Modelling Interference in a CSMA/CA Wireless Network

Athanasia Tsertou

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
December 2006

Abstract

The last couple of years were characterised by the revival of interest in the performance evaluation of wireless networks. Due to the broadcast nature of the wireless medium, interference caused by simultaneous transmissions is an important factor of throughput degradation in Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) based networks.

The present thesis aims to develop suitable mathematical modelling methods for the analysis of the network performance. In fact, the proposed models lie on either side of the line between high modelling accuracy and scalability.

Initially, a systematic characterisation of all the possible ways in which two communicating pairs of nodes can interfere with each other is made. Using this as a building block and assuming independence of the stations, an estimate for the network throughput can be derived. The latter proves to be quite accurate for networks with a symmetric connectivity graph and manages to follow the performance trends in an arbitrary network.

Following this, a more detailed Markovian-based mathematical model is proposed for the analysis of the hidden node case. This approach does not rely on common assumptions, such as renewal theory and node synchronisation, and is highly accurate, independently of the system parameters, unlike prior methods. Moreover, the usual decoupling approximation is not adopted; on the contrary, a joint view of the competing stations is taken into consideration. The model is firstly developed based on the assumption that the network stations employ a constant contention window for their backoff process. However, later in the thesis, this assumption is relaxed and performance curves are derived for the case when the stations employ the Binary Exponential Backoff scheme, as is the case in practice. The Markovian state space is kept relatively small by employing an iterative technique that computes the unknown distributions. The adoption of this technique makes the analysis computationally efficient.

Declaration of originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the Department of Electronics and Electrical Engineering at The University of Edinburgh, unless explicitly stated otherwise.

Athanasia Tsertou
December 2006

To my loved ones

Acknowledgements

This thesis aims to show the outcome of my research endeavours in attempting to solve a quite challenging problem which intrigued me towards the end of the first year of my PhD. It was a really interesting journey, through which I have, hopefully, been transformed from a passive student, that is merely asked to solve given and well-known problems, to a researcher whose primary goal is to identify problems. I know I still have a long way to go, but I feel that these three years have really changed my perception of science and the world in general.

Throughout this painstaking—I must admit—period of my life, a number of people have assisted me in my endeavours, whose help I would like to take the opportunity to acknowledge in this prologue. First of all, I would like to thank my supervisor, Dr. David Laurenson for his continuous support, optimism and valuable comments during these last three years. I know I was not an easy student to deal with. I would also like to acknowledge the assistance and sincere interest of Prof. Jane Hillston and Dr. Konstantinos Drakakis, who offered me their expertise in mathematical modelling of stochastic processes. A special thanks to both my thesis examiners, Dr. Mischa Dohler and Prof. Mike Davies for their invaluable comments, that helped me improve significantly the quality of this thesis. Moreover, I would also like to express my gratitude to the faculty members Dr. John Thompson, Dr. James Hopgood, Prof. Steve McLaughlin and Dr. Norbert Goertz for the useful discussions we have had, during or outside the group seminars. Many thanks to the NEWCOM project that has funded my internship to Politecnico di Torino, which undoubtedly had a very strong influence (both negative and positive!) on the route that this thesis has followed. Finally, I should not forget to mention the computing facilities and the people behind them, whose help has always been invaluable.

I would also like to thank my friends, that in one way or another helped me submit this thesis: First of all, Kostis—without your encouragement and motivation I would not have applied to the University of Edinburgh for a PhD. Anastasia, Maria, Anastasis, Luis, Giorgos, Ana, Michalis and Cristina—your love, friendship and sentimental support has been necessary for me to go on, when the problems seemed unsolvable and my self-confidence was below zero. Elias—your assistance and support whenever I needed them, especially in the last stage of writing the thesis, means a lot to me.

Finally, I would like to express my infinite love and gratitude to my family, and especially, my parents and my sister, because you made it possible for me to come to Edinburgh, and although you are far away, you could not be closer. Thank you for your endless sacrifices and for always giving me a reason to fight and accomplish my goals. A special thanks to my father to whom—I think—I owe the research stubborness and the love for maths.

Last but not least, I wish to thank Yannis. If it was not for you, I would have quit the effort during the hard times. You have been my source of inspiration. I really appreciate your confidence in me, and the unlimited help you gave me with your always valuable comments. I admire you both as a person and as a researcher, and I am confident that all your professional aims will be accomplished soon. Thank you for always being on my side.

Athanasia Tsertou
Edinburgh, December 2006.

Contents

List of figures

List of tables

Acronyms and abbreviations

ACK	Acknowledgement
AP	Access Point
BEB	Binary Exponential Backoff
BTMA	Busy Tone Multiple Access
CBR	Constant Bit Rate
CCA	Clear Channel Assessment
CCW	Constant Contention Window
CDMA	Code Division Multiple Access
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
CTS	Clear To Send
CW	Contention Window
DCF	Distributed Coordination Function
DIFS	Distributed Coordination Function Interframe Space
DRV	Discrete Random Variable
DSSS	Direct Sequence Spread Spectrum
DTMC	Discrete Time MC
EDCA	Enhanced Distributed Channel Access
EIFS	Extended Interframe Space
EMBRP	Enhanced MBRP
FDMA	Frequency Division Multiple Access
FHSS	Frequency Hopping Spread Spectrum
IEEE	Institute of Electrical and Electronics Engineers
ISM	Industrial Scientific and Medical band
LRL	Long Retransmission Limit
MAC	Medium Access Control
MANET	Mobile (Multihop) Ad Hoc Network
MBRP	Model-Based Resource Prediction
MC	Markov Chain

NAV	Network Allocation Vector
NS	Network Simulator
PDA	Personal Digital Assistant
PMF	Probability Mass Function
PRN	Packet Radio Network
QoS	Quality of Service
R	Receiver
RAM	Random Access Memory
RTS	Ready To Send
Rx	Receiver
SIFS	Short Interframe Space
SRL	Short Retransmission Limit
T	Transmitter
TDMA	Time Division Multiple Access
Tx	Transmitter
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Network

Nomenclature

α	probability that collision occurred in s_{n-2} before the success in s_{n-1}
β	probability that success occurred at state s_{n-2} before the success in s_{n-1}
Γ	set of absorbing states
γ_i	steady-state probability of state i
$\hat{\gamma}_i$	steady-state probability of state i from the sender's viewpoint
$\hat{\boldsymbol{\gamma}}$	steady-state probability vector from the sender's viewpoint
$\gamma(E,i,j)$	steady-state probability of state (E,i,j)
ΔC	difference of the backoff counters when one of the senders attempts transmission
Δt	arbitrary time interval
ϵ	arbitrarily small positive number
η	probability equal to $1-2p$
$\boldsymbol{\xi}$	vector of backoff-dwelling probabilities
ξ_i	probability that the node is in backoff stage i
$\pi(k)$	steady-state probability of state k
$\pi(\Delta C,i,j)$	steady-state probability of state $(\Delta C,i,j)$
$\pi_{k,m}$	probability that the stage is k and the counter is m, non-freezing
π_m	probability that the counter is equal to m, non-freezing (CCW)
$\pi^i_{k,m}$	probability that the counter of stage k and value m froze and $i-1$ slots elapsed since
π^i_m	probability that the counter of value m is frozen and $i-1$ slots elapsed since (CCW)
σ	time-slot
τ	transmission probability
τ'	transmission probability with the idle slots discretised
Ω	universal set
Ω_{TR}	set of two-hop neighbours of T or R that are not one-hop neighbours of either of them
A	term $\frac{(1-p)(L-c)+(W-1)/2}{(1-p)L+pC}$
$\mathbf{AR}$	aggregate reward vector until absorption
ar_k	aggregate reward until absorption for state k
BC	DRV that describes the value of the backoff counter
b_{ik}	steady-state probability that the station is in backoff stage i and has a counter of k

c	half of the vulnerable period
C	length of the collision period expressed in slots
$C_{A(B)}$	value of the backoff counter of station A(B) when either of them starts transmission
cs	number of collision slots in an obervation interval
D	delay that a packet experiences until it is sent successfully or it is dropped
$d_{1,2}$	number of offset slots in Figure 4.2
$d(i,j)$	physical distance between nodes i and j
d_g	length of the diagonal in the grid topology
d_s	distance between two succesive transmitters in the star topology
$\mathbf{e}$	column vector with ones
E	set that comprises the events success and collision
$E[.]$	expectation operation
E_i	event that T-R suffer from interference from an IP of type i given that T transmits to R
$\bar{E}_i$	complementary of E_i
e_{rel}	relative error of conventional methods compared to simulation results
F	length of the freezing period expressed in slots
F_I^E	PMF of the RV I conditioned on E
G	CTS + SIFS + DATA − CTS_Timeout − RTS − DIFS
$\mathbf{h}$	vector of sojourn times
$\hat{\mathbf{h}}$	vector of sojourn times from the sender's viewpoint
h_s	sojourn time of state s
$\hat{h}_s$	sojourn time of state s from the sender's viewpoint
H_R	set of one-hop neighbours of R that are hidden from T
H_T	set of one-hop neighbours of T that are hidden from R
i	general index
$i \oplus 1, i'$	(i+1)mod(RL+1)
i^*	general index for which $i^* \neq i$
I	set of initial states, general RV
$\mathbf{I}$	identity matrix
is	number of idle slots in an observation interval
j	general index
k	general index
K	number of backoff slots that elapse before the counter of a station freezes

K_c	K conditioned on a collision preceding
K_s	K conditioned on a success preceding
l	dimension of the transition matrix P
l^*	dimension of the transition matrix in the detailed model
L	length of the successful period expressed in slots
L_d	the DATA packet length
m	maximum backoff stage, general index
n	number of stations in a fully-connected network, general index
N_{ca}	number of collisions when one observes the channel around A in steady-state
N_i	set of one-hop neighbours of station i
N_{ij}	set of stations in range of both i and j
N_{sa}	number of successes when one observes the channel around A in steady-state
N_i^2	set of the two-hop neighbours of station i
p	probability that a CSMA node transmits at the start of the slot (Section 2.1.1)
p	probability of collision given that the node transmitted
$\mathbf{p}$	vector of conditional collision probabilities
P	transition matrix of the embedded DTMC of Chapter 5
$P\{.\}$	probability of the event .
$p_.(i,j)$	$p_.$ when the backoff stages are i, j
p'	probability that R observes a collision conditioned on any one of the senders transmitting
p_b	probability that the node freezes its counter when it senses the medium busy
p_c	probability that the node observed/suffered from a collision given that it saw a busy slot
p_{cc}	probability that a collision occurs given that a collision preceded
p_{co}	probability that a station did not get a response to it sending an RTS
p_{cs}	probability that a success occurs given that a collision preceded
p_{csc}	p_{sc} given that collision occurred previously
p_{dc}	probability that a double collision occurs conditioned on a collision occurring
p_{dpc}	probability of a data packet collision
p_f	probability of freezing
p_i	conditional collision probability given that the node is in stage i
p_{id}	probability that the station sees an idle slot
$P\{i\|j\}$	probability of the event i given the event j
p_{rs}	probability that T saw a success by an IP whose receiver is heard by T

p_s	probability of success
$p_{.s}^{A(B)}$	$p_{.s}$ when success of A(B) follows (CCW)
$p_{.s}^{A(B)}(i,j)$	$p_{.s}(i,j)$ when success of A(B) follows
$p_{sc}^{A(B)}$	probability of a collision given that success of A(B) preceded (CCW)
$p_{ss}^{BA(BB)}$	probability of two consecutive successes by B and then A(B) (CCW)
$p_{s.}(0,j)\|k$	$p_{s.}(0,j)$ given that the successful station A was in stage k in state s_1
$p_{s.}^{i}(0,j)\|k$	$p_{s.}(0,j)\|k$ conditioned on A being successful for the $(i+1)$st consecutive time
p_{sc}	probability that a collision occurs given that success preceded
p_{ssc}	p_{sc} given that success occurred previously
p_{ts}	probability that T saw a success by an IP whose transmitter is heard by T
$P_X(x)$	probability that the DRV X is equal to x
plr	packet loss ratio
$\breve{q}$	probability of data packet collision
Q	term $\frac{1}{(1-p)L+pC}$
$\mathbf{Q}$	substochastic matrix of the transitions among the transient states
$\breve{q}_{ik}$	probability of a data packet collision given that the backoff stage was i and the counter k
$q_k(0,j)$	probability that a station is successful for the $(k+1)$st consecutive time
q_{kn}	transition probability from transient state k to transient state n
R	the data transmission rate
$\mathbf{R}$	reward vector
r	transmission range (common for all network stations)
r_i	transmission range of station i
r_g	distance between two successive stations in the grid topology (in both dimensions)
r_k	average duration of state k in the detailed model of Chapter 5
r_s	distance between the receiver and any one of the transmitters in the star topology
RL	retry (retransmission) limit
S	saturation throughput
$\breve{s}_j$	number of IPs of type j around the pair T-R
s_n	state n of the channel evolution
S_{IP}	set of all the types of IPs
S_{IP}^*	$S_{IP}^* = S_{IP} \backslash \{g\}$
ss	number of successful slots in an obervation interval
t	time instant

t_n	time instant where the transition to state n occurs
T	class of absorbing states
$\breve{T}$	time-index space of a DTMC
T_{ack}	RTS + SIFS + CTS + SIFS + DATA + SIFS + ACK
T_c	duration of a collision period
T_{dpc}	duration of a data packet collision
T_f	duration of a freezing period, T_s − RTS − SIFS
T_n	event that the node transmits at the fixed-length slot n
$\bar{T}_n$	complementary of T_n
T_{nack}	RTS + SIFS + CTS + SIFS + DATA
T_{out}	duration of Timeout and DIFS normalised to the slot
T_s	duration of a successful transmission, T_{ack} + DIFS
T_v	duration of the vulnerable period
u_{max}	RTS + CTS + SIFS
u_{min}	SIFS
$\mathbf{v}$	steady-state probability vector of the embedded DTMC of Chapter 5
W	(constant) contention window
W_i	contention window at the backoff stage i
x	general index
$X\|cond(X)$	truncated DRV X after the condition $cond$ is applied to X
$X_{A(B)}$	DRV of the initial backoff counter value of the station A(B)
$X_i^{A(B)}$	DRV of the initial backoff counter of station A(B) given it is in stage i
X_A^*	truncated X_A equal to $X_A\|X_A > X_B + c - Y$
X_A^+	truncated X_A equal to $X_A\|X_A > X_B - K + c$
Y	DRV of the time difference of RTS packets given that collision occurs
Y_n	DRV of the time difference of RTS packets in s_n given that collision occurs
Y_n^c	Y_n conditioned on collision preceding in s_{n-1}
Y_n^s	Y_n conditioned on success preceding in s_{n-1}
Z	DRV of the value of the backoff counter after the latter is frozen
Z_n	DRV of the value of the backoff counter in s_n after it is frozen
Z_n^c	Z_n conditioned on collision preceding in s_{n-1}
$Z_n^{A(B),c}$	Z_n^c given that A(B) was successful previously
Z_n^s	Z_n conditioned on success preceding in s_{n-1}

$Z_n^{BA(BB),s}$	Z_n^s given that first B and then A(B) were successful
Z_{ij}	Z_n given that the stages of the senders are i, j
Z_{ij}^k	Z_{ij} given that it is the kth consecutive time that the same station is successful
$Z_{ij}^{A(B),k}$	Z_{ij}^k given that one refers to the remaining backoff slots of A(B)
(E, i, j)	system state described as a triple of the event E and the backoff stages i, j
$\bar{.}$	event complementary to .
$\|.\|$	absolute value of .
$RV_1 \oplus RV_2$	combination of the RVs 1 and 2

Chapter 1
Introduction

The performance evaluation of wireless networks has been at the centre of attention of the research community since the 1970's, when Packet Radio Networks were introduced. More recently, Wireless Local Area Networks (WLANs), Mobile (MultiHop) Ad Hoc Networks (MANETs) and also Wireless Mesh Networks (WMNs) have revived this trend due to their low cost of infrastructure, and their emergence as promising last-mile access technologies. Performance bounds of these networks under a certain channel access scheme are vital, in order not only to test the efficiency of the specific scheme, but also to provide a comparison measure for protocols designed for the upper layers (from routing layer to application layer).

1.1 Performance evaluation techniques

Performance evaluation can be done with three different approaches, namely by means of experiment, by means of simulation and, finally, by mathematical modelling. The first approach is the most accurate and realistic way to have a credible answer to the above problem. Multiple factors, from the physical to the application layer, can be taken into consideration, as well as their joint impact. However, scalability problems and considerable implementation cost are two very important impedance factors, that prevent researchers from adopting such methods, unless very small networks are to be tested. Simulation is a relatively straightforward way to model networks, using either commercial/open-source products, or building one independently. On the other hand, it has often been reported that simulators can not be used as a credible research tool, as a common platform is lacking, and different simulators produce different results [3]. More importantly, no real insights into the simulated phenomena can be obtained.

This thesis adopts the third approach, its aim being to develop analytical frameworks and models for the performance evaluation of wireless networks that employ a Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) [4] type of channel access protocol, such as the Distributed Coordination Function of the popular IEEE 802.11 [1]. Apart from its obvious aim,

which is to provide an answer to the fundamental question of how well can such a scheme perform in a wireless network, a mathematical framework is very useful, because it can provide deep insight into the channel contention phenomena, it can derive detailed Quality-of-Service metrics and, ultimately, it can give directions to the system designer on how to improve the algorithms and/or finetune the parameters of the protocol.

The chapter begins with stating the motivation for this thesis. Section 1.2 summarises the author's contributions. Finally, Section 1.3 provides an overview of the organisation of the thesis.

1.2 Problem statement and motivation

Nowadays, it is common knowledge that a network which only consists of stations transmitting at a minimum allowed power level, so as to ensure correct decoding at the receiver, would assist the spatial reuse in the network, a highly desirable target of our bandwidth-hungry communications. Trying to optimise the spatial reuse, an opposing problem arises, namely the interference caused by simultaneous transmissions due to the broadcast nature of the wireless medium. In fact, a common feature of all the aforementioned classes of networks is the high level of interference they exhibit. In fact, the most common Physical Layer and Medium Access Control (MAC) Protocol used both commercially and also in research institute testbeds still being IEEE 802.11 [5, 6], a significant question arises, that is whether 802.11 or any other random access based protocol is capable of performing efficiently in the presence of interference.

The interference can be of various types and, consequently, can have different effects on a wireless transmission. In an infrastructure WLAN there is a central Access Point (AP), the equivalent of a base station in a cellular network, and several stations around it, that are usually characterised by simpler hardware/software capabilities than the AP. The ideal topology for a WLAN would be a fully-connected topology, i.e. a configuration where all the stations are in transmission range of each other and of the AP and can communicate via a direct link. In this topology, the critical time duration—usually referred to as the vulnerable period—during which, stations may happen to transmit to the AP concurrently and, consequently, the AP is not able to decode either of them (this is referred to as packet collision throughout the thesis) is rather small. In fact, it is the time that it takes the signal to propagate from a station to any other

station of the WLAN[1] and for the physical layer to realise that there is an ongoing transmission, and, thus, notify the station to defer its own transmission. Consequently, the stations (including the AP) in a fully-connected network have almost (the exception is this small period of time described previously) the same view of the wireless channel. We describe these stations as synchronised[2].

The focus of the research community dealing with the performance analysis of CSMA/CA networks was, until fairly recently, on fully-connected networks, as will be explained in detail in the following chapter. However, it is quite often the case that, in a WLAN or in a MANET, the stations are not always in the transmission range of each other. This can be for example the result of a large physical distance, or of physical obstacles or, even, of heterogeneous power levels.

The most classic example of such a topology is the hidden terminal topology [7]. In the hidden terminal topology, two stations that are not in range of each other want to send data to a third station (the receiver). The problem is that, due to the absence of a direct link between the two senders, they can only infer each other's transmission indirectly, i.e. through the acknowledgement packets that are sent by the receiver as a confirmation of correct reception of data. Consequently, in this case, the critical period where packet collisions may occur can be quite long and this causes the stations' performance to degrade substantially. This degradation in their performance is not completely solved even with some recently proposed directional-antennae-based channel access schemes. In fact, recent advances in cross-layer design which have proposed directional-antennae-based MACs suffer from the directional hidden node problem [6, 8].

However, the hidden node problem is not the sole configuration where two simultaneous transmissions impede each other. In an arbitrary configuration, the interference caused is dependent on the form of the connectivity graph that connects the simultaneously transmitting stations. One of the primary aims of the author's research was to give insight into this complex phenomenon. In the research field of performance modelling of wireless networks, a systematic and comprehensive characterisation of all the possible interference topologies was an open

[1]The propagation delay in a WLAN, where the maximum distance, even in an outdoor enviroment, cannot exceed a few hundred meters, is in the order of ns; thus, it can be considered negligible, compared to the transmission delays, processing delays etc.

[2]In the literature, the term synchronisation refers to e.g. the clock synchronisation of a cluster of stations or computers, which is vital for some algorithms to operate satisfactorily, the sleeping schedule synchronisation when one refers to low-power wireless sensors, etc. In this thesis the definition is slightly different and is provided above.

problem. This systematic description is the initial problem that is discussed in this thesis, and this resulted in the proposal of a scalable modelling technique that computes the station throughput under certain assumptions, which will be enumerated in Chapter 3.

Following this, the realisation that this work, as well as several other research papers, make certain fundamental assumptions in order to keep the complexity of the solution relatively low, served as the motivation for the main contribution of this thesis. In fact, after a broad literature review, it was concluded that researchers, either focus on fully-connected networks, where the stations are synchronised, or they inherit techniques that suit a fully-connected network to the analysis of a network with an arbitrary connectivity graph. In particular, such common assumptions are the description of the channel evolution as a renewal process[3], the hypothesis that concurrent transmissions can be considered independently and/or the 'silent' assumption that the first statistical moment of the backoff distribution (i.e. the distribution of the waiting periods between successive transmissions) is always bigger than the vulnerable period. These assumptions limit the accuracy of their results only to certain configurations, e.g. absence or negligible duration of control packets, limited range of data packet lengths and/or selective range of transmission rates. In the Chapters 4 and 5 of this thesis, all the above assumptions are lifted and a novel and rigorous mathematical framework is developed, that manages to capture the complexity of the channel contention in the basic example of the classic hidden terminal topology. A summary of the contributions of this thesis is provided in the next section.

1.3 A summary of the contributions of the thesis

Several contributions regarding the modelling of interference in a wireless network that employs a random access MAC protocol are reported in this thesis:

- A systematic characterisation is performed of all the possible ways with which two pairs of stations employing a CSMA/CA access protocol can interfere with each other [9]. A modelling framework is proposed that derives the throughput of a network by decomposing it into these basic topologies.
- A comprehensive description and analysis of the assumptions that are usually adopted in

[3]A renewal process is a stochastic process, where the interevent times are independent and identically distributed (i.i.d).

the modelling of an arbitrary wireless network are made. The reasons behind the inconsistencies that previous techniques show are thoroughly explained and demonstrated via extensive examples [10].

- A rigorous and highly accurate mathematical model is introduced for the classic hidden terminal topology, that redefines the basic unit of model time and uses random variables for the quantification of the fundamental parameters that affect the channel contention [10, 11]. The model adopts a first-order (Markov) dependence for the description of the channel evolution, as opposed to the renewal process, which is usually assumed. Moreover, the independence and synchronisation assumptions are relaxed; on the contrary, the joint impact of the two senders' transmissions to the common receiver is considered, which enables the derivation of detailed metrics, such as a per-channel slot defined transmission probability, the collision probability and the throughput.

- An extension of the previous model to the Binary Exponential Backoff (BEB) mechanism proposed by the IEEE 802.11 is made. The computational complexity is kept relatively low, due to the adoption of a technique that calculates iteratively the unknown channel contention distributions [12].

1.4 Factors that influence the wireless network performance and have not been discussed in the analysis

The factors that influence the wireless transmission are numerous and much more challenging to deal with, as opposed to those affecting a wired transmission. Primarily, the wireless medium is far from perfect. The assumption that is usually made in the literature (not excluding this thesis) of channel access protocols evaluation is, that the transmission area is circular and consequently, any station/receiver located inside this area is able to decode perfectly the data sent to it by the station/transmitter, that lies in the centre (given that there is no other concurrent transmission). All stations that lie outside of the boundaries of the circular area do not receive anything. This deviates significantly from the reality. In fact, there are phenomena, such as shadowing, i.e. the distortion of a signal caused by large movements of a mobile station or/and the presence of obstacles in the propagation environment and multipath fading, i.e. the distortion caused by small movements of a mobile node when different reflections of the signal are destructively or constructively combined, making the signal fluctuate. In this thesis a static

network is assumed, therefore the above distortions are less important; nevertheless, they exist, but their quantification is outside the scope of this thesis, which considers a perfect wireless medium. However, as will be explained in Chapter 2, there is a fair amount of papers that deal with these factors, which can be integrated to the analysis in this thesis, since the event of the distortion of a signal due to both physical channel imperfections and due to a concurrent transmission can be considered independent and, consequently, the joint impact can be expressed in a product form. As a sidenote, the general impact of physical channel imperfections on the network performance is degrading; however, there can be cases, where the performance can be enhanced, since the concurrent transmission of a signal from an interferer, that would—under perfect channel conditions—destroy an ongoing transmission (and its own), is weakened by 'bad' channel conditions and the transmission can be resumed successfully.

The same applies to the integration of capture into the analysis. Capture is the effect, according to which, two or more concurrent transmissions do not necessarily destroy each other at the receiver, but, if the instantaneous power of one of them is larger than the sum of all the other interfering signal powers by more than a certain threshold, then the receiver captures, i.e. correctly decodes, the 'powerful' signal. This thesis does not take into account the capture effect and assumes that when two packets arrive concurrently at a station, they are both destroyed. A capture probability can, nevertheless, be integrated into the proposed analysis as described above. In fact, as was shown in [13], the throughput of a CSMA/CA type fully-connected network in the presence of various kinds of fading and taking capture into account, is much worse than that of the same network under perfect channel conditions in low load traffic; interestingly, the opposite occurs under high load conditions.

Moreover, throught the thesis, it is assumed that the transmission range is equal to the carrier sensing range. This is the range of a station, inside which, a signal generated by the station is not necessarily correctly decoded at another station, but, it is sensed by its antenna, causing it to defer its own transmission. Commercial wireless cards usually adjust their receiving and carrier sensing thresholds in such a way, that the carrier sensing range is larger than the transmission radius. Although this is done in order to alleviate collisions and the hidden node problem, it does not always lead to a higher throughput of the overall network, as many stations, which could, under equal ranges, resume their transmissions, are 'silenced'.

Finally, a fundamental hypothesis in the present thesis is the one of saturation conditions. In saturation conditions, it is assumed that a station (usually termed as backlogged station) has

always a packet to transmit, i.e. its buffer is always non-empty. The saturation throughput is very important in that, it describes the asymptotic behaviour of a network under conditions of the maximum sustainable load. However, because the delays incurred under these conditions are significantly high, quite often it is preferrable that a network operates under lighter traffic load. In order to model non-continuously backlogged stations, an additional distribution for the traffic arriving from the upper network layers is necessary. Therefore, a random variable that defines the packet interarrival times must be added to the variables describing the intertransmission times (i.e. the backoff distributions). This is possible in the models that are proposed in this thesis, however, with reference to Chapters 4 and 5, it may be that the complexity of the result may be substantial.

1.5 Organisation of the thesis

The thesis is organised as follows:

Chapter2

This chapter provides the general principles and background to CSMA access schemes, with particular emphasis on the p-persistent CSMA and on the IEEE 802.11. The most important research papers using various modelling tools and dealing with several aspects of the performance analysis of CSMA wireless networks are reviewed. Finally, the motivation and linkage to the main concepts of this thesis is outlined.

Chapter 3

This chapter provides a systematic and comprehensive description of all the possible interference topologies in a network with an arbitrary connectivity graph. Initially, these basic topologies are examined separately, and the respective collision probabilities are calculated under the assumption of only one interfering pair at any one time. Later in the chapter, this assumption is lifted and the analysis is extended to symmetric networks or very large randomly populated networks.

Chapter 4

In this chapter the reasons behind the limitations of previous methods are explained in depth, and the impact of their fundamental assumptions to their modelling accuracy is demonstrated.

Moreover, a novel method of modelling time in a hidden terminal topology is presented, and expressions for the transmission probability, the conditional collision probability and the saturation throughput are derived. On the other hand, the assumption that transmitters enter the contention phase synchronised is lifted, and the complexity of the channel contention is grasped successfully with the use of random variables. In this chapter, it is presumed that the network stations employ a constant contention window for their backoff distributions.

Chapter 5

In this chapter, the assumption of a constant contention window is relaxed, and the model is extended to the BEB scheme proposed by IEEE 802.11 (or any other backoff distribution). An iterative method is used to keep the state space relatively small. An exact solution is also provided; however, it is demonstrated that the iterative technique performs almost as well as the exact model, and is several orders of magnitude less complex.

Chapter 6

Finally, some concluding remarks and suggestions for future research directions are made, as well as some rough guidelines are provided, for the optimal configuration of IEEE 802.11 networks with hidden terminals.

Chapter 2
Background

2.1 From CSMA and Packet Radio Networks to IEEE 802.11 WLANs and multihop networks

The performance evaluation of multihop wireless networks has been at the centre of attention of the research community since the 1970's, when Packet Radio Networks were introduced. In the mid-1970's the research community had already explored fairly reliable and inexpensive wired communications, where circuits and transmission paths were established once for every upper-layer request for communication. Circuit-switching however was soon proved to be unable to handle computer-to-computer communications, where bursty traffic is frequently present and the ratio of peak to average traffic rate is considerably high. At that time, a new necessity seemed to arise, according to which communication between a number of users can be established without the presence of infrastructure and in a packet-switched fashion. This scenario presupposed the shared use of a common radio channel in a packet-switched mode among all the participating communication partners and the researchers defined this type of radio channel as 'packet radio channel' [4]. Consequently, such type of networks were called 'Packet Radio Networks' (PRNs) and are the ancestors of what we call today wireless multihop networks or wireless mesh networks, etc.

The above method of radio communications, however, directly raises the question as to how possible conflicts can be resolved when more than one packet is sent simultaneously by at least two participating stations through the common broadcast channel. The way to solve these conflicts is the object of the so called MAC Protocols that try to schedule the traffic in an optimal way in order to maximise the network throughput, and at the same time to not sacrifice the fairness between all identical users (note that sometimes unfairness can also be the goal, but this brings us to the issue of Quality-of-Service (QoS) and service differentiation).

In general, MAC protocols can be divided in two broad categories, i.e. reservation-based and random access-based [14]. The first class of protocols comprises multiple access schemes such

as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA) and Code Division Multiple Access (CDMA) where predetermined channel time slots, frequencies or code sequences respectively are assigned to each user, so that the transmissions are orthogonal and do not conflict with each other. This is a safe and straightforward method to resolve channel access and works very well in a high-contention environment. Given the fact that the delays that these protocols impose on packet transmissions are deterministic, they can, in general, support QoS. However, they are not suitable for bursty traffic and do not adapt quickly to topology changes.

The second category, i.e. the random access based protocols, consists of protocols where there is no such pre-allocation of the shared medium and stations compete for its occupancy in a 'real-time' manner. The first protocol of the kind was the so called 'pure-Aloha', where stations may transmit whenever they want, and through the reception or not of an acknowledgment packet, they realise whether their transmission attempt was successful or not. A similar algorithm, with the difference that time is slotted and stations may only initiate their transmission at the start of a slot is described by the subsequent 'slotted-Aloha'. Some years later, the seminal work of Kleinrock and Tobagi [4] introduced another protocol, called Carrier Sense Multiple Access (CSMA), according to which the stations attempt to avoid collisions by sensing, i.e. listening to the carrier of another station's transmission. Because of its significance in modern wireless communications—it is used, for example, in Wireless LANs, in Wireless PANs (IEEE 802.15.4) and, also, in cellular networks, during the setup phase—and of the fact that it is the main subject of this thesis, CSMA will be described in detail in Section 2.1.1.

Almost two decades after the introduction of CSMA, the IEEE has standardised the so-called 802.11 [1] protocol for WLANs, initially targeting networks with a base station and a number of wireless nodes within the transmission range of it and of each other (fully-connected networks). IEEE 802.11 is largely based on CSMA with the additional feature of Virtual Carrier Sensing, as will be explained in detailed in Section 2.1.2. Moreover, two control packets called Ready to Send (RTS) and Clear To Send (CTS) have been proposed by IEEE 802.11 to precede the data payload transmission, in order to combat the well known hidden node problem; the hidden node problem arises when two or more contending stations that are not in range of each other attempt to establish communication with a common receiver. In this case each transmitter can only infer an ongoing transmission between its competitor and the receiver indirectly, through the acknowledgment packet (or the CTS packet). The hidden node problem is a millstone in

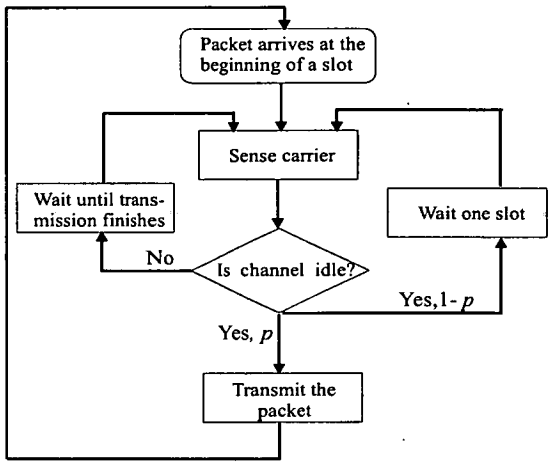


Figure 2.1: *A flow chart of the slotted p-persistent CSMA protocol.*

wireless communications and is one of the most serious and frequent interference topologies as was explained in the introduction.

2.1.1 Description of CSMA

CSMA, introduced in [4], and analysed in the presence of hidden terminals for the first time in [15], is governed by the general principle 'listen before talk'. In other words, each station tries to avoid collisions by sensing the shared channel (the carrier of other possible transmissions). Kleinrock and Tobagi proposed and analysed two large categories of CSMA protocols, namely nonpersistent and p-persistent.

In nonpersistent CSMA the idea is that every station senses the carrier before transmitting. If it finds the channel idle, it proceeds to transmit its packet. If not, it reschedules its transmission for a future time instant according to a certain distribution. At that time instant, the same procedure is repeated. Nonpersistent CSMA can also be synchronised, where the time is divided in slots and every station is forced to start transmitting only at the start of a slot. In p-persistent CSMA (see Figure 2.1) an additional parameter is introduced, namely the probability p that the node will 'persist' in transmitting at the start of the slot and, naturally $1-p$ the probability that it will postpone its transmission for the next slot. The idea is that according to the congestion of the network (e.g. the number of competing terminals) the parameter p can be suitably adjusted to minimise collisions. An optimal value of p should, on one hand, attempt to minimise interference between the competing stations, and, on the other hand, try to keep the idle periods as

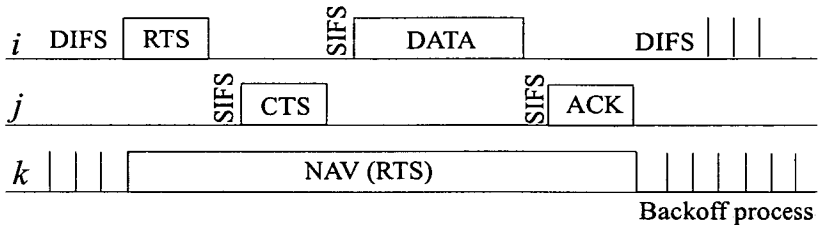


Figure 2.2: *The RTS/CTS scheme as defined in the IEEE 802.11 standard [1].*

short as possible [4]. For example, if p is very high, the station will almost certainly attempt transmissions and, thus, possibly suffer from collisions, especially if the network is densely populated. On the contrary, a very small value of p will avoid collisions, but

In fact, the popular IEEE 802.11 is inspired by the p-persistent CSMA protocol, with the Binary Exponential Backoff mechanism regulating an equivalent of the crucial parameter p according to the network congestion, and with the incorporation of some additional features, such as the use of control packets and the Virtual Carrier Sensing mechanism.

2.1.2 Description of IEEE 802.11

In this section, an outline of the most important features of the RTS/CTS and Basic Access mechanisms of IEEE 802.11 Distributed Coordination Function (DCF) is given. For a more thorough description, the reader should refer to [1]. DCF is based on CSMA. When a node wants to transmit a DATA packet, it first senses the channel. If it finds the channel idle for more than the time interval DIFS, then it sends an RTS (Ready To Send) packet, if RTS/CTS is adopted, or the DATA packet itself, if Basic Access is used. The receiver, upon error-free reception of RTS and, if, after the expiration of a Short Interframe Space (SIFS) [1] time interval, it senses the channel idle, returns another control packet, called CTS (Clear to Send) to the sender. Similarly, for Basic Access the receiver replies with an ACK (acknowledgement) if the DATA packet was received correctly. This concludes the handshake for Basic Access. For RTS/CTS, after RTS and CTS packets have been exchanged successfully, the transmitter sends the DATA packet to the receiver, who, in turn, responds with an ACK if the reception of DATA was correct. If, for any reason, the sender has not received a CTS (ACK) as a response to it sending an RTS (DATA) packet within a certain time duration (Timeout), then it reattempts transmission. In addition, RTS and CTS packets have a special field which indicates the length of the subsequent data packet. This is useful for the so called Virtual Carrier Sensing

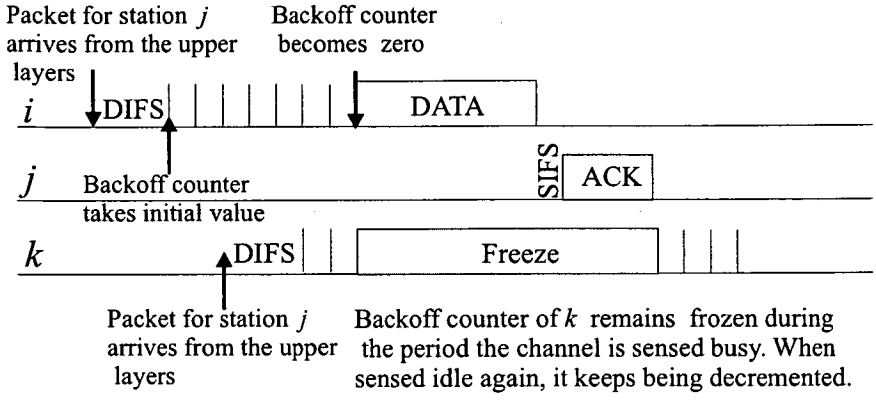


Figure 2.3: *The backoff process of IEEE 802.11 stations that use the Basic Access scheme. In particular, stations i and k want to transmit data packets to station j. For this figure it is assumed that all three stations are in range of each other.*

mechanism, where all stations that overhear the packets know the exact duration of the ongoing handshake and must defer their own transmission until the former is completed. The process is shown in Figure 2.2.

For both modes, after every successful or failed transmission, the station must wait for a time interval, called the backoff interval, before sending the next packet (or the same packet in the case where the previous transmission failed). This is done so that a single station does not monopolise the channel for a long period of time. The backoff mechanism is employed with the aid of a counter (the backoff counter) and is briefly described in Figure 2.3. After every transmission the counter takes a new value and it is decremented by one at every channel slot σ if the node senses the channel idle. On the contrary, if during its backoff, the station senses a busy channel, then the counter freezes at its current value until the channel is considered idle again (according to both the Physical Carrier Sense and the Virtual Carrier Sense mechanisms[1]. The station transmits when the value of the backoff counter has reached zero. The initial value of the counter can be selected according to a variety of schemes, but in [1] the Binary Exponential Backoff (BEB) scheme is proposed. According to BEB, the backoff counter follows a uniform distribution with its maximum possible value (named the 'contention window') being doubled after every failed transmission or being reset to an initial value W_0 after a success-

[1]The Virtual Carrier Sensing mechanism is employed via the Network Allocation Vector (NAV), which is an indicator, that determines whether there is another ongoing transmission in the channel; therefore the station cannot initiate a transmission during that time and is forced to defer. The NAV is adjusted when the station overhears a (control or data) packet, that is not destined to itself, which has a field that contains the duration of the remainder of the handshake.

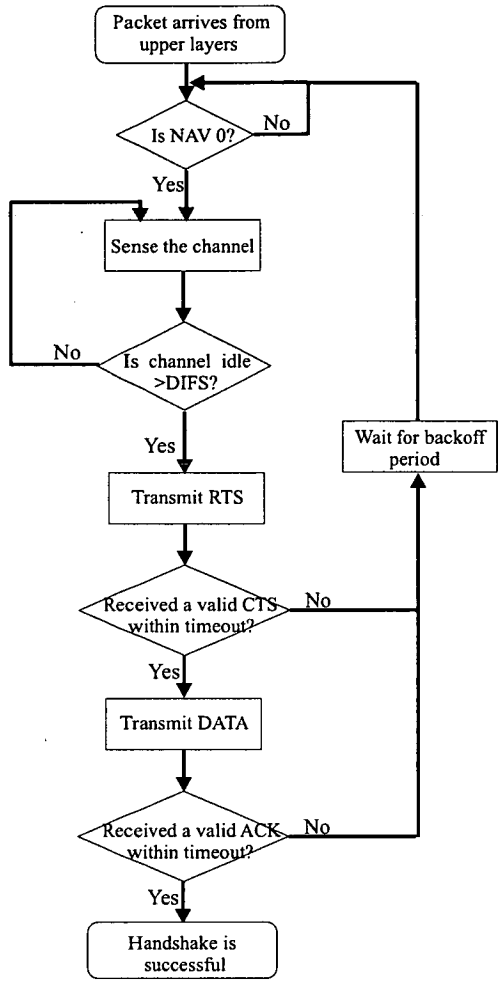


Figure 2.4: *The flow chart describes the basic operation of the RTS/CTS mechanism of IEEE 802.11 DCF.*

ful transmission. This process is repeated until the maximum retransmission limit is reached, after which the next maximum value of the counter is chosen in the interval $[0, W_0 - 1]$ independently of whether there was collision or success previously. However, the protocol allows for each station to individually and dynamically configure its contention window according to the channel conditions. Figure 2.4 presents a flow chart that summarises the operation of the RTS/CTS scheme of the DCF; the Basic Access scheme follows a very similar algorithm.

Finally, as a sidenote, as far as the physical layer part of the IEEE 802.11 protocol is concerned, the protocol determines that the signals shall be transmitted at frequencies in the Industrial Scientific and Medical (ISM) frequency band (specifically at 2.4 and 5 GHz), which is

the unlicensed band. Moreover, it suggests two possible spread-spectrum techniques, namely the Frequency Hopping Spread Spectrum (FHSS) technique and the Direct Sequence Spread Spectrum (DSSS) technique, in order to alleviate interference from non-IEEE 802.11 users. In FHSS, the carrier is rapidly switching among many frequency channels, using a pseudo-random sequence known to the IEEE 802.11 terminals. In DSSS, the data is multiplied by a pseudo-random sequence of +1 and -1 values, at a frequency much higher than that of the original signal. Throughout this thesis, the DSSS is assumed, both for the simulation runs, and also for the theoretical analysis. FHSS would be dealt with in exactly the same manner, since—from the channel access algorithms's point of view—the only difference in the two spread spectrum techniques is the numerical values of the several protocol parameters [1].

Although the terminology and some protocol details have changed, it is obvious that the problems concerning the performance analysis and enhancement of random-access based protocols, such as the original CSMA or the popular IEEE 802.11 are still present, and that there is continuity in the corresponding research areas in the last three decades, even if nowadays we speak of wireless mesh and ad hoc networks in the place of PRNs. In the sections that follow, an extensive literature review of these decades will be attempted, explaining the significant contributions of researchers that have dealt mainly with the performance analysis of the previously defined protocols in various configurations. Note that the list of references is by no means exhaustive but, to the author's knowledge, all the major contributions are included. Also, it has been attempted to categorise the papers according to a certain pattern which is neither unique nor disjoint, i.e. a paper can actually belong to two or more classes of research papers.

2.2 Early papers for the performance analysis of PRNs

Since the 1970's the interest of the research community was driven by the analysis of the throughput of multihop networks in the form of PRNs. The CSMA protocol, introduced in [4], was first analyzed theoretically in the presence of hidden nodes in [15]. Later, in 1984, Takagi *et al* [16] examined the effect of hidden nodes in a homogeneous multihop network for the slotted nonpersistent CSMA protocol without the presence of acknowledgement packets, presuming that all stations are ready to transmit at all times. As in [15], they make the assumption that the events at each station, whether an actual transmission occurs or not as a result of channel sensing throughout a sequence of slots, is a set of independent Bernoulli trials. The same assumption is used for simplicity reasons for the analysis of the performance of the nonpersis-

tent CSMA in large multihop networks and its comparison with BTMA (Busy Tone Multiple Access) by [17].

These papers make the fundamental assumption that the acknowledgment packets (ACKs) have zero length so that they can be received instantaneously, making their analysis for an IEEE 802.11 network (operating in the Basic Access scheme, let alone in the RTS/CTS scheme) inappropriate. Similar assumptions have also been made in [18], where the authors presume zero length control packets. Also, [19] extends the analysis of [17] to adjust to the IEEE 802.11 protocol, initially making the assumption that both carrier sensing and collision avoidance work perfectly, i.e. that nodes can accurately sense the channel, and that the RTS/CTS scheme can avoid the corruption of the subsequent data packets at the receivers. Later in their paper, they attempt to overcome this assumption by considering an extra state in their Markovian model, which models the collisions of the data packets incorporating an intuitive 'imperfectness factor', as they define it. Furthermore, [20] proposed a linearised model for the derivation of the transmission probability, an assumption which does not take into account the desynchronisation of the nodes that lie outside of the transmission range of each other.

In parallel, in [21], one of the early papers analysing the performance of the PRNs, the authors show that, under certain assumptions, a Markov model of a CSMA based PRN can be built which leads to product form state probabilities. The assumptions they make are:
i) zero propagation delay
ii) perfect capture, i.e. once a successful transmission is initiated, the reception is immune from any subsequent interference (this actually implies that collisions can take place in a rather small time period) and
iii) the node's transmissions to its neighbours are scheduled according to a Poisson point process. This implies that packets that failed to be successfully transmitted or were inhibited from transmission are rescheduled after a sufficiently long random time interval so that the Poisson property is preserved. In a subsequent paper [22], complementary to the former, the authors propose an algorithm that computes throughput based on their prior model. Finally, Tobagi and Brazio [23] extended the previous work to other channel access protocols, such as Aloha and a variant of BTMA and analysed the conditions under which a product form solution can be derived in a PRN.

2.3 Papers following the standardisation of IEEE 802.11

The standardised IEEE 802.11 has been analysed over the years via different paths, i.e. experimentally, via simulation tools or via analytical or semi-analytical methods. In the first category a representative example is [7], where the authors make a spherical performance analysis of the behaviour of 802.11b stations for different basic topologies and vary the value of different parameters, such as the transmission rates and the interference and carrier sensing ranges, by means of experiment.

The RTS/CTS scheme that IEEE 802.11 proposes to combat the hidden node problem is not perfect. There are a number of papers [24–27] that report cases when the RTS/CTS handshake fails to prevent the DATA packet corruption because of a collision. Reasons that are shown to be the cause of this problem are different transmission and carrier-sensing areas [24, 27], non-negligible propagation delay [25], node mobility [26] and different transmission rates [27] for control packets (RTS/CTS) and DATA packets. In a similar context, [28] explains the ineffectiveness of using RTS/CTS control packets when the transmission rate for the data packets is high and suggests that the research and standardisation community reconsider the 4-way handshake in a time-duration context rather than in a frame size context. However, as the recent paper [29] shows, there are cases when the RTS/CTS mechanism fails even under perfect channel conditions and zero mobility. [29] defines nodes that receive an RTS or a CTS packet, but cannot interpret it correctly because of another on-going transmission as *masked nodes*, and uses a queueing-theoretic approach to evaluate their effect in the performance of such a network. Similarly, in a work parallel to this thesis [9], the authors of [30] identify some of the problematic topologies where RTS/CTS performs worse than CSMA and define them as the *gagged stations*. They reach the conclusion that even if one neglects the additional overhead present in the RTS/CTS scheme, in several cases it inhibits concurrent transmissions which do not interfere with each other in reality.

2.3.1 The seminal work of Bianchi and the class of papers that directly extend it

Since the standardisation of IEEE 802.11 [1], there has been much research effort put into the theoretical analysis of fully-connected 802.11 wireless LANs, with the most prominent paper being [2]. In [2] a bidimensional Markov chain like the one of Figure 2.5 is used for

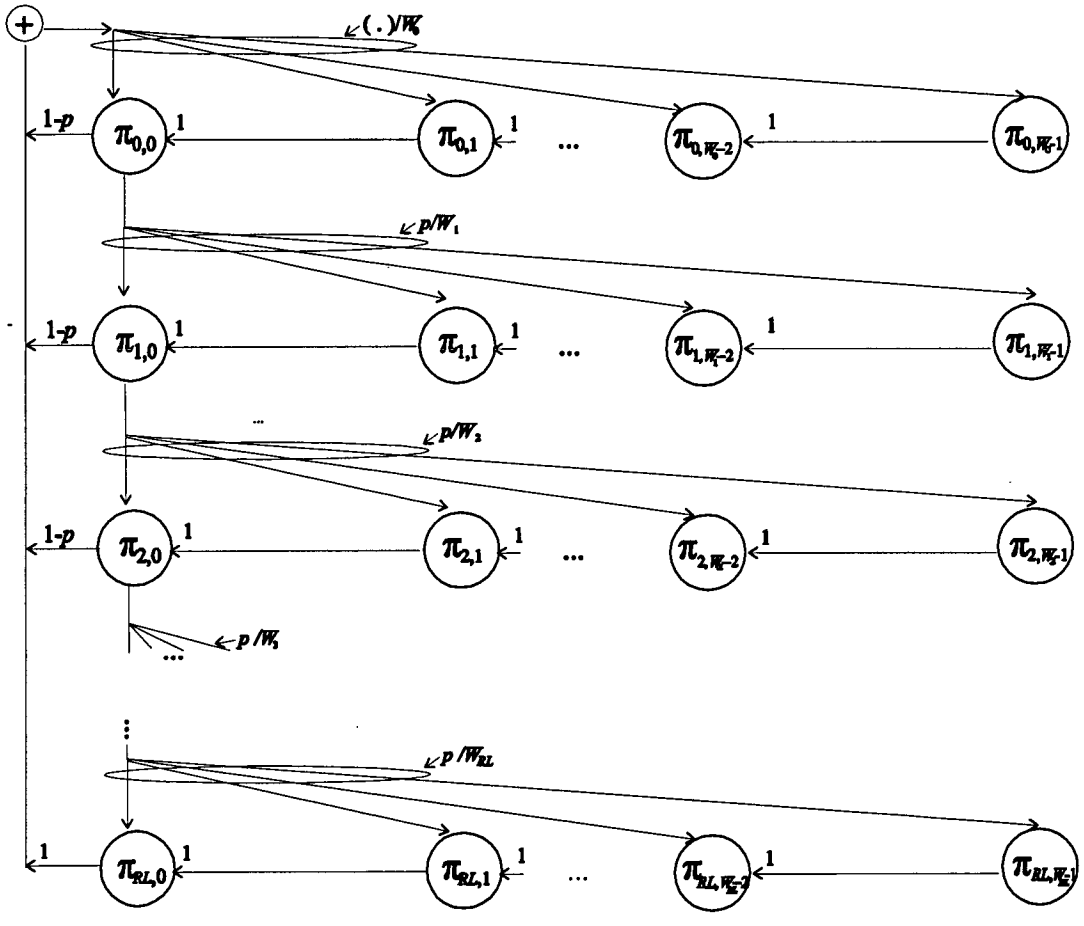


Figure 2.5: *The bidimensional Markov chain that is proposed by Bianchi to represent the back-off process of an IEEE 802.11 station.*

the description of the backoff process. In this Markov chain, the state[2] is represented by the pair (i, k), where i represents the backoff stage and k the value of the backoff counter. In every step of the process, the counter is decremented with probability 1 and when the counter becomes zero, then the station transmits. Following this, if there is a collision (this occurs with probability p), the process goes into the next backoff stage. If there is a successful transmission (probability $1 - p$), the next value of the backoff counter is chosen from the initial contention window W_0. Finally, if the packet has already been retransmitted RL times, the backoff counter is also reset to W_0, independently of whether there was collision or success. Note the uniform distribution of the backoff counter at each stage and, also, the fundamental assumption that Bianchi makes, i.e. the assumption that, regardless of the backoff stage, the station faces a collision with a constant probability p. This chain is ultimately used for the derivation of the transmission probability of a single station and, subsequently, the network throughput is derived using renewal theory arguments, as will be explained in greater detail in Chapter 3.

Moreover, [31] and [32] calculate the average packet delay for a fully-connected optical WLAN that employs the Infrared spectrum as a physical layer option, and are based on the assumption of a constant collision probability at any slot, their analysis relying on [2] and [33] respectively for their chain of arguments and problem formulation. Furthermore, [34] uses the standard framework by Bianchi for throughput analysis, but extends it to an error-prone channel and investigates the effect of packet size, the number of contending stations, transmission collision probability and channel condition of a saturated fully-connected network. Finally, there are a couple of papers that extend Bianchi's model so that they incorporate freezing the backoff counter of a station when it senses another node's transmission. In particular, in [35], the authors assign a probability, p_b, to the event that the node will freeze its backoff counter when it senses the medium busy and while being in backoff, and claim to obtain better results than [2]. More recently, [36] proposes a corrected version of [35], following the observation from the authors that the channel access probability actually depends on whether the previous period is idle or busy. Both previous papers assume saturation conditions and do not consider hidden or exposed nodes in the network.

However, as will be shown in Chapter 4, a simple extension of [2] which is followed by many

[2]The formalism followed in the literature is—usually—the graphical representation of a Markov chain by its state descriptors. In this thesis, another formalism is adopted, so that the representations of the various Markov chain models that are presented do not conflict with each other. Therefore, the states are described in the figures with their steady-state probability notation.

researchers in the field [37–40] is not appropriate for a network exhibiting hidden terminals. In particular, in [37] the authors define the 'deferral set' of neighbours that interfere with the T_x-R_x transmission (which are actually the one-hop plus the two-hop neighbours that want to communicate with the one-hop neighbours), and define the possible interferers ('equivalent competitors'). Their main simplification is that the equivalent competitors should defer transmission for a certain time interval Δt when the T_x-R_x transmission is taking place. In the analysis presented in Chapter 3 it is clear that this is not entirely true, as the effect of the transmission of an interfering pair on an active transmission can vary, and is dependent upon a) the relative position of the conflicting pairs and b) the relative time difference of their transmissions.

There are a significant number of papers that provide analytical methods for the performance of 802.11 in a multihop network, however, they assume the existence of renewal points in the analysis, i.e. that nodes enter the contention phase synchronised. Following a similar concept, [38] and [40] make a careful analysis of some interference topologies in a multihop network; however, both are based on renewal theoretic arguments for the calculation of the system throughput and use a generalised time slot as the basic unit of their models, similar to the slot introduced in [2]. As will be shown in detail in Chapter 4, all the former approaches exhibit some fundamental limitations in their accuracy. An exception to this rule is [41], that considers a scenario with three transmitting pairs, but the analysis is not computationally efficient, since it involves solving a Markovian model with several thousands of states.

In 2005, [42] proposed a queueing theoretic approach with the fundamental assumption that the time between transmissions (the backoff) is negligible, a fact that makes the analysis accurate only for long data packets combined with small transmission rates. Finally, very recently [39] has modelled the media access for several scenarios of interference in a 802.11 network. However, in order to model the classic hidden terminal problem using renewal theory, the authors consider that the nodes are synchronised when entering the contention phase. In their following paper [43] Garetto *et al.* extend the former analysis to a general multihop network, classifying traffic flows in suitable classes in order to study the starvation and unfairness phenomena that are caused by CSMA. Chapter 4 will elaborate on [38] and [39] due to their closer proximity with the case study presented in that chapter, and will also explain their main assumptions, and analyse the impact of the latter on the correct analysis of the hidden terminal effect.

2.3.2 Papers that calculate the throughput of a CSMA-based protocol but have adopted different modelling approaches

The seminal work of Bianchi has recently inspired several researchers to reproduce and verify his results under more general conditions and via different modelling methods. In particular, Kumar *et al.* in [44] introduce a fixed point analytical method for the throughput analysis of a fully-connected IEEE 802.11 network consisting of n equivalent nodes. The authors make the assumption that the aggregate transmission process of the $n-1$ stations is independent of the backoff process of a certain tagged node. They call this assumption 'the decoupling approximation' and show that it is valid for densely populated networks. Finally, the authors establish uniqueness of the solution to the fixed point equation and give expressions for the collision probability and aggregate throughput. Aiming at a similar goal, more recently the authors of [45] provide a rigorous analysis for the same problem, showing that for large networks the stochastic evolution of the backoff stages at different stations converges to a deterministic evolution and, thus, providing a theoretical justification for the decoupling assumption in [2]. Nevertheless, the authors leave the issue of hidden terminals and, in general, of interferers out of the range of the transmitters for future work. Very recently, in [46], Medepalli and Tobagi propose a unified model that attempts to take into account in the same model hidden terminals, multiple channels and/or directional transmissions and finite load traffic conditions. The authors assume independence of the events leading to a failed node transmission and assume a constant collision probability, independent of the backoff stage of the node. They ultimately calculate the average service time by numerically solving a complex system of nonlinear equations.

In [47, 48] the authors were the first that accurately dealt with the performance analysis of the newly standardised DCF function of the IEEE 802.11. Their seminal work attempts to define all the necessary and sufficient conditions for a successful data frame transmission for both access modes of 802.11, namely Basic Access and the RTS/CTS scheme. To this end, they define suitable regions of interference for each transmission and ultimately derive a closed form expression for the system throughput under general traffic conditions. In order to do so, they rely on the assumption that once the channel is sensed idle and the DIFS interval has elapsed, the time until a data frame arrives from the upper layers at station i with station j as the destination, follows an exponential distribution. Moreover, although the authors are one of the few that take the hidden terminal effect into account, they explicitly state that they do not model it accurately for the sake of tractability: "Unlike wired multiple access systems the completion

of a successful transmission or a collision is not a renewal point due to the presence of hidden stations. Obtaining statistics of the renewal interval in a WLAN environment in the presence of hidden stations is intractable; we therefore assume that the instance of the completion of a successful transmission and/or a collision is a renewal point. We expect the assumption to be satisfactory for cases where the number of hidden stations is small".

In [49], the authors analyse the performance of a p-persistent backoff algorithm and show that IEEE 802.11 deviates from the theoretical performance limit quite often. Following the former observations, they propose that the protocol uses information from the network status in order to dynamically tune the backoff contention window, and, thus, maximise the network throughput. Throughout their analysis they use the assumption of a geometric distribution for the backoff counter which enables them to consider independence from the stations' history. Note also that their analysis is only applicable to a fully-connected network. Moreover, in [50] and [51] the authors derive closed form expressions for the throughput and the delay respectively of a p-persistent CSMA multihop network. To this end, they use the assumption that the system state consists of a sequence of regeneration cycles, which is accurate, since the hidden terminal problem is neglected by supposing that all stations can sense the channel around the unique sink, and defer their transmissions when necessary.

In [52], the authors revisit the problem of analysing the performance of the Exponential Backoff scheme in a general context, using a one-dimensional Markov chain rather than the bidimensional Markov chain of [2], and, thus, being able to derive the medium access delay in addition to the throughput. Their analysis deals with infinite retransmissions as well as with a finite retry limit. However, the authors presume continuously backlogged stations, implicit transmission of acknowledgements and a fully-connected network, where all stations are synchronised.

A different class of papers that attempt to analyse the performance of CSMA/CA protocols in WLANs can be characterised by the different mathematical tools they use in order to achieve their purpose. One of the most important is [53], where the authors use Stochastic Petri Nets[3] for the derivation of the network throughput. This approach has some advantages in that many details of the protocol can be incorporated in the model and can give deeper insight to the

[3] A Petri Net is a modelling language that can describe a complex dynamic system. The state of the system is described by places, which may contain different number of tokens. According to certain rules, transitions may move the tokens and, thus, change the state of the system. Arcs are used to connect the places with the transitions or vice-versa. A Stochastic Petri Net is a special case of a Petri Net, where firing delays, as random distribution functions, are associated with every transition [53].

several phenomena that characterise channel contention. However, the current simulation tools that analyse and solve Petri Net models are unable to model all these details in practice, and the authors are forced to make several simplifications that, on one hand, have an impact on the model accuracy, and on the other hand make it extremely difficult for the model to be extended to multihop networks where hidden terminals are present.

With respect to modelling spatial reuse in a wireless multihop network, [54] introduces a purely probabilistic model for the derivation of throughput in a given wireless multihop network where the maximum number of traffic flows can be transmitted concurrently. Although the analysis is interesting, it deals with a different problem than the one investigated in this thesis, while it considers an arbitrary random access based protocol, it neglects the existence of backoff periods and control packets such as acknowledgements; thus, it cannot be applied to IEEE 802.11, for example.

Moreover, in [55] Sun *et al.* develop an analytical model called MBRP (Model-Based Resource Prediction) for resource (such as throughput and delay) estimation in a wireless network where real-time and multimedia traffic may be wanted. The model presumes Poisson distributed channel access attempts by several stations and also a constant collision rate, and manages to provide accurate results for ideal channel conditions. However, in order to deal with the effect of hidden terminals and non-ideal channels, the authors come up with an enhanced version of their model which uses the difference between an experimentally measured value and the model output as run-time feedback. The improved analysis is called Enhanced MBRP (EMBRP).

Finally, recently, [56] has presented an analytical model for the hidden terminal scenario, focusing on the short-term unfairness that the RTS/CTS scheme of IEEE 802.11 exhibits. Their analysis is more accurate than most papers, however, they, too, assume synchronisation after both successful and collision periods and they also make several simplifying assumptions concerning the backoff distributions which directly affect their results.

2.3.3 Papers that derive expressions for the service delay of a packet

In [57], the authors derive closed form expressions for the first and second order moments of the service time experienced by a packet that is generated in a fully-connected WLAN. Their model is based on the bidimensional Markov chain modelling of [2] for the transmission probability with an infinite limit of retransmissions. [58] calculates the backoff delay under

saturation conditions in a fully-connected WLAN for the recently standardised IEEE 802.11e Enhanced Distributed Channel Access (EDCA) function. The analysis is based on [2] and [59]'s expression for the transmission probability, although they have to alter the analysis in order to incorporate the prioritised channel access that EDCA determines for the various classes of nodes. Finally, the authors of [60] provide a more detailed analysis of the service time in an 802.11 WLAN by deriving the Probability Generating Function of the service time, and by not relying on the standard second order analysis often used in the literature. Nevertheless, their analysis is only applicable to a fully-connected network with no hidden or exposed nodes and assumes an infinite number of retransmissions and saturation conditions.

2.3.4 Papers that consider more complex physical layer models

Throughout the recent years there has been a set of papers that consider physical layer effects on the MAC protocol and depart from the classic two-ray ground model of circular transmission areas: In [13], the authors investigate the capture probability of an AP of a WLAN in a channel with Rayleigh fading, shadowing and near-far effects, and derive expressions for the throughput and delay of a CSMA/CA based protocol. [61] discusses the capture effect in a fully-connected IEEE 802.11 or 802.11e network as well; however, it approaches the problem via a different method, i.e. via multidimensional fixed point analysis. The authors presume in this case, too, the classic configuration of n backlogged contending stations in transmission range to each other and neglect the effects of hidden nodes. Finally, in the recent paper [62], Chang *et al.* present an iterative method in order to find the throughput of networks, where simultaneous transmissions do not necessarily result in collisions due to the capture phenomenon. However, in order to simplify their analysis, they substitute the classic equation for the transmission probability found in [2] with a linearised one, applying the least squares method to find the suitable coefficients.

2.3.5 Papers that consider finite load traffic conditions

The class of papers that do not rely on saturation conditions but consider finite load comprise the following: Barowski *et al.* [63] make the remark that for multi-hop communications the forwarding stations will not operate in saturation conditions and, thus, conclude that a finite load model is necessary. To this end, they add an extra state to Bianchi's bidimensional Markov chain, which represents the state where the station has a zero backoff counter, but does not

have a packet to transmit. Furthermore, they extend this model to three dimensions adding the queue length of the forwarding stations in order to apply to multi-hop networks. This analysis however neglects the hidden node effect that is crucial in multi-hop communications. Moreover, in [64] the authors propose a G/G/1 queueing model[4] for the analysis of 802.11 and 802.11e in a fully-connected network with finite load conditions. Finally, Garetto *et al.* [39] devote a section of their paper to the non-saturation conditions, which they incorporate in their model by defining an additional probability, i.e. the conditional probability that the transmission queue of the station is empty, given that the station can potentially start a new transmission.

2.4 An introduction to the novel features of the thesis in comparison to the literature review presented

In the first part of the thesis (Chapter 3) the author's research related to interference topologies will be presented. The analysis presented in Chapter 3 was the first to analyse and categorise all the possible interference topologies in a wireless network operating under a CSMA protocol, with particular focus on IEEE 802.11, in a detailed and comprehensive manner. After the basic interfering pairs are explained, a scalable way to calculate the throughput in a wireless network where communication is limited only to immediate neighbours (for simplicity reasons and to isolate the effect of a routing protocol) will be presented and then, the results will be extended to more complex, symmetric and randomly deployed network configurations.

Although the work presented in Chapter 3 is scalable and provides quite often trustworthy results, in Chapters 4 and 5 the classic hidden terminal problem will be investigated in more detail. It will also be explained that all prior researchers dealing with the theoretical analysis of the hidden terminal effect have used certain assumptions in order to make the modelling tractable. It will be shown that these assumptions limit the accuracy of their results only to certain configurations, e.g. absence or negligible duration of control packets, limited range of data packet sizes and/or selective range of transmission rates. To the best of the author's knowledge, the work presented in Chapter 4 is the first to present a comprehensive description and analysis of the aforementioned hypotheses which previous papers have frequently used. In particular, the reasons behind the difficulties with the assumptions previous techniques make

[4] A G/G/1 model is a queueing model, where there is one server and both the interarrival and the service times follow a general distribution.

will be explained, and some cases will be demonstrated, where this inaccuracy affects their results with extensive simulations. It will be shown that this occurs for a significant range of system parameters such as contention window size, transmission rate and data packet size.

Moreover, it will be pointed out that the basic concept that causes these limitations is the variable time-slot. This is the motivation for the second contribution, which is a novel modelling method that is based on a fixed-length unit of time. By introducing a new time scale for the modelling of channel contention, the classic metrics used in such problems, such as transmission probability, conditional collision probability and throughput, will be redefined, and relevant expressions will be derived. Moreover, standard assumptions, such as the one that transmitters enter the contention phase synchronised, will be relaxed. This assumption enabled researchers to use renewal theory for the throughput computation and, although it is valid in a fully-connected network, it introduces inaccuracies for a hidden node environment.. In this thesis a first-order Markov dependence is used instead of renewal theory for the modelling of two consecutive channel transmission states. Finally, by using random variables for the quantification of the terms that affect and determine the channel contention, in-depth insight will be provided into the complexity of the scenario under consideration. Chapter 4 will analyse the performance of IEEE 802.11 (in RTS/CTS and Basic Access mode) under the assumption of all the stations employing a constant contention window and Chapter 5 will extend this modelling method, and the relevant conclusions drawn, to the BEB scheme of [1].

Chapter 3
Interference Topologies Characterisation and Analysis—An Initial Approach

3.1 Interference in a wireless network

Unlike wired networks, the wireless medium is quite restrictive as a means of data/signal transfer. Depending on the MAC protocol that is chosen for a given application or network, there is, each time, only a limited number of concurrent transmissions that can occur around a station if orthogonal transmissions are to be achieved (e.g. in CDMA) or even only one, if a random access-based protocol such as CSMA is used. The goal of this chapter is to provide an answer to the following question:

Given a pair of wireless stations, between which there is data flow, and that have half-duplex transceivers (i.e. they can only transmit or receive at any one time), use a single channel and employ the IEEE 802.11 protocol (obviously any CSMA protocol would behave similarly), what are all the possible ways that their communication can suffer interference caused by any other station in their proximity?

People have been referring to the hidden terminal problem without distinguishing the different ways with which stations can interfere to each other's transmissions. Although there is a number of papers in the literature that define such interfering situations and often analyse their effect [15, 17, 29, 30, 37, 38, 47, 65], the author of the present thesis was the first [9] to enumerate in a detailed and comprehensive manner all the possible interfering topologies that exist in a random access wireless network. Subsequent independent work [39] reproduced analogous results providing further insight to the problems arising.

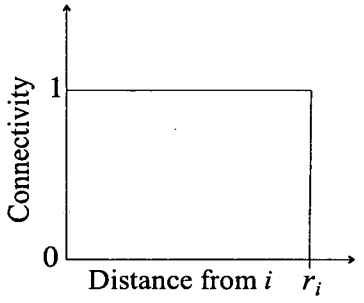


Figure 3.1: *A step-function connectivity is assumed, i.e. if the distance between station i and another station is less than a threshold, which is referred to as the transmission range r_i of i, any signal is decoded perfectly at the other station (given that there are no other simultaneous transmissions). On the contrary, if their distance exceeds r_i, then no signal can be received.*

3.2 The Model

Let us consider two wireless stations operating in the same (single) channel, employing half-duplex transceivers and let us assume that one of the stations receives from the upper layers a flow of data packets that it wants to send to the other station. From hereon the former will be referred to as 'the transmitter', denoted as T, and the latter as 'the receiver', denoted as R. It will be further assumed that the connectivity between a station i and any other station in its proximity is a step-function of the distance between the two stations as depicted in Figure 3.1. This is only an idealistic representation of the truth; in practical wireless communications there are no 'perfect circles', but the extension of the analytical modelling to more complex physical layers or to include capture, fading, shadowing, etc., has been sporadically examined in the literature [13, 61, 66] and is beyond the scope of this thesis, as was mentioned in the introduction. The present goal is to find the upper bound of a random access based protocol provided that the wireless medium is flawless.

3.2.1 Definition of the interfering pairs

Given the transmission range of the transmitter r_T, let's define the set of stations that are within the transmission range of it as the one-hop neighbours of T. They are defined as follows: $N_T = \{i : d(i,T) \leq r_T\}$, where $d(i,j)$ denotes the physical distance between nodes i and j. Equivalently, for the one-hop neighbours of the receiver R: $N_R = \{i : d(i,R) \leq r_R\}$. The two-hop neighbourhood of the two stations is defined as follows: $N_T^2 = \{i : r_T < d(i,T) \leq 2r_T\}$

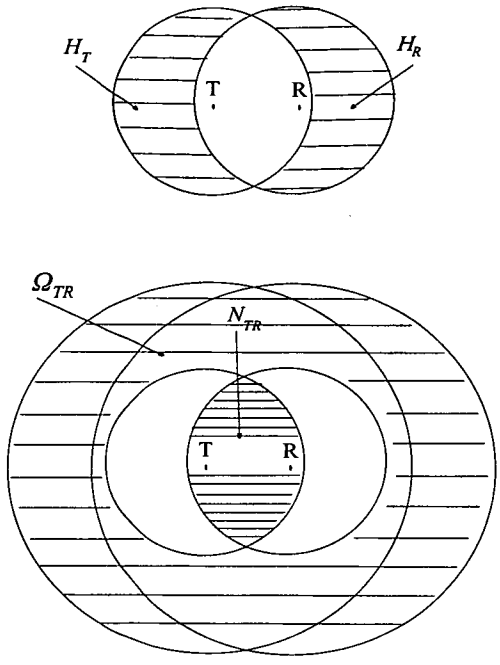


Figure 3.2: *Graphical representation of the sets of nodes that can be potential interferers to the communication pair T-R.*

and $N_R^2 = \{i : r_R < d(i,R) \leq 2r_R\}$. It is observed that the critical regions-sets of stations that can cause interference to the communication pair T-R are the following:

1)$N_{TR} = N_T \bigcap N_R$, i.e. the stations that are in the transmission range of T and the reception range of R, which are the stations of a fully-connected network that were investigated thoroughly in [2].

2) $H_R = N_R \bigcap \bar{N_T}$, i.e. the stations that are one-hop neighbours of the receiver R but are hidden from the transmitter T.

3) $H_T = N_T \bigcap \bar{N_R}$, i.e. the stations that are one-hop neighbours of T but are hidden from R.

4) $\Omega_{TR} = \left(N_T^2 \bigcup N_R^2\right) \bigcap \bar{N_T} \bigcap \bar{N_R}$, which is the set that comprises all the nodes that belong to the two-hop neighbourhood of either T or R but are not one-hop neighbours of either of them.

The regions that correspond to the above sets of interfering stations are displayed in Figure 3.2.

Note that in this chapter, the terms H_T, H_R, N_{TR} and Ω_{TR} will be used to refer to both the sets defined above or to arbitrary individual nodes belonging to the respective sets interchangeably. When the distinction is not clear from the context an explicit reference to the set or the node will be made.

For simplicity let's assume thereon that all nodes have identical transmission ranges, thus $r_T = r_R \equiv r$. It is observed that, depending on the distribution of the traffic in the network, several communication pairs can coexist with T-R. Among these there are many that do not influence the success or failure of the handshake between T and R and some others that have a direct effect on it at a higher or lower scale. As will be described immediately failure: Command not found.

afterwards, the interfering pairs that influence the outcome of the communication process between T and R are the following and only them:

(a) A station in H_R sends data to a station in N_{TR}.
(b) A station in H_R sends data to a station in H_T.
(c) A station in H_R sends data to a station in Ω_{TR}.
(d) A station in H_R sends data to a station in H_R.
(e) A station in Ω_{TR} sends data to a station in H_T.
(f) A station in Ω_{TR} sends data to a station in N_{TR}.
(g) A station in Ω_{TR} sends data to a station in H_R.
(h) A station in N_{TR} sends data to a station in N_{TR}.
(i) A station in N_{TR} sends data to a station in H_R.
(j) A station in N_{TR} sends data to a station in H_T.
(k) A station in N_{TR} sends data to a station in Ω_{TR}.
(l) A station in H_T sends data to station in N_{TR}.
(m) A station in H_T sends data to a station in H_R.
(n) A station in H_T sends data to a station in H_T.
(o) A station in H_T sends data to a station in Ω_{TR}.

The interfering pairs described above do not all behave in distinct ways but they can be categorised in several groups. In particular, the pairs (h)–(k) whose transmitting station belongs to the set N_{TR}, i.e. can be heard by both T and R have identical effects on the handshake between T and R and will be referred to thereon in a unified manner as Group (A). Similarly the pairs of type (l)–(o) whose transmitting station is a one-hop neighbour of T, but is hidden from R will be referred to as Group (B). Finally, the pairs (a) and (b) also cause the same kind of interference to T-R and will be referred to as Group (C). All the other types of interferers have distinct

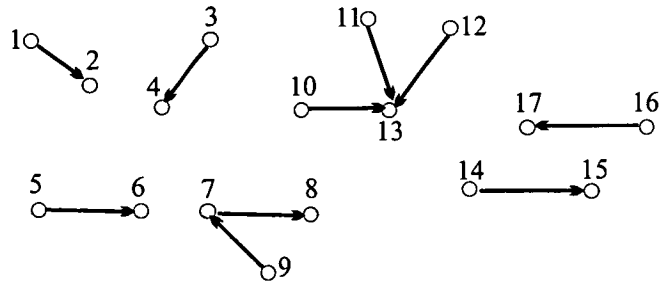


Figure 3.3: *An example of a random wireless network consisting of 17 nodes. Traffic flows only between adjacent nodes. Every node can be the transmitter for only one pair. Various types of interfering pairs exist for each such communication pair.*

characteristics and will be examined separately.

3.2.2 Spatial distribution of the interfering pairs

Another issue of primary importance is the spatial distribution of the interfering pairs, i.e. the frequency of their appearance in a given network. Obviously, in general, this can be dependent on many factors, such as the type of network, i.e. is it a WLAN, a pure ad hoc network or a WMN, the type of application which often imposes a certain node arrangement, the routing protocol or possible traffic control algorithms, which shape the packet routes in a specific way biasing the network graph, etc. In this chapter, two types of network connectivity graphs will be considered. Primarily, examples of symmetric networks will be examined, such as the star topology and the grid topology (see Section 3.3.1). Moreover, due to the fact that the purpose of this chapter is not to apply it to a specific scenario, but to remain as general as possible, the distribution of pairs in a random wireless network will be considered, where traffic flows between pairs of neighbouring nodes, so that no routing protocol is necessary. Every station can be the transmitter for only one pair (since half-duplex transceivers are assumed and are the most common in practical applications); however, a node can serve as a receiver for more than one communication pairs. An example of such a network consisting of 17 enumerated nodes is depicted in Figure 3.3. Observe for instance that node 13 is the receiver of three data flows from nodes 10, 11 and 12.

In such random configuration, the analytical derivation of the distribution of interfering pairs is basically a geometrical problem which was addressed in [39]. Here, the use of Monte-Carlo simulations is adopted which are stochastic techniques based on random number genera-

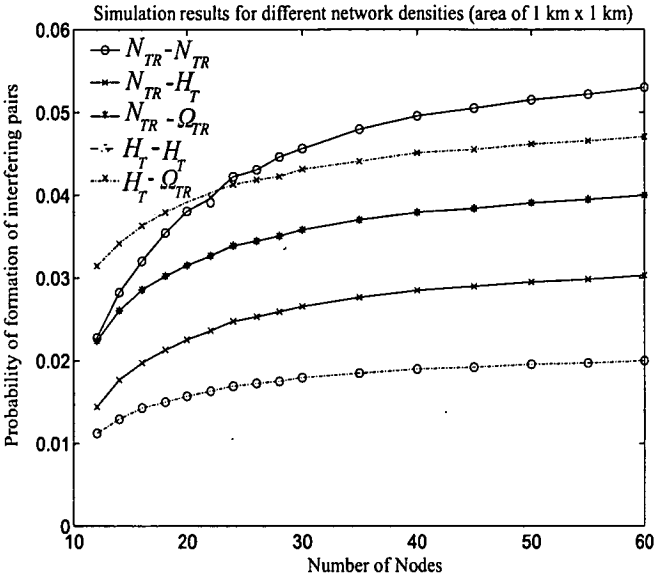


Figure 3.4: *Probability of formation of interfering pairs for different network densities.*

tors [67] for the computation of the number of competing pairs[1] for every network considered. Provided that the number of independent simulation runs is high enough (in the order of hundreds of thousands or millions of runs for each configuration), and that independent random number generation seeds are used, the objectivity of the simulation results can be guaranteed.

As an example one can refer to Figure 3.4; in this figure the normalised frequency[2] of several pairs is plotted against the number of nodes in a network of constant surface, i.e. a square area of 1000x1000 m^2. The dash between the transmitter and the receiver implies that both directions of transmission are taken into account, e.g. the pair $N_{TR}-\Omega_{TR}$ represents the spatial frequency of both pairs (f) and (k). As mentioned before, a single node can be the transmitter of only one pair, but it can be the receiver of more than one communication pairs. It can be noticed that, apart from the case of very low network density, the most dominant interfering pair is the one where, both its transmitter and its receiver are inside the range of the tagged pair ($N_{TR}-N_{TR}$). This is expected since, if two stations are located in this area, they are almost certainly going to form a communication pair due to their close proximity. Analogous arguments explain the relative frequencies of the other pairs. Finally, it is worth making the following observations:

[1]Note that the terms interfering pair and competing pair are used interchangeably.

[2]The spatial frequencies of the interfering pairs do not sum up to 1; the missing term is the relative frequency of all the communication pairs of the network that do not cause any interference to the tagged pair, which are in fact the most numerous.

1. There is an interesting symmetry—apart from the obvious, e.g. the symmetries between pairs (f) and (k), (a) and (i) or (c) and (g)—that is not apparent from Figure 3.4, as the reference to all the pairs would make the figure unreadable. In particular, it was noticed that the spatial distribution of the pairs (d) and (f) is identical. This fact becomes obvious if one observes Figure 3.2 and considers the tagged pair and the interfering pair interchangeably.
2. The normalised frequency of the pairs (b) and (m) is negligible (the probability of two nodes in these regions to be in transmission range of each other is almost zero), thus, it is omitted from the graph.

3.2.3 The ancestor model for a fully-connected network and its linkage to the model proposed in this chapter

The analysis that follows in this chapter is based on the model proposed by G. Bianchi [2] and was later extended by some researchers [37, 38] in order to characterise the performance of IEEE 802.11 in multi-hop networks. In [2] a fully-connected network is assumed, where all nodes try to access the channel in order to send data to the AP. Due to the obvious symmetry, the problem can be solved following the so-called 'decoupling approximation' (see [44]) that essentially means that every node tries to access the medium independently of its competitors and, thus, the same probabilities can be used for every station. Thus, the aggregate network throughput can be calculated on the basis of two basic quantities, namely the conditional collision probability p, which is the probability that the node will encounter a collision, conditioned on it having transmitted, and the transmission probability τ, which is the probability that in a time-slot the node will transmit a packet (successfully or unsuccessfully)[3].

It is suggested that the times at the end of every transmission (failed or successful) in a fully-connected network are renewal points [68] in the sense that the process is regenerated and the outcome, or the duration of the next event is not dependent on the previous. Thus, the analysis of all the events occurring in such a renewal interval is sufficient to derive an expression for the network throughput. Moreover, a Markov Chain is used to describe the backoff process of each station, where the transmitting states are the ones which correspond to zero backoff counter.

[3]Note that time-slot is on purpose not clearly defined here; for now, let's assume that it is the unit of time, which, however, can be of variable length. For more details on that matter the reader should refer to [2] and Chapter 4.

The transmission probability is therefore computed to be given by the following equation:

$$\tau = \frac{2\,(1-2p)}{(1-2p)\,(W_0+1)+pW_0\,(1-(2p)^m)}, \tag{3.1}$$

where W_0 is the initial backoff contention window and m is the maximum backoff stage, such that the maximum backoff value is given by $W_{max} = 2^m W_0$. Bianchi derived (3.1) under the assumption of an infinite retry limit; however, a couple of years later, Wu *et al.*, in [59], corrected the above equation for a case of finite retry limit, as is actually the case in IEEE 802.11:

$$\tau = \frac{2(1-2p)\left(1-p^{RL+1}\right)}{(1-2p)\,(1-p^{RL+1})+W_0\,(1-(2p)^{m+1})\,(1-p)+W_0 2^m p^{m+1}(1-2p)\,(1-p^{RL-m})}, \tag{3.2}$$

where RL is the retry limit (referred to as Short Retry Limit in [1]) and it is assumed that $RL > m$. The latter equation for τ is the one adopted in this chapter, as simulations with a finite retry limit that comply with the standard are run.

Furthermore, the slotted model used in [2] is based on a simple equation for the conditional collision probability

$$p = 1-(1-\tau)^{n-1}. \tag{3.3}$$

Equation (3.3) implies that the transmitted packet collides if at least one of the $n-1$ one-hop neighbours of the tagged station (it is assumed here that the fully-connected network has n station in total) transmits in the same slot. The above observation is not valid in an arbitrary topology, when interferers that belong neither to Group (A) nor (B) are present, and which cannot sense the sender's transmission directly. For the most simple scenario of this case refer to Figure 3.5.

Finally, in [2] the values of the two unknown probabilities τ and p—and, consequently, the throughput—are calculated by solving, with numerical techniques, the system of equations (3.1) and (3.3), which is proven to have a unique solution according to the fixed-point theorem [69][4].

[4] A fixed-point of a function is a point that maps itself to the function. In particular, let's consider an equation of the form $x = f(x)$. If $f : [0,1] \to [0,1]$ and, also, $f(.)$ is continuous, according to Brouwer's fixed-point theorem, f has at least one fixed point in the interval $[0,1]$, i.e. there is at least one $x \in [0,1]$ that satisfies the equation $x = f(x)$. If, in addition, $f(.)$ can be proven to be monotonically increasing (or decreasing), then this point is unique.

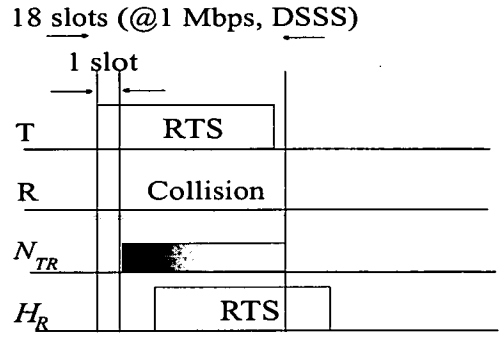


Figure 3.5: *T sends an RTS packet to R in order to establish a free channel. All nodes that belong to N_{TR} after the very short time period of one slot sense the channel busy and defer their own transmission until the channel is free again for more than DIFS [1]. On the contrary, any node that belongs to set H_R can only infer the ongoing transmission from a packet sent by R, thus, if its counter becomes zero during any of the 18 slots of RTS and, consequently, it also transmits an RTS, both packets will be destroyed at R.*

3.2.4 The main assumptions

The goal of this chapter is to propose a modelling framework that takes into account the time period during which an arbitrary transmission pair T-R is affected by each type of interfering pair (this period is usually referred to in the literature as 'vulnerable period'), and, consequently, derive expressions for the conditional collision probability and throughput. Throughout this chapter, it will be assumed that every station is examined in steady-state conditions, and that it is characterised by the same fundamental quantities—collision and transmission probability—as every other station. In other words, phenomena such as unfairness or short-term behaviour will not be considered, as the principal aim is to calculate the aggregate throughput in a random network. This is very useful when, for instance, a wireless device such as a laptop or a Personal Digital Assistant (PDA) wants to have a real-time estimate of the average throughput it will gain on entering the network[5]. One should expect that the results will converge to the actual performance for symmetric topologies, where indeed all stations are characterised by the same connectivity graph around them, thus, they have very similar, if not identical, τ and p. Another application environment where this modelling framework would be suitable is a very

[5]It is true that the addition of a new node may sometimes substantially alter the network connectivity graph, especially in small networks, and that, depending on its position, the station may experience different throughput levels. Still, the aggregate throughput that is derived in this chapter remains a useful and easily calculated quantity, which can be found in approximately real-time, i.e. in much less than a second using MATLAB®, which is not as fast as other programming languages like C/C++.

large network where boundary effects can be neglected without sacrificing the accuracy of the solution. As a sidenote, the majority of the related research makes similar hypotheses with the exception of [20] and [39].

3.2.5 Modelling interference under the assumption of at most one interfering pair at a time

Let's assume that, without loss of generality, there is a T-R communication pair that starts the RTS/CTS/DATA/ACK handshake at time $t = 0$. This communication can be disrupted in several ways by another communicating pair, i.e. the interfering pair. The effect depends on both the relative location of T-R and the competitor and, also, on the relative time difference of their handshakes. A time-frame is considered that extends one RTS/CTS/DATA/ACK transmission duration before and after the moment that T started sending the RTS control packet to R ($t = 0$). At any time during this time-frame, the interferers can initiate their transmission and influence (or not) the correct reception of DATA from R. The system is assumed to be in steady-state conditions.

To make the analysis tractable and simple, assume for the moment that, for each case considered, only one pair that can potentially cause collisions to T-R can exist. Later in this chapter this assumption will be relaxed. By this hypothesis the effect of every interfering pair can be examined independently. The case study is the RTS/CTS scheme of 802.11; however, similar arguments can be developed for any other CSMA/CA based protocol.

Provided that T starts transmitting the RTS packet at $t = 0$, only three possible events can occur:

—**Success**: In this event, the handshake between T and R is not affected by the interferers, the DATA packet can be delivered to R without errors and the communication is completed at time RTS + SIFS + CTS + SIFS + DATA + SIFS + ACK.

—**Control Packet Collision**: This event happens in the case that T does not get a response (CTS) within CTS_Timeout after it finished transmitting the RTS packet. This may be caused by a collision between T's RTS and another control or DATA packet, or because R has adjusted its Network Allocation Vector (NAV) [1] before it received the RTS. In that case, after the expiry of CTS_Timeout, T goes into backoff and retransmits at a later point in time.

—**Data Packet Collision**: As mentioned in Chapter 2, there are certain cases, even under perfect channel conditions, where the RTS/CTS scheme fails and the DATA packet that follows

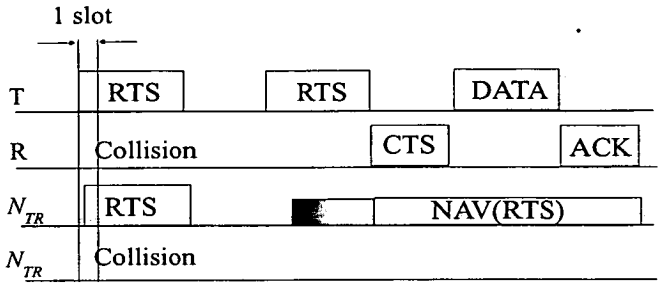


Figure 3.6: *In this case the transmitter of the interfering pair can hear both T and R. Consequently, the only case where collision occurs is when the counters of both transmitters expire in the same slot.*

collides with another transmitted packet. In such a scenario, T senses that its DATA packet was not successfully sent only after DATA + ACK_Timeout from the moment it started the DATA packet transmission. Thus, this collision incurs a high penalty in terms of delays.

3.2.5.1 Interference Group A

When the competing pair belongs to Group A, the only time period where a collision may occur is one time-slot σ. After the end of the time-slot, all the transmitters of Group A sense a busy medium, and, thus, defer their transmission until the handshake between T and R is completed. For an example of both cases refer to Figure 3.6. If E_A denotes the event that conditioned on T attempting to transmit to R, it suffers from interference from a pair that belongs to Group A, then a successful communication between T and R—despite the presence of a Group A pair—occurs with probability

$$P\left\{\bar{E_A}\right\} = (1-\tau)^{\frac{\sigma}{\sigma}} = 1-\tau, \tag{3.4}$$

where the vulnerable period has been normalised by the time-slot, which is the unit of time.

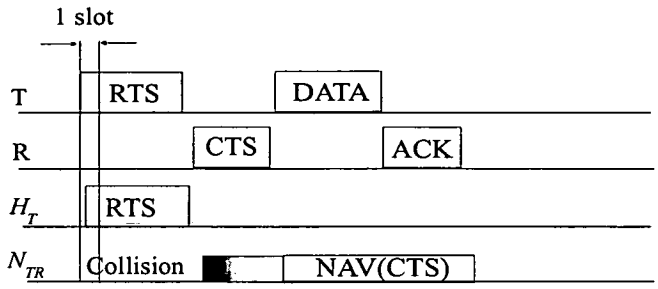


Figure 3.7: *This is a non-interfering group. Because the transmitter of the competing pair is heard by T, but not by R, even when it decides to transmit simultaneously with T, its RTS does not reach the receiver R.*

3.2.5.2 Interference Group B

Furthermore, no interference[6] is caused when the competitor belongs to Group B, as in this case the transmitters sense the channel busy as soon as T has started sending RTS (after one time-slot actually). Even in the case that their counter expires in this crucial time-slot, their RTS will not be received by R (remember that transmitters of this group are hidden from R), so the handshake between T and R will resume without interference (on the contrary, the communication of the competitor may be unsuccessful as shown in Figure 3.7). Thus, using similar notation as in the previous case

$$P\left\{\bar{E_B}\right\} = 1. \tag{3.5}$$

3.2.5.3 Interference Group C

As far as Group C is concerned, the transmitter of the competing pair is a hidden node from T. This is translated into collisions occurring when the RTS packets of T and of the interferer coincide in time as was already shown in Figure 3.5. During the critical period that is included between RTS + SIFS before and RTS + SIFS after $t = 0$, any RTS transmission from a Group C interferer will result in a collision of both RTS packets and a retransmission at a later point in time. Outside this interval, either R sends CTS to T after successfully receiving the RTS packet

[6] As a sidenote, it can be mentioned that the interfering pair (o) that belongs to this group, is what is usually termed in the literature as the exposed terminal problem. In this case, there is no interference, in the sense that the interferer listens to the RTS sent by T and, thus, defers its transmission, so there are no collisions occurring; however, a concurrent transmission by the interfering pair would not cause problems to either of the pairs, thus, resources are wasted. This effect is taken into account in the framework introduced in this chapter via equation (3.29).

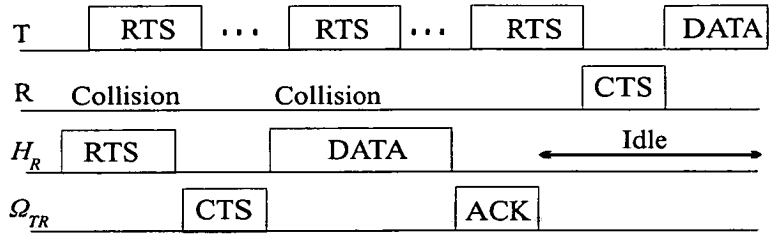


Figure 3.8: *Examples of the interference caused by a pair of type (c) are depicted for scenarios when the competing pair starts its communication earlier than T-R. The first two RTS packets are unsuccessful, as they collide with RTS and DATA respectively sent by the transmitter H_R. The final RTS manages to be successful, as it is sent during the ACK, DIFS and backoff time, where no interference is caused.*

and the interferer is silenced, or, if the transmission of the interferer precedes $t = -\text{RTS}-\text{SIFS}$, T defers and does not initiate transmission. Consequently, given that a pair of Group C is present and that T initiates transmission, a successful handshake between T and R occurs with probability

$$P\left\{\bar{E_C}\right\} = (1-\tau)^{2\frac{\text{RTS}+\text{SIFS}}{\sigma}}. \tag{3.6}$$

3.2.5.4 Interference Pair (c)

The effect of a pair of type (c) is more complicated. Let us assume again that T starts sending RTS to R at $t = 0$ and that H_R follows. One can easily verify that the vulnerable period is RTS + SIFS; after that, T-R communication may resume successfully, as in previous cases. However, in the cases when the transmission of the competing pair precedes the one between T and R ($t < 0$), then the RTS packet of T collides with either the RTS or the DATA packet sent by the interfering transmitter (see Figure 3.8). The only chance T has to get its RTS through is after DATA packet has been sent (note that the ACK packet of the interferer is sent by Ω_{TR}, which causes no interference to R). As a consequence, the probability of success, conditioned that a pair of type (c) is present, is given by

$$P\left\{\bar{E_c}\right\} = (1-\tau)^{\frac{\text{RTS}+\text{SIFS}}{\sigma}+\frac{T_{nack}}{\sigma}}, \tag{3.7}$$

where T_{nack} = RTS + SIFS + CTS + SIFS + DATA.

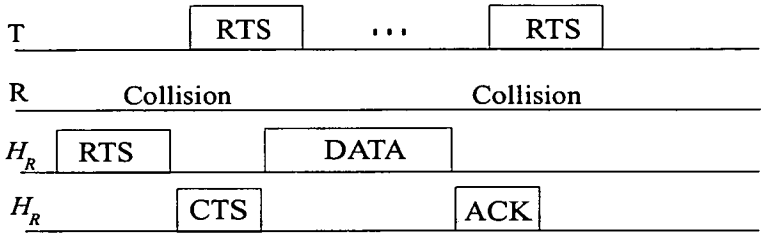


Figure 3.9: *The only difference between the present and the previous case is that collision may also occur during the ACK packet sent by the receiver of the competing pair belonging to set H_R.*

3.2.5.5 Interference Pair (d)

This interfering pair is very similar to the previous, the only difference being that packets from both the receiver and the transmitter of the pair (both belong to set H_R) can cause interference to R. In particular, the only time gap inside which the RTS from T can be successful is after transmission of ACK is concluded (see Figure 3.9). Thus, given that a pair of type (d) lies in its neighbourhood, the communication between T and R is successful with probability

$$P\left\{\bar{E}_d\right\} = (1-\tau)^{\frac{\text{RTS+SIFS}}{\sigma}+\frac{T_{ack}}{\sigma}}, \tag{3.8}$$

where T_{ack} = RTS + SIFS + CTS + SIFS + DATA + SIFS + ACK.

3.2.5.6 Interference Pair (e)

As far as the interference pairs whose transmitter belongs to set Ω_{TR} are concerned, any potential implication to the communication of T and R will be a result only of the receiver, to whom Ω_{TR} is transmitting. This is so, because any station in Ω_{TR} can be heard by neither T nor R, thus it can not have a direct effect on them, unless its receiver is interfering. Due to this argument, a pair of type (e), whose receiver belongs to the set H_T, never causes interference to T-R. As it is shown in Figure 3.10, H_T either defers because of packets sent by T, or, even if it receives the RTS correctly and responds with a CTS in parallel with a transmission of T, this CTS can not reach R. For these reasons

$$P\left\{\bar{E}_e\right\} = 1. \tag{3.9}$$

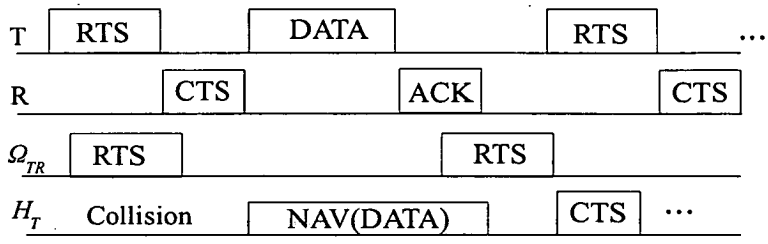


Figure 3.10: *A pair whose transmitter belongs to the set Ω_{TR} and whose receiver to H_T never causes interference to T-R.*

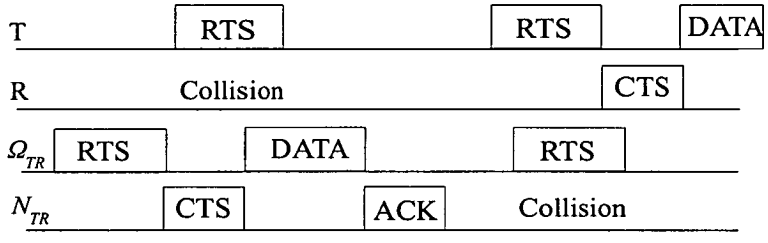


Figure 3.11: *In this case collision occurs only when H_T starts sending a CTS packet back to Ω_{TR} at the same slot when T starts sending RTS to R.*

3.2.5.7 Interference Pair (f)

From the arguments explained above, it should be clear that the only effect this competing pair has on T-R is when the event of T starting its RTS transmission occurs at the same slot as when the receiver N_{TR} starts replying to Ω_{TR} with a CTS packet (provided that Ω_{TR} has sent RTS to N_{TR} prior to T commencing its transmission). This scenario is shown in Figure 3.11 and the probability of success is

$$P\left\{\bar{E}_f\right\} = 1 - \tau. \tag{3.10}$$

3.2.5.8 Interference Pair (g)

When the interference pair comprises a node in the set Ω_{TR} as the transmitter, and a node in the set H_R as the receiver, special attention is required. To begin with, the scenarios where Ω_{TR} starts sending its RTS in the time frame between $-T_{ack}$ and $-$RTS$-$CTS$-$SIFS before time $t = 0$ (when T is supposed to start sending RTS) are considered. In this case, even though T does not sense the other pair's transmission and sends RTS, R has previously listened to the CTS packet from H_R and has adjusted its NAV; so, it does not respond to T. Remember that in this

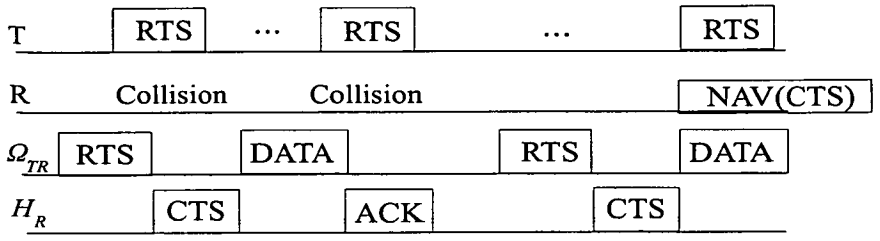


Figure 3.12: *For the scenarios where the interfering pair precedes, T does not get a response from R, either because its RTS packet directly collides with CTS from H_R, or because R defers until the handshake of the other pair is complete.*

section the assumption of only one competing pair is adopted, so CTS is definitely sent by H_R. Furthermore, when Ω_{TR} starts sending its RTS in the time frame between $-$RTS$-$CTS$-$SIFS and $-$SIFS, the CTS packet of H_R collides with the RTS sent by T[7]. These scenarios are shown in Figure 3.12.

However, apart from the possible interference described above, this pair can potentially cause another type of interference, which is defined as a data packet collision at the beginning of this chapter[8]. In particular, let's suppose that RTS is sent by T to R at time $t = 0$, and that Ω_{TR} does the same at some point in the time frame [SIFS, RTS + CTS + SIFS]. As a consequence, there is a collision of Ω_{TR}'s RTS and the CTS sent by R at H_R. Thus, Ω_{TR} does not get a response from H_R and it goes into backoff. Meanwhile, T receives the CTS from R correctly and it starts the transmission of DATA packet. Depending on the size of the DATA packet, the value of CTS_Timeout and the time which Ω_{TR} will spend in backoff, there are two possible outcomes of the above situation:

In the first, Ω_{TR} retransmits an RTS, which is correctly received by H_R, that in turn responds by sending CTS. Because of the broadcast nature of the wireless medium, this CTS is also received at R, thus corrupting the DATA packet sent by T, see Figure 3.13(a). In the other possible case (complementary to the previous) illustrated in Figure 3.13(b), the reception of DATA at R finishes before H_R transmits a CTS, so the handshake of T-R is successful.

Let $\breve{q}$ denote the probability of data packet collision occurring for the pair T-R, given that T starts sending RTS at time $t = 0$, and, also, given that Ω_{TR} starts transmitting at some point

[7]Actually, one can easily notice that in the interval [$-$SIFS, SIFS] the handshake between T and R is successful.

[8]Note that, recently, [65] considered this problematic topology and suggested a dynamically adjusted Extended Interframe Space (EIFS) [1] value in order to enhance the performance of carrier sensing. Here, no such knowledge is assumed and it is considered EIFS = SIFS + CTS + DIFS, as usually in the literature.

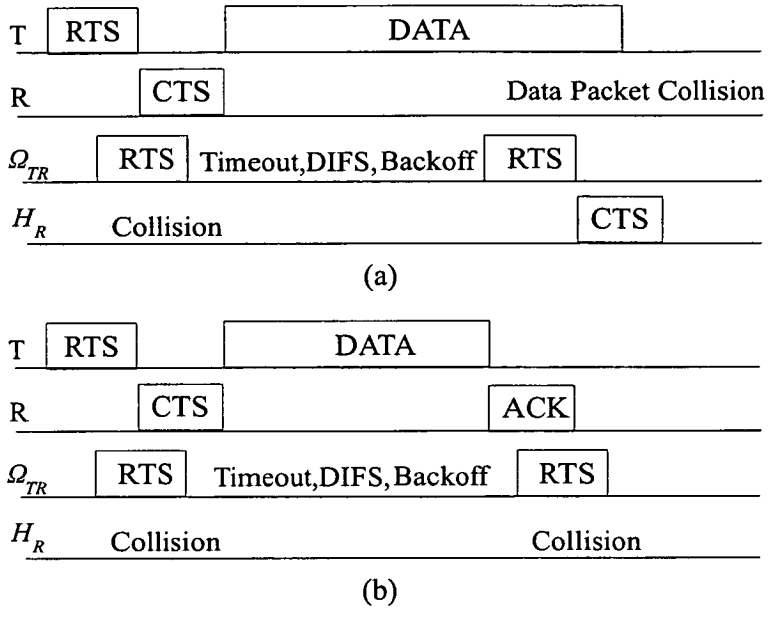


Figure 3.13: *Depending on the values of CTS_Timeout, the length of the DATA packet, the transmission rates used and the backoff of station Ω_{TR}, the outcome of Ω_{TR} initiating RTS in the interval [SIFS, RTS + CTS + SIFS] can be either data packet collision or success for the pair T-R.*

(denote this point in time with reference to $t = 0$ as t_1) in the interval [SIFS, RTS+CTS+SIFS]. The probability $\breve{q}$ can be calculated as follows: Examining Figure 3.13(a), there is a data packet collision when

$$t_1 < \text{CTS} + \text{SIFS} + \text{DATA} - \text{CTS\_Timeout} - \text{RTS} - \text{DIFS} - \text{Backoff}. \tag{3.11}$$

Let $G \triangleq \text{CTS} + \text{SIFS} + \text{DATA} - \text{CTS\_Timeout} - \text{RTS} - \text{DIFS}$, $u_{min} = \text{SIFS}$, and $u_{max} = \text{RTS} + \text{CTS} + \text{SIFS}$. The probability $\breve{q}$ can be written as:

$$\breve{q} = \sum_{i=0}^{RL} \sum_{k=0}^{W_i - 1} \breve{q}_{ik} b_{ik}, \tag{3.12}$$

where $\breve{q}_{ik}$ is the conditional probability of a data packet collision, conditioned on the backoff period of Ω_{TR} having a start-up value of backoff stage i and backoff counter k (obviously $0 \leq k \leq W_i - 1$). Also, RL is the Short Retry Limit, W_i is the contention window at stage i and b_{ik}, which is the steady-state probability that an arbitrary station is in backoff stage i and

has a counter of k, is given in [59] by the equation

$$b_{ik} = \frac{W_i - k}{W_i} \frac{2p^i\eta(1-p)}{\eta\left(1-p^{RL+1}\right) + W_0\left(1-(2p)^{m+1}\right)(1-p) + W_0 2^m p^{m+1}\eta\left(1-p^{RL-m}\right)}, \tag{3.13}$$

where $\eta \triangleq 1 - 2p$ and m is the maximum backoff stage. The conditional probabilities $\breve{q}_{ik}$ of (3.12) are equal to:

$$\breve{q}_{ik} = \begin{cases} 1 & , k \leq \frac{G-u_{max}}{\sigma} \\ \frac{G-k\cdot\sigma-u_{min}}{u_{max}-u_{min}} & , \frac{G-u_{max}}{\sigma} < k < \frac{G-u_{min}}{\sigma} \\ 0 & , k \geq \frac{G-u_{min}}{\sigma} \end{cases} \tag{3.14}$$

If one substitutes equations (3.13) and (3.14) into (3.12), the probability $\breve{q}$ is calculated. As a sidenote, equation (3.14) is derived when one co-verifies the inequality (3.11) with SIFS $<$ t_1 $<$ RTS + CTS + SIFS. In order to make the derivation of the above probabilities tractable, it is assumed that the backoff period of (3.11) is an integer multiple of the channel slot σ (i.e. Backoff $= k \cdot \sigma$, where σ in DSSS, for example, is equal to $20\mu s$) without taking into account the fact that the backoff counter may freeze and, consequently, each slot may have a longer duration than σ.

Concluding the analysis of the interference pair (g), and combining all the analysis described above, one can state that the probability of success for T-R, conditioned on a pair of type (g) being present, is given by

$$P\left\{\bar{E}_g\right\} = (1-\tau)^{\frac{T_{ack}}{\sigma}} \cdot (1-\tau\breve{q})^{\frac{\text{RTS+CTS}}{\sigma}}. \tag{3.15}$$

After the analysis of all the types of interferers described in detail above, one could state that the conditional collision probability p under the assumption of at most one interference pair at a time would be given as the probability of the complement of the junction of all the successful events mentioned previously ($\bar{E}_A$,$\bar{E}_B$,..., until $\bar{E}_g$). However, such an assumption is not realistic, thus the analysis is extended in the next section, where multiple interfering pairs are considered.

3.2.6 Modelling interference under multiple competing pairs

The key observation that has led to the incorporation of multiple competing pairs in the framework that was developed previously is simply that, when more than one pair exists in the neighbourhood of T-R simultaneously, the competing pair could face collisions from a third pair. To be more precise, consider again Figure 3.8, where the interfering pair has an H_R as a transmitter and an Ω_{TR} as a receiver. In this scenario, collision for T would not occur if the communication exchanged between H_R and Ω_{TR} was interrupted, because the RTS of H_R collided with another packet and, so, Ω_{TR} received nothing useful. Consequently, to be more precise, and if p_{co} is the probability that a station did not get a response to it sending an RTS packet[9], then the probability of success of Section 3.2.5.4 can be corrected as follows

$$P\left\{\bar{E}_c\right\} = (1-\tau)^{\frac{2\text{RTS}+\text{SIFS}}{\sigma}} \cdot \left[1-\tau(1-p_{co})\right]^{\frac{T_{nack}-\text{RTS}}{\sigma}}. \tag{3.16}$$

Following a similar argumentation, which will not be repeated here explicitly, as it is straightforward (the general rule is that when the collision of the RTS packet of T occurred with a CTS packet from the competing pair, the analysis must be reconsidered), the following corrected expressions for the corresponding probabilities of success for the several types of interferers can be obtained. Note that some remain the same, independently of the presence of one or multiple pairs.

$$P\left\{\bar{E}_A\right\} = 1-\tau. \tag{3.17}$$

$$P\left\{\bar{E}_B\right\} = 1. \tag{3.18}$$

$$P\left\{\bar{E}_C\right\} = (1-\tau)^{2\frac{\text{RTS}+\text{SIFS}}{\sigma}}. \tag{3.19}$$

$$P\left\{\bar{E}_d\right\} = (1-\tau)^{\frac{2\text{RTS}+\text{SIFS}}{\sigma}} \cdot \left[1-\tau\left(1-p_{co}\right)\right]^{\frac{T_{ack}-\text{RTS}}{\sigma}}. \tag{3.20}$$

$$P\left\{\bar{E}_e\right\} = 1. \tag{3.21}$$

$$P\left\{\bar{E}_f\right\} = 1-\tau\left(1-p_{co}\right). \tag{3.22}$$

$$P\left\{\bar{E}_g\right\} = \left[1-\tau\left(1-p_{co}\right)\right]^{\frac{T_{nack}}{\sigma}} \cdot \left[1-\tau\left(1-p\right)\right]^{\frac{\text{ACK}}{\sigma}} \cdot \left[1-\tau\breve{q}\left(1-p_{co}\right)\right]^{\frac{\text{RTS}+\text{CTS}}{\sigma}}. \tag{3.23}$$

[9]Note that p and p_{co} are slightly different. The former includes the probability of data packet collision referring to a pair of type (g), so $p \geq p_{co}$.

Furthermore, let $\check{s}_j$ denote the number of the corresponding type of pairs around the arbitrary communication pair T-R, where j depicts the type of interference pair, i.e. $j \in S_{IP} = \{\text{A,B,C,c,d,e,f,g}\}$. The quantity $\check{s}_j$ is derived by simple observation, as far as a symmetric network is concerned, or by Monte-Carlo simulations, as far as a random network is concerned, as was explained in Section 3.2.2, and is a function of the network density. To conclude the analysis, the final expression for the conditional collision probability p is given by the solution of the fixed-point equation [44,69] that follows:

$$p = 1 - \prod_{j \in S_{IP}} P\left\{\bar{E}_j\right\}^{\check{s}_j}. \tag{3.24}$$

Note that this equation implies that collision or success due to a specific pair is considered to be independent of the collision or success due to any other pair; this is why the product of the above probabilities is taken. Although this is a fundamental assumption, it enables the analysis to be tractable. In the following chapter, this assumption will be relaxed at the cost of loss of generality.

For the probability p_{co} a similar expression is found

$$p_{co} = 1 - \left\{[1-\tau(1-p_{co})]^{\frac{T_{nack}}{\sigma}} \cdot [1-\tau(1-p)]^{\frac{\text{ACK}}{\sigma}}\right\}^{\check{s}_g} \cdot \prod_{j \in S^*_{IP}} P\left\{\bar{E}_j\right\}^{\check{s}_j}, \tag{3.25}$$

where $S^*_{IP} = S_{IP} \backslash \{g\}$, i.e. the probability of a data packet collision for pair (g) has been excluded from p_{co} .

To conclude, a system of non-linear equations (3.2),(3.12),(3.24) and (3.25) with equal number of unknown parameters has been obtained. The system is solved using numerical techniques, solving the corresponding fixed-point equations (see Section 3.2.2), and numerical values for the transmission probability τ and the collision probabilities p and p_{co} are derived.

3.2.7 Derivation of the saturation throughput

For the derivation of the saturation throughput S a method similar to [2] is adopted, where renewal theory [68] is used. In particular, a renewal process is a stochastic process in which, the set of interarrival times $\{X_n; n \geq 1\}$ is a set of independent, identically distributed (i.i.d) random variables. Let's further suppose that a reward R_n is earned at the time of the nth renewal. According to the renewal reward theorem [70], the ratio $\frac{R_t}{t}$ of the aggregate reward

R_t during a period of time t, over the duration of the period of time t, approaches the ratio of the corresponding expectations $\frac{E[R]}{E[X]}$ as t approaches infinity. Consequently, if one views the throughput as the number of successful packets received by a station within a time frame, over the duration of the time frame, it is sufficient to calculate the average number of packets successfully received in a renewal interval (Bianchi's time slot) over the average duration of the renewal interval. This is the method that will be followed in this chapter.

Let L_d be the DATA packet size in bytes, which is assumed to be constant. The saturation throughput S for an arbitrary communication pair T-R is according to the previous analysis

$$S = \frac{\tau \cdot (1-p) \cdot L_d}{E[\sigma]}, \tag{3.26}$$

where $E[\sigma]$ denotes the average duration of a time-slot. This is obtained as follows: If the transmitter T sees an idle slot (with probability p_{id}) the duration of the time-slot is σ. In the opposite case (busy slot, with probability $1-p_{id}$), the duration can vary significantly, according to the event observed by station T. Namely, the slot duration in the busy case is equal to:
— a time period RTS + EIFS, if T suffered from (or observed) a control packet collision,
— the duration of a whole transmission $T_s = T_{ack}$ + DIFS, if the station itself sent a packet successfully,
— a significantly large wasted period T_{dpc} = RTS+SIFS+CTS+SIFS+DATA+ACK_Timeout, if T experienced a data packet collision,
— a time period T_s, if the station T froze its backoff counter because of the successful transmission of another pair, whose transmitter is a one-hop neighbour of T (i.e. the pairs of Groups A and B), and, finally,
— a time period $T_f = T_s$ − RTS − SIFS, if T froze its backoff counter because of the successful transmission of a pair, whose receiver and not its transmitter is a one-hop neighbour of T (i.e. the pairs of Group C and of types (e) and (f)).

The corresponding probabilities of the above events are still to be derived. The probability that the station T observes an idle slot p_{id} is derived as:

$$p_{id} = (1-\tau)(1-\tau)^{š_A}(1-\tau)^{š_B}[1-\tau(1-p_{co})]^{š_C}[1-\tau(1-p_{co})]^{š_e}[1-\tau(1-p_{co})]^{š_f}, \tag{3.27}$$

where the first term means that the node T under consideration is not currently transmitting, the terms $(1-\tau)^{š_j}$ refer to all the interfering pairs j, whose transmitters are heard by T, and the

terms $[1-\tau(1-p_{co})]^{\check{s}_j}$ refer to pairs j, whose receivers are heard by T.

The probability that the station observed a successful slot due to its own transmission, given that it observed a busy slot (a busy slot occurs with probability $1-p_{id}$), is

$$\frac{p_s}{1-p_{id}} = \frac{\tau(1-p)}{1-p_{id}}. \tag{3.28}$$

This is because the station attempts transmission with probability τ and this transmission is successful with probability $1-p$[10].

In addition, the probability that the station T observed a successful transmission by a competing pair whose transmitter is heard by T is

$$\frac{p_{ts}}{1-p_{id}} = \frac{(\check{s}_A+\check{s}_B)\tau(1-p_{co})}{1-p_{id}}, \tag{3.29}$$

whereas the corresponding probability when the receiver of the pair is heard by T is

$$\frac{p_{rs}}{1-p_{id}} = \frac{(\check{s}_C+\check{s}_e+\check{s}_f)\tau(1-p_{co})}{1-p_{id}}, \tag{3.30}$$

The probability that T experienced a data packet collision given that it observed a busy slot is

$$\frac{p_{dpc}}{1-p_{id}} = \frac{\tau(p-p_{co})}{1-p_{id}} \tag{3.31}$$

Finally, the probability that the station suffered from or observed a control packet collision, given that it observed a busy slot, can be indirectly calculated as the complementary probability of the sum of the four conditional probabilities derived above, i.e.

$$p_c = 1-\left(\frac{p_s+p_{ts}}{1-p_{id}} + \frac{p_{rs}}{1-p_{id}} + \frac{p_{dpc}}{1-p_{id}}\right). \tag{3.32}$$

The average duration of the time-slot can now be written as follows

$$E[\sigma] = p_{id}\sigma + (1-p_{id})\left[\frac{p_s+p_{ts}}{1-p_{id}}T_s + \frac{p_{rs}}{1-p_{id}}T_f + \frac{p_{dpc}}{1-p_{id}}T_{dpc} + p_c(\text{RTS}+\text{EIFS})\right] \tag{3.33}$$

Combining the equations (3.26) and (3.33), the throughput can be calculated.

[10]The probabilities τ and p are independent because, by definition, p is the probability of collision conditioned that the node transmitted.

3.3 Validation of the model—Discussion

In the present section the analytical framework defined previously is applied to several types of networks. Primarily, some examples of symmetric networks are given, and it is shown that the prediction for the saturation throughput from the proposed model is quite accurate. Moreover, the model is applied to a random network of variable population, and useful remarks and conclusions are drawn from the observation of its performance. In order to validate the model, the theoretical results are compared to simulation results obtained using the Network Simulator NS2 [71]. All network nodes employ the RTS/CTS access scheme of IEEE 802.11b [1] and, unless stated otherwise, the default parameters suggested by [1] are used in the simulations[11]. In particular, the transmission rates that are employed are the ones suggested by IEEE 802.11b, namely, the BasicRate, i.e. the rate at which the control packets are transmitted is chosen from the set $\{1, 2\}$ Mbps and the DataRate, i.e. the rate at which the data packets are transmitted, from the set $\{1, 2, 5.5, 11\}$ Mbps.

3.3.1 Symmetric networks

The first example of a symmetric network examined is the star topology depicted in Figure 3.14. Five nodes are the competing transmitters and they are all attempting to send information to the common receiver, defined as node 0, which lies in the center of the regular pentagon that is shaped. The topology is constructed in such a way that all transmitters are hidden from each other, but they lie within the transmission range of the receiver node 0. This network is entirely symmetric, in the sense that not only all the nodes have an identical number and type of interfering pairs around them, but also, the competing pairs themselves have the same connectivity graph around them. The theoretical results are presented in Figure 3.15 and are compared to the corresponding simulation results. In particular, the throughput per station is plotted against the data packet size, which varies from 256 bytes to 2294 bytes. As can be observed in Figure 3.15, the prediction of the proposed analytical model is quite accurate in this case. In fact, it is shown that for certain values of the transmission rates and the backoff-related parameters, the accuracy of the model is excellent, but for others the model prediction

[11]The only parameters, which IEEE 802.11 leaves to be configured by the wireless card manufacturer/user are the timeout values, CTS_Timeout and ACK_Timeout. In the simulations throughout this thesis, these are set equal to SIFS+CTS, e.g. with a BasicRate of 1 Mbps and DSSS as the spread spectrum technique, they are 314 μs. This is a realistic value, since, after the transmission of RTS (DATA), the station should have received a CTS (ACK) after SIFS+CTS (ACK). If it hasn't, it should go into backoff and reattempt transmission.

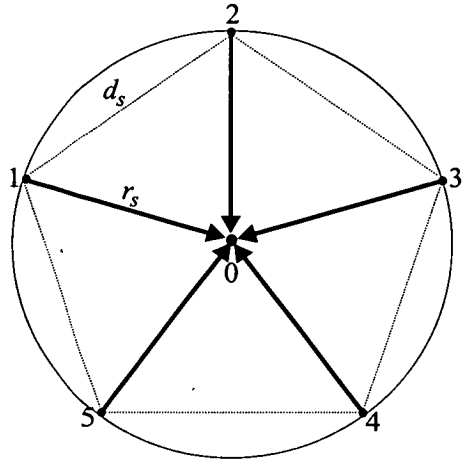


Figure 3.14: *The star topology is depicted in this figure. Nodes 1–5 are the competing transmitters who are sending data to the receiver, node 0. The distances are selected in such a way so that $r_s < r$, and at the same time $d_s > r$, where r is the transmission range, assumed to be constant. Thus, all the transmitters are hidden from each other, but they still can communicate with the receiver.*

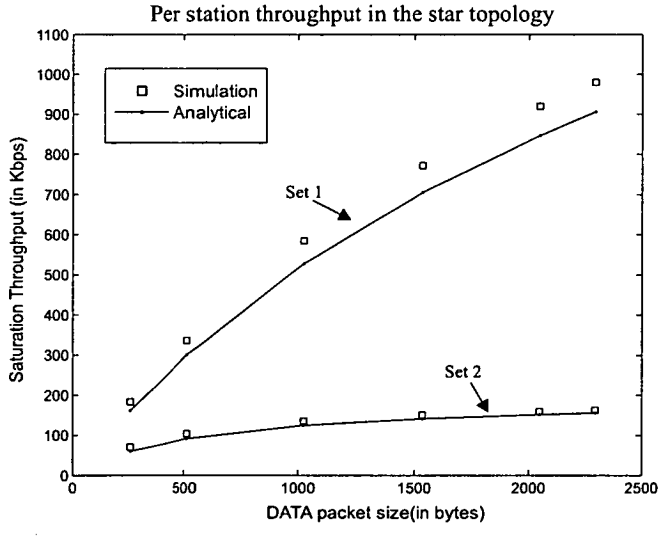


Figure 3.15: *The station throughput in the star topology is plotted against the packet size for two sets of parameters. Set 1 corresponds to a BasicRate of 2 Mbps, a DataRate of 11 Mbps, and also $RL = 6$ and $m = 5$. On the other hand, for Set 2 both transmission rates are 1 Mbps, $RL = 4$ and $m = 3$. The accuracy of the model is not constant and depends on the selection of the parameters.*

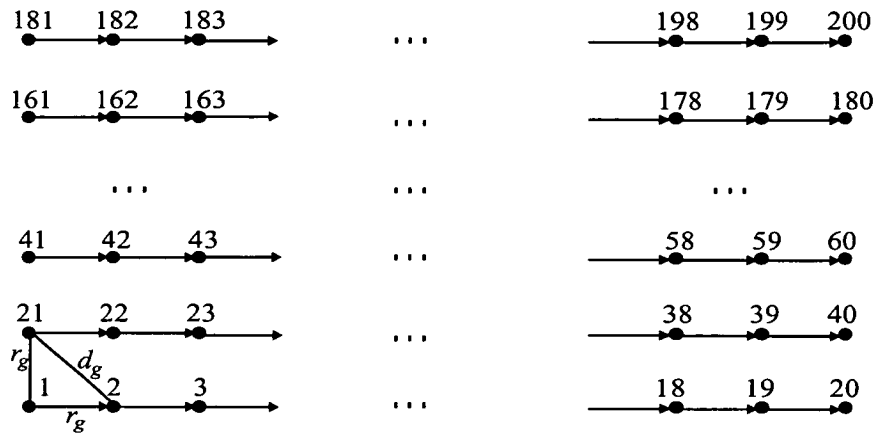


Figure 3.16: *The rectangular grid topology is demonstrated. The grid consists of 200 stations. The data flows horizontally from the left to the right, as shown. The distances are selected in such a way so that $r_g < r$, and at the same time $d_g > r$.*

deviates from the optimal. The reader should bear this observation in mind, as it is one of the crucial remarks that motivated the work of the following chapter.

Furthermore, a rectangular grid topology is considered, that comprises 200 stations and is depicted in Figure 3.16. The data flows horizontally from left to right, but it is not forwarded in intermediate stations (it is not a multi-hop communication example). Every station (apart from the ones on the far right of the connectivity graph) is a traffic source, that sends data to its destination, which is the station on its immediate right. The locations of the stations are determined in such a way that two successive nodes in either dimensions are in range of each other, but two diagonally connected nodes are hidden from each other. The theoretical model is applied to each one of the receivers, i.e. to all the nodes apart from the ones on the far left, which have zero throughput. The number and type of pairs around each receiver is found simply by observing Figure 3.16. Due to the obvious symmetry around the middle row of the grid, the throughput of only the stations 1–100 is presented.

If one observes the network connectivity graph closely, one can notice that the stations that lie in the central part of the grid, and are not located at the network boundaries, i.e. the stations 24–38, 44–58, etc. are characterised by exactly the same connectivity graph around them. This fact agrees with the assumption that was adopted in this chapter in order to make the model tractable, i.e. that the transmission and collision probabilities are the same for every node of the network. Therefore, primarily it is expected that the model will provide accurate estimates of the per

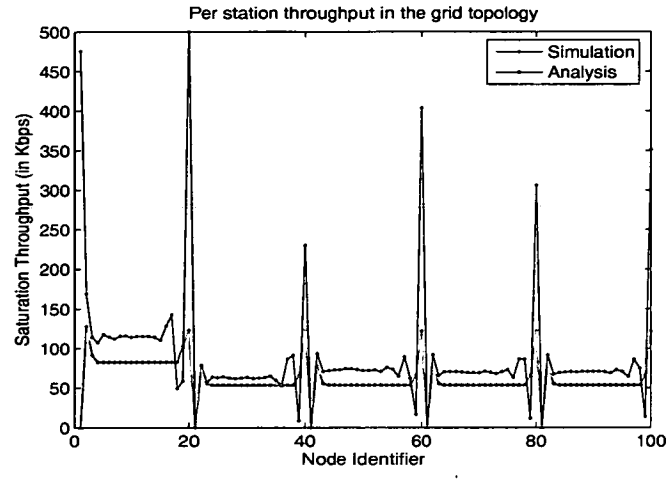


Figure 3.17: *The throughput at each receiver in the grid topology is plotted in this figure. The simulation results are marked with a black line, whereas the theoretical results are marked with a red line. Both transmission rates are set to 1 Mbps, the minimum contention window is* $W_0 = 32$, $RL = 4$, $m = 3$ *and the data packet size is selected to be equal to 1024 bytes. The model manages to give fairly accurate estimates for the 'middle' nodes, but it fails to follow the throughput of the boundary nodes.*

station throughput for these 'middle' nodes. On the contrary, the theoretical model prediction is expected to deviate from the simulation results for the boundary nodes. These observations are verified when one examines Figure 3.17. The model in general proves to be 'pessimistic' for the boundary nodes, as it relies on the hypothesis that they are characterised by the same collision and transmission probabilities as their neighbours. On the contrary, the latter (e.g. 2,20,40,60[12]) tend to exhibit particularly high throughput and lead to the starvation (i.e. a particularly low throughput) of their immediate neighbours. This phenomenon of unfairness is not grasped by the model and it requires a more detailed and computationally intensive calculation[13]. However, the model proves to be quite accurate as far as the central nodes are concerned; in fact, it has been observed that for larger grid networks, the percentage of the nodes for which the model provides accurate estimates of the throughput becomes increasingly higher, as the boundary effects are gradually alleviated.

[12] An interesting observation is that, although for example nodes 40 and 60 have exactly the same neighbourhood, their saturation throughput is very different. This is due to the fact that their neighbours/interferers have—in reality—quite different transmission and collision probabilities.

[13] For an iterative computation of the per station throughput that attempts to grasp the unfairness phenomena the reader is referred to [43].

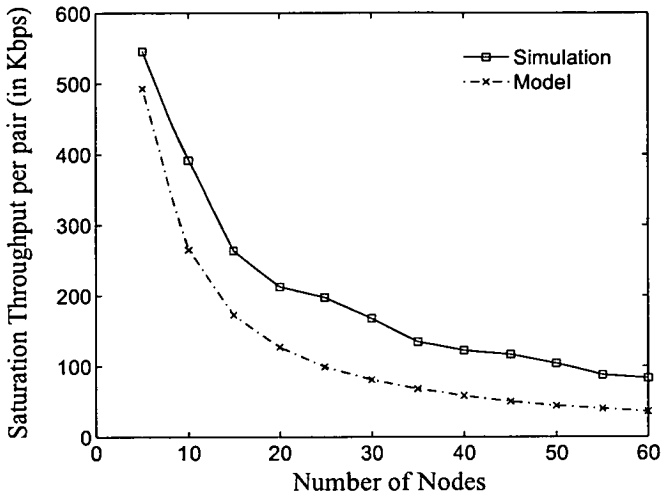


Figure 3.18: *Saturation throughput plotted against different network densities. The RTS/CTS scheme is used with 1 Mbps BasicRate and 1 Mbps DataRate. The minimum contention window is 32.*

3.3.2 Random networks

In this section the model is applied to randomly deployed networks. This thesis aims to analyse only the phenomena occurring at the medium access layer. Consequently, only direct transmissions to one-hop neighbours are considered and no routing protocol is used (actually, the absence of a routing protocol is realised in the Network Simulator NS2 by a simple protocol called 'DumbAgent'). For each simulation run, the network area is kept constant (1000m x 1000m), and the data packet rate is increased until the point of saturation. At this rate the throughput is measured. This procedure is repeated for 10 different network topologies for each trial. Note that the maximum contention window for all cases considered is 1024 and the Short Retry Limit is 7 as in [1]. The population of the network varies from 5 stations to 60 stations.

From the comparison of the analytical results to the simulation results shown in Figures 3.18 and 3.19 it can be observed that, although the model follows the trend of the aggregate per communication pair throughput quite well, it seems to underestimate the throughput. As a first remark it can be noted that the accuracy of the model is somewhat better when the transmission time of the control and data packets compared to the contention window are small. This is due to the following reason: The error in the estimation of the throughput is a result of an error in estimating the probabilities τ and p. The latter appear at both the numerator and the

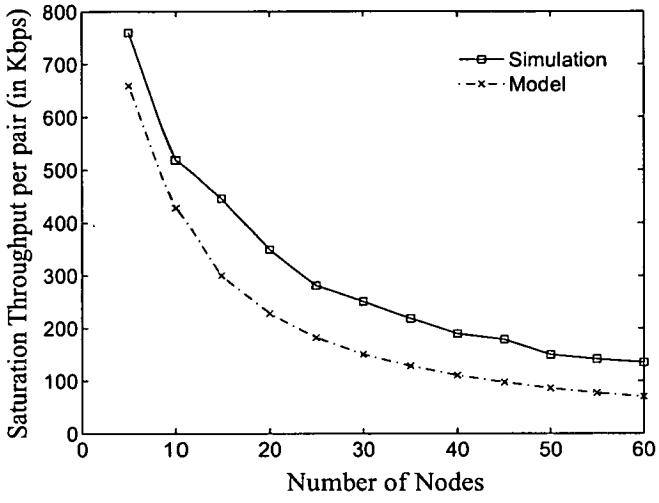


Figure 3.19: *Saturation throughput plotted against different network densities. The RTS/CTS scheme is used with 2 Mbps BasicRate and 11 Mbps DataRate. The minimum contention window is 128.*

denominator of (3.26). However, these probabilities in the denominator are multiplied by the corresponding event durations e.g. T_s, T_f, etc. and, thus, a large value of these durations renders the errors less apparent.

As a general conclusion it can be mentioned that, due to the independence assumption made in order to keep the analysis tractable, the model tends to find stricter conditions for successful transmission than in reality. Although the analysis of Section 3.2.6 partially compensated for this phenomenon, the assumption still seems to cause problems in the accuracy of the model. Consequently, although this general analysis can be a simple tool to gain some insight into the behaviour of wireless multihop networks, it still fails to be a trustworthy modelling tool. This is the motivation for the next two chapters of the thesis where all the assumptions of this chapter are lifted, and an accurate model for the classic hidden terminal scenario is presented.

Chapter 4
Revisiting the Classic Hidden Terminal Problem

4.1 Motivation: assumptions & limitations of previous methods

As far as their scalability is concerned, the analysis presented in Chapter 3, as well as several other proposed modelling techniques [20, 38–40] are quite efficient. However, a closer look at the assumptions that are being made makes it clear that the latter can lead to severe inaccuracies in a large number of cases. The purpose of this chapter is to explain the reasons that render them unable to model the hidden terminal phenomenon with a consistent accuracy, to show the relevant impact on the system performance and, more importantly, to propose a novel modelling method that relaxes most of the above assumptions, and makes a highly accurate description of the contention phenomenon. The tradeoff that arises is naturally the degradation in scalability, as an accurate analysis requires a more complex mathematical model. The present chapter, as well as Chapter 5, present the proposed model for the classic hidden terminal topology case depicted in Figure 4.1.

4.1.1 The notion of the slot

To date, the majority of papers dealing with the problem of theoretical analysis in multihop networks are based on extensions of Bianchi's bidimensional Markov Chain (MC) modelling [2]. In [2], a model that is based on the notion of a slot is considered. This slot is of variable length depending on the state of the channel. It can be of a length equal to a successful transmission T_s, e.g. equal to RTS + SIFS + CTS + SIFS + DATA + SIFS + ACK + DIFS for the RTS/CTS mechanism, when either the station itself or any of its one-hop neighbours occupy the channel successfully. Alternatively, the slot has the length of a collided transmission T_c[1] or, when the node is in backoff, it has the length σ, which is the term aSlotTime defined in [1]. aSlotTime is

[1] T_c is equal to RTS+Timeout+DIFS for the stations that were involved in the collision or RTS+DIFS for all the other stations that listened to the collision; in fact, in [2], it is assumed T_c = RTS + DIFS for all stations.

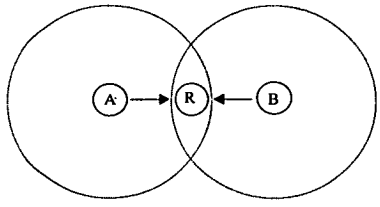


Figure 4.1: *The classic hidden terminal scenario.*

the critical time needed at any node to detect the transmission of a packet from any other station in the fully-connected network and depends on physical layer parameters (for example in the Direct Sequence Spread Spectrum (DSSS) physical layer of IEEE 802.11 it is equal to $20\mu s$). It accounts for the propagation delay (aAirPropagationTime), for the time needed to switch from the receiving to the transmitting state (aRxTxTurnaroundTime) and the minimum time that the MAC layer needs to assess the medium and detect if it is idle or busy (aCCATime).

4.1.2 Why is renewal theory suitable for single-hop but not for multi-hop networks?

In the Discrete Time Markov Chain (DTMC) of [2], all the transitions—the ones between successive decrements of the backoff counter (with or without freezing of the counter), or the ones between the state where the counter reaches zero and the node transmits, and the next state, where a new initial value of the counter is selected— are considered as equivalent for the derivation of the transmission probability τ. In fact, they are all regarded as having a unit-time duration. This is a model assumption but does not sacrifice the accuracy of the analysis in a fully-connected network. This is so, because later in [2], the saturation throughput is calculated based on an average slot and it is assumed that all the transitions are of a length equal to that average slot.

In [2], the saturation throughput is calculated using renewal theory [68] arguments. In particular, because in a fully-connected network all the nodes view the same channel state, the instants of time right after the end of a transmission are renewal points, so it is sufficient to analyze a single renewal interval (which is the variable slot) for the derivation of the saturation throughput. However, the above argument ceases to be accurate in the presence of hidden terminals due to the consequent desynchronisation of the nodes. For example, the reader can refer to Figure 4.1. This is the 'classic' hidden terminal scenario mentioned in the literature. Nodes A and B

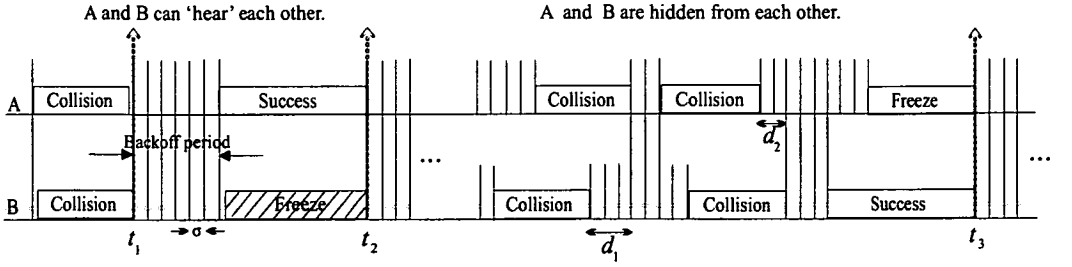


Figure 4.2: *In a fully-connected network (left) the time instants right after every transmission are renewal points. In a hidden terminal scenario (right) this is not true due to the desynchronisation among the hidden stations. In fact, the backoff counters do not start being decremented simultaneously, except for after a successful transmission. For the two collision periods in this example there are d_1 and d_2 slots 'offset' between the time the two stations start decrementing their counters respectively.*

are the two senders which lie outside of the transmission range of each other and want to send data to the receiver R. Let's suppose that node A sends RTS to the common receiver R. Node B will realise that a transmission is occurring not within the duration of one slot σ as in [2], but after the receiver R starts replying to A with a CTS packet (after 18 slots with DSSS and a bandwidth of 1 Mbps). During this time, the hidden sender B can cause a collision as it can not sense A's transmission.

Due to the desynchronisation the hidden transmitters do not start decrementing their backoff counters simultaneously as is usually assumed (see Figure 4.2). In fact, after a collision state there is an '*offset*', i.e. a time difference between the time instants that the stations 'reset' their backoff counters which can be as large as the vulnerable period. Depending on the scheme adopted (Basic Access or RTS/CTS) and the related transmission rate, this *offset* can be quite significant (e.g. in Basic Access, with DataRate of 11 Mbps and packet size of 1024 bytes, it can reach 48 slots). Consequently, the assumption of the existence of renewal points as in [47] and [39] is unrealistic even if it makes the analysis more tractable. In fact, as will be shown in Section 4.3, the evolution of the channel state in the scenario of Figure 4.1 is characterised by a first-order dependence or as it is commonly known, Markov dependence [68].

Parameters	Arithmetic Values
PHY layer	DSSS
Carrier Sensing Range	250 m
Transmission Range	250 m
Routing Protocol	DumbAgent
BasicRate	$\{1, 2\}$ Mbps
DataRate	$\{1, 2, 11\}$ Mbps
PLCP Header	192 bits
aSlotTime	20 μs
RTS	160 bits
CTS,ACK	112 bits
EIFS	SIFS+CTS+DIFS

Table 4.1: *Key NS2 parameters used in simulations.*

4.1.3 Another limitation stemming from the variable slot duration

Several previous papers, including [38] and [39], are based on the variable slot and use the relevant equation for the transmission probability τ and an expression of the form $p_s = \tau(1 - \tau)^{T_v}$ (where T_v is the duration of the vulnerable period normalised to σ) for the probability of success p_s. These modelling techniques have an important limitation. In the conventional models it is implied that if the channel time is discretised into these variable slots, then the channel state would be described as a succession of $1/\tau - 1$ non-transmitting backoff slots followed by a transmitting slot (with collision or success) and the same pattern being repeated. Let's focus on the case when nodes employ a constant contention window (CCW) of size W. It is known that for CCW $\tau = \frac{1}{1+(W-1)/2} = \frac{2}{W+1}$. Standard models do not take into account that quite often there might not be as many backoff slots as implied by the equation for p_s, not even on average. In particular, if the contention window is small, e.g. for $W = 32$, the average backoff duration is 16 slots, for $W = 64$, it is 32 slots; then it could actually be smaller than the vulnerable period, thus making the modelling problematic. Obviously, the impact of this limitation to the accuracy of the results depends on the network configurations, such as the BasicRate, the DataRate, the DATA packet size and finally the contention window used.

To give more insight to this observation, the performance of the model presented in [39] will be investigated regarding the classic hidden terminal scenario for a CCW case and for various values of the parameters mentioned above. The simulation results are obtained from NS2 and the configuration settings, if not mentioned in this section explicitly, are the ones of Table 4.1. In order to evaluate the limitations explained in this section, the relative error of the conventional modelling method compared to the simulation results will be measured, that is given by the

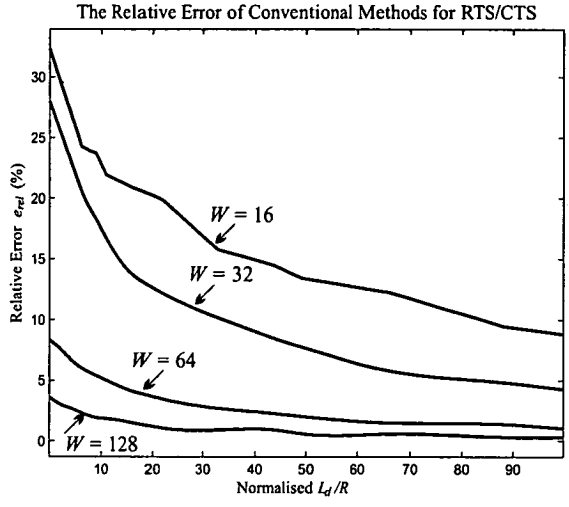


Figure 4.3: *The impact of the assumptions of conventional methods to the models' accuracy for the RTS/CTS scheme.*

following equation:

$$e_{rel} = 100 \frac{|Model - Simulation|}{Simulation} \%, \tag{4.1}$$

where $Model$ represents the value for the throughput in packets/s derived in [39] and $Simulation$ the corresponding value obtained by the average of 10 independent runs of $200s$ each in NS2 for each configuration. The operation mode of IEEE 802.11 DCF used for Figure 4.3 is the RTS/CTS scheme. In Figure 4.3, e_{rel} is plotted against the normalised fraction of the packet size L_d to the DataRate R for different contention window sizes W. This fraction is in fact analog to the time it takes to transmit the DATA packet.

One can observe that the error of the conventional modelling becomes more significant for small contention window sizes. This is because these methods imply that, for a station to be successful, the channel competitor must be silent for more backoff slots than the average backoff period. On the contrary, for larger contention windows, the error of conventional methods tends to be smaller. With the contention window kept constant, it can be observed that the error in the estimation of the throughput increases when the transmission time of the DATA packet (which is proportional to the term L_d/R) decreases. To understand this behaviour, the equation that is used in [39] for the estimation of the saturation throughput—adapted to the CCW

Packet Size	**Contention Window W**				
(in bytes)	**32**	**64**	**128**	**256**	**512**
256	18.4	7.6	2.7	0.6	0.1
512	63.7	23.9	6.8	2.4	1
1024	* >100	79.1	18.3	3.8	1.5
1536	* >100	* >100	45.2	10.1	1.9
2048	* >100	* >100	94.2	18.5	4.3
2294	* >100	* >100	* >100	20.3	4.6

Table 4.2: *The relative error (%) of conventional methods for Basic Access. The asterisked entries are cases where the Basic Access suffers from significantly low throughput. The conventional methods fail to fully capture this behaviour.*

assumption and the notation used in this thesis—is repeated for reasons of completeness:

$$S = \frac{\tau(1-\tau)^{T_v}}{(1-\tau)^2 \cdot \sigma + 2\tau(1-\tau)^{T_v} \cdot T_s + \left(1-(1-\tau)^2 - 2\tau(1-\tau)^{T_v}\right) \cdot [T_c + \sigma T_v/2]} \quad (4.2)$$

The error in the estimation of the probability $p_s = \tau(1-\tau)^{T_v}$ affects both the numerator and the denominator of S, but it becomes more significant when the term T_s is relatively small. Consequently, for long data packets and/or low values of DataRate one expects that e_{rel} will be lower, as one can easily verify from Figure 4.3. Similar trends for e_{rel} were observed when the throughput derived in [38] is compared to the simulation results.

The impact of the conventional models' inaccuracies becomes more severe in the Basic Access scheme, where the vulnerable period is of the order of the DATA packet length and can take quite high values when large packet sizes are combined with small values of DataRate. In Table 4.2 this is demonstrated by choosing some indicative configurations with the best—for the prior methods—value of DataRate, i.e. 11 Mbps. The BasicRate used is 2 Mbps.

As a conclusion, in this section some fundamental assumptions, which prior modelling techniques adopt in order to model the hidden terminal problem, were described. The actual numerical impact of these assumptions was illustrated for a large number of stations' configurations and, thus, their lack of adaptability was depicted for different random-access based protocols which would employ different backoff schemes, transmission rates, packet sizes, etc. Beyond the numerical results and the inability of the prior methods to closely follow the actual protocol performance in some cases, the most important contribution is that the reasons behind the limitations of these assumptions in a hidden node environment were explained in a detailed and comprehensive manner.

The use of a variable slot as the basis of these modelling methods renders them inherently unable to cope with all the possible existing configurations. The aim of this chapter is to introduce a modelling technique which can accurately model different CSMA/CA based schemes with dynamic tuning of the contention window and is highly reliable, independently of the packet size or the transmission rate employed. From the previous analysis it is obvious that Bianchi's notion of the variable slot should not be used in a multi-hop scenario without considering the above limitations. An accurate model of a network that contains hidden terminals must be based not on variable slots, but on a constant—and thus objective for all the nodes—notion of the slot. For this reason a model that views the channel time as a succession of 'small', fixed-length channel slots, σ, is proposed. The aSlotTime of [1], thus, becomes the unit of time in the model presented below.

4.2 The proposed method of modelling—Part I transmission probability and throughput

In this section the proposed method of modelling is described, which can apply to all random access protocols, assuming that the two transmitting nodes are continuously backlogged. However, it can be extended to the non-saturated case. This would require the definition of an additional random variable that would describe the packet arrival process (in particular, the packet interarrival times) which should then be added to the variables that define the backoff process.

4.2.1 From a variable slot to a fixed-length slot

As stated previously, the variable slot can represent different states for the channel. It can represent a successful or a collided transmission, a backoff slot of duration σ or a backoff slot where the counter freezes because the node listens to a successful transmission by another node, not destined for itself (note that in the case of two transmitting stations the freezing of the backoff counter can occur only because of a successful transmission by the other channel competitor). The proposed method will start from the conventional modelling and, stepwise, replace the variable slots with successions of fixed-length slots. As explained in the previous section, in the analysis presented in this chapter the nodes under consideration do not employ the BEB mechanism, but have a CCW. This is a common assumption in the literature [18, 41],

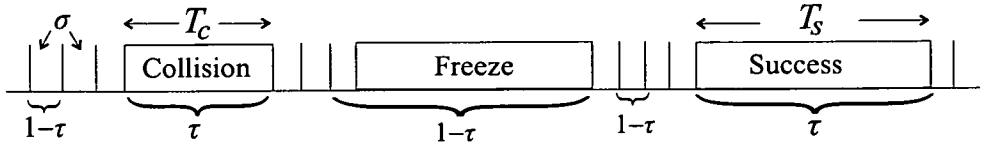


Figure 4.4: *In the conventional modelling τ refers to variable slots.*

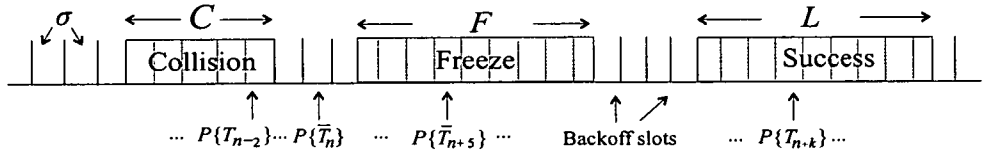


Figure 4.5: *Proposed modelling of the channel around a sender, A or B.*

it gives insight into both the problems of previous techniques and also to the philosophy behind the model, and, finally, it is not unrealistic, since in variable traffic and mobility conditions the nodes can dynamically select their maximum contention window size. Moreover, the proposed method is easily extendable to the BEB mechanism of IEEE 802.11 or other backoff schemes as will be shown in Chapter 5. For CCW the value of the backoff counter (modelled as a discrete random variable (DRV) BC) is uniformly distributed in the interval $[0, W-1]$.

The transmission probability τ, defined in [2], refers to a variable slot. As shown in Figure 4.4, during the time a node is in backoff, the same probability, $1-\tau$, describes a backoff slot σ, and also a longer backoff slot where the counter freezes. Moreover, the probability τ describes a transmitting slot, whether it represents a successful or a collided transmission. The ultimate aim of the proposed model is to substitute these variable slots with fixed-length channel slots. As a consequence, τ will be replaced by a new transmission probability $P\{T_n\}$, defined as the probability that the station is transmitting at the fixed-length channel slot n. To this end, the state of the channel around a node is modelled as successions of successful periods (whose length L is considered an integer multiple of σ), collision periods (length C), periods when the node is in backoff but its counter is frozen (F) and finally periods of backoff slots. The new model is illustrated in Figure 4.5.

As defined previously, T_n denotes the event that the node is transmitting—either successfully or facing a collision—in slot n. From the rule of total probability [72], $P\{T_n\}$ can be written as

$$P\{T_n\} = P\{T_n|T_{n-1}\}\,P\{T_{n-1}\} + P\{T_n|\overline{T}_{n-1}\}\,P\{\overline{T}_{n-1}\}\,. \tag{4.3}$$

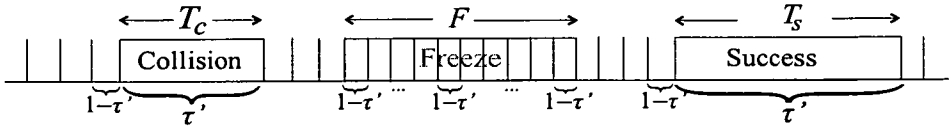


Figure 4.6: *The channel model after the discretisation of the backoff periods.*

Note that $\overline{T}_n$ denotes the complementary event that the node is not transmitting in slot n. In order to find the conditional probability $P\{T_n|T_{n-1}\}$ consider Figure 4.5. If the previous slot was a transmitting slot (belonging to a successful or to a collision period) then the next one will also be a transmitting slot. The only exception to this rule is the last slot of a succession of a successful period (length L) or a collision period (length C), which are followed by a non-transmitting, backoff slot, if a non-zero backoff counter has been selected. A non-zero backoff counter is selected with probability $1 - \frac{1}{W} = \frac{W-1}{W}$. Moreover, the last slot of a succession of transmitting slots is selected with probability $\frac{1}{(1-p)L+pC}$. These two events are independent, so the probability of the event intersection is obtained as the product of the two probabilities. Consequently, one has

$$P\{T_n|T_{n-1}\} = 1 - \frac{W-1}{W}\frac{1}{(1-p)L+pC} = 1 - \frac{W-1}{W}Q, \tag{4.4}$$

where p is the conditional collision probability defined in [2] and Q denotes the term $\frac{1}{(1-p)L+pC}$.

The other conditional probability in (4.3), $P\{T_n|\overline{T}_{n-1}\}$, represents the transitions from a non-transmitting slot to a transmitting one. Due to the fact that one is only interested in the *transitions to transmitting slots*, the transmissions can be still viewed as one large slot as explained in Figure 4.6. However, the preceding non-transmitting slots should not be viewed as in the conventional modelling, but discretised into fixed-length slots of duration σ. In order to achieve this, the classic transmission probability τ of [2] must change to a new, intermediate transmission probability τ', that has discretised the non-transmitting slots (including the freezing periods), but still considers the transmission slots as variable (see Figure 4.6). Thus, the conditional probability $P\{T_n|\overline{T}_{n-1}\}$ is equivalent to the conditional probability $\tau'|(1-\tau')$. Note that the notation $\tau'|(1-\tau')$ is used to denote the conditional probability that the station transmits, given that it did not transmit in the previous slot referring to the time-scale that corresponds to Figure 4.6.

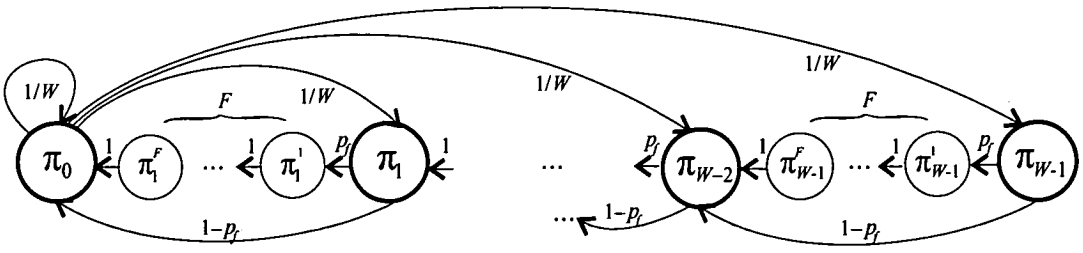


Figure 4.7: *Incorporation of the freezing of the backoff counter in the Markov Chain of [2].*

Let us suppose for the moment that τ' is known. From Bayes' Theorem one can write:

$$\tau'|(1-\tau') = 1-\tau'|\tau' \cdot \frac{\tau'}{1-\tau'} \tag{4.5}$$

As one can observe in Figure 4.6, we have $1-\tau'|\tau' = 1-\frac{1}{W}$ because the transmission occurs in one slot, which is always preceded by a backoff slot, apart from the case that a backoff counter of value 0 has been chosen (in this special case there are two back-to-back transmitting slots). Thus:

$$P\left\{T_n|\overline{T}_{n-1}\right\} = \tau'|(1-\tau') = \frac{W-1}{W} \cdot \frac{\tau'}{1-\tau'}. \tag{4.6}$$

To calculate τ', F is defined as the length of the freezing period in fixed-length slots (an integer multiple of σ) and π_m, $m \in [0, W-1]$ as the steady-state probability that the counter is equal to m, non-freezing. The corresponding states, represented as large circles in Figure 4.7, are the states that a MC, like the one in [2], consists of. In addition, π_m^i is defined as the steady-state probability that the counter of value m, $m \in [1, W-1]$, is frozen ($i \in [1, F]$ because there are F such states for each m). The corresponding states are represented as smaller circles in Figure 4.7. Let also p_f be the probability that a node in a state m listens to another transmission and freezes its counter, i.e. the transition probability from a state m to the successive m^1. The transition probabilities between two states m^i and m^{i+1}, $i \in [1, F-1]$ are equal to 1. From the analysis of the MC—which is described in detail in the Appendix B—one can derive the following steady state probabilities:

$$\begin{aligned} \pi_m &= \tfrac{W-m}{W}\pi_0 \quad &, m \in [1, W-1] \\ \pi_m^i &= p_f \pi_m \quad &, m \in [1, W-1], i \in [1, F] \end{aligned} \tag{4.7}$$

From the MC regularities one gets

$$\tau' = \pi_0 = \frac{2}{2 + (W-1)\left[1 + p_f F\right]}. \tag{4.8}$$

In order to derive an expression for the probability p_f one must think of it as the fraction of the number of 'long' freezing slots over the number of all the backoff slots. The number of freezing slots is equal to the number of times that the station froze its counter because of a successful transmission by its competitor (with probability $1 - p$, conditional on the station transmitting). The backoff slots are given by the product of the average duration $E[BC]$ of a backoff interval, multiplied by the number of times that the station goes into backoff (or, equivalently, the number of times that the station transmitted, since the backoff intervals and the transmissions always succeed one another). Thus:

$$p_f = \frac{1-p}{E[BC]} = \frac{2(1-p)}{W-1}. \tag{4.9}$$

Substituting (4.4) and (4.6) into (4.3), one gets:

$$\begin{aligned} P\left\{T_n\right\} &= \left(1 - \tfrac{W-1}{W} Q\right) P\left\{T_{n-1}\right\} + \tfrac{W-1}{W} \tfrac{\tau'}{1-\tau'} P\left\{\overline{T}_{n-1}\right\}, \\ P\left\{T_n\right\} &= \tfrac{\tau'}{\tau' + Q(1-\tau')}, \end{aligned} \tag{4.10}$$

where τ' is given by (4.8). Note that due to the steady-state assumption $P\left\{T_n\right\} = P\left\{T_{n-1}\right\} = 1 - P\left\{\overline{T}_{n-1}\right\}$ was used in order to derive (4.10). The result of the discretisation of the channel time is depicted in Figure 4.5.

4.2.2 Derivation of throughput

The throughput S (in packets per slot) is defined as the number of packets transmitted during a specific period of time divided by the duration of that period, or alternatively as the fraction of one packet over the time (expressed in slots) elapsed between the start of the backoff process for the packet and the completion of its successful transmission. With reference to Figure 4.5, this time includes periods of idle slots (is), where the counter is either frozen or not, periods of collision slots (cs) and finally the period of successful slots (ss). In order to 'pass' from the

time periods to the corresponding probabilities which were calculated, we have:

$$S = \frac{1}{is + ss + cs} = \frac{1}{ss} \cdot \frac{ss}{ss + cs} \cdot \frac{ss + cs}{is + ss + cs}. \tag{4.11}$$

The last term represents the fraction of the transmitting slots over all the slots, so:

$$\frac{ss + cs}{is + ss + cs} = P\{T_n\}. \tag{4.12}$$

Furthermore, the middle term is the ratio of successful slots to all the transmitting slots (successful and collided), so it is equal to

$$\frac{ss}{ss + cs} = \frac{L}{L + \frac{p}{1-p}C}. \tag{4.13}$$

Finally $1/ss = 1/L$. Thus, the throughput S can be written as:

$$S = P\{T_n\} \cdot \frac{(1-p)L}{(1-p)L + pC} \cdot \frac{1}{L}. \tag{4.14}$$

Consequently, the only unknown parameter that still remains to be derived so that the throughput is calculated is the conditional collision probability p.

4.3 The proposed method of modelling—Part II conditional collision probability

As it was also mentioned in [39], the classic hidden terminal scenario in Figure 4.1 is very complex due to the high correlation between the two transmitters. In the beginning of the section, the analysis is focused on the RTS/CTS scheme. Towards the end of the section, the Basic Access scheme will also be addressed; in fact, the analysis followed for both access schemes is similar, with some minor differences that will be described in Section 4.3.3.

In order for the packet sent by A to be successfully delivered to R, B must defer its transmission during a vulnerable period T_v. It is argued that T_v is not equal to RTS as it is commonly believed by many researchers [38, 47]. In these papers, the authors assume that the stations enter the contention phase synchronised, as explained in Section 4.1. Under this assumption, the period during which a collision can occur in the RTS/CTS (Basic Access) scheme is exactly

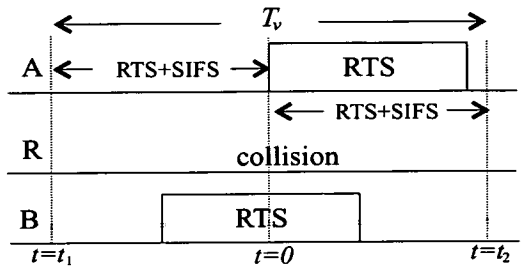


Figure 4.8: *The vulnerable period T_v for the RTS/CTS mode of IEEE 802.11 DCF.*

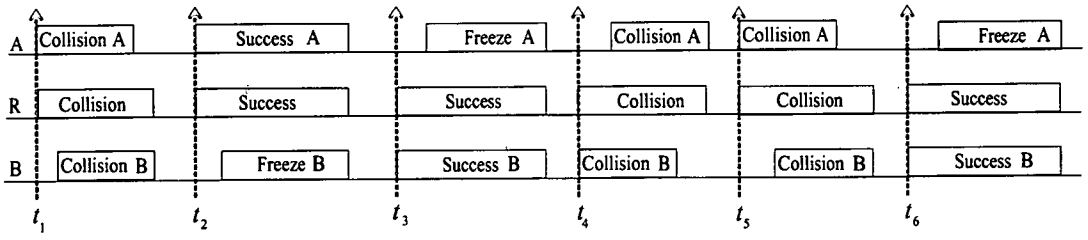


Figure 4.9: *The relation between p and p' and the time-index space of the relative Discrete Time Markov Chain.*

the duration of the RTS (DATA) packet. On the contrary, the actual vulnerable period is

$$T_v = \begin{cases} 2 \cdot (\text{RTS} + \text{SIFS}) & \text{for RTS/CTS} \\ 2 \cdot \text{DATA} & \text{for Basic Access.} \end{cases}$$

With reference to Figure 4.8, if one considers (without loss of generality) $t = 0$ the time when A starts sending RTS to R and no synchronisation of the senders is assumed, one observes that at any point in the interval $[t_1, t_2]$ an RTS transmission from B will cause collision at R.[2]

In order to facilitate the analysis, an intermediate step is used, that can incorporate and accurately describe the joint behaviour of the senders. To this end, the channel around the receiver R is modelled. R observes a collision when (both) the transmitters attempt to transmit and fail and, conversely, observes a successful transmission when one of them is successful (Figure 4.9). In fact, the number of collisions viewed by the receiver is equal to the number of

[2]During SIFS, although the RTS packet is successfully received by R, the node still senses the channel, so if it hears another transmission, it will not reply with a CTS [1]. Of course this argument assumes that the transceiver is 'almost' perfect, i.e. that aCCATime+aRxTxTurnaroundTime is very small, and at least shorter than SIFS. On the contrary, this is not the case for Basic Access, where ACK is always transmitted upon correct reception of the DATA packet, independently of the busy or idle state of the medium [1].

collisions that each transmitter faces, whereas the successful transmissions viewed by the receiver are the sum of the successful transmissions of the two senders. Let N_{ca} (N_{sa}) denote the number of collisions (successes) occurring when one observes the channel around A in steady-state. Equivalently, for B, N_{cb} and N_{sb} are defined. Also $N_{ca} = N_{cb}$ and in the steady-state $N_{sa} = N_{sb}$, since in this scenario the two transmitters share an equivalent part of the bandwidth in the long-term [39]. If the probability p' that R sees a collision (conditioned that either A or B attempted transmission) is known, then p is derived (e.g. for transmitter A) as follows:

$$\begin{aligned} p &= \frac{N_{ca}}{N_{ca}+N_{sa}} = \frac{N_{ca}}{N_{ca}+N_{sa}+N_{sb}} \cdot \frac{N_{ca}+N_{sa}+N_{sb}}{N_{ca}+N_{sa}} \\ &= p' \cdot [1+(1-p)], \end{aligned} \tag{4.15}$$

which yields

$$p = \frac{2p'}{1+p'}. \tag{4.16}$$

Thus, the problem is transformed, so that it is now sufficient to find the probability p' that the receiver observes a collision, conditioned on any of the senders attempting transmission.

4.3.1 The first-order dependence of two successive channel states

At this point, the reasons for choosing Markov chain modelling to describe the channel around the receiver will be explained. It was shown previously that the use of renewal theory is not appropriate for a hidden terminal environment. The assumption other researchers have made, i.e. that the state of the system is independent of the previous state, will prove to not be the case in the analysis that follows. In fact, the probability that in the next state the receiver will observe success (or collision) is quite different when the previous state was collision to when it was success. Thus, a first-order dependence [68] describes the channel, in other words the current state depends only on the previous state and the transition probabilities.

To this end a DTMC will be used that has only two states, success and collision. The use of discrete time implies that we are not interested in the exact time needed to switch from one state to the other and are only counting 'transition steps'. This is because the probability p' that the receiver will observe a collision, conditioned on one of the senders A, B transmitting, is of interest, so only the instances of time when the receiver starts sensing a (failed or successful) transmission are important. These instances are described in Figure 4.9 as red dashed arrows, and they are exactly the time instants that the considered DTMC is built upon. If

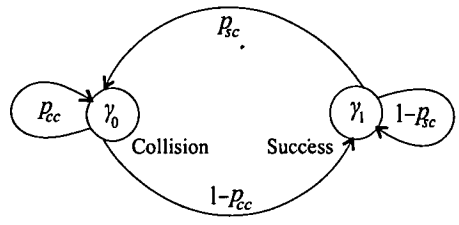


Figure 4.10: *The transmitting states of the channel around R.*

$\breve{T} = \{t_1, t_2, t_3, ...\}$ is the time-index space of the DTMC [68], then the red dashed arrows point to its elements.

Consequently, the channel around the common receiver R can be characterised by a MC (shown in Figure 4.10) that has only two states, collision (both A and B face collision) and success (one of A or B is successful). Let's define the four transition probabilities, using p_{cc} the probability that R will see a collision conditioned that it saw a collision in the previous channel state, and p_{sc} the probability that R will see a collision conditioned that it saw a successful transmission before. Then, if γ_i, $i = \{0, 1\}$, is the stationary distribution of the MC of Figure 4.10, then p' is written from the MC as

$$p' = \gamma_0 = \frac{p_{sc}}{1 - p_{cc} + p_{sc}}. \tag{4.17}$$

The problem of deriving expressions for the transition probabilities of Figure 4.10 will be solved based on an iterative approach. In order to explain the motivation behind it, it is essential to stress that the nature of the system is such that every state depends on the previous one, and there is not *de facto* information upon which one can start the analysis. As will become obvious in the paragraphs following, there are some parameters (actually they are distributions of some random variables which will shortly be defined) necessary for the derivation of the transition probabilities, which are derived iteratively as follows: the derivation starts with an arbitrary initial guess for the unknown distributions, makes a first estimate of the transition probabilities and of the distributions mentioned before, and repeats the same procedure until the iterative process converges.

4.3.2 The derivation of the transition probabilities p_{cc}, p_{sc}—Iterative method

A key and, to the best of the author's knowledge, novel point in this chapter is the use of DRVs for the representation of the various quantities which affect and, ultimately, determine

the outcome of the contention between the transmitting stations. These quantities are the values of the backoff counters of the two stations, the '*offset*' between the two transmissions as was described in Section 4.1.2, and the remaining backoff slots after a node has frozen its counter. These terms will be explained in the following paragraphs.

The analysis will initially focus on the cases where the system state changes from collision (state $n-1$, hereafter referred to as s_{n-1}) into either collision or success (state s_n). The corresponding figures are Figure 4.11(a,b). Without loss of generality, it is assumed that A starts transmitting RTS at time $t_n = 0$ in s_n, and at t_{n-1} in s_{n-1}. Because there was collision in s_{n-1}, B could not have started transmitting RTS more than $c = (\text{RTS} + \text{SIFS})/\sigma$ slots before or after t_{n-1}. With reference to Figure 4.11(a), let's define as X_A (X_B) the DRV describing the initial backoff counter value of A (B), that follows a uniform distribution in $[0, W-1]$. Moreover, Y denotes a DRV representing the (signed) time difference (expressed in slots) of the starting times of the RTS transmissions conditioned that collision follows immediately afterwards[3]. The notation Y_n is used to refer to state s_n. Note that, in steady-state conditions, the Probability Mass Function (PMF) of Y_n should be the same for all n. Y_n is drawn from the interval $[-c, +c]$ (the vulnerable period) but its exact distribution is still unknown. The condition for collision in s_n is $c + C + X_A - c \leq c + Y_{n-1} + C + X_B \leq c + C + X_A + c$, thus

$$p_{cc} = P\{-c \leq Y_{n-1} + X_B - X_A \leq c\}. \tag{4.18}$$

A conditional distribution of Y_n can also be obtained (denoted as Y_n^c) given that a collision precedes (remember that in this case both s_{n-1} and s_n are collision states). This is known as a truncated distribution in the literature [73]. In fact, Y_n^c is given by $Y_n^c = Y_{n-1} + X_B - X_A | -c \leq Y_{n-1} + X_B - X_A \leq c$.

On the contrary, the condition for A to transmit successfully in s_n is that the initial value of the backoff counter of B X_B is large enough, so that A's RTS is received correctly by R, i.e. $X_B > X_A + c - Y_{n-1}$ (Figure 4.11(b)). Note that, in this case, when B senses the CTS reply sent by R, it freezes its backoff counter. After the end of s_n (t_e) both transmitters start decrementing their counters simultaneously. However, although A starts a new backoff period, and, consequently, its initial counter value can be described by the DRV X_A, B continues decrementing its backoff counter after several backoff slots have already elapsed. Z is defined to be the DRV that shows

[3] In fact, the DRV Y quantifies the 'offset' that was described in Section 4.1.2.

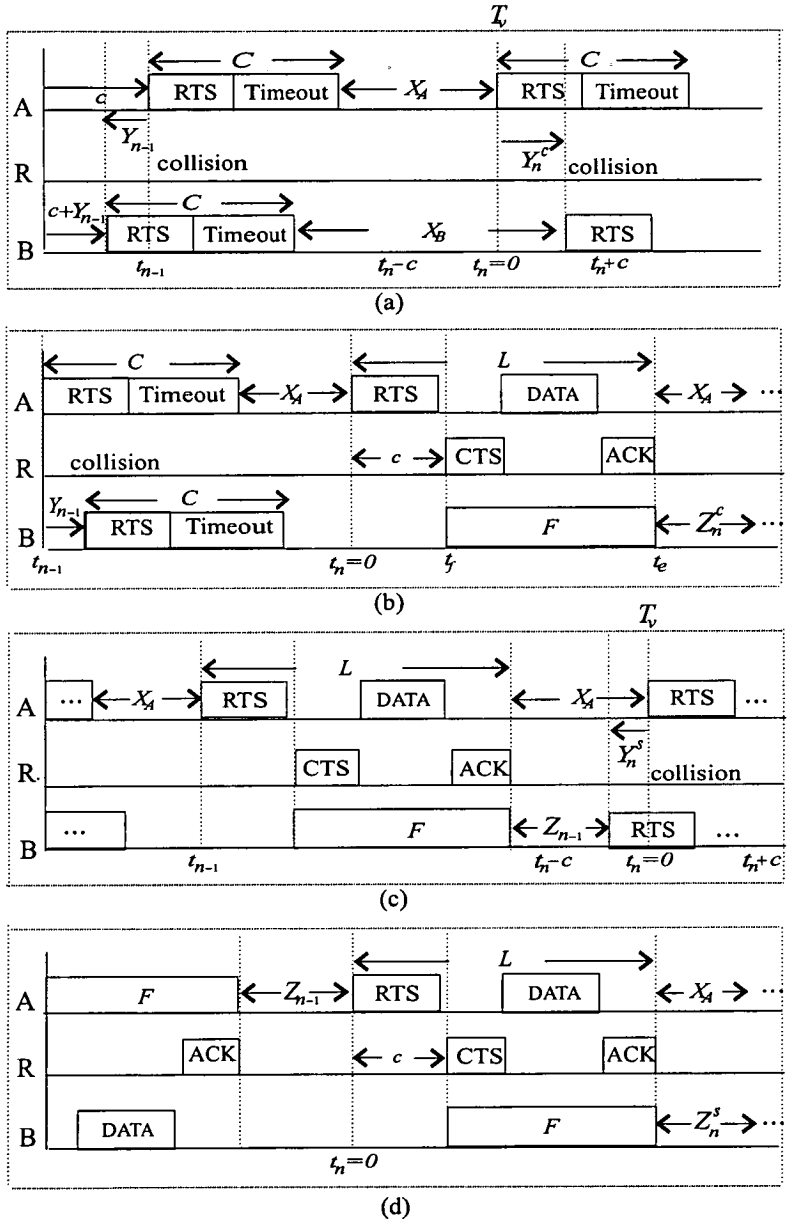


Figure 4.11: *Schematic analysis of the transitions between system transmitting states.*

$p'(0) \leftarrow$ HUGE

$Y_{n-1}(0) \leftarrow$ arbitrary distribution

$Z_{n-1}(0) \leftarrow$ arbitrary distribution

repeat

$i \leftarrow i+1$

compute $p_{cc}(i)$ from (4.18)

compute $p_{sc}(i)$ from (4.19)

estimate $Y_n(i)$ from (A.11)

estimate $Z_n(i)$ from (A.14)

$Y_n(i\text{-}1) \leftarrow Y_n(i)$

$Z_n(i\text{-}1) \leftarrow Z_n(i)$

estimate $p'(i)$ from (4.17)

until $|p'(i) - p'(i\text{-}1)| < \varepsilon$

Figure 4.12: *The pseudocode of the iterative process. HUGE and ϵ are arbitrarily large and small (respectively) numerical values. The current iteration step is denoted as i.*

the distribution of remaining backoff slots of a freezing node, after its competitor (in this case A) transmitted successfully. In this scenario (collision followed by success), a conditional distribution of Z_n given that collision precedes, say Z_n^c, is provided by the following truncated distribution with PMF $Z_n^c = X_B - (X_A + c - Y_{n-1}) \,|X_B > X_A + c - Y_{n-1}$. Of course, if B was the successful transmitter (this occurs when $X_A > X_B + c - Y_{n-1}$), Z_n^c would be $Z_n^c = X_A - (X_B + c - Y_{n-1}) \,|X_A > X_B + c - Y_{n-1}$.

The cases where the system state switches from success to either collision or success will now be examined (Figure 4.11(c,d)). The arguments followed will not be repeated in all their details, because the analysis is very similar to the previous cases. For a collision to occur in s_n the following expression must be true: $-c \leq Z_{n-1} - X_A \leq c$. Thus, the probability p_{sc} is given by

$$p_{sc} = P\{-c \leq Z_{n-1} - X_A \leq c\}. \tag{4.19}$$

A conditional distribution is also obtained for Y_n conditioned on a success preceding, say Y_n^s, that has the following PMF $Y_n^s = Z_{n-1} - X_A| - c \leq Z_{n-1} - X_A \leq c$. Finally, when

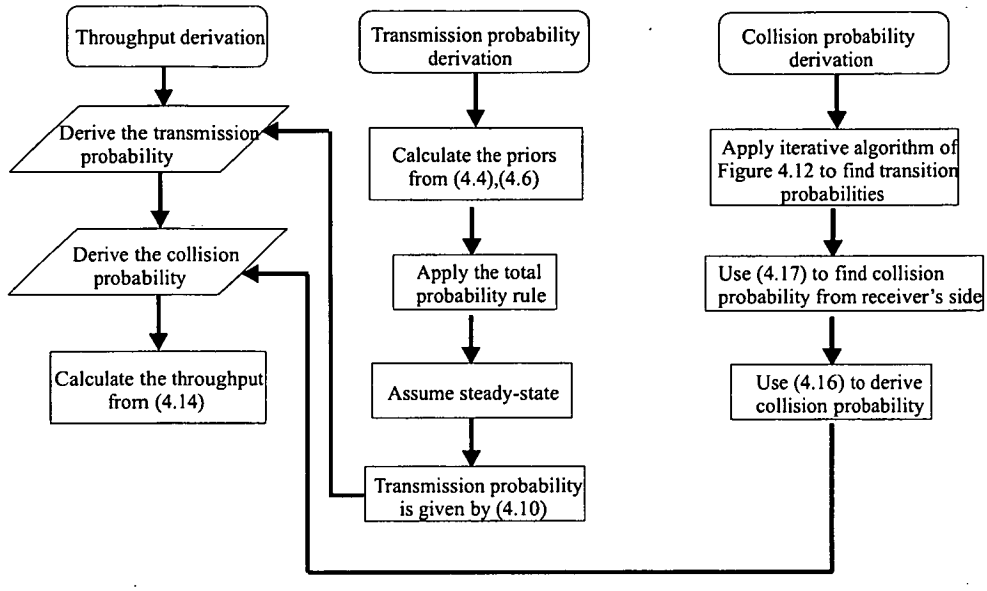


Figure 4.13: *This flow chart provides an overview of the analytical method followed, in order to derive the saturation throughput of a hidden terminal.*

success is followed by success (Figure 4.11(d)), a conditional distribution for Z_n, say Z_n^s, given that success precedes, can be derived. The latter is drawn from the truncated distribution $Z_n^s = X_B - (Z_{n-1} + c) \,|X_B > Z_{n-1} + c$. In the latter case, it was assumed that, in s_{n-1}, B was successful and, in s_n, A was successful (symbolically BA). Very similar expressions can be derived for the three other possible combinations AA, AB, BB.

Consequently, the only unknowns that have to be derived in order to find the transition probabilities of Figure 4.10, are the two distributions Y_{n-1} and Z_{n-1} (or equivalently Y_n and Z_n, since the system is supposed to be in the steady-state). It is evident from the previous analysis that an iterative method is suitable. The relevant pseudocode is presented in Figure 4.12. In particular, the procedure for the derivation of the two distributions is as follows: An initial guess for Y_{n-1}, Z_{n-1} is made, e.g. that they follow uniform distributions in some arbitrary interval. Using (4.18) and (4.19), some first estimates for the two transition probabilities p_{cc}, p_{sc} are obtained. In addition, from the analysis explained above, there are initial estimates for Y_n^c (Y_n^s), which is the distribution of Y_n given that collision (success) preceded. Using posterior probabilities and with reference to the MC of Figure 4.10, the PMF of Y_n is derived as a combination of Y_n^c and Y_n^s with weights p_{cc} and $1 - p_{cc}$ respectively [72] (see Appendix A). Similarly, for the PMF of Z_n, Z_n^c and Z_n^s are used with corresponding weights p_{sc} and $1 - p_{sc}$ respectively.

The procedure is repeated based on the new, more accurate than before, estimates for Y_n and Z_n.

Independently of the initial guess, the iteration converges in a few steps, a fact that makes the method computationally efficient. Thus, knowing the transition probabilities p_{cc}, p_{sc}, the value of p' is evaluated from (4.17), and substituting it to (4.16), the conditional collision probability p is finally obtained. For a more detailed description of the steps followed for the derivation of p the reader is referred to the Appendix A. A flow chart that summarises the method followed to derive the basic channel contention probabilities and, ultimately, the saturation throughput, is presented in Figure 4.13.

4.3.3 The Basic Access scheme

The method followed for the calculation of p in the Basic Access scheme is, in general, the same as was described previously, with the most obvious difference being the duration of the vulnerable period, which in this case is $T_v = 2 \cdot \text{DATA}$, and is, thus, dependent on the data packet length and the parameter DataRate [1]. There is, however, a point that needs to be taken into consideration in order to derive an accurate expression for p (note that for the first part of the method, presented in Section 4.2, the analysis is exactly the same for both modes of DCF).

As one can observe in Figure 4.14, in the Basic Access scheme, where the vulnerable period can be quite large, it is sometimes observed that during the same DATA packet transmission from, say, transmitter B, station A transmits two DATA packets and, consequently, faces collision twice. This phenomenon will be described with the term '*double collision*'. This phenomenon was not observed for the RTS/CTS scheme due to the small duration of the RTS transmission (with the values of BasicRate used in [1]) compared to the duration of the Timeout followed by the backoff interval. In general, however, a situation like the one in Figure 4.14 can describe a case of arbitrary duration of the first packet of the handshake (RTS or DATA).

'*Double collision*' affects the argument that was used in the beginning of the Section 4.3, i.e. that the number of collisions faced by each of the two senders is equal to the number of collisions observed by the receiver R. In the general case, and if p_{dc} denotes the probability that '*double collision*' has occurred, given that collision has occurred, one gets the revised expression for equation (4.16)

$$p = \frac{(2 - p_{dc})p'}{1 + (1 - p_{dc})p'}. \tag{4.20}$$

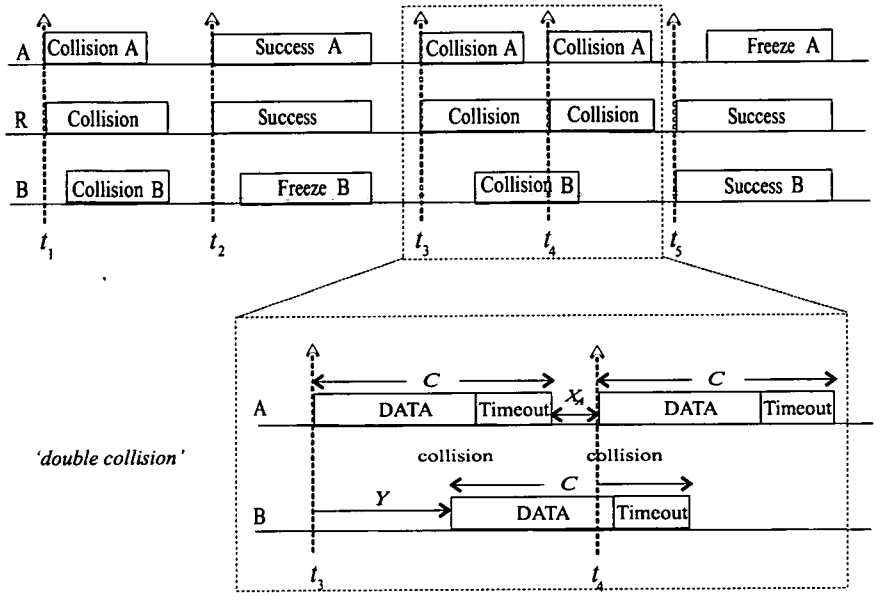


Figure 4.14: *With the term 'double collision' one refers to the scenario where two consecutive packets of the same transmitter (station A) collide with the same packet of the other transmitter (B).*

using similar arguments as for the derivation of (4.15). Furthermore, regarding the iterative method, some minor modifications are necessary in order to distinguish between the case of a normal collision and a '*double collision*'. However, the basic process is the same as already described, and only the expression for p_{dc} will be explicitly mentioned, which is

$$p_{dc} = P\{|Y| \geq X_i + T_{out}\}, \tag{4.21}$$

where $i \in \{A, B\}$ describes the associated transmitter and T_{out} is the duration of Timeout and DIFS normalised to the fixed-length slot σ.

4.3.4 The derivation of the transition probabilities p_{cc}, p_{sc}—Approximation method

In this section an alternative method for the derivation of the transition probabilities of the MC of Figure 4.10 is presented, that does not rely on iteration, and attempts to approximate the basic unknown distributions with some well known ones. Apart from this fundamental

difference, many common features with the Section 4.3.2 exist, as it will be evident in what follows. The RTS/CTS scheme will be used as a case study, although similar argumentation can be developed for any other CSMA/CA type access scheme.

The first probability to be calculated will be p_{cc} (Figure 4.15(a)), where two successive collision periods occur. Without loss of generality it is assumed that A starts transmitting RTS at time $t_n = 0$ in the current state, and at t_{n-1} in the previous state. Because there was collision in the previous state, B could not have started transmitting RTS more than $c = (\text{RTS} + \text{SIFS})/\sigma$ slots before t_{n-1}. As a result, all the possible collision–collision realisations are taken into account by considering the time observation window $[t_{n-1} - c, t_n + c]$, as it is shown in Figure 4.15(a)[4].

X_A (X_B) are defined as previously, i.e. as uniform distributions in [0,W − 1]. Y is drawn from the interval $[-c, +c]$, as the previous channel state was a collision. In s_n, B and A start sending RTS $c + Y + C + X_B$ and $c + C + X_A$ slots after the start of the observation window correspondingly. Consequently, the condition for collision in the current state is $C + X_A \leq c + Y + C + X_B \leq C + X_A + T_v$, thus

$$p_{cc} = P\{X_A \leq Y + X_B + c \leq X_A + T_v\}. \tag{4.22}$$

For the derivation of p_{sc} the reader is referred to Figure 4.15(b). In this case, the channel around R is characterised by a successful period (s_{n-1}), followed by a collision period (s_n). In particular, in s_{n-1}, it is assumed, without loss of generality, that A successfully transmits a data packet, while B, after listening to the CTS reply sent by R, freezes its backoff counter. In order to find the probability of collision in s_n, the same method as in the previous case is followed. In s_n, for collision to occur, the starting time of the RTS transmission of B must be inside the interval $[t_n - c, t_n + c]$. As one can observe in Figure 4.15(b), the transmitters start decrementing their backoff counters after the end of s_{n-1} (t_e) simultaneously. However, although A starts a new backoff period and, consequently, its initial counter value can be described by the DRV X_A, B continues decrementing its backoff counter after K backoff slots have already elapsed[5]. In particular, the backoff slots that have elapsed from the start of B's previous backoff period, until B heard CTS from the receiver R—at the time instant t_f—and froze its backoff counter,

[4]Note that the intervals SIFS and DIFS are omitted from Figures 4.15 and 4.16.

[5]Note that, for this section only, Z, which was defined earlier in this chapter, is substituted by the distribution K, which represents the backoff slots that have already elapsed before freezing, and not the remaining backoff slots as Z did.

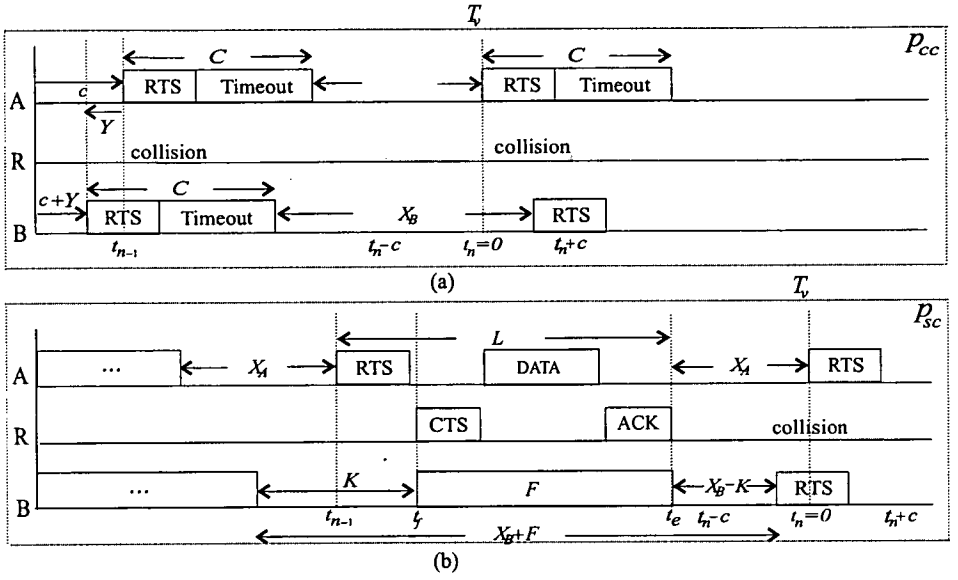


Figure 4.15: *Schematic analysis of the transition probabilities* p_{cc}, p_{sc}.

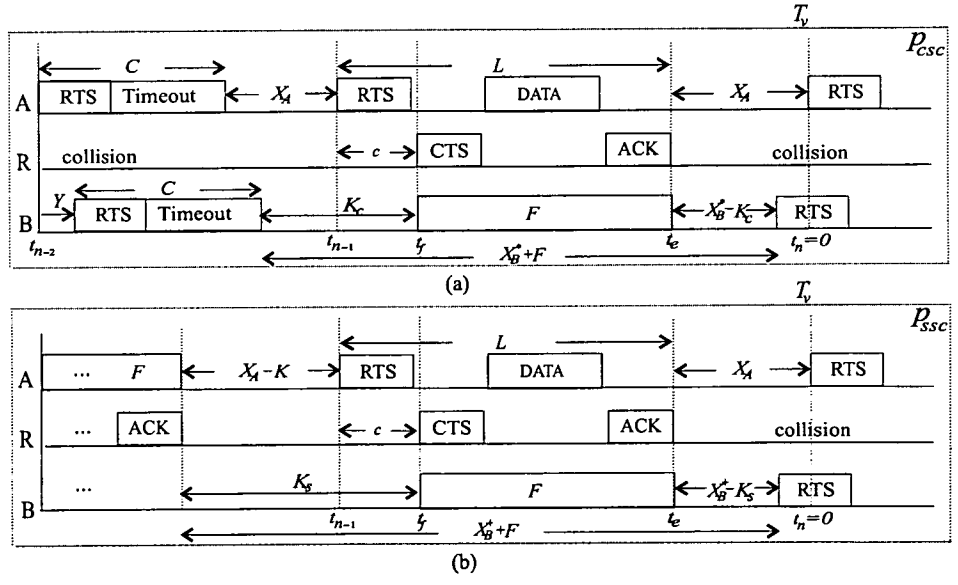


Figure 4.16: *Schematic analysis of the transition probabilities* p_{csc}, p_{ssc}.

are defined with the DRV K (its initial backoff counter value can be described with the DRV X_B). Thus, in order to find p_{sc}, it is necessary to know K. As one observes in Figure 4.15(b), K is dependent on the previous state of the channel, i.e. if there was a collision or a success prior to the success (s_{n-1}).

As a consequence, one needs to break the analysis into two cases, the case where there are successively collision−success−collision in the channel (Figure 4.16(a)) and the case where there are success−success−collision (Figure 4.16(b)). Note that the probability that, after success in s_{n-1}, collision follows in s_n, conditioned that in s_{n-2} collision (success) had occurred, is denoted as p_{csc} (p_{ssc}). Consequently, the probability p_{sc} can be written as $p_{sc} = a \cdot p_{csc} + \beta \cdot p_{ssc}$. The weight a (β) is the probability that collision (success) occurred at s_{n-2} before the success (at s_{n-1}). They can be derived as *a posteriori* probabilities according to Bayes' Theorem from the MC of Figure 4.10. In particular, one gets $a = (1 - p_{cc}) \cdot \gamma_0/\gamma_1$ and $\beta = 1 - p_{sc}$. After some algebra one obtains

$$p_{sc} = \frac{p_{ssc}}{1 - p_{csc} + p_{ssc}}, \tag{4.23}$$

where p_{csc} and p_{ssc} are yet to be calculated.

In particular, regarding the first one, as one observes in Figure 4.16(a), station B listens to the CTS packet that R sends to A $c = (\text{RTS} + \text{SIFS})/\sigma$ slots after t_{n-1}. As a result, if X_B is the initial backoff counter of B after s_{n-2}, then, in order for A to be successful, the counter of B must not expire before the time instant t_f. In other words, the required condition for success of A in s_{n-1} is $Y + C + X_B > C + X_A + c$, which yields $X_B > X_A + c - Y$. The truncated DRV $X_B | X_B > X_A + c - Y$ [73] is denoted as X_B^*. Also, the backoff slots K_c that have elapsed before the counter of B freezes are $K_c = X_A + c - Y$. From these conditions, one can calculate

$$p_{csc} = P\left\{-c \leq X_A - (X_B^* - K_c) \leq c\right\}. \tag{4.24}$$

Similarly (see Figure 4.16(b)) one has for the probability p_{ssc}

$$p_{ssc} = P\left\{-c \leq X_A - (X_B^+ - K_s) \leq c\right\}, \tag{4.25}$$

where X_B^+ is the truncated DRV $X_B | X_B > (X_A - K) + c$. Substituting (4.24) and (4.25) into (4.23), the probability p_{sc} can be obtained. Finally, with this expression and (4.22), one calculates p' and p, as was done in the Section 4.3.1.

Observing the equations (4.22),(4.24) and (4.25) it is obvious that there are still some unknown distributions -similar to the ones that in the Section 4.3.1 were derived via the iterative process. In fact, as one can observe, these distributions, Y and K, do not appear in the above equations by themselves, but as a sum or a difference with a uniform distribution. This is an interesting feature, that one can take advantage of, and approximate the latter with some known distributions.

In particular, let's focus on the derivation of p_{cc} from equation (4.22). It is known that if X is a uniform DRV defined in the interval $[0, W-1]$ and k is a constant, the PMF P_{X+k} is

$$P_{X+k}(x) = \frac{1}{W}, x \in [k, k+W-1]. \tag{4.26}$$

As was mentioned previously, the DRVs X_A, X_B follow a uniform distribution in the interval $[0, W-1]$, so their PMF can be derived according to (4.26). If the channel was monopolised by successions of collision periods, then the starting times of the corresponding transmissions of the two senders A,B would be uniformly distributed in the interval $[0, T_v]$. As a consequence, Y, as the difference between two uniform distributions would follow a triangular distribution in the interval $[-c, +c]$. This is confirmed by Figure 4.17, where both the triangular distribution and the one obtained by simulation results using $W = 32$ are plotted. In the simulations the difference of the starting times of the transmissions of A and B when collision follows is measured. One can observe a slight deviation of the latter from the theoretical distribution. This is due to the fact that, in simulations, there are not always successions of collision periods, but successful periods may also be interposed. However, in this case, the assumption of the channel being monopolised by successions of collision periods is close to the reality, because of the fact that the contention window is relatively small and the collisions are very frequent.

For larger values of W, the collision probability, as one will notice in the next section, drops, and the assumption of a triangular distributed Y is no longer valid. However, as in (4.22) one is not interested in the PMF of Y alone, but in that of $Y + X_B$; one gets:

$$P_{Y+X_B}(x) = \begin{cases} \frac{1}{W} \sum_{i=-c}^{x} P_Y(i) & x \in [-c, c-1] \\ \frac{1}{W} & x \in [c, W-c-1] \\ \frac{1}{W} \sum_{i=x-W+1}^{c} P_Y(i) & x \in [W-c, W+c-1] \end{cases} \tag{4.27}$$

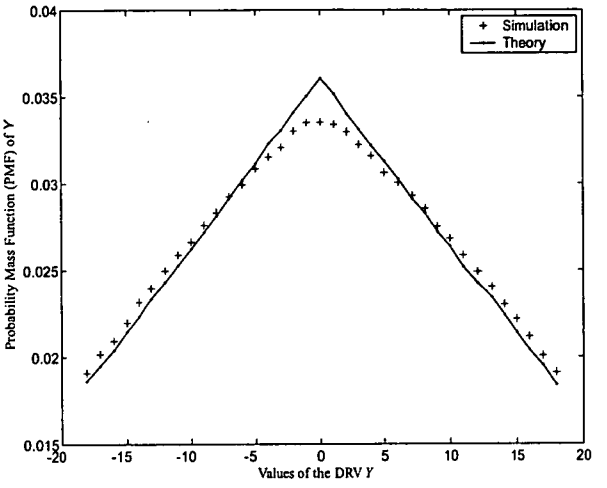


Figure 4.17: *Comparison of simulation and theoretical analysis for P_Y with $W = 32$.*

From (4.27) one observes that for large values of W, the error in the estimation of the PMF P_Y does not significantly affect the derivation of P_{Y+X_B}. This is explained because, as W increases, the middle term in (4.27) becomes dominant and P_{Y+X_B} degenerates to a uniform distribution. Analogous arguments exist for the derivation of (4.24) and (4.25). These remarks are verified by the good match of the theoretical values of all the above defined probabilities to the corresponding measured values.

4.4 Model validation

In this section, the results obtained from the proposed analytical method will be presented. Note that all the theoretical results presented are obtained via the iterative method; however, the numerical differences between the latter and the ones derived from the approximation method of Section 4.3.4 are minor. Extensive graphs are provided which show the throughput S, the conditional collision probability p and the transmission probability $P\{T_n\}$ for a very wide variety of parameters, such as contention window sizes, packet sizes, transmission rates and operation modes of the IEEE 802.11 DCF. In order to validate the accuracy of the proposed analytical model, the above theoretically obtained values are compared with their corresponding counterparts, measured from simulation runs in NS2. It will be evident from all the graphs presented that the method is highly accurate and reliable for all the sets of parameters considered, depicting a very good match to the simulation results.

The considered topology is the classic hidden terminal scenario presented in Figure 4.1, where stations A and B are continuously backlogged; in other words, at all times their buffers are non-empty. Several values of packet sizes are used for the simulations. In particular, DATA packets of size $L_d = \{256, 512, 1024, 1536, 2048, 2294\}$ bytes are used, with 2294 bytes as the MAC fragmentation threshold defined in [1]. The traffic is Constant Bit Rate (CBR) and a wide range of values of contention window are deployed, namely $W = \{32, 64, 128, 256, 512, 1024\}$. Finally, both operating modes of DCF, i.e. the RTS/CTS scheme and the Basic Access scheme are examined, with the considered transmission rates BasicRate and DataRate taking all the possible values that 802.11b suggests. Each simulation point in the graphs corresponds to the average of the measured quantity over 10 different runs of the traffic scenario for $200s$ simulation time. Other important NS2 parameters used were summarised in Table 4.1.

Figure 4.18 presents the saturation throughput of RTS/CTS for various contention window sizes W. For each graph the throughput is plotted against the DATA packet size for three different values of DataRate R. One can observe that, with W constant, throughput increases with higher transmission rates, because the transmission probability increases and, simultaneously, the collision probability p remains constant. Note that in RTS/CTS the vulnerable period and, consequently, the collision probability depend only on the RTS packet size and the value of BasicRate. Also, one can see that, with both W and R kept constant, throughput is a decreasing function of the DATA packet size. Note that throughput here is expressed in packets/s; if measured in Kbps (a metric that includes L_d) the graphs would be ascending. Finally, one observes that throughput at first increases as W increases, but there is a point where it starts decreasing again. Thus, there is an optimal value of W, which differs slightly depending on R, but can be claimed to lie between 128 and 256.

In Figure 4.19, the throughput of the Basic Access scheme is plotted against L_d for several W values. For this case, and as a complement to the results already shown in Table 4.2 of Section 4.1.3, both the proposed in this chapter method and the conventional method are compared to the simulation results; an illustrative example of the close match of the modelling technique presented in this chapter to the actual performance of the protocol, as opposed to the other technique, which significantly fails to capture the throughput values, apart from when very wide contention windows are used. The transmission rates used are 2 and 11 Mbps respectively for BasicRate and DataRate; for all the other combinations, which are omitted here due to the very low throughput they exhibit, the accuracy of the method proposed in this thesis is at least

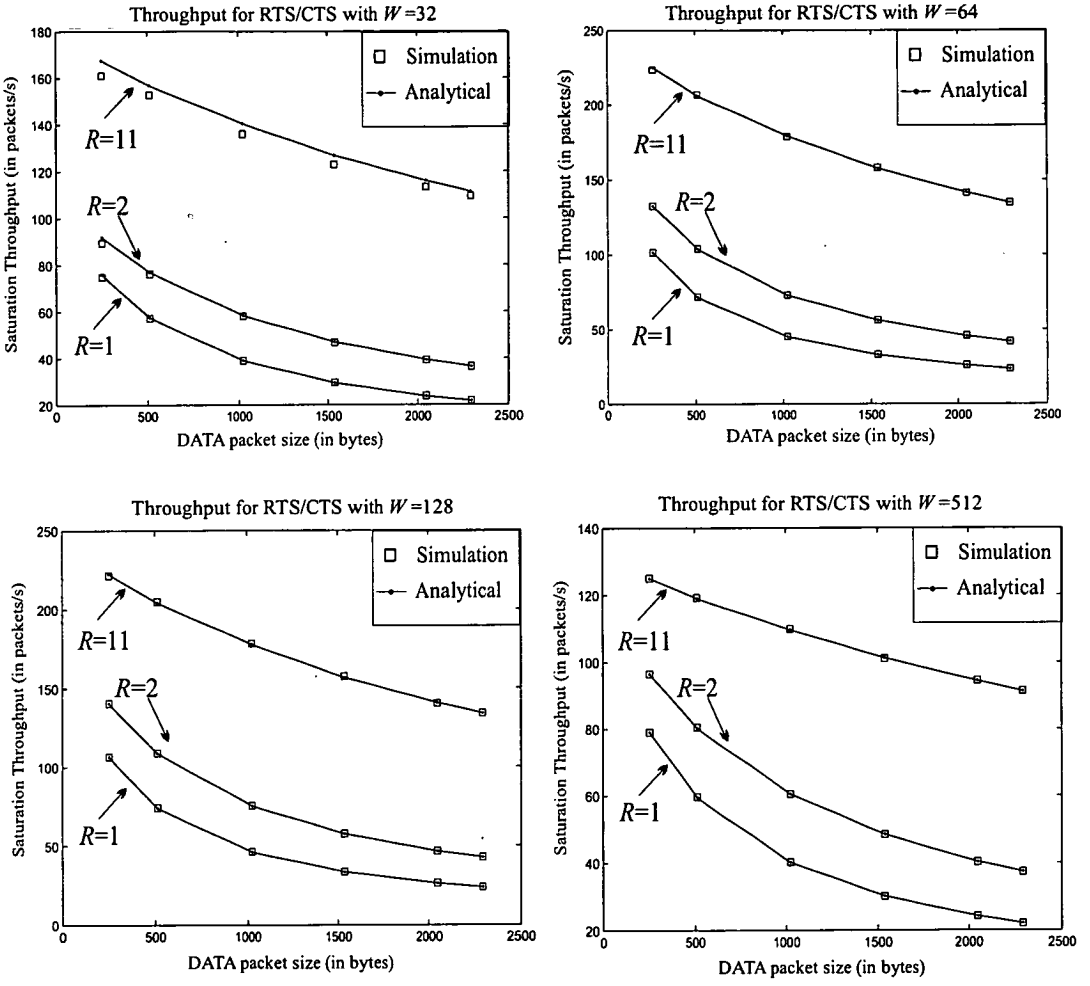


Figure 4.18: *Comparison between the analytical method and the simulation results of the saturation throughput of the RTS/CTS Scheme for several values of contention window W. For each graph the throughput against the packet size is plotted, for three distinct values of the DataRate, namely $R \in \{1, 2, 11\}$.*

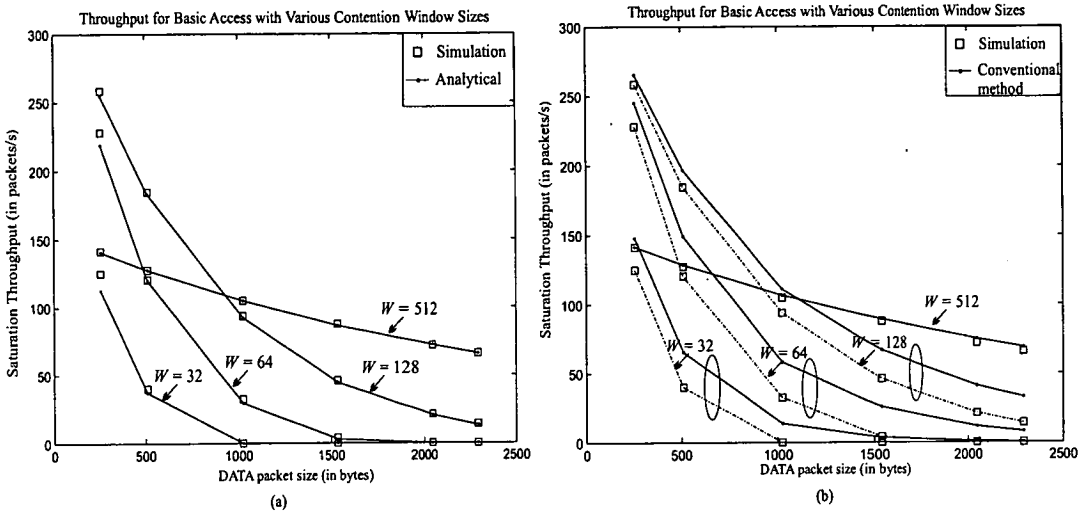


Figure 4.19: *The saturation throughput of the Basic Access Scheme is plotted against the packet size, for several values of contention window W. The BasicRate used is 2 Mbps and the DataRate is $R = 11$ Mbps. As one can observe in Figure 4.19(a), the proposed method is a very close match to the simulation results. On the contrary, the conventional method presented in Figure 4.19(b) fails to closely follow the simulation results, with the relative error being over* 100% *in many cases, as was also explained in Table 4.2 of Section 4.1.3.*

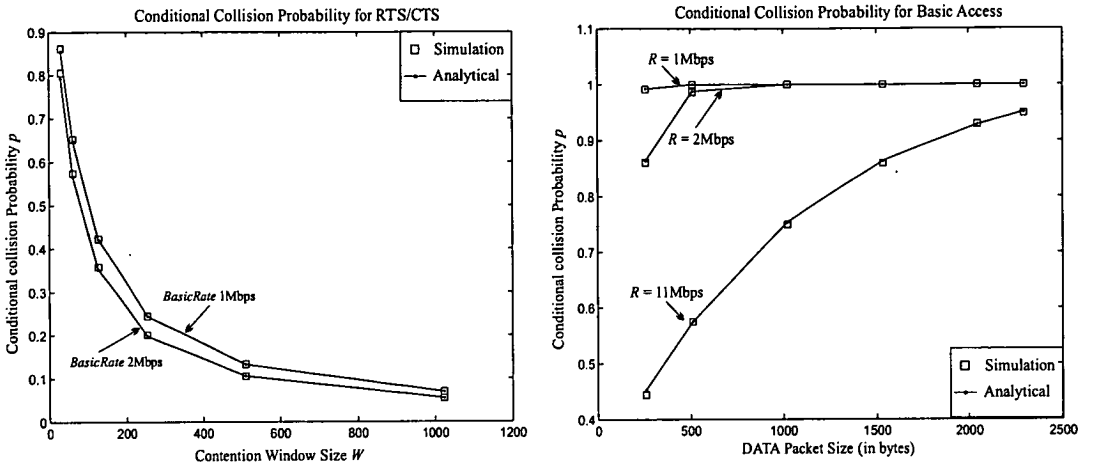


Figure 4.20: *Comparison between the analytical method and the simulation results of the conditional collision probability p. On the left, p of the RTS/CTS scheme is plotted against the contention window size (for RTS/CTS p is independent of the packet size) for two different values of BasicRate, namely 1 and 2 Mbps. On the right, p of the Basic Access scheme is plotted against the packet size for three distinct values of DataRate, namely* $R \in \{1, 2, 11\}$. *The contention window size is* $W = 128$.

equally high (in contrast to the conventional methods).

One can notice that, again, as W increases, higher values of throughput are initially observed. For very high values of contention window the throughput drops again for small packets as was the trend in RTS/CTS. On the contrary, it keeps increasing when the packet size is large (note that in Basic Access the vulnerable period and the collision probability depend on L_d). In general, one could state that the throughput for a very large contention window tends to be more 'flat' and insensitive to the variation of the packet size. This is due to the fact that, in this configuration, the process is dominated by very long idle periods, and any differences in the collision probability do not play a significant role in the actual throughput. As a final remark one could state that, for medium and large packets, Basic Access should preferably be used in combination with relatively high W; whenever this is not recommended, i.e. for delay-sensitive applications, the throughput anticipated should be quite low.

In Figure 4.20, the conditional collision probability p is plotted for RTS/CTS (left) and Basic Access (right). As far as RTS/CTS is concerned, note that p is independent of the DATA packet size and R, and only depends on the value of BasicRate adopted and the contention window size W. In fact, it is a decreasing function of both the aforementioned parameters.

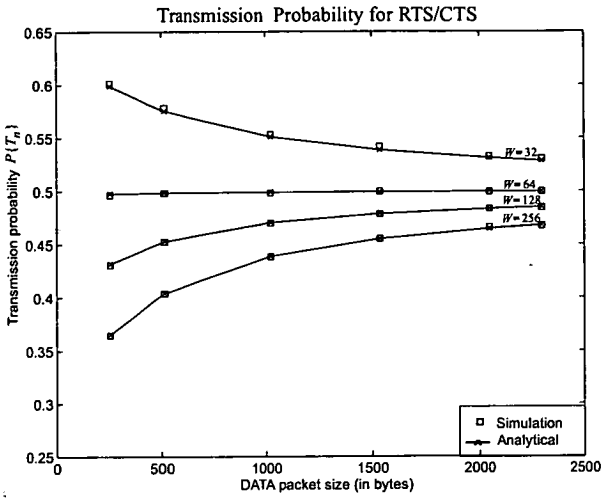


Figure 4.21: *The transmission probability $P\{T_n\}$ of RTS/CTS for different values of DATA packet sizes and contention window sizes. Both transmission rates are 1 Mbps.*

When the first parameter (BasicRate) increases, the vulnerable period becomes shorter, with a subsequent drop of p. Regarding the second parameter, one observes that as W increases and the nodes' backoff times are uniformly distributed over longer intervals, the probability of them making simultaneous transmissions inside a specific time duration (the vulnerable period) so that collision is caused, significantly drops. For Basic Access, p depends strongly on both the DATA packet size and R, which are actually the two factors that determine the vulnerable period. In fact, for the low rates 1 and 2 Mbps the collision probability is almost 1. For R =11 Mbps, it varies between 0.4 and 0.9 with an increasing trend as L_d becomes bigger.

Figure 4.21 compares the transmission probability of RTS/CTS (ratio of the transmitting slots over the sum of all the slots, both transmitting and non-transmitting) obtained from simulation with the analytically derived $P\{T_n\}$. At this point it is important to stress the fact that one of the most significant advantages of the proposed method is the accurate derivation of an objective transmission probability. All prior techniques use the conventional transmission probability τ, which is a model invented metric showing the frequency of transmission variable slots, and consequently cannot provide an accurate estimate of the channel busyness ratio, a useful Quality of Service indicator.

For this graph both transmission rates are the lowest possible, i.e. 1 Mbps; however, similar accuracy was observed for all the other values. A close match of the simulation and the theoretical values is again evident for all the contention window sizes used, i.e. $W = 32, 64, 128, 256$.

One can observe that $P\{T_n\}$ drops as W increases, as nodes spend more time being in back-off. With W constant, an interesting observation can be made. For $W = 32$ the transmission probability decreases as the packet size increases, for $W = 64$ it remains almost stable, and, finally, for larger contention window sizes, it increases as the packet size increases. This remark may appear to be counter-intuitive. However, from equations (4.8)-(4.10) and the fact that $F = L - c$ (the freezing period and the successful period differ only $c =$ (RTS+SIFS)/σ slots) the transmission probability can be written as follows:

$$P\{T_n\} = \frac{1}{1+A}, \text{ where } A = \frac{(1-p)(L-c) + \frac{W-1}{2}}{(1-p)L + pC}. \tag{4.28}$$

If one takes the derivative $\partial A/\partial L$ they can observe that indeed for $W = 32$ it is positive, and consequently $P\{T_n\}$ decreases (A increases) as the packet size increases, whereas for $W = 64$ it is almost zero and for $W = 128, 256$ it becomes negative. Qualitatively, the relative importance of the slots $c = L - F$ (these slots count as transmitting) becomes higher when both W and L are small. When either W or L increases, this term becomes less significant and, consequently, the transmission probability degrades.

As a sidenote, it is pointed out that the proposed method is also computationally efficient since it takes under a second for any scenario to produce the theoretical results in a Linux machine with two 2.4 GHz CPUs and 2 GB of Random Access Memory (RAM), using an unoptimised MATLAB®6.0 code. As a comparison, the simulation runs in NS2 take one or two orders of magnitude more time to produce statistically reliable results, even for this relatively small scenario.

4.5 Conclusions

In this chapter the hidden node problem, which arises in multihop wireless networks has been discussed. It was shown that the classical method of modelling time of [2], although it gives a correct analysis of the throughput of fully connected networks, cannot accurately describe the effect of hidden stations. It was also pointed out that prior techniques, which use Bianchi's notion of the variable slot and are based on renewal theory due to their assumption of synchronisation, exhibit an intrinsic difficulty in providing a reliable analytical tool for an arbitrary random access scheme and an arbitrary station configuration.

Therefore, a novel method of modelling was presented that relies on a constant, and thus objective for all the nodes, notion of the slot. Moreover, the analysis proposed has accurately described the dynamics of the channel by adopting a first-order (Markov) dependence between consecutive transmission states and by using random variables in order to capture the complexity of the scenario considered. Thus, the derived expressions for the throughput, the conditional collision probability and the transmission probability are a very close match with the simulation results for all the parameters and access schemes considered.

Chapter 5
Extending the Proposed Analysis to the Binary Exponential Backoff Scheme of IEEE 802.11

The analysis presented in Chapter 4 was focused on the CCW case where it was assumed that both transmitters choose the initial value of their backoff counter after each transmission according to a uniform distribution in an interval $[0, W-1]$, where W is the contention window size. This was done for a number of reasons: the assumptions and limitations of prior modelling methods are more easily grasped in a CCW environment, there is significant simplicity in introducing and explaining the proposed method, and, finally, the previous analysis can be used as a building block to derive the actual performance of the protocol when the standard backoff scheme that [1] suggests is adopted. The latter is the so called Binary Exponential Backoff (BEB) mechanism. Thus, in Chapter 5 the assumption of CCW will be relaxed and, after describing briefly how BEB works, the previously proposed method will be extended and applied to this, more complicated case.

5.1 Description of the BEB mechanism of IEEE 802.11

In IEEE 802.11 every station is refrained from continuously taking over the channel by the so called backoff mechanism, which consists of several backoff stages. After every packet transmission, and independently of whether the station was successful or faced a collision, it is forced by the protocol to wait for a period of time before retransmitting, and consequently to give the chance to the other stations-competitors to take hold of the channel as well. This period of time is equal to an integer multiple of the channel time-slot σ. In particular, the number of backoff slots it has to wait before attempting retransmission is drawn from a uniform distribution in the interval $[0, W_i - 1]$, where W_i is the contention window (CW) corresponding to the backoff stage i and W_0 is the initial CW. After every failed retransmission the station doubles its CW and moves to the next backoff stage. This process is repeated for several times

until the station reaches the so called Retransmission Limit[1] RL. At this point it resets its CW and moves to the initial backoff stage $i = 0$. Thus, the CW of a station being in the backoff stage i is given by the following relation

$$W_i = \begin{cases} 2^i W_0, & 0 \leq i \leq m \\ 2^m W_0, & m \leq i \leq RL \end{cases} \tag{5.1}$$

Typical values of the above parameters are $W_0 = 32$, $m = 5$, $SRL = 6$, but in general the IEEE 802.11 stations are able to configure them as they wish.

An important detail that should not be forgotten to be mentioned is that whenever a station is successful, it resets its CW, in other words, after a successful transmission, the station always goes to the initial backoff stage choosing its backoff counter from the interval $[0, W_0 - 1]$. This has the following very important consequence, which is the source of the very much discussed in the literature 'unfairness of IEEE 802.11'.

Let's suppose that both stations initiate a sequence of DATA packet transmissions being in the same backoff stage W_i. After the contention phase either collision will occur, or one of the transmitters, say A, will manage to take hold of the channel successfully whereas the other, say B, as soon as it senses the CTS packet from the receiver, will freeze its counter as described in previous sections. In the first case, both stations will go to the next backoff stage, thus competing with equal chances for the hold of the channel. On the contrary, in the case that A was successful, then, in the end of the successful handshake, it will reset its counter to W_0. It is clear that for large values of i this leads to the channel being repeatedly taken over by the same sender, in this case A, whereas the other station (B) repeatedly freezes its counter (see Figure 5.1). It is observed that this pattern keeps occurring, the chance of B taking hold of the channel becoming more and more difficult as it keeps entering into higher backoff stages. This unfairness is mitigated only when the retransmissions of B exceed the RL and B resets its counter, reverting to the backoff stage W_0 and having—for the next transmission attempt—a statistically equal chance of taking hold of the channel. The existence of RL is the element of

[1] In [1], two retransmission limits exist. The first one, called Short Retransmission Limit, SRL, is the maximum number of retransmissions that the first packet (RTS for the RTS/CTS scheme and the DATA packet for the Basic Access scheme) of a handshake can be transmitted. The second one, defined as the Long Retransmission Limit, LRL, refers to the maximum number of times that the data packet can be retransmitted when RTS/CTS is used, conditioned that the RTS packet is successful. LRL is only used to describe the data packet collision phenomena which were introduced in Chapter 3, and does not apply to the classic hidden terminal scenario (under the assumptions of no mobility and error-free channel). Thus, from now on, when referring to retransmission limit, the SRL is implied.

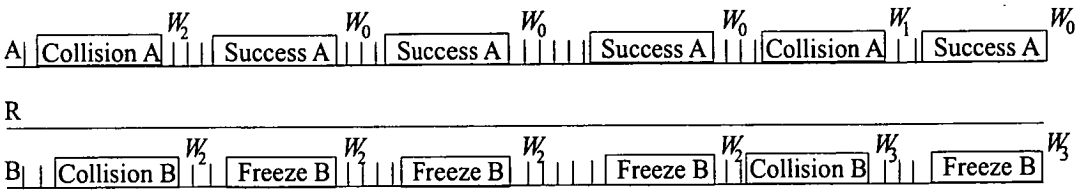


Figure 5.1: *A demonstration of the unfairness of BEB of IEEE 802.11.*

BEB which keeps the protocol fair in the long-term in the considered scenario of the classic hidden terminal case.

5.2 Adaptation of the proposed model to BEB—Part I transmission probability and throughput

The basic difference between the CCW case and the BEB mechanism is the fact that, depending on the backoff stage of the transmitter, the probability that it will face a collision, conditioned on it transmitting, is quite different. In fact, as will be evident in the following sections, the probability increases with the backoff stage, in other words the higher the backoff stage the station is in, the more probable it is that it will go into an even higher one. This is the source of the so-called unfairness of the IEEE 802.11 in a short-term basis, until the station reaches the maximum limit of retransmissions and resets its backoff window.

For this reason, the model which was introduced in Chapter 4 must be adjusted in order to incorporate the specific features of channel contention in the environment where stations employ the BEB. This can be done in the following, fairly straightforward manner, which shows that the model is indeed general and easily adjustable to BEB or an arbitrary backoff scheme. As stated previously, although the average conditional collision probability p can still be derived, the latter is not enough for an accurate analysis of the problem. A more detailed definition of collision probability is required depending on the backoff stage of the station. To this end, let's define the vector $\mathbf{p} = [p_0 p_1 ... p_{RL-1}]$ whose elements are the conditional collision probabilities p_i of a station, when the latter is in backoff stage i. In order to compute the elements of $\mathbf{p}$ the DTMC of Figure 5.2 is used which represents the possible backoff stages of a station. Let ξ_i denote the steady-state probability that the station is in backoff stage i, $i \in [0, RL]$ and let's also define the vector $\boldsymbol{\xi} = [\xi_0 \xi_1 ... \xi_{RL}]$. These probabilities will be referred to as 'backoff-dwelling'

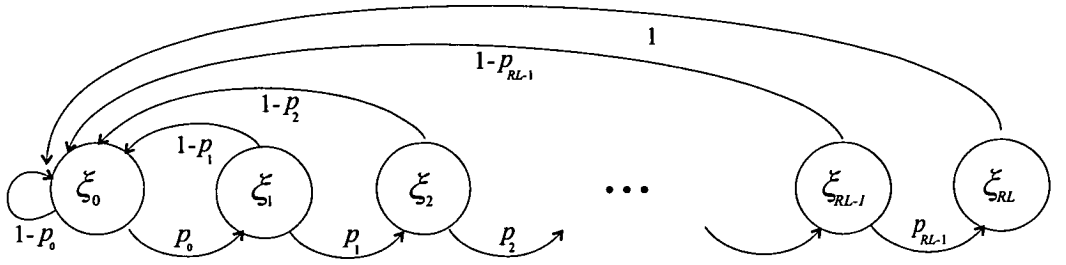


Figure 5.2: *The 'Backoff-Dwelling' distribution.*

probabilities. Due to the balance equations of the MC, if $\boldsymbol{\xi}$ is known, then the elements of the vector $\mathbf{p}$ can be calculated as follows:

$$p_i = \frac{\xi_{i+1}}{\xi_i},\ i \in [0, RL-1]. \tag{5.2}$$

In order to 'pass' from the variable slot-based model of Bianchi to the proposed fixed length slot-based model, two main changes to the MC of [2] are made. Firstly, the assumption of a constant conditional collision probability at every stage is dropped, and the previously defined p_i's are used. Secondly, freezing states are added after every backoff decrement. Following a similar method of analysis and notation with the one in the Section 4.2, the idle periods of each transmitter are discretised using the MC that is shown in Figure 5.3 and is very similar with the one of Figure 4.7, with the only difference that the station now can be in different backoff stages.

Let F denote the length of the freezing period in fixed length slots (an integer multiple of σ). Let's also define $\pi_{k,m}$, where $k \in [0, RL], m \in [0, W_k - 1]$ the steady-state probability that the backoff stage is k and the counter is equal to m, non-freezing. The corresponding states, represented as large circles in Figure 5.3, are the usual states the MC in [2] consists of. In addition, $\pi^i_{k,m}$ is the steady-state probability that the counter of stage k and value m is frozen ($i \in [1, F]$ because there are F such states for each (k, m)). These additional states are represented as smaller circles in Figure 5.3. Let also p_f be the probability that a node in a state (k, m) listens to another transmission and freezes its counter, i.e. the transition probability from a state (k, m) to the successive (k, m)[1]. The transition probabilities between two states $(k, m)^i$ and $(k, m)^{i+1}$, $i \in [1, F-1]$ are equal to 1. From the analysis of the MC one has:

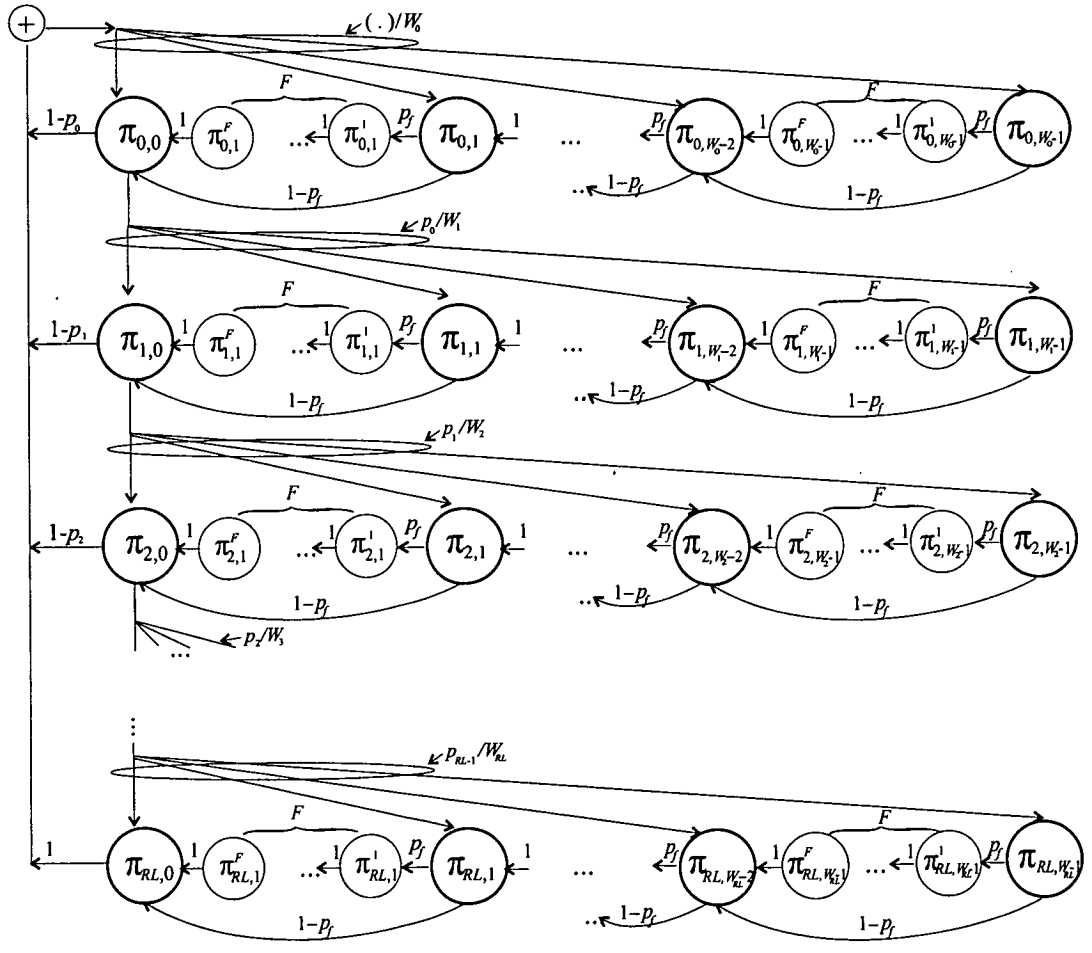


Figure 5.3: *Incorporation of the freezing of the backoff counter in the MC.*

$$\begin{aligned} \pi_{k,m} &= \tfrac{W_k - m}{W_k}\pi_{k,0} \quad , k \in [0, RL],\ m \in [1, W_k - 1] \\ \pi^i_{k,m} &= p_f \pi_{k,m} \quad , k \in [0, RL],\ m \in [1, W_k - 1], i \in [1, F] \end{aligned} \tag{5.3}$$

As far as the freezing probability is concerned, and referring the reader to Section 4.2.1, one gets a similar expression that needs to take into account that the collision probability and the average backoff duration are different depending on the backoff stage of the station. Thus, one obtains:

$$p_f = \frac{\sum_{i=0}^{RL} \xi_i (1 - p_i)}{\sum_{i=0}^{RL} \xi_i E[BC_i]}, \tag{5.4}$$

where $E[BC_i]$ is the expected value of the backoff counter when the stage of the transmitter is i.

The states where a station transmits are of the form $(i, 0)$, where $i \in [0, RL]$. So, due to the ergodicity[2] of the MC, after the discretisation of the backoff periods the steady-state probabilities of the transmitting states are derived:

$$\begin{aligned} \pi_{i,0} &= \prod_{j=0}^{i-1} p_j \pi_{0,0},\ \ i \in [1, RL] \\ \pi_{0,0} &= \left\{ \left[1 + \frac{W_0 - 1}{2}(1 + p_f F) \right] \right. \\ &\quad \left. + \sum_{i=1}^{RL} \left[1 + \frac{W_i - 1}{2}(1 + p_f F) \right] \prod_{j=0}^{i-1} p_j \right\}^{-1}. \end{aligned} \tag{5.5}$$

The expression for the transmission probability after incorporating the freezing of the counters is

$$\tau' = \sum_{i=0}^{RL} \pi_{i,0}. \tag{5.6}$$

[2]The Markov chain that was introduced only consists of states that are recurrent and aperiodic. A state is called recurrent, if the chain returns to it with probability 1 (i.e. it is non-transient). A recurrent state is called aperiodic, if the greatest common divisor of the set of positive integers n, which represent the number of steps after which the process returns to the same state, is 1. A recurrent, aperiodic state is called ergodic. A MC chain that only comprises ergodic states and, also, all its states can communicate with each other, is called ergodic [68]. It can be proved that a steady-state distribution exists for this MC.

Finally, the probability that the station is transmitting at a fixed length time slot, $P\{T_n\}$, as well as the throughput S are calculated in a similar manner as in the case of CCW. For convenience the final expressions for $P\{T_n\}$ and the throughput S are repeated:

$$P\{T_n\} = \frac{\tau'}{Q + \tau'(1-Q)}, \text{ where } Q \triangleq \frac{1}{(1-p)L + pC}, \tag{5.7}$$

and

$$S = P\{T_n\} \cdot \frac{(1-p)L}{(1-p)L + pC} \cdot \frac{1}{L} \cdot \frac{L_d}{\sigma}, \tag{5.8}$$

where p is the average conditional collision probability, L, C are correspondingly the duration of a successful and a collision period in multiples of the channel time slot σ, L_d is the DATA packet size in bytes and the term L_d/σ converts the expression for the throughput from packets per slot to bps. Note that combining the equations (5.4) and (5.5) and substituting them to (5.6), τ' is derived and is ultimately substituted in (5.7). The only unknown is the steady-state distribution ξ, which is dependent on the specific topology and will be computed separately in the next section.

5.3 Adaptation of the proposed model to BEB—Part II conditional collision probability

It is apparent from the description of BEB that in this mechanism there are no constant transition probabilities p_{cc} and p_{sc} as in the CCW assumption. On the contrary, in BEB, the probability that e.g. the current state, s_n, is a collision, conditioned that the previous, s_{n-1}, was a collision as well, is dependent on the backoff stages of both transmitters (or, to be more accurate, on the relative difference of the stages). For example, it is much more probable that collision will occur when the stages of the transmitters are identical, than when the one is small and the other one is high.

Because of that observation, the proposed DTMC of Figure 4.10 does not suffice for an accurate representation of the transitions as seen by the receiver R. In order to accurately describe the system state changes, two more variables/dimensions need to be added to the DTMC, namely the backoff stages of the transmitters. In particular, a system state is defined as a triple (E,i,j), where E is the state of the channel as observed by R and i, j the backoff stages of the transmitters A and B respectively right after the transmission of one station (in a successful case) or both

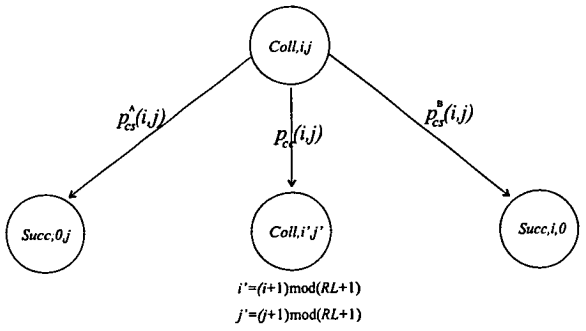


Figure 5.4: *The transitions from a collision state.*

(in a collision case), i.e. $E \in \{Succ, Coll\}$, and $i, j \in [0, RL]$. The number of states of the DTMC corresponding to collision are $(RL+1)^2$, whereas the ones corresponding to success are only $2RL+1$, because after a station transmits successfully, it resets its CW to the initial backoff stage W_0, so there are only successful states of the form $(Succ, 0, j)$ or $(Succ, i, 0)$. For a schematic presentation of the possible transitions of the embedded DTMC the reader is referred to Figures 5.4 and 5.6.

After collision has occurred, both transmitters go to the next backoff stage, with the exception of one of them reaching the RL. Then, this transmitter resets its backoff counter and moves to the backoff stage 0. Thus, an allowed transition from a collision state to another collision state can be described as follows

$$(Coll, i, j) \overset{p_{cc}(i,j)}{\rightarrow} (Coll, i \oplus 1, j \oplus 1), \tag{5.9}$$

where $i \oplus 1 \triangleq (i+1)\text{mod}(RL+1)$. The term $p_{cc}(i,j)$ describes the transition probability from a collision state $(Coll, i, j)$ to a consecutive collision state described above. Also, if after a collision state one of the transmitters is successful, there is the transition

$$(Coll, i, j) \overset{p_{cs}^{A}(i,j)}{\rightarrow} (Succ, 0, j), \tag{5.10}$$

when transmitter A is successful with probability $p_{cs}^{A}(i,j)$ and

$$(Coll, i, j) \overset{p_{cs}^{B}(i,j)}{\rightarrow} (Succ, i, 0), \tag{5.11}$$

when B is successful with probability $p_{cs}^{B}(i,j)$.

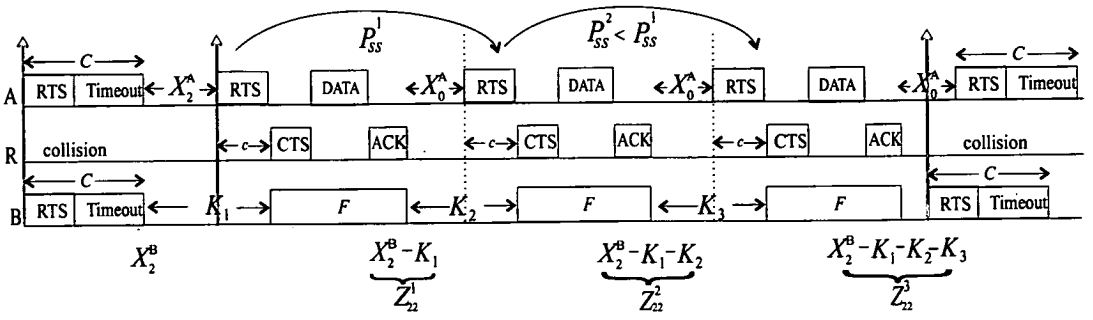


Figure 5.5: *The DTMC does not include self-transitions.*

5.3.1 Independence from the history

As far as the transitions from a successful state are concerned, the analysis is not so straightforward. This stems from the fact which was already described previously (Figure 5.1), about a station taking hold of the channel for more than one successive transmission. The analysis will elaborate on this behaviour of the BEB in order to justify the analysis that follows. The following example is presented, accompanied with Figure 5.5, where the simplifying assumption is made that, in the beginning of the 'observation window', both senders started transmitting RTS simultaneously; thus collision occurred, after which, both of them go to the backoff stage W_2. The notation X_i^A (X_i^B) is introduced and denotes the DRV that describes the initial value of the counter of A (B) when it is in the backoff stage i. Remember that X_i^A follows a uniform distribution in the interval $[0, W_i - 1]$.

Suppose that A is successful in the next state; this occurs with probability $P\{X_2^B > X_2^A + c\}$. If Z_{22}^1 denotes the DRV that represents the remaining backoff slots of B after freezing (the subscript 22 stands for the backoff stages of A and B being 2 and 2 respectively, and the superscript 1 shows that it is the first time that A manages to transmit successfully in this sequence) then one gets

$$Z_{22}^1 = X_2^B - \left(X_2^A + c\right) | X_2^B > X_2^A + c. \tag{5.12}$$

Afterwards, station A is successful for a second consecutive transmission (one is interested in the probability that it is successful, given that it was successful previously) with probability

$$P_{ss}^1 = P\left\{Z_{22}^1 > X_0^A + c\right\}. \tag{5.13}$$

Note that the DRV that represents the initial backoff counter of A, after the first time it was successful, is X_0^A, because the IEEE 802.11 stations reset their backoff counter to the initial stage after a successful transmission. Also, the new DRV showing the remaining backoff slots of B after the second successful transmission of its competitor is

$$Z_{22}^2 = Z_{22}^1 - \left(X_0^A + c\right) | Z_{22}^1 > X_0^A + c. \tag{5.14}$$

Following similar arguments and supposing that A is successful for the third time in a row

$$P_{ss}^2 = P\left\{Z_{22}^2 > X_0^A + c\right\}. \tag{5.15}$$

and $Z_{22}^3 = Z_{22}^2 - \left(X_0^A + c\right) | Z_{22}^2 > X_0^A + c$. If one observes closely the above equations (5.12)–(5.15), one can notice that the transition probability P_{ss}^1 is different; in fact, it is higher than P_{ss}^2, although they both represent the self-transitions from state $(Succ, 0, 2)$ to itself. Thus, it becomes apparent that the model which was introduced earlier does not always sustain the main characteristic of a MC, i.e. the fact that, given the present state of the system, the 'future' is independent of the 'past' [68]. As a consequence, it is necessary to make some essential modifications to the embedded channel evolution DTMC, so that it does not come in conflict with the fundamental definition of Markovian models.

Consequently, as far as the transitions from a successful state are concerned, any self-transitions must be excluded. In other words only transitions of the form are considered (see Figure 5.6):

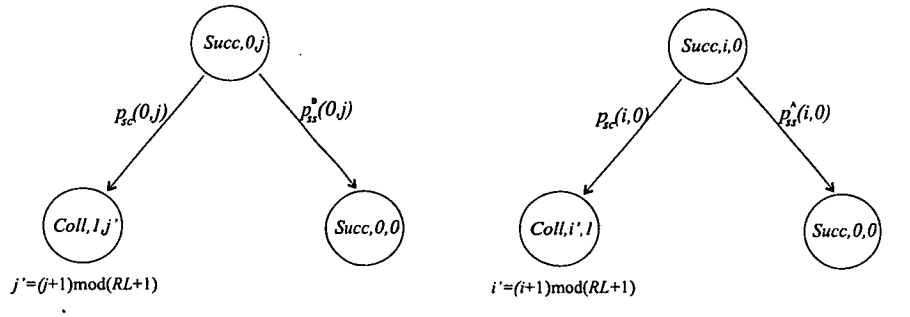
$$(Succ, 0, j) \overset{p_{sc}(0,j)}{\rightarrow} (Coll, 1, j \oplus 1), \tag{5.16}$$

in the case that collision follows after A was successful and

$$(Succ, 0, j) \overset{p_{ss}^B(0,j)}{\rightarrow} (Succ, 0, 0), \tag{5.17}$$

when success of B follows after A was successful. Obviously, the transitions from a state $(Succ, i, 0)$ where B was successful are entirely symmetrical, i.e.

$$(Succ, i, 0) \overset{p_{sc}(i,0)}{\rightarrow} (Coll, i \oplus 1, 1), \tag{5.18}$$



(a) Transitions from a state where A had been successful.

(b) Transitions from a state where B had been successful.

Figure 5.6: *The transitions from a successful state.*

when collision follows after B was successful, and finally

$$(Succ, i, 0) \overset{p_{ss}^{A}(i,0)}{\rightarrow} (Succ, 0, 0), \tag{5.19}$$

when success of A follows after B was successful. With reference again to Figure 5.5, the time-index space of the DTMC [68] includes only the transitions denoted with the red arrows and not the ones shown with dotted lines.

5.3.2 Definition of the transition probabilities

The set of transitions, which completely defines the transition matrix of the DTMC, is presented in Figures 5.4 and 5.6. In Figure 5.4 the transitions from a collision state are shown; $p_{cc}(i,j)$ defines the transition probability that collision follows, conditioned on collision occurring in the previous state, and $p_{cs}^{A}(i,j)$ ($p_{cs}^{B}(i,j)$) the probability that A (B) is successful conditioned on collision occurring in the previous state. In the former expressions, the indexing (i,j), $i,j \in [0, RL]$ is used, because, as noted in the introduction, the probability values are different for different combinations of the backoff stages of the transmitters. This is actually the reason why a simple DTMC like the one of Figure 4.10 would fail to accurately describe the BEB. Obviously there are $(RL+1)^2$ different values for each of the transition probabilities above, which account for all the possible collision states $(Coll, i, j)$.

Furthermore, with respect to the transition probabilities from a successful state (Figure 5.6), they can be divided into two very similar sets, according to which transmitter had been successful in the previous state. If A was successful (Figure 5.6(a)) then the initial state is of the form

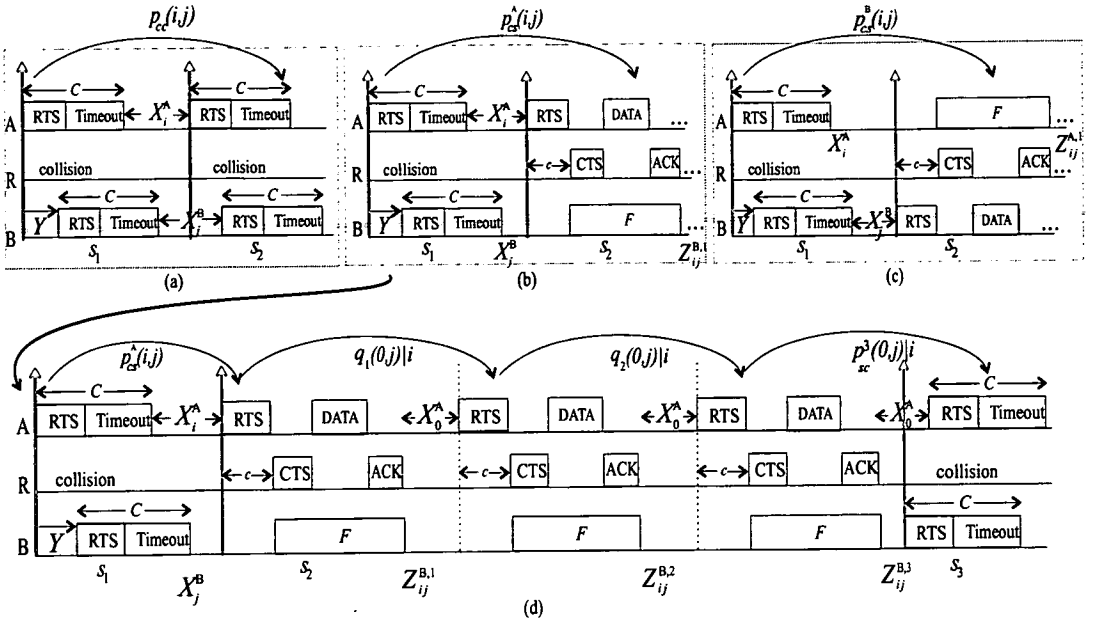


Figure 5.7: *A paradigm of the analysis of the transition probabilities.*

$(Succ, 0, j)$, and, the way the DTMC is defined, only two events are possible. The first is that collision follows; the corresponding transition probability is denoted as $p_{sc}(0, j)$, $j \in [0, RL]$. The second is that success of the other transmitter, B, follows ($p^B_{ss}(0, j)$). The transition probabilities from a system state where B is successful are analogous (see Figure 5.6(b)).

5.3.3 Calculation of the transition probabilities

The method that is used for the derivation of the transition probabilities defined above is similar to the one presented in Section 4.3.2. In particular, the analysis begins with the assumption that the system is in a collision state (state s_1) and the backoff stages of the transmitters are i and j, and the transition probabilities $p_{cc}(i, j)$, $p^A_{cs}(i, j)$ and $p^B_{cs}(i, j)$ are calculated. These three possible transitions are shown in Figures 5.7(a),(b),(c) respectively.

Consecutively and without loss of generality, the scenario of Figure 5.7(b) is isolated, where in the state s_2, A is successful, and it is used for the derivation of the probabilities $p_{sc}(0, j)|i$ and $p^B_{ss}(0, j)|i$. Note that the probabilities $p_{sc}(0, j)|i$ and $p^B_{ss}(0, j)|i$ are being used and **not** the probabilities $p_{sc}(0, j)$ and $p^B_{ss}(0, j)$, which are the ones of ultimate interest, because, in this specific configuration, the former are conditional on the fact that in the state s_1 trans-

mitter A was in the specific backoff stage i. In fact, a transition to the state $(Succ, 0, j)$ is possible to have followed after a state $(Coll, i^*, j)$, where $i^* \neq i$ and the corresponding transition probabilities $p_{sc}(0,j)|i^*$ and $p_{ss}^B(0,j)|i^*$ would be different. For example, the state $(Succ, 0, 2)$ could have been preceded by all of the following collision states: $(Coll, 0, 2)$, $(Coll, 1, 2)$, $(Coll, 2, 2)$,...,$(Coll, RL-1, 2)$, $(Coll, RL, 2)$. It will be shown later in this chapter how $p_{sc}(0,j)$ and $p_{ss}^B(0,j)$ are obtained once all the conditional probabilities $p_{sc}(0,j)|i$ and $p_{ss}^B(0,j)|i$, $i \in [0, RL]$ have been calculated. For the moment, the analysis that follows explains how the expressions $p_{cc}(i,j)$, $p_{cs}^A(i,j)$, $p_{cs}^B(i,j)$, $p_{sc}(0,j)|i$ and $p_{ss}^B(0,j)|i$ for specific values of i, j are derived.

In particular, assuming an arbitrary known distribution for the DRV Y that describes the difference of the RTS transmissions in state s_1, one obtains the following:

$$p_{cc}(i,j) = P\left\{-c \leq Y + X_j^B - X_i^A \leq c\right\}, \tag{5.20}$$

$$p_{cs}^A(i,j) = P\left\{Y + X_j^B > X_i^A + c\right\}, \text{ and} \tag{5.21}$$

$$p_{cs}^B(i,j) = P\left\{X_i^A > Y + X_j^B + c\right\}. \tag{5.22}$$

Furthermore, analysing further the second scenario, where, after collision, A transmits successfully, the transition probabilities $p_{sc}(0,j)|i$ and $p_{ss}^B(0,j)|i$ will be calculated. Note that in general, after A takes over the channel successfully, and depending on the backoff stages i, j, there is a sequence of times when A is repeatedly successful (the self-transitions mentioned previously). At some point A fails to transmit successfully again and, consecutively, there is either a collision or B transmits successfully. This constitutes a real transition of the DTMC because the previous self-transitions are not taken into account. In Figure 5.7(d) an example of such a sequence of events is depicted, when A manages to transmit successfully for three consecutive times and then collision occurs (which is referred to as state s_3 of the DTMC). The red arrows in Figure 5.7 show the time instants when the transitions of the embedded DTMC occur.

The transition probabilities of Figure 5.7(d) are calculated using DRVs and truncated distributions as was done in the previous chapter. If $Z_{ij}^{B,1}$ denotes the DRV that represents the remaining backoff slots of B after freezing (the subscript ij refers to the backoff stages of A and B being i and j respectively, and the superscript 1 shows that it is the first time that A manages to transmit

successfully in this sequence), then one obtains:

$$Z_{ij}^{B,1} = X_j^B - \left(X_i^A + c\right) | X_j^B > X_i^A + c. \tag{5.23}$$

Afterwards, station A is successful for a second consecutive transmission (one is interested in the probability that it is successful given that it was successful previously) with probability

$$q_1(0,j)|i = P\left\{Z_{ij}^{B,1} > X_0^A + c\right\}. \tag{5.24}$$

Also, the new DRV showing the remaining backoff slots of B after the second successful transmission of its competitor is

$$Z_{ij}^{B,2} = Z_{ij}^{B,1} - \left(X_0^A + c\right) | Z_{ij}^{B,1} > X_0^A + c. \tag{5.25}$$

Following similar arguments and supposing that A is successful for the third time in a row one has:

$$q_2(0,j)|i = P\left\{Z_{ij}^{B,2} > X_0^A + c\right\}. \tag{5.26}$$

and

$$Z_{ij}^{B,3} = Z_{ij}^{B,2} - \left(X_0^A + c\right) | Z_{ij}^{B,2} > X_0^A + c. \tag{5.27}$$

Then, the values of the DRVs $Z_{ij}^{B,3}$ and X_0^A are such that collision occurs with probability

$$p_{sc}^3(0,j)|i = P\left\{-c \leq Z_{ij}^{B,3} - X_0^A \leq c\right\}, \tag{5.28}$$

where the superscript 3 means that A was successful for three consecutive times before collision occurred.

Generally speaking (the reader should refer to the 'tree-structure' of Figure 5.8), after the initial collision state there are three possible states, i.e. collision, success of A and success of B. Concentrating on the success of A, this can be followed by a sequence of self-transitions, until the probability that A transmits successfully for the $(N+1)$st consecutive time becomes negligible. Let $q_k(0,j)|i,\ k \in [1,N]$[3] denote the probability that, after A was successful for the $(k+1)$st consecutive time, collision occurs as $p_{sc}^k(0,j)|i$, and the probability that, after A was successful for the $k+1$th consecutive time, B manages to transmit successfully as $p_{ss}^{k,B}(0,j)|i$. The states

[3]Note that, as defined the 'tree-structure', N is such that $q_N(0,j)|i$ is very small.

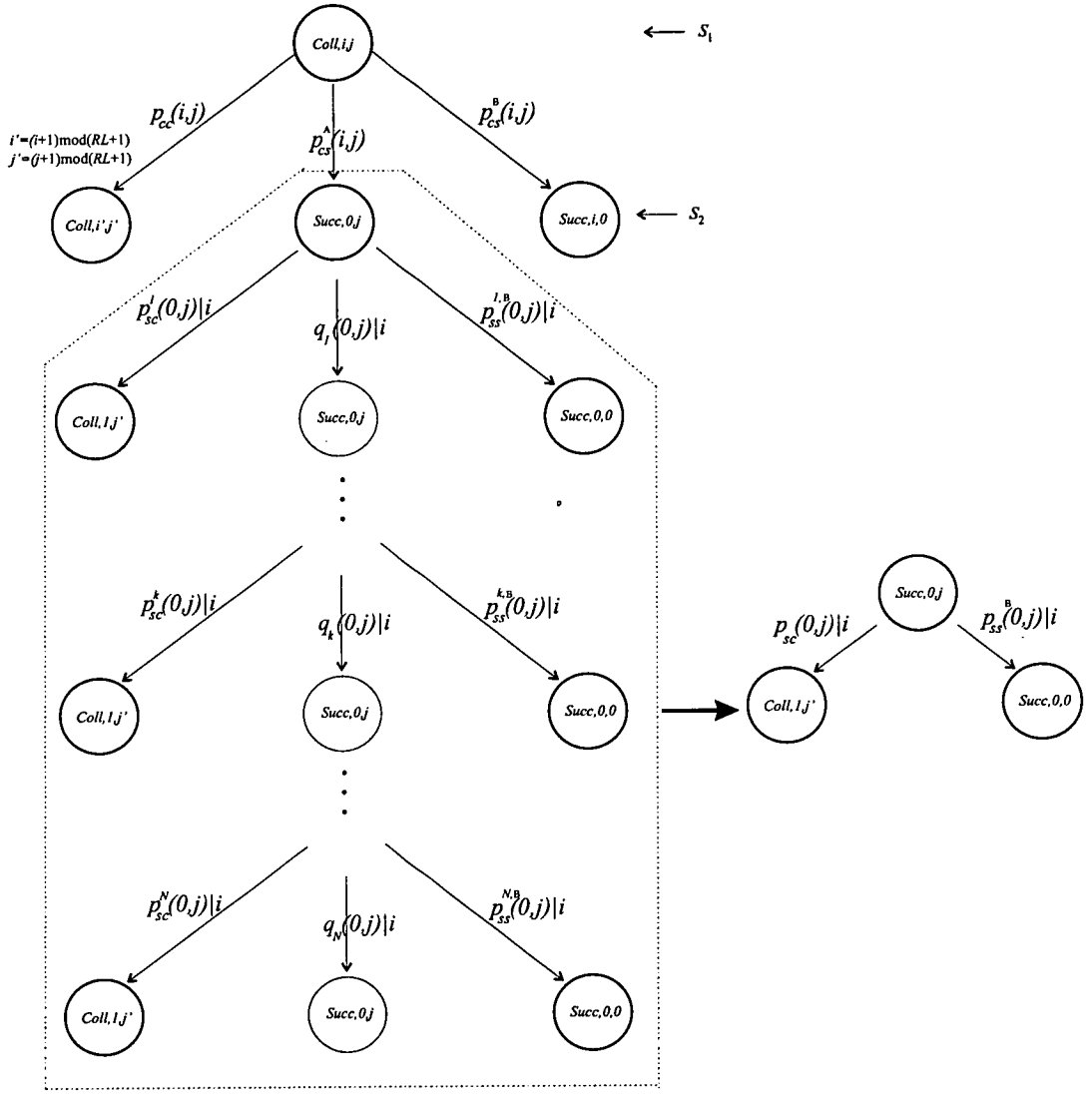


Figure 5.8: *The 'tree-structure' for the analysis of the transition probabilities.*

that are represented by grey circles are not DTMC states as mentioned previously. Also, note that all the probabilities above are conditional on i, meaning that in state s_1, A was in backoff stage i.

The total probability that collision will follow in the state s_3, conditioned on success of A occurring in the state s_2 (this can happen after several self-transitions) can be calculated as follows: It is the probability $p_{sc}^1(0,j)|i$ that collision will occur immediately after the first time A was successful, added to the probability $p_{sc}^2(0,j)|i \cdot q_1(0,j)|i$ that collision will occur after the second time A was successful, provided that A indeed transmitted successfully for a second time etc.; in other words

$$\begin{aligned} p_{sc}(0,j)|i &= p_{sc}^1(0,j)|i + p_{sc}^2(0,j)|i \cdot q_1(0,j)|i + ... \\ &+ p_{sc}^N(0,j)|i \cdot q_1(0,j)|i \cdot ... \cdot q_{N-1}(0,j)|i, \\ &= \sum_{k=1}^{N} \left[p_{sc}^k(0,j)|i \cdot \prod_{n=0}^{k-1} q_n(0,j)|i \right], \text{ where } q_0(0,j)|i \triangleq 1. \end{aligned} \tag{5.29}$$

Similarly, for the probability that success of B will take place in the state s_3, conditioned on success of A occurring in the state s_2, one gets:

$$p_{ss}^B(0,j)|i = \sum_{k=1}^{N} \left[p_{ss}^{k,B}(0,j)|i \cdot \prod_{n=0}^{k-1} q_n(0,j)|i \right]. \tag{5.30}$$

After the derivation of $p_{sc}(0,j)|i$ and $p_{ss}^B(0,j)|i$ the ultimate aim is to calculate $p_{sc}(0,j)$ and $p_{ss}^B(0,j)$ independently of i. Quoting from [72]:
"Suppose we have a set of propositions $\{A_1, ..., A_n\}$ which on information X are mutually exclusive, in other words

$$P\{A_i A_j|X\} = P\{A_i|X\} \cdot \delta_{ij}, \text{ where } \delta_{ij} \text{ is the Dirac function.} \tag{5.31}$$

Then $P\{C|A_1 + A_2 + ... + A_n X\}$ is a weighted average of the separate plausibilities $P\{C|A_i X\}$, i.e.

$$P\{C|A_1 + A_2 + ... + A_n X\} = \frac{\sum_i P\{A_i|X\} P\{C|A_i X\}}{\sum_i P\{A_i|X\}}." \tag{5.32}$$

In the current case the event C corresponds to the event of a transition from a state $(Succ, 0, j)$ to $(Coll, 1, j \oplus 1)$ when referring to the probability $p_{sc}(0,j)$, and that of a transition from a state $(Succ, 0, j)$ to $(Succ, 0, 0)$ when referring to $p_{ss}^B(0,j)$. X is the information of the previous

state (the one which was defined as s_1) and

$$\{A_1, A_2, ..., A_n\} = \{(Coll, 0, j), (Coll, 1, j), ..., (Coll, RL, j)\}, \tag{5.33}$$

mutually exclusive states. From the above analysis one obtains the unconditional transition probabilities of interest:

$$p_{sc}(0,j) = \frac{\sum_{i=0}^{RL} \gamma(Coll, i, j) \cdot p_{sc}(0,j)|i}{\sum_{i=0}^{RL} \gamma(Coll, i, j)} \text{ and} \tag{5.34}$$

$$p_{ss}^{B}(0,j) = \frac{\sum_{i=0}^{RL} \gamma(Coll, i, j) \cdot p_{ss}^{B}(0,j)|i}{\sum_{i=0}^{RL} \gamma(Coll, i, j)}, \tag{5.35}$$

where $\gamma(Coll, i, j)$ symbolises the steady-state probability that the system is in state $(Coll, i, j)$ and $j \in [0, RL]$.

However, one can observe that in the last two equations the steady-state distribution of the DTMC appears and is still unknown. To overcome this problem an iterative method like the one of Chapter 4 will be adopted, where in the first iteration the steady-state distribution $\gamma(.)$ as well as the PMF of the DRV Y are defined arbitrarily (e.g. as uniform distributions), then, the transition probabilities are derived from the equations (5.20)–(5.22) and (5.34)–(5.35) based on that initial guess and, consequently, a new estimated steady-state distribution is derived according to the method described in the next section. The iteration process converges quite fast and independently of the initial assumption for the steady-state distribution.

A final observation of this paragraph concerns the fact that, for the derivation of the transition probabilities $p_{sc}(0,j)$ and $p_{ss}^{B}(0,j)$ the analysis was based on the assumption that the system state s_1 prior to state $(Succ, 0, j)$ was collision, $(Coll, i, j)$. The results that will be depicted later show that this assumption does not sacrifice the accuracy of the proposed method. However, there is also a theoretical justification for this assumption being valid. As the DTMC was defined, the only state that can precede a successful state of the form $(Succ, 0, j), j \neq 0$ or $(Succ, i, 0), i \neq 0$ is a collision state $(Coll, i, j)$. This is so because in the definition of the

DTMC self-transitions are not allowed. The only case where there can be a successful state before a successful state is in the scenario that, e.g., A is transmitting successfully, and at some point B takes hold of the channel and is the one who transmits successfully (see Equations (5.17) and (5.19)). However, the above case has a very low probability of occurrence.

5.3.4 The steady-state distribution

After the above analysis, the transition matrix, P, of dimension (l, l), where $l \equiv (RL+1)^2 + 2RL + 1$, is well defined. One can find the steady-state distribution of the embedded DTMC with numerical methods (e.g. using the technique of finding the eigenvectors which is supported in common software tools, e.g. in MATLAB®). In particular, if the steady-state probabilities of the embedded DTMC are denoted as $\mathbf{v} = [v_0, v_1, ...]$, then $\mathbf{v}$ is calculated solving the linear system of equations:

$$\begin{aligned} \mathbf{v} &= \mathbf{v}P \\ \mathbf{ve} &= 1, \end{aligned} \tag{5.36}$$

where $\mathbf{e}$ is a column vector with ones. However, in order to calculate the probability p', all the transmission attempts of the senders should be accounted for, a fact that is in contradiction to the exclusion of the self-transitions from the embedded DTMC.

The way to come around this problem is to account for the number of 'steps' (often referred to in the literature as 'sojourn times') that the system stays in a successful state, in other words to adopt a Semi-Markovian approach. A Semi-Markov process is a generalisation of a Markov process, where, the distribution of time that the process spends in a given state is allowed to be general, as opposed to that of a DTMC, which follows a geometrical distribution [68]. Thus, in a Semi-Markov process, the transitions between the system states are governed by the same laws as the ones of the embedded DTMC; however, the sojourn times must be accounted for in order to derive the steady-state probabilities. In particular, let $\mathbf{h}$ denote the vector that defines the number of steps that the system stays in each of the l states.[4] For any collision state $s \in [1, (RL+1)^2]$ one gets $h_s = 1$. For the successful states, the number of consecutive times that e.g. station A manages to transmit successfully (assuming that the state of the DTMC is

[4]Here it is implied that the system states are enumerated sequentially from 1 to l in the following order: $(Coll, 0, 0)$,$(Coll, 0, 1)$,...,$(Coll, 1, 0)$,...,$(Coll, RL, RL)$,$(Succ, 0, 0)$,...,
$(Succ, 0, RL)$,$(Succ, 1, 0)$,...,$(Succ, RL, 0)$.

$(Succ, 0, j)$), conditioned on the backoff stage of A prior to the success being i, is

$$h_s|i = \sum_{k=0}^{N-1} \prod_{n=0}^{k} q_n(0,j)|i. \tag{5.37}$$

The unconditional number of steps h_s that the system stays in a state $s \in [(RL+1)^2+1, (RL+1)^2 + RL + 1]$, i.e. of the form $(Succ, 0, j)$, is

$$h_s = \frac{\sum_{i=0}^{RL} \gamma(Coll, i, j) \cdot h_s|i}{\sum_{i=0}^{RL} \gamma(Coll, i, j)}. \tag{5.38}$$

After finding the vector $\mathbf{h}$, the steady-state probabilities are given by:

$$\gamma_i = \frac{v_i \cdot h_i}{\sum_{j=1}^{l} v_j \cdot h_j}, \; i \in [1, l]. \tag{5.39}$$

The conditional collision probability p' is calculated as the sum of the steady-state probabilities of all the collision states

$$p' = \sum_{j=1}^{(RL+1)^2} \gamma_j. \tag{5.40}$$

As far as the backoff-dwelling probabilities ξ_i are concerned, one is interested in the system states from a transmitter's point of view. This means that, if, without loss of generality, station A is chosen, the system states $(Succ, i, 0)$, $i > 0$ should be excluded (these are not transmitting states of A). Due to the flexibility of the proposed framework this can be done easily by considering the corresponding sojourn times to be zero. All the other sojourn times as well as the transition matrix P are the same as before. Thus, if one denotes the sojourn time vector and the steady-state distribution vector from a transmitter's point of view as $\hat{\mathbf{h}}$ and $\hat{\gamma}$ respectively, then similarly

$$\hat{\gamma}_i = \frac{v_i \cdot \hat{h}_i}{\sum_{j=1}^{l} v_j \cdot \hat{h}_j}, \; i \in [1, l], \tag{5.41}$$

where $\hat{\mathbf{h}}$ differs from $\mathbf{h}$ only in that the sojourn times of the states $(Succ, i, 0)$, $i \in [1, RL]$ are zero.

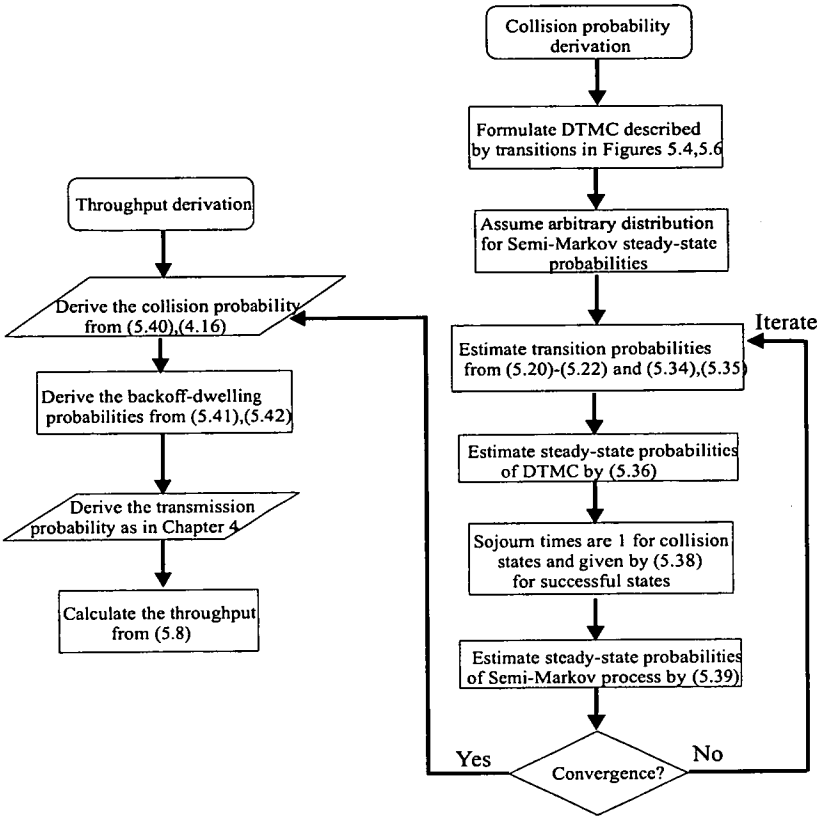


Figure 5.9: *This flow chart provides an overview of the analytical method followed, in order to derive the saturation throughput of a hidden terminal that employs BEB.*

Finally, in order to find the steady-state probability ξ_i that transmitter A, for instance, is in the backoff stage i, one simply has to add the previous steady-state probabilities that correspond to all the possible states where the stage of A is i. Thus,

$$\begin{aligned} \xi_i &= \sum_{j=0}^{RL} \hat{\gamma}(Coll, i, j), \quad i \in [1, RL] \\ \xi_0 &= \sum_{j=0}^{RL} \hat{\gamma}(Coll, 0, j) + \sum_{j=0}^{RL} \hat{\gamma}(Succ, 0, j). \end{aligned} \tag{5.42}$$

The process followed for the derivation of the collision probability, the backoff-dwelling distribution and, ultimately, the station throughput is summarised in Figure 5.9.

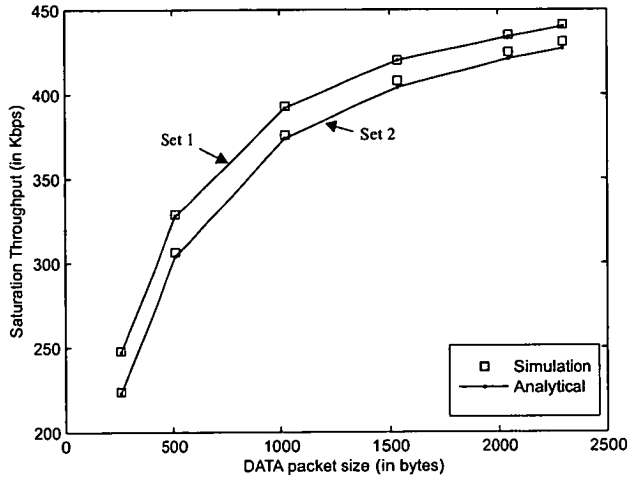


Figure 5.10: *The throughput is plotted against the DATA packet size for two different sets of parameters of the BEB. Set 1 corresponds to a maximum backoff counter value of 1024 and RL equal to 6, whereas Set 2 corresponds to a maximum counter value of 256 and RL equal to 4. For both sets of parameters the analytical model has a very good match to the simulation results. Both transmission rates are set to 1 Mbps.*

5.4 Validation of the model

In order to validate the accuracy of the proposed analytical model, it is essential to compare the above, theoretically obtained values with their corresponding counterparts, measured from simulation runs in NS2. Although only a subset of the results is presented, the method shows a consistently high agreement to the simulation results as far as throughput, collision probability and backoff-dwelling distribution are concerned for all system parameters examined. The considered topology is the classic hidden terminal scenario presented in Chapter 4 (Figure 4.1), where stations A and B are continuously backlogged. DATA packets of size $L_d = \{256, 512, 1024, 1536, 2048, 2294\}$ bytes are used. The access scheme used in this section is RTS/CTS; however, a straightforward extension to the Basic Access scheme (see Section 4.3.3) is possible. Each simulation point in the graph corresponds to the average of the corresponding metric (e.g. throughput, backoff-dwelling probability, collision probability) over 10 different trials for $200s$ simulation time.

In Figure 5.10 the throughput of one of the transmitters is plotted against the packet size. The parameters used are $RL = 6$, $W_0 = 32$ and $m = 5$ for the parameter Set 1 and $RL = 4$, $W_0 = 32$ and $m = 3$ for the parameter Set 2. Both transmission rates (BasicRate and DataRate)

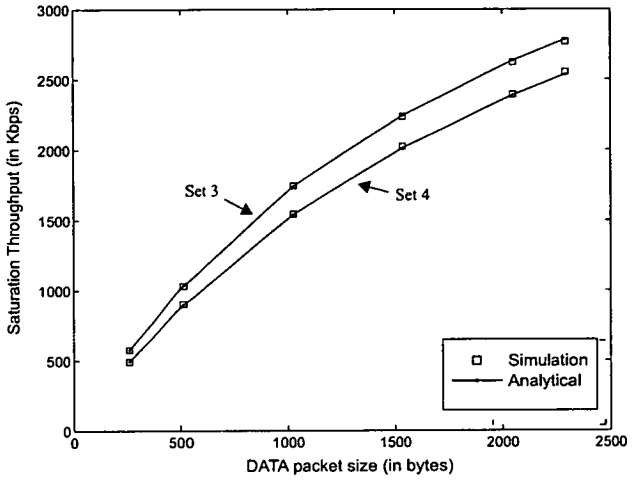
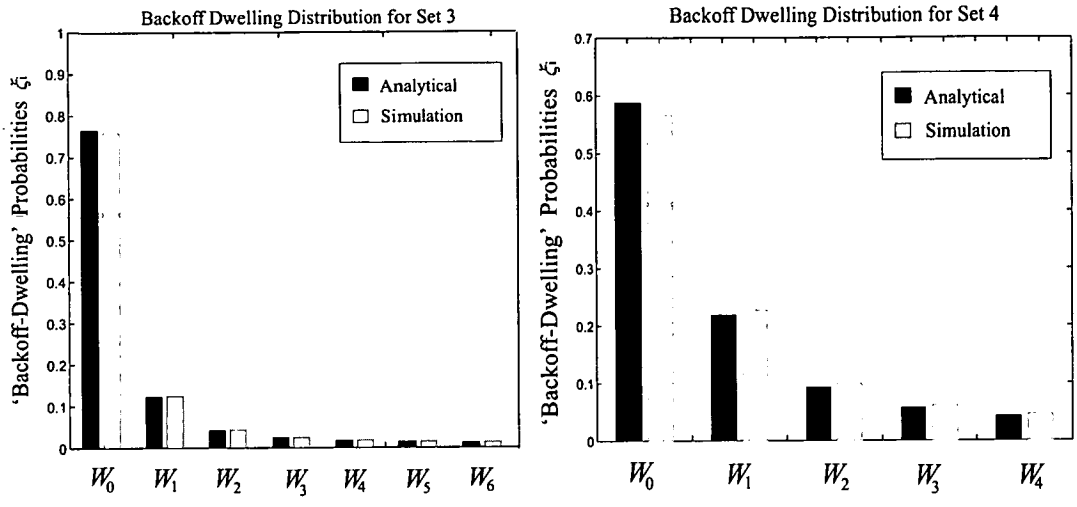


Figure 5.11: *The throughput is plotted against the DATA packet size for 2 Mbps BasicRate and 11 Mbps DataRate. Set 3 corresponds to a maximum backoff counter value of 1024 and RL equal to 6, whereas Set 4 corresponds to a maximum counter value of 256 and RL equal to 4.*

are set to 1 Mbps for this graph. Due to the obvious symmetry of the considered topology and the assumption that, both transmitters are characterised by the same protocol parameters, they share an equal part of the bandwidth in the long-term, as can be verified by the simulations. Furthermore, the throughput (expressed in Kbps) is an increasing function of the DATA packet size. One can also observe that in the case (parameter Set 2) when the nodes' backoff times are distributed over shorter intervals, the probability of them making simultaneous transmissions during the vulnerable period, so that collision is caused, increases. This has a direct decreasing effect on the station throughput. As far as Figure 5.11 is concerned, the parameters used are the same as before, with the exception of the transmission rates, which are 2 Mbps for the BasicRate and 11 Mbps for the DataRate. The model proves to be consistently accurate, when compared to the simulation results, for any combination of the parameters. As a sidenote, it can be observed that the throughput of the topology under consideration greatly improves when higher transmission rates are used. This is due to two main reasons: primarily, with the increased rates the time of packet transmission is considerably reduced, and that has an increasing effect on the station throughput, and, secondly, the vulnerable period, which is in this example dependent on the transmission time of the RTS packet, is reduced, causing less collisions and, thus, improving the throughput of both transmitters.

Taking a microscopic view of the channel contention with the examination of the backoff-



(a) The parameters used are $RL = 6$ and $W_{max} = 1024$.

(b) The parameters used are $RL = 4$ and $W_{max} = 256$.

Figure 5.12: *The Backoff-dwelling distribution for the parameter Sets 3 and 4. Note that a BasicRate of 2 Mbps and a DataRate of 11 Mbps are used for this figure. The model is consistent with the simulation results for both cases.*

dwelling distributions, a few interesting conclusions can be drawn. In particular, in Figures 5.12(a) and 5.12(b) the backoff-dwelling probability is plotted for each backoff stage for the parameter Sets 3 and 4 correspondingly. It is evident that the proportion of time spent in the different backoff stages tends to be more equally distributed for the Set 4, where the backoff counter only reaches 256. In other words, the stations are more likely to be found in the initial backoff stage (remember that a station goes in the backoff stage W_0 after a successful transmission) for 'wider' backoff distributions. This observation agrees with the results shown in Figure 5.11, where the throughput of the Set 3 is higher than that of the Set 4, even though the time period when the channel remains idle is bigger for the parameter Set 3. Finally, a very good match of the theoretically obtained probabilities to the simulation measured probabilities is evident for both cases.

Moreover, the conditional collision probabilities p_i obtained by (5.2) are compared with the ones measured in the simulations for the parameter Set 1. Again, the proposed model proves to be highly accurate. In Figure 5.13 one can confirm the observation made in Section 5.1, i.e. that the probability of collision becomes higher as the node enters into higher backoff stages in the hidden terminal topology. This justifies the use of distinct collision probabilities for the

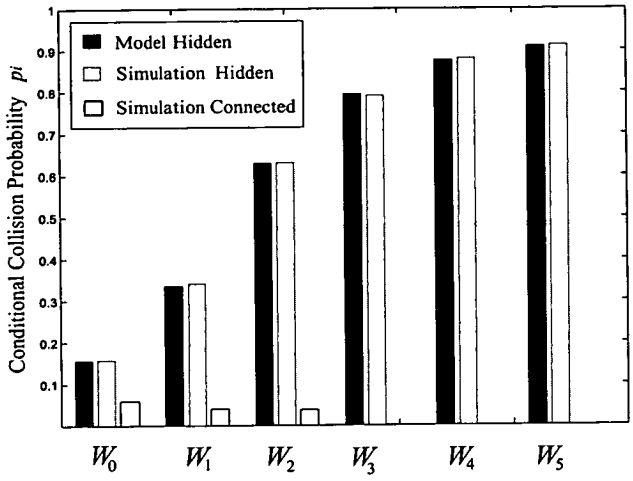


Figure 5.13: *The collision probabilities p_i vs. the backoff stage, as calculated by the analytical model for the hidden node topology, and as measured in the simulations for both hidden and connected topologies. The station of the connected topology never goes in the higher backoff stages, and, thus, the respective p_i's are not defined.*

different backoff stages of the station in the proposed model. To give more insight into this the respective collision probabilities of a fully-connected topology (stations A and B can 'hear' each other) are also plotted in Figure 5.13, keeping all the other parameters of the scenario identical. Two key observations can be made: Firstly, the conditional collision probability on average is higher in the hidden node topology than in the fully-connected topology. Actually, it was noticed in the simulations that, as far as the latter is concerned, the transmitter never goes in the higher backoff stages. Secondly, the assumption of [2] that in fully-connected networks the conditional collision probability can be considered constant and independent of the backoff stage is very close to the truth. However, this is clearly not the case in the hidden node topology[5].

5.5 A stricter but computationally inefficient mathematical solution

The consideration of a Semi-Markov chain in the above analysis and the adoption of an iterative technique in order to find the steady-state probabilities, and, subsequently, the metrics of inter-

[5]This can be noticed if the average conditional collision probability is calculated as the weighted average of the p_is for both cases (the weights are the ξ_is).

est, were necessary because of the choice of the fundamental system state as a triple (E, i, j). The first degree of freedom of the state, event E, has only two possible values, collision or success, but it is the one that forces us to exclude the self-transitions in order to preserve the Markovian characteristics of the process.

5.5.1 Introduction of the model

A stricter mathematical solution to the problem is the description of the channel contention state as a triple $(\Delta C, i, j)$, where i, j are the backoff stages of the two transmitters A and B respectively, but ΔC is the difference of the backoff counters at the time instant when one of the senders attempts transmission. In particular, if C_B (C_A) is the value of the backoff counter of B (A), then ΔC is defined as follows: $\Delta C = C_B - C_A$. This definition describes the process in detail, since collision is translated into the value of ΔC being in the interval $[-c, +c]$, where c is half the vulnerable period in slots, successful transmission of A is translated into $\Delta C > c$ and successful transmission of B is translated in $\Delta C < -c$. The transition probabilities can be derived in a similar way as has already been described in this thesis. An analytical derivation for the former, as well as a more detailed description of the model can be found in [74]. The definition and formulation of this model is not done by the author, but by Dr. Konstantinos Drakakis; however, the analysis presented in the following section that derives several metrics of interest, and the relevant results presented afterwards, are the author's contributions.

5.5.2 Derivation of metrics of interest

The collision probability as seen by the receiver, p', can be easily calculated by the model of Section 5.5.1, since p' is, by definition, the sum of the probabilities of all states leading to a collision, namely:

$$p' = \sum_{i=0}^{RL} \sum_{j=0}^{RL} \sum_{\Delta C=-c}^{c} \pi(\Delta C, i, j). \tag{5.43}$$

Moreover, the packet loss ratio, denoted as plr, can also be derived from the model. The packet loss ratio is the ratio of packets that are dropped after they have been retransmitted for RL times, over all the packets that were actually sent by the MAC layer. Equivalently, the packet loss ratio can be viewed as the probability that a transmitted packet is dropped. Assuming, without loss of generality, that the computation is carried out for node A, this is the probability

that the system reaches a state where $i = RL$ and $|\Delta C| \leq c$ over the aggregate probability of all the states the system can reach, while A attempts to transmit the packet (successfully or unsuccessfully). Therefore, the packet loss ratio is:

$$plr = \frac{\sum_{j=0}^{RL} \sum_{\Delta C=-c}^{c} \pi(\Delta C, RL, j)}{\sum_{j=0}^{RL} \sum_{\Delta C=-c}^{c} \pi(\Delta C, RL, j) + \sum_{j=0}^{RL} \sum_{i=0}^{RL} \sum_{\Delta C=c+1}^{W_j+c} \pi(\Delta C, i, j)}. \quad (5.44)$$

As far as the delay-sensitive applications are concerned (real-time traffic), the distribution of the delay is also of primary importance. Let's define $E[D]$ as the expectation of the delay that a transmitted packet experiences until it is sent successfully to its recipient, or it is ultimately dropped by its sender. In particular, focusing again without loss of generality on station A, the first moment of the delay can be viewed as the average transition time between an 'initial' state $(\Delta C, 0, j)$, where ΔC can take any allowed value (this is the state of the chain that describes the arrival of a new packet for node A), and a 'final' state, that is either of the form $(\Delta C, i, j)$, where $\Delta C > c$ (this is the state of the chain that describes a successful transmission by A) or of the form $(\Delta C, RL, j)$, where $|\Delta C| \leq c$ (this is the state where A has already transmitted this packet RL times and since collision occurs again, it has to drop the packet). Note that in all the previous expressions the backoff stage j of B can take any allowed value, i.e. $0 \leq j \leq RL$. Let I denote the set that comprises the 'initial states', and Θ denote the set of the final states.

Similar mathematical problems are solved using the so-called 'time to absorption' [70, 75]. Using this formulation, one decomposes the state space Ω to a class of transient states T and a class of absorbing states Γ, such that $\Omega = \Gamma \cup T$. The expected value of the time to absorption is the average duration between the time instant when the system is in a transient state of the system (belonging to class T), and the time instant when it exits from the class T of transient states. Clearly, in the suggested model the class of absorbing states coincides with what was defined as set Θ, and the class of transient states with $\bar{\Theta}$. For the derivation of the expected time to absorption the substochastic matrix $\mathbf{Q}$ is necessary, which is a square matrix whose elements q_{kn} are the transition probabilities from state k to state n, where both k and n belong to the set $\bar{\Theta}$. Given a vector $\mathbf{R} = (r_k)$, whose elements r_k are a reward that is assigned to the state k of the set $\bar{\Theta}$, the expected value of the aggregate reward until absorption for the state $k, k \in \bar{\Theta}$ is

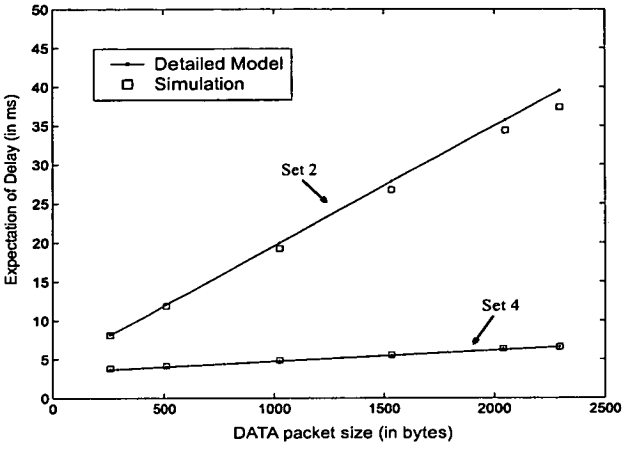


Figure 5.14: *The expected value of the delay $E[D]$ is plotted against the data packet size for the parameter Sets 2 and 4.*

given by:

$$\mathbf{AR} = (ar_k) = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}, \tag{5.45}$$

where $\mathbf{I}$ is the identity matrix.

In the present application the reward vector $\mathbf{R}$ contains the average time duration of an arbitrary state $k \equiv (\Delta C, i, j)$, i.e.

$$r_k = \begin{cases} F & , \Delta C < -c \\ C + (W_i - 1)/2 & , |\Delta C| \leq c \\ L + (W_i - 1)/2 & , \Delta C > c \end{cases}$$

Since one is interested in the expected value of the delay from an initial state belonging to the set I to an absorbing state of Θ, the desired metric can be obtained using the sum for all states k that belong to the set I, weighted by their respective steady-state probabilities as follows:

$$E[D] = \frac{\sum_{k \in I} \pi(k) a r_k}{\sum_{k \in I} \pi(k)}. \tag{5.46}$$

Note that although the steady state probabilities of the MC are independent of the data packet size and so are the collision probability and the packet loss ratio, the expected value of the delay is strongly dependent on the latter parameter.

		Detailed Model	Iterative Method	Simulation
p	Set 1	0.2548	0.2566	0.2566
p	Set 2	0.5009	0.5043	0.5020
p	Set 3	0.2495	0.2510	0.2503
p	Set 4	0.4632	0.4663	0.4660
plr	Set 1	1.95%	1.92%	1.94%
plr	Set 2	8.66%	8.35%	8.88%
plr	Set 3	1.38%	1.34%	1.37%
plr	Set 4	6.06%	5.62%	6.16%

Table 5.1: *Comparison between the mathematical models (both the detailed and the iteration-based) and the simulation results for the collision probability and the packet loss ratio.*

In order to validate the model presented in this section, performance results are derived for the three performance indices defined above. Furthermore, the theoretical results are compared with the simulation results obtained from NS2. A very good match to the simulation results is evident for all the metrics considered. Figure 5.14 depicts the expected value of the transmission delay $E[D]$ with respect to the data packet size for the parameter Sets 2 and 4. These sets have identical backoff related parameters, and are different only in the transmission rates. As expected, the average time for a completed transmission (either a successful or a failed one) is much higher for the case of lower transmission rates, both because of the longer packet transmission times and also because of more frequent collisions. An almost linear relation of the mean value of the delay with the packet size can be observed.

Furthermore, Table 5.1 presents the two other performance indices, namely the collision probability and the packet loss ratio. It can be observed that the detailed model (i.e. the one presented in Section 5.5) proves to be a very good match with the simulation results for both metrics calculated. As a sidenote, one can observe that, with the transmission rates kept constant, the packet loss ratio and the collision probability are lower for the cases of 'wider' backoff distributions (parameter Sets 1 and 3), as expected.

However, the complexity of the detailed model presented in this section is very high, as the state space grows exponentially with the number of backoff stages. In particular, the dimension of the square transition matrix of the DTMC (l^*, l^*) is given by the formula:

$$l^* = \sum_{i=0}^{RL} \sum_{j=0}^{RL} [W_i + W_j + 2c - 1]. \tag{5.47}$$

As an example, it is mentioned that for the MC that gives the solution for the parameter Set 2, the dimension is $l^* = 8035$ and for the Set 1, $l^* = 44275$. Thus, it is evident that the state space can easily reach hundreds of thousands of states or more, misusing both memory and processing time. As a comparison, the reader is reminded that the iterative method allows the dimension of the state space to be only $l = 34$ and $l = 62$, respectively, for the parameters considered above. The reduction in the processing time is considerable as well: it goes from several hours down to several seconds. If this fact is taken into consideration along with the negligible differences in the performance results (see Table 5.1), the conclusion is that the iterative solution is preferable for a real application.

5.6 Conclusions

The present chapter had the aim of showing that the analytical model presented in Chapter 4 is extendable to the BEB and, subsequently, to an arbitrary backoff distribution, with a suitable change of definition of the basic parameters and states that define the channel contention. It was shown that with the aid of an iterative method similar to the one presented in Chapter 4, the problem can be solved quite efficiently and with a high degree of accuracy. The model no longer assumes a constant collision probability throughout the backoff process, as is the usual assumption in a fully-connected network analysis, and this modelling approach is justified appropriately. A detailed Markov model is also presented, which does not rely on an iterative method, but, although it is highly accurate, it lacks computational efficiency and, thus, only serves as a reference point for the iterative method.

Chapter 6
Conclusions

This thesis has contributed to the development of suitable mathematical modelling methods for the analysis of the performance of a wireless, random access based network. In fact, the proposed models lie on either side of the line between high modelling accuracy and scalability. This concluding chapter will give a summary of the main methods and results presented in the different chapters and, hopefully, serve as a motivation for further research.

6.1 Summary of results

The problem of finding a mathematical model that accurately describes the channel contention in a CSMA network, and is at the same time sufficiently general and scalable, has concerned the research community during the last three decades. Several solutions have been proposed, each incorporating its own assumptions, and each attempting to grasp different aspects of the complex phenomena that take place when two or more stations of the same neighbourhood contend for the occupancy of the wireless medium. Hopefully, this thesis has added a small contribution to this continuing process by proposing two different methods of approach, each lying on either side of the line between high accuracy and scalability.

Until recently, research papers were dealing with interferers in a wireless network sporadically, usually considering as the only example the classic hidden node problem, or the exposed node problem. Realising this gap, the first part of the thesis was devoted to the identification and the systematic analysis of all the possible topologies that can cause interference to an ongoing transmission, when the nodes of interest employ a CSMA MAC protocol. It was observed that the effect that different interferers have can vary quite heavily, and is highly dependent on whether the transmitter or the receiver is in range of the interferer, and, also, on the time that the interfering transmission started. Thus, there are topologies that can cause very minor problems to the ongoing transmission or can even occur in parallel to the latter, others that force the transmitter to defer for a certain period of time, and, finally, others that lead to collisions. Thus, a systematic description of the respective collision probabilities is performed, and, then, a

simplified model is developed that attempts to incorporate the effect of all these different types of interferers, under the assumption of a single interfering pair at any one time. Later in the thesis, this assumption is lifted and a corrected version of the model is developed, where it is included that the interferers themselves may not be successful in their attempt.

However, the necessity to keep the model tractable, since it is desirable to keep it as much general as possible, forces the adoption of the following assumption: Every interferer of the pair under consideration is assumed to have the same transmission and collision probability as the pair itself. This hypothesis does not seriously affect the accuracy of the model when the connectivity graph of the network is symmetrical, e.g. in a star topology or in the central part of a grid topology. However, in a more general case, where there are significant boundary effects, the model fails to show high accuracy, managing only to follow the performance trends. Moreover, the independence assumption that is also adopted, as it is the only way to solve a network consisting of many nodes, has an impact on the results when random networks are examined. Finally, even in entirely symmetrical networks, there are sets of parameters that make the model more or less consistent when compared to simulation results.

The latter observations, which were also verified when reviewing and implementing several other modelling methods, served as the motivation for the main contribution of this thesis, presented in the Chapters 4 and 5. In particular, there it is explained why traditional models, that borrow tools suitable for a fully-connected network, fail to give a consistent and independent of the transmission parameters performance. Such assumptions are the well known decoupling approximation, the description of the channel evolution as a regenerative process, and the hypothesis that the first central moment of the backoff distribution is always bigger than the vulnerable period.

On the other hand, a novel technique to model the channel time is introduced, which does not rely on the former assumptions. On the contrary, it is based on a constant and objective for all the stations unit of time, i.e. the protocol defined channel slot, and redefines the transmission probability according to this new timescale. Moreover, the complexity of the interfering scenario is grasped via the examination of the joint behaviour of the contending stations, and also via the use of random variables for a detailed analysis of several crucial parameters. The case of an arbitrary length of the packet that initiates the communication handshake is addressed. Finally, it is shown that the process that describes the channel evolution is not regenerative, but is characterised by a Markov dependence. Based on this framework, expressions for the

conditional collision probability and the station throughput are derived in the saturated case, presuming that the nodes employ a constant contention window for the backoff process.

The latter assumption is later relaxed, and the model is extended to describe the BEB mechanism proposed by the IEEE 802.11 protocol, or, similarly, an arbitrary distribution of the backoff counters. The system is modelled with the aid of a Discrete Time Markov Chain embedding the channel evolution and an accompanying Semi-Markov process. The Markovian state space is kept relatively small with the adoption of an iterative technique, which, independently of the initial estimates, converges to the correct steady state distribution. The iterative technique is finally compared to an exact model of the channel evolution; however, it is observed that the former performs almost equally well with the detailed model and has a several orders of magnitude lower computational complexity. Thus, it is preferable for a real system application.

6.2 Suggestions for an optimal configuration of IEEE 802.11 stations

In this section, some useful suggestions to the system designer will be outlined, that have arisen from the insights that the proposed analytical framework provided[1]. In particular, an optimal value of contention window exists for the RTS/CTS scheme, regardless of the data packet size. The value for the configuration considered in this thesis is around 128; however, it is suspected that, when multiple hidden nodes exist, this optimal value will increase, due to more frequent collisions.

Moreover, RTS/CTS seems to be preferable over Basic Access, when low contention window sizes are to be used, e.g. for delay-sensitive applications, such as voice or video streaming. On the contrary, for large contention windows and high data rates, Basic Access outperforms RTS/CTS, especially for small packets. Regarding the data rates that should be selected, it is suggested that, when low rates are going to be used—the latter are preferred when the physical channel is unreliable, due to the greater redundancy that the modulation techniques of low rates exhibit—Basic Access should not be adopted for almost any data packet size, since it is almost impossible to avoid collisions, even with a high contention window. RTS/CTS is a far more efficient option in these situations. Furthermore, if Basic Access is to be used, the data flows

[1]The suggestions refer to networks where hidden terminals exist and cannot be applied directly to a fully-connected network, since in the latter, the behaviour may follow different rules [2].

should preferably be fragmented into smaller packets, as the latter can be transmitted more efficiently; however, for RTS/CTS, as the vulnerable period is not dependent on the data packet size, it can be verified that, the longer the data packet, the more the information bits that are delivered successfully.

Finally, as far as the BEB is concerned, it is observed that, for low transmission rates, a relatively 'wide' BEB outperforms even an optimal contention window size for the RTS/CTS scheme, as its dynamic adjustment to the channel contention seems to improve the overall throughput, whereas, for the higher transmission rates, the advantage of using the BEB is very small and is gained at the expense of the short-term fairness. If the parameters of the BEB (i.e. the retry limit and the minimum/maximum contention windows) are adjusted, so that the backoff distribution becomes more narrow, the above behaviour is alleviated; the optimal contention window performs equally well and, also, the fairness is improved.

6.3 Future research areas

There are several areas of this thesis that can be extended through future research. These research directions are outlined below:

- The model developed in the first part of the thesis, that attempts to calculate the throughput in an entire network, suffers from pessimistic results, as can easily be observed in the relevant graphs. This is due to the simplifying assumption that to keep the model scalable, it was presumed that all the neighbouring stations have identical fundamental transmission parameters, i.e. transmission and collision probability. A more complex model, incorporating the fact that some nodes may be in a better position, because they may lie for instance in the border of the network graph, seems an appropriate field for further research for networks of finite and small size.

- The models presented in the Chapters 4 and 5 may, quite straightforwardly, derive expressions for the average delay of a successfully transmitted packet (excluding the queueing delay at the station's buffer). This is easy to compute in the proposed framework, since the delay comprises a number of steps, i.e retransmissions, experienced by the packet until its successful reception. The number of these steps may be 1,2,...,$RL + 1$; therefore, since in both the CCW and also the BEB case, the respective collision probabilities, and, consequently, the probabilities of success can be computed by the models, the final

delay is simply the sum of the delays incurred in each step (C for a collision step and L for a successful step, added to an expected backoff time), weighted by the respective probabilities of the packet entering the kth step, i.e. the packet being retransmitted for the kth time.

- The proposed modelling approach regarding the classic hidden terminal topology case manages to show very high accuracy and consistency of performance evaluation under all the considered parameters. Moreover, deep insights into the channel contention are provided. However, it still remains a challenging research problem to extend the introduced method to a multiple interferers case. Due to the expected exponential growth of the state space, it is a real challenge to identify some key points where certain assumptions can be introduced, not 'light-heartedly', but in a consistent and quantifiable manner.

- The present thesis adopts the assumption that the stations are working with their buffers non-empty at all times, i.e. in the saturated region. An extension of the model to include some form of packet arrival process is desirable as a future research direction. Moreover, the cases where the transmission and the carrier-sensing range are different, or where they are not perfect circles (more realistic physical layer modelling) would also be quite interesting.

Appendix A
Detailed Description of the Iterative Method

This appendix will elaborate on the details of the iterative method described above and give the corresponding equations. Throughout the appendix the term *combination* will be used, so it would be useful to explain exactly what is meant with this term. Suppose that a random experiment with two possible outcomes, E and $\overline{E}$, which occur with corresponding probabilities $P\{E\}$ and $1 - P\{E\}$ is performed. After each trial, a RV I is observed, which follows different distributions according to the result of the trial. Suppose that F_I^E ($F_I^{\overline{E}}$), defined over a set S_E ($S_{\overline{E}}$), represents the Probability Density Function (PDF), if one refers to continuous RVs, or the PMF, if one refers to DRVs, of the RV I, conditioned that the outcome of the experiment is E ($\overline{E}$). According to probability theory [72], the unconditional PMF (or PDF) of I is given by the following equation:

$$F_I(x) = P\{E\} \cdot F_I^E(x) + (1 - P\{E\}) \cdot F_I^{\overline{E}}(x), \quad x \in S_E \cup S_{\overline{E}}. \tag{A.1}$$

Note that in the above equation $F_I^E(x) = 0$ for $x \in S_E \cup S_{\overline{E}} - S_E$. This definition is inferred each time the author refers to the *combination* of two RVs. Symbolically the notation $A \oplus B$ is used for the *combination* of RVs A, B, where e.g. $A = I|E$ and $B = I|\overline{E}$.

The case described in Figure 4.11(a), i.e. collision followed by collision, provides the probability that collision occurs, conditioned on collision preceding, so as mentioned before

$$p_{cc} = P\{-c \leq Y_{n-1} + X_B - X_A \leq c\}. \tag{A.2}$$

Furthermore, the corresponding conditional distribution Y_n^c is equal to

$$Y_n^c = Y_{n-1} + X_B - X_A | -c \leq Y_{n-1} + X_B - X_A \leq c. \tag{A.3}$$

When the case examined is collision followed by success (Figure 4.11(b)), two occurrences are possible. In the first, transmitter A is successful. The probability p_{cs}^A that A is successful,

conditioned on collision preceding, is

$$p_{cs}^{A} = P\{X_B > X_A + c - Y_{n-1}\}. \tag{A.4}$$

Also, the consecutive distribution of the remaining backoff slots for the sender B, after it froze its counter because of the successful transmission of A, is given by $Z_n^{A,c} = X_B - (X_A + c - Y_{n-1}) \,|X_B > X_A + c - Y_{n-1}$. On the contrary, if B is successful conditioned on collision occurring, then the corresponding probability would be

$$p_{cs}^{B} = P\{X_A > X_B + c + Y_{n-1}\}, \tag{A.5}$$

and the distribution of remaining backoff slots would be $Z_n^{B,c} = X_A - (X_B + c + Y_{n-1}) \,|X_A > X_B + c + Y_{n-1}$. The ultimate DRV which expresses the distribution of remaining backoff slots of the freezing station, conditioned on collision preceding, and independently of which transmitter is successful, is given by

$$Z_n^c = Z_n^{A,c} \oplus Z_n^{B,c}, \tag{A.6}$$

with corresponding weights $p_{cs}^{A}/\left(p_{cs}^{A} + p_{cs}^{B}\right)$ and $p_{cs}^{B}/\left(p_{cs}^{A} + p_{cs}^{B}\right)$.

Following the same method, one can resume the analysis with the case when success is followed by collision (Figure 4.11(c)). Supposing that the transmitter who was successful in the state s_{n-1} was A (this occurs with probability 1/2 as in this scenario, as was mentioned already, the two transmitters share an equivalent part of the bandwidth in the long-term), then the probability that collision follows, conditioned on A being successful previously, p_{sc}^{A}, is obtained by the following equation:

$$p_{sc}^{A} = P\{-c \leq Z_{n-1} - X_A \leq c\}. \tag{A.7}$$

Furthermore, the distribution of time differences among RTS packets, conditioned on success of A preceding, is $Y_n^{A,s} = Z_{n-1} - X_A| - c \leq Z_{n-1} - X_A \leq c$. On the contrary, if B was successful in s_{n-1} (probability 1/2), then:

$$p_{sc}^{B} = P\{-c \leq X_B - Z_{n-1} \leq c\}, \tag{A.8}$$

and $Y_n^{B,s} = X_B - Z_{n-1}| - c \leq X_B - Z_{n-1} \leq c$. As a consequence, the transition probability p_{sc} is given by

$$p_{sc} = \frac{1}{2} \cdot p_{sc}^{A} + \frac{1}{2} \cdot p_{sc}^{B}, \tag{A.9}$$

and the DRV Y_n^s, as defined in Section 4.3.2, is derived as

$$Y_n^s = Y_n^{A,s} \oplus Y_n^{B,s}, \tag{A.10}$$

with both weights being equal to $1/2$.

Finally, concerning the case where success is followed by success (Figure 4.11(d)), two different cases can be distinguished, i.e. BB and BA, using the notation that was introduced previously (the other two possible cases AA and AB are very similar). Between the two scenarios, BA occurs with probability

$$p_{ss}^{BA} = P\{X_B > Z_{n-1} + c\}, \tag{A.11}$$

and the corresponding DRV depicting the remaining backoff slots, conditioned on BA occurring, is $Z_n^{BA,s} = X_B - (Z_{n-1} + c) \,|X_B > Z_{n-1} + c$. Similarly, when the scenario under consideration is BB, which happens with probability

$$p_{ss}^{BB} = P\{Z_{n-1} > X_B + c\}, \tag{A.12}$$

the remaining backoff slots are given by the DRV $Z_n^{BB,s} = Z_{n-1} - (X_B + c) \,|Z_{n-1} > X_B + c$. Consequently the DRV describing the remaining backoff slots of a sender after freezing, conditioned that in the state s_{n-1} success had also occurred, is

$$Z_n^s = Z_n^{BA,s} \oplus Z_n^{BB,s}, \tag{A.13}$$

with corresponding weights $p_{ss}^{BA} / \left(p_{ss}^{BA} + p_{ss}^{BB}\right)$ and $p_{ss}^{BB} / \left(p_{ss}^{BA} + p_{ss}^{BB}\right)$.

After obtaining expressions for the conditional distributions Y_n^c, Y_n^s and Z_n^c, Z_n^s, the ultimate aim is to derive the corresponding distributions Y_n and Z_n independently of the previous state s_{n-1}. Firstly, as far as Y_n is concerned, the notion of *combination* will be applied, as defined at the beginning of Appendix A. Particularly for this case, I is the distribution of RTS differences, given that collision follows immediately afterwards, and the event E ($\overline{E}$) is the event of collision (success) in the state s_{n-1}, conditioned on collision following in s_n (let's denote the respective

probabilities as $P\{Coll_{n-1}|Coll_n\}$ and $P\{Succ_{n-1}|Coll_n\}$)[1]. Thus:

$$Y_n = Y_n^c \oplus Y_n^s. \tag{A.14}$$

As was mentioned before, the corresponding weights are obtained as posterior probabilities according to Bayes' Theorem as follows:

$$\begin{aligned} P\{Coll_{n-1}|Coll_n\} \cdot P\{Coll_n\} &= P\{Coll_n|Coll_{n-1}\} \cdot P\{Coll_{n-1}\}, \\ P\{Coll_{n-1}|Coll_n\} &= p_{cc}, \end{aligned} \tag{A.15}$$

because in the steady state $P\{Coll_n\} = P\{Coll_{n-1}\}$, and also from the definition of p_{cc} follows that $P\{Coll_n|Coll_{n-1}\} = p_{cc}$. Similarly, for the second weight one can get:

$$\begin{aligned} P\{Succ_{n-1}|Coll_n\} \cdot P\{Coll_n\} &= P\{Coll_n|Succ_{n-1}\} \cdot P\{Succ_{n-1}\}, \\ P\{Succ_{n-1}|Coll_n\} &= 1 - p_{cc}, \end{aligned} \tag{A.16}$$

(A.16) is derived because in the steady-state and from the definition of p', the following equations hold: $P\{Succ_{n-1}\} = 1-p'$ and $P\{Coll_n\} = p'$. It is also true that $P\{Coll_n|Succ_{n-1}\} = p_{sc}$, so with the help of (4.17) and after some algebra, (A.16) is derived. Equivalently, the DRV Z_n is obtained as

$$Z_n = Z_n^c \oplus Z_n^s. \tag{A.17}$$

with corresponding weights

$$P\{Coll_{n-1}|Succ_n\} = p_{sc}, \tag{A.18}$$

and

$$P\{Succ_{n-1}|Succ_n\} = 1 - p_{sc}. \tag{A.19}$$

[1]Note that here posterior probabilities are used. The expression 'conditioned on collision following', even though it may seem awkward, is very useful in the probability theory and uses information about later events (in general not necessarily later in time, but later in the logical chain, although in our case these notions coincide) to determine the conditional probability of prior events. For a very interesting simple example the reader is referred to paragraphs 3.2 and 4.1 of [72].

Appendix B
Derivation of the Steady-State Probabilities of the Markov Chain

This appendix provides the analysis of the chain of arguments that yield the steady-state probabilities of the Markov Chain presented in Figure 4.7. The same procedure is followed for the derivation of the steady-state probabilities of the MC of Figure 5.3, thus, it will not be repeated here explicitly. From the balance equations of the MC we have the following relations:

$$\pi_m^1 = \pi_m^2 = ... = \pi_m^F \, , m \in [1, W-1], \tag{B.1}$$

because the respective transition probabilities are all equal to 1. It is also true that

$$\pi_m^1 = p_f \pi_m \, , m \in [1, W-1], \tag{B.2}$$

as the station goes into the first slot of a sequence of freezing slots with probability p_f. Moreover,

$$\begin{aligned} \frac{1}{W}\pi_0 + (1-p_f)\pi_m + \pi_m^F &= \pi_{m-1} \Rightarrow \\ \frac{1}{W}\pi_0 + (1-p_f)\pi_m + p_f\pi_m &= \pi_{m-1} \Rightarrow \\ \frac{1}{W}\pi_0 + \pi_m &= \pi_{m-1}, \end{aligned} \tag{B.3}$$

where $m \in [1, W-1]$. After some simple algebra one has:

$$\pi_m = \frac{W-m}{W} \cdot \pi_0 \, , m \in [1, W-1]. \tag{B.4}$$

The normalisation condition (the sum of all the steady-state probabilities of the ergodic MC is equal to 1) is then applied, so that one more independent equation is obtained and, thus, the above system of equations can be solved, i.e.

$$\sum_{m=0}^{W-1} \pi_m + \sum_{m=1}^{W-1} \sum_{i=1}^{F} \pi_m^i = 1. \tag{B.5}$$

Replacing in (B.5) the balance equations (B.1) and (B.2) one derives:

$$\pi_0 + \sum_{m=1}^{W-1} \pi_m + F \cdot p_f \cdot \sum_{m=1}^{W-1} \pi_m = 1, \tag{B.6}$$

which, with the help of (B.4), becomes

$$\pi_0 + (F \cdot p_f + 1) \sum_{m=1}^{W-1} \frac{W-m}{W} \pi_0 = 1, \tag{B.7}$$

from which the final expression for π_0 is derived:

$$\pi_0 = \frac{2}{2 + (W-1)\left[1 + p_f F\right]}. \tag{B.8}$$

References

[1] *ANSI/IEEE Std. 802.11-1999 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* IEEE Standards, 1999.

[2] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 535–547, March 2000.

[3] D. Reddy, G. F. Riley, B. Larish, and Y. Chen, "Measuring and Explaining Differences in Wireless Simulation Models," in *Proceedings of the 14th IEEE International Symposium on Modeling, Analysis, and Simulation*, vol. 0, (Los Alamitos, CA), pp. 275–282, 2006.

[4] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part 1- Carrier Sense Multiple-Access modes and their throughput-delay characteristics," *IEEE Transactions on Communications*, vol. 23, pp. 1400–1416, December 1975.

[5] R. Bruno, M. Conti, and E. Gregori, "Mesh Networks: Commodity MultiHop Ad Hoc Networks," *IEEE Communications Magazine*, vol. 43, pp. 123–131, March 2005.

[6] I. Akyildiz and X. Wang, "A survey on wireless mesh networks," *IEEE Communications Magazine*, vol. 43, pp. S23–S30, September 2005.

[7] G. Anastasi, E. Borgia, M. Conti, and E. Gregori, "IEEE 802.11b Ad Hoc Networks: Performance Measurements," *Cluster Computing*, vol. 8, pp. 135–145, July 2005.

[8] M. Sekido, M. Takata, M. Bandai, and T. Watanabe, "A directional hidden terminal problem in ad hoc network MAC protocols with smart antennas and its solutions," in *Proceedings of IEEE GLOBECOM 2005*, vol. 5, November 2005.

[9] A. Tsertou, D. I. Laurenson, and J. S. Thompson, "A New Approach for the Throughput Analysis of IEEE 802.11 in Networks with Hidden Terminals," in *Proceedings of IWWAN 2005*, (London, UK), May 2005.

[10] A. Tsertou and D. I. Laurenson, "Revisiting the Hidden Terminal Problem in a Multihop Wireless Network," *submitted to IEEE Transactions on Mobile Computing*, May 2006.

[11] A. Tsertou and D. I. Laurenson, "Insights into the Hidden Node Problem," in *Proceedings of ACM IWCMC 2006*, (Vancouver, Canada), pp. 767–772, July 2006.

[12] A. Tsertou and D. I. Laurenson, "Modeling the Effect of BEB for a Hidden Terminal Topology from a New Perspective," in *Proceedings of IEEE IWWAN 2006*, (New York, US), June 2006.

[13] J. H. Kim and J. K. Lee, "Capture Effects of Wireless CSMA/CA Protocols in Rayleigh and Shadow Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 1277–1286, July 1999.

[14] R. Rom and M. Sidi, *Multiple Access Protocols: Performance and Analysis.* Springer Verlag, 1990.

[15] F. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part 2-The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy-Tone Solution," *IEEE Transactions on Communications*, vol. 23, pp. 1417–1433, December 1975.

[16] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Transactions on Communications*, vol. 32, pp. 246–257, March 1984.

[17] L. Wu and P. K. Varshney, "Performance Analysis of CSMA and BTMA Protocols in Multihop Networks:Part I–Single Channel Case," *Information Sciences, Elsevier Sciences Inc.*, vol. 120, pp. 159–177, 1999.

[18] X. Wang and K. Kar, "Throughput Modelling and Fairness Issues in CSMA/CA Based Ad Hoc Networks," in *Proceedings of IEEE INFOCOM 2005*, vol. 3, (Miami, FL), pp. 1997–2007, 2005.

[19] Y. Wang and J. J. Garcia-Luna-Aceves, "Performance of Collision Avoidance Protocols in Single-Channel Ad-Hoc Networks," in *Proceedings of the 2002 IEEE International Conference on Network Protocols (ICNP)*, pp. 68–77, 2002.

[20] M. Carvalho and J. J. Garcia-Luna-Aceves, "A Scalable Model for Channel Access Protocols in Multihop Ad Hoc Networks," in *Proceedings of the 2004 ACM Mobicom*, pp. 330–344, 2004.

[21] R. R. Boorstyn, A. Kerschenbaum, B. Maglaris, and V. Sahin, "Throughput Analysis in Multihop CSMA Packet Radio Networks," *IEEE Transactions on Communications*, vol. 35, pp. 267–274, March 1987.

[22] A. Kerschenbaum, R. R. Boorstyn, and M.-S. Chen, "An Algorithm for Evaluation of Throughput in Multihop Packet Radio Networks with Complex Topologies," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 1003–1012, July 1987.

[23] F. A. Tobagi and J. M. Brazio, "Throughput Analysis of Multihop Packet Radio Networks Under Various Channel Access Schemes," in *Proceedings of IEEE INFOCOM*, (San Diego, CA), pp. 381–389, April 1983.

[24] K. Xu, M. Gerla, and S. Bae, "How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks?," in *Proceedings of 2002 IEEE GLOBECOM*, vol. 1, pp. 72–76, 2002.

[25] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Floor Acquisition Multiple Access (FAMA) for packet-radio networks," in *Proceedings of SIGCOMM '95*, (Cambridge, MA), pp. 262–273, 1995.

[26] Z. J. Haas, "On the performance of a medium access control scheme for the reconfigurable wireless networks," in *Proceedings of 1997 IEEE MILCOM*, vol. 3, pp. 1558–1564, 1997.

[27] M. Borgo, A. Zanella, P. Bisaglia, and S. Merlin, "Analysis of the hidden terminal effect in multi-rate IEEE 802.11b networks," in *Proceedings of the International Symposium of Wireless Personal Multimedia Communication WPMC 2004*, September 2004.

[28] I. Tinnirello, S. Choi, and Y. Kim, "Revisit of RTS/CTS Exchange in High Speed IEEE 802.11 Networks," in *Proceedings of the IEEE WoWMoM 2005*, pp. 240–248, June 2005.

[29] S. Ray, D. Starobinski, and J. B. Carruthers, "Evaluation of the Masked Node Problem in Ad Hoc Wireless LANs," *IEEE Transactions on Mobile Computing*, vol. 4, pp. 430–442, September/October 2005.

[30] J. L. Sobrinho, R. de Haan, and J. M. Brazio, "Why RTS-CTS is not your ideal wireless LAN multiple access protocol," in *Proceedings of IEEE WCNC 2005*, vol. 1, (New Orleans, LA), pp. 81–87, March 2005.

[31] V. Vitsas and A. C. Boucouvalas, "Performance Analysis of the Advanced Infrared (AIr) CSMA/CA MAC Protocol for Wireless LANs," *Wireless Networks*, no. 9, pp. 495–507, 2003.

[32] P. Chatzimisios and A. C. Boucouvalas, "Packet Delay Analysis of the Advanced Infrared (AIr) CSMA/CA MAC Protocol in Optical Wireless LANs," *International Journal of Communication Systems*, vol. 18, pp. 307–331, April 2005.

[33] G. Bianchi and I. Tinnirello, "Remarks on IEEE 802.11 DCF Performance Analysis," *IEEE Communications Letters*, vol. 9, pp. 765–767, August 2005.

[34] J. Yin, X. Wang, and D. P. Agrawal, "Optimal Packet Size in Error-prone Channel for IEEE 802.11 Distributed Coordination Function," in *Proceedings of IEEE WCNC 2004*, pp. 1654–1659, 2004.

[35] E. Ziouva and T. Antonakopoulos, "CSMA/CA Performance Under High Traffic Conditions:Throughput and Delay Analysis," *Computer Communications*, vol. 25, pp. 313–321, February 2002.

[36] C. H. Foh and J. W. Tantra, "Comments on IEEE 802.11 Saturation Throughput Analysis with Freezing of Backoff Counters," *IEEE Communications Letters*, vol. 9, pp. 130–132, February 2005.

[37] Y. Fang and B. McDonald, "Theoretical Channel Capacity in Multi-hop Ad Hoc Networks," in *Proceedings of the 2004 IEEE LANMAN*, pp. 33–38, 2004.

[38] F. Alizadeh-Shabdiz and S. Subramaniam, "MAC layer Performance Analysis of Multi-Hop Ad Hoc Networks," in *Proceedings of the 2004 IEEE GLOBECOM*, vol. 5, pp. 2781–2785, 2004.

[39] M. Garetto, J. Shi, and E. Knightly, "Modeling Media Access in Embedded Two-Flow Topologies of Multi-Hop Wireless Networks," in *Proceedings of the 2005 ACM Mobicom*, pp. 200–214, 2005.

[40] T.-C. Hou, L.-F. Tsao, and H.-C. Liu, "Analyzing the Throughput of IEEE 802.11 DCF Scheme with Hidden Nodes," in *Proceedings of IEEE VTC 2003*, pp. 2870–2874, 2003.

[41] C. Chaudet, I. G. Lassous, E. Thierry, and B. Gaujal, "Study of the Impact of Asymmetry and Carrier Sense Mechanism in IEEE 802.11 Multi-Hop Networks Through a Basic Case," in *Proceedings of ACM PEWASUN 2004*, (Venice, Italy), pp. 1–7, 2004.

[42] S. Ray, D. Starobinski, and J. Carruthers, "Performance of Wireless Networks with Hidden Nodes: A Queueing Theoretic Analysis," *Elsevier Computer Communications Journal (Special Issue on Perf.Issues of WLANs,PANs,and Ad Hoc Networks)*, vol. 28, no. 10, pp. 1179–1192, 2005.

[43] M. Garetto, T. Salonidis, and E. Knightly, "Modeling Per-flow Throughput and Capturing Starvation in CSMA Multi-hop Wireless Networks," in *Proceedings of the IEEE INFOCOM '06*, (Barcelona, Spain), 2006.

[44] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New Insights from a Fixed Point Analysis of Single Cell IEEE 802.11 WLANs," in *Proceedings of IEEE INFOCOM '05*, vol. 3, (Miami, FL), pp. 1550–1561, 2005.

[45] G. Sharma, A. Ganesh, and P. Key, "Performance Analysis of Contention Based Medium Access Control Protocols," in *Proceedings of IEEE INFOCOM '06*, (Barcelona, Spain), April 2006.

[46] K. Medepalli and F. A. Tobagi, "Towards Performance Modeling of IEEE 802.11 based Wireless Networks:A Unified Framework and its Applications," in *Proceedings of IEEE INFOCOM '06*, (Barcelona, Spain), April 2006.

[47] H. Chhaya and K. Gupta, "Performance Modeling of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol," *IEEE Personal Communications*, vol. 3, pp. 8–15, October 1996.

[48] H. S. Chhaya and S. Gupta, "Performance Modeling of Asynchronous Data Transfer Methods of IEEE 802.11 MAC protocol," *ACM Wireless Networks*, vol. 3, no. 3, pp. 217–234, 1997.

[49] F. Cali, M. Conti, and E. Gregori, "Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit," *IEEE/ACM Transactions on Networking*, vol. 8, pp. 785–799, December 2000.

[50] A. Gkelias, M. Dohler, and H. Aghvami, "Throughput analysis for wireless multi-hop CSMA," in *Proceedings of the 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2004)*, vol. 2, pp. 979–983, September 2004.

[51] A. Gkelias, M. Dohler, V. Friderikos, and H. Aghvami, "Average packet delay of CSMA/CA with finite user population," *IEEE Communications Letters*, vol. 9, pp. 273–275, March 2005.

[52] B.-J. Kwak, N.-O. Song, and L. E. Miller, "Performance Analysis of Exponential Backoff," *IEEE/ACM Transactions on Networking*, vol. 13, pp. 343–355, April 2005.

[53] R. German and A. Heindl, "Performance Evaluation of IEEE 802.11 Wireless LANs with Stochastic Petri Nets," in *Proceedings of the 8th International Workshop on Petri Nets and Performance Models (PNPM '99)*, (Zaragoza, Spain), pp. 44–53, October 1999.

[54] F. Ye and B. Sikdar, "Evaluation of Spatial Reuse in Wireless Multiple Access Networks," in *Proceedings of IEEE VTC 2004*, vol. 6, (Los Angeles, CA), pp. 26–29, September 2004.

[55] Y. Sun, X. Gao, E. M. Belding-Royer, and J. Kempf, "Model-based Resource Prediction for Multihop Wireless Networks," in *Proceedings of the 1st IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS) '04*, (Ft.Lauderdale, FL), pp. 114–123, October 2004.

[56] Z. Li, S. Nandi, and A. K. Gupta, "Modeling the short-term unfairness of IEEE 802.11 in presence of hidden terminals," *Elsevier Performance Evaluation*, vol. 63, pp. 441–462, May 2006.

[57] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "Delay Analysis of 802.11 in Single-Hop Networks," in *Proceedings of the IEEE ICNP 2003*, pp. 146–155, November 2003.

[58] A. Banchs and L. Vollero, "A Delay Model for IEEE 802.11e EDCA," *IEEE Communications Letters*, vol. 9, pp. 508–510, June 2005.

[59] H.-T. Wu, Y. Peng, K. Long, S.-D. Cheng, and J. Ma, "Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN," in *Proceedings of IEEE INFOCOM 2002*, vol. 2, (New York, NY), pp. 599–607, June 2002.

[60] A. Zanella and F. D. Pellegrini, "Statistical Characterization of the Service Time in Saturated IEEE 802.11 Networks," *IEEE Communications Letters*, vol. 9, pp. 225–227, March 2005.

[61] V. Ramaiyan, A. Kumar, and N. Vasudevan, "Fixed Point Analysis of the Saturation Throughput of IEEE 802.11 WLANs with Capture," in *Proceedings of the National Conference on Communications (NCC)*, (IIT, Kharagpur), January 2005.

[62] H. Chang, V. Misra, and D. Rubenstein, "A General Model and Analysis of Physical Layer Capture in 802.11 Networks," in *Proceedings of IEEE INFOCOM 2006*, (Barcelona, Spain), April 2006.

[63] Y. D. Barowski, S. Biaz, and P. Agrawal, "Towards the Performance Analysis of IEEE 802.11 in Multi-hop Ad-Hoc Networks," in *Proceedings of the 2005 IEEE Wireless Communications and Networking Conference (WCNC 2005)*, vol. 1, pp. 100–106, March 2005.

[64] O. Tickoo and B. Sikdar, "Queueing Analysis and Delay Mitigation in IEEE 802.11 Random Access MAC based Wireless Networks," in *Proceedings of IEEE INFOCOM 2004*, vol. 2, (Hong-Kong), pp. 1404–1413, March 2004.

[65] Z. Li, S. Nandi, and A. K. Gupta, "An Enhanced Carrier Sensing Mechanism for Wireless Ad Hoc Networks," *Computer Communications*, vol. 28, no. 17, pp. 1970–1984, 2005.

[66] P. Stuedi, O. Chinellato, and G. Alonzo, "Connectivity in the Presence of Shadowing in 802.11 Ad Hoc Networks," in *Proceedings of the IEEE ICC 2004*, pp. 4004–4011, June 2004.

[67] J. Woller, "The Basics of Monte Carlo Simulations," in *http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html*, webpage accessed on 29/09/2006.

[68] K. S. Trivedi, *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. Wiley, 2002.

[69] J. F. Traub, *Iterative Methods for the Solution of Equations.* Prentice-Hall Series in Automatic Computation, 1964.

[70] M. Kijima, *Markov Processes for Stochastic Modeling.* Chapman & Hall, 1997.

[71] *The Network Simulator -ns-2, release ns-2.1b9a.* http://www.isi.edu/nsnam/ns/.

[72] E. T. Jaynes, *Probability Theory: The Logic of Science.* Cambridge, 2003.

[73] E. Castillo, A. S. Hadi, N. Balakrishnan, and J. M. Sarabia, *Extreme Value and Related Models with Applications in Engineering and Science.* Wiley-Interscience, 2004.

[74] K. Drakakis, A. Tsertou, and D. I. Laurenson, "An Exact Markov-chain model for the hidden terminal topology case of the 802.11 protocol," *submitted to Elsevier Journal of Computer Networks*, July 2006.

[75] S. I. Resnick, *Adventures in Stochastic Processes.* Birkhauser, 1992.