NUI MAYNOOTH

Ollscoil na hÉireann Má Nuad

Power Evaluation and Performance Enhancement of CSMA/CA based WLANs

by

Minyu Fang

Masters Thesis

Hamilton Institute

National University of Ireland, Maynooth

Co. Kildare

June 2010

Head of Department: Prof. Douglas Leith

Research Supervisor: Dr. Ken R. Duffy

Contents

List of Figures

Abstract

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) based Wireless Local Area Networks (WLANs) are becoming pervasive. As the most commonly employed standard in WLANs, IEEE 802.11 not only gives rise to health and safety concern from the general public, but also has the potential for enhanced performance. Our contributions in this thesis are twofold: (1) We extend a recently introduced model of transmitted power in WLANs to cover unsaturated conditions when stations do not always have packets to send. we implement an experimental verification of the original analytic model and the extended one. We confirm the estimated maximum power is substantially lower than the International Commission on Non-Ionizing Radiation Protection (ICNIRP) limit; (2) We propose two decentralized Multiple Access Control (MAC) schemes that converge to collision-free schedules almost surely and therefore improve throughput performance. In adopting decentralized learning techniques, the convergence times of both schemes are brief. Decentralized schedule length adaption is introduced to provide long-run fair access to the medium and scalability of the MAC schemes to networks of any size.

Acknowledgements

I would like to thank my supervisor Dr. Ken R. Duffy for his constant support, incredible patience and insightful guidance. He provides me such a priceless opportunity to do my research in Hamilton Institute. During the past one-and-half years, I have learned a lot from him, not only knowledge and technical skills, but also the most importantly rigorous scientific attitude.

I would also like to thank my co-supervisor Dr. David Malone for his warm help, constructive suggestions and comments on every aspect of work and life. I am grateful to Tianji Li, Qizhi Cao and Xiaomin Chen for their help and support. I am indebted to Rosemary Hunt, Kate Moriarty and Ruth Middleton for their assistant.

Thanks also go to my parents, Jianfei Fang and Lianjun Chen for their selfless love. Finally, special thanks go to my fiancee Xi Yu who made all of this possible.

Chapter 1

Introduction

Wireless networks for local area communication have become increasingly ubiquitous during recent years, and have received considerable research attention [1][2]. A study group 802.11 was formed under IEEE (Institute of Electrical and Electronics Engineers) project 802 to recommend an international standard for Wireless Local Area Networks (WLANs), which seek to ensure standardized protocols across different manufacturers. In WLANs, the Multiple Access Control (MAC) uses a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol with binary exponential backoff to regulate how participants communicate via a common physical medium. In this thesis, we consider how to experimentally evaluate transmitted power for WLANs based on mathematical models of the 802.11 MAC in saturated conditions [3][4] and unsaturated conditions [5], as well as how to improve network performance for CSMA/CA based WLANs.

1.1 An overview of WLANs

A series of CSMA/CA based standards [6] are specified by IEEE to define the medium access function, which is primarily designed for WLANs but also encompassed wireless mesh networks. A brief overview of theses standards is given in this section.

1.1.1 IEEE 802.11

Wireless is a broadcast medium, with transmissions received by every station for whom the transmitter is within its carrier sensing range. IEEE 802.11 wireless networks with current hardware cannot support transmission and reception at the same

time, because the difference between the power of reception and of transmission of a packet is significant. Each station has to share time on the medium with those in its local area. This medium-sharing scheme is referred to as Carrier Sense Multiple Access (CSMA), which is a listen-before-talk protocol. That is, a station cannot transmit until it senses the channel is idle. Collisions take place when more than one station starts to transmit simultaneously after sensing the medium to be idle. Failure to receive an acknowledgement (ACK) within a specified ACK timeout period will cause the source station to retransmit the packet. A scheme called CA (Collision Avoidance), which we will explain in detail later, is employed to decrease the likelihood of the stations colliding repeatedly in the presence of contention.

IEEE 802.11 MAC defines two different access mechanisms, the mandatory Distributed Coordination Function (DCF) which provides channel access based on decentralized CSMA/CA principles, and the optional Point Coordination Function (PCF) which provides centrally controlled channel access through polling.

DCF access scheme

DCF is the basic access mechanism of 802.11 and is based on CSMA/CA. Before transmitting a frame, the station senses the medium (carrier sensing). If the medium is found idle for at least a DCF Inter-Frame Space (DIFS) time period ($34\mu s$ in 802.11a, $50\mu s$ in 802.11b and $28\mu s$ in 802.11g), the station starts to transmit if no other stations are transmitting. The other stations sense the medium is busy and wait until it becomes idle at least for the DIFS time period before attempting to win access. Two methods are employed to determine whether the medium is busy or not. With physical carrier sensing, the wireless medium is sensed at the physical layer. On the other hand, virtual carrier sensing works at the MAC layer. If a station receives a frame that is not destined for it, it checks the duration field in the frame header, which records the time of one successful transmission (including ACK, frame spacing, etc), and then defers the access to the medium for the duration time period. When the destination station receives a frame, it sends back an ACK frame to the source station after a Short Inter-Frame Space (SIFS) time period ($16\mu s$ in 802.11a, $20\mu s$ in 802.11b, and $10\mu s$ in 802.11g). SIFS is shorter than DIFS, so no other stations can gain access to the medium. A typical operation is illustrated in Fig 1.1, where "*data*" includes MAC header, physical header and a packet.

The above example leads to collisions if more than one station senses the medium

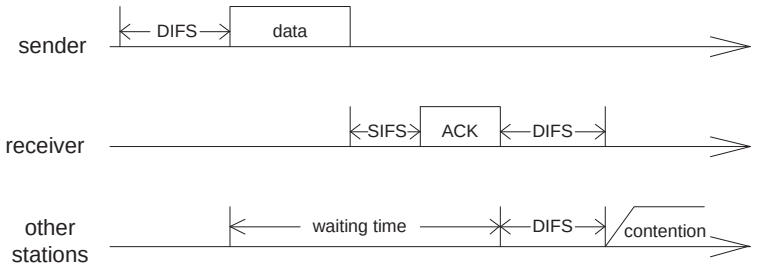


Figure 1.1: Basic DCF access scheme

is idle for the DIFS period and start to transmit at the same time. In order to avoid such collisions, stations have to wait for an additional time beyond DIFS before transmitting if the medium is sensed idle. This additional time is random, referred to the backoff time, which is the time of integer backoff counter multiplying the idle slot of fixed length $\sigma\mu s$ ($9\mu s$ in 802.11a, $20\mu s$ in 802.11b and $9\mu s$ in 802.11g).

CA regulates this additional random delay to help to avoid collisions, otherwise all stations would try to transmit and collide with each other again as soon as the medium becomes idle for the next DIFS period. As the medium is sensed idle for a DIFS period, the station starts to decrease its backoff counter by one for each idle slot time. If the medium becomes busy during this backoff process, it pauses its backoff counter, and resumes it after sensing the idle medium for the DIFS time period. When the counter decreases to zero, the station starts to transmit, after which the station having packets to send, regardless of success or failure, starts a new backoff process by uniformly choosing a new random backoff value to update its backoff counter from the interval $[0, CW - 1]$, called the current Contention Window. At the first transmission attempt or after a successful transmission, CW is set at its minimal value CW_{min}. After each failure, CW is doubled until it reaches its maximum value CW_{max}. For the most commonly used physical layer, 802.11b and the Direct Sequence Spread Spectrum (DSSS) PHY, CW_{min} and CW_{max} are set at 32 and 1024 respectively. DCF also specifies a maximum number of retransmissions for one packet as retry limit. If the number equals the retry limit, the packet is discarded, and CW is reset to CW_{min}.

When a sending station does not receive an ACK in the ACK timeout period, it assumes that a collision has occurred and enters into the backoff stage again after the medium is sensed idle for the DIFS period, such that the new backoff value is chosen uniformly at random using a CW that is twice the previous CW.

Due to random chosen backoff value for all stations, the DCF scheme does not completely eliminate the collisions. There is a persistent non-zero likelihood of collisions [3] even after a long period operation.

PCF access scheme

In PCF, a Point Coordinator (PC), which is typically located at the Access Point (AP), manages the medium access via a polling scheme, such that the PC polls individual station to assign access to the medium depending on their requirements. It combines a contention period with a contention free period, at the beginning of which the PC sends a beacon frame. All stations listen to this frame containing the maximum duration of the contention free period. All contention free transmissions are separated only by SIFS and the PCF interframe space. Both of these are shorter than the DIFS, so DCF stations cannot gain access to the medium. The PC can also send a beacon to end the contention free period. Stations receiving this will then revert to using the DCF scheme. As in the PCF scheme, stations do not contend for the medium and instead transmit at the allocated time controlled centrally by PC, the access scheme is referred as contention-free channel access. The disadvantage of this scheme is its centralized nature. It has polling overheads and is less robust to partial information. Consequently, it is not implemented on most commodity platforms. It also requires the network to work in an infrastructure mode with an AP and stations, and so cannot be applied to other wireless networks such as an wireless multi-hop mesh network.

1.1.2 other standards

The maximum PHY rate of the original protocol 802.11 [6] is only 2 Mbps, which is too slow to meet the requirements of modern applications. Therefore, to support higher throughput, a series of extensions to the original 802.11 protocol are introduced in PHY layer which leave the MAC layer unchanged. For 802.11b [7] and 802.11g [8], a PHY layer standard in 2.4 GHz radio band is specified. Three orthogonal channels are supported and maximum rates are 11 Mbps and 54 Mbps per channel respectively. In addition to this, 802.11b employs complementary code keying (CCK) modulation, while 802.11g also uses orthogonal frequency division multiplexing (OFDM) modulation as well. IEEE 802.11a [9] specifies a 5 GHz radio band in PHY layer, and is not interoperable with 802.11b/g. It also supports three orthogonal channels with the

AC	CW_{min}	CW_{max}	AIFSN	TXOP limit
AC_BK	CW_{min}	CW_{max}	7	0
AC_BE	CW_{min}	CW_{max}	3	0
AC_VI	$CW_{min}/2$	CW_{max}	2	3.008ms
AC_VO	$CW_{min}/4$	$CW_{max}/2$	2	1.504ms

Table 1.1: Default ACs Parameter values

maximum rate of 54 Mbps per channel based on OFDM modulation.

IEEE 802.11e [10] aims to support quality of service (QoS) by making changes to the MAC layer of 802.11. It defines two access schemes: the decentralized contention-based Enhanced Distributed Channel Access (EDCA) and the centrally controlled contention-free Hybrid Coordination Function Controlled Channel Access (HCCA). We only briefly describe the EDCA as the HCCA is not considered in this thesis.

For each station, four first-in-first-out (FIFO) queues called Access Categories (AC) are introduced to provide different priorities to access the medium. The four Access Categories are named AC_BK (background), AC_BE (Best effort), AC_VI (Video) and AC_VO (Voice), where AC_BK has the lowest and AC_VO has the highest priority. Incoming packets from upper layers are mapped into different ACs according to their priorities. Each AC behaves similarly to a single DCF station contending for the medium. The different priorities of them are specified by the relative CSMA/CA parameters CW_{min}, CW_{max}, AIFS (Arbitration Inter-Frame Space), and TXOP (transmission opportunity). CW_{min} and CW_{max} have been defined in the previous section as the minimal and maximum values of contention window. AIFS takes the role of DIFS with $AIFS = AIFSN \times \sigma + SIFS$. TXOP specifies the time duration an queue may transmit after winning access to the medium. TXOP is defined by a maximum duration called the TXOP limit. Multiple frames can be transmitted in a TXOP burst, if the transmission duration does not exceed the TXOP limit and the frames belong to the same queue. This process is called as contention free bursting (CFB), where the consecutive frames are separated by SIFS time periods instead of AIFS plus the post backoff period. The default values for all four ACs can be seen in Table 1.1. When queues at the same station attempt to transmit at the same time, a *virtual collision* occurs and the queue with higher priority is allowed to transmit.

1.2 The Contributions in this thesis

We make two contributions in this thesis. Firstly, IEEE 802.11 networks use radio frequency (RF) energy and regulations stipulating the maximum transmit power used by WLANs are set by the Federal Communication Commission (FCC) and European Conference of Postal and Telecommunications Administrations (CEPT). The International Commission on Non-Ionizing Radiation Protection (ICNIRP) [11] defined acceptable thresholds for absorbed radiation power at 80 mW kg^{-1} for the general public and 400 mW kg^{-1} for occupational exposure (whole body). Preliminary results suggested that exposure from standard deployments of WLANs are well within internationally accepted ICNIRP guidelines [12][13]. While WLANs have not attracted the same level of interest as mobile phone networks, there still exists public concern regarding health and safety issues, particularly in schools [14] but also in homes and offices [15][16]. The trend is toward denser Wi-Fi deployments, such as in hotspots in urban areas. A house or apartment could have ten WLANs devices, including broadband routers, laptops, phones, PDAs, games consoles and media players. It is therefore useful to model and evaluate how the radiated power scales with the number of stations and level of activity in order to determine if radiation levels are within acceptable limits. This information, based on analytic models, may be of use for both retrospectively assessing RF levels or for planning of future WLAN use, where measurement is not possible. In this thesis we introduce models for estimating the power output of WLANs in unsaturated conditions based on the existing power model in saturated conditions [17] and conduct an experimental verification of the the original analytic model and the extended one. Results confirm that the estimated maximum power is substantially lower than the acceptable limit given by ICNIRP. Parts of this work have been published in the journal Health Physics by Fang and Malone [18].

Our second contribution is related to improving network performance. As the most commonly employed MAC in WLANs, DCF and its 802.11e counterpart EDCA regulate the random backoff process to decrease the likelihood of the stations colliding again. In a network with more than one transmitter, a significant disadvantage of the DCF is that there is a persistent non-zero chance of collision. The network throughput is substantially degraded when a large number of stations are contending for the medium [3]. In this thesis we propose completely decentralized learning MAC schemes, which can be regarded as evolutions of the Learning Binary Expo-

nential Backoff (L-BEB) [19] and Zero Collision (ZC) [20] incorporating ideas from the self-managed decentralized channel selection algorithm [21]. We prove that these new access methods converge to a collision-free schedule, if one exists. By avoiding collisions, network throughput is significantly higher than DCF. In particular, reducing the convergence time to collision-free operation offers improved performance for delay-sensitive periodic traffic such as the voice and game data or when there are many stations and the IEEE 802.11 collision rate is likely to be large. Faster convergence times also accommodate changing network conditions. Finally, we propose a schedule length adaption scheme that is decentralized, but retains long-run fairness properties. This enables scalability of these new MACs to networks of any size.

1.3 Outline

This thesis is organized as follows. In Chapter 2, we present the analytical models of transmitted power for WLANs, and conduct the experimental verifications of them. In Chapter 3, learning MAC schemes are proposed to achieve collision-free access in WLANs. Chapter 4 shows the performance evaluation of these MAC schemes. Chapter 5 draws conclusions and discusses future work.

Chapter 2

Mathematical Modeling and Experimental Verification of Transmitted power of WLANs

2.1 Introduction

Wireless Local Area Networks (WLANs) are now common in many places. WLANs use Radio Frequency (RF) energy and regulations stipulating the maximum transmit power are set by the FCC (Federal Communications Commission) and CEPT (European Conference of Postal and Telecommunications Administrations). Signals are transmitted at low powers, typically 0.1 W for both computers and access points (APs) in the ISM (industrial, scientific and medical) band at 2.4 GHz or at 5 GHz. Although WLANs have not attracted the same level of interest as mobile phone networks, sometimes concern arises among the public regarding health and safety issues related to exposure to radio frequency (RF) energy, particularly in schools [14] but also in homes and offices [15][16].

The impact of electromagnetic fields on health is reviewed by the International Commission on Non-Ionizing Radiation Protection (ICNIRP) with a long literature review [11]. Acceptable thresholds for absorbed radiated power given by the ICNIRP

are 80mW kg^{-1} for the general public and 400mW kg^{-1} for occupational exposure. Experimental results [12][13] so far show that exposures from standard deployments of WLANs are well within internationally accepted ICNIRP guidelines. The chairman of the UK's Health Protection Agency (a body established to protect the public from environmental hazards, including non-ionizing radiation) has stated it would be timely to carry out further research as this technology is rolled out. The trend is toward denser WLANs deployments, such as extremely dense hotspots in urban areas. A recent report[1] released by a commercial firm shows that there are presently more than 289,000 public WLANs hot spots in operation in 140 countries. A house or apartment could have ten wireless devices, including broadband routers, laptops, phones, PDAs, games consoles and media players. Classrooms or conference halls could have larger numbers of devices with high levels of activity. It is therefore useful to model and evaluate how the radiated power scales with the number of stations and level of activity to determine if radiation levels are within acceptable international limits. Such models may be of use for both retrospectively assessing RF levels or for planning of future WLANs where measurement is not possible in advance.

Malone and Malone [17] proposed a model of total transmitted power which is a function of the number of stations in WLANs that are assumed to always be busy. In this thesis we consider how to estimate the power output of WLANs that are not always busy. As noted by various authors [12][13], WLANs transmissions are intermittent and time-averaged powers depend on the amount of data transferred. It is this issue that we consider. Factors such as the speed of broadband access links and the speed at which people can navigate the network serve to restrict how busy WLANs can become. For example, a photographer might send large files to clients each day, but the speed is restricted by a broadband link. Alternatively, someone watching YouTube videos will not tend to download faster than they can watch them. Of course, the wireless link may become congested in cases with fast links (e.g. a large school/campus) or where network transfers are local (e.g. backing data up to a local server). Thus we will look at both cases of unsaturated and saturated networks.

Most deployments of WLANs are based on the widely used IEEE 802.11 technology. In this thesis, we concentrate on WLANs based on the IEEE 802.11 standard and its amendments. The technology is popularly known by the name Wi-Fi and is supported by an industry group, the Wi-Fi Alliance. We do not consider other tech-

[1] http://www.jiwire.com/downloads/pdf/JiWire_MobileAudienceInsights_Q409.pdf

nologies such as WiMAX (Worldwide Interoperability for Microwave Access) or 3G. In the most typical WLAN scenario (known as an infrastructure mode network), every station communicates with an access point (AP) connected to the wired Internet. Widely deployed IEEE 802.11b and IEEE 802.11g networks operate in the unlicensed spectrum at 2.4 GHz; IEEE 802.11a utilizes spectrum around 5 GHz. We focus on 802.11b in the 2.4G Hz band, because it is supported by almost all existing hardware. Nonetheless, our theoretical results extend directly to 802.11a and 802.11g.

WLANs devices only transmit and radiate power when they have packets to send and when they are permitted to do so by the IEEE 802.11 protocol. The 802.11 Medium Access Control (MAC) protocol regulates the channel access scheme. We are interested in the impact of this MAC on transmitted power. If more than one device transmits at a time the result is a collision, where no data is successfully transferred. The MAC attempts to control transmissions so that there is a high likelihood that only one device transmits at a time. This is achieved by the MAC by inserting random time gaps (called backoff periods) after transmissions and collisions. Thus, the MAC has a significant impact on transmitted power by 802.11 devices, achieving a middle ground between all devices transmitting at once and just one device transmitting at a time. An IEEE 802.11 MAC model is proposed in [3] to predict the network performance such as throughput. From that model, the probability that a station attempts to transmit in a typical slot can be calculated, as can the probability that transmission results in a collision. Theses quantities are determined as a function of the number of stations and the network parameters, assuming that each device always has a packet to send. Based on this analytical model, Malone and Malone [17] estimate the mean transmitted power as a function of the number of stations under the same assumptions and study the total power emitted in error-free, error-prone, broadcast and unicast networks.

To consider unsaturated networks, we extend the model of power output beyond the saturated situation using the unsaturated models [5][22]. These models allow the amount of network traffic at each device to be varied. The models consider two extremes: Duffy et al. [22] assume that traffic arriving while the device is busy will be queued until it can be transmitted, whereas Malone et al. [5] assume such traffic is discarded without being transmitted. We call the first one the infinite buffer model and the latter the no buffer model. We also present experimental results to compare the predictions with the saturated model presented in [17] and the unsaturated models

that are described in this thesis. We apply these models to consider the power output associated in a number of scenarios.

Note that we calculate the sum of the power of all stations in the network, rather than the exposure at a particular point. As in [17] we omit some factors for calculating exposure, such as the distances between devices, reception errors caused by absorption/reflection in the environment or interference from other devices sharing the same frequency and so on. Similarly, we assume that maximum transmitter power approximates actual transmit power. These assumptions provide upper bounds of the transmitted power involved in exposure and are considered further in our discussion.

2.2 Modeling transmitted power

The 802.11 MAC defines two different access mechanisms, the mandatory Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The PCF provides centrally controlled channel access via a polling scheme, but is rarely used in practice. They are described in Section 1.1 of this thesis. The models we are interested in are of the DCF.

2.2.1 Bianchi's Model of the DCF

In [3][4], a mean field Markov Chain model was established to evaluate the performance of 802.11 DCF as a function of the number of station in saturated conditions. This model describes the behaviour of an 802.11 MAC as a discrete time Markov Chain and allows various important properties to be predicted, such as each station's throughput and their collision probabilities. Here we provide a detailed review of Bianchi's model as it is the model we use for the DCF in saturated conditions. We then briefly describe some of its extensions to this model developed by [22][5], and use these models into the extended models of transmitted power.

The Markov Chain [23] of interest to us is a discrete-time stochastic process with finite states. As it is a Markov Chain, the probability of choosing any given state only depends on the current state and is independent of the prior history. The long-term properties of Markov Chains are connected with its invariant distribution (its stationary distribution). Given any initial state, each state in the Markov Chain is visited with its stationary probability after the system has been running for a infinite number of iterations, if a stationary distribution exists.

In [3][5], a 2-dimensional stochastic process $(s(t), b(t))$ is introduced to model the backoff behaviour of the DCF scheme for a stations what always has a packet to send. Let $b(t)$ denote the station's backoff counter, which is decremented at the end of each slot. The slot used in this model represents either an idle slot, a slot with a successful transmission or a slot in which a collision occurs. For an idle slot, the duration lasts a physical slot time. For a successful transmission or collision slot, $b(t)$ is frozen for the duration of a transmission, and is subsequently decreased by one. When $b(t)$ reaches zero, a transmission occurs and a new backoff starts. If the result of this transmission is failure, the packet is retained for retransmission or discarded if the retry limit M is reached. On collision, the contention window doubles if it does not reach its maximum value W_m. If it reaches, W_m is then used as the new contention window size where m is the maximum backoff stage. The backoff value is randomly and uniformly from $\{0, \cdots, W_i - 1\}$ at backoff stage i. If the result of the transmission is successful, the contention window is reset to its minimal value W_0 and the backoff value is uniformly and randomly assigned from $\{0, \cdots, W_0 - 1\}$. This implies $b(t)$ depends on the history of collision. The variable $s(t) \in \{0, \cdots, M\}$ is used to record how many collisions the present packet has experienced, where the retry limit M is bigger than m, the maximum backoff stage.

Bianchi [3] makes certain judicious simplifying assumptions to create an analytically tractable model:

1. Failed transmissions only occur as a consequence of collision;

2. All stations are saturated, always having packets to send;

3. For any given station, conditional on transmission attempt, the probability of collision, p, is constant and independent of the station's collision history of the station and all other stations. In particular, it does not depend on the backoff stage at which the transmission is made.

4. In the mean field approximation the impact of each station is small compared to the entire WLAN, which generates the self-consistent equation. With this in mind we observe that the probability that each station does not experience a collision given that it is attempting to transmit is the probability that no other station is transmitting. Let $\tau(p)$ be the stations's stationary attempted transmission probability given the conditional collision probability is p and n

be the number of stations. This logic gives the following relationship:

$$1 - p = (1 - \tau(p))^{n-1}. \tag{2.1}$$

The assumptions 1, 2, 3 and 4 have been relaxed and investigated by subsequent research such as [24][25]. Under Bianchi's assumptions, the 2-dimensional stochastic process $(s(t), b(t))$ forms a discrete-time Markov Chain as illustrated in Figure 2.1, where the conditional collision probability p is assumed known and W_i is defined as the contention window size at the backoff stage i as

$$W_i = 2^{\min(i,m)} W_0. \tag{2.2}$$

In this Markov Chain, all states are aperiodic, recurrent and non-null, and thus the process is ergodic [23] and a stationary distribution exists. We define $\pi(s(t) = i, b(t) = k)$ to be the probability that the backoff counter is k for the backoff stage i at the time t. We let $b_{i,k} = \lim_{t\to\infty} \pi(s(t) = i, b(t) = k)$ be the stationary distribution of this Markov Chain at the state (i, k). Since each station only attempts to transmit when the backoff counter reaches zero, the transmission attempt probability $\tau(p)$ is

$$\tau(p) = \sum_{i=0}^{m} b_{i,0}. \tag{2.3}$$

We are interested in the $\tau(p)$ given the stationary probability $b_{i,k}$ of the Markov Chain. With a little elementary algebra, $b_{i,k}$ can be calculated explicitly and $\tau(p)$ is then obtained [3]. Providing more insight, as observed by Kumar, et al. [26][27] and reviewed by Duffy [28], $\tau(p)$ can be obtained directly as a deduction from the renewal reward theorem.

For a station $i \in \{1, \cdots, n\}$ in the WLAN, define $C_k := 1$ if the k^{th} transmission by station i results in a collision and $C_k := 0$ if it results in a success. By assumption 3, $P(C_k = 1) = p$. We also define U_j as the uniform distribution on $\{1, \cdots, W_j\}$ so that this variable counts how many idle slots a station observes before transmission rather than its backoff counter. Thus, viewing the number of transmission attempts for each packet as a reward associated with the renewal cycle of the number of slots a packet spends being processed in the MAC, we yield the following formula for the transmission probability for a given conditional collision probability p.

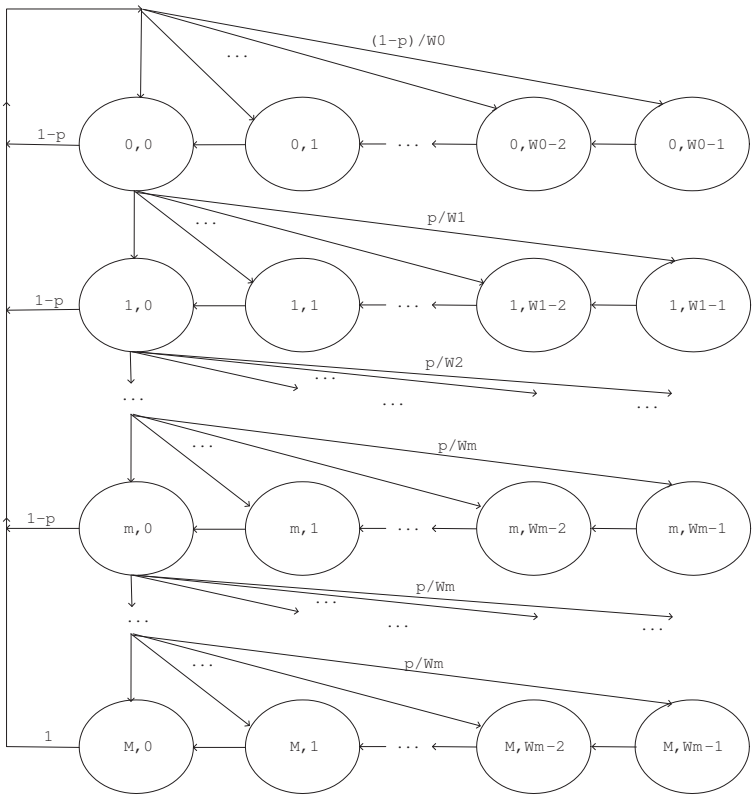


Figure 2.1: Markov Chain proposed in Bianchi's Model

$$\begin{aligned}\tau(p) &= \frac{E(1 + C_0 + C_0C_1 + \cdots + C_0C_1 \cdots C_M)}{E(U_0 + C_0U_1 + C_0C_1U_2 + \cdots + C_0C_1 \cdots C_MU_M)} \\ &= \frac{1 + p + p^2 + \cdots + p^M}{W_0/2 + pW_1/2 + \cdots + p^MW_M/2}. \end{aligned} \tag{2.4}$$

The renewal reward theorem is employed to illustrate that the model is insensitive to the backoff distribution at any given backoff stage. This means the uniform distribution $\{U_j\}$ only enter into the formula in equation (2.4) through their expectations, so that any other collection of backoff distribution with the same mean number of slots to be counted down, $W_j/2$, at each stage j, will give rise to the same stationary attempt probability $\tau(p)$. Combining with equation (2.2) and the network parameters, we simplify equation (2.4) into

$$\tau(p) = \frac{2(1 - p^{M+1})}{W_0(1 - p - p(2p)^m)/(1 - 2p) - W_0 2^m p^{M+1}}. \tag{2.5}$$

This is the expression used in [17] to calculate τ in saturated conditions.

2.2.2 Saturated Model of Transmitted Model

Based on the analysis of DCF's performance, the transmission probability τ can be calculated by combining equation (2.5) and equation (2.1) given the number of stations n and network parameters. If no stations transmit, then no energy is transmitted. If a single station transmits, then we will have a successful transmission. If $r \geqslant 2$ station transmits, we will have r simultaneous transmissions followed by a timeout while the transmission wait for an ACK (Acknowledgement) but do not receive one. Thus the average energy transmitted, as sum of these three contributions, during one slot time is obtained as below:

$$\begin{aligned}\mathcal{E} &= 0(1-\tau)^n + \mathcal{E}_s n\tau(1-\tau)^{n-1} + \mathcal{E}_c \sum_{r=2}^{n} rC_r^n \tau^r (1-\tau)^{n-r} \\ &= (\mathcal{E}_s - \mathcal{E}_c) n\tau(1-\tau)^{n-1} + \mathcal{E}_c n\tau,\end{aligned}$$

where C_r^n is the usual binomial coefficient, $\mathcal{E}_s$ and $\mathcal{E}_c$ are the mean energies consumed by a successful transmission and a collided transmission respectively. These quantities can be estimated by introducing the nominal power output P_0 (say 100mW) and the mean time the medium is busy for a successful transmission, $T_{\mathcal{E}_s}$, and collision transmission, $T_{\mathcal{E}_c}$, which exclude the idle-medium time such as SIFS (Short Interframe Space). They are given as below:

$$\mathcal{E}_s = P_0 T_{\mathcal{E}_s} = P_0(2\ preamble + (header + payload)/rate + ACK) \tag{2.6}$$

and

$$\mathcal{E}_c = P_0 T_{\mathcal{E}_c} = P_0(preamble + (header + payload)/rate + ACK), \tag{2.7}$$

where *preamble*, *header*, *payload* and *ack* are the times/sizes used for each of these transmissions and *rate* is the speed at which data is transmitted, see Table 2.1 for more details. The calculation of these quantities is described in detail in the appendix of [17], and depends on packet lengths, protocol constants and so on. Quantities such as the signal propagation delay also have a small impact. The values used in this thesis are shown in Table 2.1 mirroring 802.11b which is employed by our experimental testbed.

We also consider saturated broadcast traffic. The calculation for broadcast traffic

Physical rate=11Mbps Payload size = 1400 bytes preamble = $192\mu s$ header=28 bytes		
W_0	32	Contention Window
m	5	maximum Backoff stage
M	7	retry limit
SIFS	$10\mu s$	Short Interframe Space
δ	$1\mu s$	Propagation Delay
ACK	$202\mu s$	Acknowledgement
T_i	$20\mu s$	Idle Slot Time
T_s	$1515\mu s$	Average time of a Successful Transmission
T_c	$1281\mu s$	Average time of a Collision
$\mathcal{E}_s$	$145\mu s$	Mean Energy of a Successful Transmission
$\mathcal{E}_c$	$123\mu s$	Mean Energy of a collision

Table 2.1: MAC and PHY parameters mirroring 802.11b

only differs because an ACK is never sent in response to the reception of a packet. The value of collision time and relative energy remain unchanged. We obtain E_s as

$$\mathcal{E}_s = P_0 T_{\mathcal{E}_s} = P_0(2\ preamble + (header + payload)/rate). \tag{2.8}$$

The medium is idle, when none of stations are transmitting with the probability $(1-\tau)^n$. A successful transmission occurs, when only one station is transmitting with the probability $\tau(1-\tau)^{n-1}$. Then the probability that a collision occurs is $1-(1-\tau)^n - n\tau(1-\tau)^{n-1}$. Thus we give the average length of a slot time T as

$$T = T_i(1-\tau)^n + T_s n\tau(1-\tau)^{n-1} + T_c(1-(1-\tau)^n - n\tau(1-\tau)^{n-1}), \tag{2.9}$$

where T_i is the length of an idle state, T_s is the mean time for a successful transmission and T_c is the mean time for a collided transmission. They are easily calculated from the IEEE 802.11 standards and network settings.

Then the mean power is obtained by [17] as

$$P = \frac{\mathcal{E}}{T} = \frac{(\mathcal{E}_s - \mathcal{E}_c)n\tau(1-\tau)^{n-1} + \mathcal{E}_c n\tau}{T_i(1-\tau)^n + T_s n\tau(1-\tau)^{n-1} + T_c(1-(1-\tau)^n - n\tau(1-\tau)^{n-1})} \tag{2.10}$$

We can also give an expression for the duty cycle for RF energy from WLANs devices based on IEEE 802.11. The duty cycle (or duty factor) is a measure of the fraction of the time during which a RF is transmitting. It is an useful factor because it is related to peak and average power in the determination of total energy output. For WLANs it equals the fraction of the time during which at least one station is

transmitting. We obtain the duty cycle for the whole network as:

$$D_{net} = \frac{T_{\mathcal{E}_s} n\tau(1-\tau)^{n-1} + T_{\mathcal{E}_c}(1-(1-\tau)^n - n\tau(1-\tau)^{n-1})}{T_i(1-\tau)^n + T_s n\tau(1-\tau)^{n-1} + T_c(1-(1-\tau)^n - n\tau(1-\tau)^{n-1})}. \quad (2.11)$$

The duty cycle for a single station's activity will be

$$D_{sta} = \frac{T_{\mathcal{E}_s} \tau(1-p) + T_{\mathcal{E}_c} \tau p}{T_i(1-\tau)^n + T_s n\tau(1-\tau)^{n-1} + T_c(1-(1-\tau)^n - n\tau(1-\tau)^{n-1})}. \quad (2.12)$$

We note that the energy associated with a given station can be obtained by multiplying the duty cycle by the nominal power. The power for the network can then be obtained by summing the power outputs over all stations, to give the same result as equation (2.10).

2.2.3 Unsaturated Model of Transmitted Model

In the unsaturated case, we aim to obtain the transmission probability τ given n and the level of the traffic by considering two unsaturated models [5][22] based on different assumption of buffer size. After obtaining the transmission probability, the transmitted power can be predicted by equation (2.10). Firstly, we let q be the probability of a packet arriving at the MAC during an average slot time, and r be the probability that at least one packet in the buffer after the station successfully transmits or discards a packet. The main difference of the unsaturated model from the saturated one is the introduction of new state called "Idle state" corresponding to the station which has no packet to send when a backoff process is finished. By assuming all packets experience a backoff period, we neglect the states of post backoff as most authors do [5]. The augmented Markov Chain is depicted in Figure 2.2. We consider the transitions from $(i, 0),\ i \in [0, m]$ where the counter reaches zero: if there is no packet available to send, it goes to Idle state; otherwise a new backoff stage at 0 process is triggered. Starting from Idle state, the change of state depends on whether at least one packet arrives while the station is idle: if no packet arrives, we stay in this state, otherwise it starts a new backoff stage at 0.

As in Section 2.2.1, we employ the renewal reward theorem to directly generate an expression for the transmission probability $\tau := \tau(p, q, r)$. The assumptions 2, 3, and 4 in Bianchi's model are still applied in this case. For a given station we define the following sequence: $Q_k = 1$ if after the k^{th} successful transmission (or a packet discard

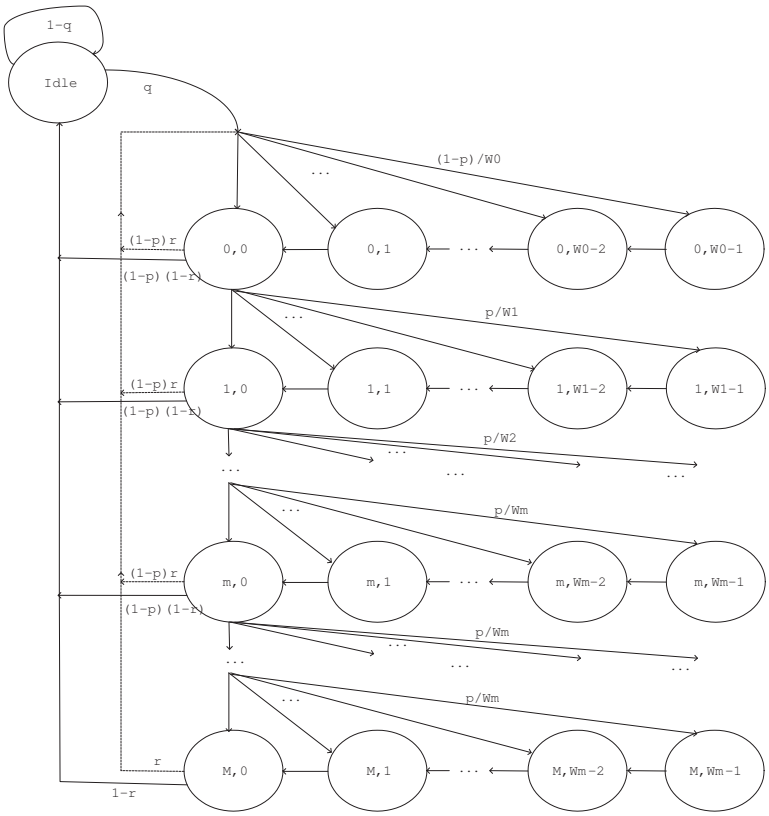


Figure 2.2: Unsaturated Markov Chain

as a consequence of retry limit exceedance) there is a packet waiting for sending and $Q_k = 0$ if there is no packet available for the station. We make the assumption that $P(Q_k = 1) = r$ and $P(Q_k = 0) = 1 - r$. These transitions are present in Figure 2.2. After a successful transmission (or discard), the expected number of slots $(1 - r)/q$ until another packet is available is the probability that there is no packet awaiting transmission $1 - r$, times the expected number of slots until a packet arrives $1/q$. Then we have that

$$\begin{aligned}\tau(p,q,r) &= \frac{E(1 + C_0 + C_0C_1 + \cdots + C_0C_1 \cdots C_M)}{E(U_0 + C_0U_1 + C_0C_1U_2 + \cdots + C_0C_1 \cdots C_MU_M) + (1-r)/q} \\ &= \frac{1 + p + p^2 + \cdots + p^M}{W_0/2 + pW_1/2 + \cdots + p^MW_M/2 + (1-r)/q}. \end{aligned} \quad (2.13)$$

and with equation (2.2), it is followed by

$$\tau(p,q,r) = \frac{2(1-p^{M+1})}{W_0(1-p-p(2p)^m)/(1-2p) - W_0 2^m p^{M+1} + 2(1-p)(1-r)/q}. \quad (2.14)$$

We consider a traffic arrival model using a Poisson traffic model [29]. That is, if packets arrive at the MAC in a Poisson manner with a rate λ packets per second, then $e^{\lambda T}$ is the probability that no packets arrive in an expected slot time T according to the probability mass function of the Poisson distribution. It then generates q as $1 - e^{-\lambda T}$, the probability that one or more packets arrive in an expected slot time. An expression is given for q for a traffic arrival rate of λ packets per second using a Poisson traffic model:

$$q = 1 - e^{-\lambda T}, \tag{2.15}$$

where T is the mean slot time in equation (2.9). In [17], other traffic models are considered, but they are found to have similar performance in terms of throughput and conditional collision probability. Hence we only consider the Poisson model here. An expression for r is discussed based on two assumptions of buffer size.

Infinite buffer model

Firstly, we consider one unsaturated model with infinite buffer introduced in [22]. In order to obtain r, we consider the distribution $B(p)$ of the number of states in the Markov Chain that pass for each packet prior to a successful transmission given the conditional collision probability p. [28] gives the expectation of $B(p)$ as

$$E(B(p)) = \frac{W_0(1 - p - p(2p)^m)}{2(1 - 2p)(1 - p)} - \frac{W_0 2^m p^{M+1}}{2(1 - p)}. \tag{2.16}$$

In this case, each station is effectively an M/G/1 queue [30]. We use a common assumption [22]: after a packet is sent, the probability that the queue is busy can be approximated by the stationary probability that the buffer is not empty for the M/G/1 queue. This is , from standard results in queueing theory, the stationary probability r that at least one packet is available after a transmission is $\min(1, \lambda E(B(p))T)$. Combining equation (2.15), it gives

$$r = \min\left(1, -E(B(p))\log(1 - q)\right). \tag{2.17}$$

No buffer model

Secondly, another unsaturated model of 802.11 DCF [5] is introduced based on idealized assumption of no buffer. That is, no packets can be buffered until the end of a service period and the likelihood of a packet arrival during a slot of average length is

the probability that an inter-arrival time is shorter than the average slot time. It is achieved by setting $r = 0$ with q given by equation (2.15).

In conclusion, for saturated traffic, the relationship between the transmission probability and the conditional collision probability can be obtained by combining equation (2.1) and equation (2.5). From equation (2.10), the transmitted power is estimated for a given number of stations. In the non-saturated infinite buffer case, we use equation (2.1) equation (2.14), equation (2.15) and equation (2.17) to find τ and then the transmitted power can be calculated theoretically using equation (2.10) for different numbers of stations and traffic loads. Similarly, estimations for the no buffer case can be obtained using equation (2.14) and equation (2.15).

2.3 Experimental Verification

2.3.1 Experimental Apparatus

Our experiments are carried out based on our IEEE 802.11 wireless testbed as shown in Figure 2.3, which is configured in the infrastructure mode. This is a similar configuration to that which might be found with a number of devices in a home or in a public hot spot. Each station denotes a laptop or other wireless equipment using IEEE 802.11 WLANs. The testbed includes 9 stations that are a collection of PC-based embedded Linux boxes based on Soekris net4801 [31], a desktop PC (Personal Computer) acting as a client station and another desktop PC acting as an access point. All devices are installed with an Atheros AR5215 802.11b/g PCI card with an external antenna. All stations, including the AP, run on a Linux 2.6.8.1 kernel and a version of the MADWIFI [32] wireless driver which we have modified to allow greater logging and control. The desktop PC is employed as a station to record detailed per-packet statistics. An advantage of this PC acting as a station is that there is adequate storage space, competent random access memory (RAM) and central processing unit (CPU) for the collection of statistics. All stations are connected through 100 Mbps wired cables to a desktop PC that controls the testbed system.

We use a number of common tools in the traffic engineering community in our testbed. We use *systcl* to specify traffic parameters, such as fixed data rate. We use *ssh* and *scp* are used to manage the network and control traffic sources over wired Ethernet ports. The *megn* tool is used to generate User Datagram Protocol (UDP) traffic [33]. Another *athstats* tool is also used to collect statistics from the wireless

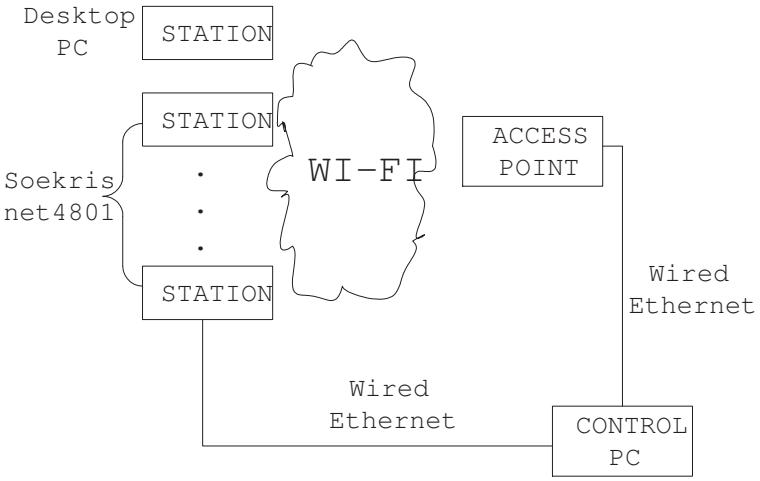


Figure 2.3: Schematic diagram of the wireless testbed

driver.

We implement our experiments based on two measurement methods. For the first method we record the number of successful transmissions and collisions by analyzing a trace file produced by the modified driver and stored only in the desktop PC station. We then scale up the number of transmissions/collisions by the number of stations in the wireless network to approximate results for the whole network. We are assuming that the network is symmetric and so the performance of other stations is the same as the desktop PC. This technique only requires us to record data at one PC, but we can only expect accurate results if the traffic load on the stations is symmetric and are in a symmetric environment.

Our second measurement technique uses *athstats*, which records basic statistics relating to the wireless card. We focus on the number of transmitted frames, the number of retries and the number of failed transmissions. We record these statistics for each station in our testbed. Compared to the first method, we expect higher accuracy, as we now have a picture of the whole system's performance. We will compare results generated by both of methods in the Section 2.3.2.

In our experiments, we try to configure all stations identically to make the network symmetric. Regardless, there still exist some differences due to the environment. An example is depicted in Figure 2.4 and Figure 2.5, which show the total number of retries and transmissions as the load is varied. The stars are the results for the desktop PC station, with the lines for the other 8 stations from the infinite buffer experiment based on 11 Mbps data rate and 100s experiment time. Results are

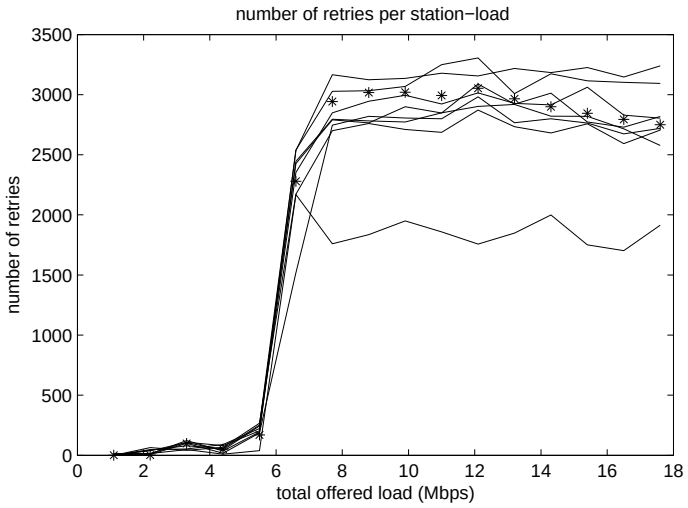


Figure 2.4: Number of Retries for each station vs. offered load

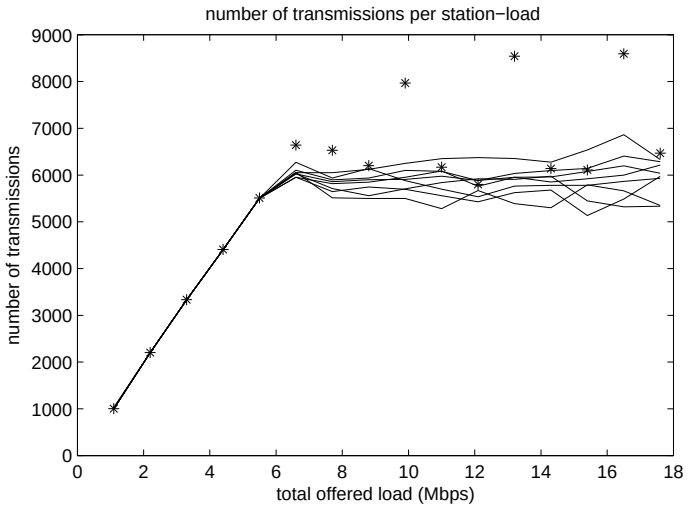


Figure 2.5: Number of Transmissions for each station vs. offered load

shown for 9 stations operating at the same time with the same network parameters. It is evident that most of stations are similar in terms of transmission transmissions and retry times. One station shows a much smaller number of retries. We will see in Section 2.3.2 that this symmetry actually has a small impact in the prediction of transmitted power which, in this case, is dominated by the number of successful times rather than the number of retries.

After obtaining the desired statistics n_s (the number of successful transmissions) and n_c (the number of collisions), the transmitted power is calculated as

$$P = \frac{n_s \mathcal{E}_S + n_c \mathcal{E}_C}{T_{total}} \tag{2.18}$$

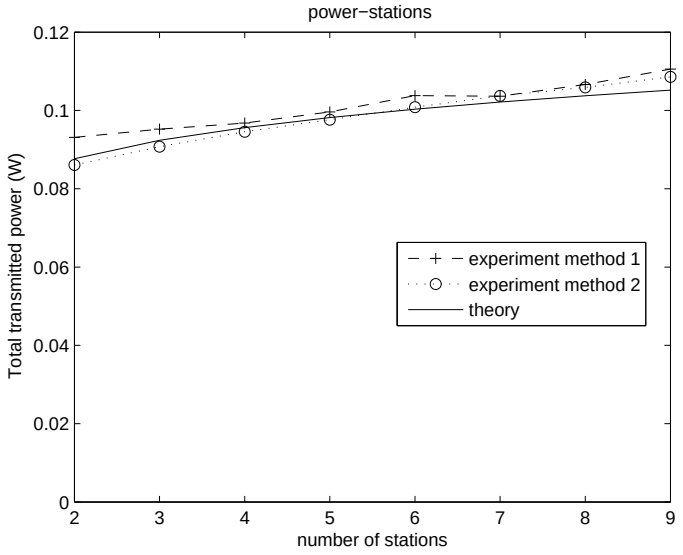


Figure 2.6: Transmitted power vs. Number of stations for unicast network

where E_S and E_C are the mean energy associated with a successful transmission and a collision; T_{total} is the time for the whole experiment, which is calculated accurately by subtracting first in-queue time by last in-queue time. They are calculated in the same way as their theoretical counterparts as described in Section 2.2.

2.3.2 Results

We use the parameter values of our network in Table 2.1 combined with the analysis in Section 2.2 to compare the theoretical transmitted power with our measured estimates as we vary the number of stations or the offered load. In the experiments, UDP is generated at 11 Mbps to saturate the network. The experiment is run for 100 seconds.

Figure 2.6 shows the transmitted power comparison between theory and experiments in saturated conditions. As expected, the power increases for larger number of stations. We see good match between theory and experiments regardless of our measurement method: as predicted the power goes from slightly below the nominal value to around the nominal value as the number of stations is increased.

Note that the results of experimental method 1 are slightly more variable. This is because of the network is not symmetric in practice as shown in Figure 2.4 and Figure 2.5. We see better agreement with model predictions using method 2, which more accurately reflects the total power actually transmitted.

Broadcast packets are considered in [17], because the 802.11 backoff mechanism

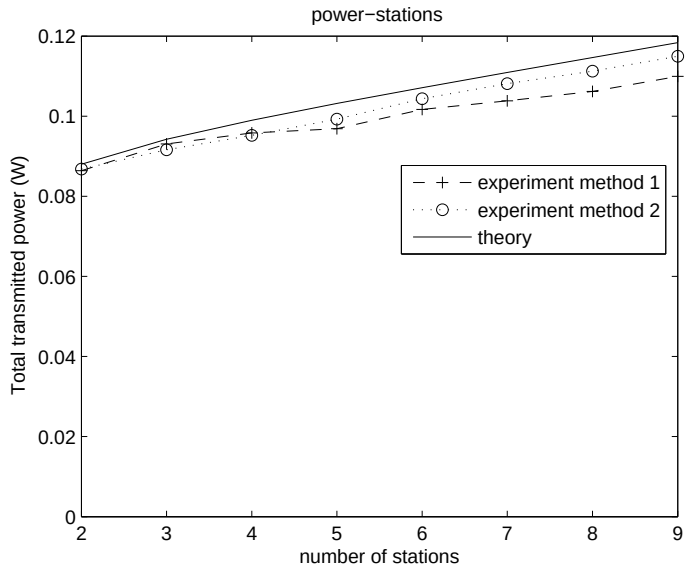


Figure 2.7: Transmitted power vs. Number of stations for broadcast network

operates differently for packets that are destined to groups of devices. The differences arise from the fact that no ACK packet is sent, because no one host can know if the whole group has received the packet. We compare the predictions of the model with results in our testbed in Figure 2.7. As expected, we see slightly higher power output than in Figure 2.6, and the match between the theory and the experimental result remains good.

In Figure 2.7 and Figure 2.6, we observe the power being an approximately linear function of the number of saturated stations. The model has captured the intercept (80 mW) and slope (around 2 or 4 mW per station for non-broadcast/broadcast). For non-broadcast packets, we expect this slope to decrease for larger numbers of active stations, as the MAC's backoff will tend to reduce the transmission rate. This behavior is predicted by the model, as shown by [17], but our testbed is not large enough to verify the result. In the next section, we will focus on results obtained by the measurement method 2 due to its better accuracy.

We are also interested in the transmitted power of unsaturated traffic for WLANs. The experimental results are shown in Figure 2.8 and Figure 2.9 for the big buffer and the no buffer respectively. We look at the non-saturated case with a 11Mbps data rate and 100s experiment time. In practical implementations, we approximate the infinite buffer with 200 packet buffer and the no buffer with a one packet buffer. We show the results of method 2 and equivalent model predictions for 2, 5 and 9 devices

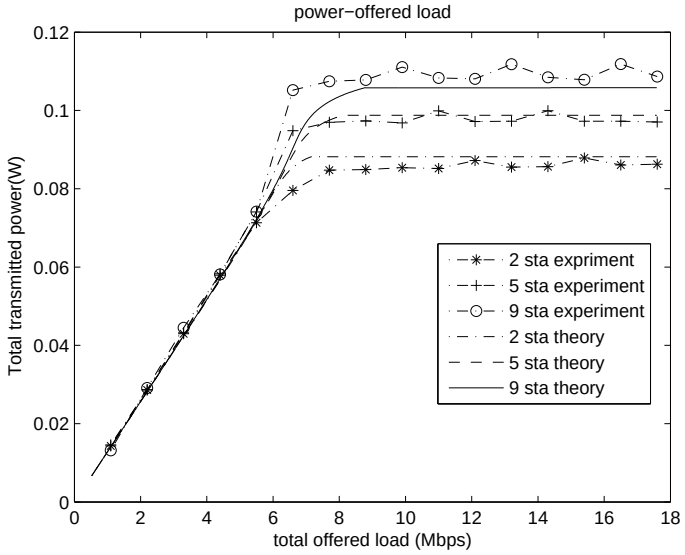


Figure 2.8: Transmitted power vs. offered load. Infinite buffer

as we vary the load. The match between theory and experiments is excellent over lower traffic loads.

There is also good agreement as the network becomes saturated. For heavy load, the small buffer shows an almost perfect match for the cases of 2 stations and 5 stations, but underestimates by about 7% for 9 stations. In contrast, the big buffer presents a better match with theory in the case of 5 stations, but slightly overestimates the power for 2 and 9 stations. In the intermediate region, larger discrepancies are evident.

We also present the results of our duty cycle calculations. Figure 2.10 demonstrates the difference between the duty cycle of the entire network and the duty cycle summed over the stations as predicted by the model described in Section 2.2.2. Collisions allow the duty cycle summed over all stations to exceed 100%, which leads to the power exceeding the nominal value. Since our testbed results are per-station statistics, they only allow us to compare the duty cycle summed over stations with the model (Figure 2.11 and Figure 2.12).

Figure 2.11 shows that the duty cycle of the saturated network increases quickly from approximately 75% at 1 station to approximately 105% at 9 stations. Figure 2.12 shows how the duty cycle is small when the non-saturated 0.5 Mbps traffic is used, and increases linearly to around 6 Mbps. As these results are essentially rescaled versions of our power graphs, we see similarly good matches between model predictions and experimental results.

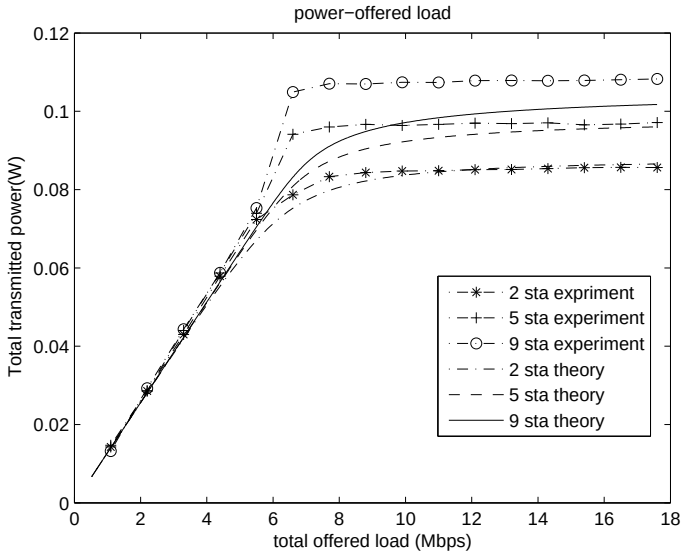


Figure 2.9: Transmitted power vs. offered load. No buffer

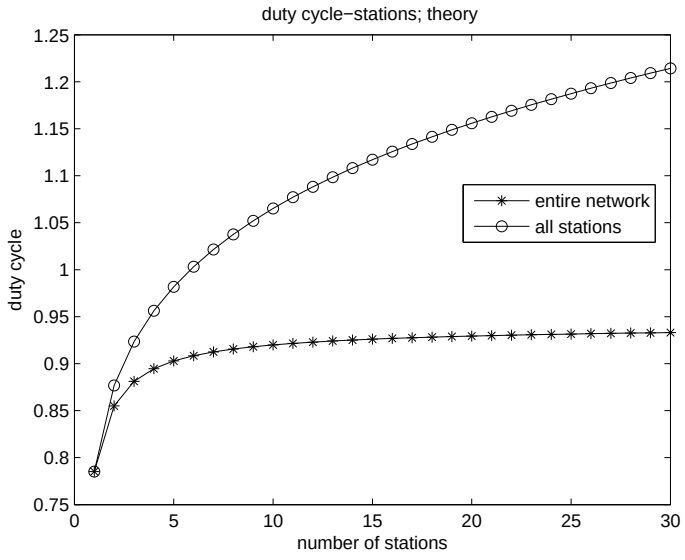


Figure 2.10: Duty Cycle vs. Number of stations. as predicted by the model

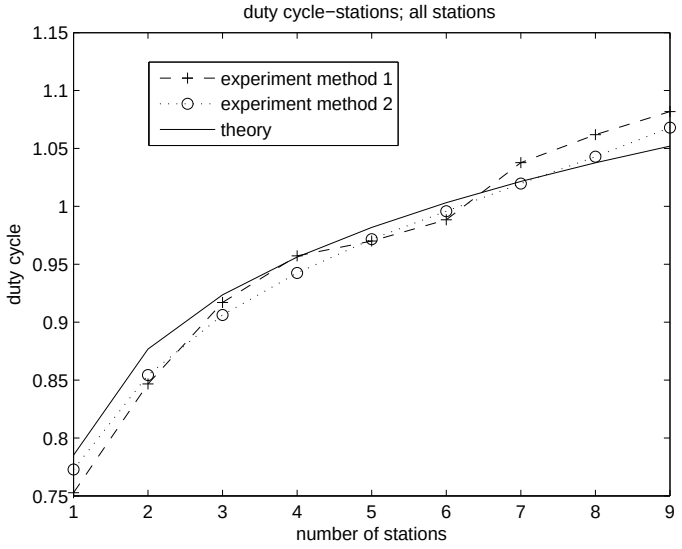


Figure 2.11: Duty Cycle vs. Number of stations for the saturated network

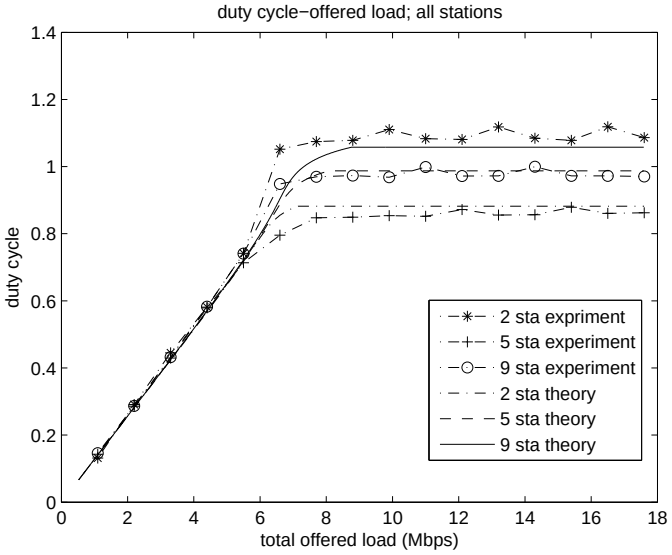


Figure 2.12: Duty Cycle vs. offered load. Infinite Buffer

These differences observed between theory and experiments may be attributed to the assumptions of the theoretical model: the network is not completely symmetric and we approximate an infinite buffer or no buffer with 200 packets or 1 packet. While the models are clearly not capturing the physical systems exactly, the power predictions are still quite satisfactory. Overall, the infinite buffer model's predictions appear better and are likely to better reflect the configuration of actual devices.

2.3.3 Discussion

When estimating the power output of a network, it may be useful to be able to estimate the largest possible power, regardless of traffic conditions. Intuition would suggest that the most power will be output when the network has the most to send, and is likely to be an implicit assumption of experimental studies. However, one feature of random-access MAC systems, such as 802.11, is that better data throughput can sometimes be achieved before the network becomes saturated. This is demonstrated, for example, in [5]: for larger numbers of stations as load is increased the network's throughput increases to a peak and then decreases to its saturated level. Interestingly, we see no power pre-saturation peak. We believe this is because, for realistic parameters, the expression for power (equation (2.10)) is a strictly increasing function of transmission probability, unlike the expression for throughput [3]. This suggests that the upper-limit on outputted power of a network can be reasonably approximated by calculating the throughput when the network is saturated.

Another interesting observation from the graphs that the power is a linear function of the offered load when there is a small amount of traffic in the network. This arises because the number of collisions for light loaded traffic is small, and so each packet is transmitted just once successfully. Since 802.11 has a per-packet power overhead (for *preamble*, *headers* and ACK) and then a per-byte power cost (for transmitting the actual data) we may approximate the power as:

$$P = P_0(pps)(preamble + header/rate + ACK) + (bps)/rate \tag{2.19}$$

where *pps* is the number of packets per second and *bps* is the number of bytes per second. Figure 2.13 shows the results of applying this rule of thumb to our experimental data. Note that the predictions are independent of the number of stations, and actually match well until the network reaches saturation. This suggests that this

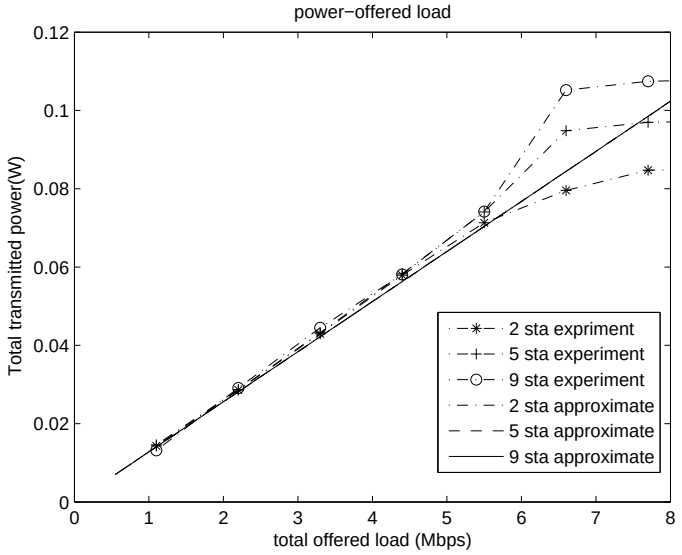


Figure 2.13: Power vs. offered load with simple lightly-loaded approximation

will be a useful approximation for lightly loaded networks.

Combining these two observations gives a simpler technique for estimating the power, if we know the amount of traffic. First, we use equation (2.19) to predict the power. Then we compare this to the power for a saturated network, and take the minimum. As examples, consider the following situations.

1. An architect who uploads a large amount of data through their 1Mbps broadband link is concerned about their RF exposure. We need to determine the total number of bytes per second and packets per second being sent over the network so that we can use equation (2.19). We note that there will actually be two senders in this wireless network: the station uploading the data and the access point, which will be sending higher level response packets. Protocols are usually designed so that these response packets are sent for every one or two packets, but will be much smaller (60 or 70 bytes). To estimate the number of packets per second that can be sent over a 1Mbps link, we need to know the packet size in bits. Packet sizes of 1400-1500 bytes are typical on modern broadband networks, so we use 1400 bytes = 11200 bits. This gives a figure of 1Mbps/11200 = 89 packets per second in one direction. We double this, to allow for the responses in the other direction. The number of bits per second will be 1Mbps in one direction and roughly (1Mbps) (70bytes)/1400bytes = 0.05Mbps in the other direction. For a Wi-Fi rate of 11Mbps, we can use equation (16) to

estimate the power as about 17mW. This is well below the saturated power of just over 80mW, so we do not need to make any adjustment.

2. An office worker registers a complaint about a colleague who spends their lunch breaks watching YouTube videos, and is worried about the impact of the continuous downloads. This example is very similar to the previous example but the traffic now flows from AP to station. The constraint is how fast the user needs to download video in order to watch it for a period of time. Gill et al. [34] show that most YouTube videos are encoded at a rate between 300 and 400Kbps, with the mean and median falling in this range. Starting with a rate of 400Kbps rather than 1Mbps, we may repeat the above calculation to get a value around 7mW.

3. In a high-school class, 30 students are encouraged to watch a short documentary from YouTube on their laptops at the end of each class. Parents express concern about 30 wireless devices being used at the same time.In this case we now have 400KB * 30 users traffic from the access point to the laptops, plus the response packets from the laptops to the access point. Calculating as above, we get a power estimate of around 203mW. Checking Figure 2.10, we find that the summed duty cycle for 30 saturated nodes is just over 1.2, suggesting that power actually saturates around 120mW, rather than the potential 30*100mW.

Unsurprisingly, these powers are low when compared to the ICNIRP limit of 80mW kg^{-1}. However we now have a quick way to estimate power, provided that some information about the traffic is available in the WLANs.

2.4 Summary

In conclusion, we have extended the recently introduced model of transmitted power [17] from saturated conditions to unsaturated conditions. Through our testbed experiments, we have verified this model and see a close match between theoretical predictions and experimental results. We find that the power of a saturated network is a reasonable upper bound on the power of an unsaturated network. For lightly loaded networks, we also offer a simple but accurate technique for approximating the power output. Finally, we give some examples of how these techniques might be applied, and confirm that the estimated maximum power is substantially lower than the

internationally acceptable limit given by ICNIRP.

In the thesis, we have focused on the total transmitted power of WLANs by assuming the nominal power 0.1 W. When more information is available, for example mixed output powers, distances from devices, antenna details, reflection patterns, etc, these could be incorporated into the model via the per-node duty cycle in equation (2.12). By weighting this duty cycle with per-node factors, such as output power, antenna gains and power decline due to distance, more exact dose calculations could be performed. Forster [13] provides a detailed discussion of factors that could be accounted for. Some of the work contained in this chapter appears in Fang and Malone [18].

Chapter 3

Design of Decentralized Learning MAC Schemes for Collision-free Access in WLANs

3.1 Introduction

As the most commonly employed Multiple Access Control (MAC) scheme in WLANs, the IEEE 802.11 Distributed Coordination Function (DCF) as described in Section 1.1.1 regulates access to the medium based on the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA). In this MAC, time on the medium is divided into idle slots of fixed length σ μs, and busy slots of variable length caused by transmission. Frames are positively acknowledged to allow retransmission on failure. In a network with more than one transmitter using the DCF, there exists a persistent non-zero likelihood of collisions. In particular, the throughput performance is substantially degraded when a large number of stations are contending for the medium [3]. As we have discussed in Section 2.2, collisions can also increase the transmitted power output.

In contrast, Time Division Multiple Access (TDMA) based MACs can make better use of the radio channel by eliminating collisions. However, traditional TDMA has

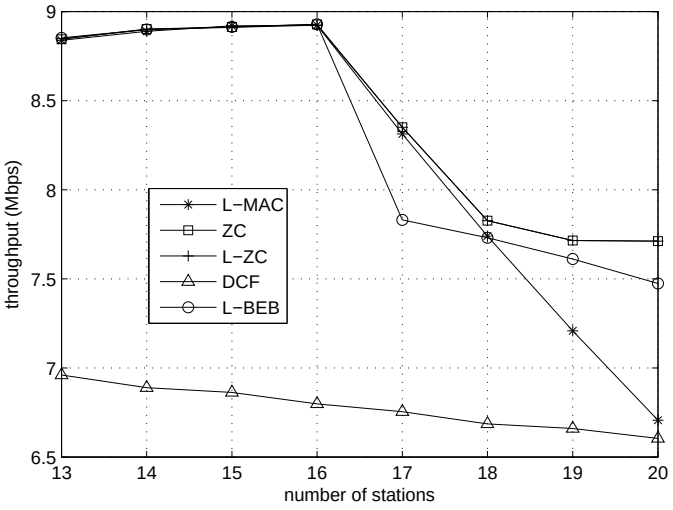


Figure 3.1: Network throughput vs. number of stations, comparison of MACs. Schedule length $C = 16$. ns-2 simulations. L-ZC overlays ZC.

drawbacks, typically employ a central controller that must maintain detailed knowledge of each station queue occupancy and their topology which requires extra data exchange.

Combining advantages of both TDMA and CSMA/CA, new hybrid MAC protocols have recently been proposed. For example, Figure 3.1 shows the throughput performance of a number of MACs that we will discuss, which can be seen to outperform DCF by almost 30% by avoiding collisions. For example, Learning Binary Exponential Backoff (L-BEB) [19] uses a fixed or reselected random backoff value to achieve collision-avoidance. Like 802.11's DCF, it chooses backoff values based on the success or failure of the last transmission, making it amenable to implementation on existing platforms. Other similar schemes have also been proposed, see Section 3.2 for a brief review. Another collision-free scheme, ZC (Zero-Collision), was proposed in [20]. In contrast to DCF and L-BEB, this MAC requires information about every slot, not just those where it transmits. By using this additional information about slot occupancy, ZC achieves significantly faster convergence than L-BEB, but is not as readily implementable on existing hardware.

In this thesis, we propose two learning MACs for collision-free access as evolutions of the L-BEB and the ZC in terms of two positive aspects:

1. We incorporate ideas from a decentralized channel selection algorithm introduced in [21][35], inspired by learning automata [36], to improve convergence

times of these MACs. In particular, we propose a fully decentralized Learning MAC (L-MAC) that uses the same information as L-BEB, but achieves convergence orders of magnitude more quickly. Similar ideas are also applied to ZC, and we demonstrate a learning version, L-ZC that provides convergence that is faster than ZC.

2. In Figure 3.1, throughput begins to fall when the number of stations exceeds 16, a fixed schedule length employed by these learning MACs resulting in unavoidable collisions. We introduce a mechanism that automatically adapts schedule length in a decentralized fashion that does not require agreement between stations while crucially retaining fairness properties expected of the MAC. This allows that MACs to scale to any number of stations.

These MACs are fully decentralized and do not require information exchange among transmitting stations or additional control frames that would increase system complexity. L-MAC only uses feedback concerning whether each transmission is successful or not. This information is already provided by IEEE 802.11 hardware and, thus, L-MAC can be implemented with relatively minor changes on a flexible MAC platform. In contrast both ZC and L-ZC require additional information on each slot on the medium, restricting their implementation to future hardware, but provide enhanced performance.

We prove that these new access methods converge to a collision-free schedule, if one exists. In doing so, we address a lacuna[1] in the analysis in [20]. We show how to set the learning parameters of these algorithms. In the case of L-MAC, we use simulations to choose parameters that offer a good balance between fairness and capacity. For L-ZC we provide mathematical analysis of the convergence that enables analytic optimization of the algorithms parameters.

By avoiding collisions, network throughput is significantly higher than DCF. In particular, reducing the convergence time to collision-free operation offers improved performance for delay-sensitive periodic traffic such as the voice and game data in addition to many station where the IEEE 802.11 collision rate is likely to be large [3]. Faster convergence times also accommodate changing network conditions. Finally, by addressing the fundamental issue of how to adapt the schedule lengths for these access schemes in a decentralized way, while retaining fairness, this enables scalability

[1]The number of stations colliding is assumed to form a Markov chain, however additional state is needed for this to be true.

to networks of any size.

This work was performed in collaboration with K. Duffy, D. Malone and D. Leith (NUIM) and is being prepared for submission for publication.

3.2 Related Work

Z-MAC [37] is a hybrid protocol that combines TDMA with CSMA. Z-MAC assigns each station a slot, but other stations can borrow the slot, with contention, if its owner has no data to send; the collision-free MAC proposed in [38] has less communication complexity. Both of these MACs experience the same drawback that extra information exchange beacons are required. These introduce additional system complexity, including neighbour discovery, local frame exchange and global time synchronization.

A collision-free MAC is introduced in [39] for wireless mesh backbones that reduces the control overhead greatly when compared to the DCF protocol. It guarantees priority access for real-time traffic, but it is restricted to a fixed wireless network and requires extra control overhead for every transmission. Ordered CSMA [40] uses a centralized controller to allocate packet transmission slots. It ensures that each station transmits immediately after the data frame transmission of previous station. Its drawback is that the requirement of having a centralized controller with its associated coordination overhead.

Recently, Barcelo et al. [19] proposed Learning-BEB, based on a modification of the conventional 802.11 DCF. In a decentralized fashion, it ultimately achieves collision-free TDMA-like operation for all stations which occupy different slots from each other. The basic principle of its operation is that all stations choose a fixed, rather than random, backoff value after a success. After a failure, they choose a slot uniformly at random, as does DCF. We can think of this as each station randomly choosing a slot in a schedule, until they all choose a distinct slot. Arriving at this collision-free schedule can take a substantial period of time. In particular, when the number of slots in a schedule is close to the number of stations, it will take an extremely long time to converge to collision-free scenario. The authors of [41] propose a scheme, SRB (Semi-Random Backoff), that is similar in spirit to L-BEB.

In hashing backoff [42] each station chooses its backoff value by using asymptotically orthogonal hashing functions. Its aim is to converge to a collision-free state. One structural difference from L-BEB [19] is that hashing backoff [42] introduces an

algorithm to dynamically adapt the schedule length using a technique similar to Idle Sense [43]. This length adaption scheme requires additional information such as the estimation of idle rate. The broad principles of these MAC protocols are similar and both have the drawbacks of slow convergence speed to a collision-free state and relatively poor robustness to new entrants to the wireless network.

ZC is proposed in [20]. We can regard ZC as being similar to L-BEB in that on success it effectively chooses a fixed backoff. On failure, however, a station looks at the occupancy of slots in the previous schedule. The station chooses uniformly between the slot it failed on previously and the slots that were idle in the last schedule. By avoiding other busy slots, which other stations have 'claimed', ZC finds a collision-free allocation more quickly than other schemes. Through the use of this additional information, ZC converges more quickly than the schemes mentioned above.

3.3 Learning MAC and Learning ZC

3.3.1 The L-MAC protocol

Here we propose a decentralized Learning MAC (L-MAC), which can be regarded as an evolution of the L-BEB [19] incorporating ideas from the self-managed decentralized channel selection algorithm [21]. The primary difference between L-MAC and L-BEB is in the latter collisions cause memory to be lost of the current schedule. In contrast, L-MAC keeps some state: each station that has found a slot that previously did not have competition is likely to persist with that slot even after a small number of collisions. A probability distribution is introduced as internal state for each station. It determines the likelihood of choosing each slot in a periodic schedule $\{1, \cdots, C\}$. Note, that no agreement is required between the stations on the labeling of the slots, only on their boundaries, just as in DCF. The advantage of learning is that it introduces a stickiness that improves the speed of convergence to a collision-free transmission schedule and facilitates quick re-convergence to a new schedule when additional stations join an existing network.

L-MAC's slot selection algorithm is parameterized by two numbers: the number of slots in a schedule, C, and the learning strength, $\beta \in (0, 1)$. All stations listen to the medium when not transmitting so that busy slots can be of arbitrary length, but idle slots are of fixed length σ. For each station, L-MAC is defined as follows for each station.

1. The probability distribution $\mathbf{p}(0)$ is initialized at time 0 to the uniform distribution,

$$\mathbf{p}(0) = [p_1(0), \ldots, p_C(0)] = \left[\frac{1}{C}, \ldots, \frac{1}{C}\right].$$

and a slot $s(0)$ is selected in $\{1, \ldots, C\}$ according to $\mathbf{p}(0)$.

2. Let $s(n)$ denote the slot selected for transmission in the n'th schedule. If the station has a packet to send and is successful or if it has no packet to send and observes the medium to be idle during slot $s(n)$, then the probability distribution is updated to $\mathbf{p}(n+1)$ defined by

$$\begin{aligned} p_{s(n)}(n+1) &= 1 \\ p_j(n+1) &= 0 \end{aligned}$$

for all $j \neq s(n)$, $j \in \{1, \ldots, C\}$. That is, after selecting a non-colliding slot in the schedule, the station will persist with the same slot $s(n)$ in the following schedule.

If transmitting in slot $s(n)$ result in a collision or if the station has no packet to send and observes the medium to be busy during slot $s(n)$, the probability distribution is updated to $\mathbf{p}(n+1)$ defined by

$$\begin{aligned} p_{s(n)}(n+1) &= \beta p_{s(n)}(n) \\ p_j(n+1) &= \beta p_j(n) + \frac{1-\beta}{C-1} \end{aligned}$$

for all $j \neq s(n)$, $j \in \{1, \ldots, C\}$. That is, after a failed transmission, a station changes its probability distribution to make it less likely that it selects the same slot again, but it does so in a way that reflects how confident the station was that the previously selected slot would not result in a collision.

The station then selects a new slot $s(n+1)$ from the probability distribution $\mathbf{p}(n+1)$ and will next transmit after $C - s(n) + s(n+1)$ slots. That is, at slot $s(n+1)$ in the next schedule.

3. Return to step 2).

Before identifying good choices of L-MAC's two parameters, C and β, we state

the following theorem that proves that L-MAC converges to a collision-free schedule if one exists.

Theorem 1. *Suppose that all stations employ the decentralized L-MAC. Assuming that the number of stations is not more than C, for any $\beta \in (0,1)$ the network converges in finite time to a collision-free schedule with probability one.*

Proof. By adapting ideas from [35], we will show that from any state in any two steps of the algorithm, there is a probability of convergence that is bounded away from zero. The probability of selecting a slot can become arbitrarily small if the station has been colliding on the same slot for many schedules, so we must construct a sequence of events that avoids this possibility.

Suppose the WLAN consists of N stations. Define $\mathbf{p}^{(i)}(n) \in [0,1]^C$ to be station i's probability distribution in the n'th schedule and $s^{(i)}(n) \in \{1, 2, \cdots, C\}$ to be its slot chosen for transmission.

If we have $s^{(i)}(n) \neq s^{(j)}(n)$, $\forall i \neq j \in \{1, \ldots, N\}$, then the network has already found a collision-free schedule and there is nothing to prove. If, at schedule n, there was at least one collision, then as $C \geq N$, there must be some slot i^*, which has been selected by none of the stations. At schedule $n+1$, for any station k colliding at slot $i \neq i^*$ in schedule n, the probabilities of moving to i^* is

$$p_{i^*}^{(k)}(n+1) = \beta p_{i^*}^{(k)}(n) + \frac{1-\beta}{C-1} \geqslant \frac{1-\beta}{C-1}.$$

Thus the probability that all the stations that collided in schedule n then, in schedule $n+1$, choose i^* is at least $((1-\beta)/(C-1))^N$.

In schedule $n+2$, the probability a station k that collides in schedule $n+1$ now picks any slot j is bounded by below by

$$p_j^{(k)}(n+2) = \beta p_j^{(k)}(n+1) + \frac{1-\beta}{C-1} \geqslant \frac{\beta(1-\beta)}{C-1}.$$

Since there is at least one non-colliding configuration, the probability of jumping to this is at least

$$\left(\frac{\beta(1-\beta)}{C-1}\right)^N.$$

In summary, no matter what the slot-selection conditions for stations are in sched-

ule n, the probability of schedule $n+2$ being collision-free, $P(\vec{\mathbf{p}}(n+2) \in A)$, is bounded below by:

$$K := \left(\frac{1-\beta}{C-1}\right)^N \left(\frac{\beta(1-\beta)}{C-1}\right)^N > 0$$

Let τ be the first time a collision-free schedule is found, we want to show $P(\tau < \infty) = 1$. At time $2n$, the probability of arriving at collision-free schedule for the first time is:

$$P(\tau \geqslant 2n) \leqslant (1-K)^n. \tag{3.1}$$

Thus, as $n \to \infty$ for any $(1-K) \in (0,1)$, this equation implies:

$$\lim_{n\to\infty} P(\tau \geqslant n) = \lim_{n\to\infty} (1-K)^n = 0.$$

and so $P(\tau < \infty) = 1$. Note that equation (3.1) upper bounds the stopping time τ by a geometric distribution and, therefore, all of this stopping time's moments (mean, variance, etc.) are finite.

□

3.3.2 The L-ZC protocol

L-ZC is a modification of the ZC protocol proposed in [20]. In ZC, each station initially chooses randomly and uniformly from the all virtual slots $\{1, 2, \ldots, C\}$. If it is successful, it chooses a backoff value of C to use the same slot in the next schedule. Otherwise, it notes the n_i idle slots from the previous schedule and the slot that resulted in a collision, and chooses among these with a uniform probability $1/(n_i + 1)$.

In L-ZC we introduce an additional parameter γ, that will control probability that we choose the same slot after a collision. Instead of using the uniform probability distribution, we will use $(1-\gamma)/n_{i+1}$ for idle slots and γ for the previously selected slot. The rationale is that different numbers of stations see particular slots as available for choice, depending on if a slot was idle, busy or chosen by a particular station in the previous schedule. By controlling the weight assigned to collision slots, we are able to improve convergence times.

L-ZC uses the same information that ZC does. It needs to know if its own trans-

mission was successful and which of the previous schedule's slots were idle. We have a result analogous to Theorem 1.

Theorem 2. *Suppose that all stations employ the decentralized L-ZC. Assuming that the number of stations is not more than C, for any $\gamma \in (0,1)$ the network converges with probability one in finite time to a collision-free schedule.*

Proof. The number of colliding stations in next schedule only depends on current number of colliding stations and the slots they collide on, hence we build a Markov chain model to study this stochastic process. We have N stations in the same channel without hidden nodes, and $C \geqslant N$ per schedule to ensure a collision-free schedule exists. We let $N_{(C)}$ be the number of stations experiencing schedule, n_C be the number of slots with collisions, and then $n_I = C - N + N_{(C)} - n_C$ is the number of idle slots. We can immediately establish our result by noting that the probability that $N_{(C)} > 0$ decreases is lower bounded by $(1-\gamma)\gamma^{N-1}/C$, the probability that one station jumps to an idle slot, but all others remain fixed. □

3.4 Schedule Length Adaptation

L-MAC is a two-parameter algorithm: the learning strength β and the schedule length C. Similarly, L-ZC is a two parameter scheme, γ and C. Learning strengths will be discussed in Section 4.2. In this section we will show how the schedule length can be adapted in a decentralized fashion while retaining throughput efficiency and fairness. Adapting the schedule length in a decentralized way, while retaining fairness, is particularly challenging. A decentralized scheme may stabilize at different schedule lengths, either because of differences in environment or available history (say, because the station is a new entrant to the network). This can result in unfairness or even failure to converge to a collision-free state, because of schedules drifting out of phase.

As the challenges for L-MAC and L-ZC are similar, we will employ similar schemes for both. We begin with an analysis of of how schedule length impacts on efficiency, and then describe the schemes for L-MAC and L-ZC.

3.4.1 The Impact of Schedule Length on Efficiency

As C is the number of available slots for all stations to achieve the collision-free schedule, naturally this is only feasible if the number of stations are not more than C.

date rate= 11Mbps	basic rate=11Mbps
PHY header= 24 bytes MAC header= 32 bytes payload=1020 bytes	SIFS=10μs DIFS=50μs idle slot time= $\sigma = 20\mu s$
header=(MAC header)/(date rate)+(PHY header)/(basic rate)	
ACK=(MAC header)/(date rate)+(14)(8)/(date rate)	
E_p=(payload size)(8)/(date rate)	
T_S=DIFS+(slot time)+header+E_p+SIFS+ACK	
T_C=DIFS+(slot time)+header+E_p+DIFS	

Table 3.1: MAC/PHY values mirroring 802.11b, E_p is the time spent transmitting payload, T_S is a successful transmission slot length and T_C is a collision slot length

We will begin by comparing the long-run throughput when the number of stations is less than or greater than C.

Assuming that the number of stations is N, all of which are saturated, we partition C into C_{suc}, C_{col} and C_{idle}, which denote the number of the successful slots, slots with collisions and the idle slots respectively. For an 802.11-like protocol, Table 3.1 shows parameters such as the length of idles and busy slots (see papers such as [3][5] to see how these are derived). Note that idle slots are an order of magnitude shorter than successful or collision slots.

When the number of stations $N \leq C$, then, once we have achieved a collision-free schedule, C_{col} equals zero and C_{suc} equals N. Hence, we get $C_{idle} = C - N$. Then we get a theoretical normalized throughput of

$$S = \frac{NE_p}{NT_S + (C-N)\sigma}. \tag{3.2}$$

When $N > C$, it is infeasible to achieve a collision-free state. We carry out an approximate analysis of throughput under an assumption of strong stickiness (i.e., big β or full L-ZC schedule) and a moderate number of excess stations. We assume that that each slot will have a single station 'stuck' to it and that the remaining $N - C$ stations are allocated to slots randomly with probability $1/C$. The number of slots occupied by the $N - C$ stations will be the number slots experiencing collisions, C_{col}. The problem becomes a balls-in-bins problem, where we are assigning $N - C$ balls to C bins, and so the mean number of occupied bins will be

$$E(C_{col}) = C\left(1 - \left(1 - \frac{1}{C}\right)^{N-C}\right). \tag{3.3}$$

With this estimate of C_{col} and $C_{suc} = C - C_{col}$, we obtain the normalized throughput

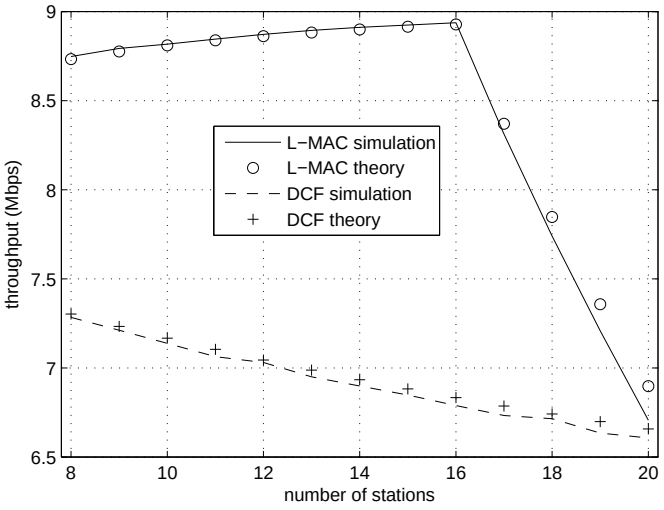


Figure 3.2: Network throughput vs number of stations, comparison between the theoretical model and the simulation results. $\beta = 0.95$. ns-2 simulations and theory based estimates

as

$$S = \frac{C_{suc}E_p}{C_{suc}T_S + C_{col}T_C}.$$

For example, consider the throughput as N changes and C is fixed at 16, as shown in Figure 3.2, where we set $\beta = 0.95$ for reasons that will be described in Chapter 4. For comparison DCF's throughput is also shown (the theoretical results for DCF are produced using well-known model from [3]). We note that a good match exists between the values predicted by theory and simulation results. Observe that L-MAC's throughput gradually increases as we increase the number of stations N to be the same as the number of slots. This is because we are eliminating short idle slots. A further increase results in a rapid decrease in throughput. This is because we now replace successful slots with long collision slots. Despite this, L-MAC continues to outperform DCF until $N = 20$ stations. In conclusion, the maximum throughput is achieved when N equals C, and a slightly smaller throughput is maintained when N is smaller than C as busy slots are of considerably longer duration than idle slots.

3.4.2 Choice of schedule length for L-MAC

When choosing a value for C it is better to overestimate the number of required slots in a schedule. Indeed, Figure 3.2 shows that even with too many stations (i.e.

$N = 17$), there can be a greater loss in throughput than having half the slots idle (i.e. $N = 8$). We use this to suggest two simple schemes to choose the value of C. One option is to use a static value of C that is larger than the largest number of simultaneously contending stations expected in the network. For example, a value of $C = 16$ would be reasonable for many small networks, and this is the scheme adopted by L-BEB.

A second option adapts the value of C. If stations operate with different values of C, two problems may arise. First, stations are trying to learn a good periodic schedule and so stations' schedules must not drift with respect to one another. Second, a station transmits once in every $1/C$ slots when a collision-free schedule is found, so fairness issues can arise.

We address the first problem by by using schedules of length $2^n B$, where B is a base schedule length. Thus, if two stations have different schedule lengths then the shorter divides evenly into the longer one. Consequently, the station with the longer schedule length sees a schedule where the other station seems to have claimed several slots.

To address the second fairness-related problem, we can choose to transmit multiple packets in a single slot using a technique such as 802.11e's TXOP mechanism [10]. Here, a station transmits multiple packet/ACK pairs separated by a short interframe space (SIFS). This time is short enough that other stations observing the medium will not consider it to have been idle and so backoff processes remain suspended. Thus we can avoid (long-term) fairness issues by allowing a station operating at $C = 2^n B$ to transmit 2^n packets. Short-term fairness issues will be over a time-scale of shorter than C_{max}/B schedules.

This suggests using an MIMD scheme where if a station finds that the schedule length is too short to accommodate all N stations it doubles the value of C being used. If the schedule length is much too large then C is halved. It remains to specify a mechanism that will trigger increases and decreases. As our scheme has been designed so that it does not require the values of C to be the same at all stations to provide fairness, this gives us increased flexibility in our choices, as we will not require the MIMD scheme to arrive at a consensus value of C, or even the same mean value. The trigger we use for doubling C is based on $f(C)$, the number of schedules we need for $C - 1$ stations starting in a random configuration to have converged with 0.95 probability. After arriving at a schedule length of C, the station checks every $f(C)$

schedules, checks to see if there collisions in that schedule. If it sees collisions C is doubled, otherwise C is unchanged.

As we've noted above, throughput can be quite satisfactory, even when C is larger than necessary, so we expect that reducing C will mainly contribute to improving short-term fairness. Conversely, having too small a C value can result in significantly reduced throughput. For this reason, we aim to probe smaller C values with a frequency that ensures we achieve 90% throughput possible at the current C value. To achieve this, we wait for $0.9f(C/2)(1-0.9)$ successful schedules at schedule length C before probing a schedule length of $C/2$. This ensures that if even all transmissions at the shorter schedule length fail, we will still see the desired throughput. In practice, we expect to see even higher throughput.

3.4.3 Choice of schedule length C for L-ZC

We being by noting that while L-ZC uses more information than L-MAC, once converged it behaves in a similar way to L-MAC. Thus, our analysis for L-MAC of the $N \leq C$ case above applies directly to L-ZC. While the exact details of what happens when $N > C$ are different, the broad principles are similar: as collision slots are longer than idle slots, it will be more desirable to have idle slots than collision slots.

This suggests that we can reuse the schemes suggested for L-MAC. Again, a fixed value of $C = 16$ may work well in small networks. An adaptive MIMD scheme can use the same principles as L-MAC, however L-ZC takes advantage of the of positions idle slots in the previous schedule, which allows us to use a more sophisticated trigger for MIMD. We suggest that an adaptive MIMD scheme that doubles C when there are no idle slots remaining. We decrease half C when the number of idle slots is at least half the schedule. In order to avoid decreasing C while L-ZC is converging and collisions are still ongoing, we wait until the the number of busy slots has stabilized and we see two consecutive schedules with the same number of busy slots before we consider a possible decrease.

3.5 Summary

Here we have proposed two techniques to improve MACs that discover collision-free schedules. By applying learning, we aim to reduce convergence times compared to L-BEB's. Crucially, we have also made these MACs scalable beyond a fixed number of

stations by showing how schedule length adaptation can be achieved, while retaining fair decentralized operation. Of our two proposed MACs, L-MAC uses the same information as DCF, making it amenable to implementation on existing platforms. L-ZC uses additional information to obtain improved performance, at the cost of restricting its implementation to more future hardware.

Chapter 4

Performance Evaluation of Decentralized Learning MAC Schemes for Collision-free Access in WLANs

4.1 Introduction

In Chapter 3, we discussed the design of Learning Multiple Access Control (L-MAC) and learning zero collision (L-ZC). L-MAC is a two-parameter algorithm: the learning strength β and the schedule length C. Similarly, L-ZC is a two parameter scheme, γ and C. The adaptive schedule length schemes have been proposed in Section 3.4. In the following sections we will identify reasonable values for β and γ. For L-MAC we will use simulations to consider factors such as transient fairness and achievable throughput, as well as convergence time, ultimately choosing $\beta = 0.95$. For L-ZC convergence times are much shorter and we will use mathematical analysis to show convergence times are minimized by selecting $\gamma = 1/(C - N + 2)$. We then carry out various simulations to compare the performance of learning MACs and existing MACs in terms of speed of convergence, long term throughput, robustness to new entrants, transmitted power and coexistence with 802.11 Distribution Coordination Function (DCF).

4.2 Parameter Choice

4.2.1 Choosing the learning strength β in L-MAC

The learning parameter β has an important impact on the convergence speed, the pre-convergence access fairness, achievable throughput when the network is oversubscribed and reconvergence after a change in network conditions. We will see that while large β offer fastest convergence, there is a value for β where convergence is fast while almost optimal fairness, oversubscribed throughput and reconvergence are achieved.

Speed of Convergence

First, consider the case where there are $N = 16$ stations that, in the terminology of [3], are saturated so that they always have packets to send. The schedule length, C, is also set to 16, allowing 16 idle, successful and collision slots. As $N = C$, this is the most challenging situation where a collision-free schedule exists. Other network parameters are as in Table 3.1.

Figure 4.1 shows the number of schedules required for convergence for a range of different β, with 95% confidence interval shown based on a Gaussian approximation. Note the larger graph is on a log scale, while the inset graph is on a linear scale. It can be seen that larger values of β give smaller number of schedules (i.e., faster convergence times). The value of β that gives the fastest convergence time is approximately 1.0. For $\beta > 0.4$ the time to converge to a collision free schedule is substantially shorter than that of L-BEB (learning binary exponential backoff).

Pre-convergence short-term fairness

A second factor that influences the choice of β is its impact on short-term fairness prior to converging to a collision-free schedule. This is a consideration, as convergence may require tens of schedules. As we aim for a symmetric sharing of throughput, we employ Jain's fairness index [44][45][46] to evaluate fairness.

Fairness is solely a function of the sequence of successful transmissions. Consider a network of stations labeled $\{1, \ldots, N\}$. For each simulation we generate the subsequence of K successful slots prior to convergence to a collision-free schedule. We record the the sequence of stations that have successful transmissions, $X_1, \ldots, X_K$, where $X_j \in \{1, \ldots, N\}$. For each $m \in \{1, 2, \ldots, \lfloor K/N \rfloor\}$, where $\lfloor x \rfloor$ denotes the

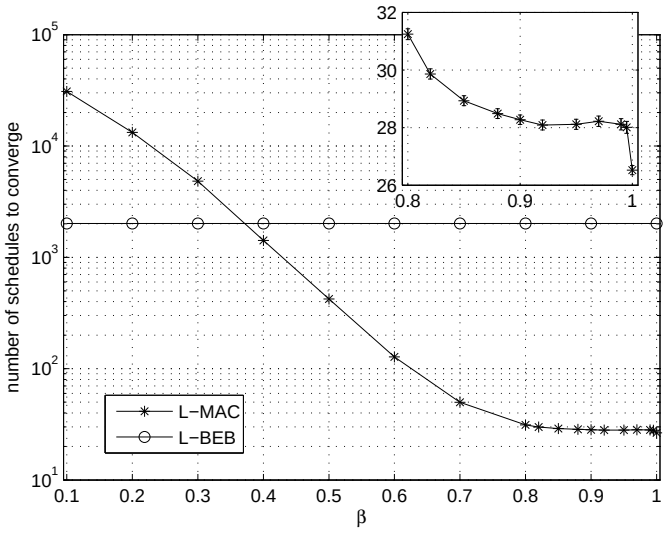


Figure 4.1: L-MAC's convergence time for a range of learning strengths, β, and L-BEB on log scale. $C = 16$, 16 stations. The inset graph shows the detail for $\beta \in (0.8, 1)$ on a linear scale. ns-2 simulations

greatest integer less than x, we consider fairness over windows of $w = mN$ successful transmissions. For each station i and window k of length w, we look at the ratio of the actual number of successes to the number in a perfectly fair allocation:

$$\nu_i(w,k) = \frac{N}{w} \sum_{j=(k-1)w+1}^{kw} 1_{\{X_j=i\}}.$$

Then, for each window, Jain's index is given by

$$F(w,k) = \frac{\left(\sum_{i=1}^{N} \nu_i(w,k)\right)^2}{N \sum_{i=1}^{N} \nu_i(w,k)^2}.$$

Finally we evaluate the empirical average fairness over all windows in the successful transmission sequence:

$$F(w) = \frac{1}{\lfloor K/w \rfloor} \sum_{k=0}^{\lfloor K/w \rfloor - 1} F(w,k).$$

When $F(w) = 1/N$ this corresponds to the worst unfairness. Perfect fairness is obtained when $F(w) = 1$. Note that perfect fairness is achieved by a collision-free schedule and that is why we concentrate on fairness prior to convergence.

In the same scenario as the comparison of the convergence time, a comparison of Jain's fairness index is shown in Figure 4.2. In general, we see that smaller values of β, lead to better fairness, though the relationship is not monotone, as 0.95 and 1

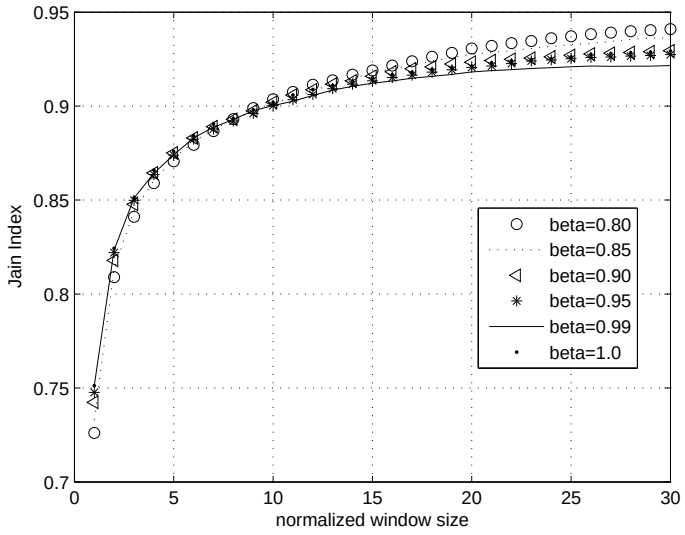


Figure 4.2: Jain's index vs normalized window size, m, L-MAC with different learning strengths β, $C = 16$, 16 stations. ns-2 simulations

both offer better fairness than 0.99. We have seen similar trends in other network configurations, including oversubscribed network where $N > C$ as shown in Figure 4.3.

Rate region of oversubscribed network

We also wish to have reasonable performance when there are more stations than slots. This can allow the network to have reasonable performance while a schedule length adaptation is in progress. We will look at how β effects the achievable throughput during this period. In MACs where there can be a trade off between idle slots and collision slots that the maximum throughput may not be when all stations are saturated. Thus, we consider an unsaturated network with Poisson arrivals at each station and estimate each station's traffic intensity,

$$\rho = \frac{\text{expected service time}}{\text{expected inter-arrival time}}.$$

Note, both arrival times and service times are stochastic. We vary the arrival rate λ and look for the largest λ that gives $\rho < 1$ for all stations [30]. This identifies the stability region when the network is symmetrically loaded. Figure 4.4 shows this upper value of λ as β is varied. This suggests that for an unsaturated network, the largest achievable throughput is available around $\beta = 0.95$.

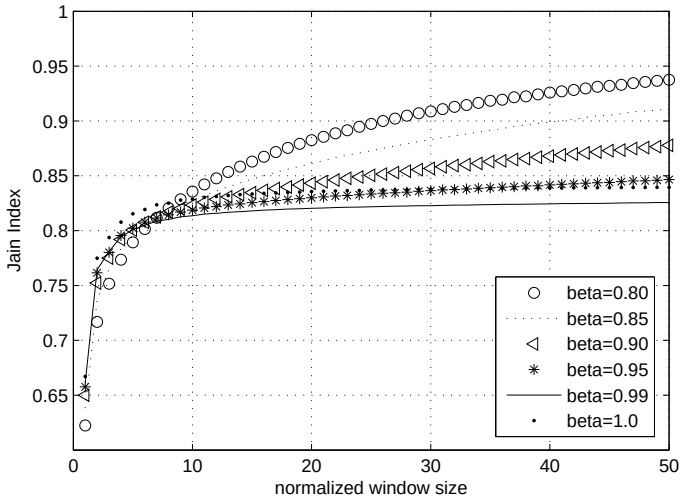


Figure 4.3: Jain's index vs normalized window size (m), L-MAC with different learning strengths β, $C = 16$, 19 stations, ns-2 simulations

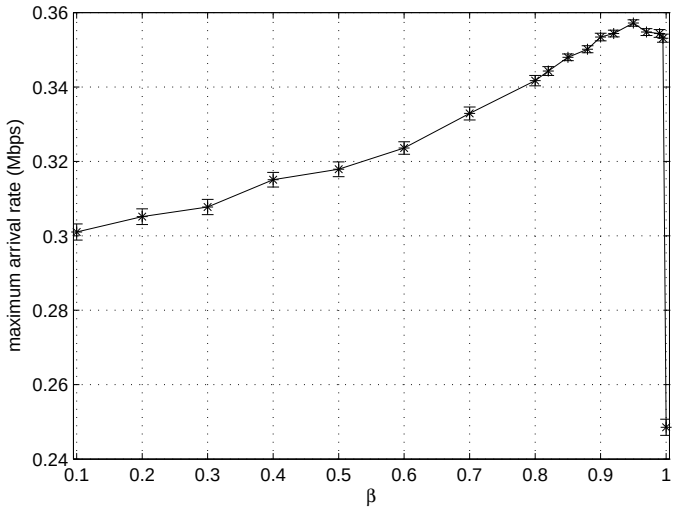


Figure 4.4: Achievable symmetric rate for different values of β. L-MAC, $C = 16$, $N = 20$ stations, ns-2 simulations

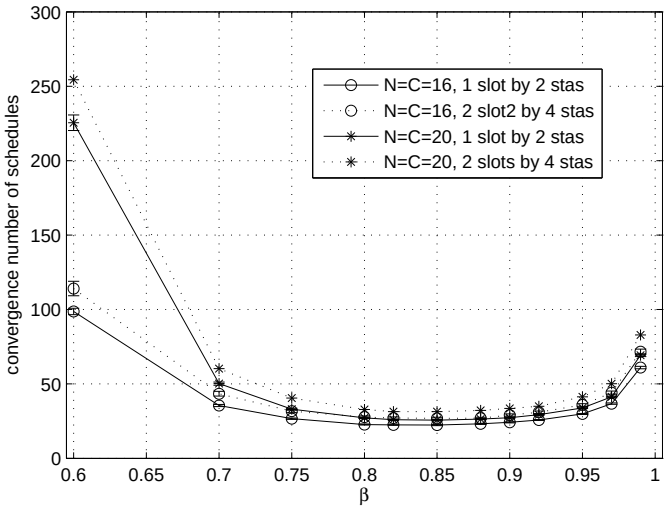


Figure 4.5: Re-convergence number of schedules between L-MAC for a range of learning strengths, β. C=16.

Speed of Reconvergence

Finally, we assess the re-convergence properties as a function of β. An obvious problem with $\beta = 1$ is that if a station ever has a success in a slot, it will stay in the same place in the schedule indefinitely. If two stations believe they have had a success in the same slot in the schedule, which is possible if they are not fully loaded, then they may both try to share a slot, even though free slots may be available. To assess this, we allow a group of stations to converge to a collision-free state. They we add two additional stations at the same moment, which choose the same free slot. We then look at the length of time for the whole network to re-converge to being collision-free. Though this situation is somewhat unlikely, it represents a challenging scenarios for L-MAC. Figure 4.5 shows the results of these simulations as β is varied. No value is shown for $\beta = 1$, because the network will never converge to being collision free. Smaller re-convergence times are possible for a range of β from about 0.75 to 0.95. We also show the results if we introduce two pairs of two stations, and see broadly similar behaviour.

To summarize, convergence time is optimized when $\beta = 1$, but there is only a small reduction for choosing a value in $(0.9, 0.99)$. In contrast, lower β values generally lead to better fairness before convergence, with values at 0.95 and 1 being comparable. When we look at the value of β that maximizes the throughput region when the network is oversubscribed, we find a value around 0.95 is best, though performance

is relatively flat between 0.9 and 1. When we look at the time to reconverge when colliding stations are introduced to an otherwise converge network, we find that rapid convergence is given for $\beta = 0.85$, but again there is a flat region from 0.75 to 0.95.

Consequently, we suggest that L-MAC use $\beta = 0.95$. This offers a good compromize between convergence time, fairness and achievable throughput. We have checked a range of schedule lengths with these metrics, and find that $\beta = 0.95$ remains an appropriate compromize.

4.2.2 Choosing the collision weight γ in L-ZC

We give a more refined analysis that enables us to determine the optimal learning parameter. For each $N_{(C)}$ different configurations of collisions are possible, so we label these by a sequence $S_{(N_{(C)},i)} = (I_1, I_2, \cdots, I_{n_C})$ where i indexes the different states and I_j is the number of stations transmitting in slot j. By relabeling the slots, we only need to consider the case where $I_{j-1} \leqslant I_j$ and we omit slots which have no collision (i.e. $I_j < 2$). For example, for two colliding stations, the only possible state is $S_{(2,1)} = (2)$. When $N_{(C)} = 5$, there are two possible states $S_{(5,1)} = (5), S_{(5,2)} = (2, 3)$. We denote $\overline{S}_{N_{(C)}} := \{S_{(N_{(C)},i)} : i\}$ and $\overline{S} := \bigcup_{N_{(C)}=2}^{N} \overline{S}_{N_{(C)}}$. These sets can be identified by combinatorial search.

These sequences, $S_{(N_{(C)},i)}$, are the states of our Markov chain as depicted in Figure 4.6. We add an initial state IS (N stations start to transmit) and an absorbing state 0 representing collision-free schedules. Note that in this discrete-time Markov chain $S_{(N_{(C)},i)}$ has non-zero probability to transition to state $S_{(k,j)}$ if $k \leqslant N_{(C)}$ and the state IS has positive probability to transfer to all states except itself.

Note that the transition probability from $S_{(N_{(C)},i)}$ to $S_{(k,i)}$ is zero if $k > N_{(C)}$, because $N_{(C)}$ is non-increasing in the next schedule by design. Assume that $G_{N_{(C)}}$ is a $|\overline{S}_{N_{(C)}}| \times |\overline{S}_{N_{(C)}}|$ matrix of transition probabilities among states in $\overline{S}_{N_{(C)}}$ with the same number of colliding stations. Considering the state IS and the absorbing state, we obtain the $(|\overline{S}| + 2) \times (|\overline{S}| + 2)$ full transition matrix Π in upper-triangular block form,

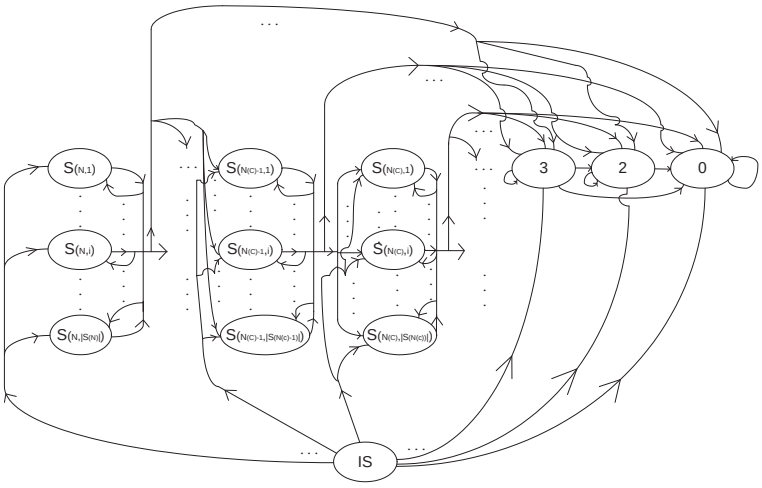


Figure 4.6: Markov chain for L-ZC

$$\Pi = \begin{bmatrix} 0 & P_{12} & \cdot & \cdot & \cdot & \cdot & P_{1(2+|\overline{S}|)} \\ 0 & G_N & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & G_{N_{(C)}} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & G_2 & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix}. \tag{4.1}$$

The initial probability measure for all states $\Phi_{(0)} := [1, 0, \cdots, 0]$, at the n'th schedule $\Phi_{(n)} = \Pi^n \Phi_{(0)}$, and stationary measure is $[0, \cdots, 0, 1]$ due to the absorbing state 0. The convergence speed depends on the second largest eigenvalue λ^* of the transition matrix: the smaller λ^*, the quicker convergence speed. As Π is a upper triangular matrix, the determinant of $\lambda I - \Pi$ is the product of determinants of its diagonal entries, (4.2).

$$|\lambda I - \Pi| = \lambda \prod_{N_{(C)}=2}^{N} |\lambda I - G_{N_{(C)}}|(\lambda - 1). \tag{4.2}$$

It is evident that $\lambda_0 = 0$ and $\lambda_{2+|\overline{S}|} = 1$. In order to get the rest eigenvalues λ, we will evaluate the transition matrix G_{N_C}, and obtain the largest eigenvalue of those matrices which is second largest eigenvalue λ^* of Π.

Let $\pi_{kl}^{N_{(C)}}$ be the entry of $G_{N_{(C)}}$ corresponding to the probability of moving from the state $S_{(N_{(C)},k)} = (K_1, \cdots)$ to state $S_{(N_{(C)},l)} = (L_1, \cdots)$. Let n_C^k and n_C^l be the the number of slots experiencing a collision in these states respectively. Consider colliding stations that choose to remain fixed in the same slot. Since other stations will have seen that slot as busy, no additional stations will be able to move into this slot. This if some of the K_j stations remain fixed, they must correspond to a slot j' with $L_{j'} \leq K_j$. Let $\Omega \subset \{1, \dots n_C^k\}$ represent slots that will have some fixed station and let

$$M(\Omega) := \left\{\sigma : \Omega \to \{1, \dots n_C^l\} : L_{\sigma(j)} \leq K_j, \forall j \in \Omega \text{ and } \sigma \text{ is one-to-one.}\right\}$$

Note that $M(\Omega)$ may be empty. Let $\{j_1, j_2, \dots\} := \{1, \dots n_C^l\} \backslash \sigma(\Omega)$ be the indices of collision slots not arising from fixed stations. The number of stations moving to previously idle slots to produce these collision slots will be

$$m(\Omega, \sigma) := \sum_{j \in \{j_1, j_2, \dots\}} L_j,$$

and the number of ways we can choose the idle slots will be

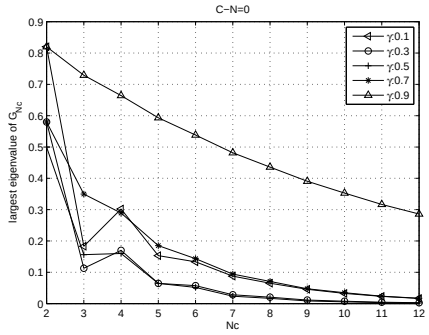
$$P(n_I^k, n_C^L - |\Omega|) := \frac{n_I^k!}{(n_I^k - n_C^L + |\Omega|)!}.$$

So, we may write the transition probability as

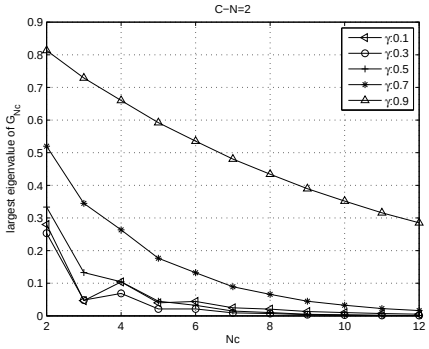
$$\begin{aligned} \pi_{kl}^{N_{(C)}} = & \sum_{\Omega \subset \{1, \dots n_C^k\}} \sum_{\sigma \in M(\Omega)} \left[\prod_{j \in \Omega} \binom{K_j}{L_{\sigma(j)}} \gamma^{L_{\sigma(j)}} \right] \\ & \left[\binom{m(\Omega, \sigma)}{j_1 \ j_2 \ \dots} \left(\frac{1-\gamma}{n_I^k} \right)^{m(\Omega, \sigma)} \right] \frac{P(n_I^k, n_C^l - |\Omega|)}{R}, \end{aligned} \tag{4.3}$$

where R is the number of permutations of the sequence $S_{(N_{(C)},l)}$ that result in the same state. For particular $N_{(C)} \in [2, N]$ and $\gamma \in (0, 1)$, we can obtain the full set of states $\overline{S}_{N_{(C)}}$, obtain the transition matrix $G_{N_{(C)}}$ based on equation (4.3), and then calculate the largest eigenvalue $\lambda^*_{(N_{(C)})}$ of $G_{N_{(C)}}$. Then the second largest eigenvalue will be

$$\lambda^* = \max_{N_C \in [2,N]} [\lambda^*_{(N_C)}]. \tag{4.4}$$



(a) C=N



(b) C=N+2

Figure 4.7: Largest eigenvalue vs $N_{(C)}$ for L-ZC, various γ values, numerical results

Based on this analysis, Figure 4.7(a) and Figure 4.7(b) show the largest eigenvalue of matrix at different $N_{(C)}$ when $N \leqslant C$. We observe that largest eigenvalue λ^* is achieved at $N_{(C)} = 2$ at the same γ. Hence we obtain:

$$\lambda^* = \gamma^2 + \frac{(1-\gamma)^2}{C-N+1}. \tag{4.5}$$

The minimum λ^* is obtained by setting $\gamma = \frac{1}{C-N+2}$. When $N = C$, γ is set at 0.5 for the faster convergence speed for L-ZC.

Using this Markov chain, we can also predict the number of schedules until collision-free schedule is obtained, assuming that all stations start to transmit at the same time. Let Π_T be the transition matrix between all transient states. We have already obtained the diagonal, $G_{N_{(C)}}$ in equation (4.3).

Considering the transition probability from $S_{(N_m,i_m)}$ with n_I^m idle slots to $S_{(N_n,i_n)}$ when $N_m > N_n$, $N_m - N_n$ stations choose successfully their own slots, and rest N_n stations maintain the same collision cases as $S_{(N_n,i_n)}$ in next schedule. Let $\mu \subset \{1, \cdots, N_m\}$ represent stations that will be successful with size $|\mu| = N_m - N_n$. Let

n_C^m is the number of slots experiencing a collision for $S_{(N_m,i_m)}$. We define $A(\mu) := \{n_c^i : j \in n_c^i, \forall j \in \mu\}$ as slots where at least one station will be successful. Let $\mu_s \subset A(\mu)$ be collided slots that will be occupied successfully by stations in μ. Note that stations in μ_s can only choose idle slots because previous collided slots will be successfully occupied. As a complement to μ_s, μ_s^c is available collided slots which can be chosen to achieve collision case as $S_{(N_n,i_n)}$, and the number of their stations is $\mathbf{I}_{\mu_\mathbf{s}^\mathbf{c}} = \sum_{j\in\mu_s^c} I_j^m$. Then the transition probability π_{mn} is the probability $\pi_{N_n}^{rest}$ that N_n stations achieving the same collision cases as $S_{(N_n,i_n)}$, on condition that $N_m - N_n$ stations choose successful slots. $\pi_{N_n}^{rest}$ can be obtained by equation [8] by setting the starting state as $\{I_j : I_j \in \mu_s^c\}$ with rest $N_m - \mathbf{I}_{\mu_\mathbf{s}^\mathbf{c}}$ stations only choosing idle slots.

$$\pi_{mn} = \sum_{\mu\subset\{1,\cdots,N_m\}} \sum_{\mu_s\subset A(\mu)} \gamma^{|\mu_s|}(\frac{1-\gamma}{n_I^m})^{N_m-N_n-|\mu_s|} P(n_I^m, N_m - N_n - |\mu_s|)\ \pi_{N_n}^{rest} \tag{4.6}$$

We do have to calculate the first row of Π_T, representing transition probabilities from IS into other states $S_{(N_{(C)},i)}$. If $N - N_{(C)}$ stations choose their own successful $N_{(C)}$ slots, and n_C slots are chosen from rest $C - N + N_{(C)}$ slots to obtain the same collision case as $S_{(N_{(C)},i)}$ and the probability of choosing each slot is initially $\frac{1}{C}$. Thus we get the transition probability from IS to $S_{(N_{(C)},i)}$ is

$$\pi_{IS,S_{(N_{(C)},i)}} = \binom{C}{N-N_{(C)}} P_{(N,N-N_{(C)})} \binom{C-N+N_{(C)}}{n_C} / R\binom{N_{(C)}}{I_1\ I_2\ \ldots} \left(\frac{1}{C}\right)^N \tag{4.7}$$

where again, R is number of permutations of $S_{(N_{(C)},i)}$ that result in the same collision state.

Let $\kappa_{(S_{(N_{(C)}),i)}}$ denote the number of schedules elapsed before the network reaching collision-free schedule given the initial state $S_{(N_{(C)},i)}$, and $\kappa_{(IS)}$ denote the number of schedules elapsed from state IS. Using standard Markov chain results, the mean number of convergence schedules from initial state IS is obtained as

$$E(\kappa_{(IS)}) = [1, 0 \ldots 0](I - \Pi_T)^{-1}[1, 1 \ldots 1]^T. \tag{4.8}$$

This mathematical analysis of L-ZC allows us to predict the convergence times

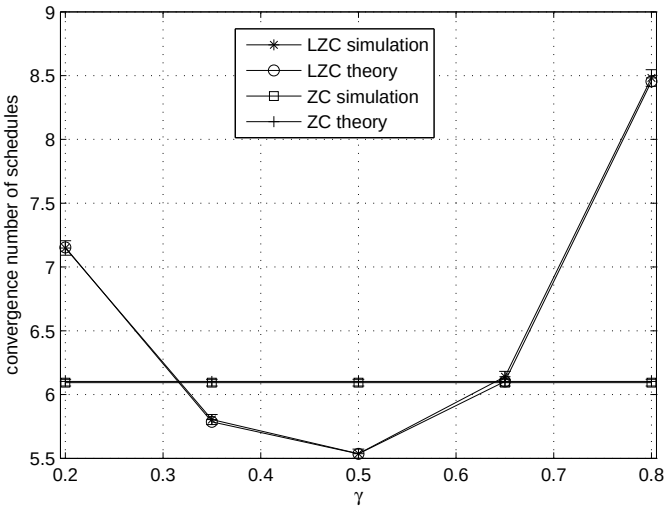


Figure 4.8: Comparison between L-ZC's convergence rate, for a range of γ values, and ZC's convergence rate. $C = 16$, 16 stations, ns-2 simulations and theory

for different values of γ. Figure 4.8 shows how mean convergence times vary as a function of γ. Note that this analysis predicts simulated times accurately, and that the confidence intervals are tight. The optimal value of γ can be explicitly calculated, which is $\gamma^* = 1/(C - N + 2)$; note that when $N = C$ the graph confirms that the shortest convergence time is when $\gamma = 1/2$.

We base our choice of γ purely on optimizing convergence time, because it is so short. Reconvergence of of L-ZC and ZC (Zero Collision) amounts to convergence starting with a smaller number of stations. Thus reconvergence is optimized by optimizing convergence. There will be a period of unfairness during any convergence, but because of the fast convergence, this period will be of no concern.

For a station to choose the optimal γ, it must know $C - N$, which corresponds to the number of idle slots when the scheme converges. This number may be known from above the MAC layer, in which case the exact value can be used. Alternatively, the station might try to estimate this value. It must already know C. If it can sense the difference between collision and successful slots then it could estimate N as the number of successful slots plus twice the number of collision slots. If the station only knows about its own collision slot and estimates N as the number of busy slots plus one, then the L-ZC scheme reverts to the original ZC. For the remainder of the paper, we assume L-ZC knows the value of $C - N$ and use $\gamma = 1/(C - N + 2)$.

4.3 Performance Evaluation

We have implemented these MAC protocols in *ns-2*, as described in Section 3.3. Unless otherwise noted, all stations are transmitting saturated UDP traffic (with payload 1000 bytes) and a PHY rate of 11Mbps. All stations share the same physical channel, where each station can hear each other and there are no hidden nodes. A static routing (NOAH) is employed. When simulating DCF parameters are as for 802.11b. We know from Bianchi's model [3] that for lower collision rates the transmission probability will be approximately $2/(CW + 1)$ close to $1/16$. Thus, when working with a fixed schedule length, we work with $C = 16$, to allow comparisons with DCF. All simulation results are obtained as mean values over repeated simulations with different seeds. Error bars based on the central limit theorem are not shown on the graphs as they are on a similar scale to the symbols used for plotting points.

We expect results from DCF, L-BEB and L-MAC to be comparable, as they work with the same information. Likewise, we also expect ZC and L-ZC to be comparable, because they both leverage extra information not available to the other MACs. When applying the adaptive schedule length schemes, we will refer to the MACs as A-L-MAC, A-ZC and A-L-ZC.

4.3.1 Speed of Convergence

In a static network we record the elapsed time before the collision-free schemes reach a collision-free state (no results are shown for DCF, as it does not converge). Figure 4.9 shows this as the ratio N/C is varied. We see that for small numbers of stations, all the algorithms converge in a fraction of a second. However, when the number of stations becomes close to C, we can see the advantages of the L-MAC over L-BEB. In particularly, observe that using learning has reduced the convergence time of hundreds of seconds for L-BEB to under a second for L-MAC. We can see advantage of the ZC-based schemes over both L-BEB and L-MAC. In fact, on the log scale shown, the performance of L-ZC and ZC are comparable. L-MAC, in particular, is performing well for an algorithm working with less information.

4.3.2 Long Term Throughput

For the non-adaptive schemes, when $N > C$ collisions are unavoidable. In Figure 4.10 we compare the collision rates of conventional DCF and the learning schemes. Note

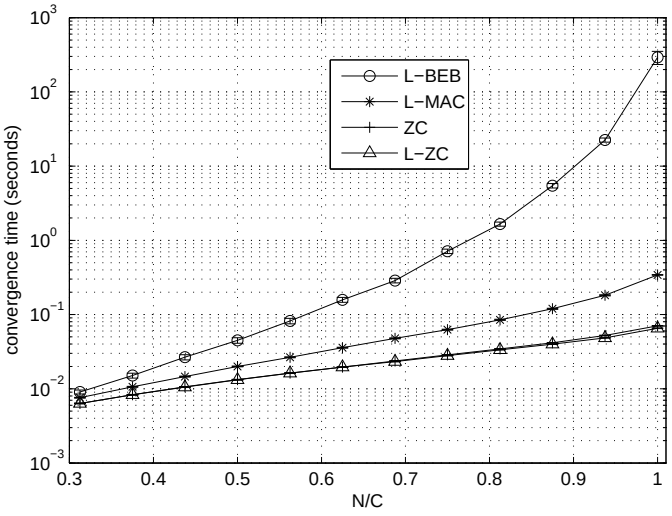


Figure 4.9: Convergence Time vs ratio number of stations to C, Comparison between learning MACs, from 5 stations to 16 stations. ns-2 simulations, error bars too small to be shown

that L-BEB has twice the collision rate as DCF in a comparable. In contrast, L-MAC degrades gradually with a lower collision rate than DCF's while the number of stations is from 17 to 19. ZC and L-ZC offer a slightly lower collision rate again Figure 3.1 shows the corresponding results for throughput. This demonstrates that our learning MACs can achieve good channel utilization with lower collision probability than CSMA, even if collisions persist.

We also investigate the performance of the adaptive schemes for more than 16 stations. As expected A-ZC and A-L-ZC, they all achieve a long-term collision rate of zero. Figure 4.11 shows that A-ZC and A-L-ZC have essentially the same performance, and A-L-MAC lags only slightly behind. Both adaptive learning schemes offer substantially higher than that of DCF. Comparing Figure 3.1 and Figure 4.11, we see how adapting the schedule length allows the performance of the schemes to scale beyond a fixed schedule length. While A-L-MAC shows a slight decline in throughput for $N > 16$, it outperforms all the non-adaptive schemes. A-L-ZC's performance continues to increase as the relative proportion of idle slot decreases.

4.3.3 Transmitted power

This section considers the total transmitted power of those MAC schemes and DCF as depicted in Figure 4.12 and Figure 4.13. For fixed schedule length, by avoiding collisions, the learning MACs' transmitted power is less than DCF's counterpart when

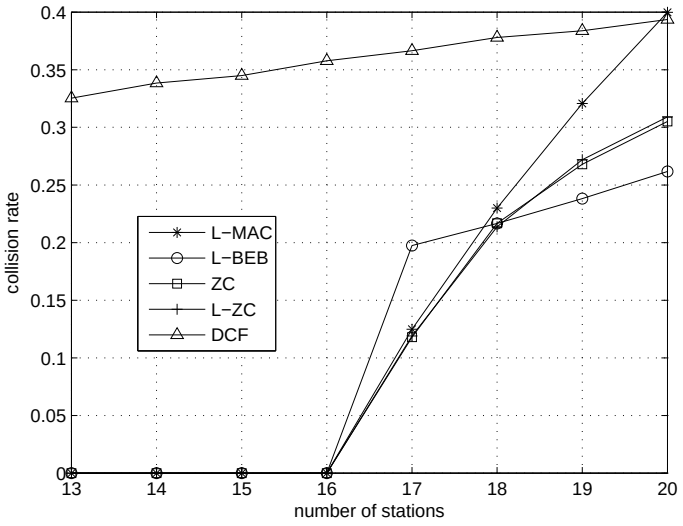


Figure 4.10: Collision Rate vs. number of stations, comparison of MACs, ns-2 simulations

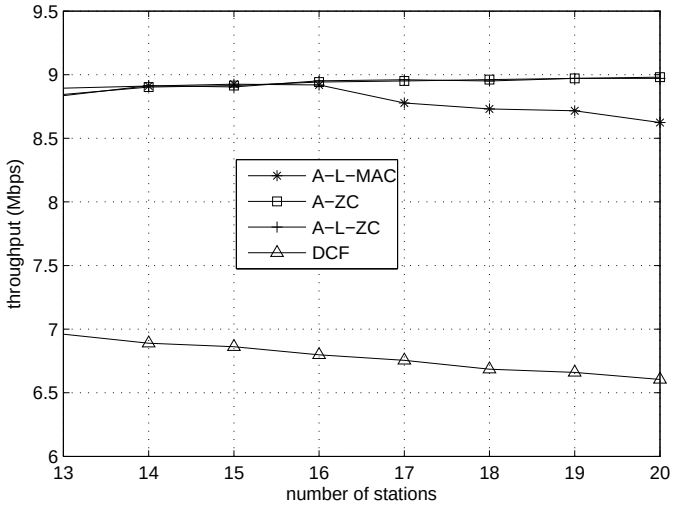


Figure 4.11: Network throughput vs. number of stations, comparison of learning MACs with adaptive schedule length, ns-2 simulations

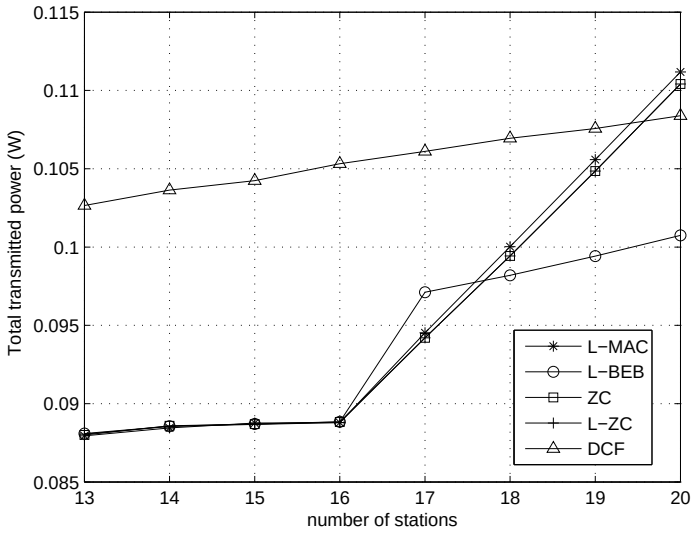


Figure 4.12: Transmitted power vs. number of stations, comparison of MACs, ns-2 simulations

$N \leqslant 19$, and increase substantially when $N > C$. We also investigate the power of the adaptive schemes for more than 16 stations. As A-ZC and A-L-ZC achieve a collision-free schedule, the transmitted power of them is similar and slightly lower than the power output of A-L-MAC which could not avoid collisions completely. All adaptive learning schemes offer lower power output than DCF.

4.3.4 Performance in presence of errors

In previous graphs we have considered the case of a clean channel where no packets are lost to noise or interference, and all losses are due to collisions. A more realistic setting is considered by introducing errors caused by a fading channel [47]. We consider a simple model where errors are introduced at a particular rate (1% and 10%). Errors present an interesting challenge to the learning schemes, because they use transmission failure as an indication of a slot being occupied.

Figure 4.14 shows the achieved throughputs for the fixed schedule length learning MACs. We note that DCF's performance is only slightly degraded by the presence of errors. As all of L-BEB's state is related to the success of the current transmission, if suffers quite badly in the presence of errors and its performance can fall below that of DCF. L-MAC, ZC and L-ZC are more robust to the presence of errors because their memory is not limited to the success of a single slot. L-MAC's learning memory will tend to restore the correct schedule after an error, whereas ZC and L-ZC can see that

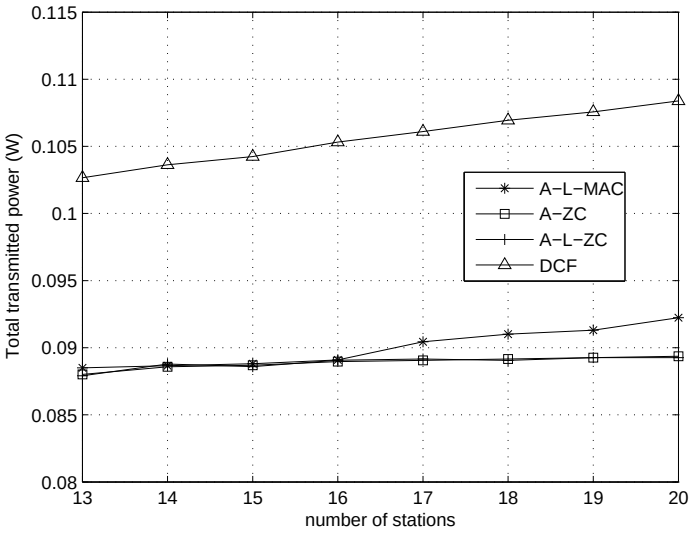


Figure 4.13: Transmitted power vs. number of stations, comparison of MACs with adaptive schedule length, ns-2 simulations

other slots have been allocated and do not move to other slots. Indeed, we can see that $N = 15$ is one of the most challenging cases for L-ZC and ZC, because there will typically be one slot available, which several stations will be drawn to in the case of multiple errors in the same schedule.

We have also investigated the performance of the adaptive schedule length schemes as depicted in Figure 4.15. As expected, the adaptive schemes offer comparable performance to their non-adaptive equivalents, with a small increase in performance after 16 stations, where the extra slots help accommodate churn caused by random losses.

4.3.5 Robustness to New Entrants

In this section, we briefly consider what happens when the network has converged, and then more stations are added. We naturally expect that the improved convergence will extend to quick convergence when more stations are added to the network. Figure 4.16 shows the time to reconverge to a collision-free schedule after new stations are added to a collision-free schedule with 8 stations. As expected, we see rapid convergence, of around one second, even when 8 stations are added to the network at the same time.

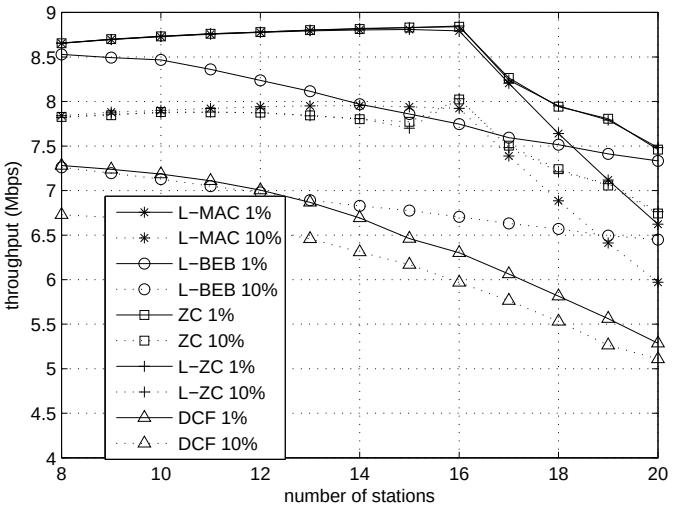


Figure 4.14: Network throughput vs number of stations with errors, comparison between DCF and MACs with fixed schedule length, ns-2 simulations

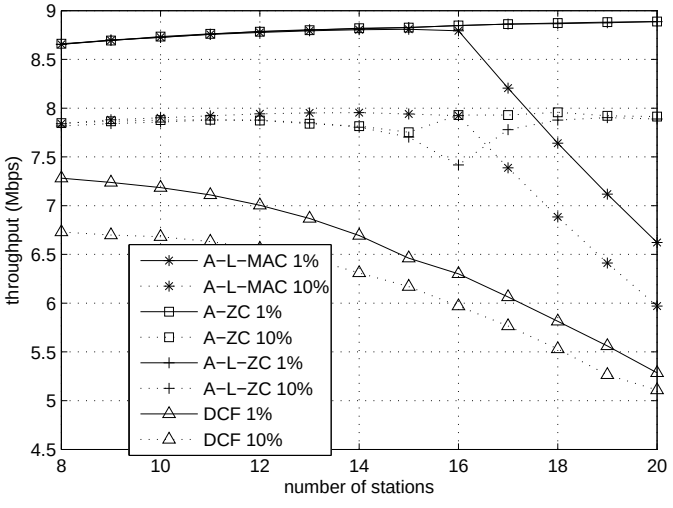


Figure 4.15: Network throughput vs number of stations with errors, comparison between DCF and MACs with adaptive schedule length, ns-2 simulations

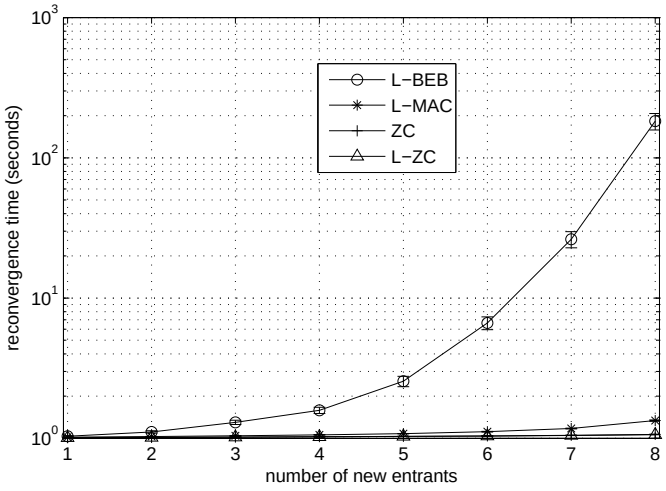


Figure 4.16: Reconvergence time when eight stations are in the network and a variable number of stations are added. ns-2 simulations

4.3.6 Coexistence with 802.11 DCF

This section considers the performance of multiple MAC protocols employed simultaneously by the stations which share the same wireless channel. In particular, we note that as all these MACs are based on the basic channel-sensing techniques of DCF, any of our MACs should be able to coexist with DCF. Coexistence is a significant feature of the MACs, because it allows incremental deployment.

We consider a scenario where we have $2K$ stations in the network. Of these stations K use the DCF protocol and K use another protocol. All the stations are saturated. Figure 4.17 shows the aggregate network throughput achieved as K is varied. The line for DCF+DCF is our baseline, where all stations use the DCF protocol. We see that the mixed networks outperform DCF alone for small numbers of stations. The mix of L-BEB+DCF's performance decreases below that of DCF alone when $K > 8$, when there will be 16 stations in the network. The other schemes, as they have more learning state, retain their performance until $K = 16$. Here, the schedule will be full for the non-adaptive schemes. We can also see that the adaptive schemes offer slightly lower throughput compared to the adaptive ones just below $K = 16$, because they begin to increase their schedule length.

The question of how this throughput is shared is also important. The throughput achieved by the DCF stations is shown in Figure 4.18, compared to the second MAC stations' counterparts depicted in Figure 4.19. We see that DCF throughput is substantially reduced by the presence of large numbers of stations using a different

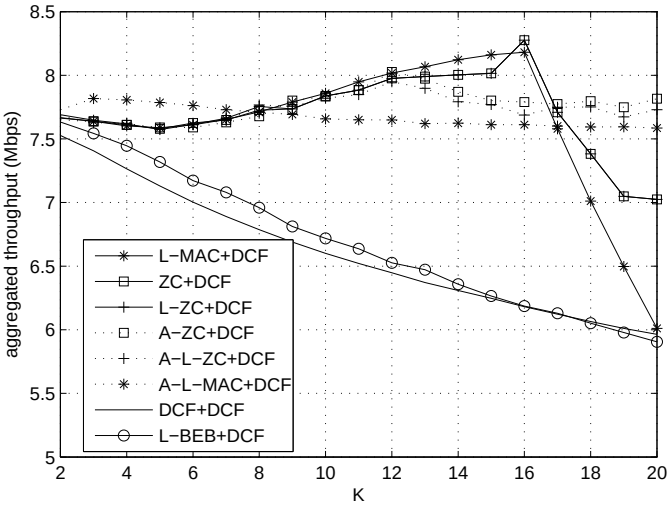


Figure 4.17: Network throughput for a mixed network of mixed MACs. Half the stations use DCF, and the other half a second MAC. ns-2 simulations

MAC, compared to other stations running DCF. Their only respite is when the adaptive schemes begin to increase schedule length, making space for the DCF stations to transmit. A-L-MAC responds to the persistent collisions similarly to a DCF backoff, and so shares more evenly with DCF.

These results suggest that incremental deployment of these new MAC protocols would be possible, at the cost of reduced performance for legacy DCF equipment.

4.4 Summary

In this Chapter we have proposed the choice of learning parameter β for L-MAC and γ for L-ZC in order to optimize their performance. For L-MAC we will use simulations to consider factors such as transient fairness and achievable throughput, as well as convergence time, ultimately choosing $\beta = 0.95$. For L-ZC convergence times are much shorter and we will use mathematical analysis to show convergence times are minimized by selecting $\gamma = 1/(C - N + 2)$. We also show a wide variety of simulation results to compare the performance of learning MACs and existing MACs. Improvements achieved by L-MAC and L-ZC over DCF and even L-BEB are substantial, with reduced convergence times, graceful degradation in the presence of too many stations, reduced transmitted power and improved robustness to channel errors.

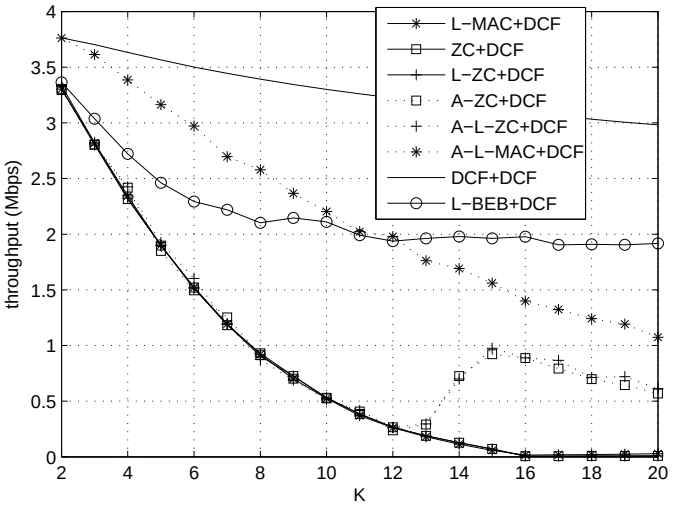


Figure 4.18: Achieved throughput of DCF stations in a mixed network of mixed MACs. Half the stations use DCF, and the other half a second MAC. ns-2 simulations

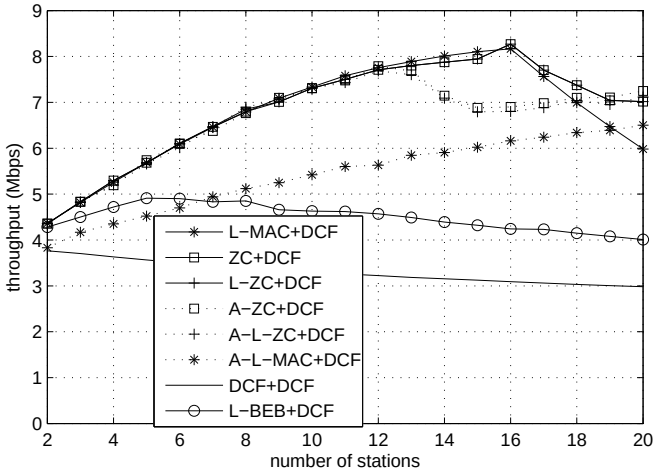


Figure 4.19: Achieved throughput of second MAC stations in a mixed network of mixed MACs. Half the stations use DCF, and the other half a second MAC. ns-2 simulations

Chapter 5

Conclusions

The objective of this thesis is to investigate two issues related to the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) based Wireless Local Area Networks (WLANs): evaluation of transmitted power in IEEE 802.11 WLANs, and performance enhancement of CSMA/CA based WLANs. Our contributions are twofold:

1. Considering the public concern regarding health and safety issues related to exposure to radio frequency (RF) energy by IEEE 802.11 WLANs, in Chapter 2 we introduced models for predicting the power output of WLANs in unsaturated conditions by extending the existing power model in [17] and carried out experimental measurements to verify the theoretical models. Experimental results demonstrate that there is a good match between the theoretical predictions and the experimental results. Moreover, we give a simple technique to quickly predict power output based on traffic levels. Most importantly, the results confirm that the estimated maximum power is substantially lower than the acceptable international limit given by International Commission on Non-Ionizing Radiation Protection (ICNIRP). This work was performed in collaboration with D. Malone (NUIM) and is published, in part, in Fang and malone [18].

2. As the most commonly employed Multiple Access Control (MAC) in the practical hardware of WLANs, the IEEE 802.11 Distributed Coordination Function (DCF) regulates the random backoff process to decrease the likelihood that the stations collide again. Its disadvantage is a persistent non-zero chance of collision. In particular, when a large number of stations are accessing the medium, the network throughput is substantially degraded [3]. In order to overcome this

disadvantage, in Chapter 3 we proposed totally decentralized learning MAC schemes by adopting the self-managed decentralized channel selection algorithm [21] based on recently introduced L-BEB (Learning Binary Exponential Back-off) [19] and ZC (Zero collision) [20]. We proved the convergence of new access methods to the collision-free operation, if one exists. By eliminating collisions, the network throughput is substantially higher than DCF. By applying learning, we have been able to reduce convergence times by several orders of magnitude. Finally, a schedule length adaption scheme is introduced in a decentralized manner, but achieving long-run fairness and scalability of these new MACs beyond a fixed number of stations. This work was performed in collaboration with K. Duffy, D. Malone and D. Leith (NUIM). It is being prepared for submission for publication.

There are interesting issues yet to be resolved in the future. Firstly, rate control methodologies are currently available in 802.11 network cards. The impact of rate control on the transmitted power has not yet been considered; Secondly, we choose β for Learning MAC (L-MAC) using simulations. An analytical study of convergence times of L-MAC is not given but could add interesting insight. The analysis of adaptive schemes are not considered as well; Finally, it would be interesting to see the experimental performance of learning MAC schemes in our IEEE 802.11 testbed.

Bibliography

[1] I.F. Akyildiz, X. Wang, and W. Wang. Wireless mesh networks: a survey. *Computer Networks*, 47(4):445–487, 2005.

[2] V.K. Garg. Wireless communications and networking. 2007.

[3] G. Bianchi. Performance analysis of the IEEE 802. 11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3):535–547, 2000.

[4] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma. Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement. In *IEEE INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, volume 2, 2002.

[5] D. Malone, K. Duffy, and D. Leith. Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions. *IEEE/ACM Transactions on Networking*, 15(1):159–172, 2007.

[6] LAN Wireless. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications. *IEEE std*, 802.11, 1999.

[7] LAN Wireless. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Higher-speed physical layer extension in the 2.4 GHz band. *IEEE std*, 802.11b, 1999.

[8] LAN Wireless. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Further Higher Date Rate Extension in the 2.4 GHz band. *IEEE std*, 802.11g, 2003.

[9] LAN Wireless. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Higher-speed physical layer extension in the 5 GHz band. *IEEE std*, 802.11a, 1999.

[10] LAN Wireless. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: medium access control (MAC) quality of service (QoS) enhancements. *IEEE std*, 802.11e, 2004.

[11] International Commission on Non-Ionizing Radiation Protection. Exposure to high frequency elextromagnetic field, biological effects and health consequences (100 kHz-300 GHz). ICNIRP 16/2009. Available at `http://www.icnirp.org/documents/RFRwview.pdf`, Accessed August 2009.

[12] G. Schmid, P. Preiner, and D. Lager. Exposure of the general public due to wireless LAN applications in public places. *Radiation Protection Dosimetry*, 124(1):48, 2007.

[13] K.R. Foster. Radiofrequency exposure from wireless LANs utilizing Wi-Fi technology. *Health Physics*, 92(3):280, 2007.

[14] A. Peyman, C. Calderon, S. Mann, M. Khalid, D. Addison, and T. Mee. Evaluation of Exposure of School Children to Electromagnetic Fields from Wireless Computer Networks (Wi-Fi): Phase 1 Laboratory Measurements, 2009.

[15] Australia Center for RF Bioeffects Research. EME in the Home. SW2008 Wireless and Health Unplugged and Uncertain. Brain Science Institute, Swinburne University, 2008.

[16] Kramer A Khn S, Lott U and Kuster N. Assessment of Human Exposure to Electromagnetic Radiation from Wireless Devices in Home and Office Environments. Available at `http://www.who.int/pen-emf/meetings/archive/bsw_kuster.pdf`, Accessed August 2009.

[17] D. Malone and L.A. Malone. Ambient radiofrequency power: the impact of the number of devices in a Wi-Fi network. *Health physics*, 96(6):629, 2009.

[18] M. Fang and D. Malone. Experimental Verification of A Radiofrequency Power Model for Wi-Fi Technology. *Health Physics*, 98(4):574, 2010.

[19] J. Barcelo, B. Bellalta, C. Cano, and M. Oliver. Learning-BEB: Avoiding Collisions in WLAN. *Eunice Summer School*, 2008.

[20] J. Lee and J. Walrand. Analysis and design of a zero collision mac protocol. Technical Report UCB/EECS-2007-63, UC Berkeley, 2007.

[21] D.J. Leith and P. Clifford. A self-managed distributed channel selection algorithm for WLANs. In *Proceedings of International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pages 1–9, 2006.

[22] K.R. Duffy and A.J. Ganesh. Modeling the impact of buffering on 802.11. *IEEE Communications Letters*, 11(2):219–221, February 2007.

[23] J.R. Norris. *Markov Chains*. Combridge University Press, 1998.

[24] K.D. Huang, K.R. Duffy, D. Malone, and D.J. Leith. Investigating the validity of IEEE 802.11 MAC modeling hypotheses. In *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008*, pages 1–6, 2008.

[25] K.D. Huang, K.R. Duffy, and D. Malone. On the validity of IEEE 802.11 MAC modeling hypotheses. Technical report, To appear in IEEE/ACM Transactions on Networking.

[26] A. Kumar, E. Altman, D. Miorandi, and M. Goyal. New insights from a fixed point analysis of single cell IEEE 802.11 WLANs. In *IEEE INFOCOM*, volume 3, page 1550, 2005.

[27] V. Ramaiyan, A. Kumar, and E. Altman. Fixed point analysis of single cell IEEE 802.11e WLANs: uniqueness and multistability. *IEEE/ACM Transactions on networking*, 16(5):1080–1093, 2008.

[28] K.R. Duffy. Mean Field Markov Models of Wireless Local Area Networks. *To appear in Markov Process and Related Fields*, 2009.

[29] D. P. Bertsekas and R. G. Gallagher. *Data Networks*. Longman Higher Education, 1986.

[30] S. Asmussen. *Applied Probability and Queues*. Springer, second edition, 2003.

[31] Soekris Engineering. net4801 series boards and systems user's manual. `http://www.soekris.com/`, Accessed August 2009.

[32] Multiband Atheros Driver for WiFi (MADWiFi). `http://sourceforge.net/projects/madwifi/`, Accessed August 2009.

[33] Navy Research Lab. Multi-generator (mgen) version 4.2. `http://pf.itd.nrl.navy.mil/mgen/mgen.html`, Accessed August 2009.

[34] P. Gill, M. Arlitt, Z. Li, and A. Mahanti. Youtube traffic characterization: a view from the edge. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. ACM, 2007.

[35] K. R. Duffy, N. O'Connell, and A. Sapozhnikov. Complexity analysis of a decentralised graph colouring algorithm. *Information Processing Letters*, 107(2):60–63, 2008.

[36] K. Narendra and M. A. L. Thathachar. *Learning Automata: An Introduction*. Prentice Hall, 1989.

[37] I. Rhee, A. Warrier, M. Aia, J. Min, and M.L. Sichitiu. Z-MAC: a hybrid MAC for wireless sensor networks. *IEEE/ACM Transactions on Networking*, 16(3):511–524, 2008.

[38] C. Busch, M. Magdon-Ismail, F. Sivrikaya, and B. Yener. Contention-free MAC protocols for wireless sensor networks. *Lecture Notes in Computer Science*, 3274:245–259, 2004.

[39] P. Wang and W. Zhuang. A collision-free MAC scheme for multimedia wireless mesh backbone. In *IEEE International Conference on Communications, 2008. ICC'08*, pages 4708–4712, 2008.

[40] Y.J. Chen and H.L. Wang. Ordered CSMA: a collision-free MAC protocol for underwater acoustic networks. *Oceans 2007*, pages 1–6, 2007.

[41] H. Yong, Y. Ruxi, S. Jie, and G. Weibo. Semi-Random Backoff: Towards Resource Reservation for Channel Access in Wireless LANs. In *Proceedings of International ICNP*, 2009.

[42] P. Starzetz, M. Heusse, F. Rousseau, and A. Duda. Hashing Backoff: A Collision-Free Wireless Access Method. In *Proceedings of IFIP Networking*, volume 5550, pages 11–15. Springer, 2009.

[43] M. Heusse, F. Rousseau, R. Guillier, and A. Duda. Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless LANs. *ACM SIGCOMM Computer Communication Review*, 35(4):121–132, 2005.

[44] R. Jain, D. Chiu, and W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. *DEC Research Report TR-301*, 1984.

[45] C.E. Koksal, H. Kassab, and H. Balakrishnan. An analysis of short-term fairness in wireless media access protocols (poster session). In *Proceedings of the 2000 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 118–119. ACM New York, NY, USA, 2000.

[46] G. Berger-Sabbatel, A. Duda, O. Gaudoin, M. Heusse, and F. Rousseau. Fairness and its impact on delay in 802.11 Networks. In *IEEE GLOBECOM*, 2004.

[47] Q. Ni, T. Li, T. Turletti, and Y. Xiao. Saturation throughput analysis of error-prone 802.11 wireless networks. *Wireless Communications and Mobile Computing*, 5(8):945–956, 2005.