OPTIMAL SPATIAL REUSE IN MOBILE AD HOC NETWORKS

LUBO SONG

Bachelor of Computer Science
Chongqing University
June, 1993

Master of Electrical Engineering
Chongqing University
June, 1996

submitted in partial fulfillment of requirements for the degree

DOCTOR OF ENGINEERING

at the

CLEVELAND STATE UNIVERSITY

May, 2007

This dissertation has been approved
for the Department of Electrical and Computer Engineering
and the College of Graduate Studies by

Dissertation Committee Chairperson, Dr. Chansu Yu

Department/Date

Dr. Janche Sang

Department/Date

Dr. Yongjian Fu

Department/Date

Dr. Nigamanth Sridhar

Department/Date

Dr. Wenbing Zhao

Department/Date

ACKNOWLEDGEMENTS

Many thanks go to my advisor, Dr. Chansu Yu, for giving me such a challenging opportunity, for his inspiring direction for the research and for his great support in all aspects of my study.

I would like to thank Dr. Janche Sang, Dr. Dan Simon, Dr. Yongjian Fu, Dr. Nigamanth Sridhar and Dr. Wenbing Zhao, who are on my committee, for their time in reviewing and evaluating this dissertation.

Thank you to my family and friends for standing by me while I went crazy and stayed up days and nights on end trying to finish this on time.

OPTIMAL SPATIAL REUSE IN MOBILE AD HOC NETWORKS

LUBO SONG

ABSTRACT

Carrier-Sense (CS) *Medium Access Control* (MAC) protocols such as IEEE 802.11 MAC, called *Distributed Coordination Function* (DCF), avoid collisions by holding up pending transmissions when the carrier signal is observed above a certain threshold. However, this often results in unnecessarily conservative communication, thus making it difficult to maximally reuse the spatial spectral resource in *Mobile Ad hoc Networks* (MANETs).

This dissertation shows that more communications can be simultaneously successful when the communication distance or the spatial reservation is adjustable. Two corresponding solutions are proposed: *Multiple Access with Salvation Army* (MASA) and *Collision-Aware DCF* (CAD). MASA adopts less sensitive carrier sensing, equivalently a higher threshold, to promote more concurrent communications. While this potentially increases the collision probability, MASA effectively addresses this problem by adjusting the communication distance adaptively via “packet salvaging” at the MAC layer. In comparison to MASA, CAD’s approach is proactive in the sense that it tries to efficiently utilize the available spatial resource through “collision prediction”. In other words, MASA and CAD are not incompatible with each other and can be integrated into a single MAC protocol when more optimized performance is desired. Extensive simulation based on ns-2 has shown that they substantially outperform the DCF.

Keywords: Carrier Sense, Mobile Ad hoc Networks, Medium Access Control, Capture Effect, DCF, Spatial Reuse.

TABLE OF CONTENTS

Page

NOMENCLATURE

AP:	Access Point
ACK:	Acknowledgement
CAD:	Collision-Aware DCF
CCA:	Clear Channel Assessment
CDF:	Cumulative Distribution Function
CFP:	Collision-Free communication Pair
CS:	Carrier Sense
CSMA:	Carrier Sense Multiple Access
CSMA/CA:	CSMA with Collision Avoidance
CSMA/CD:	CSMA with Collision Detection
CSR:	Carrier Sense Range
CTS:	Clear-To-Send
DCF:	Distributed Coordination Function
DIFS:	Inter-Frame Space
DSSS:	Direct Sequence Spread Spectrum
ED:	Energy Detection
EIFS:	Extended Inter-Frame Space

FHSS:	Frequency Hopping Spread Spectrum
IFR:	Infrared
IFS:	Inter-Frame Space
IR:	Interference Range
LAN:	Local Area Network
LLC:	Logic Link Control
LOS:	Line-Of-Sight
MAC:	Medium Access Control
MANET:	Mobile Ad hoc Network
MASA:	Multiple Access with Salvation Army
MIMO:	Multiple Input Multiple Output
MPDU:	MAC Protocol Data Unit
NAV:	Network Allocation Vector
NLOS:	No-Line-Of-Sight
OFDM:	Orthogonal Frequency Division Multiplexing
PC:	Point Coordinator
PCF:	Point Coordination Function
PDA:	Personal Digital Assistant
PDF:	Probability Distribution Function

PDR: Packet Delivery Ratio

PIFS: PCF Inter-Frame Space

PLCP: Physical Layer Convergence Protocol

QoS: Quality of Service

RSSI: Received Signal Strength Index

RSR: Required Spatial Reservation

RTS: Request-To-Send

RTT: Round Trip Time

SACK: Salvaging ACK

SDATA: Salvaging DATA

SIFS: Short Inter-Frame Space

SINR: Signal to Interference and Noise Ratio

SR: Spatial Reservation

TDMA: Time Division Multiple Access

TR: Transmission Range

WLAN: Wireless LAN

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The information in this chapter includes a brief overview of the history of wireless communication, popular wireless networks such as *Wireless Local Area Network* (WLAN) and *Mobile Ad hoc Network* (MANET) and the existing *Medium Access Control* (MAC) protocols for wireless networks. In addition, the motivation for the research conducted in this dissertation will be explained. Two proposed solutions, *Multiple Access with Salvation Army* (MASA) and *Collision-Aware DCF* (CAD) will be simply introduced. Finally the organization of the dissertation will be presented.

1.1 A Short Overview of Wireless Communication History

The history of wireless communication begins as early as two centuries ago. In the early 1800s, scientists such as Michael Faraday and Heinrich Rudolf Hertz envisioned the possbility of wireless communication [1]. In the 1830s, William Cooke and Charles Wheatstone developed the first electric telegraph for commercial service [2]. In 1901, Guglielmo Marconi successfully transmitted a radio signal across Atlantic Ocean from

Cornwall to Newfoundland [3]. By 1920s, mobile wireless receivers had been developed and installed in police cars in Detroit [4].

In the 1960s, the thoughts about networks were at hand in the United States, due to the launching of the Sputnik in the USSR [5]. In 1971, when networking technologies met wireless communication, ALOHNET, the first wireless network was born at the University of Hawaii [6]. In 1990, the IEEE 802 Executive Committee began to define standards for wireless networking. The recent rapid development of radio communication technologies and wireless networking has become a major element of the IT revolution.

1.2 Wireless LAN and MANET

A wireless network is a collection of wireless devices such as laptops, *Personal Digital Assistants* (PDAs) [7, 8], smart IP phones [9] and fixed devices that are equipped with wireless network interface cards. These wireless devices communicate via radio, exchange information, and share resources such as printers, files, internet access, etc. From the perspective of network connectivity, the wireless network can be classified into single-hop and multi-hop networks. *Wireless LAN* (WLAN) [10] is a typical single-hop wireless network, in which a mobile user can connect to the Internet via an *Access Point* (AP), as shown in Figure 1.

The WLAN technology was standardized as Wi-Fi by the IEEE 802.11 [11, 12] (abbreviated as 802.11 later) group and is widely applied in schools, hospitals, homes and businesses. According to the research report from Strategy Analytics Inc. [13] in 2006, more than 80 percent of laptops in the US have Wi-Fi, with 88 percent of business professionals actively using this feature. Experts have forecasted that in the next five years a cumulative 940 million wireless devices will be needed for use at home.

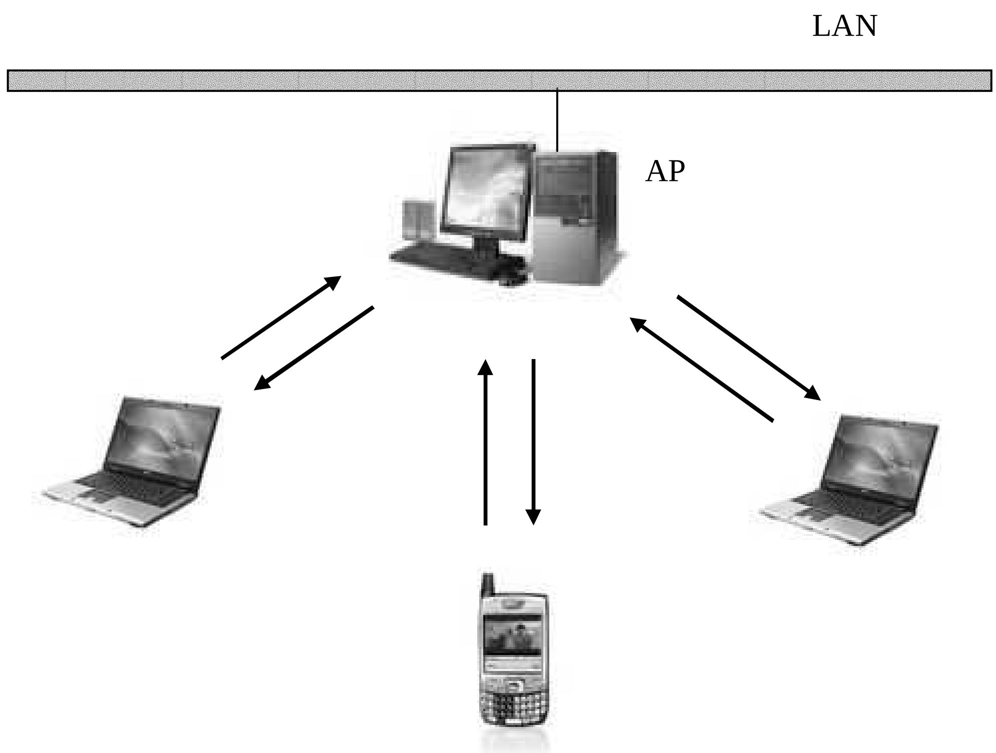


Figure 1: Single-hop Wireless LAN

The WLAN based on 802.11 has been a great success; however, it has some significant limitations. First, it is a single-hop wireless network, where the mobility of wireless stations are limited within the AP's coverage, usually is about 45m (150ft) indoors and 90m (300ft) outdoors. In addition, the WLAN cannot be deployed or applied in all environments, namely, where the infrastructure (AP) is difficult or very costly to install, such as deserts, battle fields, and fire spots. In these environments, wireless devices are required to coordinate, exchange and share information and resources in a distributed way. Every device communicates with each other in a peer-to-peer fashion.

Due to the movements of the devices, the network topology varies over time. This type of network is known as a *Mobile Ad hoc Network* (MANET) [14-17] as shown in Figure 2.

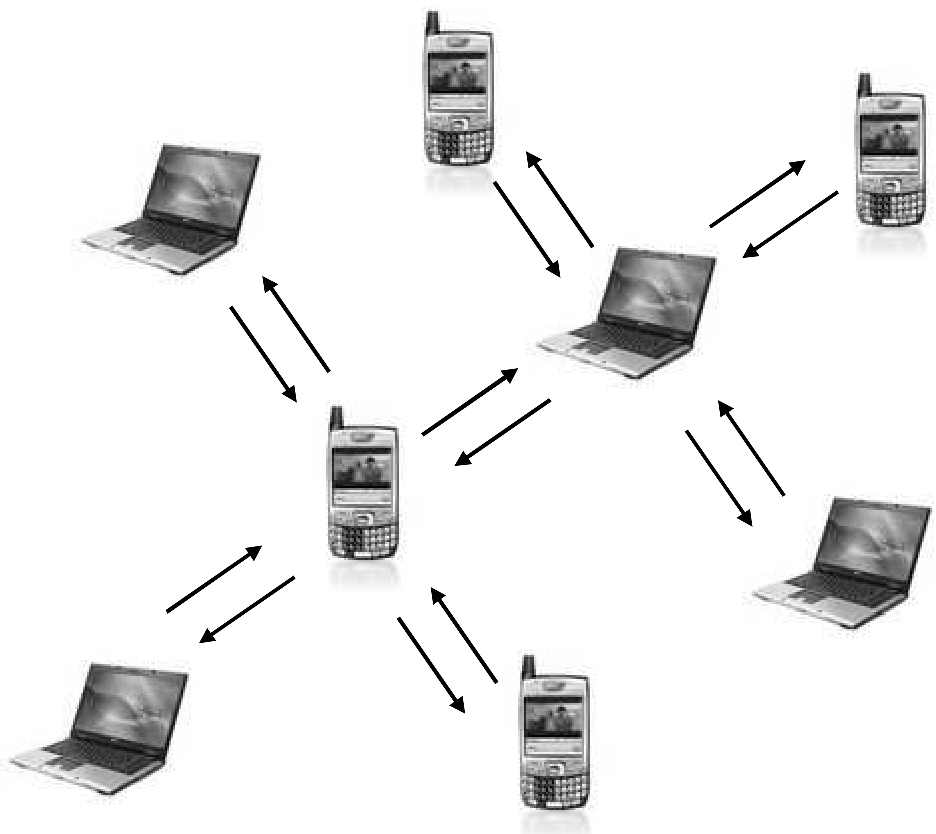


Figure 2: Mobile Ad hoc Network

In general, a MANET is a group of mobile wireless nodes that are self-configuring and infrastructure-independent. They are connected by wireless links, the union of which forms a communication network with arbitrary topology. Each node in the network acts not only as an end system, but also as a router to forward packets. In summary, a MANET is a peer-to-peer multi-hop mobile wireless network, which requires a different set of operation principles than those developed for WLANs.

In comparison to WLANs, MANETs are much more flexible; this means, however, they are also more complicated. Although it has been studied for more than a decade, there are still a number of open issues such as energy efficiency, effective multi-hop routing, wireless communication security, and etc. This dissertation focuses on the problem of spatial spectral efficiency at the MAC layer.

1.3 CSMA Protocols

A MAC protocol provides the arbitration method for efficient use of the shared medium among multiple stations or network nodes. Typical wired MAC protocols are *Carrier Sense Multiple Access with Collision Detection* (CSMA/CD, IEEE 802.3) [18], token ring (IEEE802.4) [19] and token bus (IEEE802.5) [20]. Typical MAC mechanisms for wireless networks include *Time Division Multiple Access* (TDMA) [21] and *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) [22].

CSMA/CA and CSMA/CD are the most popular MAC mechanisms in wired and wireless networks, respectively. They are both based on carrier sensing and are dedicated to prevent collisions. The *Carrier Sense* (CS) in the CSMA describes the fact that a transmitter listens for carrier signals before transmitting. If a carrier is sensed, the node waits for the ongoing transmission to be finished before starting its own transmission. *Multiple Access* (MA) describes the fact that multiple nodes share and access the common medium.

Since the transmission from a node is supposed to be received by all other nodes using the same medium [23], even wth CSMA collisions are still possible if more than one node begins their transmissions at the same time. In wired networks, *Collision Detection* (CD) is employed to improve CSMA performance by stopping transmission as soon as a collision is detected. However, this technique cannot be used in the wireless networks. First, it is impossible to listen while transmitting in some wireless networks like WLANs. Second, the collision at the receiver is very difficult to detect at the transmitter because of the physical seperation of the two nodes. In wireless networks, instead of CD, *Collision Avoidance* (CA) is employed to prevent collisions. According to CA, a pair of communication nodes should exchange control signals before transmitting data frames. A node should defer its transmission if one of the control signals is sensed. The deferment details will be described in the next chapter.

1.4 Motivation

The IEEE 802.11 MAC [24] adopts the CSMA/CA and achieved a great success. Due to this success as well as to the absence of industrial standards for MANETs, 802.11 MAC is generally accepted for use in MANETs as well. However, 802.11 MAC may not be appropriate for MANETs because originally this standard was designed for the single-hop WLAN. The multi-hop MANET environment complicates the communication situation and thus it is not straighforward to apply WLAN technologies in the MANETs.

The motivation of this dissertation is to reconsider design choices made for a single-hop wireless network in the context of a multi-hop wireless environment.

More specifically in this dissertation, the carrier sense mechanism will be reconsidered in an attempt to fine tune it for use in the multi-hop MANET environment. Note that 802.11 MAC does CS by comparing the strength of incoming signal with a pre-set threshold called *CS threshold* [25]. If the signal strength is weaker than the threshold, then the medium is considered idle. Otherwise, the medium is considered busy and nodes should defer their transmissions. A higher CS threshold makes a node less sensitive to the carrier signals from other nodes, especially signals from distant nodes. On the other hand, a lower CS threshold makes the node more sensitive. In a WLAN every wireless node is supposed to directly communicate with the AP only; if one node is communicating with the AP, all the other nodes are not supposed to interrupt them. In order to be aware of the presence of the ongoing communication the CS threshold is configured conservatively.

However, this does not work well in the MANETs, where nodes operate in a peer-to-peer way, delivering data packets through multiple hops. The spatial area becomes very valuable as the reuse of the spatial spectral resource is in fact the essential idea of the multi-hop networks. To make more nodes able to transmit their packets simultaneously, it seems that a larger CS threshold should be configured. However, this would result in more interference and more collisions. Obviously, there is a tradeoff between the spatial spectral utilization and collisions. The purpose of this dissertation is to find some efficient ways to improve the spatial reusability while the collision problems are effectively handled.

1.5 Proposed Solutions

This dissertation proposes two different solutions to efficiently improve the spatial reusability. Both of these solutions will allow for more concurrent communications by addressing the collision problem either after or before it happens. The solutions are referred to as *Multiple Access with Salvation Army* (MASA) and *Collision-Aware DCF* (CAD).

The key idea of MASA is to encourage more concurrent transmissions by employing a higher CS threshold first. If the collision does not occur, the spatial reusability is improved anyway. If the collision happens, MASA salvages the collided packet by some node between the sender and the receiver. This packet salvage automatically divides a long-distance link into two short-distance links, which are more robust against interferences.

While MASA is a post-collision method, CAD is a pre-collision method. The key idea of CAD is to predict collisions before transmission. In the CAD, at first each node estimates its reservation requirements based on the interference level and the communication distance. It then broadcasts the reservation requirements. Every overhearing node makes its transmission decision according to the received reservation requirements as well as its own reservation requirements.

Since 802.11 MAC has been widely adopted, the proposed MASA and CAD are implemented based on 802.11 specifications and do not demand any incompatible changes for immediate employment.

1.6 Organization of the Dissertation

This dissertation contains six chapters. Chapter I introduces the background, motivation and proposed solutions of the dissertation. Chapter II discusses the IEEE 802.11 in detail, with a focus on its MAC, DCF, which is the baseline model in this dissertation. Chapter III describes the propagation models and gives a thorough analysis on spatial spectral utilization with carrier-sense based MAC. Chapters IV and V presents the proposed solutions, MASA and CAD. Finally, Chapter VI draws conclusions and describes future work that can be based on the research presented in this dissertation.

CHAPTER II

IEEE 802.11

This chapter discusses the IEEE 802.11 in details. The discussion covers the overview and architecture of 802.11, and emphasizes on its MAC mechanism, *Distributed Coordination Function* (DCF).

2.1 Overview of IEEE 802.11

The IEEE 802.11 is a set of WLAN standards developed by working group 11 of the IEEE Standards Committee. The term of 802.11x is used to denote the variations and extensions of the original 802.11 (called 802.11legacy). The 802.11 legacy was documented in 1999. It only provides two transmit rates (1 and 2 Mbps). Its extension 802.11b improved the rate up to 11Mbps in 2003. Now 802.11a and 802.11g providing 54Mbps transmit rate have been widely accepted and used. And 802.11n that supports 540Mbps transmit rate by using *Multiple Input Multiple Output* (MIMO) antenna is being designed and will be available very soon.

TABLE I: IEEE 802.11 STANDARD FAMILY

Protocol	**Description**	**Release Date**
802.11legacy	Original 1 and 2 Mbps, 2.4 GHz RF and IR standard	1999
802.11a	54 Mbps, 5 GHz standard	1999
802.11b	Enhancements to 802.11 to support 5.5 and 11 Mbps	1999
802.11c	Network bridge operations	1998
802.11d	Operations in additional regulatory domains	2001
802.11e	QoS enhancements	2005
802.11f	Inter-Access Point Protocol	2003
802.11g	54 Mbps, 2.4 GHz standard	2003
802.11h	Spectrum and transmit power management extensions	2003
802.11i	Security enhancements	2004
802.11j	Extensions for Japan	2004
802.11k	Radio resource measurement enhancements	2007
802.11l	Reserved	Future
802.11m	Maintenance of the standard	Ongoing
802.11n	540 Mbps, 2.4 and 5 GHz standard (using MIMO)	2008
802.11o	Reserved	Future
802.11p	Wireless access for vehicular environment	2008
802.11q	Reserved	Future
802.11r	Fast roaming	2007
802.11s	ESS mesh networking	2008
802.11t	Wireless performance	2009
802.11u	Inter-working with non-802 networks	2008
802.11v	Wireless network management	2009
802.11w	Protected Management Frames	2008
802.11x	Reserved	Future
802.11y	3650-3700 Operation in the U.S.	2008
802.11z	Reserved	Future

Besides these modulation and air interface standards, 802.11 working group also specify other standards to enhance and extend services. For example, 802.11e specifies the standard for *Quality of Service* (QoS). 802.11i is an amendment to 802.11 legacy to improve the security of wireless access. 802.11p focuses on wireless access for the Vehicular Environment. 802.11s extends the services to ESS mesh networks. 802.11u works on inter-working with non-802 networks such as cellular and bluetooth. The details are shown in Table I [26].

2.2 Architecture of IEEE 802.11

In the 802.11 the standards specify several modulation and over-the-air techniques at physical layer such as *Frequency Hopping Spread Spectrum* (FHSS) [27, 28], *Direct Sequence Spread Spectrum* (DSSS) [29, 30], *Infrared* (IR) [31, 32] and *Orthogonal Frequency Division Multiplexing* (OFDM) [33, 34]. But they all are used by the same MAC protocol. Figure 3 shows the protocol stack of 802.11.

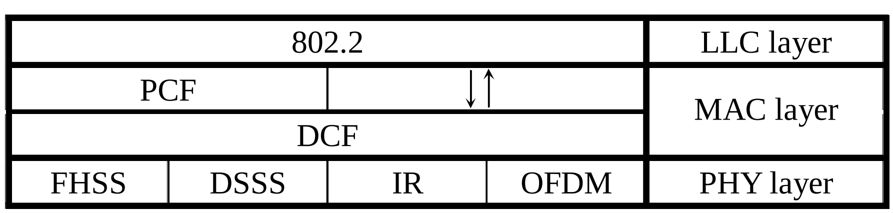


Figure 3: Protocol Stack of IEEE 802.11

At the MAC layer, 802.11 provides two access control functions, *Point Coordination Function* (PCF) and *Distributed Coordination Function* (DCF). In the DCF every node accesses the medium in a distributed way. But in the PCF a node is required to operate as *Point Coordinator* (PC) to schedule the medium access for all the nodes. So basically DCF is a contention-based method while PCF provides contention-free data transfer. However, they can coexist in the network and operate as shown in Figure 4. PCF and DCF control the medium access alternatively.

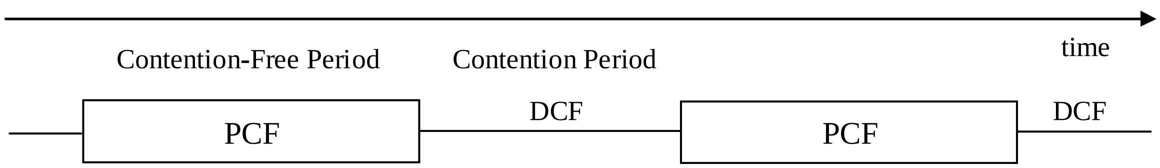


Figure 4: Coexistence of PCF and DCF

As shown in Figure 3, PCF resides on the top of DCF. That means 802.11 MAC provides PCF through the services of DCF. And, because of the requirement of the PC usually PCF is only usable in an infrastructure-available network such as WLAN, where the AP acts as the PC. In other words, in MANETs PCF is not applicable. Since MASA and CAD are all MAC solutions for MANETs both of them take DCF as their baseline mode.

2.3 Distributed Coordination Function (DCF)

DCF is a distributed MAC protocol that schedules automatic medium access among multiple wireless nodes. The basic access mechanism in DCF is CSMA plus CA. In addition, with consideration of unreliable links in wireless communication an acknowledgement frame (ACK) from the receiver is required in order to make sure that the Data frame is successfully delivered. In the case that ACK is not received the sender will think the delivery fails and make schedule for retransmission. This section will discuss these mechanisms in details.

2.3.1 Carrier Sense Multiple Access (CSMA)

DCF performs *Carrier Sense* (CS) through both *Physical Carrier Sense* (PCS) and *Virtual Carrier Sense* (PCS). According to the PCS a node is required to hold up its transmission if it detects the presence of other communications. The PCS in DCF is handled by the *Clear Channel Assessment* (CCA) function. The DSSS physical layer provides three CCA modes [24]. CCA mode 1: energy above threshold. CCA reports medium busy if the detected energy is above the *Energy Detection* (ED) threshold. More exactly, the *Received Signal Strength Index* (RSSI) is above the ED_THREHOLD (CS threshold). CCA mode 2: carrier sense only. CCA reports a busy medium only if it successfully detects a signal. CCA mode 3 is the combination of the mode 1 and 2. For the convenience CCA mode1 is employed in the dissertation for analysis and simulations.

The *Virtual Carrier Sense* (VCS) mechanism is achieved by distributing medium reservation information that informs the impending use of the medium. The reservation information is contained in the Duration/ID field of MAC header that defines the period of time needed to finish the whole communication procedure. To distribute the reserved period two control frames *Request-To-Send* (RTS) and *Clear-To-Send* (CTS) are exchanged before transmitting Data frame. Every overhearing node sets its *Network*

Allocation Vector (NAV) according to the value of the Duration/ID field in the received frame. The NAV works like a time counter, and nonzero of NAV indicates the impending use of the medium. An example of VCS and NAV is shown and explained in the Figure 9 in the subsection 2.3.6.

The DCF combines the NAV state and CCA indication to determine the medium state. The medium is regarded as idle only when the CCA indication is idle (for CCA mode 1, RSSI is less than CS threshold) and the NAV is zero. A pure CSMA mechanism allows a node to start its transmission when it senses the medium idle.

2.3.2 Inter-Frame Space (IFS)

In the DCF, when a node senses the medium idle it does not initialize its transmission immediately because this immediate transmission might interrupt some ongoing communication. Fox example, as shown in Figure 5, suppose the sender starts to send Data frame at t_0 and the transmission ends at t_1. Then after some period (SIFS, explain later) at t_3 the receiver is supposed to reply ACK frame to sender. However, if at t_2 node X senses the medium idle and initializes its transmission, at t_3 the receiver will not reply ACK back because it will sense the medium busy then. As a result, node X's transmission interrupts the ongoing communication between the sender and the receiver.

In order to prevent ongoing communications from interruption or to provide medium access priority, *Inter-Frame Space* (IFS) that is the time interval between frames is employed. In the 802.11 MAC there are four different IFS: *Short IFS* (SIFS), *DCF IFS*

(DIFS), *PCF IFS* (PIFS), and *Extended IFS* (EIFS), listed in the order from the shortest to the longest.

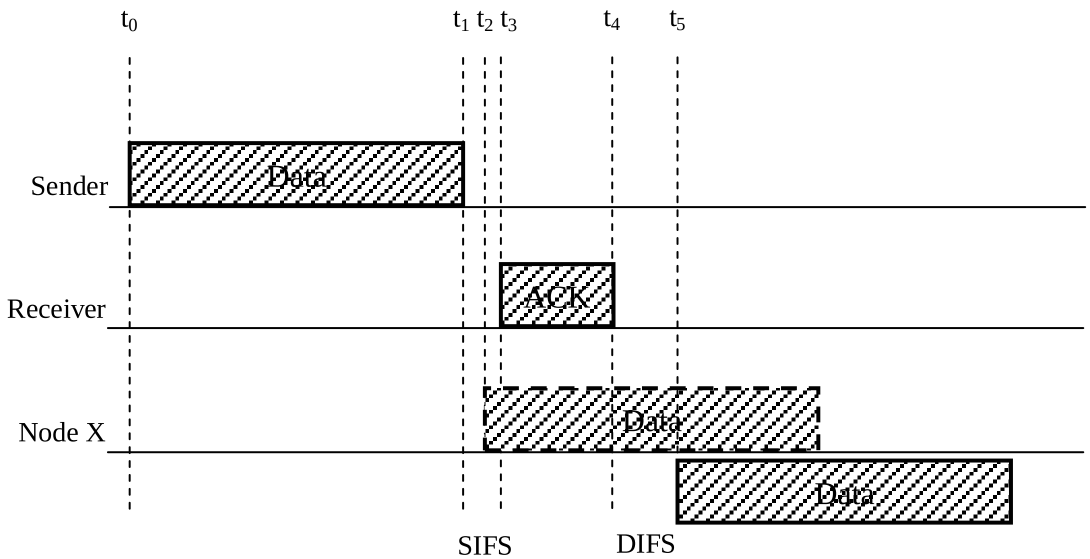


Figure 5: Interruption of Ongoing Communication

SIFS is the shortest one among them. It is used as the time interval of exchanging frames between the sender and the receiver. For example, as shown in Figure 5, after the sender ends its Data frame transmission, the receiver waits for SIFS and then replies ACK frame. The purpose of SIFS is to prevent the ongoing frame exchanges from interruption.

DIFS is the second shortest one and is used by nodes operating under the DCF to access the medium. More specifically, even if the medium is idle the node defers DIFS before starting its transmission. During the deferment if the medium keeps idle, the node can start its transmission after the deferment is finished. Otherwise it has to wait till the medium becomes idle and then defers DIFS again. By doing so, the ongoing communication can be protected from interruption. Consider the interruption case shown in the Figure 5 again, since DIFS > SIFS, node X never gets a chance to initialize its transmission during the SIFS period from t_1 to t_3. In other words, the receiver always has a higher priority to seize the medium than the waiting node, node X. Actually in this case, node X has to wait till the ACK transmission is finished, then defers DIFS period, and finally can start its transmission at t_5.

PIFS is the third shortest IFS. Like usage of DIFS, even if the medium is sensed idle a node operating under PCF should defer PIFS before starting its transmission. Because PIFS > DIFS a DCF node has a higher priority to gain medium access than a PCF node. But since in the PCF the PC exchanges frames with the wireless nodes using SIFS, once the PC has seized the medium DCF nodes have to wait till *Contention-Free Period* is finished and compete for medium access in the *Contention Period*, as shown in the Figure 4.

EIFS is the longest IFS and is applied by the DCF nodes whenever they receive an erroneous MAC frame. The major purpose of using EIFS is to protect ACK reception from collision.

2.3.3 Collision Avoidance (CA)

In order to avoid collisions the CS mechanism requires a node to check the medium status before transmission. In order to prevent ongoing frame exchanges from interruption IFS control is employed. But only the CS and IFS are not good enough to protect transmissions from collisions because they have no control in the case that multiple nodes wait for the busy medium and simultaneously access the medium after the ongoing communication and DIFS deferment are finished. For example, as shown in the Figure 6, the sender and the receiver exchange Data and ACK frames between t_0 and t_3. Suppose that nodes A and B have their Data frames ready at t_1 and t_2 respectively. According to the pure CSMA and IFS control mechanisms nodes A and B will simultaneously start their transmissions at t_4. Obviously nodes A and B will interfere with each other and possibly cause collisions at their intended receivers. And, since the wireless transmitters do not detect collisions they cannot stop transmitting until the entire frame is sent out and thus waste bandwidth.

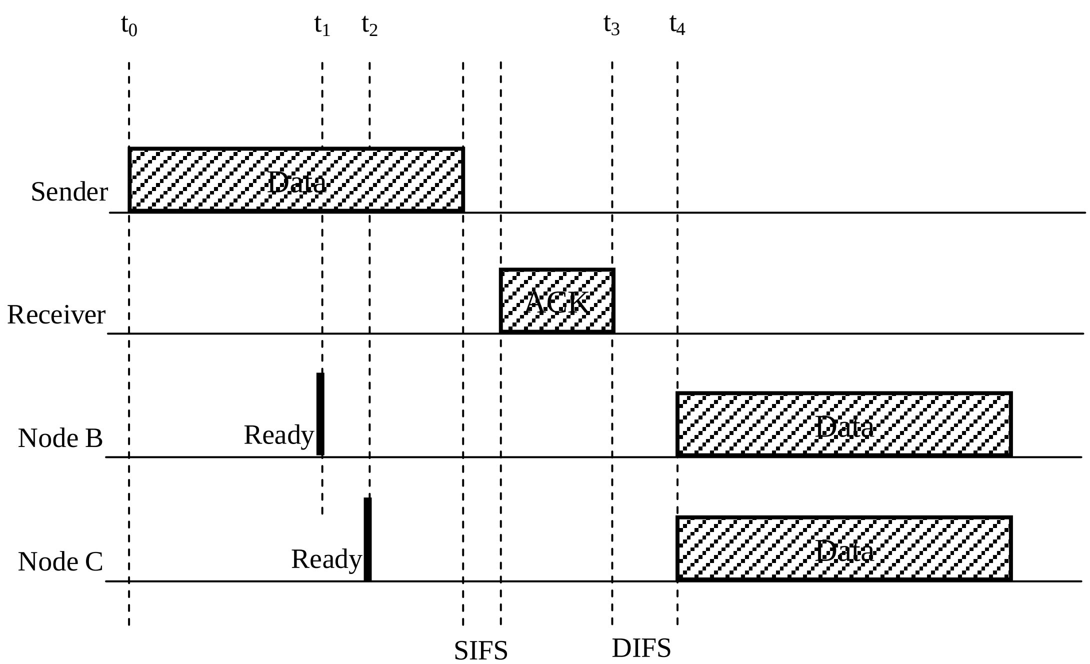


Figure 6: Collision Case in Pure CSMA

CA is designed to avoid this kind of collisions caused by the unwanted concurrent transmissions. According to the CA mechanism, the node does random backoff procedure after deferring DIFS. The backoff procedure pauses if medium becomes busy and resumes when the medium becomes idle again. Only when the backoff procedure is finished and medium is idle the node can initialize its transfer.

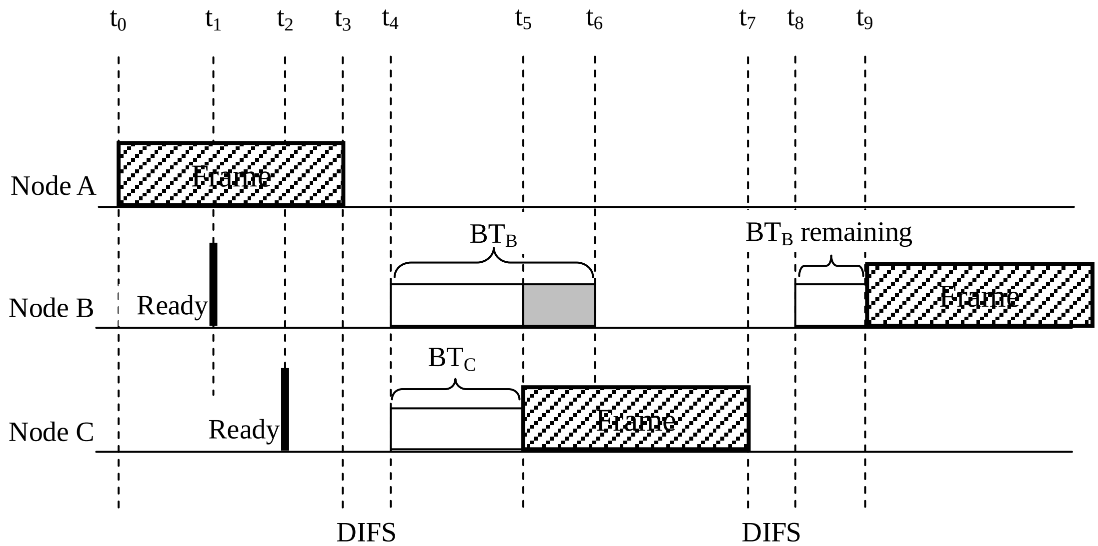


Figure 7: Backoff Procedure

Figure 7 shows an example of medium access with backoff procedure. Suppose that node A starts its transmission at t_0 and ends it at t_3, and nodes B and C have frames ready at t_1 and t_2 respectively. After node A finishes its communication both nodes B and C defer DIFS and then starts their backoff procedures at t_4. BT_B and BT_C stand for their backoff times respectively. Since BT_C is shorter than BT_B node C finishes the backoff procedure earlier than node C and gains the medium access at t_5. Because at that moment the medium becomes busy the backoff procedure of node B is paused. The shaded area shows the remaining backoff time of node B. In the next, after node C finishes its transmission node B defers DIFS and resumes its backoff procedure at t_8. After the backoff procedure is completely finished node B initiates its transmission at t_9.

The backoff time like BT_B and BT_C is a random period, which is calculated by the following equation.

$$\text{Backoff Time} = \text{Random()} \times \text{Slot Time} \tag{2.1}$$

where slot time is a system value determined by the wireless device. Random() generates a random integer that satisfies a uniform distribution over the interval [0, CW]. CW is called contention window size. Obviously the larger the CW is the longer backoff time is generated with higher probability. The longer backoff time is definitely more helpful to avoid collisions. But on the other hand it has the disadvantage of wasting bandwidth if network traffic is light (less number of nodes have transmission demand, lower collision probability). So ideally the CW size should vary over the network traffic. In the DCF, a method of *Exponential Backoff* is employed to dynamically adjust it. The details will be discussed in the next section.

2.3.4 Retransmission and Exponential Backoff

Because of interferences the wireless link is much less reliable than the wired one. In order to ensure that the Data frame is successfully delivered the receiver is required to reply a positive acknowledgement frame (ACK) after it successfully received the Data frame. If ACK is not received by the sender the transmission is regarded as failed and then retransmission is scheduled. If the retransmission fails again another one will be scheduled till the number of retransmission (denoted as *nRetr*) reaches the system retransmission limit (denoted as *nRetrLimit*). In this case the transmitted frame will be discarded anyway. Every time when a new frame is transmitted the retransmission count (*nRetr*) should be reset.

If the network traffic is heavy more collisions are caused, thus more retransmissions happen. So retransmission is a good indicator of the network traffic. DCF uses the number of consecutive retransmissions to dynamically adjust the CW size. The CW is initialized as CW_{min}. Every time a retransmission happens CW is almost roughly except when CW reaches CW_{max}. On the other hand, every time the data transfer is successful (the ACK is received) the CW is reset as CW_{min}. Since doubling the CW makes it increase exponentially this retransmission schedule is also referred as *Exponential Backoff*. The details on CW control due to successful transmission and the number of consecutive retransmissions are shown in the Figure 8.

2.3.5 Procedure of Medium Access and Frame Exchanges

Based on the MAC mechanisms introduced in the previous subsections, Figure 9 shows the whole procedure of medium access and frame exchanges. Suppose that at the beginning the medium is busy and the sender has a Data frame ready to send. First of all, the sender waits for the medium to become idle, and then defers DIFS and processes random backoff procedure. When the backoff procedure is finished the communication (frame exchanges) between the sender and the receiver begins. The exchanged frame sequence is RTS by the sender, CTS by the receiver, Data by the sender, and finally ACK by the receiver. To grab the medium access, SIFS is applied between the exchanged frames.

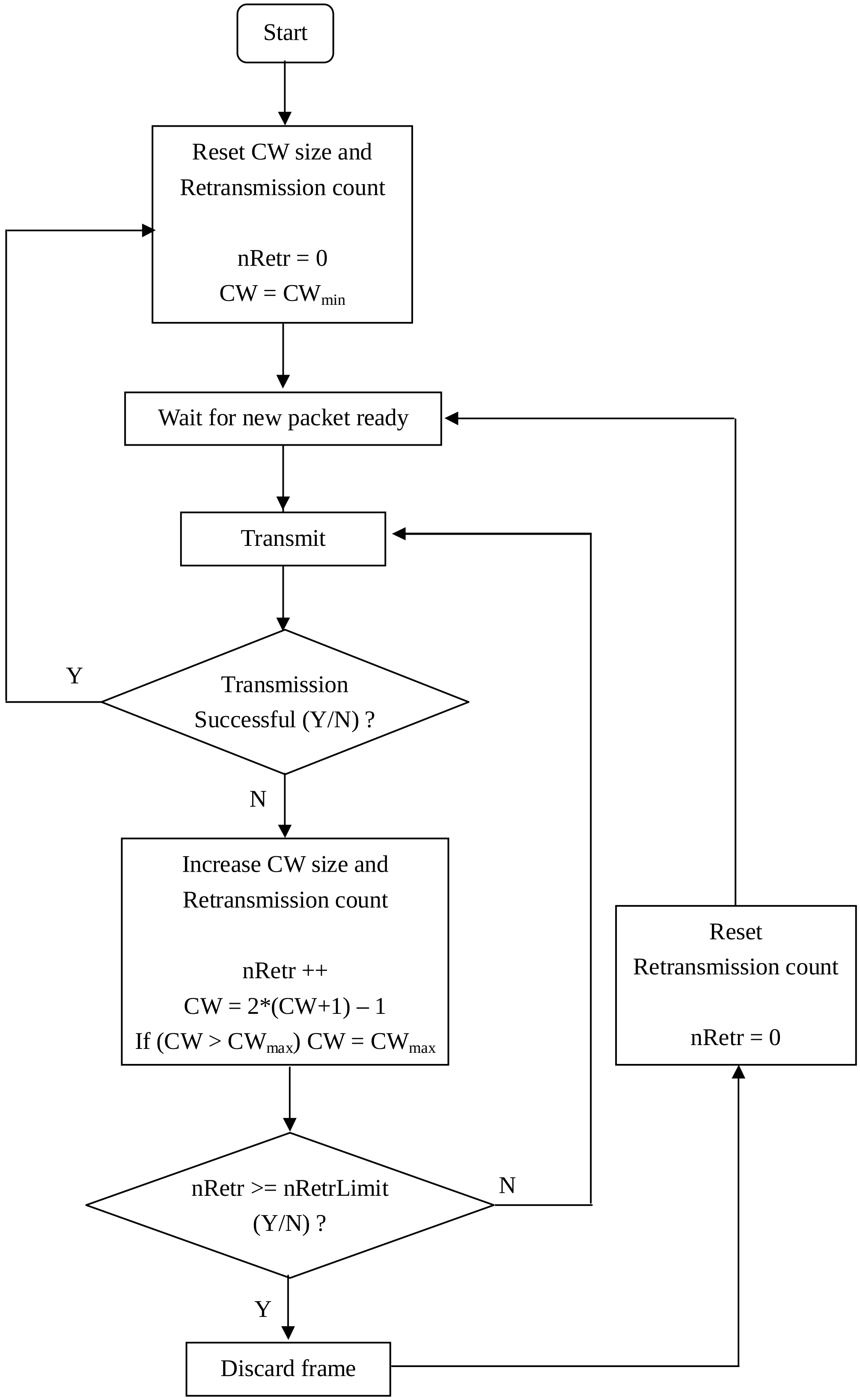


Figure 8: Adjust CW Due to Retransmissions

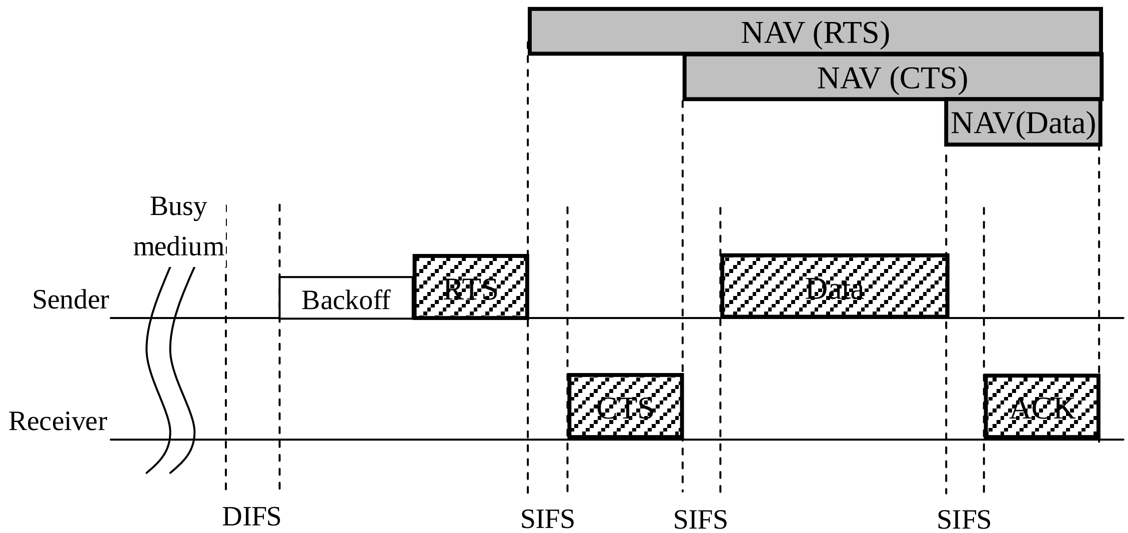


Figure 9: Frame Exchanges

Subsection 2.3.1 mentions that the NAV reservation is embedded in the Duration/ID field of MAC frame to announce the impending use of the medium. The information must be decoded after the whole frame is received. So the NAV information carried in a frame should be the time period for the rest of the communication. For example, as shown in the Figure 9, the NAV in RTS is the duration between the moment when the RTS is received and the moment when the ACK is received. Similarly, the NAVs for CTS and Data are estimated according to when ACK is received. Since ACK is the last exchanged frame, the NAV in ACK is 0, which means it does not have to make any reservation.

According to the VCS, the RTS/CTS exchange is used to distribute the NAV information. In addition, these two small control frame exchange can be used to perform a fast collision check. If CTS is not received that means either the RTS or the CTS reception fails. No matter whichever it is, if the Data frame instead of the RTS is sent, the ACK cannot be received by the sender. But the failure of Data transmission wastes more bandwidth than the RTS transmission failure. However, if the Data frame is small the RTS/CTS exchange is not necessary because it just wastes bandwidth. Because of this reason, the RTS/CTS exchange is optional in the DCF. The use of RTS/CTS exchange is under control of a threshold (RTSThreshold). If the data frame is larger than the RTSThreshold, the RTS/CTS exchange is initiated. Otherwise, Data and ACK are exchanged directly.

CHAPTER III

SPATIAL SPECTRAL UTILIZATION WITH DCF

In *MANETs*, the spatial area is a valuable resource in addition to the shared radio spectrum [35]. Delivering a data packet from one end-node to another consumes the precious spatial spectral resource in the proximity of the source and the destination as well as the intermediate forwarding nodes. Multi-hopping delivery contributes to improve the overall spectral utilization because a series of indirect communications requires less combined spatial footprint than a single direct communication. However, the corresponding benefit of supporting more concurrent data transfers is limited by collisions and interference. More concurrent communications would cause more collisions. There is a tradeoff between them.

The MAC protocol such as DCF is used to schedule or coordinate as many collision-free accesses to the shared medium as possible. In other words, it is essentially responsible for affecting the spatial spectrum utilization. This chapter will analyze how the spatial area is reused with DCF. Before presenting the analysis as an analytical basis, the signal propagation and reception model will be introduced first.

3.1 Signal Propagation in Wireless Environment

Radio propagation in mobile wireless channel is described by means of three effects: attenuation due to path loss, shadowing due to obstacles, and fading due to multiple paths. These three affects are described by path loss model, shadowing mode and fading model respectively. This section will introduce the first two in details.

3.1.1 Path Loss Model

In this model the path loss $L_p(d) \propto d^{\alpha}$, where d is the distance and α is path loss exponent. The path loss exponent varies in different environments. Table II [36] shows the variations.

TABLE II: PATH LOSS EXPONENT WITH ENVIRONMENT

Environment	**Path loss exponent,** α
Free space	2
Urban cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5
In building with LOS	1.6 to 1.8
Obstructed in building	4 to 6

Free Space Propagation Model and Two-Ray Ground Reflection Propagation Model are two common models describing path loss. Free space propagation is regarded as an ideal model, which assumes that the transmitter and the receiver are both located in free space and there is only one clear *Line-Of-Sight* (LOS) path between them. This model does not consider other sources of loss such as reflections, cable, etc. The receive power is not dependent on antenna heights. The signal is attenuated slowly ($\alpha = 2$). The following equation is used to estimate the receive power [37].

$$P_r(d) = P_t G_t G_r \left(\frac{\lambda}{4\pi d} \right)^2 \tag{3.1}$$

where $P_r(d)$ is the receive power with distance d, P_t is the transmit power, G_t and G_r are the gains of the transmitter and the receiver respectively, and λ is the wave length.

As the single LOS is rarely the only path between the transmitter and the receiver, the Two-Ray Ground Reflection Propagation Model considers both the direct LOS path and a ground reflection path. According to this model, the antenna heights are taken into account and the receive power is estimated using the following equation.

$$P_r(d) = P_t G_t G_r \frac{(h_t h_r)^2}{d^4} \tag{3.2}$$

where h_t and h_r are the antenna heights of the transmitter and the receiver respectively.

Notice that the Two-Ray Ground Reflection Propagation Model shows a quicker path loss than the Free Space Propagation Model as the distance increases. It gives more accurate results when the distance is long [38], but does not give good predictions when the distance is short because of the oscillation caused by the constructive and destructive combination of the two rays. When d is small the Free Space Propagation Model is used instead.

Therefore, there is a *cross-over distance* d_c. When $d \le d_c$, the Free Space Propagation Model is used. Otherwise, the Two-Ray Ground Reflection Propagation

Model is employed. Comparing Equations (3.1) and (3.2) the d_c is computed as the following equation.

$$d_c = (4\pi h_t h_r)/\lambda \tag{3.3}$$

According to the path loss model, given the distance d, the receive power is deterministic. On the other hand, based on the receive power it is possible to estimate the distance to the transmitter. Equations (3.4) and (3.5) show the distance estimation due to the Free Space Propagation Model and the Two-Ray Ground Reflection Propagation Model respectively.

$$d(P_r) = \frac{\lambda}{4\pi}\sqrt{\frac{P_t}{P_r}G_t G_r} \tag{3.4}$$

$$d(P_r) = \sqrt[4]{\frac{P_t}{P_r}G_t G_r (h_t h_r)^2} \tag{3.5}$$

Similar to the *cross-over distance* d_c, there is a *cross-over power* P_c, which is computed as Equation (3.6). When estimating distance d from receive power P_r, if $P_r \le P_c$, Equation (3.4) is used. Otherwise, Equation (3.5) is applied.

$$P_c = P_r(d_c) = \left(\frac{\lambda}{4\pi}\right)^4 \cdot \frac{P_t G_t G_r}{(h_t h_r)^2} \tag{3.6}$$

3.1.2 Shadowing Model

Either the Free Space Propagation Model or the Two-Ray Ground Reflection Propagation Model estimates the receive power as a deterministic function of the distance. But in reality the receive power at a given distance is random because of multi-path propagation effects. To describe this randomness the Shadowing Model is proposed [38].

According to the Shadowing Model, the receive power consists of two parts. The first one is known as path loss model, which predicts the mean received power at distance d, denoted by $\overline{P_r(d)}$. It uses a distance d_0 as a reference. $\overline{P_r(d)}$ is computed relative to $P_r(d_0)$ as follows.

$$\frac{P_r(d_0)}{\overline{P_r(d)}} = \left(\frac{d}{d_0}\right)^{\alpha} \tag{3.7}$$

where $P_r(d_0)$ is the relative receive power, which is estimated by using the Free Space Propagation Model or the Two-Ray Ground Reflection Propagation Model according to the referred distance d_0.

$$\left[\frac{\overline{P_r(d)}}{P_r(d_0)}\right]_{dB} = -10\alpha \log\left(\frac{d}{d_0}\right) \tag{3.8}$$

The second part of the Shadowing Model reflects the variation of the received power at the given distance. It is a log-normal random variable; that is, it is of Gaussian distribution if measured in dB. The overall shadowing model is represented by

$$\left[\frac{P_r(d)}{P_r(d_0)}\right]_{dB} = -10\alpha \log\left(\frac{d}{d_0}\right) + X_{dB} \tag{3.9}$$

where X_{dB} is a Gaussian random variable with zero mean and standard deviation σ_{dB}. σ_{dB} is called the shadowing deviation, and the above equation is also known as a log-normal shadowing model.

In this dissertation, to simplify the spectral utilization analysis the path loss model is employed for analysis. Both the path loss mode and the Shadowing model will be used in simulations for the evaluations and comparisons of the proposed solutions.

3.2 Signal Reception and Capture Effect

To successfully receive a transmission the following two conditions have to be satisfied. First, the receive power must be equal or larger than the receive sensitivity. Second, the receive power must be strong enough to overcome the influence of the noise and interference. This section will discuss them in detail.

3.2.1 Transmission Range and Carrier Sense Range

While a signal propagates in the wireless environment nodes at different locations may receive different levels of power. For analytical convenience suppose that the signal is propagated in an open area, thus the signal attenuation is only due to the path loss. The nodes that are at different locations but at the same distance to the transmitter receive the same power level.

According to the first condition of the signal reception model, to successfully decode a received signal the receive power must be equal or larger than the receive sensitivity, denoted as P_{tr} , which can be translated as a distance called *Transmission Range* (TR), $\text{TR} = d(P_{tr})$. As shown in the Figure 10, suppose that node i is sending a packet to node j. Only the nodes within the TR_i such as node A can decode the transmission. The other nodes like B and C cannot decode it. Given the system parameters like transmit power, gains and antenna heights, the size of TR is dependent on the receive sensitivity, which is determined by the transmit rate (modulation). Usually lower rate makes larger TR.

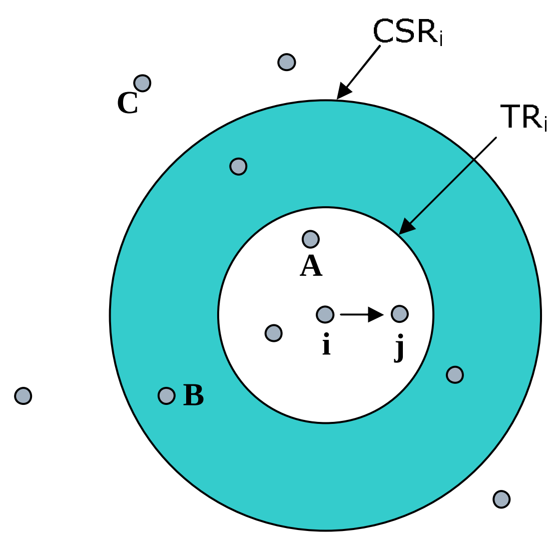


Figure 10: TR and CSR

The receive sensitivity is a power threshold that determines if an incoming signal can be successfully decoded or not. In the DCF the CS threshold, denoted as P_{cs} , is another power threshold that determines if a transmission can be sensed or not. This threshold can be translated as another distance called *Carrier Sense Range* (CSR),

$\text{CSR} = d(P_{cs})$. Also as shown in the Figure 10, all the nodes out of the CSR_i such as node C cannot sense the transmission. For the nodes that are within the shaded area (out of the TR_i but within the CSR_i) like node B can sense the transmission but cannot decode it. According to the PCS, all the nodes within the CSR_i, no matter whether they can decode the signal or not, are required to hold up their transmissions to protect node j's reception. Given the system parameters like transmit power, gains and antenna heights, the size of CSR is dependent on the CS threshold. The higher CS threshold makes a node less sensitive to ongoing transmissions, thus causing more collisions. On the other hand, the lower threshold can avoid more collisions. But it is at the cost of degrading spatial reuse.

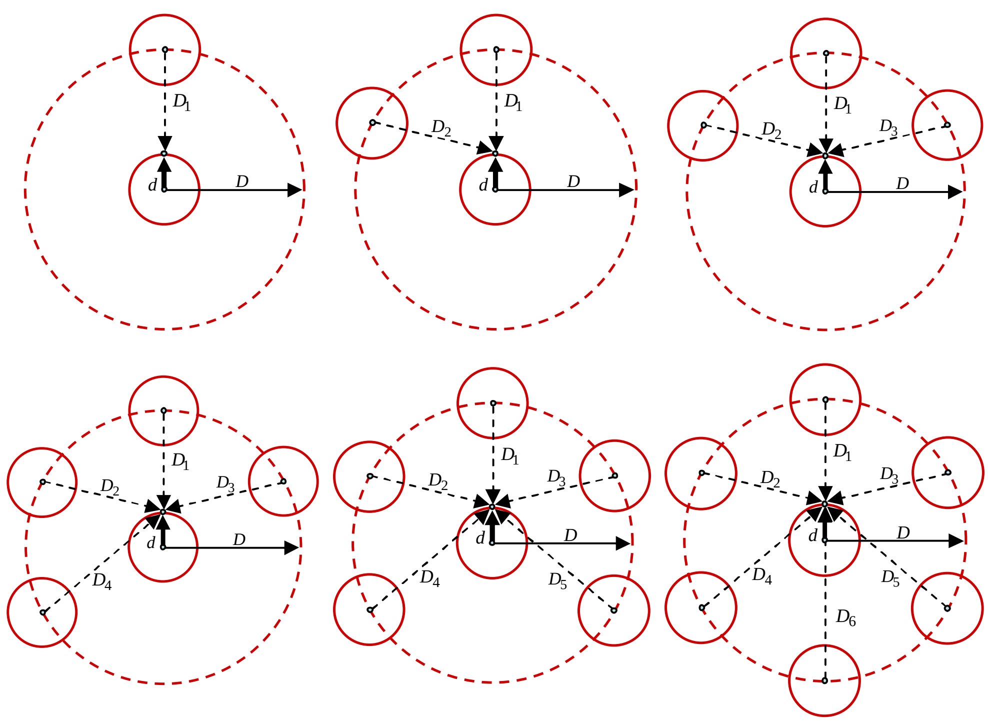
3.2.2 Capture Effect and Interference Range

If the receiver is within the TR of the transmitter, then it is supposed to successfully decode the transmitter's signal if there is no interference. But it may not if it is interfered by other transmissions. Whether the signal can be decoded or not is also dependent on if the receive power is strong enough to overcome the influence of the noise and interference. This condition is described by the following *Signal to Interference and Noise Ratio* (SINR) model.

$$\text{SINR} = \frac{P_r}{N_0 + \sum I} \geq Z_0 \tag{3.10}$$

where N_0 is the background noise, $\sum I$ is the interference from all other simultaneous transmissions, and Z_0 is the minimum required SINR ratio, or *capture ratio*. The SINR model suggests that even if more than one signal overlaps at the receiver, one of them could survive if it is much stronger than the others. This is called the *capture effect* [39].

From the Inequality (3.10), given a receive power P_r (communication distance) it is not difficult to find out the maximum tolerable interference. This maximum tolerable interference can be translated as a maximum distance from the receiver to an interfering node that could cause collision to the receiver. This maximum distance is called *Interference Range* (IR). Only the nodes within this range could cause collision to the receiver. In other words, for the nodes out of the IR, they may sense the transmission but they would not cause collision to the receiver. The size of IR is dependent on many factors such as communication distance (d) between the sender and the receiver, capture ratio (Z_0), as well as the number of interferers (k) and their locations.



$$(D_1 = D - d \ , \ D_2 = D_3 = \sqrt{D^2 + d^2 - Dd} \ , \ D_4 = D_5 = \sqrt{D^2 + d^2 + Dd} \ , \ D_6 = D + d \ )$$

Figure 11: Worst-Case Interference Scenarios with *k* (1~6) First Tier Interferers

To derive the IR, the worst-case interference scenarios with different number of the first tier interferers ($1 \le k \le 6$) are considered, as shown in the Figure 11. Only the first-tier interferers are considered [40] because their influence is dominant. Let D be the separating distance between the sender and an interferer and D_i be the distance between the receiver and interferer i.

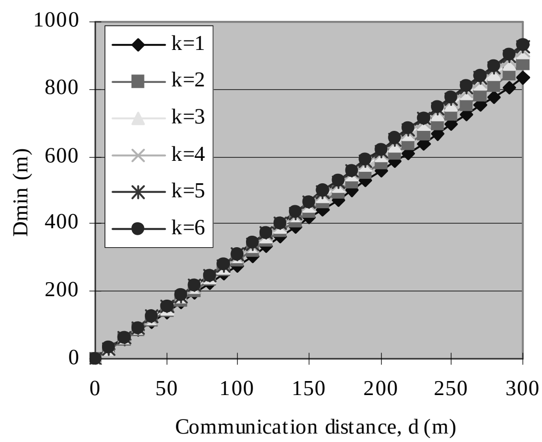


Figure 12: $D_{\min}$ versus d with varying number of interferers (k)

Now, given d, k, Z_0 and N_0, it is not difficult to find the minimum D (denote as $D_{\min}$) that satisfies the inequality (3.10). Figure 12 plots $D_{\min}$ versus the communication distance (d) with differing number of interferers (k) assuming that N_0 is ignorable. It is surprising that from the Figure 12 that d almost dominates the influence. For example, when d is 150m, the variation of $D_{\min}$ with different k is at most 7%. This small variation is because the signal attenuates very quickly with distance and thus the topmost interferer in each of the six figures in the Figure 11 (with the shortest distance to the receiver) dominates the interference. Therefore, it is quite reasonable to assume that $k = 1$.

And, $D_{\min} = \left(\sqrt[4]{Z_0} + 1\right) \cdot d$ and $\mathrm{IR} = D_{\min} - d = \sqrt[4]{Z_0} \cdot d$.

3.3 Dilemma of CSMA/CA

As introduced before, there is a tradeoff between the spatial reuse and collision. This tradeoff can be expressed by the hidden and exposed terminal problems. The hidden terminal problem argues that a node does not sense a transmission but could cause collision to it. On the contrary, the exposed terminal problem says that a node senses a transmission but would not cause collision to it even if the node initiates its own transmission.

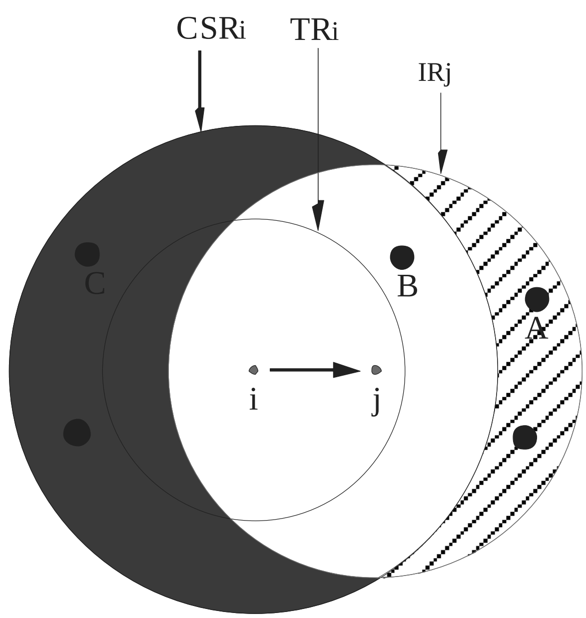


Figure 13: Hidden and Exposed Terminal Problems

Figure 13 illustrates these two problems. Suppose that node i is sending a packet to node j. TR_i and CSR_i are the TR and CSR of node i respectively. And IR_j is the IR of node j. Due to the PCS, nodes within CSR_i like B and C should keep silent while node i is sending. Because of capture effect only nodes within IR_j like A and B would cause collision to node j. Node B as an interferer is required to hold up its transmission as expected. But node A is a potential interferer but cannot sense node i's transmission, thus is a hidden terminal. Similarly, node C is an exposed terminal because it is not an interferer but does sense the transmission.

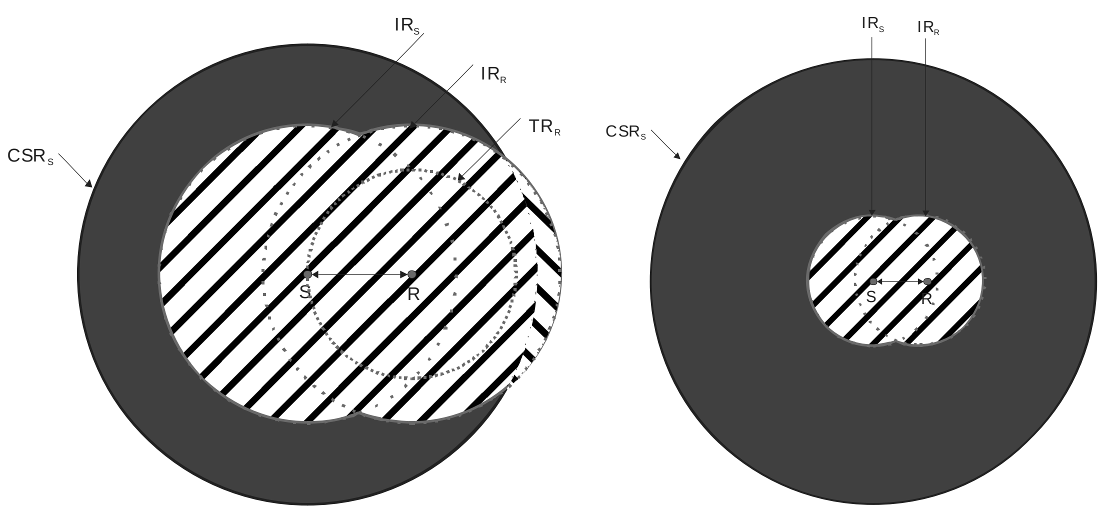
The area where the hidden terminals are located is called *Vulnerable Space* (VS), the hatched areas shown in the Figure 13. The place where the exposed terminals are positioned is called *Wasted Space* (WS), the shaded area in the figure. From the figure it is clear to see that the sizes of VS and WS are determined by CSR_i and IR_j. More specifically, VS = $IR_j - CSR_i$ and WS = $CSR_i - IR_j$. The dilemma of CSMA/CA based MAC mechanism is: In order to reduce collision (make VS smaller) it has to increase CSR. However, increasing CSR will result in larger WS and make the exposed terminal problem more serious. On the other hand, reducing CSR encourages more concurrent transmissions but causes serious collision problem.

Figure 13 shows a simple example of VS and WS. If the whole frame exchanges of DCF are considered the situation becomes much more complicated. The details will be presented in the next section.

3.4 Spatial Reservation in DCF

According to the frame exchange sequence described in the Subsection 2.3.5, before transmitting DATA frame the sender transmits RTS first. Then the receiver replies CTS. And then the sender starts DATA frame transmission. Finally the receiver replies ACK. This section discusses the spatial reservation for each stage of the communication procedure. To be convenient denote SR be the spatial reservation, in other words the reserved spatial area, and RSR be the required spatial reservation, so in general VS = RSR – SR and WS = SR – RSR.

Figure 14 shows the spatial reservation while RTS is transmitted. Since the sender just initiates the communication the reserved spatial area is only the CS range of the sender. In other words, SR = CSR_S. Actually while RTS is transmitted only the reception at the receiver needs protection. That means RSR = IR_R. But with consideration that the sender will receive CTS and ACK, IR_S should be reserved too. In other words, RSR = $IR_R \cup IR_S$. Therefore, VS = $CSR_S - (IR_R \cup IR_S)$ and WS = $(IR_R \cup IR_S) - CSR_S$.

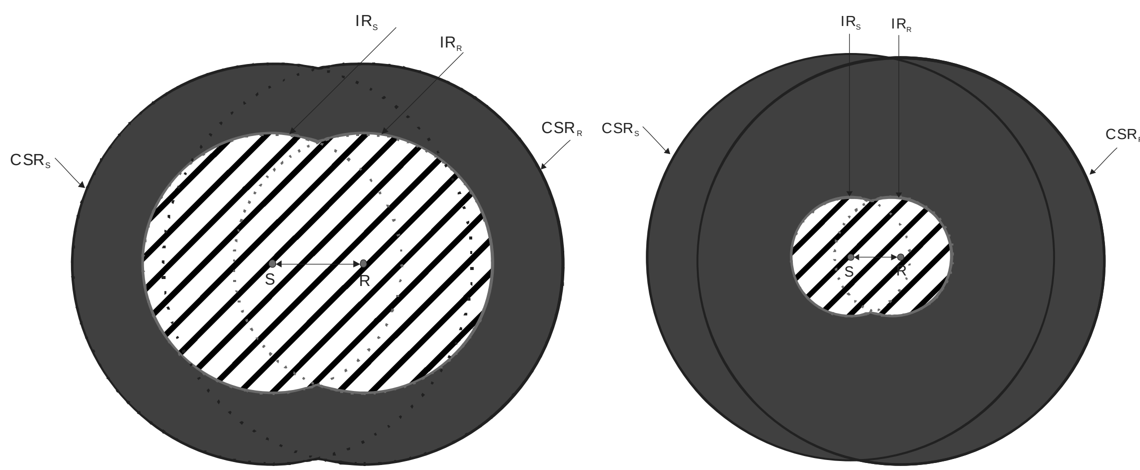


(a) With Long-Distance Link (b) With Short-Distance Link

Figure 14: Spatial Reservation while RTS or DATA is Transmitted

Figure 14(a) shows the spatial reservation while the sender and the receiver communicate with a long-distance link. In this case, VS is the right-hatched area shown in the figure. Because VS is not empty this transmission still suffers from the *Hidden-Terminal Problem*. Figure 14(b) shows the spatial reservation when the communication distance is short. The *Hidden-Terminal Problem* is eliminated, but the *Exposed-Terminal Problem* becomes more serious because of the large WS, the shaded area in the figure.

As introduced before, RTS/CTS exchange is optional in the DCF. If the sender transmits the DATA frame directly, then the spatial reservation is as same as that of transmitting RTS, shown in Figure 14.



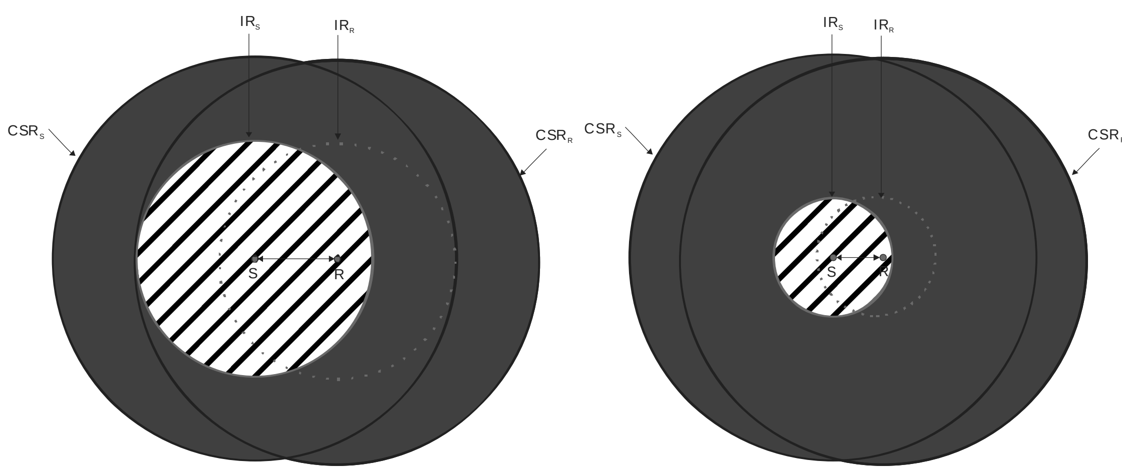
(a) With Long-Distance Link (b) With Short-Distance Link

Figure 15: Spatial Reservation while CTS or DATA is Transmitted

Figure 15 shows the spatial reservation while CTS is transmitted. Because of the VCS mechanism of DCF, the TR_S is reserved. Because EIFS is longer than the CTS transmission time, the nodes that are out of TR_S but within CSR_S also keep silent. So during CTS transmission $SR=CSR_S \cup CSR_R$. Thus, $VS = (CSR_S \cup CSR_R) - (IR_R \cup IR_S)$

and WS = ($IR_R \cup IR_S$) – ($CSR_S \cup CSR_R$). Since CSR_R is reserved, the VS is completely eliminated even with long-distance link. But it does not ensure that VS is empty while DATA is transmitted later because this reservation may not be able to cover the whole DATA transmission. In addition, the CSR_R makes the *Exposed-Terminal Problem* even worse when communication distance is short.

While DATA frame is transmitted, the spatial reservation is a little complicated. At first, if a small DATA frame is transmitted before the EIFS timers of the nodes within CSR_R expire the reserved spatial area is same as the case of replying CTS, as shown in the Figure 15. If a large DATA frame is transmitted after the EIFS timer expires the reserved spatial SR becomes $CSR_S \cup TR_R$. However, since CSR is usually two times as large as TR, $TR_R \subset CSR_S$. Thus, SR = CSR_S. In this case the spatial reservation is as same as the case of sending RTS, as shown in the Figure 14.



(a) With Long-Distance Link (b) With Short-Distance Link

Figure 16: Spatial Reservation while ACK is Transmitted

Finally, while ACK is replied, because the receiver will not receive frame from the sender during this communication procedure, it is not necessary to reserve IR_R anymore. In other words, RSR is reduced to IR_S. But the SR is still $CSR_S \cup CSR_R$. So WS = $(CSR_S \cup CSR_R) - IR_S$, which suggests more spatial area is wasted, as shown in Figure 16. Moreover, the spatial area of $CSR_R - TR_R$ will be reserved even after ACK transmission is finished because the nodes within that area do not know the transmission is done and still wait due to the EIFS mechanism.

The spatial reservations shown in Figures 14~16 expose the inefficiency of spatial reuse in DCF. The root reason lies in the CS mechanism and the fixed CS threshold of DCF. Is it possible to only reserve the spatial area that is required? Chapter V will introduce the proposed solution *Collision-Aware DCF* (CAD), which can achieve this goal while DATA and ACK frames are transmitted.

3.5 Analysis on Maximizing Network Throughput

This section gives an analysis on maximizing network end-to-end throughput and discusses how the throughput is affected by the CS threshold and communication distance.

3.5.1 Upper Bound of Network End-To-End Throughput

Maximizing network throughput in MANETs is equivalent to find the maximum number of *Collision-Free communication Pairs* (CFPs). In other words, the maximum total end-to-end throughput, T_e, is attained when the number of senders that can simultaneously transfer data is maximized. Multiplying this number by the wireless link bandwidth and then dividing by the average number of hops between the source and the destination will yield an estimate of T_e. In the following analysis, assume a heavily-loaded network in which each node is always backlogged and has a packet to transmit whenever it is allowed. Perfect MAC-layer coordination is assumed without collision so that spatial spectrum utilization is maximized as similarly assumed in [40].

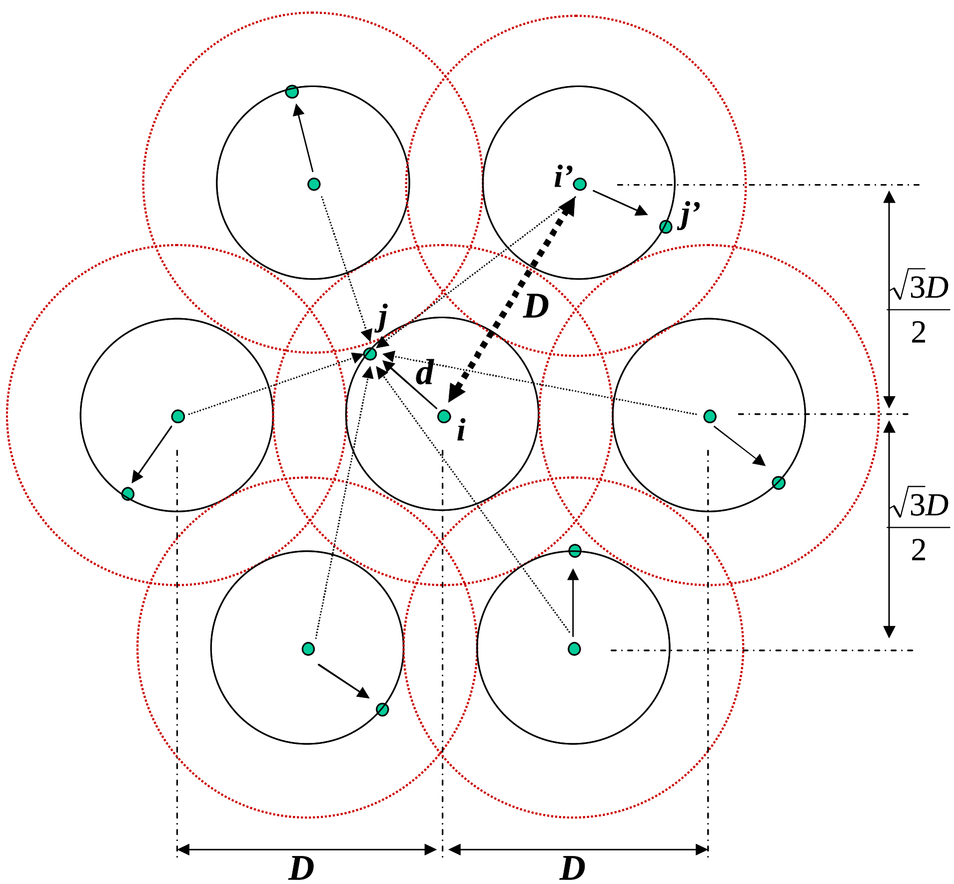


Figure 17: Constellation of Senders for Maximum Throughput

The number of senders can be maximized when they are located as close to each other as possible without interfering with each other's data transfer. This is similar to the co-channel interference problem in cellular networks [41]. Consider the constellation of senders as in Figure 17, which is the densest arrangement of senders. Assume that each communication distance is *d*, the purpose is to find the sender-to-sender distance *D* that allows all data transfers to be simultaneously successful. In this analysis only the six first-tier interferers are considered because the interference from them is much stronger than that from second-tier interferers and beyond. Now, the worst-case interference to the communication between nodes *i* and *j* happens when the six interferers are $D-d$, $\sqrt{D^2+d^2-Dd}$, $\sqrt{D^2+d^2-Dd}$, $\sqrt{D^2+d^2+Dd}$, $\sqrt{D^2+d^2+Dd}$, and $D+d$ apart to the receiver *j*, respectively.

Therefore, ignoring while noise N_0 and applying the signal propagation and reception models obtain the following inequality.

$$\mathrm{SINR}=\frac{P_r(d)}{P_r(D-d)+2P_r\left(\sqrt{D^2+d^2-Dd}\right)+2P_r\left(\sqrt{D^2+d^2+Dd}\right)+P_r(D+d)}\geq z_0 \quad (3.11)$$

If $D_{\min}$ is the minimum *D* that satisfies equation (3.11), the maximum number of concurrent successful data transfers in an $L\times L$ square network area is

$$\frac{L}{D_{\min}}\times\frac{L}{\sqrt{3}/2\,D_{\min}}=\frac{2L^2}{\sqrt{3}D_{\min}{}^2} \quad (3.12)$$

Since the average distance between a source-destination pair in the $L\times L$ square network is about $0.616L$ (refer to the Appendix A), the average hop count is $0.616L/d$. Therefore, T_e is

$$T_e = \frac{2L^2 b}{\sqrt{3}D_{\min}{}^2} \div \frac{0.616L}{d} \approx \frac{1.875Lbd}{D_{\min}{}^2} \qquad \text{when } d_{CS} < D_{\min} \tag{3.13}$$

where d_{CS} is the CS range and b is the bandwidth. Note that when $d_{CS} \geq D_{\min}$, senders would be separated by d_{CS} instead of $D_{\min}$ due to aggressive carrier sensing, and thus, equation (3.13) becomes

$$T_e = \frac{1.875Lbd}{d_{CS}{}^2} \qquad \text{when } d_{CS} \geq D_{\min} \tag{3.14}$$

3.5.2 Effect of Carrier Sense Range and Communication Distance

The previous subsection obtains the analytical form of overall network throughput. This subsection will further analyze how communication distance and carrier sense range affect the network throughput.

Equation (3.13) becomes clearer if simply assume that the six interferers are all D apart from the receiver j. Then, Inequality (3.11) becomes

$$\frac{P_r(d)}{6P_r(D_{\min})} = z_0 \tag{3.15}$$

Suppose that α is the path loss exponent. Equations (3.15) and (3.13) become

$$\frac{d^{-\alpha}}{6D_{\min}{}^{-\alpha}} = z_0 \Rightarrow D_{\min} = \sqrt[\alpha]{6z_0}\,d \tag{3.16}$$

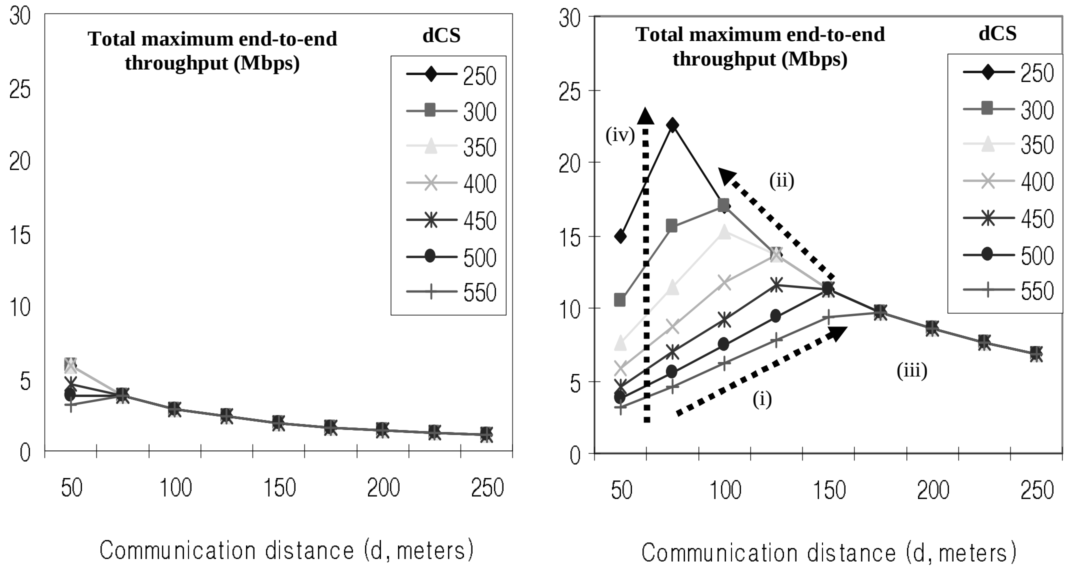
and

$$T_e = \frac{1.875Lbd}{\left(\sqrt[\alpha]{6z_0}\,d\right)^2} = \frac{1.875Lb}{\left(\sqrt[\alpha]{6z_0}\right)^2} \times \frac{1}{d} \qquad \text{when } d_{CS} < D_{\min} \tag{3.17}$$

In other words, T_e increases as the communication distance d decreases as predicted in [42]. When $d_{CS} \geq D_{\min}$, T_e is estimated as Equation (3.14) meaning that T_e increases as the communication distance d increases. In this case the d_{CS} makes spatial reservation more than necessary. In other words, the spatial reuse is not efficient. The reservation has room to make longer communication distance successful. The longer communication distance can make more progress for data delivery and thus improves the overall network throughput.

Therefore, given the communication distance, the optimal carrier sense range should be $D_{\min}$. The reason is straightforward. If $d_{CS} > D_{\min}$ the spatial reuse is not efficient; If $d_{CS} < D_{\min}$ the spatial reservation is not large enough to prevent the communication from collision.

Figure 16 shows T_e versus d for different d_{CS} values based on equations (3.11), (3.13) and (3.14). When the path loss exponent is 2, the communication signal travels farther and causes stronger interference to other communications and much smaller T_e as in Figure 18(a). However, when it is 4 as in a land mobile environment, the effect becomes significant as in Figure 18(b). From the d_{CS}'s point of view, when d_{CS} is large enough, it is better to exploit the CS-protected area and deliver data packets as far as possible within the CS range (large communication distance d). See mark (i) in Figure 16(b). When d_{CS} is not large, a better performance can be obtained by shortening the communication distance even though it increases the hop count between the source and the destination (mark (ii)). From the communication distance's perspective, when short communications are frequent, the $D_{\min}$ required is smaller and Equation (3.14) applies. T_e increases as d_{CS} decreases (or less sensitive carrier sensing) as indicated (iv) in the figure.



(a) α = 2.0 (b) α = 4.0

(b=1Mbps, L=10km, z_0=10dB)

Figure 18: Maximum Total End-To-End Throughput

CHAPTER IV

IMPROVING SPATIAL REUSE VIA PACKET SALVAGE

One important observation in Section 3.5 is that the performance of a multi-hop network greatly depends on the CS threshold and communication distance. This chapter proposes the *Multiple Access with Salvation Army* (MASA) protocol that uses a fixed, higher CS threshold (smaller d_{CS}) to increase the spatial reusability and solves the collision problem from hidden terminals via packet salvaging. It essentially reduces the communication distance on-the-fly by breaking one hop communication into two shorter-hop communications when it is beneficial.

This chapter is organized into three sections. In Section 4.1, some related work including the existing packet-salvaging schemes at the network and MAC layers, are discussed. In Section 4.2, the details of the proposed packet-salvaging MAC algorithm, MASA, are presented. Finally, in Section 4.3, the results of extensive simulation based on ns-2 [43], which has been conducted to evaluate various performances using metrics such as packet delay, packet delivery ratio, routing control overhead, and packet queuing requirement, will be reported.

4.1 Packet Salvage

This section provides an overview of the existing packet-salvaging schemes at the network and MAC layer in Subsections 4.1.1 and 4.1.2, respectively.

4.1.1 Routing-Layer Packet Salvage

For a collided packet, one possible salvaging solution at the network layer is to relay it via an alternative path in order to avoid the congested area and to exploit unused area. This may improve the performance significantly because a link breakage, even though it is temporary, could cause serious performance degradation if it is misinterpreted as a permanent link error. A number of packets already in flight could be lost and a routing protocol, *e.g.*, DSR [44], would initiate a new route-discovery procedure that basically floods the network with control messages, making the situation worse or the problem more likely to persist.

In the DSR, an optimization technique known as “packet salvaging” [44] is used so that the node encountering the forwarding failure may search its local storage for alternative routes. If a route is found, it is used to forward the undeliverable packets without resorting to an expensive route-discovery procedure. The “local repair” mechanism in the AODV routing protocol [45] Uses a similar technique. Valera *et al.* suggested a distributed packet salvaging scheme for more improvement [46]: every node maintains a small buffer for caching data packets that pass through it and at least two routes to every active destination. When a downstream node encounters a forwarding error, an upstream node with an alternative route as well as the pertinent data in its buffer can be used to retransmit the data packets.

However, the above-mentioned packet-salvaging schemes do not keep the sender from initiating an expensive route-discovery procedure because the original goal of these schemes is to save packets in flight. Moreover, these schemes are triggered only after a lower-level protocol has attempted a number of times without a success. For example, the DCF [24] retransmits four times before the link error is reported to the higher-level protocol. Each retransmission not only wastes resources such as node energy and channel resource but also extends the packet delay. Shortest-path routing protocols aggravate the situation because they prefer a longer per-hop communication distance, and the corresponding wireless links are more prone to temporary breakages [47].

4.1.2 MAC-Layer Packet Salvage

Non-deterministic packet salvaging at the MAC layer has recently received significant attention to deal with frequent, temporary link errors quickly and efficiently [48-52]. It is more direct and efficient than routing layer packet salvage because each hop connection is established for communication at the link layer. This subsection overviews four MAC-layer packet salvaging schemes.

- Biswas and Morris proposed *Extremely Opportunistic Routin*g (ExOR), which defers the choice of the next-hop node among the pre-computed candidates until after the previous node has transmitted the packet via its radio interface [48]. Based on the number of hops to the final destination and the past history of delivery ratios, the sender prioritizes the candidates and includes the list in the packet header. Each candidate competes to become a receiver by delaying its reply for the amount of time determined by its priority in the list.

- Blum *et al.* proposed *Implicit Geographic Forwarding* (IGF) which is also a non-deterministic algorithm [49]. As in *Geographic Forwarding* (GF) [53], the sender has position information of its neighbors as well as the final destination node of its packet. However, unlike in GF, the choice of the next-hop node is not determined by the sender but by competition among the candidates as in the ExOR scheme. The sender transmits an *Open RTS* (no intended receiver is specified) and each candidate delays its response (Clear-to-Send or CTS) for an amount of time determined by the distance to the destination and the remaining node energy.

- Zorzi and Rao presented *Geographic Random Forwarding* (*GeRaF*), which is basically the same as IGF but the competition is coordinated by the sender with two control messages, called CONTINUE and COLLISION, in addition to RTS and CTS messages [52]. In GeRaF, the transmission coverage area of a sender, only in the direction of the final destination, is divided into a number of regions. When a sender transmits an RTS, any node in the closest region to the destination responds with a CTS. When no CTS is heard, the sender transmits a CONTINUE message so that the nodes in the next region can respond. When more than one CTS are sent, the sender hears a signal but is unable to detect a meaningful message. In this case, the sender transmits a COLLISION message, which will trigger a collision-resolution algorithm [52].

- In the *Stateless Non-deterministic Geographic Forwarding* (SNGF) algorithm, which is part of the sensor network protocol *SPEED* [50], each node computes the forwarding candidate set for each destination, a member node of which is a neighbor and is closer to the destination than the node itself. Location information of the node as well as the destination is necessary in SNGF.

The above-mentioned schemes depend either on location information [49, 50, 52] or use a link-state flooding scheme [48] to help determine the salvager among multiple candidates, which may not be feasible in real implementations. The MASA algorithm presented in this dissertation is a practical non-deterministic MAC algorithm that requires neither the location information nor the link state propagation. Note that MAC-layer packet salvaging targets temporary link breakages assuming that the current routing path is still usable while network-layer packet salvaging attempts to save packets in transit (and initiates a new route discovery as in conventional routing algorithm) assuming that the routing path is no longer usable. If a communication attempt fails due to a short-lived temporary problem, a new route discovery is not necessary at all, thus favoring MAC-layer salvaging. However, if a communication attempt fails due to a permanent problem such as node mobility, MAC-layer salvaging may be able to save the current packet but not the next one because the receiver moves farther away from the sender. Network-layer salvaging is invoked by saving the packet at hand as well as those in transit along the routing path. In other words, they play roles in different areas and improve the packet-delivery capability synergistically if both of them are employed.

4.2 Multiple Access with Salvaging Army (MASA)

This section presents details of the proposed packet-salvaging MAC algorithm, *Multiple Access with Salvation Army* (MASA). The issues such as how to salvage the collided packets and how to elect the salvaging node will be addressed in this section.

4.2.1 Overview of MASA

In order to increase spatial reuse MASA adopts a higher CS threshold. In order to mitigate the interference problem caused by the higher CS threshold MASA adjusts the communication distance on-the-fly by salvaging packets at the MAC layer. A key idea of MASA is that even if an intended receiver could not receive a data packet due to interference, a third party node among those in between the sender and the receiver, called the *salvation army*, "captures" or "salvages" the packet and then forwards the salvaged packet to the receiver.

MASA has several significant benefits. First, since the salvaging node, referred to as the salvager, is between the sender and the receiver, it is obvious that MASA's salvage makes delivery progress to the receiver. This salvage also makes the subsequent packet forwarding from the salvager more robust against interference than the retransmission from the original sender. Secondly, MASA tries to exploit the benefits of both shortest path and short-distance communication. When no collision occurs MASA delivers the packet by using the shortest path obtained by the routing algorithm. If collision occurs MASA breaks the shortest-path link into two shorter-distance communications in order to salvage the collided packet. Finally, MASA can reduce the false alarm for live link. If MASA does not salvage the packet, the sender will retransmit. Because of the interference to the receiver, the retransmission could fail again. After the failure of several attempts of retransmission, it will report to the upper layer that the link is broken. However, that link is still alive. Since MASA reduces the false alarm it prevents the traffic source from broadcasting routing requests to cause unnecessary routing overhead. The salvage in MASA is also helpful for offloading the sender's pending packets, and thus for reducing the packet queue size and the packet's waiting time in the queue. These benefits will be verified in the Section 4.3 through simulation.

4.2.2 Packet Salvage in MASA

MASA is based on the DCF but does not use the optional RTS/CTS exchange because collisions in the absence of RTS/CTS can also be effectively masked by packet salvaging. The MASA algorithm includes two new frame types, called *Salvaging ACK* (SACK) and *Salvaging DATA* (SDATA) as will be explained later in this subsection.

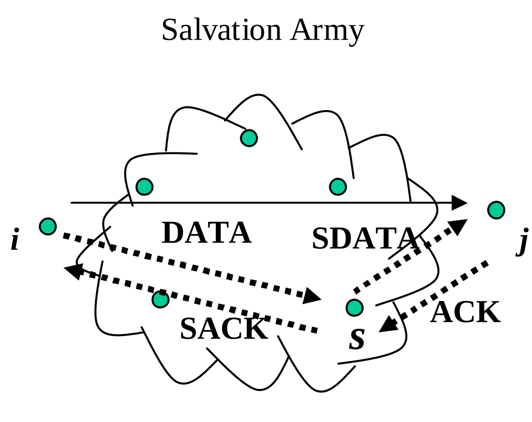


Figure 19: Salvation Army

In wireless networks, nodes use broadcast as opposed to point-to-point communication, and therefore, data packets are typically sent to multiple nodes in the proximity of the sender at no extra cost. The set of those overhearing nodes are called the *Salvation Army*. A key idea in the proposed MASA protocol is that a third party node (say, node *s*) in the *Salvation Army* captures or salvages a data packet that has collided at the intended receiver and then lets the packet make progress toward the receiver. This procedure is shown in the Figure 19. Since the sender-salvager distance is smaller than the sender-receiver distance, there is a higher probability that the salvager *s* will receive the packet successfully with and then completes the communication session by replying SACK to node *i*. The salvager *s* then forwards the data packet (SDATA) to the original receiver *j* based on the usual defer and backoff procedure. Note that while ACK is transmitted regardless of the status of the medium, SACK is transmitted only when the medium is free. This occurs in order to address the potential collision problem. The modified MAC behaviors at the salvager (*s*), the sender (*i*) and the receiver (*j*) are described below.

As shown in the Figure 20, first, at the sender (*i*), when an ACK is not received during *ACKTimeout* interval, the sender concludes that the transmission has failed and invokes its backoff procedure to re-transmit the packet. In MASA, the sender cancels the backoff procedure when it receives SACK, even after the ACKTimeout interval. Second, at the salvager (*s*), it waits for an *SIFS* upon successful reception of a data packet and checks the channel status (BUSY or IDLE), using the PCS supported by the IEEE 802.11-conformant hardware [24]. This determines whether it is necessary to salvage the packet or not. If ACK is received (more accurately, if the channel status changes to

BUSY), it cancels its salvaging activity. Otherwise, it starts its *salvaging backoff procedure* (to be explained shortly), and accordingly transmits SACK to the sender. Then, it starts its normal backoff procedure to forward the data packet (SDATA) to the receiver (*j*) who then replies with an ACK to the salvager after an SIFS period. Both the sender and the salvager would retransmit the same packet a pre-specified number of times as defined in the DCF if they do not receive ACK or SACK. Note that MASA does not allow a salvaged packet to be salvaged again. This is because consecutive salvages of a packet make it travel along a longer detour path, thereby potentially losing the benefit of MASA. Third, at the receiver (*j*), it may receive the same data packet more than once from more than one salvager. How this problem (duplicate reception) is handled in MASA will be explained in the next Subsection.

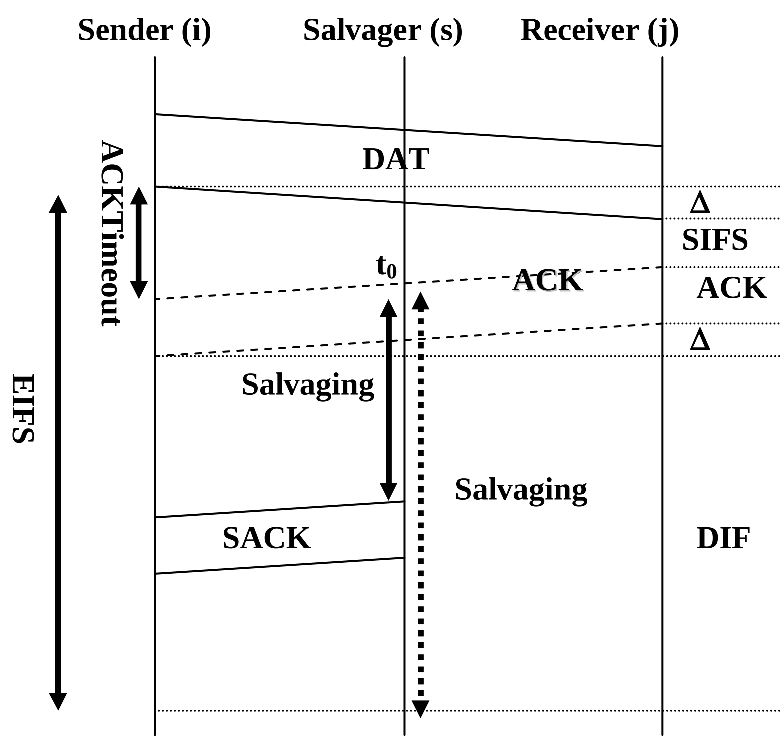
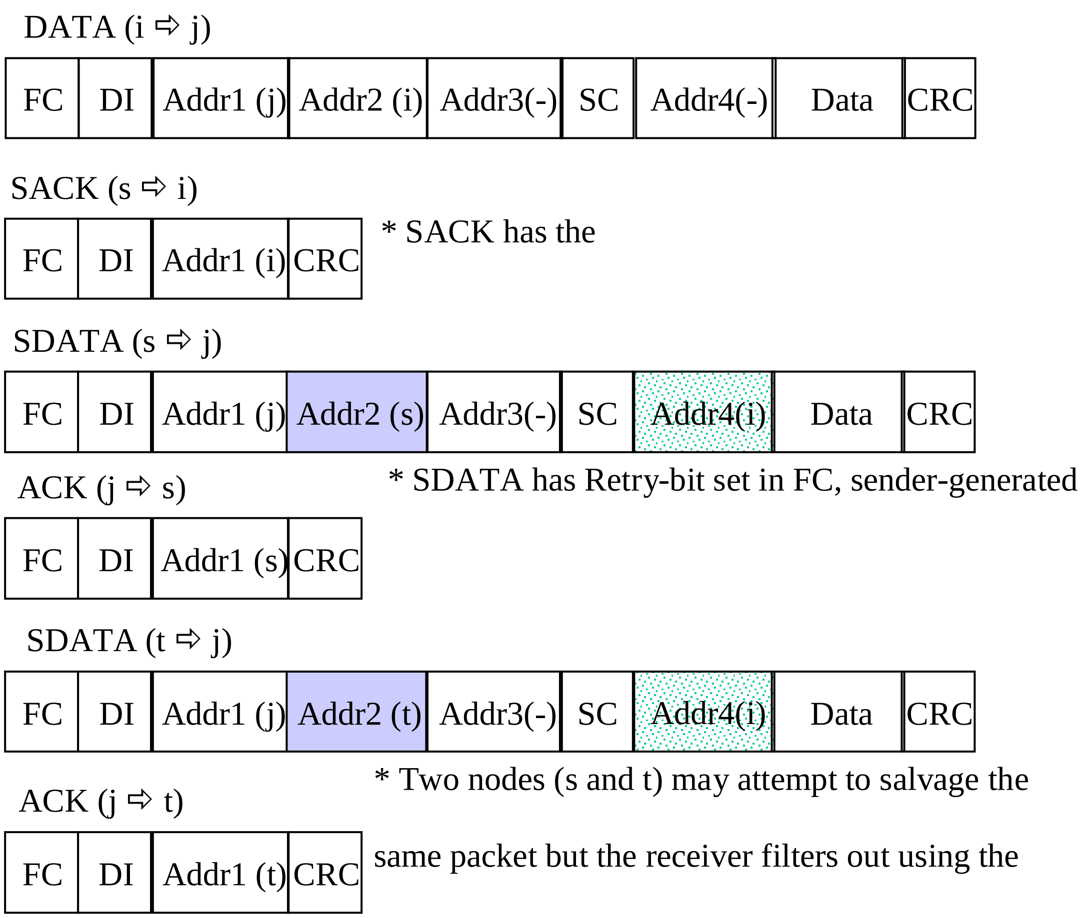


Figure 20: Salvaging Procedure

4.2.3 Suppression of Duplicate Salvage



(FC: Frame control, DI: Duration/ Connection ID, SC: Sequence control)

Figure 21: Format of MPDU Frames in the MASA Protocol

It is possible that more than one node salvages the same packet and the receiver receives the same packet more than once. Such duplicate packets can be filtered out within the receiver MAC based on the original functionality of the DCF, called *Duplicate Packet Filtering* [24]. This algorithm matches the sender address (*Addr2* in Figure 21) and the sender-generated *Sequence Control Number* (*SCN*) of a new packet against those of previously-received ones. If there is a match, the receiver transmits ACK but does not forward the packets. This does not solve the above-mentioned problem in MASA because duplicate packets from different salvagers (*s* and *t*) include different identities than node *i* in Addr2 field. The approach in MASA is to use a new data type, SDATA, which includes the original sender's address in *Addr4* (*logical address field*) so that the receiver can use this address rather than the salvager address (Addr2) when it compares against the stored information.

4.2.4 Determination of a Salvager among Salvaging Army

When more than one node is able to salvage a packet in collision, the candidate that can make the greatest progress should be selected. For this purpose, assume that each node maintains neighbor list and signal quality information for its neighbors. It is not difficult to keep track of the node's neighbors because each node overhears every other neighbor's communications. The signal quality for each neighbor can be obtained using the previous signal it received from the particular neighbor. The functionality of PHY layer of IEEE 802.11 is modified to support this. PHY layer of the IEEE 802.11 checks the *Received Signal Strength Index* (RSSI) of the signal to inform the channel status to the MAC layer (CCA signal) [24]. In MASA, the PHY layer is assumed to inform not only the channel status, but also the RSSI information to the MAC. When a sender transmits a MAC frame, assume that the frame includes the signal quality information for the receiver.

When node *s* receives a data packet that is not intended for it, the node evaluates its eligibility as a salvager using the following five rules. (i) The specified sender as well as the receiver must be in the neighbor list of node *s*. (ii) When node *s* overhears a SACK for the packet it is about to salvage, it should cancel its salvaging activity for that particular packet. (iii) In order for node *s* to make progress toward the receiver, it must be nearer the receiver than the sender. Node *s* speculates this condition based on signal strength information as mentioned earlier. (iv) Node *s* must not have a pending packet at its MAC-layer software. (v) Node *s* does not have a recent history that it failed to forward a packet after salvaging for the same pair of nodes. The neighbor list mentioned earlier can be used for this purpose as well. This is important because a node that is ignorant of a broken link might consecutively fail to forward packets but continue to salvage them.

If a node is considered a legitimate candidate, it starts its salvaging activity at time t_0 after waiting for an ACKTimeout interval as shown in Figure 18. Then, it chooses its *salvaging backoff time* (t_S) within the *salvaging interval* (T_{SI}), during which it is allowed to salvage the packet.

- T_{SI} can be considered the opportunity window open to salvagers, which begins at t_0 and must end before the next data transfer begins. Based on the operation principle of the DCF, T_{SI} = *ACK transmission time* + *DIFS* as shown in Figure 20. This is because nodes in the proximity of the communication between nodes *i* and *j* would wait for ACKTimeout in order to allow the pair to complete their communication. An additional DIFS is available because it is required for a new data transfer to start. Nodes outside of TR_i may corrupt the salvaging activity by transmitting their own packets during salvaging.

However, based on the DCF specification, they would wait *EIFS* (*Extended IFS*) before starting their own transmission [24], which turns out to be the same opportunity window to salvagers because EIFS is set to *SIFS + ACK transmission time + DIFS*. For simplicity, the propagation delay, Δ, is not included because it is relatively small and can be ignored.

- t_S is considered a priority among multiple candidates. The node that is closer to the receiver should be elected as the salvager because it can make greater progress. The proposed MASA uses the signal quality to determine the salvager. In other words, node *s* calculates t_S, using both the signal quality from the sender (q_{is}) and from the receiver (q_{js}), *i.e.*, $t_s = q_{is} / q_{ji} \times T_{SI}$. This is based on the assumption that the signal quality directly corresponds to distance. Even if the assumption is not valid, this arbitration rule still works well and it simply becomes a randomized algorithm.

Figure 22 summarizes the proposed MASA algorithm.

```
// MASA (Multiple Access with Salvation Army) at node s

// when it receives a frame
Upon receiving a frame from node i (Addr2) to node j (Addr1)
    if (s == j)  // node s is an intended receiver
    {
        transmit (ACK);
        if ((SC, Addr2)-pair ∈ pairs of recently received packets) return;
                                                   // duplicate reception of a frame (DATA)
        if ((SC, Addr4)-pair ∈ pairs of recently received packets) return;
                                                   // duplicate reception of a frame (SDATA)
        send up to network layer (frame);          // let it forward to the next intermediate node
    }

    else  // node s is potentially a salvager
    {
        if (Addr4 != EMPTY) return;                // do not salvage a salvaged packet
        if (i or j ∉ neighbor set of s) return;    // eligibility test
        if (pending packet in the queue) return;
        if (failed to forward for the same node pair (i, j) recently) return;
        if (ACK received from j during ACKTimeout) return;
                                                   // do not salvage if  the receiver gets it successfully
        // communication was not successful; salvaging activity starts
        t_s = random(0, T_SI);                     // or t_S = q_is/q_ji × T_SI , determine t_s within T_SI
        if (SACK received during t_s) return;      // do not salvage if it is salvaged by another node
        transmit (SACK) to i;                      // now, it's time to salvage
        enqueue (frame);                           // put into the packet queue
    }

// when packet queue is not empty
Upon being ready to transmit a frame
    dequeue (frame);                               // retrieve from the packet queue
    Addr1 = j; Addr2 = s;                          // receiver & sender
    if (the frame is a salvaged one for node pair (i, j)) Addr4 = i;
    else Addr4 = EMPTY;                            // for duplicate packet filtering
    transmit (frame);
```

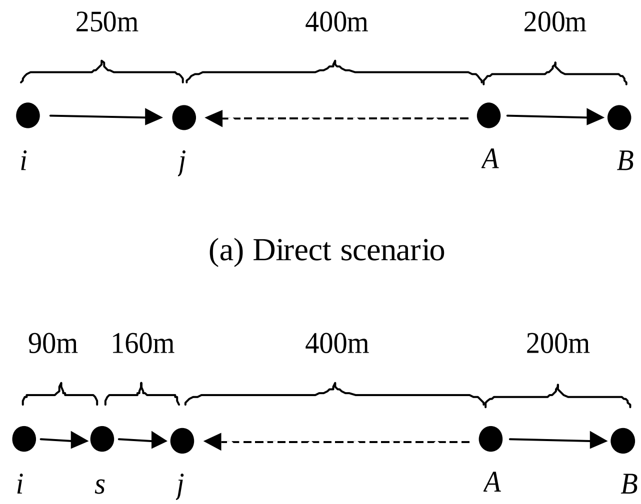
Figure 22: MASA Algorithm

4.3 Simulation and Evaluation

The performance of the MASA algorithm is evaluated using the *ns-2* [43], which simulates node mobility, radio network interfaces, and the DCF protocol. The *Two-Ray Ground Propagation Model* is assumed with a radio transmission range of 250 m and a data rate of 2 Mbps. In order to show the benefits of the packet salvaging, Subsection 4.3.1 presents the simulation result of a simple 4- and 5- node scenario with a single interferer. More realistic scenarios with more than 50 nodes and the corresponding simulation results are presented in Subsections 4.3.2 and 4.3.3, respectively.

4.3.1 Benefit of Packet Salvage with a Single Interferer

Figure 23 shows a simple communication scenario with 4 and 5 nodes. Node pair *i*-*j* is the primary focus while node pair *A*-*B* provides interfering signals. Node *i* sends 512-byte *Constant Bit Rate* (CBR) or TCP packets to node *j*. Node *A* also sends 512-byte CBR or TCP packets to node *B*. In the Direct scenario in Figure 23(a), there exists no salvager candidate between nodes *i* and *j*; thus SINR at node *j* is always low and the communication is easily subjective to interference from node *A*. On the other hand, in the Salvaging scenario in Figure 23(b), node *s* is capable of capturing and salvaging a collided packet at node *j*; thus node *j* receives a stronger signal with high SINR from node *s*. SINR at node *j* in the Direct scenario is $(400/250)^4$ or 8.16 dB for the packet from node *i*, which is smaller than Z_0. In the Salvaging scenario, however, it is $(400/160)^4$ or 15.92 dB for the packet that has been salvaged by *s*, which is larger than Z_0.

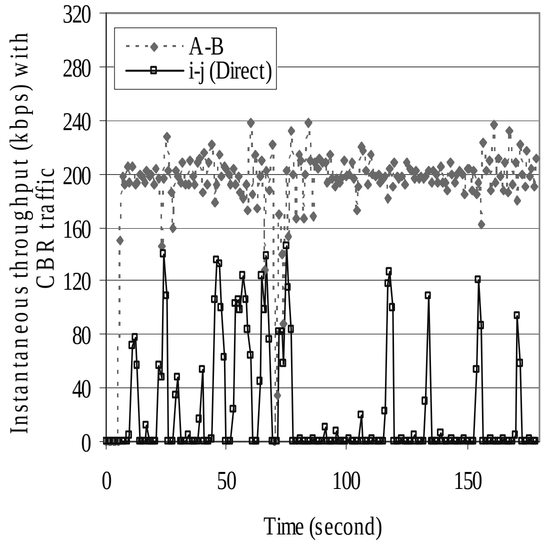


(a) Direct scenario

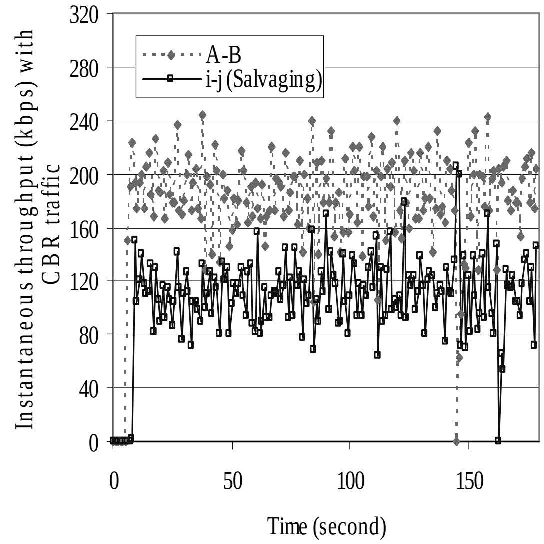
(b) Salvaging scenario

Figure 23: Simple Salvaging Communication Scenarios

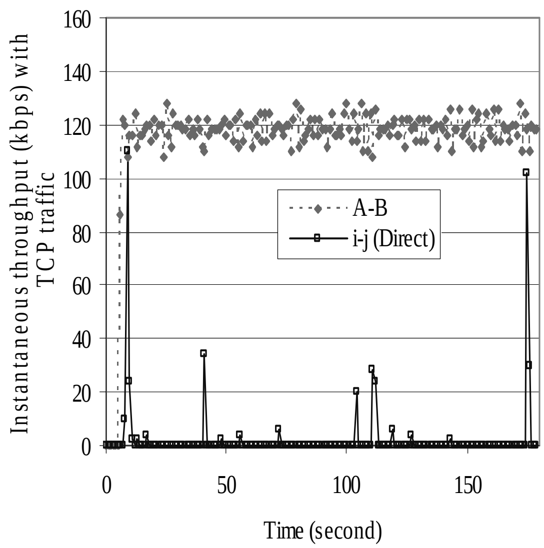
Figure 24 compares instantaneous throughput, measured at every simulated second, with CBR and TCP traffic. As shown in the Figures. 24(a) and (b), the Salvaging scenario offers a higher aggregate throughput than the Direct scenario with CBR traffic even though the average number of hops between the communication pair (*i*-*j*) is larger. This is also true with TCP traffic as drawn in the Figures 24(c) and (d). Moreover, the Direct scenario exhibits unacceptably serious unfairness, which is a well-researched phenomenon observed by Xu *et al*. [54]. According to their observation, the throughput of one TCP session can be almost zero while the other TCP session monopolizes the channel bandwidth. The simulation results confirm that this is also the case with CBR traffic and infer that the capture effect and packet salvaging may alleviate the unfairness as well as the performance problem.



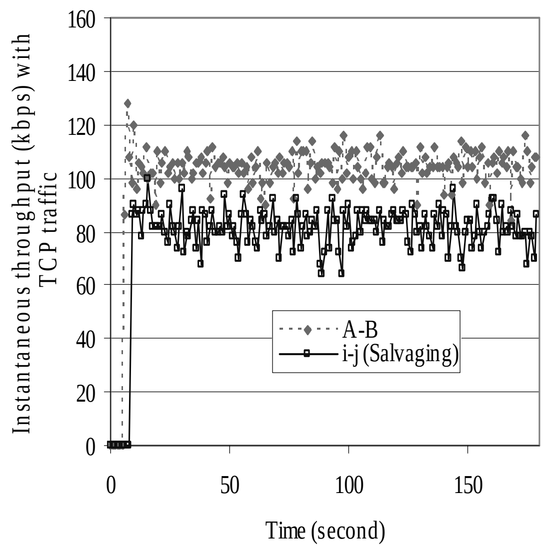
(a) Direct with CBR



(b) Salvaging with CBR



(c) Direct with TCP



(d) Salvaging with TCP

Figure 24: Effect of Packet Salvaging with Simple Communication Scenarios

4.3.2 Simulation Environment with Multiple Interferers

The previous subsection shows the benefit of packet salvaging in MASA on a small network with a single interferer. The following two subsections present the merits of the proposed MASA algorithm in more complex and larger network scenarios. Protocols to be compared are MASA, DCF2 (DCF without RTS/CTS) and DCF4 (DCF with RTS/CTS). DCF2 is included because MASA does not incorporate the RTS/CTS handshake, either. Note that, in general, DCF2 outperforms DCF4. This is counter-intuitive but has been predicted by a number of researchers [55] and has also been observed in this dissertation. It is observed, however, that DCF2 degrades more significantly in comparison to DCF4 with the *shadowing radio-propagation model*. Randomness in radio propagation makes the RTS/CTS handshake more useful. This issue will be discussed later in the next subsection.

The performance evaluation is based on the simulation of 100 mobile nodes located in an area of 300×1500 m^2. The CS distance is assumed to be 550m and 350m with the DCF and the MASA, respectively. AODV routing algorithm [45] is used to find and maintain the routes between two end-nodes. The data traffic used in the simulation is CBR and TCP traffic. In case of CBR, 40 sources generate three 256-byte data packets every second. Destination nodes are selected randomly. The *Random Waypoint Mobility Model* is used in the experiments with the maximum node speed of 5 m/s and the pause time of 100 seconds. The simulation is run for 900 seconds and each simulation scenario is repeated ten times to obtain steady-state performance metrics. For more accurate performance evaluation, different routing algorithms (DSR [44]), and different propagation models are used. Various traffic intensities in terms of packet rate and the number of sources and various numbers of nodes are also used to observe the performance scalability of the DCF and the MASA.

In the experiments, the following aspects of signal capture are assumed.

- When two packets arrive, if the first signal is 10dB (Z_0) stronger than the second, then the first signal can be successfully received. However, if the second signal is 10dB stronger than the first, neither packet is successful because the receiving node has already started decoding the first signal and cannot switch to the second immediately. This is, in fact, the way that the ns-2 is implemented. However, in the latter case, if the first signal is weaker than the receive threshold but larger than the CS threshold, the receiver can receive the second signal successfully. Since ns-2 still drops both packets in this case, it is modified to reflect this fact.
- The SINR computation requires two samples of the signal, the desired signal and the signal with interference, and their availability is assumed for computation.

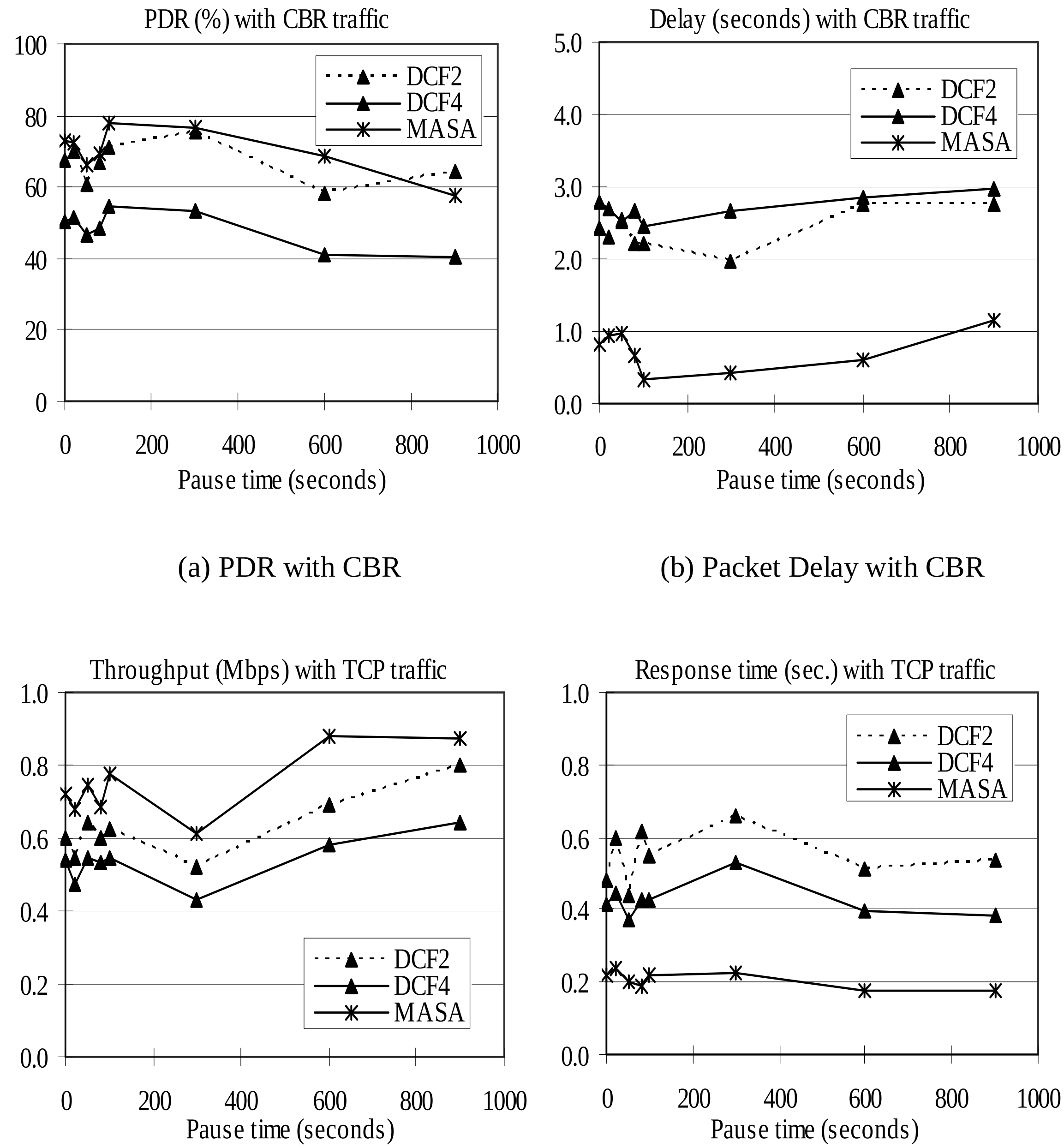
The signal strength comparison for determining capturing is on a per-packet basis in ns-2. That is, if multiple interfering packets were to be received, they are only compared individually, not their combinations. Ns-2 is modified to simulate *additive interference* if concurrent multiple interfering signals exist.

4.3.3 Results and Discussion

General Network Performance

Figure 25 shows the network performance with respect to node mobility represented by pause time. Note that 900 seconds of pause time means a static scenario while 0 seconds mean a constant-moving scenario. Figures 25(a) and (b) show the *Packet Delivery Ratio* (PDR) and packet delay with CBR traffic. While the PDR of MASA is equal to that of DCF2 as shown in the Figure 25(a), it is clear from the Figure 25(b) that MASA outperforms DCF2 and DCF4 in terms of packet delay, showing a 53~85% and a 59~86% reduction, respectively. A major factor that contributes to reduction in packet delay is fewer false alarms for live links. Each link error report in AODV triggers a route-discovery procedure causing the packets in transit as well as the following packets to experience a large delay until a new routing path is found. It also causes network-wide flooding of RREQ packets that waste a substantial amount of wireless bandwidth.

The large reduction in packet delay with the CBR traffic motivated us to experiment with TCP traffic because TCP behaves adaptively according to *Round Trip Time* (RTT) estimate. 40 TCP connections are simulated in the same ad hoc network environment. The aggregate end-to-end throughput and response time are plotted in the Figures 25(c) and (d), respectively. As shown in the figures, the MASA achieves as much as 27% and 45% higher throughput than DCF2 and DCF4. Response time is reduced by 70% and 58%, respectively, as seen in Figure 25(d). It is concluded from the figures that in general the MASA protocol and its MAC-layer packet salvaging mechanism improve the network performance, particularly for TCP-based applications. More importantly, the MASA would be best suited in application scenarios where delay is a primary concern.



(a) PDR with CBR

(b) Packet Delay with CBR

(c) Throughput with TCP

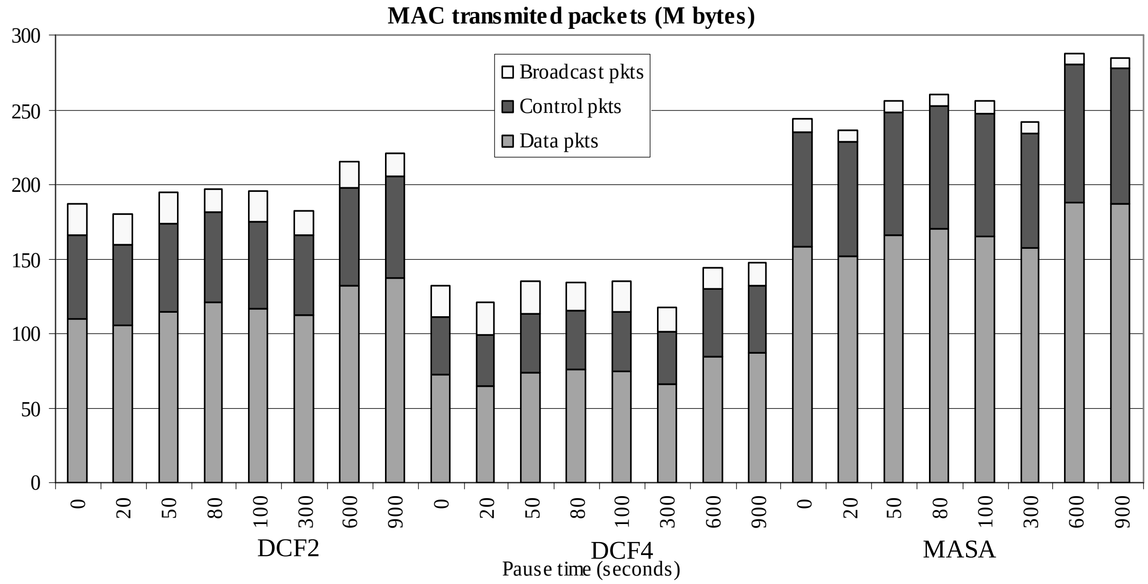
(d) Packet Delay with TCP

Figure 25: Performance Comparison with Mobility

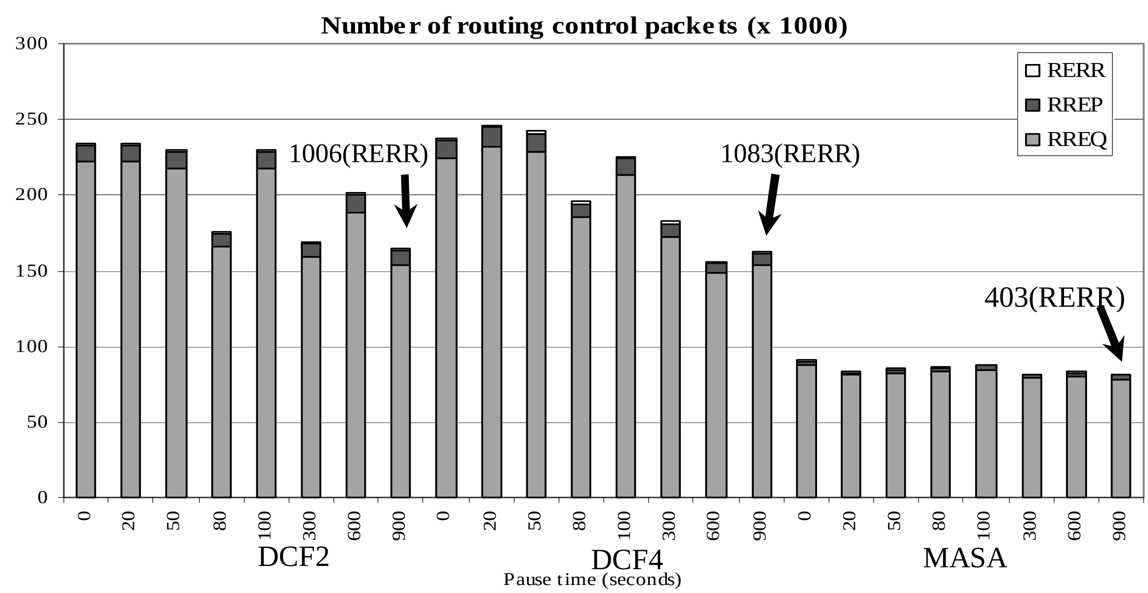
An interesting observation from this simulation result is that performance degrades as node mobility decreases (during 100~900 seconds with CBR traffic in Figure 25(a) and (b), and during 100-300 seconds with TCP traffic as drawn in Figure 25(c) and (d)). The same phenomenon was also reported in [56], the authors of which explained that this is due to a higher level of network congestion and multiple access interferences at certain regions of the ad hoc network. With moderate node mobility, every node experiences overloading when it happens to be in the center but the problem disappears when it moves away from the center. With less mobility, the same set of nodes in the center stay overloaded and thus, they become serious bottlenecks in the network. However, as node mobility decreases even further, link errors are reduced significantly and thus the negative effect is cancelled out. When additive interference is considered, as explained in the previous subsection, overload will be more significant and the corresponding negative effect will continue well beyond the case of unmodified ns-2 simulation without additive interference.

Overhead Analysis

MAC and routing overhead, data overhead, and packet queue size have been measured during the simulation. Figure 26 shows the overhead analysis results with TCP traffic. First, Figure 26(a) presents various overhead traffic: *Address Resolution Protocol* (ARP) traffic (almost negligible), MAC layer control traffic (RTS, CTS and ACK), routing control traffic (RREQ, RREP and RERR) and DATA traffic (TCP data and TCP Ack). Since MASA encourages more concurrent transmissions due to its lower carrier sense range, it shows more DATA traffic, indicating that MASA uses more bandwidth for useful data transmission than DCF2 and DCF4. For instance, with the pause time of 0 seconds, data traffic takes up 91% of entire traffic in case of MASA while it is 83% and 71% in DCF2 and DCF4, respectively. Like DCF2, it shows less MAC layer control traffic than DCF4 because it does not use the RTS/CTS handshake. MASA generates the least routing control traffic, which is detailed in Figure 26(b).



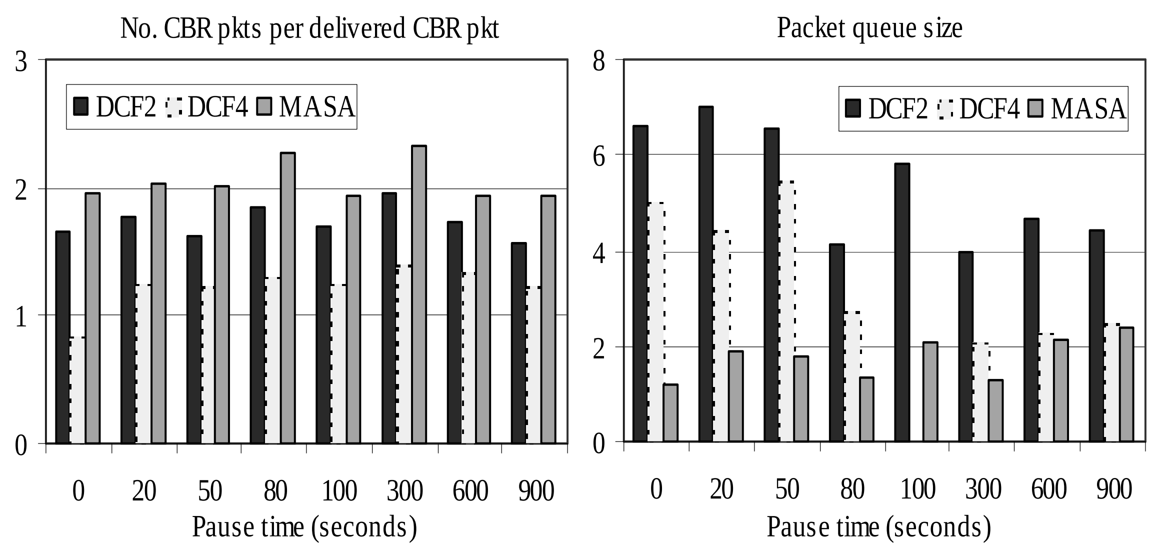
(a) MAC layer overhead



(b) Routing layer overhead

Figure 26: Overhead Analysis with TCP Traffic

However, as far as the data transmission overhead (retransmissions) is concerned, MASA is disadvantageous. Figure 27(a) shows the number of TCP packets transmitted at the MAC layer for each successfully-delivered TCP packet. They are 1.65, 0.84, and 2.08 packets for DCF2, DCF4, and the MASA, respectively, with the pause time of 0 seconds. Since the DCF4 algorithm employs the RTS/CTS exchange before transmitting a data packet, it results in fewer collisions on data packets and thus reduces the number of retransmissions compared to DCF2 and MASA. In comparison with DCF2, the MASA algorithm incurs more overhead mainly because of the reduced CS zone. Nonetheless, it does not overshadow the advantage of MASA as already seen in Figure 25.



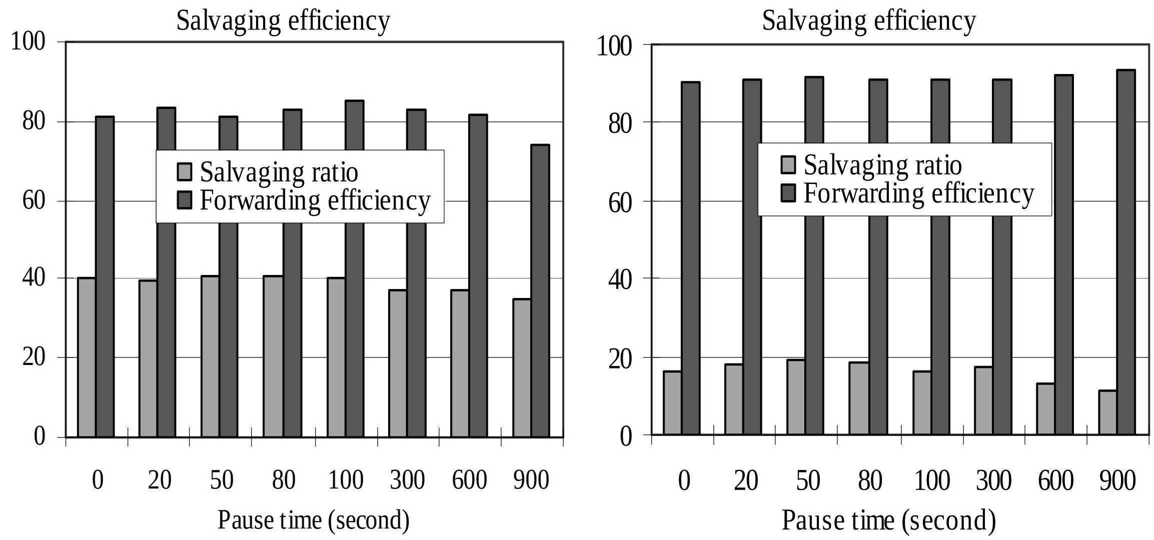
(a) Normalized data overhead (b) Packet queue

Figure 27: Another Overhead Analysis with TCP Traffic

A primary advantage of MASA is short packet delay. The investigation shows that packet queuing delay is an important ingredient for this. Once again, making progress via packet salvaging facilitates a mobile node's quick offloading of the pending packets and therefore, it helps to keep its packet queue at the routing layer as short as possible. In each of 900 seconds of simulation runs, the information about packet queue size is collected every 10 seconds at each node and the average statistics across all mobile nodes in the network are calculated. As shown in Figure 27(b), each node has, on average, about 5.39 and 3.06 packets in its queue with DCF2 and DCF4, respectively, while this number is 1.57 with MASA. Similar observations have been made with CBR traffic.

Salvaging Efficiency

Since MASA salvages collided packets, it would be interesting to know how many packets are actually salvaged, which is called *Salvaging Ratio*, and how many of them are successfully forwarded to the original receivers, which is called *Forwarding Efficiency*. Figures 28(a) and (b) show them with CBR and TCP traffic respectively with the pause time of 100 seconds and the maximum node speed of 5 m/s. The salvaging ratio is about 31% with CBR and 11% with TCP. This higher percentage of packets that is salvaged with CBR means that there are more packet collisions at their first transmission attempts. This is partly because TCP generates less traffic than CBR in this scenario. The light traffic causes little interference and fewer packet drops, which leads to fewer salvaging. On the other hand, the forwarding efficiency is 80% with CBR and 91% with TCP. Most of the salvaged packets are forwarded successfully but TCP traffic results in better efficiency. This is also attributed to the traffic volume they generate.



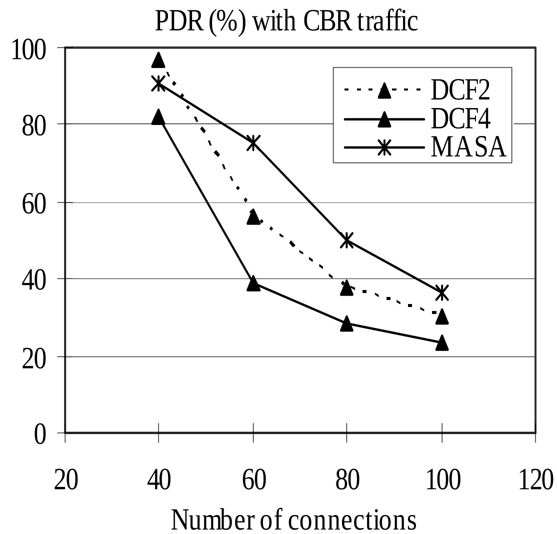
(a) Salvaging Efficiency with CBR (b) Salvaging Efficiency with TCP

(3 packets/second for CBR traffic)

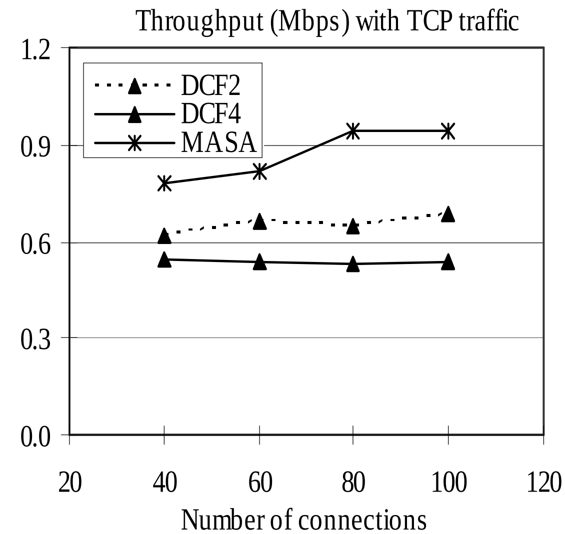
Figure 28: Salvaging Efficiency of MASA

Scalability

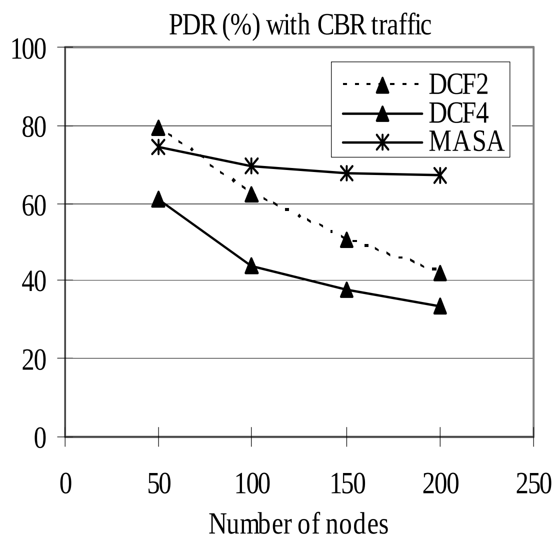
Scalability of MASA is evaluated in terms of different numbers of nodes (node density) and communication pairs (traffic intensity). Figures 29(a) and (b) show the PDR and throughput with different numbers of connections. It can be inferred from the figures, particularly from Figure 29(b), that the proposed MASA consistently outperforms DCF2 and DCF4 regardless of the traffic intensity. In fact, the advantage of MASA becomes more pronounced as the number of connections increases. This is because the MASA encourages more spatial reuse and thus is more beneficial if backlogged nodes can be found in any of the reusable spatial area.



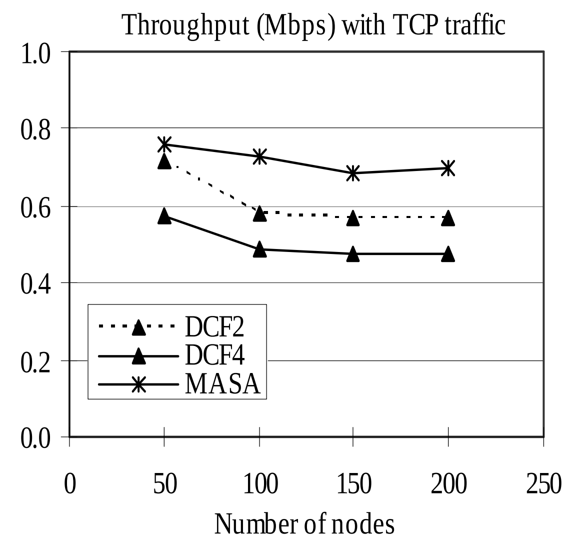
(a) Traffic Intensity with CBR



(b) Traffic Intensity with TCP



(c) Node Density with CBR



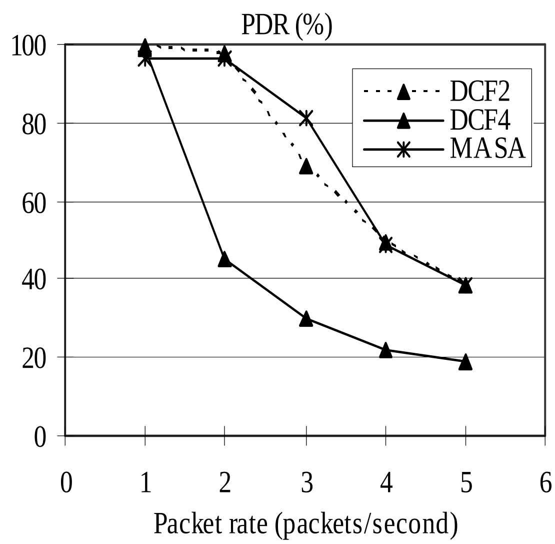
(d) Node Density with TCP

(2 packets/second for CBR traffic)

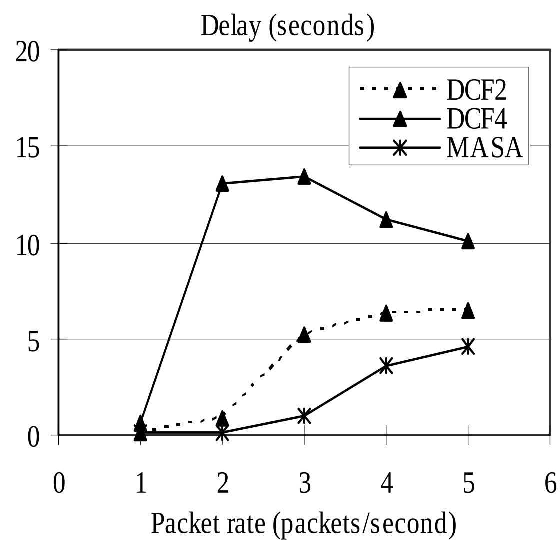
Figure 29: Effect of Traffic Intensity and Node Density

Increasing the number of nodes is especially helpful in MASA as shown in Figures 29(c) and (d) because more candidates are willing to help others by salvaging collided packets. Because DCF2 and DCF4 generally cause more route-discoveries than MASA, increasing the number of nodes directly translates to the exponential increase of routing control overhead. In Figure 29(c), MASA shows higher PDR than the other two, irrespective of the number of nodes; however the gap is more prominent as node density increases. Since TCP sources adapt their data rate based on network feedback, the network performance will not be drastically degraded; however, MASA still outperforms DCF2 and DCF4 as in Figure 29(d).

Effect of Routing Protocols



(a) PDR with DSR



(b) Packet Delay with DSR

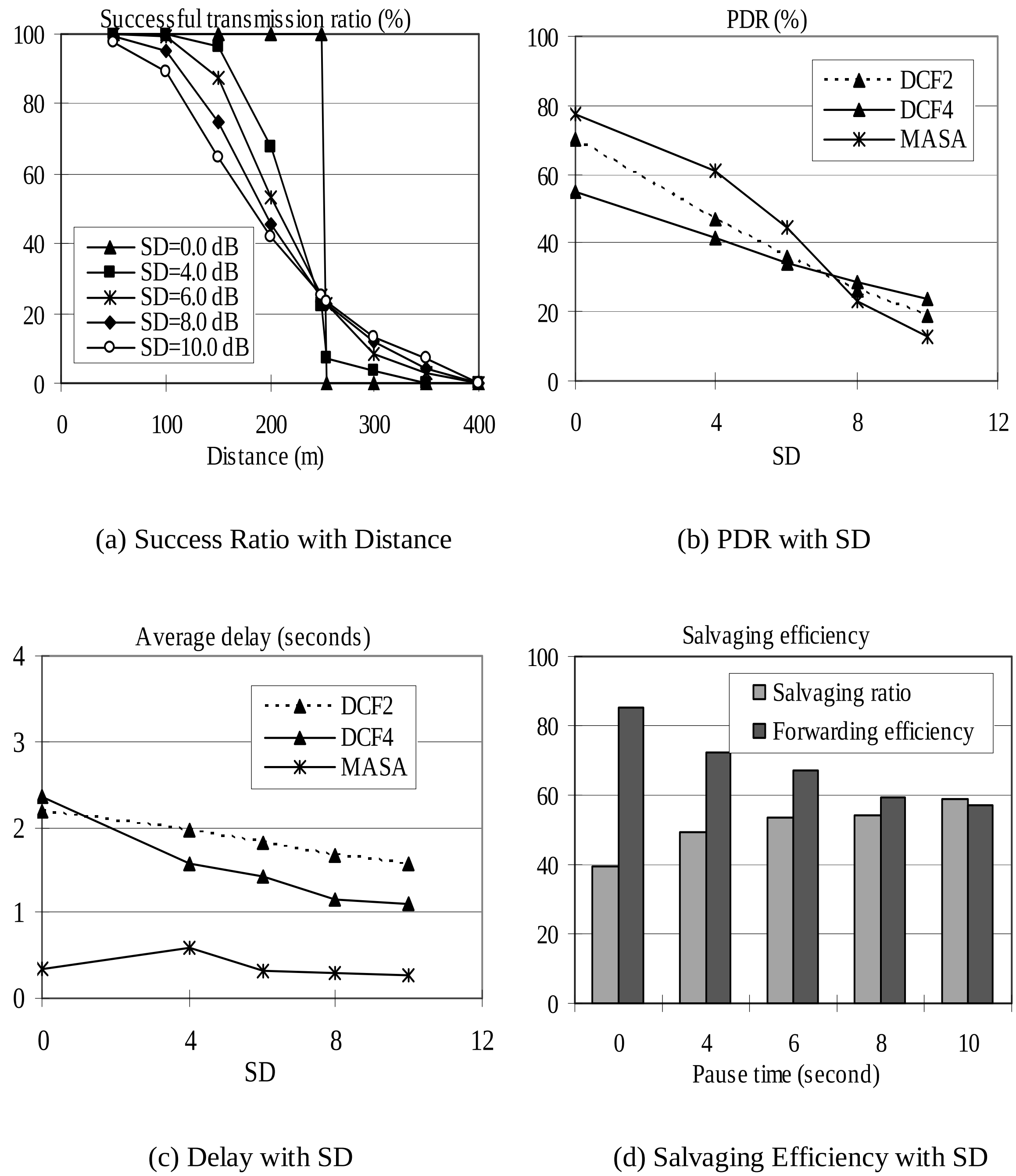
Figure 30: Effect of Routing Protocols

One of the main differentiating characteristics of the proposed MASA protocol is its independence from upper layer protocols. So the effect of routing protocols is discussed here. Figures 30(a) and (b) show the performance evaluation with a different routing algorithm, DSR. 40 CBR sources generate 1~5 packets every second in this simulation. The simulation results show that the performance advantage of the MASA is consistent regardless of the routing algorithm employed at the network layer.

Effect of Unreliable Links

Recent experimental studies show that the shortest (hop count) path does not always provide the best performance because it usually consists of longer hop communications, each of which is easily subjective to interference with a small SINR [47, 57, 58]. In order to see how MASA performs in a practical environment, a set of experiments has been conducted with the *Shadowing Propagation Model* instead of the conventional *Two-Ray Ground Propagation Model*. As introduced in Subsection 3.1.2, shadowing is caused by the lack of visibility between two communicating nodes and it causes slow variations over the mean received power. The mean received power is calculated deterministically based on the communication distance. The randomness of channel is described by a log-normal random variable, the distribution function of which is Gaussian with zero mean and a specified *Standard Deviation* (SD). MASA is expected to be more advantageous over a random channel because of its adaptability.

Before presenting the simulation results, Figure 31(a) shows how the radio channel behaves with the shadowing model by presenting the success ratio versus communication distance using ns-2. In case of SD of 0.0 dB, the shadowing model is equivalent to the deterministic two-ray ground model and thus the success ratio is 100% if the distance is less than 250m (the transmission range). Otherwise, it is 0%. As SD increases, more communications fail even if the distance is less than 250m, and more communications succeed even if the distance is longer than 250m. When the communication distance is 200m, the success ratio is 42% with SD of 10dB. Less than half of the transmission attempts can be successful even if the communication distance is shorter than the transmission range.



(a) Success Ratio with Distance

(b) PDR with SD

(c) Delay with SD

(d) Salvaging Efficiency with SD

Figure 31: Performance with Shadowing Model with CBR Traffic

Figures 31(b) and (c) show the effect of channel randomness on the network performance such as PDR and packet delay with the CBR traffic. MASA consistently outperforms DCF2 and DCF4 in terms of packet delay, as shown in Figure 31(c). However, this is not always the case with PDR, as shown in Figure 31(b). It loses its advantage when SD becomes extremely large, such as 10dB.

This can be explained with Figure 31(d). Since MASA salvages collided packets, it would be interesting to know how many packets are actually salvaged (*salvaging ratio)* and how many of them are successfully forwarded to the original receivers (*forwarding efficiency)*. Figure 31(d) shows that the salvaging ratio and forwarding efficiency are about 40% and 80%, respectively, when SD is 0dB. More than a third of the packets are salvaged (since they are collided) and most of them are forwarded successfully, demonstrating the effectiveness of the MASA algorithm. When SD is 10dB, the salvaging ratio is as high as 59% but the forwarding efficiency is as low as 57%. Only a half of the salvaged packets are forwarded successfully due to the low success ratio, e.g., 42% as explained earlier. Even though some of the lost packets are salvaged and forwarded successfully to the next-hop node ($59\% \times 57\% = 33.6\%$ of packets), many others are ultimately lost in spite of neighbors' help to salvage them. Their help in this case makes the channel contention even worse, decreasing the network performance without yielding any benefit. Packet salvaging does not help when SD is 10dB but the performance benefit of MASA is observed up to a SD of 8dB. Packet delay in Figure 31(c) decreases when the network environment is more random. This should not be interpreted as an improvement because fewer packets are delivered to the desired destinations.

CHAPTER V

IMPROVING SPATIAL REUSE VIA COLLISION AWARENESS

Chapter IV exhibits how MASA improves the spatial reusability by encouraging more concurrent communications while salvaging the collided packets. This mechanism takes effect after collisions have occurred. Essentially, it is a post-collision method. In comparison to MASA, this chapter proposes a pre-collision method called *Collision-Aware DCF* (CAD) that is another enhancement of DCF. CAD is a collision-avoidance mechanism like DCF. However, it takes into account additional factors such as communication distance, packet size, transmit rate, and interference. The collision avoidance mechanism of CAD is much more efficient than that of DCF.

Section 5.1 begins the chapter with a summary of the existing collision avoidance techniques including transmit power control, directional antenna control and carrier sense control. Section 5.2 follows with a presentation of the details of the proposed collision-aware MAC algorithm, CAD. Section 5.3 concludes the chapter by reporting on an extensive simulation based on ns-2 [43].

5.1 Survey on Collision Avoidance Techniques

This section summarizes the recent work addressing the collision avoidance issue based on *Transmit Power Control* (TPC), *Directional Antenna Control* (DAC) and *Carrier Sense Control* (CSC) techniques.

5.1.1 Transmit Power Control

TPC is designed to apply the lowest transmit power necessary to maintain the communication between the sender and the receiver. This lowest transmit power insures that the transmission minimizes the interference to other communications; thus, the collision can be alleviated.

In general, a TPC scheme requires a node to dynamically adjust the transmit power according to the link quality, which is often expressed by link distance, signal strength or successful transmission ratio. Typical DCF based TPC schemes exploit RTS/CTS exchange to detect the link quality and then determine the optimal power level used to transmit Data frames. In order to avoid collisions, maximum transmit power is usually used for the RTS/CTS handshake while minimum necessary transmit power is applied for Data transmission. *Power Controlled Medium Access* (PCMA) protocol proposed in [59, 60] employs a variation of RTS/CTS exchange, called *Request-Power-To-Send* (RPTS)/*Acceptable-Power-To-Send* (APTS), to negotiate the transmit power for Data transmission. In addition, a *power-based interference graph* in [59] is built to smartly control the transmit timing and power. However, a serious problem with most

TPC-based protocols is that TPC causes *asymmetric links* that would fail the symmetric-links-required routing protocols. Moreover, the *asymmetric links* might make it more difficult to handle collision avoidance. [61] introduces *Power-Stepped Protocol* (PSP) to eliminate the *asymmetric link* effects, but it is limited by some requirements of network topology, node density and mobility.

5.1.2 Directional Antenna Control

Applications of directional antenna can significantly improve spatial reuse since it makes more concurrent transmissions possible in the same interference domain. To exploit the benefits of directional antenna, *Directional Virtual Carrier Sensing* (DVCS) technique is introduced in [62], and some MAC protocols using directional antennas are proposed in [63-65]. Unfortunately, directional transmissions could make the hidden terminal problem more serious than omni-directional transmissions because more potential interferers are deaf to the ongoing communication. Additionally, a sender using directional antenna must know the location of the intended receiver to turn over the beam to the right direction. Furthermore, directional transmissions require the line of sight propagation environment. The *Multiple-Input Multiple-Output* (MIMO) system [66] is introduced to overcome these problems. However, it makes the system much more costly and dramatically increases the complexity of the transmitter and the receiver.

5.1.3 Carrier Sense Control

Recent literature indicates that CSC has been considered as an alternative solution based on the assumption that the CS threshold is tunable within the detect sensitivity of the hardware [67]. A higher CS threshold can encourage more concurrent transmissions, but at the cost of more collisions. On the other hand, a lower CS threshold reduces the collision probability but requires a larger spatial footprint and prevents simultaneous transmissions from occurring. It potentially limits the network throughput. Obviously, there is a tradeoff between high spatial reuse and increased chances of collisions [67].

Fuemmeler, *et al.* studied the collision prevention conditions in this context and concluded that the product of transmit power and CS threshold should be kept to be a fixed constant [68]. Zhu, *et al.* used an analytical model to determine the optimal CS threshold [69]. However, their analytical model does not consider the influence of MAC overhead and transmit rate, which has been addressed by Yang and Vaidya [71] and Zhai and Fang [72].

While these studies focus on analytical models for obtaining aggregate throughput, Zhu, *et al.* proposed a distributed algorithm, called *Adaptive Physical Carrier Sensing* (APCS), which dynamically adjusts the CS threshold in 802.11 mesh networks [70]. This scheme was improved recently by adding receive sensitivity adaptation [73]. It is considered as a *receiver technique* in the sense that a node is not allowed to receive a weak signal even though it is stronger than the CS threshold. The node's radio cannot be *locked* onto the first signal and thus becomes available to receive any late-arriving signals [73]. In contrast, the proposed approach in this chapter is considered as a *transmitter technique* because a node is allowed to transmit its pending frame even in the presence of carrier signal.

5.2 Collision-Aware DCF (CAD)

This section presents the details of the proposed collision-avoidance MAC algorithm, *Collision-Aware DCF* (CAD). The collision-awareness of CAD is due to sharing the spatial and time reservation requirements among neighboring nodes. Three issues will be addressed in this section. First, how are the reservation requirements estimated? Second, how are they prepared and distributed? Third, how are they received and handled?

5.2.1 Overview of CAD

As introduced in Chapter II, 802.11 DCF renders a node to defer its communication if it senses that the medium is busy. For the duration of deferment, each packet carries in its MAC header a 16-bit number in microseconds during which the overhearing nodes must defer. However, even if the carrier signal is detected, both ongoing and new communications can be simultaneously successful depending on their relative positions in the network and the status of the communication channel. In other words, DCF defers more communications than necessary in favor of simplicity. In addition, in the DCF, the time duration information mentioned above is not delivered to all potential interferers, particularly those that are far from the sender.

The proposed method in this chapter, CAD, efficiently utilizes the available channel resource along both the spatial and time dimensions. In the CAD, a transmitter estimates the optimal spatial and time deferment requirements, which are adaptively based on the communication distance, packet size, transmit rate and the status of the medium. The transmitter then propagates this information by embedding it in the transmitted frame. A potential interferer in the proximity estimates its own optimal reservation and confirms two things before transmitting its frame: whether or not it disrupts the ongoing communication and whether or not the ongoing transmission disrupts its communication. An important design decision in the CAD is to embed the spatial reservation and desired defer duration in the PHY header instead of the MAC header. This is beneficial because a larger group of neighbors receive this information and behave more consistently.

The details of estimating, distributing and handling the spatial and time reservation requirements will come up in the following subsections.

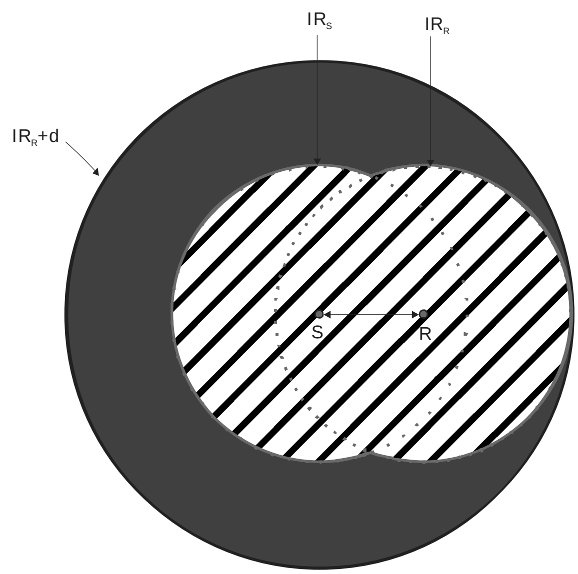
5.2.2 Estimating Spatial and Time Reservation Requirements

The time reservation requirement embedded in a packet is the time period required to protect the communication of the issuing node. This is similar to Duration/ID field in each *MAC Protocol Data Unit* (MPDU) packet. It is 16-bit duration information measured in microseconds and is used by each node to maintain the status of the medium [24]. CAD does the same thing, but based on the information included in the *Physical Layer Convergence Protocol* (PLCP) header of the frame that it receives. Since a MAC packet is not considered legitimate until the whole packet is received and its CRC checksum is confirmed, the time reservation in the DCF does not take the present packet into consideration. However, in the CAD, since the PLCP header is received and confirmed for its integrity even before a node starts receiving the PHY payload, the time reservation information includes the duration for the present packet as well.

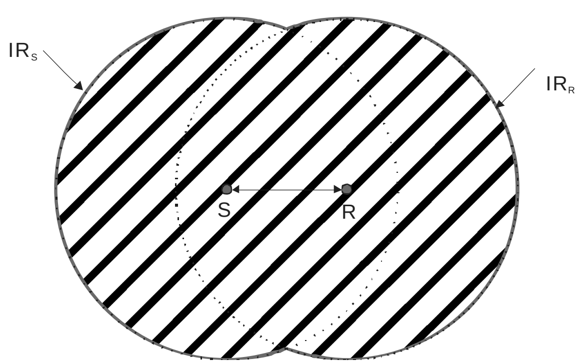
The spatial reservation requirement in the CAD is based on the IR estimation discussed in Chapter III. It becomes more optimal as each step of the 4-way handshaking procedure progresses because the IR estimation becomes more accurate with additional information obtained from the previous steps. Figure 32 shows the spatial reservation of CAD. In the figure, node S is the sender and node R is the receiver. They are separated with distance *d*. IR_S and IR_R are their interference range, respectively.

- While an RTS frame is transmitted, IR_R must be reserved for the duration of the RTS frame but IR_S must be protected until the sender receives the CTS in return. Since the sender just initiates the communication it is responsible for making spatial reservation to protect the RTS reception at the receiver. Therefore, the sender calculates the spatial reservation requirements as $\mathrm{SR} = d + \mathrm{IR}_{\mathrm{R}} = (1 + \sqrt[4]{Z_0}) \cdot d$, as shown in the Figure 32(a). The communication distance *d* is assumed to be available based on the signal strength measurements in the past. The time reservation requirement is thus the RTS and CTS transmission time plus additional overhead such as inter-packet gap (known as SIFS) and propagation delay (Δ).

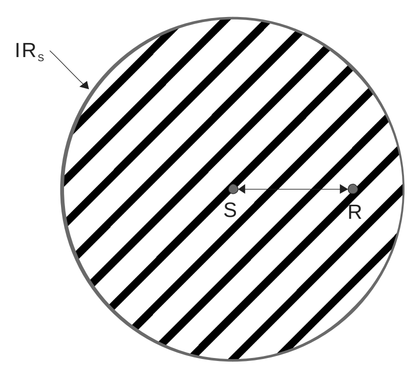
- While a CTS frame is replied, IR_S must be reserved for the duration of the CTS frame and IR_R must be protected until the receiver receives a DATA frame. However, since the CTS reception has already been protected based on the information in the RTS frame, the receiver only needs to concern about its reception of the DATA frame. Therefore, the spatial reservation requirement in the CTS is reduced to $\mathrm{IR_R} = \sqrt[4]{Z_0} \cdot d$. However, during CTS transmission the spatial reservation made by RTS is still effective, so the overall spatial reservation is as same as that during RTS transmission, as shown in the Figure 32(a). The time reservation requirement in the CTS is the CTS and DATA transmission time plus additional overhead mentioned above.



(a) While RTS or CTS is Transmitted



(b) While Data is Transmitted



(c) While ACK is Transmitted

Figure 32: Spatial Reservation in CAD

- When a DATA frame is transmitted, the spatial reservation requirement is shown in Figure 32(b). Since the sender does not have to make the reservation for the receiver's reception, it only makes reservation for its ACK reception. Therefore, the spatial reservation requirement in the DATA is reduced to $\mathrm{IR_S} = \sqrt[4]{Z_0} \cdot d$. Due to this reduction, during DATA frame transmission the spatial reservation SR = $\mathrm{IR_R} \cup \mathrm{IR_S}$, is minimized. In other words, both WS = SR – $\mathrm{IR_R} \cup \mathrm{IR_S}$ and VS = $\mathrm{IR_R} \cup \mathrm{IR_S}$ – SR are Ø. As shown in the Figure 32(b), the exposed and hidden terminal problems are nicely handled. For the time reservation requirement in the DATA frame it is the DATA and ACK transmission time plus additional overhead.

- Finally, when an ACK frame is replied, the receiver does not have any reservation requirement. Only the sender's reservation in the previous DATA fame is effective as shown in Figure 32(c).

TABLE III: SPATIAL AND TIME RESERVATIONS IN CAD

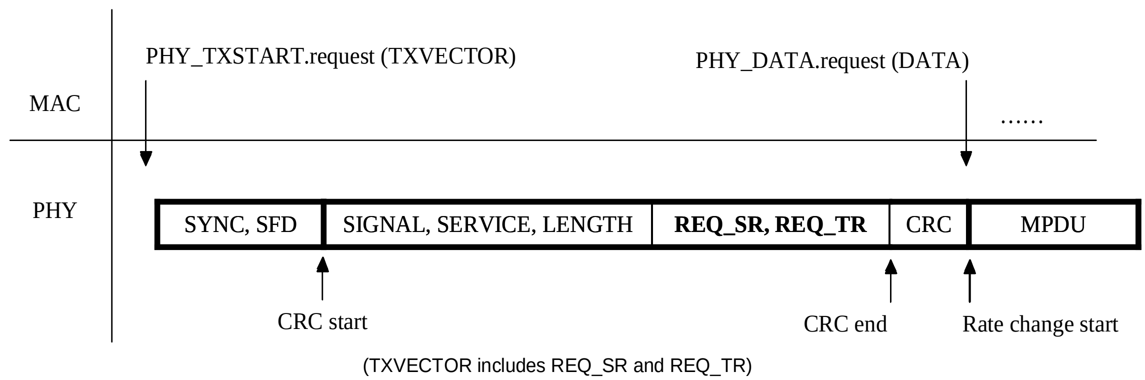
Frame	**Spatial Reservation**	**Time Reservation**
RTS	$d + \mathrm{IR_R} = (1+\sqrt[4]{Z_0}) \cdot d$	T_{RTS}+SIFS+T_{CTS}+Δ
CTS	$\mathrm{IR_R} = \sqrt[4]{Z_0} \cdot d$	T_{CTS}+SIFS+T_{DATA}+Δ
DATA	$\mathrm{IR_S} = \sqrt[4]{Z_0} \cdot d$	T_{DATA}+SIFS+T_{ACK}+Δ
ACK	0	0

As a summary, Table III shows the spatial and time reservation requirements of each frame in the CAD, where Δ is the propagation delay, and T_{RTS}, T_{CTS}, T_{DATA} and T_{ACK} are the transmission time for the RTS, CTS, DATA and ACK, respectively. Since the transmission time of a frame is determined by the transmission rate and the frame size, it is not difficult to estimate.

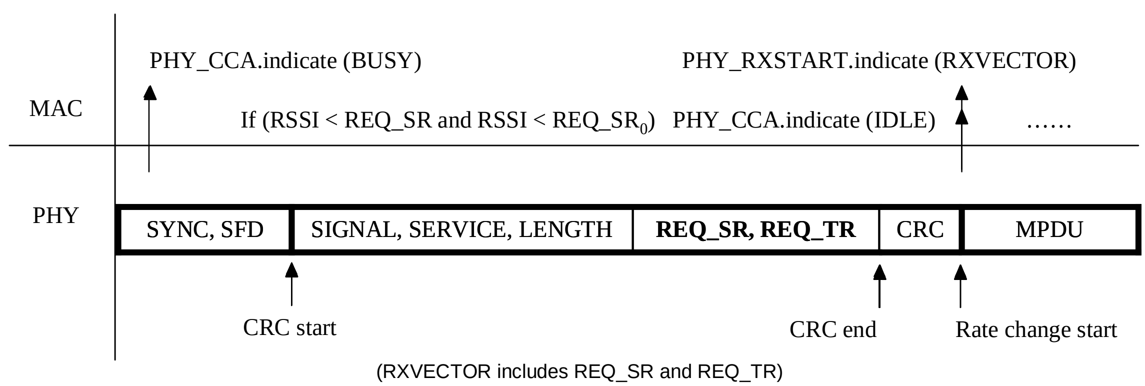
5.2.3 Preparing and Distributing Reservation Requirement

The previous section introduced the spatial and time reservation requirements, denoted here as REQ_SR and REQ_TR respectively. For the convenience of operation the spatial reservation is translated into power. For example, in the case of an RTS frame, REQ_SR $= P_r\left(\left(\sqrt[4]{Z_0}+1\right)\cdot d\right)$. For other frames, REQ_SR is $P_r\left(\left(\sqrt[4]{Z_0}\right)\cdot d\right)$. To get the communication distance d, the CAD requires that each node maintains a neighboring list that contains the relevant signal strength information (*e.g.* recent *Received Signal Strength Indication* (RSSI)) of each neighbor. In case no signal strength information is available, the maximum communication distance corresponding to the transmit rate is assumed to estimate REQ_SR. Estimation of REQ_TR is based on frame length, frame type and data rate.

When a backlogged node estimates REQ_SR and REQ_TR at the MAC layer, this information is added in the TXVECTOR and passed to the PHY layer along with PHY_TXSTART.request [24] as shown in Figure 33(a). PHY layer prepares the PLCP header as in the Figure 34, where REQ_SR and REQ_TR are embedded. It then transmits the frame according to the transmit procedure specified in the 802.11 standard [24], which not only delivers the header to the receiver but also distributes the reservation requirements to the neighboring potential interferers so that they can optimally decide whether or not to comply. Data rate used for MPDU transmission is indicated by the 8-bit SIGNAL field and the MPDU is transmitted upon PHY_DATA.request as in Figure 33(a).



(a) Transmit Procedure



(b) Receive Procedure

Figure 33: Transmit and Receive Procedures in the 802.11 Standards

Embedding the reservation requirement of a transmission in the PLCP header has two benefits: First, a neighbor can immediately determine if it is the potential interferer of the ongoing communication when it receives the PLCP header. Second, since the PLCP header is transmitted at the lowest data rate, it reaches nodes in farther distance as in Table I. While neighbors in direct transmit range would not cause any trouble because they hear the ongoing communication anyway, those in farther distance are the most critical potential interferers. In 802.11, they are inhibited partly by RTS/CTS exchange and partly by EIFS as discussed earlier. However, CAD embeds the requirements in the PLCP header so that they are delivered to a larger group of neighbors in a simpler manner.

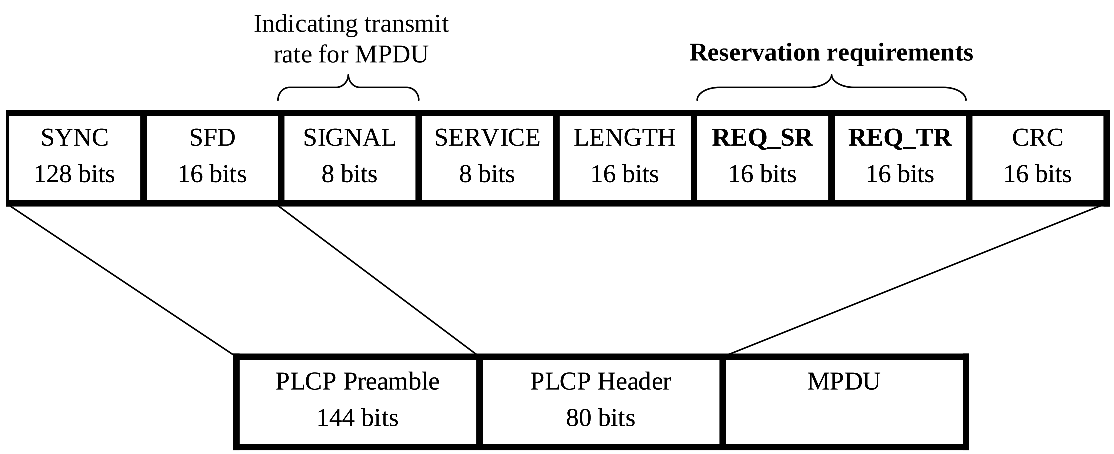


Figure 34: PLCP Frame Format (REQ_SR and REQ_TR are added)

5.2.4 Receiving and Handling Reservation Requirements

Figure 33(b) shows the PLCP receive procedure. Upon detecting a coming signal, PHY_CCA.indicate (BUSY) will be issued to the MAC layer if the signal strength is higher than CS threshold. Then PHY will begin searching for *Start Frame Delimiter* (SFD) and start to receive a PLCP header. If a PLCP header is successfully received (CRC check passes), the PHY_RXSTART (RXVECTOR) will be issued to the MAC layer according to the 802.11. The RXVECTOR contains the information of SIGNAL field, SERVICE field, LENGTH field, RSSI, signal quality, and antenna used for receive.

In the CAD, RXVECTOR includes REQ_SR and REQ_TR in addition to the information mentioned above. When a backlogged node receives a PLCP header successfully, it has to make two decisions: Whether or not its communication is successful if it transmits concurrently with the current data transfer and whether or not the current communication will be successful if it transmits. For the former question, the node compares the RSSI of the incoming signal with its own spatial requirement, denoted as REQ_SR_0. In other words, the node defers if $RSSI \geq REQ\_SR_0$ because the strength of the incoming signal exceeds the maximum interference level that its outgoing transmission can tolerate. For the latter, the node compares the RSSI of the incoming signal with REQ_SR of the current transmission, *i.e.*, the node defers if $RSSI \geq REQ\_SR$ because the current communication would fail if the node transmits. This is based on the assumption that the link is symmetric; the RSSI of the incoming signal is equal to the RSSI (interference) that the node would cause to the ongoing transmission.

In summary, if $RSSI \geq REQ\_SR_0$ or $RSSI \geq REQ\_SR$, the medium is considered busy and the node holds up its transmission. In this case, PHY will continue to receive MPDU but NAV is set to a new value REQ_TR, which is obtained from the incoming PLCP header. On the other hand, if $RSSI < REQ\_SR_0$ and $RSSI < REQ\_SR$, the medium is considered idle. In this case, PHY will issue PHY_CCA.indicate(IDLE) to the MAC layer so that the node can transmit its frame even though there is an ongoing communication. Therefore, CAD encourages more concurrent communications as long as they do not interfere with each other and thus increases the network throughput.

5.3 Simulation and Evaluation

In order to evaluate the performance improvement of CAD, this section compares CAD with DCF based on *ns-2* [43]. The simulation environment is explained in Subsection 5.3.1. Simulation results and discussions are presented in Subsection 5.3.2.

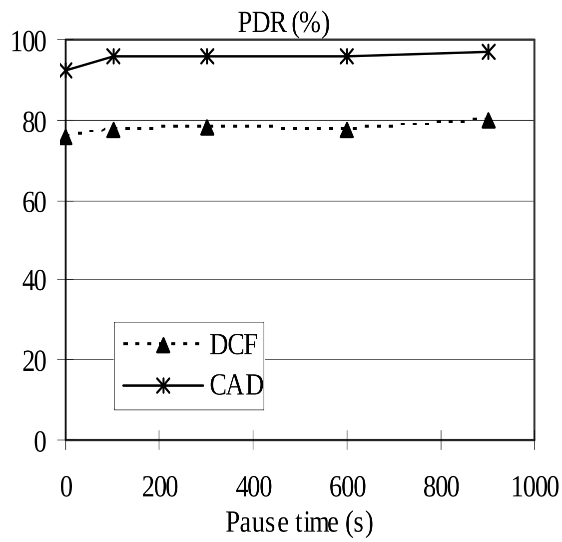
5.3.1 Simulation Environment

The performance study is based on *ns-2* simulation of 50 mobile nodes that are distributed in a 300×1500 m^2 area. The radio propagation model used is the *path loss model* discussed in Subsection 3.1.1 and equations (3.1) and (3.2). Transmit range (TR) of 250m and carrier sense range (CR) of 550m is assumed. Capture ratio (Z_0) of 10dB is used in the performance study. Regarding signal transmission and capture, ns-2 has been extended as described in Subsection 4.3.2.

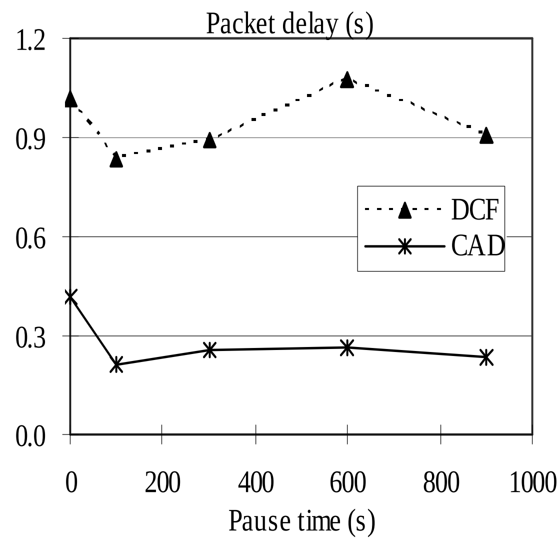
The movement of the nodes is described by the *random waypoint mobility model* with the maximum speed of 5m/s and with the pause time of 0~900 seconds. 10~50 *Constant Bit Rate* (CBR) traffic is used to simulate the network traffic. *Ad-hoc On-demand Distance Vector* (AODV) routing protocol [45] is used in the simulation to determine the routing path between the source and the destination. The simulation time is 900 seconds and each simulation scenario is tested with five runs to obtain the average performance measures.

5.3.2 Results and Discussion

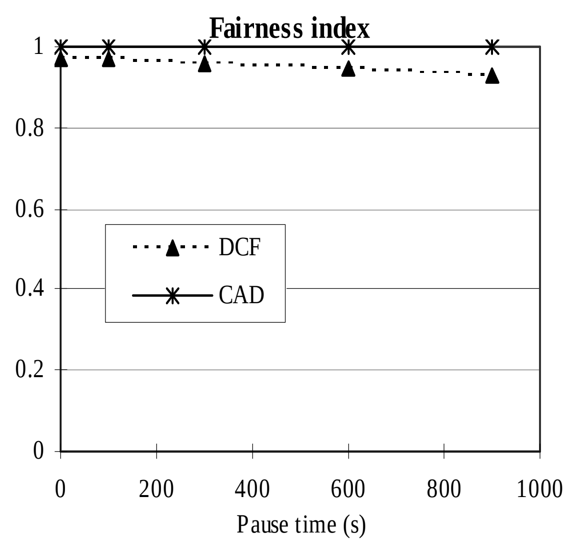
General Network Performance



(a) Packet Delivery Ratio (%)



(b) Packet Delay (seconds)



(c) Fairness Index

Figure 35: Performance Comparison with Mobility

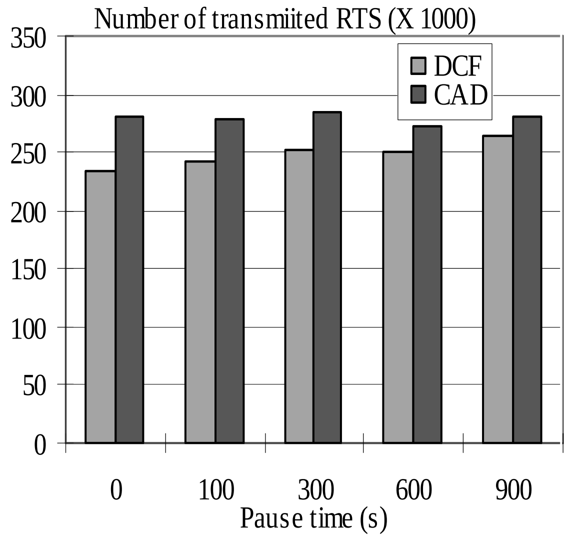
This subsection presents simulation results comparing the performance of the proposed CAD with DCF. *Packet Delivery Ratio* (PDR) and packet delay are used as primary performance metrics as shown in Figure 35. The pause time varies between 0 to 900 seconds. Note that pause time of 900s translates to a static network where nodes do not move because the simulation time is 900s. On the other hand, pause time of 0s corresponds to a constant moving scenario. 30 CBR connections are simulated where source and destination nodes are chosen randomly among the 50 mobile nodes. Each traffic source generates three 1024-byte packets every second. As clearly seen in Figures 35(a) and (b), CAD significantly outperforms DCF. For example, when the pause time is 300s, CAD achieves 17% higher PDR and 72% lower packet delay than DCF.

Figure 35(c) investigates the fairness problem. It is measured by *Fairness Index* that was originally proposed in [74, 75], here its extended definition used in [76] is applied. It is defined as follows.

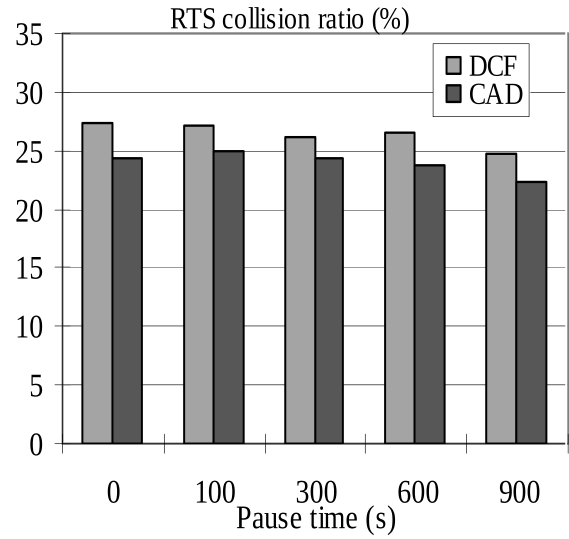
$$F = \frac{\left(\sum_{i=1}^{N} \gamma_i\right)^2}{N \sum_{i=1}^{N} \gamma_i^2} \tag{5.1}$$

where N is the number of connections and γ_i is the number of received packets for connection i. The value of this index is between 0 (completely unfair) and 1 (perfectly fair). According to the CAD, the nodes may not block their transmissions even they detect the medium busy. This might cause the problem of unfair communications. However, the Figure 35(c) shows that the CAD does not hurt the fairness of communications. On the contrary it improves the fairness than the original DCF.

Transmission Concurrency and Collision Analysis



(a) Number of RTS Packets (per second)



(b) RTS Collision Ratio (%)

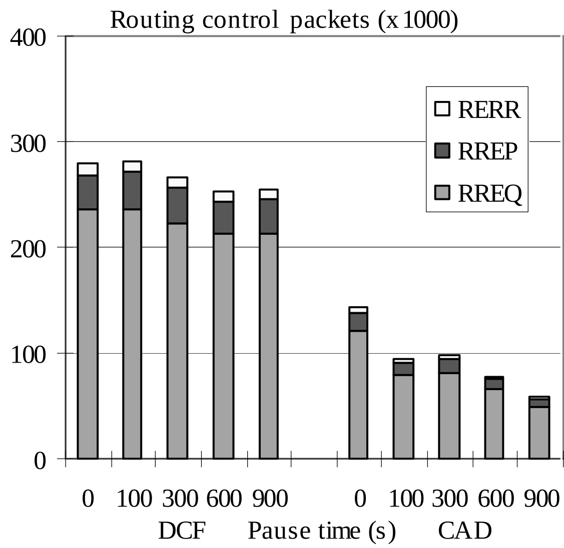
Figure 36: Transmission Concurrency and Collision Analysis

The dramatic performance improvement of CAD over DCF is attributed to the higher concurrency and to its superior capability in avoiding collisions. While nodes make transmission decisions depending on the carrier signal and the pre-determined CS threshold in DCF, the CAD protocol allows the nodes to make more intelligent decisions based on information from their neighbors. As a result, CAD produces more communication opportunities but reduces collisions, as evident in Figures 36(a) and (b), respectively. Note that the number of RTS transmissions and their collisions are investigated for this purpose because every routing or data packet is preceded by a RTS packet. Figure 36(a) shows the number of RTS packets transmitted during the simulation period of 900s. With CAD, nodes send 6~20% more RTS packets (communication opportunities) than with DCF. However, CAD results in 7.7~10.0% less collisions on RTS packets as shown in Figure 36(b). Note that Figure 36(a) does not count the retransmitted RTS packets because more retransmissions mean more collisions rather than more communication opportunities. On the other hand, Figure 36(b) includes both initial and retransmitted RTS packets because the collision probability in general is good to know.

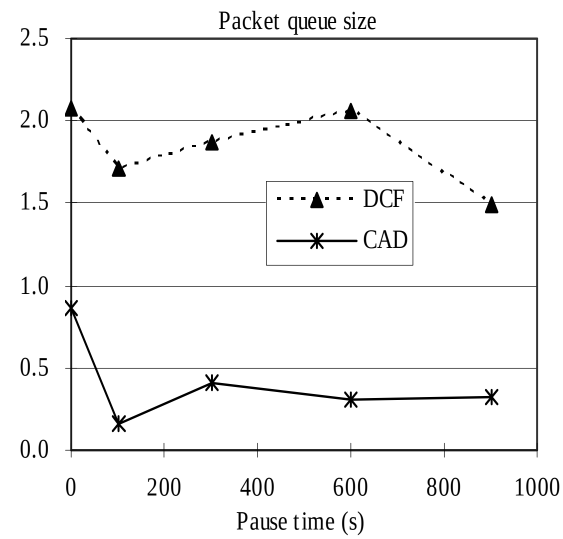
Overhead Analysis

In order to understand how CAD improves the performance, some overhead are measured. These measured overhead include routing layer and MAC layer control overhead as well as the packet queue size as shown in the Figure 37. Figure 37(a) shows that DCF generates 1.9~4.3 times more routing control traffic than CAD. At the pause time of 900s (static network), no RERR packets are expected but DCF still generates 10.4 RERR packets on the average, which must be contrasted to 2.5 such packets with CAD as in Figure 36(a). Intelligent spatial and time reservation in CAD increases the resistance of communication links to interference. This significantly reduces the occurrence of false alarms that cause the unnecessary routing control overhead.

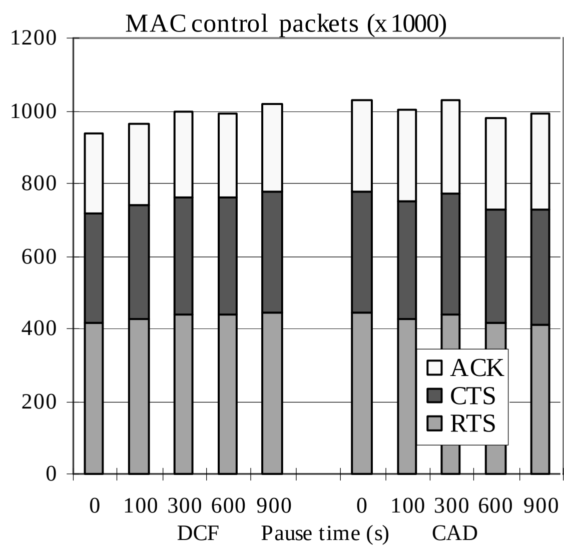
One important advantage of CAD is short packet delay. The investigation shows that packet queuing delay is an important ingredient for this. Having more communication opportunities in CAD facilitates a mobile node to quickly offload pending packets and it therefore, helps to keep its packet queue at the routing layer as short as possible. In each of the 900 seconds of simulation runs, the information of packet queue size is collected every 10 seconds at each node and the average statistics is calculated across all mobile nodes in the network. As shown in Figure 37(b), each node has about 1.5~2.1 packets in its queue on the average with DCF while it is 0.2~0.9 with CAD.



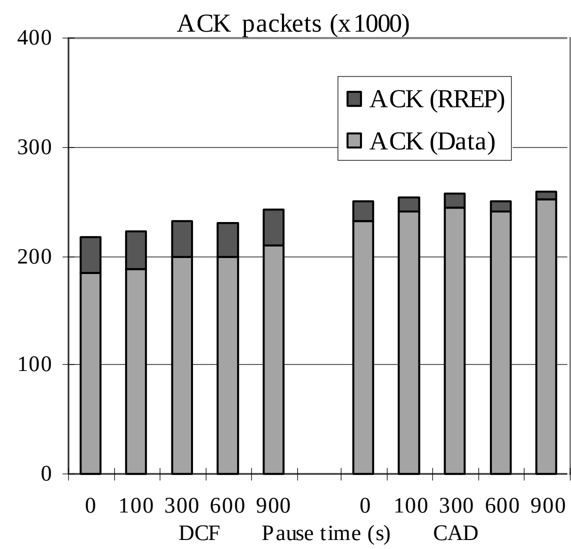
(a) Routing Control Overhead



(b) Packet Queue Size



(c) MAC Control Overhead



(d) A detailed Analysis of ACK Packets

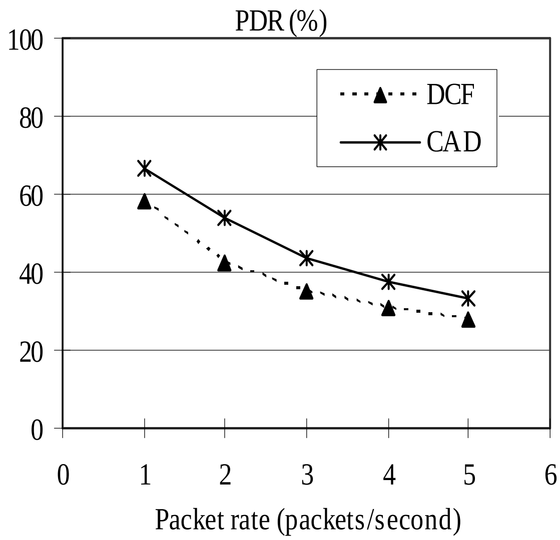
Figure 37: Overhead Analysis

Figure 37(c) presents MAC layer control traffic (RTS, CTS and ACK) in CAD and DCF. Although CAD generates about the same amount of MAC layer overhead, its detailed figures are quite different. For example, the number of RTS packets is almost the same in CAD and DCF but this is only true for the combined initial and retransmitted RTS packets. While CAD allows more number of initial RTS packets as already seen in Figure 36(a), it causes less RTS retransmissions than DCF (not shown here for brevity). Similarly, the total number of ACK packets is comparable between CAD and DCF. However, CAD results in more ACKs than DCF in response to DATA packets while it is exactly the opposite for ACKs in response to RREP packets as shown in Figure 37(d). This indicates that CAD uses more bandwidth for useful data transmission than DCF.

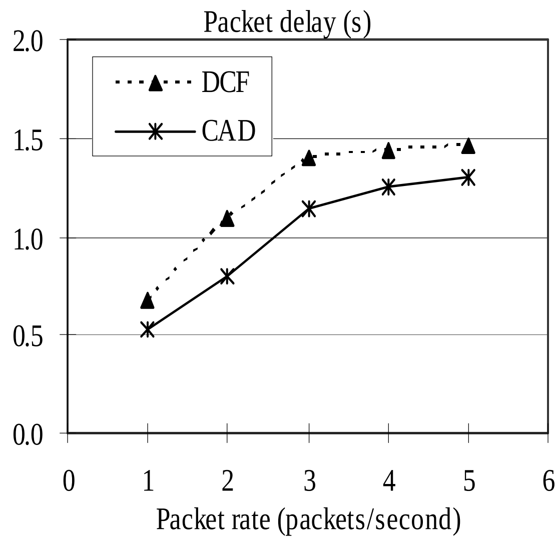
Effect of Unreliable Links

In order to see how CAD performs in a more realistic environment, a set of experiments has been conducted with the *Shadowing Propagation Model* instead of the conventional *Path Loss Model*. The randomness of the channel is described by a specified standard deviation (SD). The effects that the SD has on the channel are presented in the Figure 31(a). When SD is zero, the channel is no longer random. In other words, it converges into the *Path Loss Model*. Figures 38(a) and (b) show the effect of packet rate when SD=4.0. CAD outperforms DCF in PDR with about 10% and packet delay with about 0.5 second.

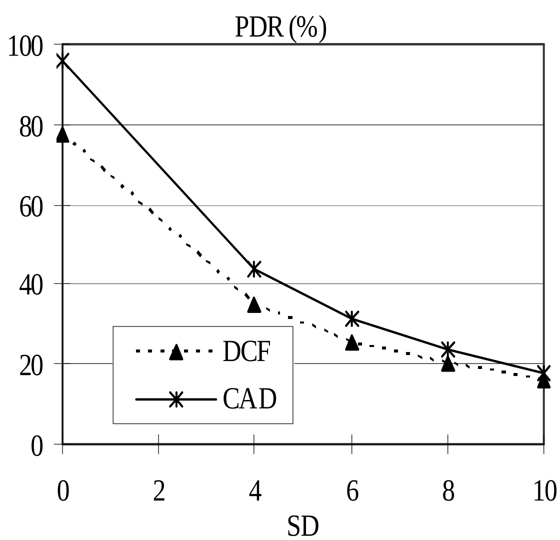
Figures 38(c) and (d) show the effect of channel randomness in term of SD on the network performance. CAD consistently outperforms DCF in terms of PDR. However, the PDR margin between CAD and DCF decreases when SD is getting larger. This is because in CAD the REQ_SR calculation is deterministic and based on the communication distance only. In other words, the REQ_SR estimation is not accurate in the random channel. The more random the channel it is, the more error the estimation contains, and the performance margin is reduced. Efficient operation in the presence of randomness of the communication channel comprises one of the future works.



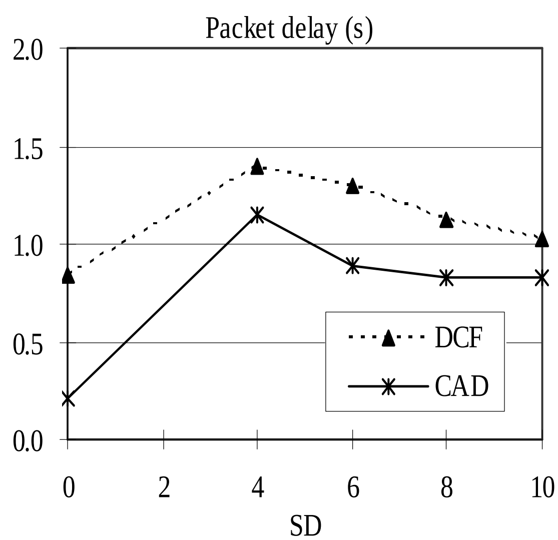
(a) PDR with Packet Rate



(b) Packet Delay with Packet Rate



(c) PDR with SD



(d) Packet Delay with SD

Figure 38: Effect of Random Channel

CHAPTER VI

CONCLUSIONS AND FUTURE WORK

This chapter summarizes the dissertation with the highlights of the proposed two proposed solutions MASA and CAD. Additionally, the future work will also be discussed here.

6.1 Conclusions

Carrier sensing MAC protocols such as IEEE 802.11 DCF avoid collisions by employing aggressive carrier sensing. However, this makes them unable to maximize the spatial spectral utilization. This dissertation thoroughly discusses the spatial reuse in DCF and analyzes the upper bound network throughput with a carrier sensing MAC. Based on the analysis, two different solutions (MASA and CAD) are proposed to improve the spatial reusability.

The MASA algorithm adopts a fixed, small carrier sense range but adaptively adjusts the communication distance via salvaging packets. While the former encourages more concurrent communications, the latter alleviates the collision problem. MASA is a pure MAC solution. It is more efficient than those at the routing layer because the packet salvage at the routing layer occurs after the repeated failures of retransmissions at the MAC layer. The salvaging mechanism in MASA essentially ensures that the salvage is better than retransmission in DCF. The extensive simulation study showed that MASA enhances the network performance regardless of mobility, traffic intensity and the routing algorithm used. In particular, it reduces packet delay significantly.

While the DCF avoids collisions based on a pre-determined carrier sense threshold (physical carrier sense) and the advertisement of the communication duration embedded in the MAC header (virtual carrier sense), both methods often fail to reach the maximum achievable performance, particularly in a multi-hop network environment. In the CAD, each node estimates the range that it wishes to reserve for its data transfer (spatial reservation requirement) and the time duration (time reservation requirement) based on the communication distance, transmit rate, packet type and size. The analysis shows that after RTS/CTS exchange, the CAD algorithm only reserves the spatial area that is required (both VS and WS become empty). Because the reservation requirements are embedded in the PLCP header, a larger group of potential interferers become aware of it. The simulation study based on ns-2 shows that CAD significantly improves the network performance in terms of packet delivery ratio and packet delay. The benefit of CAD is derived from more number of concurrent transmissions and smaller collision ratio, which in fact is the original goal of the CAD mechanism.

6.2 Future Work

According to the simulation results in Chapter IV, the MASA algorithm is considered the most preferable in a wireless ad hoc network where a large number of nodes exchange small packets, as is typically the case in *wireless sensor networks*. Applying MASA in this area is one of the future works. In the MASA, the salvager is elected based on a backoff procedure. In some particular scenarios, more than one node may salvage a same packet then cause duplicate receptions at the receiver. Even though the duplicate packets can be filtered out, bandwidth is wasted to deliver them. Electing a salvager deterministically rather than randomly between each pair of communicating nodes is another future work of MASA.

The benefits of CAD come from the accurate estimation and notification of collisions. However, the estimation accuracy is limited by the randomness of the channel, and the notification range is limited by the transmit rate. Study on how to improve the performance in the random channel or in the transmission with low data rate is one of the future works.

Basically, MASA is a post-collision solution because it is engaged after collision occurs while CAD is a pre-collision solution. Research on how to integrate them is an interesting work in future. In addition, CAD is designed to be compatible with TPC and TRC capability in the sense that estimation of the spatial reservation requirement can easily accommodate the transmit power and transmit rate information. This issue requires further study in future work.

REFERENCES

[1] “Wireless Telegraphy,” http://en.wikipedia.org/wiki/Wireless_telegraphy.

[2] T. H. White, “United States Early Radio History,” http://earlyradiohistory.us/sec002.htm.

[3] “Guglielmo Marconi,” http://en.wikipedia.org/wiki/Guglielmo_Marconi.

[4] J. Shea, “Brief History of Wireless Communications,” http://wireless.ece.ufl.edu/~jshea/ eel6509/misc/ history.html.

[5] P. Noboa, “World of Wireless Networking,” http://www.sinc.sunysb.edu/Stu/pnoboa.

[6] “History of Wireless,” http://www.jhsph.edu/wireless/history.html.

[7] “Personal Digital Assistant,” http://en.wikipedia.org/wiki/Personal_digital_assistant.

[8] “Personal Digital Assistant (PDA) Tutorial,” http://gsm.utmck.edu/library/pda/tutorial/introduction.htm.

[9] “Smartphone,” http://en.wikipedia.org/wiki/Smart_phone.

[10] “Wireless LAN,” http://en.wikipedia.org/wiki/Wireless_LAN.

[11] “IEEE 802.11,” http://en.wikipedia.org/wiki/IEEE_802.11.

[12] IEEE 802.11 Group, http://grouper.ieee.org/groups/802/11.

[13] Strategy Analytics, http://www.strategyanalytics.com/press/PR00342.htm.

[14] O. K. Tonguz and G. Ferrari, “Ad Hoc Wireless Networks: A Communication-Theoreteic Perspective,” May 2006.

[15] “Mobile Ad Hoc Network,” http://en.wikipedia.org/wiki/MANET.

[16] M. Guarnera, M. Villari, A. Zaia and A. Puliafito, “MANET: possible applications with PDA in wireless image environment,” *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Lisboa , Portugal, September 2002.

[17] Madhavi W.Subbarao. “Mobile Ad Hoc Networks for emergency preparedness Telecommunications – Dynamic Power-Conscious Routing Concepts,” *Wireless Communications Technologies Group, National Institute of Standards and Technology*, February 2000.

[18] IEEE Std 802.3-2002, “Local and Metropolitan Area Network, Specific Requirements, Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications,” http://standards.ieee.org/ getieee802/download/802.3-2002.pdf.

[19] I IEEE Std 802.4-1990, “Local and Metropolitan Area Network, Specific Requirements, Part 4: Token ring access method and Physical Layer specifications,” http://standards.ieee.org/ getieee802/ download/802.4-1990.pdf.

[20] IEEE Std 802.5-1998, “Local and Metropolitan Area Network, Specific Requirements, Part 5: Token ring access method and Physical Layer specifications,” http://standards.ieee.org/ getieee802/ download/802.5-1998.pdf.

[21] L. Harte, A. Smith and C. A. Jacobs, “IS-136 TDMA technology, economics, and services,” Boston : Artech House, 1998.

[22] “Carrier Sense Multiple Access with Collision Avoidance,” http://en.wikipedia.org/ wiki/CSMA/CA.

[23] “Carrier Sense Multiple Access,” http://en.wikipedia.org/wiki/Carrier_Sense_Multiple_ Access.

[24] IEEE Std 802.11-1999, “Local and Metropolitan Area Network, Specific Requirements, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,” http://standards.ieee.org/ getieee802/download/802.11-1999.pdf.

[25] D. Bertsekas and R. Gallager, “Multiaccess Communication,” *Ch. 4, Data Networks, 2nd Edition*, Prentice Hall PTR, Upper Saddle River, New Jersey, 1992.

[26] IEEE 802.11 Group, “Official IEEE 802.11 Working Group Project Timelines,” http://grouper.ieee.org/ groups/802/11/Reports/802.11_Timelines.htm.

[27] K. Cheun and W. Stark, "Probability of Error in Frequency-hop Spread-Spectrum Multiple-Access Communication Systems with Noncoherent Reception," IEEE Trans. Commun., vol.39, no.9, pp.1400–1410, Sept. 1991.

[28] K. Cheun and K. Choi, "Performance of FHSS Multiple-Access Networks Using MFSK Modulation," IEEE Trans. Commun., vol.44, no.11, pp.1514–1526, Nov. 1996.

[29] "An Introduction to Direct-Sequence Spread-Spectrum Communications," http://www. Maxim-ic.com/ appnotes.cfm/appnote_number/1890/.

[30] R. L. Peterson, R. E. Ziemer, and D. E. Borth, "Introduction to Spread-Spectrum Systems," Englewood Cliffs, NJ: Prentice-Hall, 1995.

[31] K. K. Maureen, and M. E. Novak, "An Introduction to Infrared Technology: "Applications in the Home, Classroom, Workplace, and Beyond ..., " http://trace.wisc.edu/docs/ir_intro/ir_intro.htm.

[32] J. A. Harrington, "Infrared Fibers," in *Handbook of Optics: Fiber and Integrated Optics*, Vol. 4, McGraw-Hill, New York, 2000.

[33] R. Prasad, OFDM for Wireless Communications Systems, Artech House, ISBN: 1-58053-796-0, 2004.

[34] J. Heiskala and J. Terry, OFDM Wireless LANs: A Theoretical and Practical Guide, Sams, December 2001.

[35] Narayanaswamy, S., Kawadia, V., Sreenivas, R. S., Kumar, P. R., "Power Control in Ad-Hoc Networks: Theory, Architecture, Algorithm and Implementation of the COMPOW Protocol," *European Wireless*, pp. 156-162, 2002.

[36] J. W. Mark and W. Zhuang, "Wireless Communications and Networks", Pearson Education, Inc., 2003.

[37] H. T. Friis, "A note on a simple transmission formula," *Proc. IRE*, 34, 1946.

[38] T. S. Rappaport. *Wireless communications, principles and practice*. Prentice Hall, 1996.

[39] M. Zorzi and R. Rao, "Capture and retransmission control in mobile radio," *IEEE Journal on Selected Areas in Communications*, 12(8), pp. 1289-1298, 1994.

[40] Yang, X., and Vaidya, N. "On the Physical Carrier Sense in Wireless Ad Hoc Networks," *IEEE INFOCOM*, 2005.

[41] K. Pahlavan and P. Krishnamurthy, "Network Planning. Ch. 5," *Principles of Wireless Networks*, Prentice Hall PTR, Upper Saddle River, New Jersey, 2002.

[42] M. Grossglauser and D. Tse, "Mobility Increases the Capacity of Ad-hoc Wireless Networks," *IEEE INFOCOM*, 2001.

[43] The *ns* Manual, http://www.isi.edu/nsnam/ns/doc/ns_doc.pdf.

[44] D. B. Johnson and D. Maltz, "Dynamic Source Routing in Ad-Hoc Wireless Networks," Ch. 5, *Mobile Computing*, edited by Imielinski, T., and Korth, H., Kluwer Academic Publishers, pp. 153-181, 1996.

[45] C. E.. Perkins and E. Royer, "Ad-hoc On-Demand Distance Vector Routing," *IEEE Workshop on Mobile Computing Systems and Applications*, pp. 90-100, 1999.

[46] A. Valera, W. Seah,and S. V. Rao, "Cooperative Packet Caching and Shortest Multipath Routing In Mobile Ad hoc Networks.," *IEEE INFOCOM*, 2003.

[47] D. S. J. Couto, D. Aguayo, B. A. Chambers and R. Morris, "Performance of Multihop Wireless Networks: Shortest Path is Not Enough," *The First Workshop on Hot Topics in Networks (HotNets-I)*, 2002.

[48] S. Biswas and R. Morris, "Opportunistic Routing in Multi-Hop Wireless Networks," *HotNets-II*, 2003.

[49] B. Blum, T. He, S. Son and J. Stankovic, "IGF: A State-Free Robust Communication Protocol for Wireless Sensor Networks," *Technical Report CS-2003-11*, University of Virginia, April 21, 2003.

[50] T. He, J. Stankovic,, C. Lu and T. F. Abdelzaher, "SPEED: A Stateless Protocol for Real-Time Communication in Sensor Networks.," *IEEE ICDCS*, May 2003.

[51] P. Skraba,, H. Aghajan and A. Bahai, "Distributed Passive Routing Decisions in Mobile Ad-Hoc Networks.," *IEEE Vehicular technology Conference (VTC'04)*, Sept. 2004.

[52] M. Zorzi and R. Rao, "Geographic Random Forwarding (GeRaF) for Ad hoc and Sensor Networks: Energy and Latency Performance," *IEEE Tran. Mobile Computing*, 2(4), 2003.

[53] B. Karp, "Geographic Routing for Wireless Networks.," Ph.D. Dissertation, Harvard University, October, 2000.

[54] S. Xu,and T. Saadawi, "Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks? "*IEEE Communications Magazine*, pp.130-137, Jun. 2001.

[55] L. E Miller, "Carrier Sense Threshold/Range Control: Compilation of MANET email messages," http://www.antd.nist.gov/wctg/manet/docs/carriersense.pdf, Jan. 2004.

[56] S. R.Das, C. E. Perkins and E. M. Royer, "Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks," *IEEE INFOCOM*, 2000.

[57] J. Bicket, D. Aguayo, S. Biswas and R. Morris, "Architecture and Evaluation of an Unplanned 802.11b Mesh Network," *IEEE/ACM MobiCom*, 2005.

[58] R. Draves, J. Padhye and B. Zill, “Comparison of Routing Metrics for Static Multi-Hop Wireless Networks,” *ACM SIGCOMM*, 2004

[59] J. P. Monks, V. Bharaghavan and W. W. Hwu, “A Power Controlled Multiple Access Protocol for Wireless Packet Networks,” *Proceedings of the Conf. on Computer Communications* (IEEE INFOCOM 2001), 2001

[60] J. P. Monks, J.-P. Ebert, W.-M. W. Hwu and A. Wolisz, “Energy saving and capacity improvement potential of power control in multi-hop wireless networks,” *Computer Networks*, pp. 313–330, 41 (2003), 2003.

[61] C. Yu, K. G. Shin, and B. Lee, “Power-Stepped Protocol: Enhancing Spatial Utilization in a Clustered Mobile Ad Hoc Network,” *IEEE Journal of Selected Areas in Communications* (J-SAC), Vol. 22, No. 7, pp. 1322-1334, Sep. 2004.

[62] Y.B. Ko, V. Shankarkumar and N.H. Vaidya, “Medium Access control protocols using directional antennas in ad hoc networks,” *In Proceedings of INFOCOM*, 2000.

[63] M. Takai, J. Martin, R. Bagrodia, and A. Ren, “Directional virtual carrier sensing for directional antennas in mobile ad hoc networks,” *In Proceedings of ACM MOBIHOC*, 2002.

[64] Y. Wang and J. J. Garcia-Luna-Aceves, “Spatial Reuse and Collision Avoidance in Ad Hoc Networks with Directional Antennas,” *http://www.soe.ucsc.edu/ccrg/publications/ywang.globecom02.pdf*, Mar. 2005.

[65] T. Korakis, G. Jakllari, and L. Tassiulas, “A MAC protocol for full exploitation of Directional Antennas in Ad-hoc Wireless Networks,” *In Proceedings of ACM MOBIHOC*, 2003.

[66] J.-S. Park, A. Nandan, M. Gerla and H. Lee, “SPACE-MAC Enabling Spatial Reuse Using MIMO Channel-aware MAC,” *http://www.cs.ucla.edu/NRL/wireless/uploads/spacemac_icc.pdf*, Mar. 2005.

[67] F. Ye, S. Yi, and B. Sikdar, “Improving Spatial reuse of IEEE 802.11 Based Ad Hoc Networks,” *IEEE GLOBECOM*, 2003.

[68] J. Fuemmeler, N. Vaidya, and V. Veeravalli, “Selecting Transmit Powers and Carrier Sense Thresholds for CSMA Protocols,” *Technique report, University of Illinois at Urbana-Champaign*, October 2004.

[69] J. Zhu, X. Guo, L. L. Yang, and W. S. Conner, “Leveraging Spatial Reuse in 802.11 Mesh Networks with Enhanced Physical Carrier Sensing,” *IEEE ICC*, pp. 4004–4011, 2004.

[70] J. Zhu, X. Guo, L. L. Yang, W. S. Conner, S. Roy and M. M. Hazra, “Adapting physical carrier sensing to maximize spatial reuse in 802.11 mesh networks,” *Wireless Communications and Mobile Computing*, (4), pp. 933-946, 2004.

[71] X. Yang and N. Vaidya, “On the Physical Carrier Sense in Wireless Ad Hoc Networks,” *IEEE INFOCOM*, 2005.

[72] H. Zhai and Y. Fang, “Physical Carrier Sensing and Spatial Reuse in Multirate and Multihop Wireless Ad Hoc Networks,” *IEEE INFOCOM*, 2006 (to appear).

[73] J. Zhu, B. Metzler, X. Guo, and Y. Liu, “Adaptive CSMA for Scalable Network Capacity in High-Density WLAN: a Hardware Prototyping Approach, *IEEE INFOCOM*, 2006.

[74] D. M. Chiu and R. Jain, “Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks”, *Computer Networks and ISDN Systems*, vol. 17, pp. 1–14, 1989.

[75] R Jain, “The Art of Computer Systems Performance Analysis”, John Wiley and Sons, 1991.

[76] T. Nadeem, L. Ji, A. Agrawala and J. Agre, “Location Enhancement to IEEE 802.11 DCF”, *In Proceedings of IEEE INFOCOM*, March 2004.

[77] C. Yu, K. G. Shin, and L. Song, "Maximizing Communication Concurrency via Link-Layer Packet Salvaging in Mobile Ad Hoc Networks," IEEE Trans. Mobile Computing (to appear).

[78] L. Song and C. Yu, "Improving Spatial Reuse with Collision-Aware DCF in Mobile Ad Hoc Networks," 35th International Conference on Parallel Processing (ICPP), 2006.

[79] C. Yu, K. G. Shin, and L. Song, "Link-Layer Salvaging for Making Routing Progress in Mobile Ad Hoc Networks," The Sixth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), 2005.

[80] L. Song and C. Yu, “Minimizing Spatial and Time Reservation with Collision-Aware DCF in Mobile Ad Hoc Networks,” submitted to Ad Hoc Networks Journal.

APPENDIX A. AVERAGE END-TO-END DISTANCE IN MANETS

In the MANET, nodes randomly move around. It is reasonable to assume that nodes are uniformly distributed in the network. To simplify the problem (it does not affect the validation of the analysis in Chapter III), suppose the network is an $L \times L$ network. Nodes are randomly distributed. What is the average distance between two nodes?

Suppose the coordinates of the two nodes are $(\mathbf{X_1}, \mathbf{Y_1})$ and $(\mathbf{X}_2, \mathbf{Y}_2)$ respectively, where $\mathbf{X_1}$, $\mathbf{Y_1}$, $\mathbf{X_2}$, and $\mathbf{Y_2}$ can be considered as independent random variables with uniform distribution $U(0, L)$. The problem is to find the mean value of a random variable $\mathbf{Z}$, $E\{\mathbf{Z}\}$, where

$$\mathbf{Z} = \sqrt{(\mathbf{X_1} - \mathbf{X_2})^2 + (\mathbf{Y_1} - \mathbf{Y_2})^2} \tag{A.1}$$

The *Probability Distribution Function* (PDF) and *Cumulative Distribution Function* (CDF) of $\mathbf{X_1}$, $\mathbf{Y_1}$, $\mathbf{X_2}$, and $\mathbf{Y_2}$ are as follows.

$$f_{\mathbf{X}}(x) = \frac{1}{L} \qquad \text{where } 0 \le x \le L \tag{A.2}$$

$$F_{\mathbf{X}}(x) = \frac{x}{L} \qquad \text{where } 0 \le x \le L \tag{A.3}$$

Use two temporary random variables $\mathbf{U}$ and $\mathbf{V}$ where

$$\mathbf{U} = |\mathbf{X_1} - \mathbf{X_2}| \qquad \text{and} \qquad \mathbf{V} = |\mathbf{Y_1} - \mathbf{Y_2}| \tag{A.4}$$

So $\mathbf{Z} = \sqrt{\mathbf{U}^2 + \mathbf{V}^2}$. Consider the random variable $\mathbf{U}$, its CDF becomes

$$F_{\mathbf{U}}(u) = \Pr ob(\mathbf{U} \le u) = \Pr ob(\left|\mathbf{X_1} - \mathbf{X_2}\right| \le u) \tag{A.5}$$

Since both of $\mathbf{X_1}$ and $\mathbf{X_2}$ satisfy the same uniform distribution $U(0, L)$, as shown in the Figure 39, $F_U(u)$ is equal to the ratio of the shaded area to the $L \times L$ rectangular area.

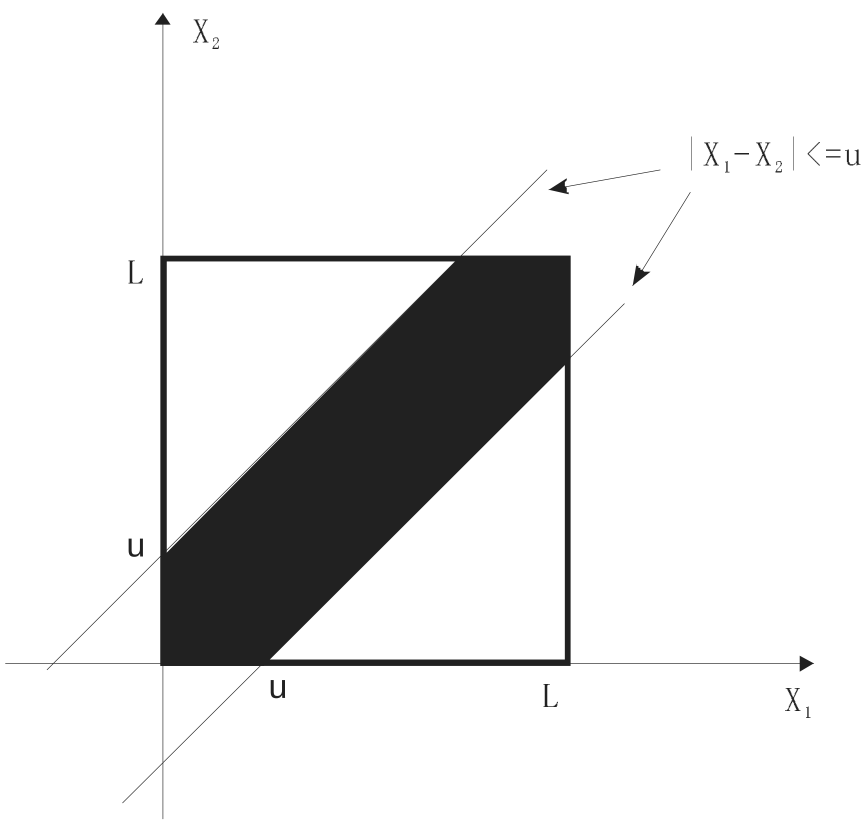


Figure 39: Probability of $\left|\mathbf{X_1} - \mathbf{X_2}\right| \le u$

Therefore,

$$F_{\mathbf{U}}(u) = \frac{L^2 - (L-u)^2}{L^2} \qquad 0 \le u \le L \tag{A.6}$$

$$f_{\mathbf{U}}(u) = \frac{2(L-u)}{L^2} \qquad 0 \le u \le L \qquad \text{(A.7)}$$

Similarly,

$$F_{\mathbf{V}}(v) = \frac{L^2 - (L-v)^2}{L^2} \qquad 0 \le v \le L \qquad \text{(A.8)}$$

$$f_{\mathbf{V}}(v) = \frac{2(L-v)}{L^2} \qquad 0 \le v \le L \qquad \text{(A.9)}$$

Now, consider $\mathbf{Z} = \sqrt{\mathbf{U}^2 + \mathbf{V}^2}$, its PDF becomes

$$F_{\mathbf{Z}}(z) = \Pr ob(\mathbf{Z} \le z) = \Pr ob(\sqrt{\mathbf{U}^2 + \mathbf{V}^2} \le z) \qquad 0 \le u, v \le L \qquad \text{(A.10)}$$

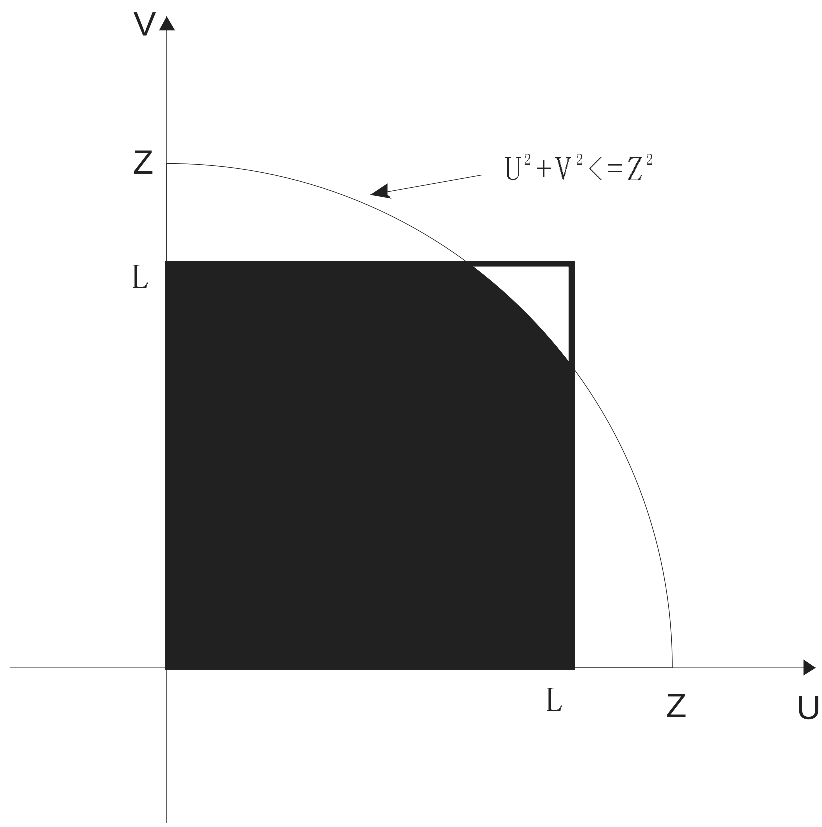


Figure 40: Probability of $\sqrt{\mathbf{U}^2 + \mathbf{V}^2} \le z$

As shown in the Figure 40, do integration over $\mathbf{U}$ and $\mathbf{V}$, the PDF of $\mathbf{Z}$ can be obtained as

$$F_{\mathbf{Z}}(z) = \Pr ob(\mathbf{Z} \le z) = \Pr ob(\sqrt{\mathbf{U}^2 + \mathbf{V}^2} \le z) = \iint f_{\mathbf{U}}(u) f_{\mathbf{V}}(v) du dv$$

$$= \begin{cases} 0 & z < 0 \\ \int_0^z \int_0^{\sqrt{z^2-u^2}} f_{\mathbf{U}}(u) f_{\mathbf{V}}(v) du dv & 0 \le z < L \\ \int_0^{\sqrt{z^2-L^2}} \int_0^L f_{\mathbf{U}}(u) f_{\mathbf{V}}(v) du dv + \int_{\sqrt{z^2-L^2}}^L \int_0^{\sqrt{z^2-u^2}} f_{\mathbf{U}}(u) f_{\mathbf{V}}(v) du dv & L \le z < \sqrt{2}L \\ 1 & z \ge \sqrt{2}L \end{cases}$$

$$= \begin{cases} 0 & z < 0 \\ \frac{1}{L^4}\left(\frac{1}{2}z^4 - \frac{8}{3}Lz^3 + \pi L^2 z^2\right) & 0 \le z < L \\ \frac{2L\sqrt{z^2-L^2} - z^2 + L^2}{L^2} + \frac{1}{L^4}\left[2L^2 z^2 \left(\arcsin\frac{L}{z} - \arcsin\frac{\sqrt{z^2-L^2}}{z} - \frac{1}{2} \right) + 2Lz^2\sqrt{z^2-L^2} + \frac{2}{3}L\left[(z^2-L^2)^{\frac{3}{2}} - L^3 \right] - \frac{1}{2}z^4 \right] & L \le z < \sqrt{2}L \\ 1 & z \ge \sqrt{2}L \end{cases}$$

$$
= \begin{cases} 0 & z < 0 \\ \dfrac{1}{L^4}\left(\dfrac{1}{2}z^4 - \dfrac{8}{3}Lz^3 + \pi L^2 z^2\right) & 0 \le z < L \\ \dfrac{2L\sqrt{z^2 - L^2} - z^2 + L^2}{L^2} + \dfrac{1}{L^4}\left[\begin{array}{l} 2L^2 z^2 \left(2\arcsin\dfrac{L}{z} - \dfrac{\pi}{2} - \dfrac{1}{2}\right) + \\ 2Lz^2\sqrt{z^2 - L^2} + \dfrac{2}{3}L\left[(z^2 - L^2)^{\frac{3}{2}} - L^3\right] - \dfrac{1}{2}z^4 \end{array}\right] & L \le z < \sqrt{2}L \\ 1 & z \ge \sqrt{2}L \end{cases}
$$

(A.11)

So,

$$
f_Z(z) = \begin{cases} 0 & z < 0 \\ \dfrac{2}{L^4}(z^3 - 4Lz^2 + \pi L^2 z) & 0 \le z < L \\ \dfrac{2z}{L^2}\left(\dfrac{L}{\sqrt{z^2 - L^2}} - 1\right) + \dfrac{1}{L^4}\left[\begin{array}{l} 4L^2 z\left(2\arcsin\dfrac{L}{z} - \dfrac{\pi}{2} - \dfrac{1}{2}\right) + \\ 8Lz\sqrt{z^2 - L^2} - \dfrac{2L^3 z}{\sqrt{z^2 - L^2}} - 2z^3 \end{array}\right] & L \le z < \sqrt{2}L \\ 0 & z \ge \sqrt{2}L \end{cases}
$$

$$
= \begin{cases} 0 & z < 0 \\ \dfrac{2}{L^4}(z^3 - 4Lz^2 + \pi L^2 z) & 0 \le z < L \\ -\dfrac{2z}{L^2} + \dfrac{1}{L^4}\left[\begin{array}{l} 4L^2 z\left(2\arcsin\dfrac{L}{z} - \dfrac{\pi}{2} - \dfrac{1}{2}\right) \\ + 8Lz\sqrt{z^2 - L^2} - 2z^3 \end{array}\right] & L \le z < \sqrt{2}L \\ 0 & z \ge \sqrt{2}L \end{cases}
$$

(A.12)

Therefore, the average distance is

$$
d = E\{\mathbf{Z}\} = \int_{-\infty}^{+\infty} z \cdot f_{\mathbf{Z}}(z) dz
$$

$$
= \int_0^L z \cdot \frac{2}{L^4}(z^3 - 4Lz^2 + \pi L^2 z) dz - \frac{2}{L^2}\int_L^{\sqrt{2}L} z^2 dz
$$
$$
+ \frac{1}{L^4}\int_L^{\sqrt{2}L} z \cdot \left[4L^2 z\left(2\arcsin\frac{L}{z} - \frac{\pi}{2} - \frac{1}{2}\right) + 8Lz\sqrt{z^2 - L^2} - 2z^3\right] dz
$$

$$
= \int_0^L z \cdot \frac{2}{L^4}(z^3 - 4Lz^2 + \pi L^2 z) dz - \frac{2}{L^2}\int_L^{\sqrt{2}L} z^2 dz + \frac{8}{L^2}\int_L^{\sqrt{2}L} z^2 \cdot \arcsin\frac{L}{z} dz
$$
$$
- \frac{2}{L^2}\int_L^{\sqrt{2}L} z^2 \cdot (\pi + 1) dz + \frac{8}{L^3}\int_L^{\sqrt{2}L} z^2 \cdot \sqrt{z^2 - L^2} dz - \frac{2}{L^4}\int_L^{\sqrt{2}L} z^4 \cdot dz
$$

$$
= \int_0^L z \cdot \frac{2}{L^4}(z^3 - 4Lz^2 + \pi L^2 z) dz - \frac{2}{L^2}\int_L^{\sqrt{2}L} z^2 \cdot (\pi + 2) dz
$$
$$
+ \frac{8}{L^2}\int_L^{\sqrt{2}L} z^2 \cdot \arcsin\frac{L}{z} dz + \frac{8}{L^3}\int_L^{\sqrt{2}L} z^2 \cdot \sqrt{z^2 - L^2} dz - \frac{2}{L^4}\int_L^{\sqrt{2}L} z^4 \cdot dz
$$

$$
= \left(\frac{\pi}{3} - \frac{4}{5}\right)2L - \frac{2}{3}(\pi + 2)(2\sqrt{2} - 1)L + \frac{4}{3}\left((\sqrt{2} - 1)\pi + \sqrt{2} - \ln(\sqrt{2} - 1)\right)L
$$
$$
+ \left(3\sqrt{2} - \ln(\sqrt{2} + 1)\right)L - \frac{2}{5}(4\sqrt{2} - 1)L
$$

$$= \frac{1}{15}\left[2 + \sqrt{2} - 5\ln(\sqrt{2} - 1)\right] \cdot L$$

$$\approx 0.616L$$

(A.13)

Note:

$$\int z^2 \cdot \sqrt{z^2 - L^2}\, dz = \frac{z}{4}\sqrt{(z^2 - L^2)^3} + \frac{L^2 z}{8}\sqrt{z^2 - L^2} - \frac{L^4}{8}\ln\left(z + \sqrt{z^2 - L^2}\right)$$

$$\int z^2 \cdot \arcsin\frac{L}{z}\, dz = \frac{L^3}{3}\left(\frac{z^3}{L^3}\arcsin\frac{L}{z} + \frac{z\sqrt{z^2 - L^2}}{2L^2} - \frac{1}{2}\ln\sqrt{\frac{z - \sqrt{z^2 - L^2}}{z + \sqrt{z^2 - L^2}}}\right)$$