end-to-end throughput seen in simulations are shown in Figure 3.16 with respect p_T, parameterized by n. Observe that the performance increases rapidly with $p_T < 0.1$ and tapers off as $p_T \to 1$. In contrast with the performance of slotted Aloha seen in Figure 3.7, the performance of CSMA is shown to be stable in response to an increasing network load.

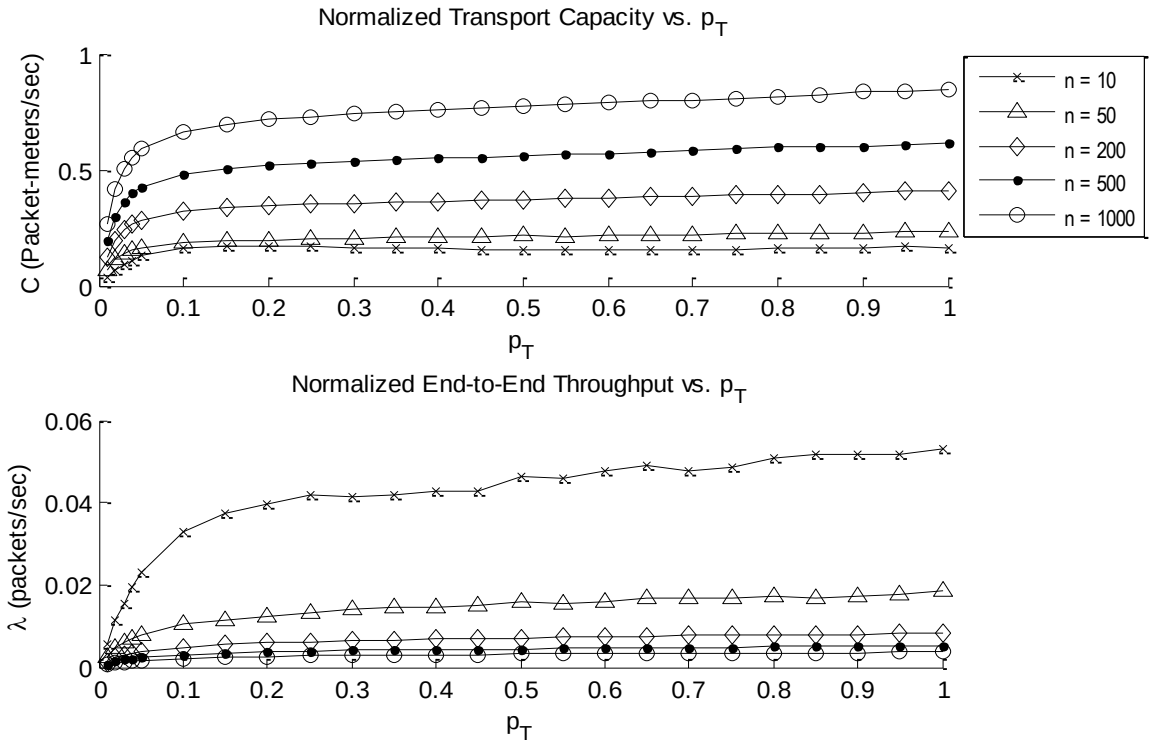

Figure 3.16 Transport capacity and end-to-end throughput vs. p_T, CSMA.

Figure 3.17 presents the performance for select values of p_T as the number of users in the network scales. Observe that the transport capacity of CSMA is much lower than the ideal MAC, yet it still grows with the number of nodes. The diminishing trend in end-to-end throughput is also witnessed as the network scales. As far as scalability is concerned, the simulations illustrate that CSMA reduces network performance, but may preserve the scaling relationships for end-to-end throughput and transport capacity. As discussed earlier, DSA and CSMA operate on a similar principle of sensing and opportunistically transmitting on a free channel. The results indicate that even with carrier sensing, collisions still occur in ad hoc networks because of interference due to hidden nodes and other reasons discussed previously.

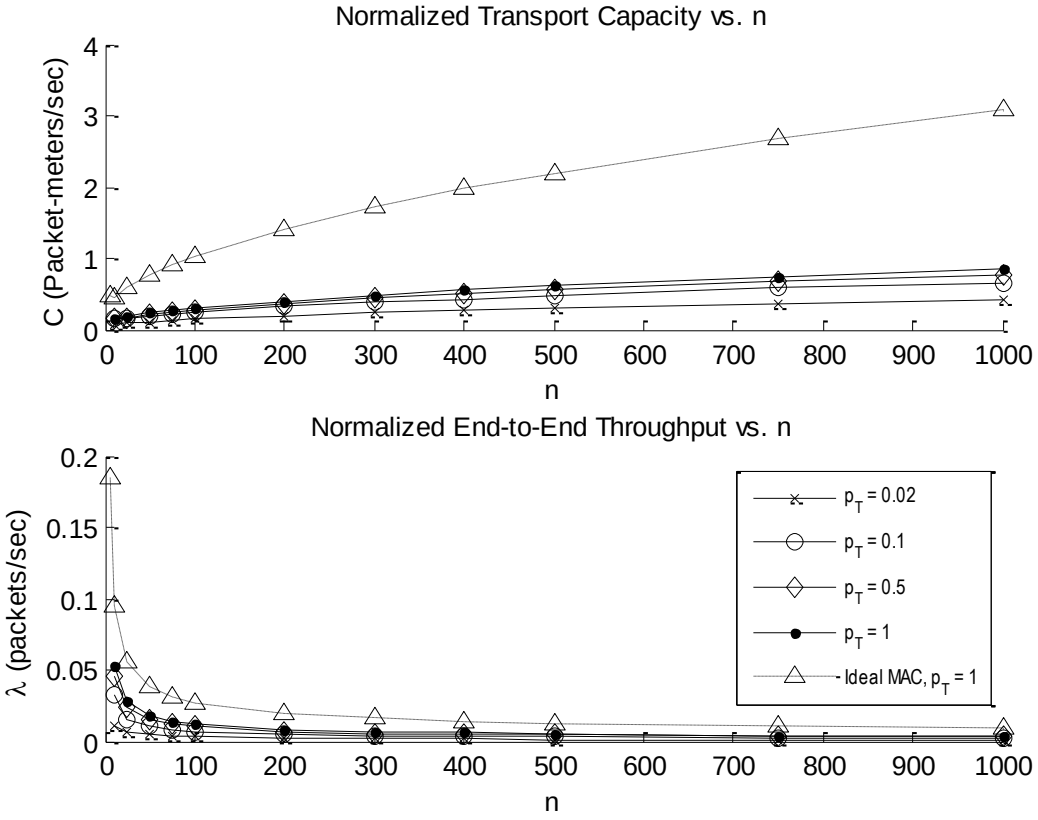

Figure 3.17 Network performance vs. *n*, CSMA.

Figure 3.18 plots the average number of simultaneous successful (top right) and attempted transmissions (top left), along with the average length of each attempted (bottom left) and successful hop (bottom right). The number of simultaneous transmissions increases rapidly with the network load and begins to stabilize for $p_T > 0.2$. The two top plots show that approximately $1/3$ of the transmissions are lost due to collisions. The two lower plots show that both d_H and d_{Hsucc} decrease as p_T increases. While both d_H and d_{Hsucc} remain constant in the ideal MAC simulations, only d_{Hsucc} decreases in the Aloha simulations as seen in Figure 3.3 and Figure 3.9, respectively. The average attempted hop distance may decrease with the network load because the sensing radius is proportional to the transmission distance, and longer transmissions are more likely to be delayed due to a larger sensing region.

Figure 3.19 plots the average N_{Tsucc} (top) and d_{Hsucc} (bottom), for select values of p_T, with respect to the network density. Although the average number is lower, the linear growth in N_{Tsucc} is consistent with the ideal MAC. The lower plot shows that the $\sqrt{logn/n}$ scaling behavior in successful hop distance is preserved with CSMA although the average length decreases as the network load increases for reasons discussed previously.

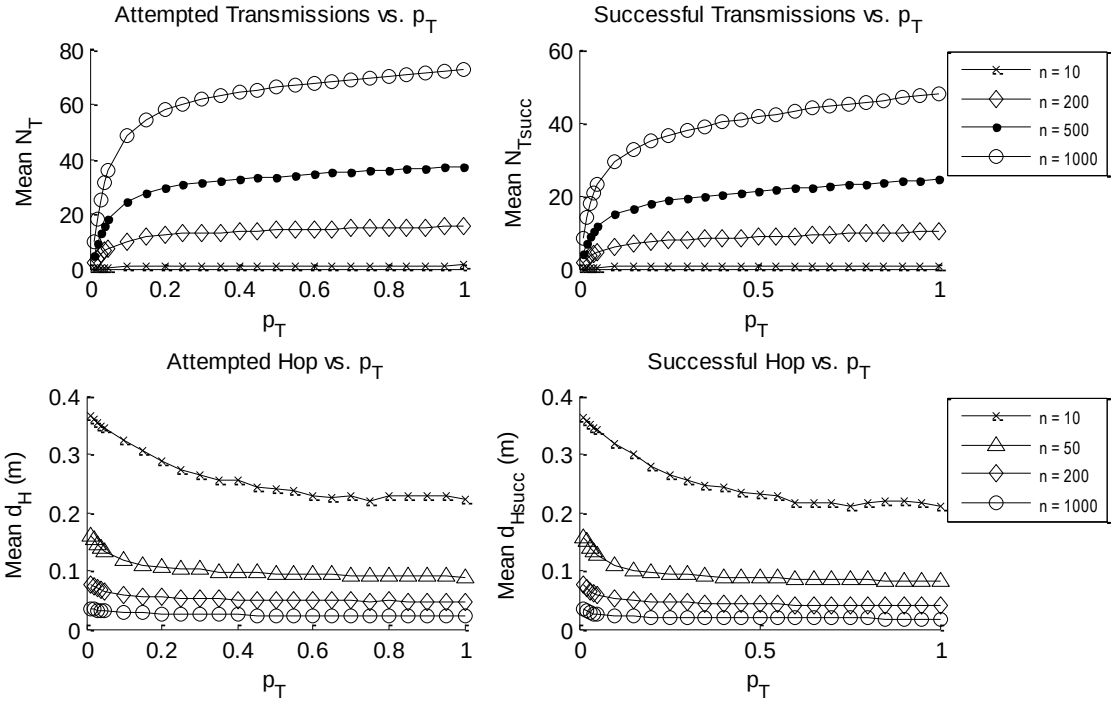

Figure 3.18 Simultaneous transmissions and transmission length vs. p_T, CSMA.

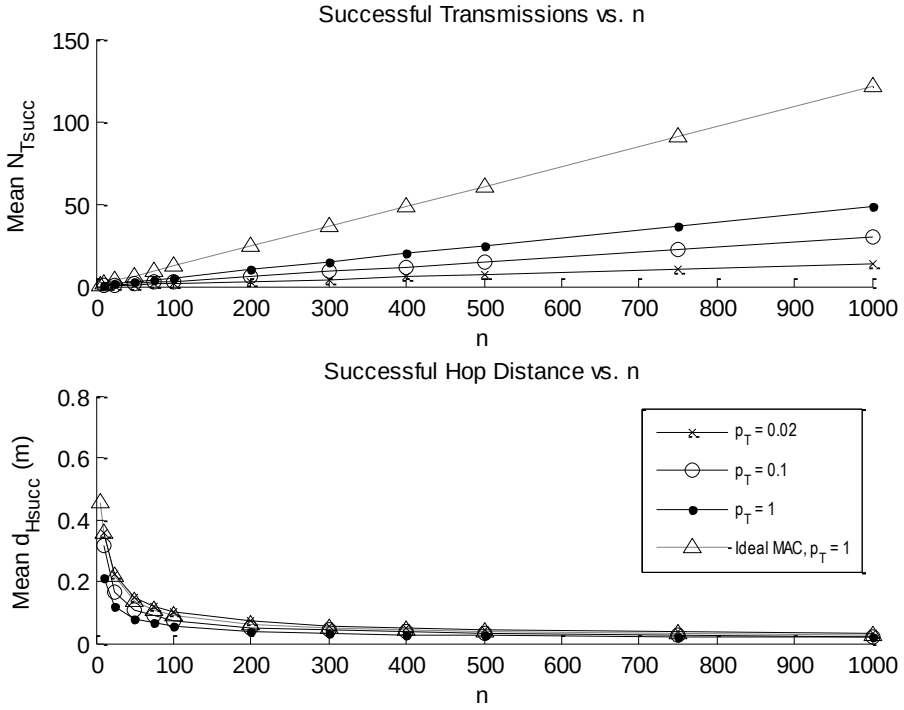

Figure 3.19 Simultaneous transmissions and transmission length vs. *n*, CSMA.

Figure 3.20 illustrates the probability of success of a packet with respect to p_T (top) and n (bottom). At low network loads ($p_T < 0.1$) the probability of success is very high and falls as the network load increases. As the load builds up ($p_T > 0.2$) the probability of success stabilizes and slightly increases as more transmissions are queued by the sensing mechanism. This increase in performance is likely because the average transmission range decreases as the load increases, which also increases the transport capacity as discussed in Chapter 2. The behavior of P_{SUCC} as the network load increases for CSMA contrasts sharply with slotted Aloha because, while imperfect, carrier sensing limits the transmission of packets in a congested network. The lower plot illustrates that the probability of success is largely independent of the network density for $n > 25$, while we saw previously in Figure 3.11 that the probability of success for slotted Aloha diminishes as more transmissions are scheduled.

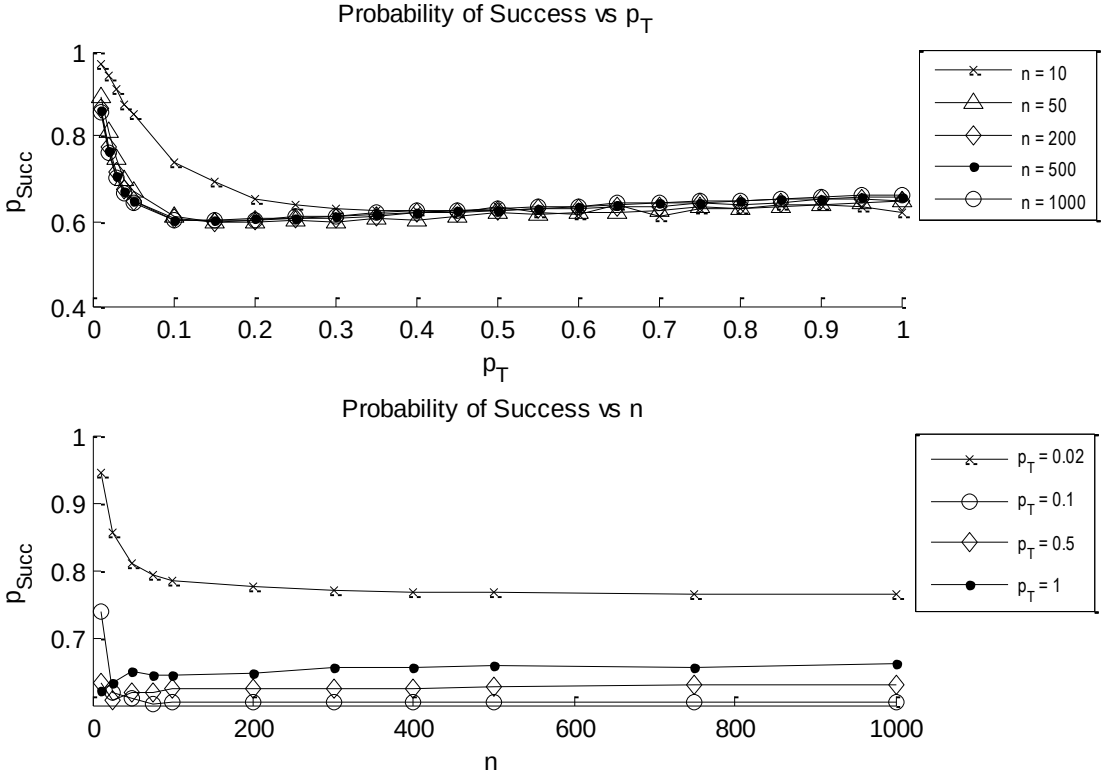

Figure 3.20 Probability of success, CSMA.

3.3.3 CSMA Discussion

CSMA is shown to be stable and mitigate the congestion problem seen in Aloha—the throughput and capacity do not decrease with higher traffic but increase slightly as medium access probability increases. Figure 3.21 plots the transport capacity and end-to-end throughput seen in simulation for $n = 100$. The performance of CSMA and slotted Aloha is essentially identical for $p_T \leq 0.08$, after which the performance of slotted Aloha begins to decrease while the performance of CSMA continues to increase at a small rate.

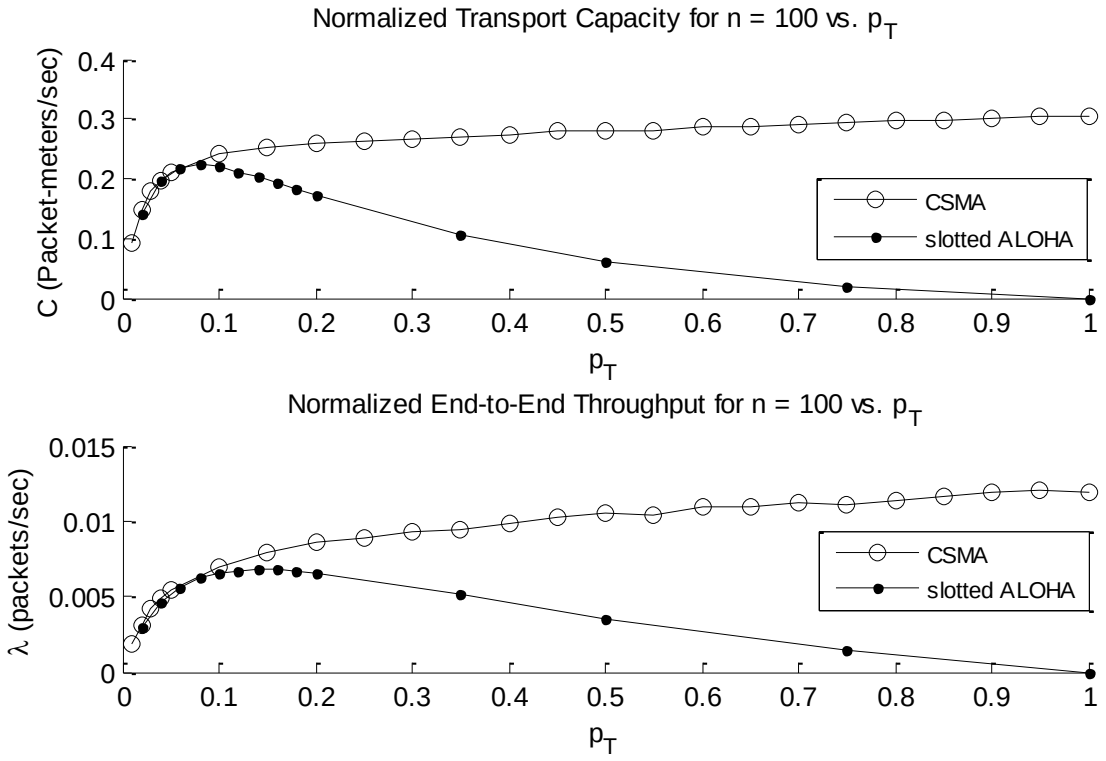

Figure 3.21 Direct comparison of CSMA and slotted Aloha for varying network load.

Although CSMA is less sensitive to changes in the medium access probability, CSMA fails to reach even half of the performance of an ideal MAC as seen in Figure 3.17, even under ideal assumptions of zero sensing delay and zero propagation time. The poor performance of CSMA in ad hoc networks is mainly due to the presence of hidden and exposed nodes, as discussed in Section 3.3.1.

Performance suffers when a transmitter fails to accurately sense the level of interference at the receiving node, which happens quite often since nodes can only estimate the interference level at another node by extrapolating the interference measured at their own location. Figure 3.22 illustrates the accuracy of the sensing mechanism determined through simulation, averaged for networks with $n \geq 50$. The percent of sensing attempts which lead to a failed transmission due to a pre-existing hidden node, and the percent of sensing attempts which were unnecessarily blocked due to an exposed node, are shown against p_T. For $p_T < 0.2$, the relative impact of pre-existing and exposed nodes is greater because the network is operating under capacity. As p_T increases, other sources exert a greater influence in causing failed and blocked transmissions. Both statistics stabilize for $p_T > 0.2$, the percentage of attempts blocked by exposed nodes (only) is approximately 7%, while the percentage of attempts resulting in a collision due to hidden nodes (only) is approximately 6.5%.

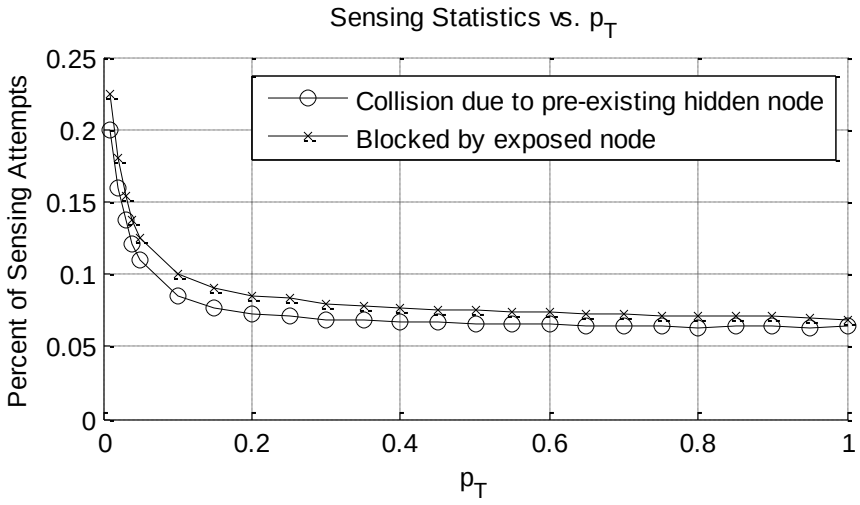

Figure 3.22 CSMA sensing performance vs. p_T.

3.4 Conclusions

Chapter 2 investigated the theoretical scaling behavior of an ad hoc network, which deals with the unavoidable limits to scalability due to physics and the nature of interference. We also simulated ad hoc network performance in the presence of an ideal MAC scheme. We have extended our simulations another level in this chapter to analyze the performance impact of random access MAC schemes for ad hoc networks. We did not consider overhead associated with the MAC layer (such as ACK packets) and assumed zero propagation and sensing delay, yet we saw severe degradation in capacity and throughput. The simulations clearly show that the network performance degrades heavily when all of the nodes must share a single channel without a centralized controller.

We observed that Aloha and slotted Aloha protocols' performance was highly dependent on the medium access probability. While CSMA mitigates the congestion problem, it does not approach the maximum end-to-end throughput and transport capacity of an ideal MAC scheme. Figure 3.23 illustrates the peak performance of CSMA, slotted Aloha, Aloha and an ideal MAC. Slotted Aloha shows a 75% gain in capacity over the un-slotted form, while CSMA is shown to achieve about 50% more transport capacity than slotted Aloha.

In order to compare the scaling behavior of transport capacity with random access MAC protocols to the known $C(n) = \Theta\left(\sqrt{n/logn}\right)$ behavior of the ideal case, the peak performance of each is plotted in Figure 3.24 as a percentage of the maximum achieved under ideal case. For low density networks ($n < 50$) the relative performance of each protocol is higher but decreases greatly as the density increases due to collisions and inefficient operation. For high density networks, the relative performance stabilizes at approximately 27% for CSMA, 20% for slotted Aloha and 11.5% for Aloha.

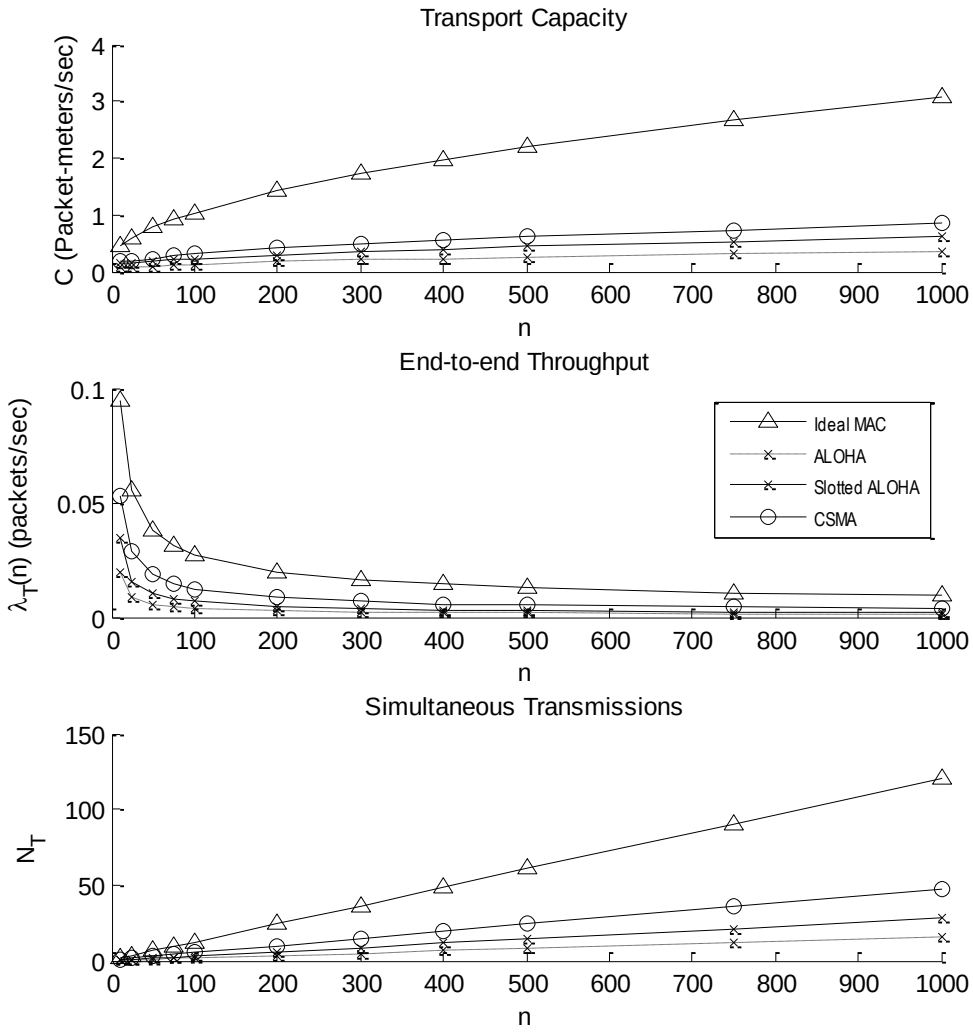

Figure 3.23 CSMA, slotted Aloha, Aloha and Ideal MAC simulations results.

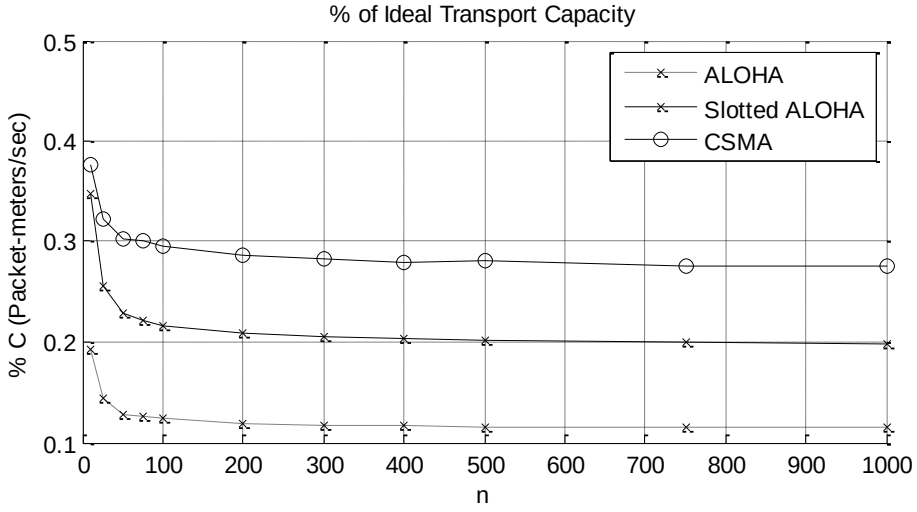

Figure 3.24 Random access MAC performance compared to Ideal MAC.

In Chapter 2 we saw that the scalability of ad hoc networks is limited by the interference model, and as noted by Gupta and Kumar, "the additional burden in coordinating access to the wireless channel … can only further throttle capacity" [9]. The performance Aloha and CSMA protocols discussed in this chapter can be improved with techniques such as optimizing the sensing radius of CSMA; however the improvement is independent of the user density n and not expected to change the scaling behavior. Figure 3.24 shows that the performance, as a percentage of the ideal case, is nearly constant as the number of nodes increases for $n \geq 50$. This observation implies that random and opportunistic MAC schemes have a linear degradation on performance with respect to the number of nodes. In other words, the random MAC schemes investigated in this chapter achieve the $C(n) = \Theta\left(\sqrt{n/logn}\right)$ scaling law proposed by Gupta and Kumar in [9].

Other forms of medium access such as TDMA or FDMA may produce better performance at the expense of increased overhead and network planning. However, our results show that the scalability bounds for ad hoc networks can be met with simple random access techniques. For DSA, the results in this chapter show that hidden node problem in sensing has a big impact on performance. DSA is expected to run into similar problems when sensing for free channels, especially when considering errors in the sensing process such as missed detections and false alarms. Additionally, so far we have concentrated on single channel networks but DSA must coordinate access to multiple channels. DSA must provide a solution for the *multi-channel rendezvous* problem, where one channel may be free at one node but may not be free at another node. This problem may be aided by *distributed sensing* [22] but that also introduces additional overhead in the network. The next chapter studies performance analysis of DSA networks.

We also examine an analytical method for performance evaluation of single channel CSMA ad hoc networks in Appendix B. The results demonstrate that it is possible to arrive at similar throughput and capacity results through analysis. Analytical models also allow for a higher degree of flexibility and provide a deeper insight into the effect of each variable. Therefore, we expand the analytical performance evaluation model to ad hoc DSA networks in Chapter 4.

4 SCALABILITY OF AD HOC DSA NETWORKS

This chapter focuses on the performance and scalability of ad hoc DSA networks. As discussed in Chapter 1, DSA MAC protocols—at a minimum—must support spectrum sensing, spectrum characterization, and channel selection in addition to traditional MAC layer operations such as collision avoidance and coordinating access to the channel. Therefore, the performance of a DSA network depends greatly on the sensing mechanism and accuracy, spectrum characteristics such as interference level and path loss, and primary user activity. In this chapter, we create a flexible analytical framework for performance analysis of ad hoc DSA networks. We also create a DSA access scheme that is well-suited for ad hoc networks and performs basic DSA functions such as spectrum sensing, spectrum sharing, and spectrum access decision. We use a generalized version of the random access and hybrid DSA MAC protocols surveyed in Section 1.1.7. Our scheme relies on a dedicated control channel for rendezvous, as our survey showed that most current DSA designs rely on a dedicated control channel. We use our model to evaluate the average throughput achieved per node under varying spectrum availability, secondary user density, and level of primary user activity. We then mesh the analysis with the Gupta and Kumar definition of network capacity [9] and derive scalability trends in end-to-end throughput per node and network transport capacity.

In Section 4.1, we review previous work in DSA and ad hoc network performance analysis and delineate our contributions. In Section 4.2, we develop a network model for a heterogeneous network of SUs and PUs. In Section 4.3, we derive a sensing model using energy detection in Poisson distributed networks. We also derive detection probabilities at DSA nodes neighboring an active transmission. In Section 4.4, we describe our DSA access scheme and a semi-Markov analysis procedure. We also derive transition probabilities based on the interference model, and solve for the steady state equations. In Section 4.5 we show performance results using an iterative root-finding algorithm and analyze scalability trends. And finally, in Section 4.6, we discuss important conclusions.

4.1 Related Work

Mo et al. developed an analytical framework for performance evaluation of various general multi-channel MAC protocols over a single collision domain [49]. Single collision domain analysis simplifies interference modeling by assuming the transmission range of each radio exceeds the total network area so all nodes can receive every other nodes' transmission. Single collision domain analysis removes the hidden node problem and does not model spatial reuse expected in ad hoc networks. Traditional multi-channel MAC protocols address the channel selection and rendezvous problems but do not take spectrum sensing and PU detection into consideration. Pawelczak et al. [50] expand the analysis in [49] for various classes of DSA MAC networks and evaluate the performance in the presence of primary users. Their analysis gives insight into the behavior of DSA as a function of spectrum sensing performance and primary user activity; nevertheless, it considers the performance in a single collision domain only, thus the analysis does not immediately transfer to ad hoc DSA networks.

Recently, Al-Mahdi et. al. [51] investigated an analytical model using a three-dimensional Markov chain to evaluate the performance of SUs operating in licensed and unlicensed channels. Their study focuses on the impact of PU traffic intensity on SU blocking probability, dropping probability and throughput. As expected, results are presented which show an improvement in SU throughput and spectrum usage as SUs are able to access licensed channels. The analysis presented in [51] does not consider rendezvous or spectrum sensing issues, such as the hidden node problem.

The performance of CSMA in the presence of hidden nodes is considered in [52-55]. The analysis estimates the long run average throughput at a single node by modeling the state of the node over time as a semi-Markov process. Wu and Vershaney [52], [53] additionally, consider busy tone channels and multiple channels as ways of mitigating the hidden node problem. Hoang and Iltis build on [52], [53] and propose a generalized analytical framework for single channel CSMA networks which can also handle fading [54]. They also develop analytical models for performance evaluation with the four-way RTS/CTS handshake and beamforming in [55]. We use the analysis and ideas presented in [54], [55] as a starting point to evaluate the performance of ad hoc DSA networks. We use a semi-Markov process to model the state of a node, and also consider spectrum sensing, primary user activity, and availability of multiple data channels. Additionally, we integrate a realistic sensing model based on energy detection and measure its impact on performance. Energy detection is cost and computationally efficient, and it is the most popular method for sensing in DSA, as discussed in Chapter 1.

4.2 Network Model

4.2.1 SU Network Model

We assume an infinite planar network with nodes distributed according to a 2-dimensional Poisson point process with intensity n. The probability of finding k nodes in an area S is [45]

$$P(k) = e^{-nS}\frac{(nS)^k}{k!}. \tag{4.1}$$

The distance and angle from the origin to a node located inside a circle of radius R is given by the joint probability density function

$$f(r, \emptyset) = \frac{r}{\pi R^2}, \tag{4.2}$$

where $0 < r < R$ and $0 < \emptyset < 2\pi$. The density function of the distance from the origin of a point uniformly distributed in a circle of radius R is

$$f(r) = \frac{2r}{R^2}, r < R. \tag{4.3}$$

Each node randomly selects a receiver among its nearest neighbors within a radius r_s. The PDF of the distance r_0 between a node and its chosen destination is

$$f(r_0) = \frac{2r_0}{{r_s}^2}, r_0 < r_s. \tag{4.4}$$

We assume a Rayleigh fading channel with path loss and unitary transmit and receive antenna gains. The received SNR for a transmitter at distance r with transmit power P_T is given by

$$\gamma = \frac{AP_T\alpha^2}{N_o W r^\beta}, \tag{4.5}$$

where $\sqrt{A}$ is the channel gain, α^2 is an exponential random variable with unit mean, N_o is the noise variance, W is the channel bandwidth, and β is the path loss coefficient. Each SU transmits at a fixed power P_0, such that the received SNR is above a given threshold at the maximum receiver distance r_s. Each SU uses a random access DSA scheme, to be described in Section 4.4.1, for data transmission.

4.2.2 PU Network Model

The DSA network is presumed to be co-located with a general primary user network. We assume the primary users are randomly distributed in the same area as the secondary user network with density m (on each channel) and fixed transmit power P_1. Primary users do not perform sensing and may occupy any of the M data channels with probability q_p, which models the long-term probability of a PU being active. For inactive primary users, we model PU traffic, on each data channel, as Poisson with turn-on probability $q_p{}'$.

4.3 Sensing in Poisson Networks

As discussed in Chapter 1, when additional information about a signal is not known the optimal signal detector is an *energy detector* [56]. The probability of false alarm is fixed on detection parameters but the probability of detection and the probability of a missed detection depend on transmitter density, sensing radius, and transmission powers. In this section, we examine energy

detection in a general Poisson network. We derive the overall probability of detection $PD(\lambda, R, P_T)$ and probability of miss $PM(\lambda, R, P_T)$ for a class of nodes with transmitter density λ, sensing radius R, and transmission power P_T.

In a non-fading environment where the channel gain is deterministic, the detection performance for an isolated transmitter is given by the probability of detection P_d and the probability of false alarm P_f, given by

$$P_d = Q_{\lfloor sW \rfloor}\left(\sqrt{2\gamma}, \sqrt{\theta}\right), \text{ and} \tag{4.6}$$

$$P_f = \frac{\Gamma\left(\lfloor sW \rfloor, \frac{\theta}{2}\right)}{\Gamma(\lfloor sW \rfloor)}, \tag{4.7}$$

respectively, where $\Gamma(.)$ and $\Gamma(.,.)$ are complete and incomplete gamma functions and $Q_m(.,.)$ is the generalized Marcum Q-function [56]. P_f depends only on the time-bandwidth product $\lfloor sW \rfloor$ and the detection chosen threshold θ. P_d depends on $\lfloor sW \rfloor$, θ, and the instantaneous signal to noise ratio γ. The detection threshold is usually fixed to give a certain probability of false alarm, so for a given $\lfloor sW \rfloor$ and θ the probability of detection is simply a function of the received SNR.

When nodes are Poisson distributed on a plane with the sensing node at the center, the probability of detection is a function of the combined received SNR from all active transmitters. Since we are only interested in detecting transmissions from nearby nodes, we assume that the signal powers from transmitters inside a radius R sum while the total interference coming from outside of radius R is summed with the thermal noise in to the noise variance N_o.

The radius R is chosen as a function of the transmit power, such that the average received SNR is greater than γ' from a single transmitter located inside R. The sensing radii for secondary users R and primary users R_P are

$$R = \left(\frac{AP_0\alpha^2}{N_o W \gamma'}\right)^{\frac{1}{\beta}} \text{ and} \tag{4.8}$$

$$R_P = \left(\frac{AP_1\alpha^2}{N_o W \gamma'}\right)^{\frac{1}{\beta}}, \tag{4.9}$$

respectively. The transmission scenario is shown in Figure 4.1 where a node P desires to communicate with a node Q at distance r_0. The circle of radius R centered at P marks the interference and sensing area of P, labeled as S_P and likewise S_Q for Q. We will refer to S_{PQ}, the union of the S_P and S_Q regions, as the combined area "used" by the communication link between P and Q. The shaded area $S_1(r_0)$ is the *hidden area* as P cannot detect any transmissions in that area and any transmitters in $S_1(r_0)$ cannot detect P. The region labeled $S_3(r_0)$ is the *exposed area* as P can sense transmitters in $S_3(r_0)$ but they do not interfere with reception at Q.

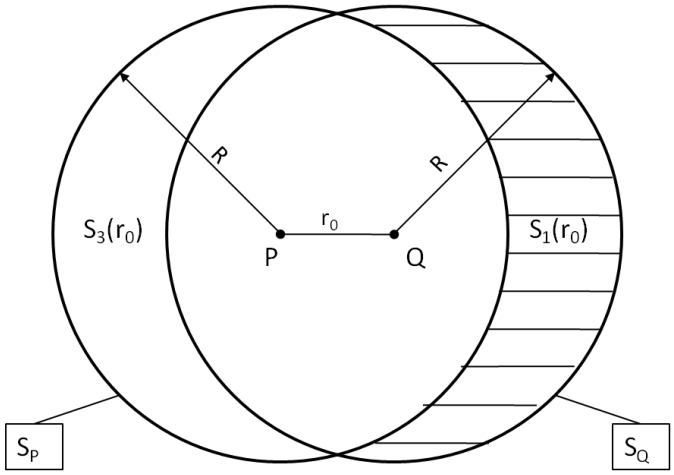

Figure 4.1 Transmission scenario.

Under slow fading and path loss, the average probability of detection for a single transmitting node uniformly distributed inside radius R, transmitting at power P_T is

$$\bar{P}_d = \int_0^R \int_0^\infty Q_{\lfloor sW \rfloor}\left(\sqrt{2\frac{AP_T\alpha^2}{N_oWr^\beta}}, \sqrt{\theta}\right) e^{-\alpha^2} f(r) d\alpha^2 dr. \tag{4.10}$$

If there are k transmitters present in S_P, each at a distance $\{r_i < R, i = 1, 2, \ldots, k\}$ and fade $\{\alpha_i^2, i = 1,2,..k\}$, respectively, and all of their signal powers sum at the receiver, then the probability of detection is

$$P_d = Q_{\lfloor sW \rfloor}\left(\sqrt{2\sum_i^k \frac{AP_T\alpha_i^2}{N_oWr_i^\beta}}, \sqrt{\theta}\right). \tag{4.11}$$

The average probability of detection for a node given that $k \geq 1$ transmitters are uniformly distributed in radius R is

$$\bar{P}_d(k|R, P_T) = \int_0^\infty Q_{\lfloor sW \rfloor}\left(\sqrt{2\frac{AP_T z_k}{N_oW}}, \sqrt{\theta}\right) f(z_k) dz_k, \tag{4.12}$$

where the random variable $z_k = \sum_{i=1}^k \frac{\alpha_i^2}{r_i{}^\beta}$. Let x_i be a set of i.i.d random variables $x_i = \frac{\alpha_i^2}{r_i{}^\beta}$, for $\beta > 2$ and $i = 1,..,k,$ then we can write the sum $z_k = \sum_i^k x_i$. Recalling that α_i^2 is an exponential random variable with unit mean and the PDF of r_i is given by (4.3), the PDF of x_i is

$$f_x(x_i) = \frac{2}{R^2\beta}\int_0^{R^\beta} (r^\beta)^{\frac{2}{\beta}} e^{-x_i r^\beta} dr^\beta, \tag{4.13}$$

for $x_i \geq 0$ and the PDF of z_k is the k-th convolution of $f_x(x_i)$, given by

$$f(z_k) = f_x(x_1) * f_x(x_2) * \ldots * f_x(x_k). \tag{4.14}$$

For a class of nodes on a channel with density λ, transmit power $P_{T,}$ and sensing radius R the overall probability that an energy detector senses the energy as greater than the given threshold is

$$PD(\lambda, R, P_T) = \sum_{k=0}^{\infty} e^{\lambda\pi R^2} \frac{(\lambda\pi R^2)^k}{k!} \bar{P}_d(k|R, P_T), where\ \bar{P}_d(0) = P_f. \tag{4.15}$$

The probability that one or more transmitters inside radius R from P are missed is

$$PM(\lambda, R, P_T) = \sum_{k=1}^{\infty} e^{\lambda\pi R^2} \frac{(\lambda\pi R^2)^k}{k!} (1 - \bar{P}_d(k|R, P_T)). \tag{4.16}$$

The average probability that a single node is missed, conditioned on the distance $\tilde{r}$ to the transmitter is

$$P\{Miss \mid \tilde{r} = r\} = 1 - \int_0^{\infty} Q_{\lfloor sW \rfloor}\left(\sqrt{2\frac{AP_T\alpha^2}{N_o W r^{\beta}}}, \sqrt{\theta}\right) e^{-\alpha^2} d\alpha^2. \tag{4.17}$$

The PDF of the distance to a missed transmitter from P, given there is a missed detection is found using Bayes' theorem [57]:

$$f(r|Miss) = \frac{f(r)P\{Miss \mid \tilde{r} = r\}}{\int_0^R f(r)P\{Miss \mid \tilde{r} = r\}dr}. \tag{4.18}$$

Since the probability of a miss is a function of distance only, the PDF of the angle $\emptyset$, to the missed node is uniform. Therefore, the joint PDF of the distance and angle to missed nodes is

$$f(r, \emptyset|Miss) = f(r|Miss) * \frac{1}{2\pi}, 0 < \phi \leq 2\pi. \tag{4.19}$$

Finally, we are interested in the distribution of the distance to the missed transmitter from the receiver Q. We do a transformation on $f(r, \emptyset|Miss)$ to find $f(r_m, \emptyset_m)$ where r_m and $\emptyset_m$ are the distance and angle between the missed transmitter and Q. As illustrated in Figure 4.2, we can use law of cosines and law of sines to express the angle $\emptyset$ and distance r from P of a missed transmitter M as follows.

$$r(r_m, \emptyset_m) = \sqrt{r_0^2 + r_m^2 - 2r_0 r_m cos(\pi - \emptyset_m)} \tag{4.20}$$

$$\emptyset(r_m, \emptyset_m) = sin^{-1}\left(\frac{r_m}{r(r_m, \emptyset_m)} sin(\pi - \phi_m)\right) \tag{4.21}$$

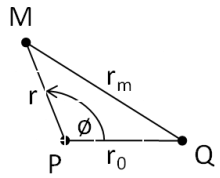

Figure 4.2 Law of cosines.

The joint density function for the distance and angle to the missed transmitter from Q is

$$f(r_m, \emptyset_m) = f(\, r(r_m, \emptyset_m), \emptyset(r_m, \emptyset_m, r) | Miss\,). \tag{4.22}$$

When there are k transmitters present, the probability that all are missed by P, conditioned on the respective distances r_i is

$$P\{\, Miss \mid \widetilde{r_1} = r_1, \ldots, \widetilde{r_k} = r_k\} = 1 - \int_0^\infty \int_0^\infty \ldots \int_0^\infty Q_{\lfloor sW \rfloor}\left(\sqrt{\sum_{i=1}^{k} 2\frac{AP_T{\alpha_i}^2}{N_o W {r_i}^\beta}}, \sqrt{\theta}\right) \prod_{i=1}^{k} e^{-{\alpha_i}^2} d{\alpha_i}^2. \tag{4.23}$$

The PDF of the distance to a missed transmitter from P, given there is a missed detection with k nodes present is

$$f(r|Miss) = \frac{f(r)\int_0^R \int_0^R \ldots \int_0^R P\{\, Miss \,|\widetilde{r_1} = r_1, \ldots, \widetilde{r_{k-1}} = r_{k-1}\} \prod_{i=1}^{k-1} f(r_i) dr_i}{\int_0^R \int_0^R \ldots \int_0^R P\{\, Miss \,|\widetilde{r_1} = r_1, \ldots, \widetilde{r_k} = r_k\} \prod_{i=1}^{k} f(r_i) dr_i}. \tag{4.24}$$

Calculation of this density function is time consuming due to multiple embedded integrals. Therefore, we approximate it using the PDF for a single missed transmitter given by (4.22). Given that k transmitters are present in S_P and missed by P, the probability that at least one transmitter is inside S_Q is 1 minus the probability that all *k* transmitters are no in S_Q:

$$P_{M1}(k, R, r_0) \cong 1 - \left(\int_R^{R+r_0} \int_0^{2\pi} f(r_m, \emptyset_m) d\emptyset_m \, dr_m\right)^k. \tag{4.25}$$

Similarly, the probability that all k missed transmitters are inside the exposed region $S_3(r_0)$, and do not interfere with reception at Q, is

$$P_{M2}(k, R, r_0) \cong \left(\int_R^{R+r_0} \int_0^{2\pi} f(r_m, \emptyset_m) d\emptyset_m \, dr_m\right)^k. \tag{4.26}$$

The overall probability that a transmitter is inside S_Q and missed by P is the total probability that there are k transmitters in S_P, all k transmitters are missed, and at least one is inside S_Q:

$$PM1(\lambda, R, r_0, P_T) = \sum_{k=1}^{\infty} e^{\lambda\pi R^2} \frac{(\lambda\pi R^2)^k}{k!} (1 - \bar{P}_d(k|R, P_T)) P_{M1}(k, R, r_0). \quad (4.27)$$

Likewise, the probability that a missed detection occurs and the missed transmitter is not in S_Q is

$$PM2(\lambda, R, r_0, P_T) = \sum_{k=1}^{\infty} e^{\lambda\pi R^2} \frac{(\lambda\pi R^2)^k}{k!} (1 - \bar{P}_d(k|R, P_T)) P_{M2}(k, R, r_0). \quad (4.28)$$

4.3.1 Detection at Nearby SU

In this section, we examine the detection probabilities at a nearby SU immediately before and during a transmission from P to Q, as shown in Figure 4.1. We assume that other DSA nodes also use sensing in a listen-before-talk way to control their transmissions. Assuming there are no hidden nodes transmitting when P performs detection, if P finds the channel to be clear without a missed detection, then there are no transmitting nodes in S_{PQ}. The interference scenario where an incidental DSA node X turns on inside S_{PQ} is shown in Figure 4.3. $S_X(\phi_X, r_X, r_0, R)$ is the shaded sensing region of X which is non-overlapping with those of P and Q. As seen in Figure 4.3, ϕ_x is the angle from the *PQ* axis to *X*, and r_x is the distance from *X* to *Q*. The distance to interferer is from the receiver Q unless specified otherwise. The node X can turn on at the same time as P or at some point during P's transmission.

Since X also performs sensing, the turn-on probability of X depends on the presence of nearby transmitters. Assuming that there are no nodes transmitting inside S_{PQ}, other than P, and any active transmitters outside of S_{PQ} are distributed according to the Poisson process, the probability that X senses the channel to be clear depends on X's proximity to P and probability of finding an active transmitter in the region $S_X(\phi_X, r_X, r_0, R)$. There is also a possibility of false alarm and missed detection at X.

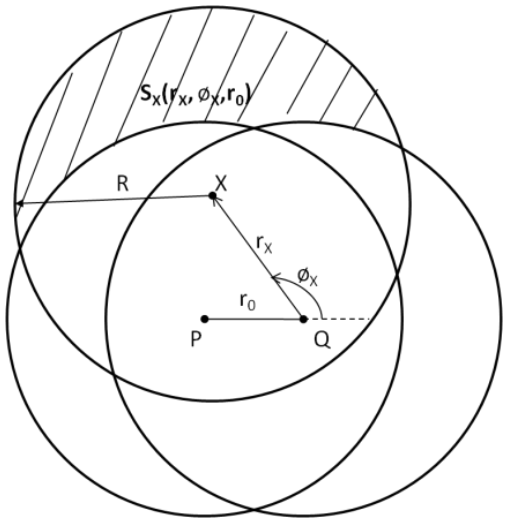

Figure 4.3 Incidental Node X.

The probability of detection at X is found by first finding the density of the distance from X to a potential transmitter in the region $S_X(\phi_X, r_X, r_0, R)$. The density of the distance from X to a fourth node Y, given that Y is present in the region $S_X(\phi_X, r_X, r_0, R)$ is found using Bayes' theorem [57]:

$$f(r_Y \mid r_x, \phi_x) = \frac{\int^{S_X(r_x,\phi_x,r_0)} f(r,\emptyset)d\emptyset}{\iint^{(S_X(r_x,\phi_x,r_0)} f(r,\emptyset)drd\emptyset}, \tag{4.29}$$

where $f(r,\emptyset)$ is given by (4.2). When there are no transmitters present inside S_{PQ}, the probability of detection at X is

$$\bar{P}_{dX1}(k \mid r_x, \phi_x, R, P_T) = \int_0^\infty Q_{\lfloor sW \rfloor}\left(\sqrt{2\frac{AP_T z_{kY}}{N_o W}}, \sqrt{\theta}\right) f(z_{kY}|r_x, \phi_x)dz_{kY}, \tag{4.30}$$

for $k > 0$, where $z_{kY} = \sum_{i=1}^{k} \frac{{\alpha_i}^2}{{r_{Y_i}}^2}$. Similar to $f(z_k)$ in (4.14), the PDF of z_{kY} is found through k convolutions of the PDF of x_i conditioned on the location of X, which is

$$f(x|r_x, \phi_x) = \frac{1}{\beta}\int_0^{R^\beta} \frac{r_Y^\beta}{{r_Y}^{1+\frac{1}{\beta}}} f(r_Y|r_x, \phi_x)e^{-xr_Y^\beta} dr_Y^\beta, \tag{4.31}$$

for $x \geq 0$ where $f(r_Y|r_x, \phi_x)$ is given by (4.29). When $k = 0$, $\bar{P}_{dX1}$ is equal to the probability of false alarm. For a class of nodes on a channel with density λ, transmit power P_T, and sensing radius R the overall probability that a third node X located at distance r_x and angle $\emptyset_x$ senses the channel as busy, when P is not transmitting, is

$$PDX1(\lambda, R, P_T, r_x, \phi_x) = \sum_{k=0}^{\infty} e^{\lambda S_X(r_x,\phi_x,r_0)} \frac{(\lambda S_X(r_x, \phi_x, r_0))^k}{k!} \bar{P}_{dX1}(k \mid r_x, \phi_x, R, P_T). \tag{4.32}$$

When P is transmitting the average probability of detection at X is a function of the combined SNR from P and k other transmitters in the region S_X is

$$\bar{P}_{dX2}(k \mid r_x, \phi_x, R, P_T) = \int_0^\infty \int_0^\infty Q_{\lfloor sW \rfloor}\left(\sqrt{2\frac{A}{N_o W}\left(P_T z_{kY} + \frac{P_0 \alpha^2}{r_{xp}}\right)}, \sqrt{\theta}\right) e^{-\alpha^2} f(z_{kY})d\alpha^2 dz_{kY}, \tag{4.33}$$

where $r_{xp} = \sqrt{r_0^2 + r_x^2 - 2r_0 r_x cos(\pi - \emptyset_x)}$ for $r_{xp} < R$. The overall probability that X senses the channel as busy, when P is transmitting, is

$$PDX2(\lambda, R, P_T, r_x, \phi_x) = \sum_{k=0}^{\infty} e^{\lambda S_X(r_x,\phi_x,r_0)} \frac{(\lambda S_X(r_x, \phi_x, r_0))^k}{k!} \bar{P}_{dX2}(k \mid r_x, \phi_x, R, P_T). \tag{4.34}$$

4.4 System Model

4.4.1 DSA Access Scheme

The DSA protocol uses a dedicated control channel and random access for both control and data channels, which is best suited for ad hoc networks. We assume each node is equipped with only one transceiver which monitors a common control channel and then switches to one of M data channels for data transmissions. All data channels have common bandwidth W. We assume that all nodes see equivalent RF conditions on each channel, so channel selection occurs with uniform probability, $1/M$. All of the data channels are equally likely to be occupied by one or more primary users, while the control channel is dedicated to the secondary DSA network.

If a node has data to transmit it sends an ATM (Ad hoc Traffic indicator Message) packet, which contains the intended receiver's address and a selected data channel, on the control channel using p-persistent CSMA. The transmitter then waits for a SIFS (Short Inter-Frame Space) period for acknowledgement from the receiver. Upon successfully receiving an ACK from the intended receiver, both transmitter and receiver sense the selected data channel for primary or secondary users. Sensing is performed using energy detection, as described in Section 4.3, for duration L_{SENSE}. To ensure that the channel is free on both ends, the nodes use a four-way RTS/CTS handshake to transmit one data packet of length L_{DATA}.

Since a primary user could come on at any time, both nodes pause and sense the channel for primary and secondary users before each data packet. We assume the length of one data packet to be less than or equal to the periodic sensing duration required by the FCC. After transmitting the desired number of data packets, or upon finding the data channel busy at either node, the nodes return to the control channel and start the procedure again. Also, we assume all of the control packets have equal lengths such that $L_{ATM} = L_{RTS} = L_{CTS} = L_{ACK.}$ The access procedure is illustrated in Figure 4.4 for two node pairs which select two different channels.

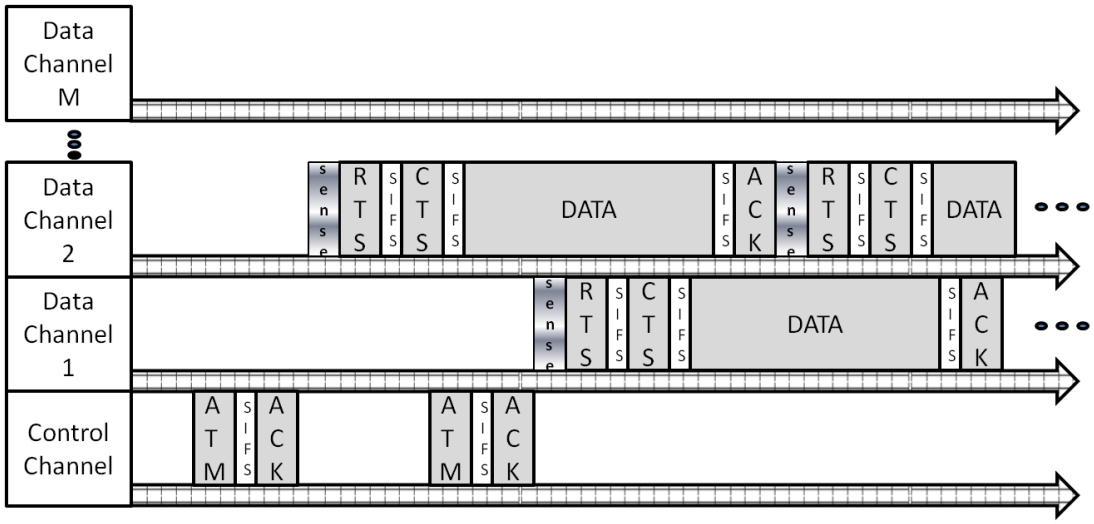

Figure 4.4 DSA access scheme.

4.4.2 Markov Model

As in [54], we approximate an asynchronous network by a slotted system, where each slot has duration L_{SLOT}, which is much greater than the propagation delay between transmitter and receiver. Given the access procedure described in Section 4.4.1, and following the method proposed in [54] for single channel CSMA networks, observe that a node may be in one of the following seven states at any given point in time:

***I*: Idle** - A node has no traffic or is waiting for the control channel to clear before transmitting. This state has duration equal to one slot, $D_I = L_{SLOT}$.

***F_C*: Failure on control channel** – A node is unsuccessfully transmitting over the control channel. This state starts when the control channel is sensed clear and an ATM is sent; if the ATM packet fails it ends after the L_{SIFS} period expires. If the ATM packet is successful but the ACK fails it ends when the ACK would have finished. This state has duration $D_{F_C}{}^1 = L_{SENSE} + L_{ATM} + L_{SIFS}$ or $D_{F_C}{}^2 = L_{SENSE} + L_{ATM} + L_{SIFS} + L_{ACK}$.

***S_C*: Success on control channel** – A node is successfully transmitting over the control channel. This state starts when the control channel is sensed clear and an ATM is sent and ends when the ACK is received. This state has duration $D_{S_C} = L_{SENSE} + L_{ATM} + L_{SIFS} + L_{ACK}$.

***Sense*: Sensing on chosen data channel** – After a success on the control channel both the transmitter and receiver sense the chosen data channel for duration L_{SENSE}. Since we assume cooperative sensing through the exchange of RTS and CTS messages, this state has multiple possible durations $D_{Sense}{}^1 = L_{SENSE}$, $D_{Sense}{}^2 = L_{SENSE} + L_{RTS} + L_{SIFS}$, or $D_{Sense}{}^3 = L_{SENSE} + L_{RTS} + L_{CTS} + 2L_{SIFS}$.

***P.Sense*: Periodic Sensing on chosen data channel** – After a successful transmission of a data packet, if P has more data for Q the nodes stop transmission and cooperatively sense the channel through a RTS/CTS exchange. This state has the same possible durations as the *Sense* state.

***S_D*: Success on Data Channel** – A node is successfully transmitting a data packet over a data channel after sensing it as clear. This state includes the data and the acknowledgement packets, having duration $D_{S_D} = L_{DATA} + L_{SIFS} + L_{ACK}$.

***F_D*: Failure on Data channel** – A node is unsuccessfully transmitting over a data channel. If the data packet fails, it ends L_{SIFS} after the data packet. If the data packet is successful but the ACK fails it ends when the ACK ends. This state has duration $D_{F_D}{}^1 = L_{SIFS} + L_{DATA}$, or $D_{F_D}{}^2 = 2L_{SIFS} + L_{DATA} + L_{ACK}$.

Note that L_{SENSE} is included in D_{S_C} and D_{F_C} in order to attribute the time that is spent immediately sensing the control channel before a transmission to one of the transmission states (S_C or F_C). Figure 4.5 illustrates the states of a node-pair which is idle for the first 3 slots, then transitions to a successful transmission on the control channel, then to a sensing state followed by a successful transmission on the data channel. The pair will then periodically go from sensing to transmitting until it runs out of data packets to send or one of the data packets fails due to interference from primary or secondary users.

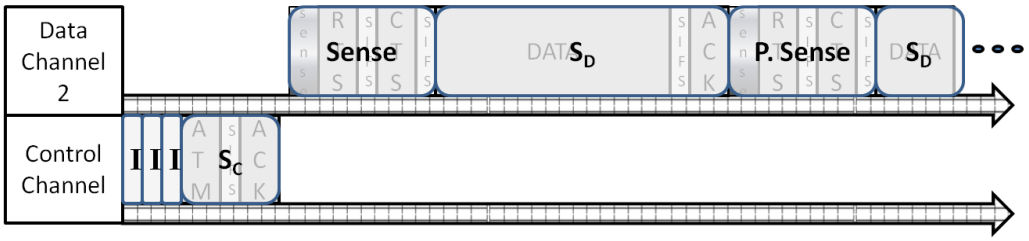

Figure 4.5 Temporal illustration of Markov states.

The states form an embedded Markov chain for the semi-Markov process, for which the state diagram is illustrated in Figure 4.6. At each slot, a node in the Idle state may stay in the Idle state with probability P_{II}, or transition to either the F_C or S_C state with probability P_{IFC} or P_{ISC}, respectively. Once in the F_C state a node transitions to the Idle state with probability 1 after the appropriate duration, and from the S_C state it transitions to the *Sense* state with probability 1 after D_{SC} slots. A node from the Sense state may transition to S_D, F_D or the Idle state with probability $P_{SenseSD}$, $P_{SenseFD}$, or P_{SenseI}, respectively and the appropriate duration. If the node transitions to the S_D state, it may keep cycling between S_D and *P.Sense* states until the data session ends, the spectrum is no longer available, or the transmission fails, and then return to the Idle state. If at any time the transmission fails the node goes from the F_D state back to the Idle state with probability 1. The transition probabilities between the states are derived in the Sections 4.4.2.1 to 4.4.2.5 as functions of the receiver distance, network density, the traffic load and other network parameters. In Section 4.4.2.5, we derive steady state and limiting probabilities for the Markov model described in this section.

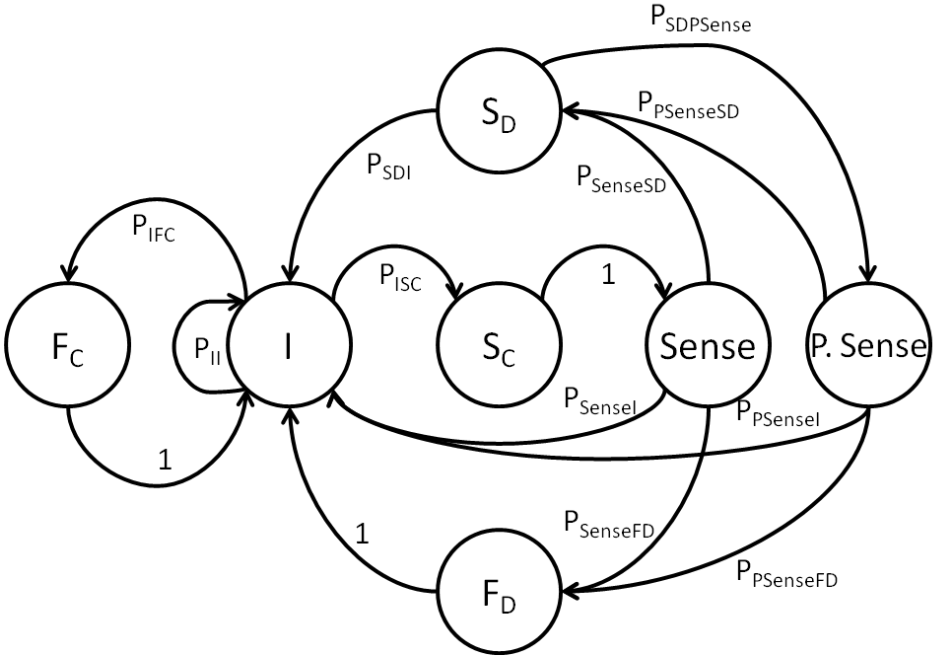

Figure 4.6 State diagram for embedded Markov chain.

4.4.2.1 Transitioning from the Idle State

The following analysis derives the transition probability for a node P, which wishes to communicate with a node Q at distance r_0 as shown in Figure 4.1. We assume the DSA system is dedicated an orthogonal channel for session management, therefore only SUs are present for sensing and interference on the control channel. A node in the idle state transmits on the control channel if it has traffic and senses the channel to be clear with probability

$$p_C = p_T\left(1 - PD\left(n\left(P_{S_C} + P_{F_C}\right), R, P_0\right)\right), \tag{4.35}$$

where *PD* is given by (4.15). Therefore, at each slot, the probability that a node does not transition out of the Idle state is

$$P_{II} = 1 - p_C. \tag{4.36}$$

At each slot a node in the idle state transmits on the control channel with probability p_c, whether this transmission fails or succeeds determines if the transition is into the S_C or F_C state. We adapt an interference model similar to one used in [54] for CSMA networks, where the transmission fails with probability 1 if another node inside radius *R* from the receiver transmits at any point during the transmission. We classify interfering users as either *pre-existing* or *incidental*. Pre-existing interferers are transmitting on the channel prior to P's transmission but are not detected by P because they are out of its sensing range (hidden) or missed by the detector. Incidental interferers turn on at the same time as P, or after if they fail to detect P's transmission. Since other DSA nodes also employ sensing, the turn on probability of an incidental node is dependent on its location.

Given that P senses the channel to be clear, there are no missed detections or hidden nodes, the probability that Q transmits in the same slot is

$$P_Q = p_T(1 - P_f). \tag{4.37}$$

The probability that a node transitions from the idle state to the S_C state is the total probability of the following series of events:

- P transmits to Q with probability p_C.
- No hidden nodes are present with probability P_{HCC}.
- No missed detections occur with probability $1 - P_{M1} - P_{M2}$.
- Q does not transmit in the same slot as P with probability $1 - P_Q$.
- No incidental nodes in S_Q turn on in the same slot as P with probability P_{X1CC}.
- No incidental nodes in S_Q turn on during P's transmission with probability ${P_{X2CC}}^{L_{ATM}}$.
- No incidental nodes in S_P turn on while Q transmits the ACK with probability ${P_{X2CC}}^{L_{ACK}}$.
- No incidental nodes in $S_3(r_0)$ turn on in the same slot as P with probability P_{X3CC}.
- No incidental nodes in $S_3(r_0)$ turn on during P's transmission with probability ${P_{X4CC}}^{L_{ATM}}$.

While the individual components are derived next, the transition probability is

$$P_{IS_C} = p_C P_{HCC}(1 - P_{M1} - P_{M2})(1 - P_Q) P_{X1CC} {P_{X2CC}}^{L_{ATM}+L_{ACK}} P_{X3CC} {P_{X4CC}}^{L_{ATM}+L_{ACK}}. \quad (4.38)$$

Similarly, the probability that a node transitions from the idle state to the F_C state due to ATM packet failure is the total probability of the following series of events:

- P transmits to Q with probability p_C.
- A hidden node is present with probability $1 - P_{HCC}$, or no hidden nodes are present with probability P_{HCC}, and
- an active transmitter in the region S_Q is missed by P with probability P_{M1}, or there is no active transmitter in the region S_Q that is missed by P with probability $1 - P_{M1}$, and
- Q transmits in the same slot as P with probability P_Q, or Q does not transmit in the same slot as P with probability $1 - P_Q$, and
- one or more incidental nodes in S_Q turn on in the same slot as P with probability $1 - P_{X1CC}$, or no incidental nodes in S_Q turn on in the same slot as P with probability P_{X1CC}, and
- one or more incidental nodes in S_Q turn on during the ATM packet with probability $1 - {P_{X2CC}}^{L_{ATM}}$.

Therefore, the partial transition probability from the idle to the F_C due to ATM failure is

$$P_{IF_C}{}^1 = p_C \Big[1 - P_{HCC} + P_{HCC} \Big[P_{M1} + (1 - P_{M1}) \Big[P_Q + (1 - P_Q)\big[1 - P_{X1CC} + P_{X1CC}\big(1 - {P_{X2CC}}^{L_{ATM}}\big)\big]\Big]\Big]\Big]. \quad (4.39)$$

Similarly, the probability that a node transitions from the idle state to the F_C state due to ACK failure is

$$P_{IF_C}{}^2 = p_c P_{HCC}(1 - P_Q) P_{X1CC} {P_{X2CC}}^{L_{ATM}} \Big[P_{M2} + (1 - P_{M2}) \Big[1 - P_{X3CC} + P_{X3CC}\big[1 - {P_{X4CC}}^{L_{ATM}} + {P_{X4CC}}^{L_{ATM}}\big(1 - {P_{X2CC}}^{L_{ACK}}\big)\big]\Big]\Big]. \quad (4.40)$$

Thus, the probability that a node transitions from idle to F_C is the sum

$$P_{IF_C} = P_{IF_C}{}^1 + P_{IF_C}{}^2. \quad (4.41)$$

Entering F_C due to ATM failure gives the duration $D_{F_C}{}^1$, while ACK failure gives the duration $D_{F_C}{}^2$. Therefore, for each transition into F_C, the average duration is the mean of the two cases, divided by the sum probability of transitioning in to the F_C state:

$$D_{F_C} = \frac{D_{F_C}{}^1 P_{IF_C}{}^1 + D_{F_C}{}^2 P_{IF_C}{}^2}{P_{IF_C}{}^1 + P_{IF_C}{}^2}. \quad (4.42)$$

Pre-existing Interferers

In the event of a missed detection at P, the transmission fails immediately if the missed transmitter(s) is in S_Q, which happens with probability

$$P_{M1} = PM1(n(P_{SC} + P_{FC}), R, r_0, P_0), \tag{4.43}$$

where $PM1$ is given by (4.27). Otherwise the return ACK fails at P with probability

$$P_{M2} = PM2(n(P_{SC} + P_{FC}), R, r_0, P_0), \tag{4.44}$$

where $PM2$ is given by (4.28). As seen in Figure 4.1, active nodes in the region $S_1(r_0)$ may cause interference to Q but are hidden from P. The probability that there are no hidden nodes present when P starts transmitting is the probability of finding 0 Poisson distributed points in the hidden area, which is

$$P_{HCC} = e^{-|S_1(r_0)|n(P_{SC}+P_{FC})}, \tag{4.45}$$

where $|S_1(r_0)|$ is the hidden area as shown in Figure 4.1

Incidental Interferers

If no pre-existing nodes are active when P transmits, the transmission may fail due to incidental nodes. Immediately prior to the slot where P begins transmission, P is off and there are no other nodes inside R from P and Q. Then, the turn on probability of an incidental node X at angle ϕ_x and distance r_x (see Figure 4.3) is

$$p_2(r_x, \phi_x, r_0) = p_T * PDX1(n(P_{SC} + P_{FC}), R, P_0, r_x, \phi_x), \tag{4.46}$$

where the probability *PDX1* is given by (4.32). If no other node turns on at the same time as P, the probability that a node X at angle ϕ_x and distance r_x turns on at a slot while P is on is

$$p_3(r_x, \phi_x, r_0) = p_T * PDX2(n(P_{SC} + P_{FC}), R, P_0, r_x, \phi_x), \tag{4.47}$$

where the probability *PDX2* is given by (4.34). Since X has Poisson distribution, the average turn on probability of a single node inside of radius R, assuming the turn on probability at each point is $p(r_x, \phi_x)$, is the area integral

$$\iint_{0\,0}^{2\pi\,R} f(r_x, \emptyset_x) p(r_x, \phi_x) dr_x d\phi_x. \tag{4.48}$$

Therefore, the probability that zero nodes inside S_Q turn on at the same time as P is

$$P_{X1CC} = e^{-n\pi R^2 \iint_{0\,0}^{2\pi\,R} f(r_x,\emptyset_x) p_2(r_x,\phi_x,r_0) dr_x d\phi_x}. \tag{4.49}$$

And the probability that no nodes inside S_Q turn on in a slot during P's transmission is

$$P_{X2CC} = e^{-n\pi R^2 \iint_{0,0}^{2\pi,R} f(r_x,\emptyset_x)p_3(r_x,\phi_x,r_0)dr_x d\phi_x}. \quad (4.50)$$

Similarly, the probability that none of the nodes in the exposed region $S_3(r_0)$ turn on in the same slot as P and interfere with the return ACK is

$$P_{X3CC} = e^{-n|S_3(r_0)| \iint^{S_3(r_0)} f(r_x,\emptyset_x)p_2(r_x,\phi_x,r_0)dr_x d\phi_x}. \quad (4.51)$$

The probability that none of the nodes in the exposed region $S_3(r_0)$ turn on in a slot during P's transmission and interfere with the return ACK is

$$P_{X4CC} = e^{-n|S_3(r_0)| \iint^{S_3(r_0)} f(r_x,\emptyset_x)p_3(r_x,\phi_x,r_0)dr_x d\phi_x}. \quad (4.52)$$

4.4.2.2 Transitioning from the Sense State

Once a pair of nodes successfully rendezvous on the control channel, both nodes tune to one of M data channels to transfer the payload. Since data channels are shared, each node independently senses the channel to avoid interference to primary users. If both nodes sense the channel clear, they handshake on the data channel and proceed to transfer data. The node pair periodically shifts from sensing to data transfer states until the data session ends or fails due to interference, after which the nodes return to the Idle state. Sensing on the data channels is similar to the control channel with the addition of primary users. We assume nodes perform sensing using energy detection with threshold θ_D and equivalent noise on each channel.

Active primary users are distributed by a Poisson process with intensity mq on any given data channel. A secondary user is may be transmitting on a data channel with probability $P_{S_D} + P_{Sense} + P_{F_D}$. If there are M data channels, the active secondary users on any data channel are distributed by a Poisson process with intensity

$$n' = n\left(\frac{P_{S_D} + P_{F_D} + P_{Sense} + P_{PSense}}{M}\right) \quad (4.53)$$

when there are no primary users present. When primary users are present, secondary users are likely to be present with a much lower density since each secondary user performs sensing, however modeling the density and distribution of a SU network in the presence of PUs is difficult. We assume the worst case for energy detection on data channels where there are either secondary or primary users present, *but never both*. Upon having a successful handshake on the control channel and randomly selecting a data channel, the probability that P senses a data channel clear is found as follows using (4.15):

$$p_D = e^{-mq\pi {R_p}^2} PD(n', R, P_0) + PD(mq, R_p, P_1). \quad (4.54)$$

If the channel is sensed clear, P sends a RTS packet to check the status at Q. The RTS/CTS exchange must be successful before the data packet is sent. If the channel is sensed busy or if the RTS/CTS handshake fails the node returns to the Idle state with probability

$$P_{SenseI} = (1 - p_D) + P_{RTSf} + P_{RTSs}P_{CTSf}, \tag{4.55}$$

where

- P_{RTSf} is the probability that P sends an RTS but no CTS is sent in return,
- P_{RTSs} is the probability that the RTS is successful and a CTS packet is sent,
- P_{CTSf} is the probability that the CTS packet is unsuccessful,
- although not used in (4.55), P_{CTSs} is the probability that the CTS is successful.

Given the above probabilities, the average duration of the *Sense* state is

$$D_{Sense} = \frac{(1 - p_D){D_{Sense}}^1 + P_{RTSf}{D_{Sense}}^2 + P_{RTSs}{D_{Sense}}^3}{(1 - p_D) + P_{RTSf} + P_{RTSs}}. \tag{4.56}$$

Deriving P_{RTSf}, P_{RTSs}, P_{CTSf}, and P_{CTSs} requires the following individual probabilities for possible sources of interference:

- $P_{MS1} + P_{MP1}$, the probability of a missed SU or PU causing RTS failure.
- $P_{MS2} + P_{MP2}$, the probability of a missed SU or PU causing CTS failure.
- P_{HPU}, the probability there is no hidden PU on data channel.
- P_{HSU}, the probability there is no hidden SU on data channel.
- P_{X1DC}, the probability no SU in S_Q turns on at same slot as P.
- P_{X2DC}, the probability no SU in S_Q turns on at a slot during P's transmission.
- P_{X3DC}, the probability no SU in $S_3(r_0)$ turns on at same slot as P.
- P_{X4DC}, the probability no SU in $S_3(r_0)$ turns on at a slot during P's transmission.
- P_{PU1}, the probability that no PU inside radius R_p from Q turns on in a slot during P's transmission.
- P_{PU2}, the probability that no PU in $S_3(r_0)$ (with radius R_p) turns on in a slot during P's transmission.

These probabilities are derived in the subsequent subsections, while here we present the probabilities associated with the RTS/CTS handshake and the final transition probabilities from the Sense state.

The probability that P senses the channel to be clear and sends an RTS but no CTS is sent in return, either because the RTS fails due to interference or Q has a false alarm and believes the channel to be occupied is

$$P_{RTSf} = p_D\Big[1 - P_{HPU} + P_{HPU}\Big[1 - P_{HSU} + P_{HSU}\Big[P_{MS1} + P_{MP1} + (1 - P_{MS1} - P_{MP1})\Big[1 - P_{X1DC} + P_{X1DC}\Big[1 - {P_{PU1}}^{L\_RTS} + {P_{PU1}}^{L\_RTS}\big[1 - {P_{X2DC}}^{L_{RTS}} + {P_{X2DC}}^{L_{RTS}}P_f\big]\Big]\Big]\Big]\Big]\Big]. \quad (4.57)$$

The probability that an RTS packet is received successfully, and a CTS packet is sent, is

$$P_{RTSs} = p_D P_{HPU} P_{HSU}(1 - P_{MS1} - P_{MP1}) P_{X1DC} {P_{PU1}}^{L_{RTS}} {P_{X2DC}}^{L_{RTS}}(1 - P_f). \quad (4.58)$$

The probability that a CTS packet is unsuccessful is

$$P_{CTSf} = P_{MS2} + P_{MP2} + (1 - P_{MS2} - P_{MP2})\Big[1 - P_{X3DC} + P_{X3DC}\Big[1 - {P_{PU2}}^{L_{RTS}} + {P_{PU2}}^{L_{RTS}}\Big[1 - {P_{X4DC}}^{L_{RTS}} + {P_{X4DC}}^{L_{RTS}}\big[1 - {P_{PU1}}^{L_{CTS}} + {P_{PU1}}^{L_{CTS}}\big(1 - {P_{X2DC}}^{L_{CTS}}\big)\big]\Big]\Big]\Big]. \quad (4.59)$$

The probability that a CTS packet is successful is

$$P_{CTSs} = (1 - P_{MS2} - P_{MP2}) P_{X3DC} {P_{X4DC}}^{L_{RTS}} {P_{PU2}}^{L_{RTS}} {P_{PU1}}^{L_{CTS}} {P_{X2DC}}^{L_{CTS}}. \quad (4.60)$$

If P and Q have a successful handshake on the chosen data channel, the probability that a node transitions from the *Sense* state to the S_D state is

$$P_{SenseS_D} = P_{RTSs} P_{CTSs} {P_{PU1}}^{L_{DATA}+L_{ACK}} {P_{X2DC}}^{L_{DATA}+L_{ACK}} {P_{PU2}}^{L_{DATA}} {P_{X4DC}}^{L_{DATA}}. \quad (4.61)$$

The probability that a node transitions from the *Sense* state to the F_D state due a data packet failure is

$${P_{SenseF_D}}^{1} = P_{RTSs} P_{CTSs}\Big[1 - {P_{PU1}}^{L_{DATA}} + {P_{PU1}}^{L_{DATA}}\big[1 - {P_{X2DC}}^{L_{DATA}}\big]\Big]. \quad (4.62)$$

Similarly, the probability that a node transitions from *Sense* to F_D due to ACK failure is

$${P_{SenseF_D}}^{2} = P_{RTSs} P_{CTSs} {P_{PU1}}^{L_{DATA}} {P_{X2DC}}^{L_{DATA}}\Big[1 - {P_{PU2}}^{L_{DATA}} + {P_{PU2}}^{L_{DATA}}\Big[1 - {P_{X4DC}}^{L_{DATA}} + {P_{X4DC}}^{L_{DATA}}\Big[1 - {P_{PU1}}^{L_{ACK}} + {P_{PU1}}^{L_{ACK}}\big[1 - {P_{X2DC}}^{L_{ACK}}\big]\Big]\Big]\Big]. \quad (4.63)$$

The overall probability of transitioning from *Sense* to F_D is

$$P_{SenseF_D} = P_{SenseF_D}{}^{1} + P_{SenseF_D}{}^{2} \tag{4.64}$$

Pre-existing Interferers

In the event of a missed detection at P, the RTS packet fails with probability

$$P_{MS1} = PM1(n', R, P_0) \tag{4.65}$$

if there is an SU inside radius R from Q, and with probability

$$P_{MP1} = PM1(mq, R_P, P_1) \tag{4.66}$$

if there is a PU inside radius R_p from Q. Recall that the conditional probability $PM1$ and $PM2$ are given in (4.27) and (4.29). A missed detection may cause the returning CTS packet to fail at P with the sum probability:

$$P_{MS2} + P_{MP2} = PM2(n', R, P_0) + PM2\big(mq, R_p, P_1\big). \tag{4.67}$$

The probability that there is no hidden PU present when P starts transmitting is

$$P_{HPU} = e^{-|S_1(r_0)|mq}. \tag{4.68}$$

Given that there is no hidden PU present, the probability that there is no hidden SU present when P starts transmitting is

$$P_{HSU} = e^{-|S_1(r_0)|n'}. \tag{4.69}$$

Incidental Secondary Interferers

Given that P senses the channel clear and transmits an RTS to Q, a third secondary user X inside the sensing radius of Q may desire to turn on in a slot with probability $\frac{P_{S_C}}{MD_{S_C}}$.

When P is off, there are no other active SUs inside S_P and S_Q, and no PUs inside radius R_P, the turn on probability of a node X conditioned on its location is

$$\begin{aligned} p_4(r_x, \phi_x, r_0) = \frac{P_{S_C}}{MD_{S_C}} \big[& e^{-mq|S_x(r_0, r_x, \emptyset_x, R_p)|} PDX1(n', R, \mathrm{P}_0, r_x, \phi_x) \\ & + PDX1\big(mq, R_p, \mathrm{P}_1, r_x, \phi_x\big)\big]. \end{aligned} \tag{4.70}$$

The turn on probability of X when P is on is

$$\begin{aligned} p_5(r_x, \phi_x, r_0) = \frac{P_{S_C}}{MD_{S_C}} \big[& e^{-mq|S_x(r_0, r_x, \emptyset_x, R_p)|} PDX2(n', R, \mathrm{P}_0, r_x, \phi_x) \\ & + PDX2\big(mq, R_p, \mathrm{P}_1, r_x, \phi_x\big)\big]. \end{aligned} \tag{4.71}$$

Considering the density of the network and its Poisson distribution, the probability that no SU inside S_Q turns on at the same time as P is

$$P_{X1DC} = e^{-n\pi {R_s}^2 \iint_{0\,0}^{2\pi\, R_S} f_{\tilde{r}\tilde{\emptyset}}(r_x,\emptyset_x) p_4(r_x,\phi_x,r_0) dr_x d\phi_x}. \quad (4.72)$$

And the probability that no SUs inside radius S_Q turn on in a slot during P's transmission is

$$P_{X2DC} = e^{-n\pi {R_s}^2 \iint_{0,0}^{2\pi,R_S} f_{\tilde{r}\tilde{\emptyset}}(r_x,\emptyset_x) p_5(r_x,\phi_x,r_0) dr_x d\phi_x}. \quad (4.73)$$

Similarly, the probability that no SU in the exposed region $S_3(r_0)$ turns on in the same slot as P and interferes with the return ACK is

$$P_{X3DC} = e^{-n|S_3(r_0)| \iint^{S_3(r_0)} f_{\tilde{r}\tilde{\emptyset}}(r_x,\emptyset_x) p_4(r_x,\phi_x,r_0) dr_x d\phi_x}. \quad (4.74)$$

The probability that no SU in the exposed region $S_3(r_0)$ turns on in a slot during P's transmission and interferes with the return ACK is

$$P_{X4DC} = e^{-n|S_3(r_0)| \iint^{S_3(r_0)} f_{\tilde{r}\tilde{\emptyset}}(r_x,\emptyset_x) p_5(r_x,\phi_x,r_0) dr_x d\phi_x}. \quad (4.75)$$

Incidental Primary Users

The probability that no primary users turn on and cause interference to an ongoing transmission is

$$P_{PU1} = e^{-mq\pi {R_p}^2}. \quad (4.76)$$

The probability that a PU turns on in the exposed region and causes interference to the returning packet is

$$P_{PU2} = e^{-mq|S_3(r_0,R_p)|}. \quad (4.77)$$

4.4.2.3 Transitioning from the S_D State

We assume that, after successfully transmitting the initial data packet, the number of packets each node desires to send is independent and geometrically distributed with parameter x (so a data session has a mean desired length of $1 + 1/x$ data packets). Therefore, the probability that a node moves from S_D to the *P.Sense* state is

$$P_{S_D Psense} = x, \quad (4.78)$$

and the probability that a node transitions to the Idle state is

$$P_{S_D I} = 1 - x. \quad (4.79)$$

4.4.2.4 Transitioning from the P.Sense State

The transition probabilities from the P.Sense state are found similar to the ones from the Sense state with the following changes:

- Since P must have successfully received an ACK from Q in the slot prior to transitioning to the *P.Sense* state, the probability that P senses the channel to be clear and transmits an RTS is the probability that it does not have a false alarm: $p_D{}' = 1 - P_f$.
- The probability that there is no hidden PU present when P starts transmitting is the probability that no PU turned on in Q's exposed PU region during L_{ACK}: $P'_{HPU} = {P_{PU2}}^{L_{ACK}}$.
- Given that there is no hidden PU present, the probability that there is no hidden SU present when P starts transmitting is: $P'_{HSU} = {P_{X4DC}}^{L_{ACK}}$.

Therefore, the probability that P senses the channel to be clear and sends an RTS, but no CTS is sent in return, is

$$P'_{RTSf} = p_D{}'\Bigg[1 - P'_{HPU} + P'_{HPU}\Bigg[1 - P'_{HSU} + P'_{HSU}\Big[1 - P_{X1DC} + P_{X1DC}\Big[1 - {P_{PU1}}^{L_{RTS}} + {P_{PU1}}^{L_{RTS}}\big[1 - {P_{X2DC}}^{L_{RTS}} + {P_{X2DC}}^{L_{RTS}}P_f\big]\Big]\Big]\Bigg]\Bigg]. \tag{4.80}$$

The probability that an RTS packet is successful, and a CTS packet is sent, is

$$P'_{RTSs} = p_D{}'P'_{HPU}P'_{HSU}P_{X1DC}{P_{PU1}}^{L_{RTS}}{P_{X2DC}}^{L_{RTS}}(1 - P_f). \tag{4.81}$$

The probability that a CTS packet is unsuccessful is

$$P'_{CTSf} = 1 - P_{X3DC} + P_{X3DC}\Big[1 - {P_{PU2}}^{L_{RTS}} + {P_{PU2}}^{L_{RTS}}\Big[1 - {P_{X4DC}}^{L_{RTS}} + {P_{X4DC}}^{L_{RTS}}\big[1 - {P_{PU1}}^{L_{CTS}} + {P_{PU1}}^{L_{CTS}}\big(1 - {P_{X2DC}}^{L_{CTS}}\big)\big]\Big]\Big]. \tag{4.82}$$

The probability that a CTS packet is successful is

$$P'_{CTSs} = P_{X3DC}{P_{X4DC}}^{L_{RTS}}{P_{PU1}}^{L_{CTS}}{P_{PU2}}^{L_{RTS}}{P_{X2DC}}^{L_{CTS}}. \tag{4.83}$$

The probability that the node pair transitions from *P.Sense* to *I* is

$$P_{PSenseI} = (1 - p_D{}') + P'_{RTSf} + P'_{RTSs}P'_{CTSf}. \tag{4.84}$$

The average duration of *P.Sense* is

$$D_{Psense} = \frac{\big(1 - p_D{}'\big)D_{Sense}{}^1 + P'_{RTSf}D_{Sense}{}^2 + P'_{RTSs}D_{Sense}{}^3}{\big(1 - p_D{}'\big) + P'_{RTSf} + P'_{RTSs}} \tag{4.85}$$

If P and Q have a successful handshake on the chosen data channel, the probability that a node transitions from *P.Sense* to S_D is

$$P_{PSenseS_D} = {P'}_{RTSs}{P'}_{CTSs}{P_{PU1}}^{L_{DATA}+L_{ACK}}{P_{X2DC}}^{L_{DATA}+L_{ACK}}{P_{X4DC}}^{L_{DATA}}. \tag{4.86}$$

The probability that a node transitions from *P.Sense* to F_D due a data packet failure is

$$P_{PSenseF_D}{}^{1} = {P'}_{RTSs}{P'}_{CTSs}\left[1 - {P_{PU1}}^{L_{DATA}} + {P_{PU1}}^{L_{DATA}}\left[1 - {P_{X2DC}}^{L_{DATA}}\right]\right]. \tag{4.87}$$

Similarly, the probability that a node transitions from *P.Sense* to F_D due to ACK failure is

$$\begin{aligned} P_{PSenseF_D}{}^{2} = {P'}_{RTSs}{P'}_{CTSs}{P_{PU1}}^{L_{DATA}}{P_{X2DC}}^{L_{DATA}}\Big[1 - {P_{X4DC}}^{L_{DATA}} \\ + {P_{X4DC}}^{L_{DATA}}\left[1 - {P_{PU1}}^{L_{ACK}} + {P_{PU1}}^{L\_ACK}\left[1 - {P_{X2DC}}^{L_{ACK}}\right]\right]\Big]. \end{aligned} \tag{4.88}$$

The overall probability of transitioning from *P.Sense* to F_D is the sum

$$P_{PSenseF_D} = P_{PSenseF_D}{}^{1} + P_{PSenseF_D}{}^{2} \tag{4.89}$$

For each transition into F_D from *Sense* or *P.Sense*, the average duration of F_D is

$$D_{F_D} = \frac{D_{F_D}{}^{2}\left(P_{SenseF_D}{}^{1} + P_{PSenseF_D}{}^{1}\right) + D_{F_D}{}^{2}\left(P_{SenseF_D}{}^{2} + P_{PSenseF_D}{}^{2}\right).}{P_{SenseF_D} + P_{PSenseF_D}} \tag{4.90}$$

4.4.2.5 Steady State Analysis

In this section we derive state probabilities and limiting probabilities for our Markov model. The method followed in this section borrows mostly from [54] and [55]. The state diagram seen in Figure 4.6 leads to steady state probabilities given by (4.91)-(4.97).

$$\pi_I + \pi_{F_C} + \pi_{S_C} + \pi_{Sense} + \pi_{PSense} + \pi_{F_D} + \pi_{S_D} = 1 \tag{4.91}$$

$$\pi_{F_C} = \pi_I * P_{IF_C} \tag{4.92}$$

$$\pi_{S_C} = \pi_I * P_{IS_C} \tag{4.93}$$

$$\pi_{Sense} = \pi_{S_C} \tag{4.94}$$

$$\pi_{PSense} = \pi_{S_D} * P_{S_DPSense} \tag{4.95}$$

$$\pi_{F_D} = \pi_{Sense} * P_{SenseF_D} + \pi_{PSense} * P_{PSenseF_D} \tag{4.96}$$

$$\pi_{S_D} = \pi_{Sense} * P_{SenseS_D} + \pi_{PSense} * P_{PSenseS_D} \tag{4.97}$$

We solve the system of equations for π_{SD}, π_{FD} and π_I using substitution, resulting in (4.98)-(4.100). Note that the transition probabilities derived in the previous sections are inherently functions of the receiver distance although not explicitly indicated by our notation.

$$\pi_{S_D} = \frac{\pi_{Sense} * P_{SenseS_D}}{1 - P_{PSenseS_D} P_{S_D fPSense}} \tag{4.98}$$

$$\pi_{F_D} = \pi_{Sense}\left(P_{SenseF_D} + \frac{P_{SenseS_D} P_{S_D PSense} P_{PSenseF_D}}{1 - P_{PSenseS_D} P_{S_D PSense}}\right) \tag{4.99}$$

$$\pi_I = \frac{1}{1 + P_{IF_C} + P_{IS_C}\left(2 + P_{SenseF_D} + P_{SenseS_D}\left(\frac{1 + P_{S_D PSense}\left(1 + P_{PSenseF_D}\right)}{1 - P_{PSenseS_D} P_{S_D PSense}}\right)\right)} \tag{4.100}$$

After calculating the transition probabilities as a function of the receiver distance r_0, we use (4.92)-(4.100) to find the steady state probabilities $\pi_I(r_0)$, $\pi_{F_C}(r_0)$, $\pi_{S_C}(r_0)$, $\pi_{S_D}(r_0)$, $\pi_{F_D}(r_0)$, and $\pi_{Sense}(r_0)$. For semi-Markov processes, the steady state probability of a state j (π_j) represents the long-run proportion of transitions into state j [58]. Let μ_j represents is mean time spent in j per transition, then the limiting probability (P_j), which represents the long-run proportion of time in state j is [58]

$$P_j = \frac{\pi_j \mu_j}{\sum_{i=0}^{\infty} \pi_i \mu_i}. \tag{4.101}$$

Using the steady state probabilities we can calculate the limiting probabilities $P_I(r_0), P_{F_C}(r_0), P_{S_C}(r_0), P_{S_D}(r_0)$, $P_{F_D}(r_0), P_{Sense}(r_0)$, and $P_{PSense}(r_0)$ which give the long-run average proportion of the time that a node spends in each state, as a function of the distance to the intended receiver r_0. The calculation results in

$$P_{F_D}(r_0) = \frac{\pi_{F_D}(r_0) D_{F_D}(r_0)}{D(r_0)}, \tag{4.102}$$

$$P_{S_D}(r_0) = \frac{\pi_{S_D}(r_0) D_{S_D}}{D(r_0)}, \tag{4.103}$$

$$P_{Sense}(r_0) = \frac{\pi_{Sense}(r_0) D_{sense}(r_0)}{D(r_0)}, \tag{4.104}$$

$$P_{PSense}(r_0) = \frac{\pi_{PSense}(r_0) D_{Psense}(r_0)}{D(r_0)}, \tag{4.105}$$

$$P_I(r_0) = \frac{\pi_I(r_0) D_I}{D(r_0)}, \tag{4.106}$$

$$P_{S_C}(r_0) = \frac{\pi_{S_C}(r_0) D_{S_C}}{D(r_0)}, \tag{4.107}$$

$$P_{F_C}(r_0) = \frac{\pi_{F_C}(r_0) D_{F_C}(r_0)}{D(r_0)}, \tag{4.108}$$

where

$$\begin{aligned} D(r_0) = {} & \pi_I(r_0) D_I + \pi_{Sense}(r_0) D_{Sense} + \pi_{PSense}(r_0) D_{PSense} + \pi_{S_C}(r_0) D_{S_C} \\ & + \pi_{S_D}(r_0) D_{S_D} + \pi_{F_C}(r_0) D_{F_C}(r_0) + \pi_{FD}(r_0) D_{F_D}(r_0). \end{aligned} \tag{4.109}$$

After finding the limiting probabilities, we can calculate the average probability of being in each state by integrating over the PDF of r_0 given in (4.4). The long-term proportion of time a node spends successfully transmitting a data packet is given by

$$P_{S_D} = \int_0^{\mathrm{r_s}} P_{S_D}(r_0)\, \mathrm{f}(r_0) dr_0, \tag{4.110}$$

where $P_{S_D}(r_0)$ is given by (4.103) and $f(r_0)$ is given by (4.4). We define the throughput at a single node as the number of packets transmitted per packet duration (L_{DATA}), which includes original data and any relayed data but no overhead or control information. The average throughput achieved by a single node is given by

$$THR = P_{S_D} \frac{L_{DATA}}{D_{S_D}}. \tag{4.111}$$

4.4.3 Network Performance

Ad hoc networks typically employ a decode-and-forward transport model, where each node acts as a source and a relay for other nodes in the network. Network transport capacity is generally defined as the product of the bit-rate and the distance progressed towards the final destination [9]. We assume each node relays the packet to a nearest neighbor which minimizes the distance to the final destination. Since only the projection of r_0 in the direction of the final destination is important in the network capacity, an upper bound for network capacity for a network of n nodes can be obtained as

$$C(n) \leq n \frac{L_{DATA}}{D_{SD}} \int_0^{r_s} r_0 P_{S_D}(r_0)\, f(r_0) dr_0, \tag{4.112}$$

where $P_{S_D}(r_0)$ is given by (4.103) and $f(r_0)$ is given by (4.4). The transport capacity is distributed among n nodes desiring to communicate at a rate $\lambda(n)$ over a mean end-to-end path length of $\bar{L}$ meters. In Chapter 2, we saw that the end-to-end throughput per node is upper bounded by

$$\lambda(n) \leq \frac{C(n)}{n\bar{L}}. \tag{4.113}$$

4.5 Numerical Results

Hoang and Iltis [54], [55] define an iterative root finding algorithm to find the limiting probabilities for a given r_0, which we adapt to our new derivation. The pseudo code for the root finding algorithm is presented in Table 4.1.

Set network parameters $\{n, m, P_0, P_1, r_s, R, R_p, M, q, q_T, \beta\}$

Set detection parameters$\{N_o, s, W, \theta\}$

μ = 0.1;

ε = 10^{-9};

e = 1;

iters = 0;

z = [0.6, 0.1, 0.05, 0.1, 0.05, 0.05, 0.05]

while (e ≥ ε or iters ≤ max_iters)

{

Calculate state transition probabilities using (4.36), (4.38), (4.41), (4.55), (4.61), (4.64), (4.78), (4.79), (4.84), (4.86), and (4.89)

Calculate average state durations using (4.42), (4.56), (4.85), and (4.90)

Calculate steady state probabilities: $\pi_I(r_0)$, $\pi_{F_C}(r_0)$, $\pi_{S_C}(r_0)$, $\pi_{S_D}(r_0)$, $\pi_{F_D}(r_0)$, and $\pi_{Sense}(r_0)$ using (4.92)-(4.100)

f (z) = $[P_I(r_0), P_{F_C}(r_0), P_{S_C}(r_0), P_{S_D}(r_0), P_{F_D}(r_0), P_{Sense}(r_0), P_{PSense}(r_0)]$ using (4.102)-(4.108)

e ←SUM((f (z) – z).^2)

z = z + μ(f (z) – z);

iters=iters+1;

}

Table 4.1 Root finding algorithm.

The numerical results that follow are obtained by using the parameters given in Table 4.2. The detection parameters used are listed in Table 4.3. We use an energy detector with approximately 0.07 probability of false alarm and 0.72 probability of detection for a single transmitter at the edge.

Parameter	**Value**
Maximum Receiver Distance (r_s)	0.1 meters
SU Tx Power (P_0)	1.0000e-007 Watts
SNR at r_s	10 dB
SU Sensing Radius (R)	0.2 meters
SU Density (n)	{16, 40, 80, 120} (users/m^2)
Expected Number SU in R_s (N)	{2, 5, 10, 15} (users)
SU Transmission Probability on CC (p_T)	{ 0.0010, 0.0050, 0.0100, 0.0300, 0.0500, 0.1000}
PU Tx Power (P_1)	5.0000e-007 Watts
PU Density (m)	3 (users/m^2)
Percent of Active PU (q)	5 %
Expected Number PU in R_p	1.885 (users)
Expected Number Active PU in R_p	0.094248 (users)
Inactive PU Transmission Probability (q_T)	0.001
L_{DATA}	$100L_{SLOT}$
$L_{ATM}, L_{RTS}, L_{CTS}, L_{ACK}$	$5L_{SLOT}$
L_{SENSE}	L_{SLOT}
Number of Data Channels (M)	{1, 4, 8, 12, 16}

Table 4.2 Numerical parameters.

Detection Parameter	**Value**
Noise Power (N_O)	1e-12 Watts
Detection Threshold (θ)	9.5 dB
Detection Time (s)	4e-6 seconds
Bandwidth (W) (Link Rate)	1e6 Hz (bps)
Path Loss (β)	2
Channel Gain ($\sqrt{A}$)	1
P_d of SU at r_s	0.99289
P_d of SU at R	0.72893
P_d of PU at R_p	0.72893
P_f	0.079083

Table 4.3 Energy detection parameters.

4.5.1 Performance Analysis

Wireless networks are primarily interference limited. A DSA network is limited by self-interference from other secondary users as well as interference from primary users. Additionally, since SUs access the spectrum opportunistically, the amount and availability of spectrum also limits the performance of DSA systems. Here we obtain results for and analyze the performance of an ad hoc DSA system using the DSA access scheme described in Section 4.4.1 against network density, network load, number of available data channels, and primary user characteristics.

Figure 4.7 illustrates the relationship between the limiting probability of a single node being in anyone of seven possible states over a long period of time and the transmission probability per slot on the control channel (p_T) for various SU densities. The limiting probability of a node being in the *idle* state (P_I) decreases with p_T, while the probability of being in the failure on the control channel *(*P_{F_C}*)* state increases with p_T. The limiting probability of a node successfully handshaking on the control channel initially increases with p_T and begins to fall as congestion builds up. Since a node must rendezvous on the control channel before transmitting on the data channels, the limiting probability of a node sensing a data channel (P_{Sense}), and successfully (P_{S_D}) or unsuccessfully (P_{F_D}) transmitting on one of the *M* data channels follow the same trend as P_{S_C}. The behavior is expected and similar to the performance of CSMA networks where performance is interference limited. The results re-enforce the importance of careful control channel design, as DSA networks are vulnerable to congestion on the control channel. Figure 4.7 also shows that increased SU shifts the p_T value which optimizes P_{SD} closer to the origin. p_T values of 0.01, 0.02, 0.03, and 0.05 optimize the proportion of time spent transmitting data for *N* values of 2, 5, 10 and 15, respectively.

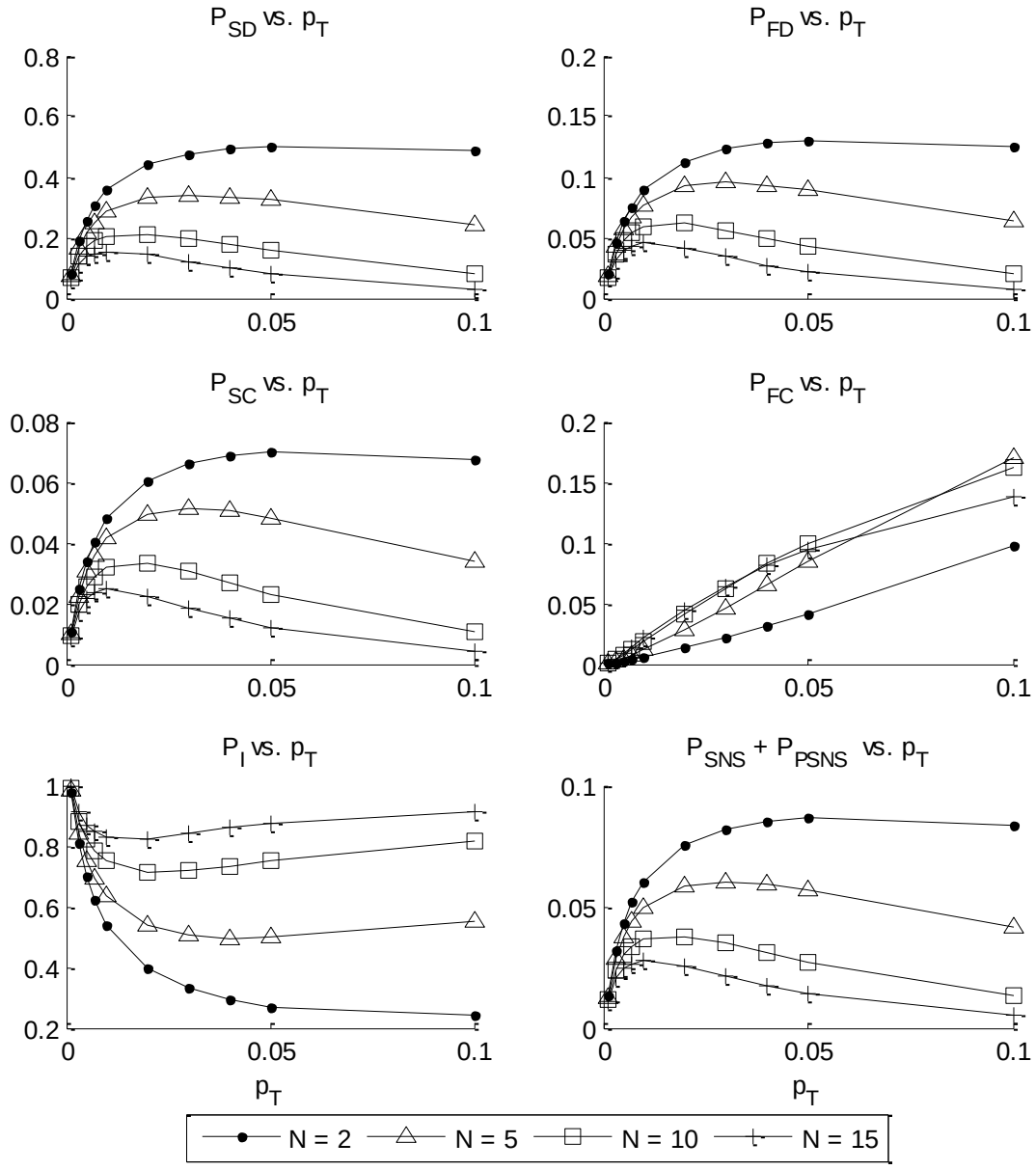

Figure 4.7 Limiting probabilities vs. p_T.

Figure 4.8 shows the effect of node density on the distribution of average time spent in different states, for $M = 20$ and $p_T = 0.03$ which is approximately the average optimal p_T. As N increases the overall probability of a node being in the idle state increases due to control channel saturation. The amount of time spent successfully transmitting on the control channel decreases while the amount of time spent unsuccessfully transmitting increases. P_{F_C} does appear to stabilize for $N > 10$, indicating that sensing is working. Interestingly $P_{S_D}, P_{F_D}, P_{Sense}$ and P_{PSense} behave in a similar trend as P_{S_C} since an initial transition to a data channel requires a success on the control channel.

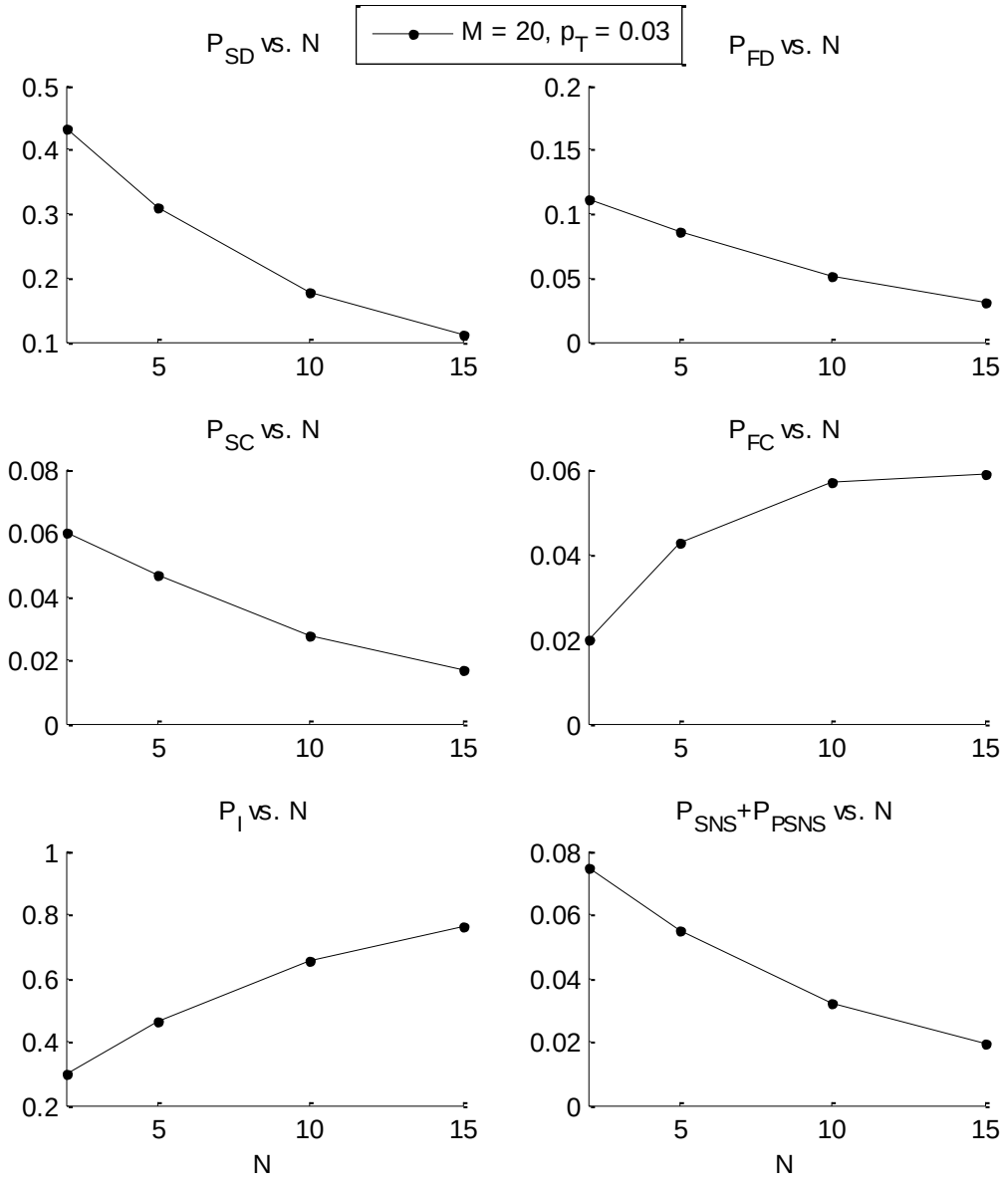

Figure 4.8 Limiting probabilities vs. *N*.

Figure 4.9 illustrates the behavior of a node with respect to the number of data channels M, for $N = 5$ and $p_T = 0.03$. As expected, P_{SD} increases with M since more simultaneous transmissions are possible in the same area. Nevertheless, it plateaus after $M > 16$ due to primary user interference and control channel saturation. On the other hand, P_{FD} decreases with M, since secondary users are less likely to choose channels which are occupied by other secondary users. Interestingly, P_I, P_{S_C} and P_{F_C} also decrease with M, showing that nodes waste less time being idle or transmitting on the control channel when there are more data channels available.

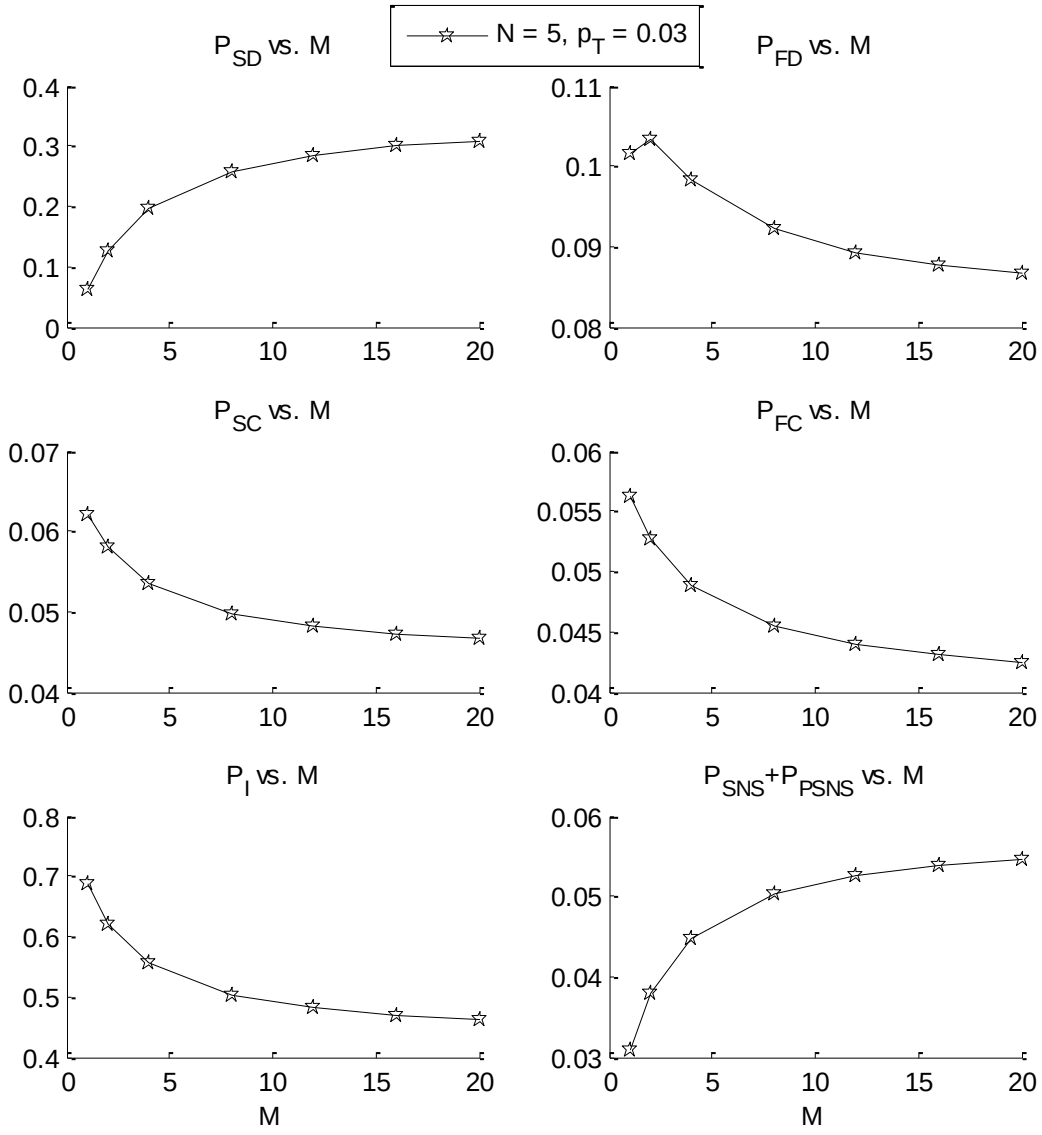

Figure 4.9 Limiting probabilities vs. *M*, *N* = 5.

4.5.2 Sensing and Interference

Figure 4.10 shows the relationship between the number of data channels and the likelihood of a node sensing a data channel as busy, failing the initial handshake, successfully and unsuccessfully transmitting a data packet from the Sense and P.Sense states. The plots show that a node is less likely to sense the channel as busy, or have a RTS/CTS or data failure as *M* increases. The second plot shows that the node is more likely to have a successful transmission after completing a successful transmission.

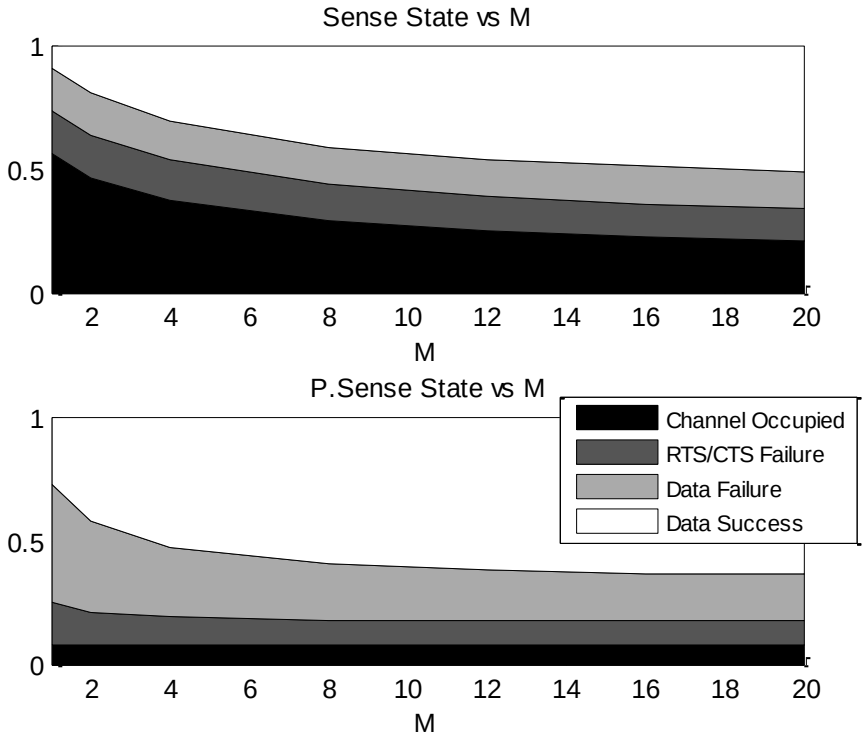

Figure 4.10 Sensing on data channel, N = 5, p_T = 0.03.

Figure 4.11 shows the relationship between M and the causes of failure on the data channels. Interference on a data channel can come from secondary and primary users in the form of to missed nodes, hidden nodes, and incidental nodes. It shows that the percentage of failures due to primary users are independent of M, while the percentage of failures due to secondary users decrease with M as the users are spread over more channels. The plot also explains that the increase in performance with M tapers off after a certain point because the SU interference is mitigated.

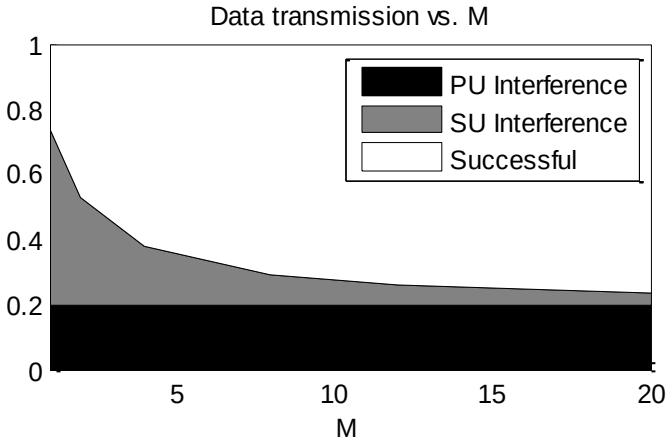

Figure 4.11 Interference on data channel vs. M, N = 5, p_T = 0.03.

4.5.3 Single Node Performance

The expected throughput available at a given node, for original as well as relayed traffic, is found by (4.111). Single node performance is evaluated against increasing SU density for fixed M, and when M increases with N. The latter case is included to evaluate the performance of a scenario where the amount of accessible spectrum is proportional to the network density. Figure 4.12 plots the expected throughput as a function of N for the probability of transmission which optimizes throughput. The throughput decreases with N for both cases, though the loss is lower for the case where M increases with N. The throughput does not go to zero as N increases but plateaus around 0.1 packets per packet duration for both cases.

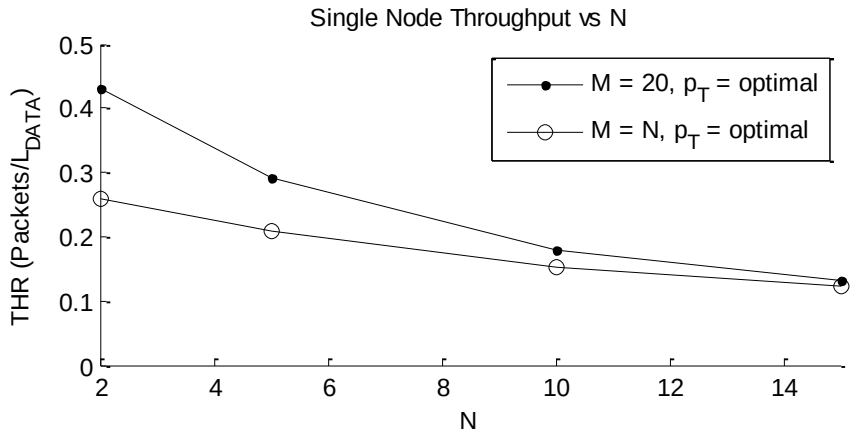

Figure 4.12 Single node throughput vs. *N*.

Figure 4.13 shows throughput as a function of increasing data channels for two scenarios: (i) the number of SUs is fixed, $N = 5$; (ii) the number of SUs increases with $M, N = M$. When N is fixed, the throughput increases with M but plateaus around 0.3 packets per packet duration. This indicates that other factors limit the throughput, such as primary user interference, sensing errors, control channel failures, and control channel wait times. The plot also shows that when N increases proportionally with the number of accessible data channels, the throughput decreases but appears to stabilize at around 0.1 packets per packet duration.

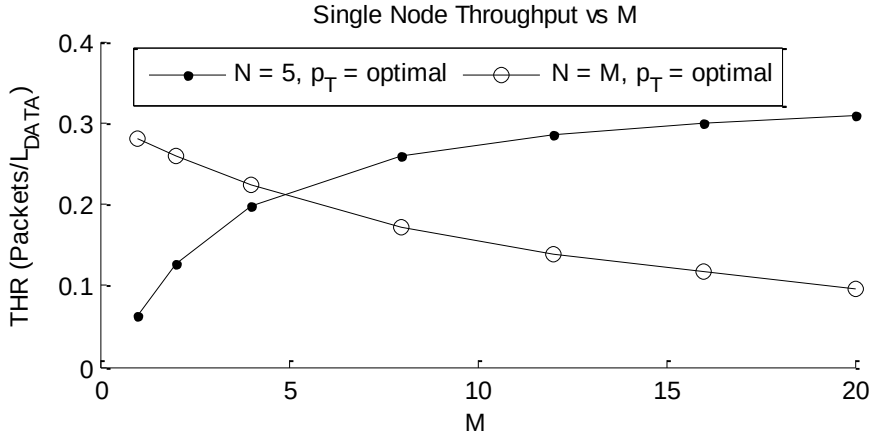

Figure 4.13 Single node throughput vs. *M*.

Figure 4.14 shows the distribution and sum total of the proportion of time that a node is in one of the *active* states, in which a node is transmitting or sensing but not idle. The first plot shows that the proportion of time that a node spends in each state is relatively constant, with the exception of F_C which increases and S_D which decreases slightly. The time that a node spends in communicating on the control channel, sensing a data channel, and unsuccessfully communicating on the control or data channels can be considered a form of overhead. The relative proportion of overhead may be reduced further as it is largely determined by the access scheme and interference tolerance capabilities of the network. Nevertheless, the first plot shows that *the relative overhead required by the DSA access scheme does not scale up with N* which is a very useful result. The second plot shows that the total time that a node is active decreases with N and explains that the drop in single node throughput with N is largely due to a larger proportion of time in the idle state, and *not because of additional interference or overhead.*

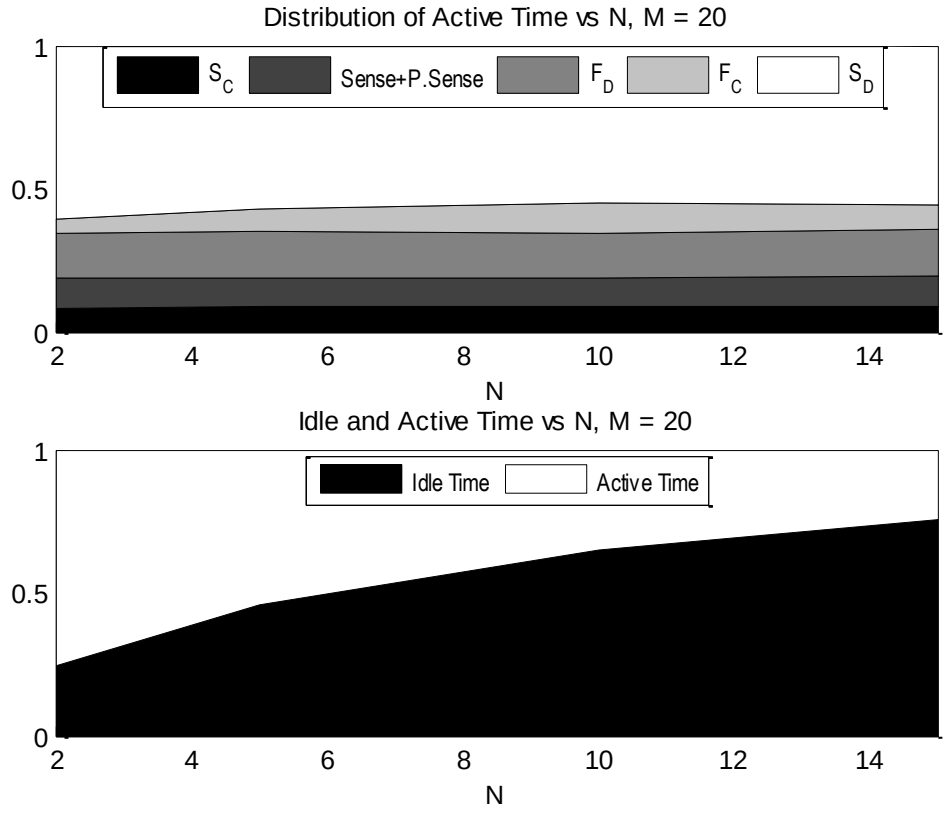

Figure 4.14 Overhead analysis and active time vs. *N*, *M* = 20.

Figure 4.15 shows the distribution of the proportion of time that a node is spends in each state as a function of the number of data channels M for a fixed node density N. The plot shows that the amount of time that a node spends in overhead states is constant for all M, while increasing M allows a node to spend more time transmitting data. The results show our DSA scheme benefits from being able to access more channels.

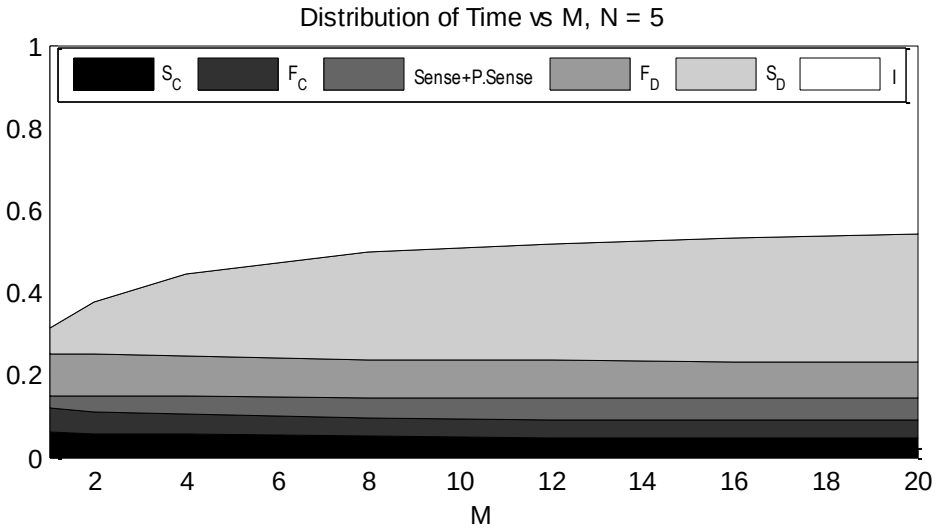

Figure 4.15 Overhead analysis and active time vs. *M*, *N* = 5.

4.5.4 Network Performance and Scalability

The data presented in Figure 4.7 is processed using (4.112) and (4.113) to yield upper bounds for network transport capacity and end-to-end throughput, respectively. Figure 4.16 shows the expected network performance with respect to SU density for optimal p_T and the maximum number of data channels, $M = 20$. The network capacity is observed to increase sub-linearly with the node density, while the end-to-end throughput decreases as SU density increases since $\lambda \propto \frac{C}{n}$ as seen in (4.113). We observe that the capacity scales similarly to the Gupta and Kumar scaling law [9], however there is a fundamental difference in the system model. As discussed in Chapter 2, Gupta and Kumar were able to achieve a sub-linear increase in capacity with the number of nodes, where the transmission range of each node scales down as the density increases. Gupta and Kumar scale the transmission range in proportion with a critical radius for connectivity which ensures network wide connectivity. In our analysis, we relax the connectivity constraint and keep the transmission range constant while increasing density, yet we still see an increase in capacity. If the transmission range is scaled down with density, we would likely see a higher increase in capacity than our current results.

Considering that certain DSA networks may collectively sense and allocate access to spectrum, suggests that the number of data channels available for access will likely be correlated with the SU density. Therefore, Figure 4.16 also illustrates the scalability of network capacity and end-to-end throughput, where the number of channels, M, increases linearly with N. We observe, in this case, that the transport capacity shows a nearly linear increase with N, and end-to-end throughput is nearly constant. The result shows that *DSA can be useful in maintaining a certain throughput in dense networks by spreading the users over a wider range of frequencies.*

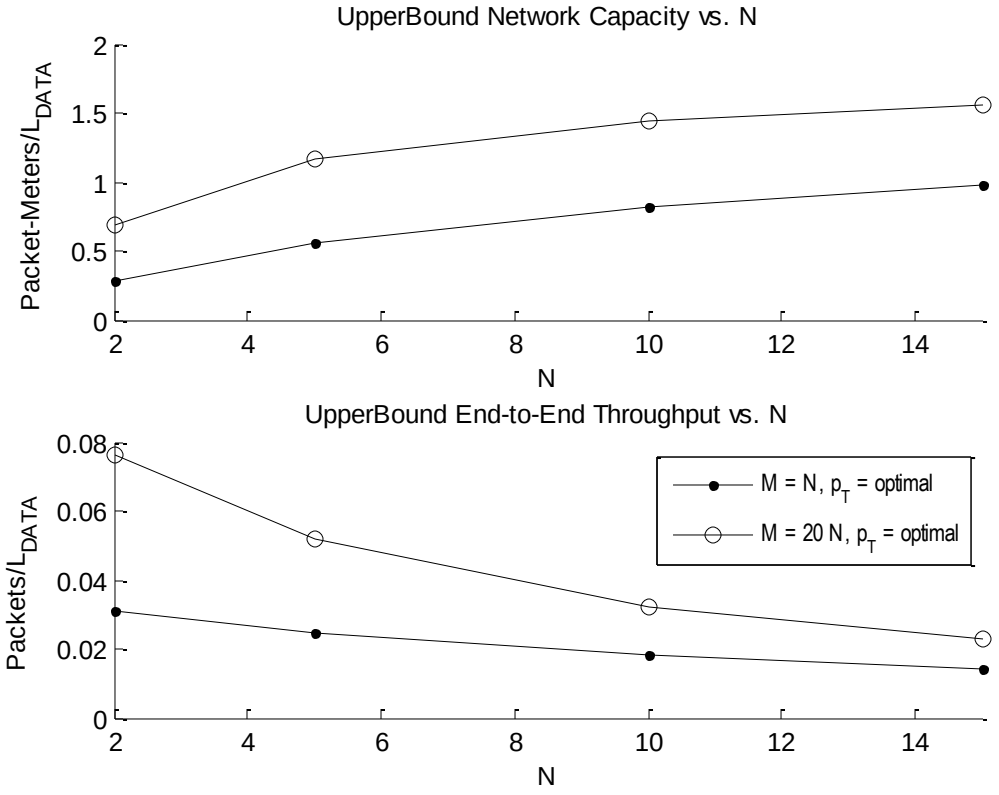

Figure 4.16 Scalability of capacity and end-to-end throughput vs. *N*.

Figure 4.17 illustrates the scalability of transport capacity and end-to-end throughput with respect to the amount of spectrum available. The capacity and throughput increase with the number of channels, however, the performance tapers off as *M* increases showing that indefinitely increasing access to the spectrum does not indefinitely increase the capacity. As discussed previously, aside from spectrum that other factors limit the performance, such as primary user interference, sensing errors, control channel failures, and control channel wait times.

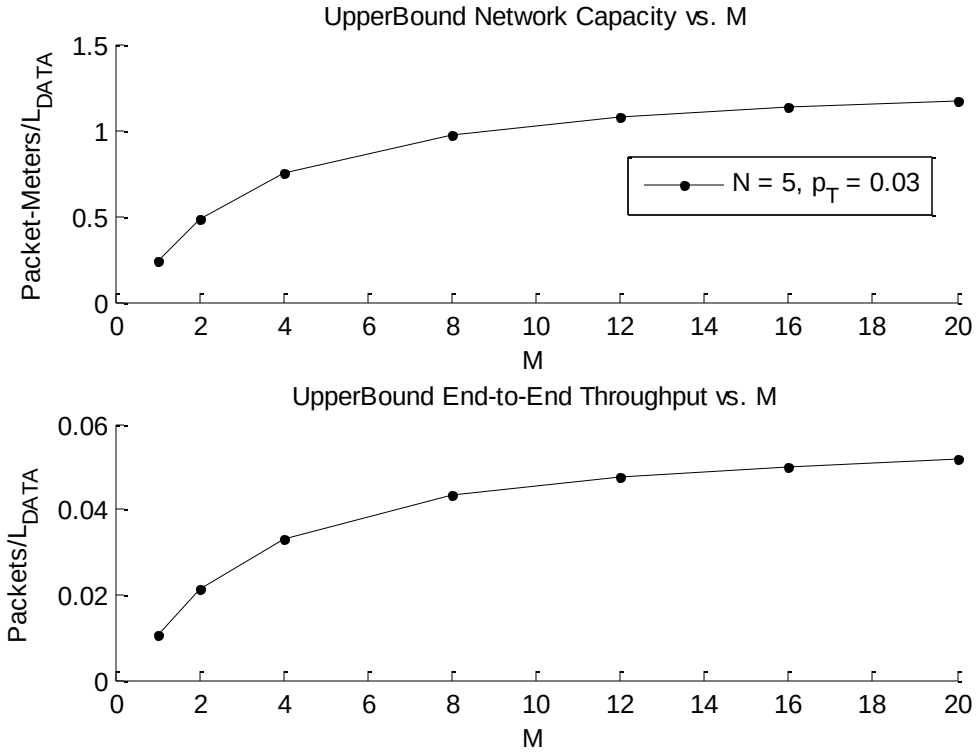

Figure 4.17 Scalability of transport capacity and end-to-end throughput vs. *M*, *N* = 5.

We also evaluate the impact on DSA performance under a varying PU network. The set of plots in Figure 4.18 show the impact of PU density (m) on the distribution of the average time a DSA node spends in each state. Increasing PU density leads to a higher number of occupied data channels and more data failures due to primary user interference. Therefore, P_{S_D} shows a significant decrease, while P_{F_D} shows a significant increase as m increases. The total time spent in the sensing states ($P_{Sense} + P_{PSense}$) also decreases as data channels are more likely to be sensed busy. As an interesting side effect, the time spent on the control channel (P_{F_C} and P_{S_C}), and in the idle state (P_I) also shows a slight increase. Since D_{F_D} is smaller than D_{S_C} on average, and the increase is P_{F_D} is smaller than the decrease in P_{S_D} there is necessarily an increase in the idle time. This increase in the idle time likely causes P_{F_C} and P_{S_C} to increase as well since attempts on the control channel are initiated from the idle state.

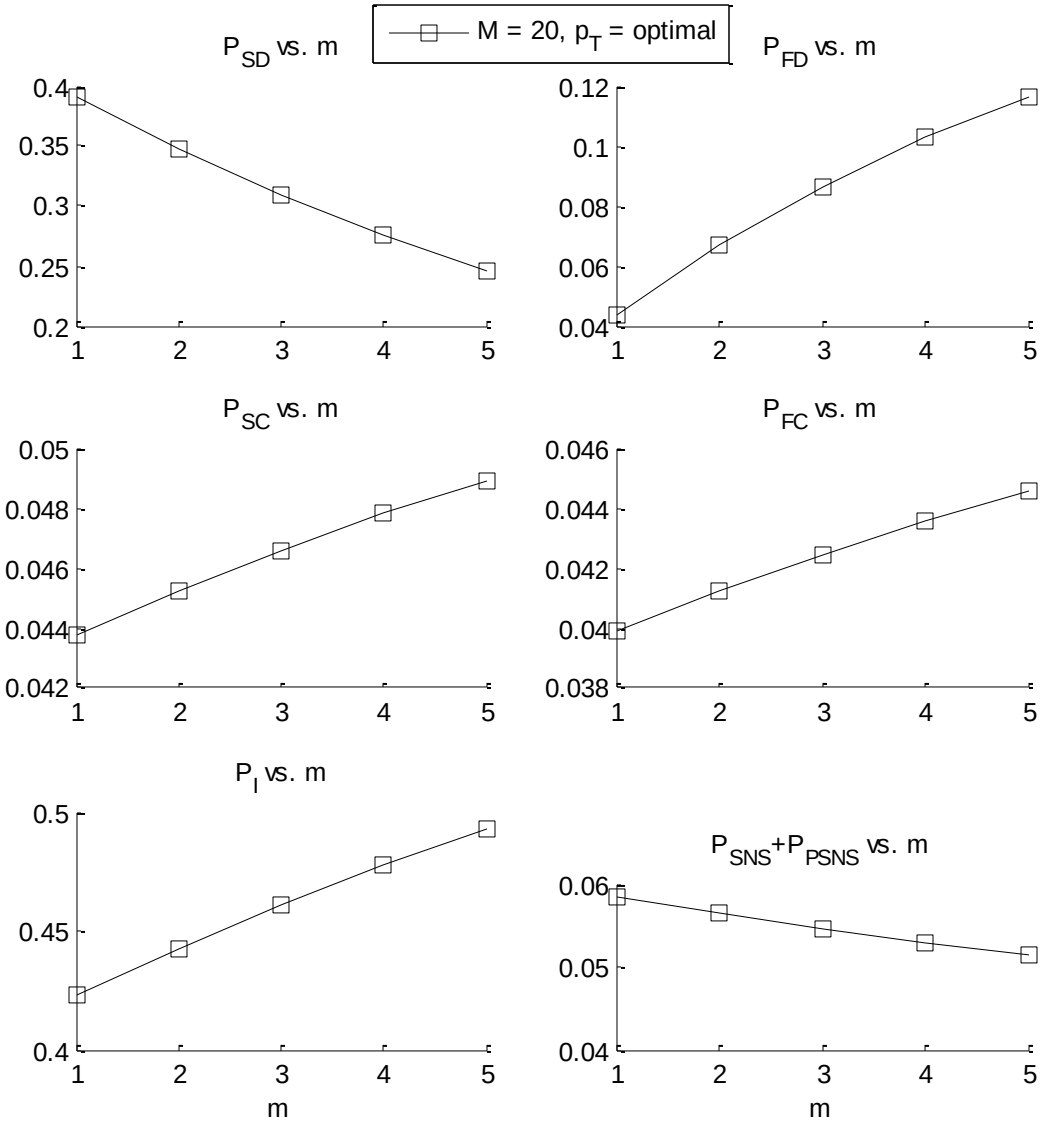

Figure 4.18 Limiting probabilities vs. PU density, N = 5.

Figure 4.19 shows the distribution and sum total of the proportion of time that a node is in one of the active states. The first plot shows that the proportion of time that a node spends in each state is relatively constant, except the time in F_D increases with m. The first plot shows that *the relative overhead required by the DSA access scheme increases with m.* The second plot shows that the total time that a node is active decreases only slightly with m, and explains that the drop in single node throughput with m is largely due to a larger proportion of failures on the data channel due to additional PU interference. In contrast with the change in SU density, change in PU density linearly reduces active time and also increases overhead due to data failures.

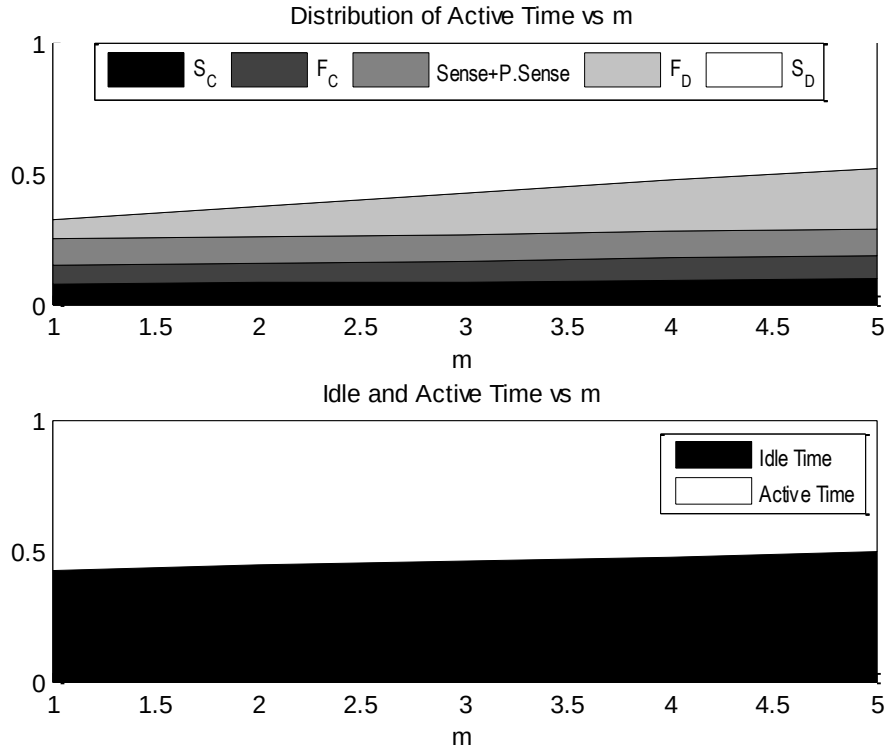

Figure 4.19 Overhead analysis and active time vs. PU density, $N = 5$, $M = 20$.

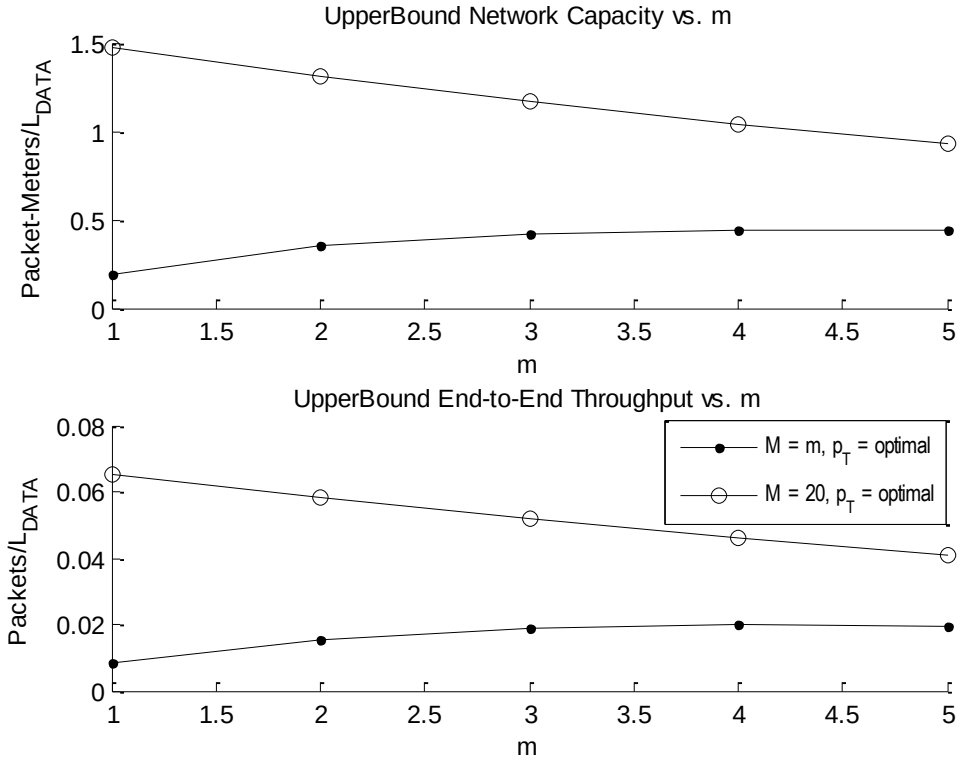

Figure 4.20 Network Performance vs. PU Density

When the number of data channels M is fixed, transport capacity and end-to-end throughput decrease linearly as PU density increases. It is reasonable to assume that DSA networks operating in regions with heavy primary user activity will be designed to access unused spectrum among a larger set of frequency bands. As can be seen in the Figure 4.20, when M increases with the PU density, the network capacity and end-to-end throughput also increase or stay constant. This result shows that flexibility in DSA design can overcome unfavorable network conditions.

4.5.5 Impact of Detection

This section examines the performance impact of detection accuracy by comparing results with an energy detector (ED) to results with perfect detection (PD). A DSA network which can perfectly sense the presence of secondary or primary users within its sensing (and interference) radius is expected to outperform those with energy detection since there are no false alarms or missed detections. Figure 4.21 compares optimal p_T on the control channel for nodes with perfect detection and energy detection. Networks with perfect detection can operate at a higher p_T since there are fewer collisions.

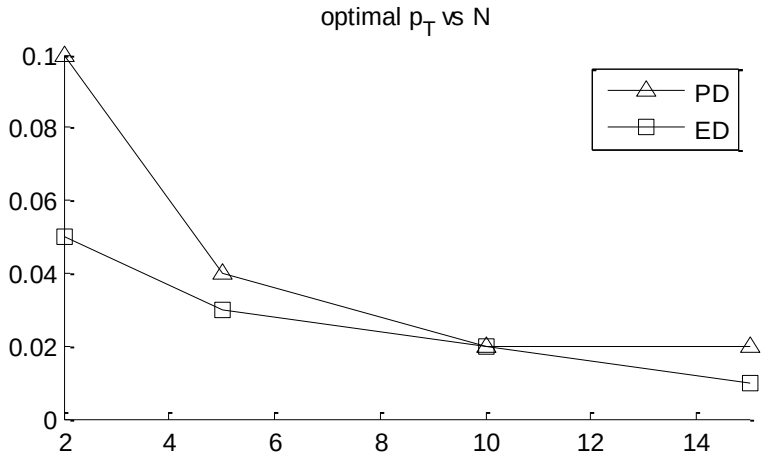

Figure 4.21 Optimal p_T for PD and ED.

Figure 4.22 shows the change in limiting probabilities with respect to N for PD and ED at the respective average optimal p_T. The limiting probabilities show similar trends for both cases as N increases. The largest difference is seen in P_{S_D} and P_I, as nodes with PD spend more time engaging in successful data transmissions and less time being idle. Incidentally, nodes with PD also spend more time in F_D and the sensing states because they are coupled to the S_D state. Each data packet attempt requires a sensing block and with each attempt additional attempt there is a chance of failure. The behavior on the control channel remains about the same for both cases, but nodes with PD spend slightly more time on the control channel, most likely because there are no false alarms preventing access.

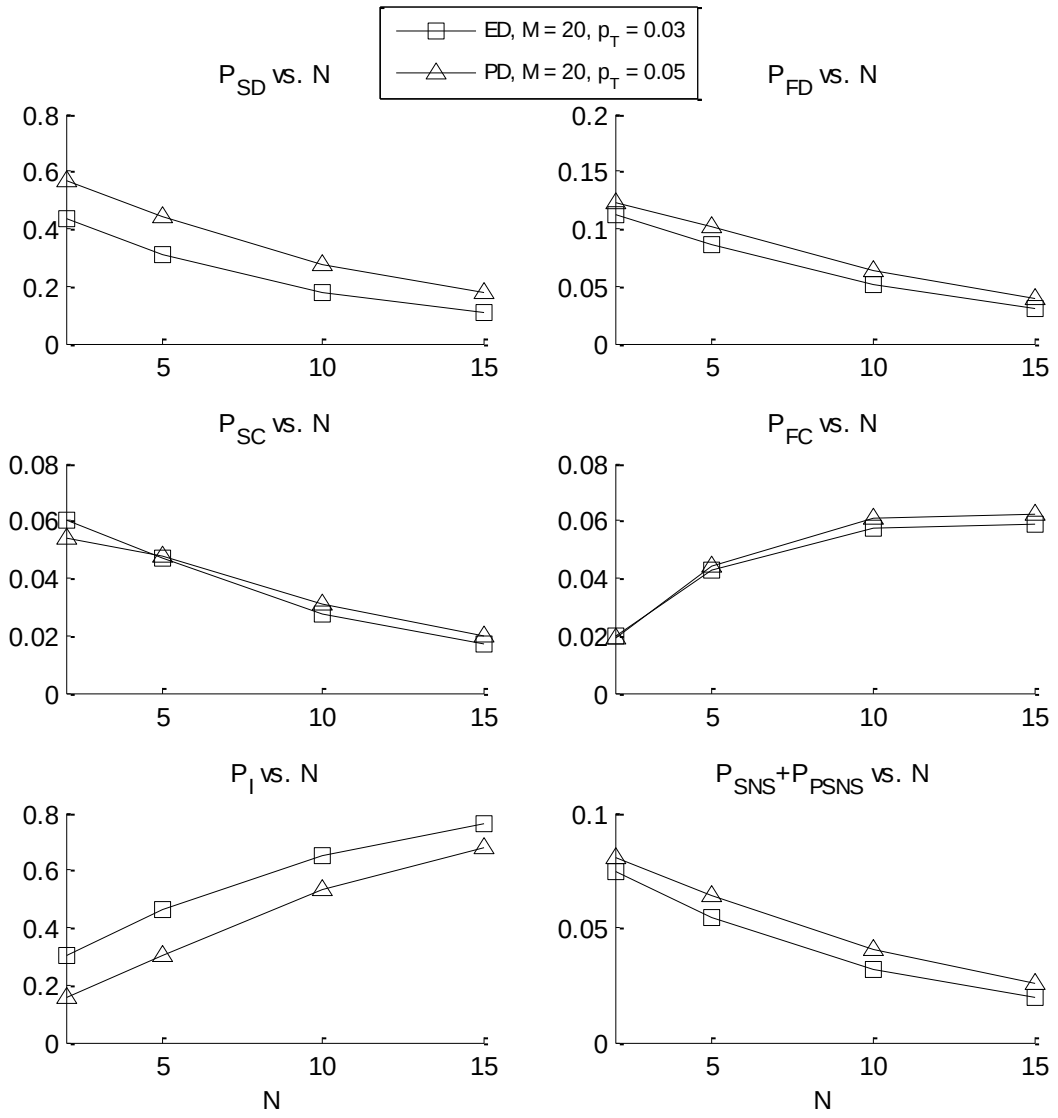

Figure 4.22 Limiting probabilities vs. *N* for PD and ED.

Figure 4.23 illustrates the average behavior of a node in the *Sense* and *P.Sense* states with respect to the number of channels. The overall trends are the same for energy detection as shown Figure 4.10, yet the probability of RTS/CTS and data failure are lower with perfect detection. Additionally, the probability of sensing the channel occupied in the P.Sense state due to a false alarm is zero for perfect detection. Since the trends are the same as M increases suggests that detector accuracy has a linear impact on performance.

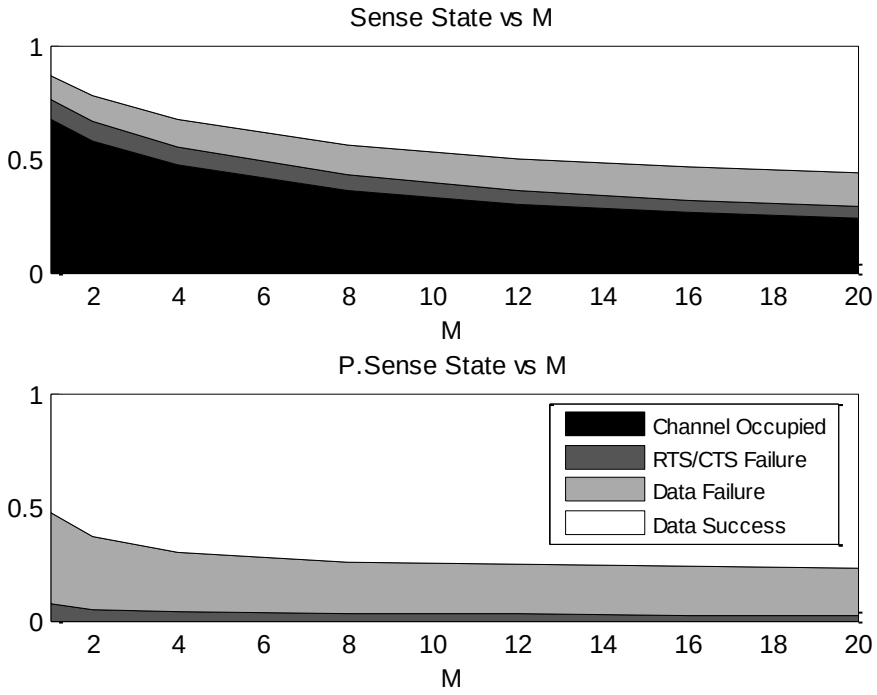

Figure 4.23 Sense state behavior vs. *M* for PD.

Figure 4.24 illustrates the expected outcome once a node begins to transmit a data packet on a data channel. As *M* increases the probability of a successful data transmission increases because the probability of seeing interference from an SU decreases. The behavior is qualitatively the same as with energy detection as seen in Figure 4.11, although less packets are lost due to SU interference, which is minimized due to perfect detection.

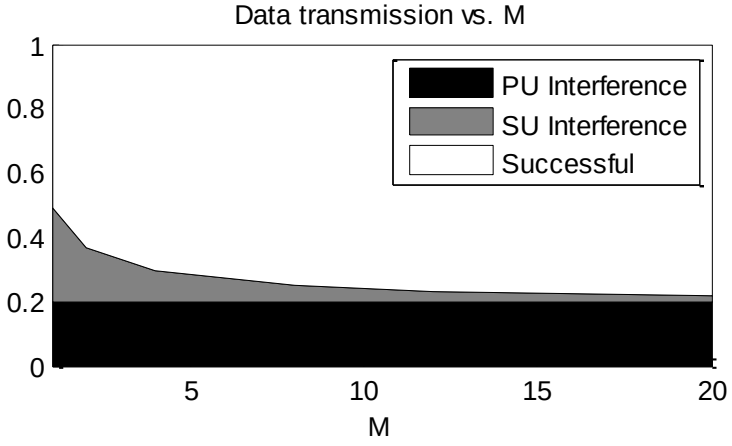

Figure 4.24 Data transmission vs. *M*, $N = 5$, $p_T = 0.04$, with PD.

Finally, the network transport capacity and end-to-end throughput achieved with perfect detection are compared with those with energy detection. Figure 4.25 shows that perfect detection achieves a higher network transport capacity and end-to-end throughput. The gain in capacity, however, appears to be linear. This is further shown by the second plot, as the end-to-

end throughput curve has the same slope but is only shifted up for perfect detection. In other words, detector accuracy may be simplified as a constant multiplier when computing network transport capacity.

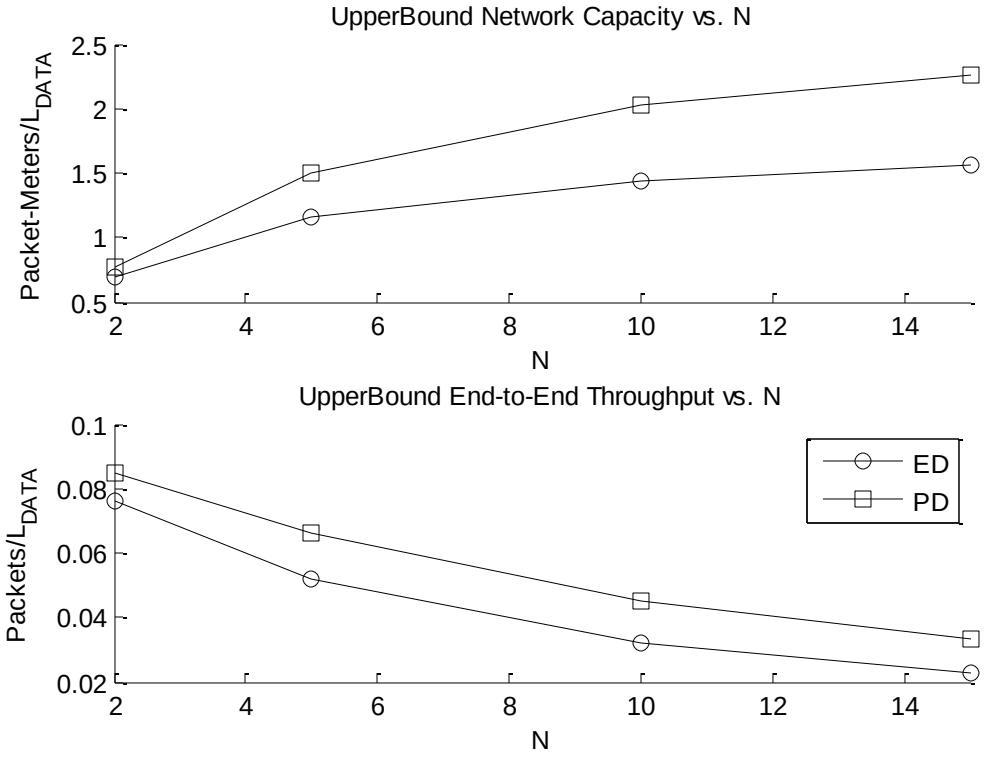

Figure 4.25 Transport capacity and end-to-end throughput with PD and ED, *M = 20*.

4.6 Conclusion

In this chapter, we developed an analytical procedure to evaluate the performance of ad hoc DSA networks using Markov models. The procedure was inspired by the work in [54], amongst others, for single channel CSMA networks. Our contribution to the work included expanding the model to an opportunistic DSA network operating alongside primary users. We also developed a sensing model for heterogeneous Poisson networks using energy detection in the presence of path loss and fading, which we integrated into the performance evaluation model. Our analytical framework is modular and flexible in order to be adaptable to different DSA implementations in future studies. For example, we can readily evaluate the performance impact of using a shared control channel rather than a dedicated control channel. Additionally, we can evaluate the value of cognition by comparing the performance of DSA with a uniform spectrum decision method to that with a hierarchal spectrum decision method which leverages statistical PU traffic data.

Using our model, we obtained results for the performance of a random access DSA network with one transceiver and a dedicated control channel. We obtained scalability trends by evaluating the performance of such a DSA network under varying network densities and spectrum availability. We saw interesting results that provide insight into the expected performance of

DSA systems and show that such analytical methods may be useful evaluation tools for design of DSA systems. As discussed in Chapter 2, network capacity scales according to $\sqrt{n/logn}$ in ad hoc wireless networks because self-interference in the network increases faster than the network size and traffic. In this chapter, we saw that for fixed bandwidth cases, ad hoc DSA networks also scale in a similar fashion. Additionally, we saw that DSA can improve network performance by allowing nodes to access more spectrum bands while still providing a mechanism for spectrum sharing and maintaining network wide connectivity. The scalability trends in Figure 4.16 illustrate this point, as denser networks with access to more spectrum can still maintain relatively constant end-to-end throughput and show a nearly linear increase in network capacity. This is a considerable scalability improvement from the fixed-bandwidth case. This result shows that DSA ad hoc networks can be scalable if DSA provides nodes with accessible spectrum proportional to the increase in network density.

We observed a linear relationship between DSA network performance and the primary user density, which also limit the amount of usable spectrum. We also observed that the percentage of relative network overhead at the medium access layer, which does not include routing and network layer overheard, stays constant and does not scale up with the number of users. This result is important and answers the fundamental question of the scaling behavior of the network overhead in addition to throughput. Overall, the scalability trends discussed in this chapter show that DSA, even in the form of a minimal random access protocol, shows potential for scalability.

5 CONCLUSION

The primary goal of our research was to characterize the performance and scalability of ad hoc DSA networks. Our study focused primarily on the scaling behavior of network transport capacity and end-to-end throughput per node with respect to the network density. We saw in Gupta and Kumar's work that ad hoc networks suffer from major scalability issues. The network capacity is not order optimal with the number of nodes and the end-to-end throughput diminishes with increasing user density. The scaling behavior of network transport capacity is interference limited, and the diminishing behavior of the end-to-end throughput results from interference and the multi-hop burden due to the decode-and-forward form of cooperation.

We also developed a MATLAB simulation of random networks under an ideal medium access scheme, and common random access schemes. Using our simulator, we reproduced the theoretical bounds on scalability of network capacity and throughput of ad hoc networks, and studied the performance impact at the medium access layer. Simulations showed a severe reduction in performance with realistic MAC schemes from the ideal case. For Aloha and slotted Aloha, we saw that the performance was highly dependent on the medium access probability due to congestion. CSMA mitigated the congestion problem through sensing and opportunistic use of the channel; however, its performance was limited because of hidden nodes. Our scalability study revealed that both CSMA and Aloha showed a linear degradation at peak performance. In other words, although absolute performance was lower, the ideal scaling law for network capacity, $C_T = \Theta\left(\sqrt{n/logn}\right)$ [9], was also observed with random access MAC schemes.

For ad hoc DSA networks, we examined scalability with respect to secondary and primary user density, spectrum availability, and detection accuracy. We developed an analytical procedure for the performance evaluation of ad hoc DSA networks using Markov analysis. DSA operation and performance is closely coupled to sensing, so we also developed a detailed sensing model for planar Poisson networks using energy detection and integrated it into the performance evaluation model. The analytical framework is flexible and can be adapted to evaluate different DSA

schemes. Since our study focused on scalability; we chose a straightforward DSA scheme to arrive at a lower bound on the scalability of DSA networks. We characterized the performance of a generic random access DSA network with one transceiver, uniform channel selection, and a dedicated control channel.

It is well known that traditional ad hoc networks suffer from scalability issues that stem from self-interference and accumulated traffic due to relaying. DSA networks are proposed as a solution to increase performance by accessing additional spectrum and controlling interference by spreading users in frequency. DSA, effectively, buys more spectrum but at the cost of added overhead due to spectrum sensing and spectrum management. Our results show that for a fixed number of accessible channels, DSA scalability rivals the ideal scaling bounds for ad hoc networks derived by Gupta and Kumar in [9]. We show that DSA successfully allows for opportunistic use of spectrum through sensing and spectrum management without changing the scaling behavior. Additionally, we show that DSA can improve network scalability by allowing nodes to access more spectrum bands without an accompanying increase in overhead. Our results show that ad hoc DSA networks can be scalable as long as the DSA mechanism can meet the required increase in accessible spectrum. If the amount of accessible spectrum increases proportionally with the network density, DSA nodes can maintain relatively constant end-to-end throughput and show a nearly linear increase in network capacity. This is a positive result because it eases fears that DSA overhead may grow as fast as or faster than the spectrum accessed. By no means do our results indicate that DSA networks can scale infinitely, but on a practically level they can greatly increase the number of nodes that a network may support until the entire RF spectrum is occupied. This solves the current issue of spectrum scarcity, which as we saw earlier is not due to lack of spectrum but inadequate spectrum allocation.

APPENDIX

A MORE ON SCHEDULING SCHEME

This section discusses the global multiple access scheme used to schedule transmission in Chapter 2. The scheme was intended to (i) allow all nodes a statistically fair chance to transmit, (ii) ensure there are no collisions, and (iii) ensure high spatial reuse such that at each time slot a combination of nodes is picked such that no other node can transmit without collisions. For each time slot τ, non-interfering nodes are randomly selected to transmit. The node selection algorithm used for scheduling randomly selects a node as a transmitter first and then randomly selects a relaying node as the intended receiver from a subset of its relays determined according to the interference in the network and ongoing transmissions. The random selection algorithm attempts to maximize spatial reuse and capacity by maximizing the number of simultaneous, non-interfering transmissions for a given network topology. The scheduling scheme picks transmitting nodes randomly ensuring that each node has a statistically equal chance of being picked. However it does not ensure that each relay associated with that node is appropriated a statistically equal amount of slots to receive.

The average hop length seen in simulations, shown by the "Simulated" curve, is plotted in Figure A.1 against the expected average hop length which is shown by the "Average Nearest Neighbor" curve. A direct comparison of the curve showing the average distance to the nearest neighbors with the average hop length reveals that the average hop length seen in simulations is consistently smaller. Additionally, Figure A.2 shows a histogram of the number of times a link was activated plotted against the link length. Figure A.2 clearly shows that longer links are activated less than shorter links.

Consequently, the transmission ranges seen in actual simulation runs are inherently smaller than the average distance between nearest neighbors, in order to minimize interference to the network and still take advantage of spatial reuse. In other words, the random scheduling algorithm—compared to a fixed, iterative, selection algorithm such as that proposed by Gupta and Kumar—

chooses shorter links more often in order to maximize capacity. Implementing a fixed selection algorithm, where each link is explicitly guaranteed an equal number of slots, would ensure fairness, at the expense of sub-optimal spatial reuse. Therefore, to simulate maximum network capacity we use random scheduling. As discussed in Chapter 2, our choice for using differentiated transmission ranges maximizes spatial reuse and illustrates that the Gupta and Kumar bounds hold regardless. This example also illustrates the trade-offs that exist between the routing scheme, throughput and delay. If a routing scheme chooses shorter links which minimize interference, then the total capacity increases at the expense of a higher delay as packets must make more hops as they travel from source to destination.

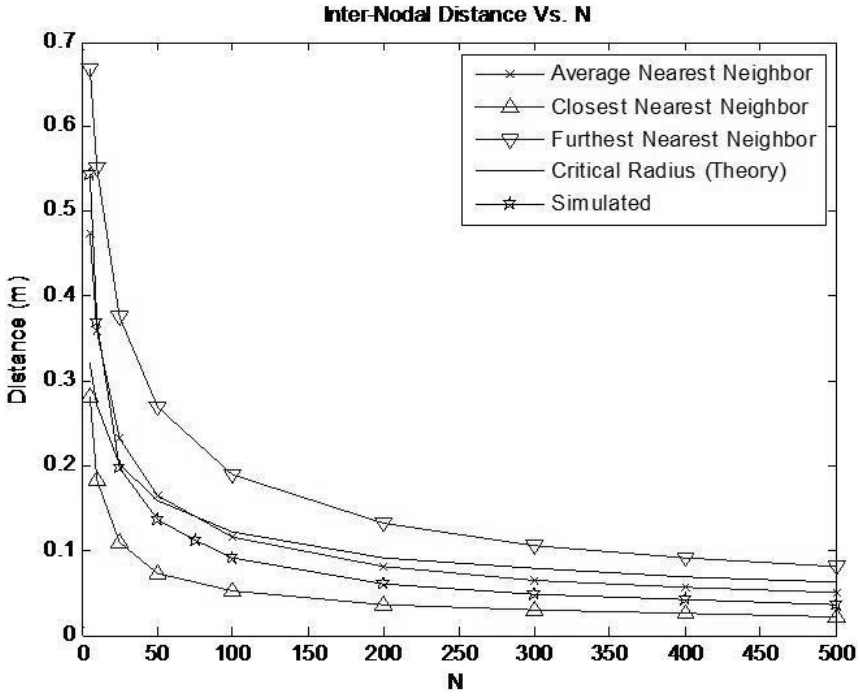

Figure A.1 Simulated transmission ranges vs. average distance to nearest neighbor.

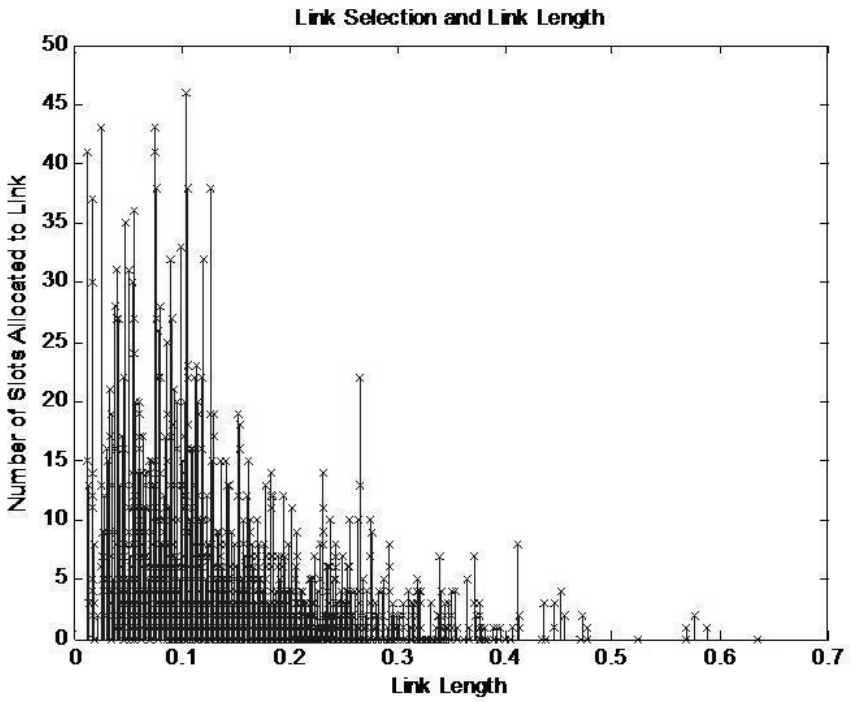

Figure A.2 Histogram of link selection versus link length.

APPENDIX

B ANALYTICAL ANALYSIS OF CSMA

In Chapter 2, we extended our system model to simulate single-channel random access MAC protocols. In order to gain a deeper understanding of the performance of ad-hoc networks, we examine an analytical framework to evaluate the performance of CSMA in ad hoc networks. The framework allows us to take a deeper look in to the throughput and transport capacity of single channel CSMA networks and analytically approximate the performance seen in our simulations. We adapt this framework to model multi-channel DSA networks in Chapter 4.

We leverage previous work by Wu and Varshney [52], [53], Iltis and Hoang [54], [55], and Takagi and Kleinrock [45] in development of analytical models for throughput analysis of CSMA and other opportunistic multi-hop MAC protocols. Takagi and Kleinrock present one of the first analyses of ALOHA and CSMA in multi-hop networks in [45]. Wu and Varshney expand the analytical model in [45] to evaluate the throughput of single channel non-persistent CSMA and BTMA (Busy tone multiple access) in [52] and extend their analysis to the multi-channel case in [53]. Hoang and Iltis in [54] develop a unified framework for performance evaluation of single-channel ad-hoc networks under various generalized access schemes including p-persistent CSMA. [55] also incorporates capture and fading effects, which we do not explicitly consider here.

B.1 System Model

The following analysis is based on the work presented in [54], adapted to approximate the CSMA system model used in Section 3.3.1. As in [54] we approximate an asynchronous network by a slotted system, where each slot has duration L_{SLOT}. Additionally, L_{SLOT} must be greater than the round trip propagation delay from transmitter to receiver and the device delay at transmitter and receiver. The main differences between our system model and the one used in [54] are:

- We adopted a non-persistent CSMA scheme while [54] analyzes a p-persistent protocol.
- We assume zero propagation and sensing delay while [54] assumes that the propagation and device delay equals the duration of one slot.
- We do not model the return ACK from the receiver while [54] models the return ACK.
- [54] assumes that the sensing radius is fixed and the attempted transmission range varies uniformly between 0 and the sensing radius. We assume that the sensing radius is dynamic based on the attempted transmission range, as described in Chapter 2. Also we assume the attempted hop length follows a Rayleigh distribution, while [54] assumes it to be uniform.

As in [54], we model the behavior of a single node in the network as an independent Semi-Markov process. At any instance a node in the network falls in to one of the following states:

- **S**: the node is transmitting a packet which is successfully transmitted according to the protocol model. The state begins as the transmission begins and lasts for $D_S = 10$ slots.
- **F**: the node transmits a packet which results in a collision. The F state also lasts for $D_F = 10$ slots.
- **I**: the node is either idle and does not have traffic to transmit or has traffic to transmit but is waiting for the channel to clear before transmitting. The duration of the I state is one slot.

As in [54], we assume that once a node finishes a transmission it returns to the idle state with probability 1, therefore $P_{SI} = P_{CI} = 1$. The resulting state transition diagram of a node is shown in Figure B.1.

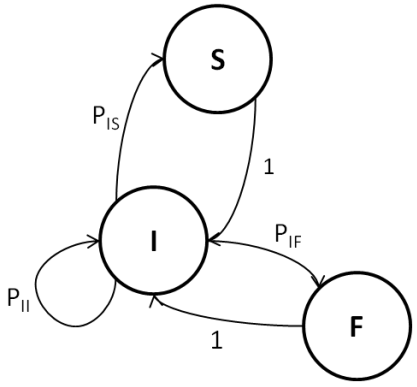

Figure B.1 State transition diagram for a node

Assume that P is a transmitting node and Q is the receiving node separated by a hop distance of d_H. The analysis of the above Markov chain provides us with the *long-run probability* of P being idle (P_I), transmitting successfully (P_S), or transmitting unsuccessfully (P_F). The long-run probability (P_X) denotes the average proportion of time a node is in state X. As discussed in Section 2.3.2, the distribution of the hop distance seen in simulations $f(d_H; n)$, is approximated by a Rayleigh distribution with Raleigh parameter $b(n) = r(n)\sqrt{\frac{2}{\pi}}$, where $r(n) = \sqrt{\frac{log\,(n)}{\pi n}}$. We calculate the expected number of successful transmissions and the transport capacity achieved by

the overall network by assuming each of the *n* nodes to be independent. Therefore, the number of nodes expected to be successfully transmitting a packet in a network of *n* nodes is

$$N_T(n) = n * \int_0^{r(n)} P_S(d_H)\ f(d_H; n) dd_H. \tag{B.1}$$

The network capacity is the product of the bit rate and the distance transported towards the final destination, which is difficult to calculate. The network capacity of a network of *n* nodes is upper bounded by

$$C_T(n) \leq n * \int_0^{r(n)} P_S(d_H) *\ d_H * f(d_H; n) dd_H \tag{B.2}$$

B.1.1 Transition Probabilities

For random distribution of nodes according to a 2-dimensional Poisson point process with intensity λ, the probability of finding k nodes in an area S is $P(k) = e^{-\lambda S} \frac{(\lambda S)^k}{k!}$ [45]. A 2-dimensional Poisson process describes a random scattering of points in an infinite plane. The following analysis uses this equation to find the probability of *k* nodes within the sensing and interference region of a node, which is much smaller than the total network area of 1 m^2. Figure B.2 illustrates the transmission scenario of P transmitting to Q, where $S_1(d_H)$ is the *hidden area* to node P. The hidden area to node P is the portion of the interference region of Q where P is unable to sense any transmitters.

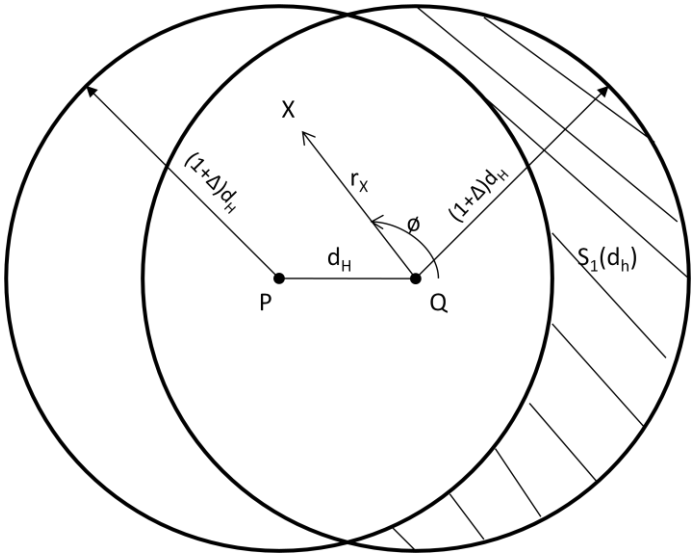

Figure B.2 Transmission scenario.

When a node P has data to transmit to a node Q at a distance of d_H, it senses the channel for traffic with sensing radius $R_{CS} = (1 + \Delta)d_H$. The probability that all nodes within a range of R_{CS} from P are idle is

$$P(I|R_{CS}) = \sum_{k=0}^{\infty} (P_I)^k \; e^{-N} \frac{N^k}{k!} = e^{-N(1-P_I)}, where\ N = n\pi {R_{CS}}^2. \tag{B.3}$$

When P has traffic to transmit to node Q, it senses the channel up to a radius R_{CS} and if the channel is clear it transmits to Q. The probability that node P transmits from the idle state, as a function of d_H, is

$$p(d_H) = \; p_T e^{-N(1-P_I)}. \tag{B.4}$$

In order for Q to receive the packet successfully, the nodes in the interference region of Q must also be idle. The probability that nodes in the interference region of Q are idle, given that P senses the channel as idle, is the probability that nodes in the hidden area $S_1(d_H)$ are idle:

$$P_{S_1}(d_H) = \; e^{-|\, S_1(d_H)|\, n(1-P_I)}, \tag{B.5}$$

where | | denotes area. Referring to [45], the hidden area is given by $|S_1(d_H)| = \pi {R_{CS}}^2 - 2{R_{CS}}^2 q\left(\frac{1}{2(1+\Delta)}\right)$, where $q(t) = \arccos(t) - \; t\sqrt{1-t^2}$. Next, given that all nodes inside the union of the circles centered at P and Q with radius R_{CS} are idle at the beginning of the slot where P begins to transmit, a collision could still occur if a third node X inside the interference region of Q also senses the channel as idle and begins transmission in the same slot. The probability that another node X begins transmission in the same slot, to a node at a distance of d_{H_x} is

$$p'_2\left(\phi, r_{x,} d_H, d_{H_x}\right) = \; p_T e^{-|\, S_2\left(\phi, r_{x,} d_H, d_{H_x}\right)| n(1-P_I)}. \tag{B.6}$$

As seen in Figure B.3, ϕ is the angle from the PQ axis to X and r_x is the distance from X to Q and the expression $S_2\left(\phi, r_{x,} d_H, d_{H_x}\right)$ denotes the portion of the sensing region of X which does not overlap with the sensing region of P or the interference region of Q.

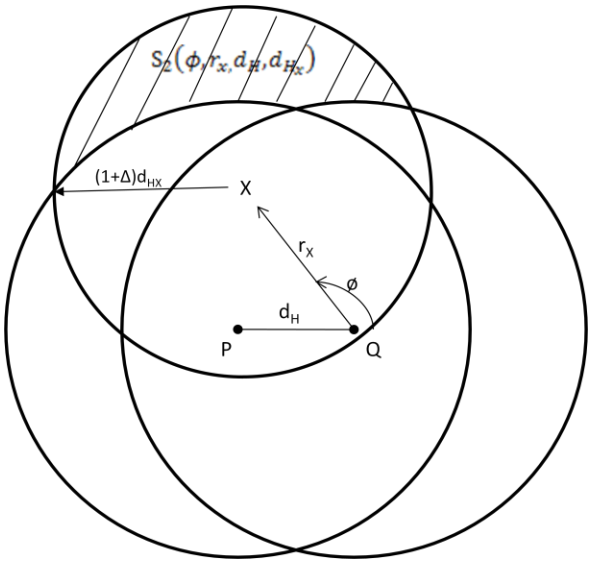

Figure B.3 Interference scenario.

Since the attempted hop distance d_{H_x} also varies according to $f^n_{\widetilde{d_H}}(d_{H_x})$, the overall probability that a node third node X begins transmission in the same slot as P is

$$p_2(\phi, r_x, d_H) = \int_0^{r(n)} p'_2(\phi, r_x, d_H, d_{H_x})\, f^n_{\widetilde{d_H}}(d_{H_x}) dd_{H_x}. \tag{B.7}$$

Integrating over all possibilities of ϕ and r_x yields the overall probability that no nodes in the circle with center Q and radius R_{CS} transmit in the same slot as P:

$$P_{L_1}(d_H) = e^{-n\int_0^{2\pi}\int_0^{R_{CS}} r_x p_2(\phi, r_x, d_H) dr_x d\phi}. \tag{B.8}$$

The probability that no nodes in the interference area of Q start transmission during the duration of P's transmission is found in a similar way. The probability that a node X begins transmission while P is transmitting is

$$p_3(\phi, r_x, d_H) = \int_0^{r_{xp}/(1+\Delta)} p'_2(\phi, r_x, d_H, d_{H_x})\, f^n_{\widetilde{d_H}}(d_{H_x}) dd_{H_x}. \tag{B.9}$$

Where r_{xp} is the distance between X and P, and the intended transmission range of X is capped at $\frac{r_{xp}}{1+\Delta}$ since X will sense P with probability one if $d_{H_x} \geq \frac{r_{xp}}{1+\Delta}$. Therefore, the probability that no nodes in the interference region of Q transmit in a single slot during P's transmission is

$$P'_{L_2}(d_H) = e^{-n\int_0^{2\pi}\int_0^{R_{CS}} r_x p_3(\phi, r_x, d_H) dr_x d\phi}. \tag{B.10}$$

The overall probability that no nodes in the hidden region transmit within the duration of the transmission D_P while P is transmitting is

$$P_{L_2}(d_H) = P'_{L_2}(d_H)^{D_P}. \quad \text{(B.11)}$$

In summary, the transition probability that node P changes from state I to state S given that Q is a distance of d_H is the overall probability that:

- P transmits to Q with probability $p(d_H)$.
- All nodes in hidden area of P are idle with probability $P_{S_1}(d_H)$.
- Q does not transmit with probability $(1 - p_T)$.
- No other nodes inside the interference region of Q begin transmission at the same slot with probability $P_{L_1}(d_H)$.
- No other nodes inside the interference region of Q transmit during D_P with probability $P_{L_2}(d_H)$.

If all of the above conditions are met a node successfully transitions from state I to state S, therefore the transition probability $P_{IS}(d_H)$ is

$$P_{IS}(d_H) = p(d_H) * P_{S_1}(d_H) * (1 - p_T) * P_{L_1}(d_H) * P_{L_2}(d_H). \quad \text{(B.12)}$$

Similarly, the probability that P transmits but the transmission fails, the transition probability from state I to state C is

$$P_{IC}(d_H) = p(d_H) * (1 - P_{S_1}(d_H)) + p(d_H) * P_{S_1}(d_H) * p_T + p(d_H) * P_{S_1}(d_H) * (1 - p_T) * \left(1 - P_{L_1}(d_H) * P_{L_2}(d_H)\right). \quad \text{(B.13)}$$

B.1.2 Long-run Probabilities

As in [54] the steady state probabilities of the node P being in states I, S, or C are as follows.

$$\pi_I(d_H) = \frac{1}{P_{IS}(d_H) + P_{IC}(d_H) + 1} \quad \text{(B.14)}$$

$$\pi_S(d_H) = \frac{P_{IS}(d_H)}{P_{IS}(d_H) + P_{IC}(d_H) + 1} \quad \text{(B.15)}$$

$$\pi_C(d_H) = \frac{P_{IC}(d_H)}{P_{IS}(d_H) + P_{IC}(d_H) + 1} \quad \text{(B.16)}$$

As discussed earlier we model each node as a semi-Markov process which is essentially a Markov process where the duration of each state is variable. If the expected duration of a state i is μ_i, the limiting probability of a node being in state i is $P_i = \frac{\pi_i \mu_i}{\sum_j \pi_j \mu_j}$ [58].

Since the states S and C last for D_P slots and the state I lasts for 1 slot, the probability that node P is in state I is

$$P_I(d_H) = \frac{\pi_I(d_H)}{\pi_I(d_H) + \big(\pi_S(d_H) + \pi_C(d_H)\big)D_p}. \quad \text{(B.17)}$$

Similarly, the probability that node P is in state S is

$$P_S(d_H) = \frac{\pi_S(d_H)D_p}{\pi_I(d_H) + \big(\pi_S(d_H) + \pi_C(d_H)\big)D_p}. \quad \text{(B.18)}$$

The long-run probabilities are numerically calculated for a given d_H using a modified version of the root finding algorithm presented in [54].

B.2 Analytical Results vs. Simulation

Figure B.4 compares the results with the results seen in simulations Section 3.3.2 for $n = 100$ nodes. The simulation and theoretical results match relatively well for the number of simultaneous transmissions. The analytical results for transport capacity show a similar trend but do not match as closely. There are certain differences in the analysis and the simulations. For example, the analysis does not model any retransmissions due to the channel being busy, while we simulate retransmissions with a uniform random delay up to $2*D_{P.}$

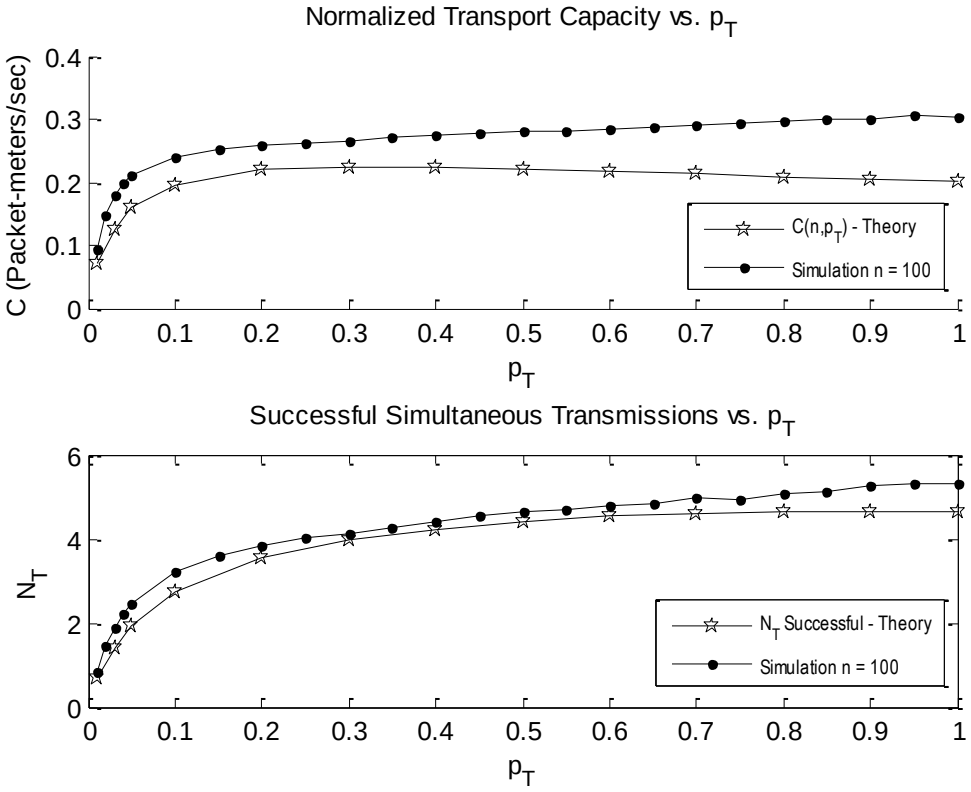

Figure B.4 Analytical vs. Simulation Results for CSMA, *n* = 100 nodes.

BIBLIOGRAPHY

[1] FCC, Docket No 03-222 Notice of proposed rule making and order, December 2003.

[2] F. H. Sanders, "Broadband spectrum surveys in Denver, CO, San Diego, CA, and Los Angeles, CA: methodology, analysis, and comparative results," in *IEEE International Symposium on Electromagnetic Compatibility, 1998*, pp. 988-993 vol.2.

[3] R. B. Bacchus, A. J. Fertner, C. S. Hood, and D. A. Roberson, "Long-Term, Wide-Band Spectral Monitoring in Support of Dynamic Spectrum Access Networks at the IIT Spectrum Observatory," in *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008*, pp. 1-10.

[4] P. Mohapatra and S. Krishnamurthy, *AD HOC NETWORKS: technologies and protocols*: Springer, 2005.

[5] I. Stojmenovic, *Handbook of wireless networks and mobile computing*: Wiley-Interscience, 2002.

[6] L. Gavrilovska and R. Prasad, *Ad hoc networking towards seamless communications*: Springer Verlag, 2006.

[7] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062-3080, Dec. 2004.

[8] H. Xiaoyan, X. Kaixin, and M. Gerla, "Scalable routing protocols for mobile ad hoc networks," *IEEE Network,* vol. 16, pp. 11-21, Aug. 2002.

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

[10] N. Abramson, "Development of the ALOHANET," *IEEE Trans. Inform. Theory,* vol. 31, pp. 119-123, Mar. 1985.

[11] R. Rom and M. Sidi, *Multiple access protocols: performance and analysis*: Citeseer, 1990.

[12] K. Römer, "Time synchronization in ad hoc networks," in *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing, 2001*, pp. 173-182.

[13] L. Kleinrock and F. Tobagi, "Packet Switching in Radio Channels: Part I--Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics," *IEEE Trans. Commun.,* vol. 23, pp. 1400-1416, Dec. 1975.

[14] F. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II--The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution," *IEEE Trans. Commun.,* vol. 23, pp. 1417-1433, Dec. 1975.

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

[16] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Int. J. Comput. Telecommun. Netw.,* vol. 50, pp. 2127-2159, Sept. 2006.

[17] J. Mitola, III and G. Q. Maguire, Jr., "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.,* vol. 6, pp. 13-18, Aug. 1999.

[18] C. Cormio and K. R. Chowdhury, "A survey on MAC protocols for cognitive radio networks," *Ad Hoc Netw.,* vol. 7, pp. 1315-1329, Sept. 2009.

[19] W. Qiwei and Z. Haitao, "Route and spectrum selection in dynamic spectrum networks," in *3rd IEEE Consumer Communications and Networking Conference, 2006. CCNC 2006*, pp. 625-629.

[20] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Commun. Mag.,* vol. 46, pp. 32-39, Apr. 2008.

[21] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004*, pp. 772-776 Vol.1.

[22] C. R. C. da Silva, C. Brian, and K. Kyouwoong, "Distributed Spectrum Sensing for Cognitive Radio Systems," in *Information Theory and Applications Workshop, 2007*, pp. 120-123.

[23] C. Cordeiro and K. Challapali, "C-MAC: A Cognitive MAC Protocol for Multi-Channel Wireless Networks," in *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007*, pp. 147-157.

[24] W. Hui, Q. Hang, and Z. Li, "A Survey on MAC Protocols for Opportunistic Spectrum Access in Cognitive Radio Networks," in *International Conference on Computer Science and Software Engineering, 2008*, pp. 214-218.

[25] L. Ma, X. Han, and C. C. Shen, "Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*, pp. 203-213.

[26] M. Liangping, S. Chien-Chung, and R. Bo, "Single-Radio Adaptive Channel Algorithm for Spectrum Agile Wireless Ad Hoc Networks," in *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007*, pp. 547-558.

[27] J. Juncheng, Z. Qian, and S. Xuemin, "HC-MAC: A Hardware-Constrained Cognitive MAC for Efficient Spectrum Management," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 106-117, Jan. 2008.

[28] Y. R. Kondareddy and P. Agrawal, "Synchronized MAC Protocol For Multi-Hop Cognitive Radio Networks," in *IEEE International Conference on Communications, 2008. ICC '08*, pp. 3198-3202.

[29] A. Chia-Chun Hsu, D. S. L. Weit, and C. C. J. Kuo, "A Cognitive MAC Protocol Using Statistical Channel Allocation for Wireless Ad-Hoc Networks," in *IEEE Wireless Communications and Networking Conference, 2007. WCNC 2007*, pp. 105-110.

[30] P. Pawelczak, R. Venkatesha Prasad, L. Xia, and I. G. M. M. Niemegeers, "Cognitive radio emergency networks - requirements and design," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005*, pp. 601-606.

[31] Z. Qing, T. Lang, S. Ananthram, and C. Yunxia, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 589-600, Apr. 2007.

[32] B. Hamdaoui and K. G. Shin, "OS-MAC: An Efficient MAC Protocol for Spectrum-Agile Wireless Networks," *IEEE Trans. Mobile Comput.*, vol. 7, pp. 915-930, Aug. 2008.

[33] S. Hang and Z. Xi, "Opportunistic MAC Protocols for Cognitive Radio Based Wireless Networks," in *41st Annual Conference on Information Sciences and Systems, 2007. CISS '07*, pp. 363-368.

[34] L. Long and E. Hossain, "A MAC Protocol for Opportunistic Spectrum Access in Cognitive Radio Networks," in *IEEE Wireless Communications and Networking Conference, 2008. WCNC 2008*, pp. 1426-1430.

[35] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. INFOCOM 2004*, p. 475.

[36] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay scaling in wireless networks with constant-size packets," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005*, pp. 1329-1333.

[37] P. Kyasanur and N. H. Vaidya, "Capacity of Multichannel Wireless Networks Under the Protocol Model," *IEEE/ACM Trans. Netw.*, vol. 17, pp. 515-527, Apr. 2009.

[38] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, pp. 477-486, Aug. 2002.

[39] F. Xue and P. Kumar, *Scaling laws for ad-hoc wireless networks: An information theoretic approach*: Now Pub, 2006.

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

[41] M. I. Shamos and D. Hoey, "Closest-point problems," in *16th Annual Symposium on Foundations of Computer Science, 1975*, pp. 151-162.

[42] M. A. Marsan and D. Roffinella, "Multichannel Local Area Network Protocols," *IEEE J. Sel. Areas Commun.*, vol. 1, pp. 885-897, Nov. 1983.

[43] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity," in *Proceedings of the 37th IEEE Conference on Decision and Control, 1998*, pp. 1106-1110 vol.1.

[44] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, pp. 421-436, Feb. 2006.

[45] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Trans. Commun.*, vol. 32, pp. 246-257, Mar. 1984.

[46] Y. Zhongbang, V. O. K. Li, and C. Zhigang, "Maximum throughput analysis and enhancement of slotted ALOHA for multihop ad hoc networks," in *IEEE International Conference on Communications, 2004*, pp. 4162-4166 Vol.7.

[47] L. Hyun-Kwan and K. Seong-Lyun, "Network Coded ALOHA for Wireless Multihop Networks," in *IEEE Wireless Communications and Networking Conference, 2009. WCNC 2009*, pp. 1-5.

[48] X. Min and M. Haenggi, "Delay performance of different MAC schemes for multihop wireless networks," in *IEEE Global Telecommunications Conference, 2005. GLOBECOM '05*, pp. 5 pp.-3427.

[49] J. Mo, H. S. W. So, and J. Walrand, "Comparison of Multichannel MAC Protocols," *IEEE Trans. Mobile Comput.,* vol. 7, pp. 50-65, Jan. 2008.

[50] P. Pawelczak, S. Pollin, H. S. W. So, A. Bahai, R. V. Prasad, and R. Hekmat, "Performance Analysis of Multichannel Medium Access Control Algorithms for Opportunistic Spectrum Access," *IEEE Trans. Veh. Technol.,* vol. 58, pp. 3014-3031, Jul. 2009.

[51] H. Al-Mahdi, M. Kalil, F. Liers, and A. Mitschele-Thiel, "Increasing spectrum capacity for ad hoc networks using cognitive radios: an analytical model," *Communications Letters, IEEE,* vol. 13, pp. 676-678, 2009.

[52] L. Wu and P. Varshney, "Performance analysis of CSMA and BTMA protocols in multihop networks (I). single channel case," *Inform. Sci,* vol. 120, pp. 159-177, Nov. 1999.

[53] L. Wu and P. K. Varshney, "Performance analysis of CSMA and BTMA protocols in multihop networks (II). Multiple channel case," *Inform. Sci,* vol. 120, pp. 179-195, Nov. 1999.

[54] D. Hoang and R. A. Iltis, "A unified framework for performance evaluation of multi-hop ad hoc networks," in *IEEE Military Communications Conference, 2005. MILCOM 2005*, pp. 2544-2550 Vol. 4.

[55] D. Hoang and R. A. lltis, "Performance evaluation of multi-hop csma/ca networks in fading environments," *IEEE Trans. Commun.,* vol. 56, pp. 112-125, Jan. 2008.

[56] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels," *IEEE Trans. Commun.,* vol. 55, pp. 21-24, Jan. 2007.

[57] A. Papoulis, *Probability, random variables, and stochastic processes*, Third ed.: McGraw-Hill New York, 1991.

[58] S. Ross, *Applied probability models with optimization applications*: Holden-Day Series in Management Sciences, San Francisco etc, 1970.