Kaveh Ghaboosi

INTELLIGENT MEDIUM ACCESS CONTROL FOR THE FUTURE WIRELESS NETWORKS

FACULTY OF TECHNOLOGY,
DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING,
CENTRE FOR WIRELESS COMMUNICATIONS,
UNIVERSITY OF OULU
KAVEH GHABOOSI

INTELLIGENT MEDIUM ACCESS CONTROL FOR THE FUTURE WIRELESS NETWORKS

Academic dissertation to be presented with the assent of the Faculty of Technology of the University of Oulu for public defence in Auditorium IT116, Linnanmaa, on 29 October 2009, at 12 noon

OULUN YLIOPISTO, OULU 2009
Copyright © 2009
Acta Univ. Oul. C 335, 2009

Supervised by
Professor Matti Latva-aho

Reviewed by
Professor Giuseppe Bianchi
Professor Roberto Verdone

ISBN 978-951-42-9217-0 (Paperback)
ISBN 978-951-42-9218-7 (PDF)
http://herkules.oulu.fi/isbn9789514292187/
ISSN 0355-3213 (Printed)
ISSN 1796-2226 (Online)
http://herkules.oulu.fi/issn03553213/

Cover design
Raimo Ahonen

OULU UNIVERSITY PRESS
OULU 2009
Ghaboosi, Kaveh, Intelligent medium access control for the future wireless networks.
Faculty of Technology, Department of Electrical and Information Engineering, Centre for Wireless Communications, University of Oulu, P.O. Box 4500, FI-90014 University of Oulu, Finland
Acta Univ. Oul. C 335, 2009

Abstract
Medium access control (MAC) in wireless ad hoc networks has received considerable attention for almost a couple of decades; however, there are still open problems which deserve thorough study in order to facilitate migration to the next generation broadband wireless communication systems. In ad hoc networks, a detected frame collision can be due to the so-called unreachability problem, where the destination station is situated either in the transmission or interference range of an emitting station and is unable to receive connection establishment frames from any of its neighboring stations. Unreachability might also be due to the inability of a radio station to respond to any connection establishment request, though when the unreachable station receives the connection establishment requests, however, it is prohibited from responding to the requests due to being situated in the interference range of the emitting neighbor.

To investigate the impact of this problem, we have to be equipped with a proper analytical framework; therefore, as the first part of this thesis, a scalable framework called Parallel Space – Time Markov chain (PSTMC) is proposed, through which a finite load non-saturated ad hoc network can be easily modeled. At the first step, a single-hop ad hoc network is considered and the accuracy of the model is evaluated using extensive numerical results. Subsequently, the proposed framework is further extended to model multi-hop ad hoc networks. Several discussions are also given on how the framework can be deployed for an arbitrary network topology. One of the main key features of the PSTMC model is its remarkable scalability in modeling complex network configurations. In fact, it is shown that multi-hop ad hoc networks have bounded complexity in being modeled by the PSTMC framework due to its spectacular specifications. These features lead us to a powerful tool by which an arbitrary network topology can be studied. In addition, the proposed models clearly facilitate demonstrating the impact of the unreachability problem on the performance of multi-hop networks.

The introduced framework shows how the unreachability problem degrades the achieved throughput and channel capacity by the contending radio stations depending on the deployed network topology.

In the remainder of the thesis the unreachability problem in mobile ad hoc networks is tackled and a new MAC protocol to enhance the performance of the network is proposed. This MAC scheme is equipped with smart decision-making algorithms as well as adaptive management mechanisms to reduce the impact of the unreachability problem in single channel scenarios. Subsequently, the problem of concurrent radio resource management and contention resolution in multi-channel cognitive ad hoc networks is considered. In particular, a multi-channel technique for traffic distribution among a set of data channels without centralized control, which is enabled by a probabilistic channel selection algorithm as well as a multi-channel binary exponential backoff mechanism, is proposed. It is shown through simulations that the suggested scheme outperforms the existing MAC protocols in multi-channel environments as well as cognitive networks coexisting with primary users. A mathematical model is also introduced to study the performance of the multi-channel MAC protocol in a single-hop non-saturated wireless network.

Keywords: ad hoc networks, cognitive radio, finite load analysis, hidden terminal problem, medium access control, Parallel Space – Time Markov chain, queuing theory, unreachability problem
To my family & Sam
Preface

Research work related to this thesis has been carried out at the Centre for Wireless Communications (CWC), Department of Electrical and Information Engineering, University of Oulu, Oulu, Finland, during the years 2006–2009. I wish to thank the director of CWC, Lic. Tech. Ari Pouttu, for giving me the opportunity to work in such a wonderful and inspiring research unit.

First of all, I would like to express my gratitude to my supervisor Professor Matti Latva-aho for giving me the opportunity to pursue my postgraduate studies at CWC, for being so supportive whenever I was totally frustrated, for being so kind and affable with me, and for his priceless guidelines on how to become a cool guy. It is my pleasure to accommodate our friendship in my heart and keep it forever.

Furthermore, I am indeed grateful to Professor Yang Xiao from the Department of Computer Science, the University of Alabama, United States, as well as his kind wife, for being implausibly kind with me during my visit to the University of Alabama. Without his support and invaluable comments, this thesis would have never been finalized and, therefore, my warmest regards go to him for being always kind and helpful at any time. I would like to express thanks to Professor Qian Zhang from the Department of Computer Science, the Hong Kong University of Science and Technology, Hong Kong, for her invaluable comments on the eMAC protocol proposal and her support and contribution to the corresponding journal article.

I do love to thank Professor Allen B. MacKenzie and Professor Luiz A. DaSilva from the Bradley Department of Electrical and Computer Engineering, Virginia Tech., Blacksburg, Virginia, United States, who were my supervisors while I was visiting their laboratory at Virginia Tech. I admire their unrelenting pursuit of the highest standards in research. I am deeply grateful for their advices. Our regular teleconferences indeed help me to keep on working on new ideas while reminding me of joyful moments that I have had in Blacksburg.

I am most grateful to my reviewers Professor Giuseppe Bianchi (Università degli Studi di Roma) and Professor Roberto Verdone (Università di Bologna) for their thorough examination of the manuscript as well as to my opponents Professor Roberto Verdone (Università di Bologna) and Professor Rahim Tafazolli (University of Surrey) in my public defense.

I also thank Ms. Anna Shepherd and Ms. Noora Valkama for the language revision. Nevertheless, any mistakes that remain are my own.
Partial funding of this thesis is due to various research projects in which I have had the pleasure of being involved over the last number of years. This research work would not have been possible without long-term funding by the Finnish Technology Agency TEKES, Elektrobit, Finnish Defense Forces, Nokia, and Nokia Siemens Networks through the Packet Access Networks with Flexible Spectrum Use (PANU) project as well as the Cognitive and Opportunistic Wireless Communication Networks (Cognac). I would like to thank the fellow researchers and steering and technical group members in these projects for their contributions and technical advice. This work has also been supported with scholarships by Nokia, Elisa, Tauno Tönning Foundation, as well as Infotech, which are hereby gratefully acknowledged.

I am indebted to all my friends from the CWC, Telecommunication Laboratory, as well as Virginia Tech. for the spectacular working atmosphere and cordial cooperation. Among all my friends, at first, I would like to express my deep gratitude to Helal Chowdhury, Attaphongse Taparugssanagorn (Dr. Pong and his lovely wife Aui), Chathuranga Weeraddana (Chathu), Nima Naimi Akbar (Oulun Prince of Persia), Mohammad Sayyad Haghighi (msh, everywhere in cyberspace), Raghavendra Ramakrishna Madanahally (Raghu), Raghavendra Sathyanarayana (Ragha), Antti Tölli (Dadollah), Hanna Saarela (Hanna Severella), Mehdi Bennis, Ikram Ashraf (iKram), Heli Niva-Puuperä (Heli NP-Hard), Marian Codreanu (cvx v2.0), Juha Karjalainen (Dr. Matsumotosis, first proposed by Dadollah), Animesh Yadav (Anish), Vamsi Krishna, Axel Dauch (Dr. Bosch), Brett Kaufman (Sir Brett and his parents, Kay & Rob), Leonardo Goratti (Leo), Aristotelis Kechagias (Ari), Li Wei (Dr. Business Class), Daniel H. Friend, Umesh Shukla (My Miaw mate at VT), Romi Thakur (VT:n Princess), Alex Romagna (the first rock dude with whom I shared my office), Valeria Vitale (Vali, according to iKram), Francesco Pantisano (Party-sano), Antti Koiranen, Tuomas Haataja (Dude Matlabinen), Pen-Shun Lu, Pedro Nardelli, Johanna Vartiainen, Juho Määtä, Timo Hongel, Ari Isola (Mr. Tableman), with whom I could share all my happiness and sorrows. Most of us spent part of our daily hours in TS414 at the Centre for Wireless Communications, the place from which TS414 Spirit comes. I hope that all individuals who are going to be accommodated in TS414 keep it alive and preserve its warm and noisy atmosphere.

Moreover, my hat goes off to my dearest friends Marcos Katz, Prof. Kaveh Pahlavan, Prof. Behnaam Aazhang, Prof. Carlos Pomalaza-Ráez, Tadashi Matsumoto (Tad), Matti Isohookana, and Timo Bräysy for their kind support and motivation throughout my studies.
I would like to thank Ashley Lee (Ash) and Jennifer Fancher Bonney with whom I had lovely and joyful moments in Blacksburg, United States. I will never forget the huge pizzas we used to eat together and get more than 27 pounds extra fat. I still hear Ashley playing video games and making lots of noise in our town house in Blacksburg.

I extend my special appreciation to the administrative personnel at the CWC. I am indeed grateful to Elina Komminaho, Timo Äikäs, Jari Sillanpää, Tero Suutari, Sari Luukkonen, and Antero Kangas for their invaluable help and providing a pleasant working environment and the fun annual activities. Among all, I am most grateful to Timo for being the most supportive and helpful friend that I could have ever had in Oulu. I am indeed pleased that I came to know his lovely family, with whom I spent great moments that I will never forget.

My warmest thanks belong to my parents Mahin and Changiz, as well as my lovely sister Akasha, for a lifetime of love, care, and support. I have always been missing my mother’s warm hugs. I can always hear her voice, telling me to work harder and not to give up because of being far from my holy country, IRAN, and all things that I have been missing. Without her support and continuous encouragement this work would not have been finished. For all these reasons, this thesis is dedicated to my family.
List of symbols and abbreviations

$a$ Vector of all tunable factors in ICP broadcast probability function $\beta = \beta(N, N', a)$

$a_{k}$ Tunable factors in ICP broadcast probability function $\beta$

$a_{0}$ Vector of all tunable factors in ICP broadcast probability function $\beta$

$a_{z}$ Initial value of data channel $C_z$’s counter

$[a_j]$ Vector of the initial values of data channels $C_j$’s counter, given that their remaining busy period is zero

$a_{+}$ Tunable coefficient in ICP broadcast probability function $\beta$

$a_{-}$ Tunable coefficient in ICP broadcast probability function $\beta$

$A_n(t)$ Probability of $n$ MAC frame(s) entrance into the transmission queue during a time period of $t$

$b$ Number of master clock beats

$b(t)$ Stochastic process describing the current step of the backoff process

$b_{k}$ Stationary post-backoff state for which the post-backoff procedure resides in the $k$th step while the transmission queue is empty

$b_{x}$ Busy tone associated to the data channel $C_x$

$\delta_{z}$ Non-negative variable specifying the remaining busy period of data channel $C_x$ in number of time slots

$B_{i,k,j}$ Stationary backoff state for which the backoff procedure resides in the $i$th stage and $k$th step while the transmission queue holds $j$ MAC frame(s) prepared for delivery on the wireless media

$c$ Constant value used to define eMAC table rebroadcast repetition frequency

$C$ Geometric random variable representing the number of unsuccessful frame (re)transmissions due to collision

$C_i$ $i$th fully-connected complete cluster

$C_z$ $z$th data channel

$d$ Dummy variable representing the station to which the tagged MAC frame is addressed

$d$ Random variable representing MAC frame service time, i.e., the time taken by the MAC layer control entity to successfully deliver a frame over the radio channel
\(d'\) Dummy variable representing the station to which the tagged MAC frame is addressed

\(D\) Total MAC frame delay

\(\bar{D}\) Dummy random variable for which \(\sigma_\bar{D}^2 = \sigma_D^2\) and \(E[D] = E[\bar{D}] + T_s\)

\(e\) Non-negative variable specifying the control channel remaining busy period in number of slots

\(E\) Empty state of MAC layer transmission queue

\(E[P]\) Average MAC frame payload size

\(\mathcal{E}(t)\) Stochastic process representing the remaining busy period of the control channel at \(t\)

\(F_j\) Asynchronous frame transmission state in which the MAC frame that is on the head of the MAC layer queue is to be delivered on the wireless media since upon its entrance the previous post-backoff procedure has been already finished while the transmission queue was empty

\(\mathcal{G}(t)\) Stochastic process representing the number of remaining busy time slots of the data channel occupied by the tagged communication pair at \(t\)

\(h\) Temporary variable used to describe the concept of intermediate state

\(h_t\) Hash function used in BTMC protocol

\(\eta\) Probability that a communication pair is utilizing one of the data channels, given the channel status of the multi-channel system

\(i, k, j\) Backoff stage, backoff step, and the number of MAC frame(s) buffered in the transmission queue

\(J\) Objective function used for optimizing ICP broadcast process outcome

\(J_0\) Objective function used by the tagged ICP candidate for optimizing broadcast process outcome

\(K_c\) Number of time slots needed to elapse for channel busy period due to frame collision to be over \(T_c = K_c \times \sigma\)

\(K_i\) eMAC table rebroadcast repetition frequency in number of beacon intervals calculated in the \(i\)th iteration

\(K_{max}\) Constant value used to define eMAC table rebroadcast repetition frequency

\(K_{min}\) Constant value used to define eMAC table rebroadcast repetition frequency
\( K_s \)  
Number of time slots needed to elapse for channel busy period due to successful frame transmission to be over \( T_s = K_s \times \sigma \)

\( L \)  
Transmission queue maximum capacity

\( L_0 \)  
Window size of the set from which data channels’ counters content are loaded, i.e., \{1, 2, ..., \( L_0 \)\}

\( L_n \)  
Window size of the set from which a data channel’s counter content is loaded after \( n \) unsuccessful candidacies during the same backoff process, i.e., \( 2^n \times L_0 \)

\( m \)  
Maximum allowable (re)transmissions per frame \( (m' \leq m) \)

\( m' \)  
Number of different contention window sizes

\( m(i, j) \)  
\( j \)th flag of \( i \)th record in eMAC table

\( M \)  
Number of data channels in the system

\( M \)  
Set of all \( M \) data channels, i.e., \( \{C_1, C_2, ..., C_M\} \)

\( n \)  
Variable used to represent 1) the number of stations in the network, 2) the number of frame arrivals to the transmission queue, or 3) the number of contending secondary communication pairs

\( n_c \)  
Random variable representing the number of collisions a MAC frame undergoes before being successfully received by the destination station

\( N \)  
Number of stations that have simultaneous friendship with both the unreachability cause and the ICP candidate (Unreachability of Type I), or the number of stations that have simultaneous friendship with any unreachable stations of Type I and the ICP candidate (Unreachability of Type II)

\( N_{ff} \)  
Number of the unreachability cause’s friends which are also the tagged ICP candidate’s friends

\( N_{f\bar{x}} \)  
Number of the unreachability cause’s friends that are \( \bar{X} \) for the tagged ICP candidate

\( N_{xf} \)  
Number of X stations for the unreachability cause which are also the tagged ICP candidate’s friends

\( N_{x\bar{x}} \)  
Number of X stations for the unreachability cause which are \( \bar{X} \) for the tagged candidate

\( N_i \)  
Number of stations that have simultaneous friendship with both the unreachability cause and the \( i \)th neighbor of the tagged ICP candidate (Unreachability of Type I), or the number of stations that have simultaneous friendship with any unreachable stations of Type
I and the \( i \)th neighbor of the tagged ICP candidate (Unreachability of Type II)

\( N_0 \) Number of stations that have simultaneous friendship with both the unreachability cause and the tagged ICP candidate (Unreachability of Type I), or the number of stations that have simultaneous friendship with any unreachable stations of Type I and the tagged ICP candidate (Unreachability of Type II)

\( N' \) Number of an ICP candidate’s neighbors that should be informed by ICP

\( p \) Conditional collision probability including the tagged station

\( p_{inc} \) Probability that an incumbent occupies a data channel, given that the channel is free of primary users during the previous time slot

\( p_{c;nb} \) Probability of observing a frame collision on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking

\( p_{t;b} \) Probability of observing an empty slot on the control channel, given that the tagged communication pair is pursuing its backoff process and all contending communication pairs are blocked due to lack of radio resources while the minimum remaining busy period in the entire system is \( b \)

\( p_{t;nb} \) Probability of observing an empty slot on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking

\( p_{s;nb} \) Probability of observing a successful channel acquisition on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking

\( p_z \) Estimated a priori probability that channel \( C_z \) will be occupied by the incumbent during the transmission of an average sized MPDU in addition to an ACK frame at the lowest possible bit rate

\( P_{\text{block}} \) Blocking probability due to lack of radio resources (i.e., unoccupied data channels)

\( P_{\text{coll}} \) Probability of collision due to simultaneous transmission of more than one radios among contending communication pairs on the control channel in a randomly chosen time slot, given the channel status of the multi-channel system
\(P_{ntt}\) Probability that none of the contending communication pairs transmit on the control channel in a randomly chosen time slot, given the channel status of the multi-channel system

\(P_s\) Probability of a successful transmission

\(P_s'\) Probability of a successful transmission due to a wireless node excluding the tagged station

\(P_e\) Probability that only one of contending communication pairs transmits on the control channel in a randomly chosen time slot, given the channel status of the multi-channel system

\(P_{tr}\) Probability of at least one channel access in a randomly chosen time slot due to any combination of wireless nodes including the tagged station

\(P_{tr}'\) Probability of at least one channel access in a randomly chosen time slot due to any combination of wireless nodes excluding the tagged station

\(P_{trt}\) Probability that at least one of the contending communication pairs transmits on the control channel in a randomly chosen time slot, given the channel status of the multi-channel system

\(P_{u,c}\) Probability of unreachability of members in sub-cluster \(\cap_j C_j\) as seen by members in \(C_i - \bigcup_{j \neq i} C_j\)

\(q(t)\) Stochastic process describing the current number of buffered frames in the transmission queue

\(q^x(a_x)\) Probability that data channel \(C_x\)'s counter with initial value \(a_x\) reaches zero exactly at \(b\) master clock beats

\(\bar{q}^x(a_x)\) Probability that data channel \(C_x\)'s counter with initial value \(a_x\) does not reach zero by \(b\) master clock beats

\(Q^x([a_x])\) Probability that data channel \(C_x\)'s counter with initial value \(a_x\) reaches zero exactly at \(b\) master clock beats before the counter of any other data channel

\(p_{e,s}^x([a_x])\) Probability that data channel \(C_x\)'s counter with initial value \(a_x\) wins the stochastic channel selection given the initial content of all unoccupied data channels’ counter

\(p_{e,s}^y([C_y])\) Probability that in the stochastic channel selection data channel \(C_y\) is chosen among all unutilized data channels

\(s_h\) Non-negative variable specifying the number of data channels with \(h\) remaining busy time slots occupied by the secondary users
\( \delta_h \)  Non-negative variable specifying the number of data channels with \( h \) remaining busy time slots at the starting point of a state transition occupied by the secondary users

\( s(i) \)  Backoff stage after \( i \) collisions

\( s(t) \)  Stochastic process describing the current stage of the backoff process

\( S_z(t) \)  Stochastic process representing the number of data channels with \( z \) remaining busy time slots occupied by the secondary users at \( t \)

\( S_j \)  \( j \)th Space representing the backoff procedure of the local station while its transmission queue holds \( j \) MAC frame(s) prepared for delivery on the wireless media

\( t \)  Temporary variable used to represent the current time instant

\( \tau_{i,j} \)  Random variable representing the duration of a slot time, defined as the time between two successive decrements of the backoff counter

\( tDIFS\text{time} \)  Channel time representing a DIFS interval

\( tEIFStime \)  Channel time representing an EIFS interval

\( tSIFS\text{time} \)  Channel time representing an SIFS interval

\( T_{ACK} \)  Channel busy period due to an ACK frame transmission

\( T_{CTS} \)  Channel busy period due to a CTS collision in which the tagged station has not been involved minus the timeout interval

\( T_{RTS} \)  Channel busy period due to an RTS collision in which the tagged station has not been involved minus the timeout interval

\( T_{CTS} \)  Channel busy period due to a CTS frame transmission

\( T_{DATA} \)  Channel busy period due to a DATA frame transmission

\( T_{RTS} \)  Channel busy period due to an RTS frame transmission

\( T_s \)  Channel busy period due to a successful transmission

\( T_{CTS} \)  Channel busy period due to a successful transmission seen by the destination station

\( T_{RTS} \)  Channel busy period due to a successful transmission seen by the source station

\( T_{\text{max}} \)  The maximum possible number of busy time slots in a data channel due to an incumbent transmission

\( T_{\text{timeout}} \)  Channel busy period due to a timeout

\( T_R \)  Random variable representing the contribution to the delay due to the backoff procedure

\( T_{r,t} \)  Random variable representing the time taken by an RTS frame collision on the control channel

16
Random variable representing the time taken by a successful frame transmission, including RTS/CTS handshake on the control channel as well as DATA/ACK exchange on a randomly chosen data channel

Conditional transmission probability that a communication pair transmits in a randomly chosen slot, given the channel status of the multi-channel system

Conditional transmission deferral probability of a communication pair due to an ongoing backoff process, given the channel status of the multi-channel system

Stationary probability distribution of the $i$th backoff stage and $k$th backoff step, given the entire system channel status

Stationary probability distribution of the frame delivery state, given the entire system channel status

Non-negative variable specifying the number of data channels with $h$ remaining busy time slots occupied by the primary users

Non-negative variable specifying the number of data channels with $h$ remaining busy time slots at the starting point of a state transition occupied by the primary users

Stochastic process representing the number of data channels with $z$ remaining busy time slots occupied by the primary users at $t$

Contention window size at the first stage of the backoff procedure or the largest possible value in the stochastic multi-channel load balancing scheme with which a counter can be loaded

Contention window size at the $i$th stage of the backoff procedure (for post-backoff, $i$ is always 0) or the largest possible value in the stochastic multi-channel load balancing scheme with which a counter can be loaded after $i$ unsuccessful attempts

Temporary variable used to describe the concept of intermediate state

Random variable representing the number of slot times spent in backoff stage $s(i)$

Random variable representing service time of head-of-line MAC frame for which there has been no preceding frame since it has entered the transmission queue

Random variable representing the time period experienced during $i$th backoff stage
\( X_{us} \) Random variable representing the service time of the head-of-line MAC frame in the transmission queue

\( \beta \) Probability with which an ICP frame is going to be broadcast by an ICP candidate

\( \delta \) Propagation delay

\( \Delta \) Channel busy period due to a collision in which the tagged station has not been involved

\( \Delta_c \) Channel busy period due to a collision in which the tagged station has been involved where \( \Delta_c = T_{\text{timeout}} + T_{\text{RTS}} \)

\( \eta_k(x, y) \) Stationary intermediate post-backoff state located between \( b_k \) and \( b_{k-1} \)

\( \theta_i(x, y) \) Stationary intermediate vacancy state located in \( E \) — plan

\( \theta_j(x, y) \) Stationary intermediate pre-backoff states located in \( S_j \) — plan

\( \theta_{i(k,j)}(x, y) \) Stationary intermediate backoff state located between \( B_{i(k,j)} \) and \( B_{i(k-1,j)} \)

\( \lambda \) Traffic arrival rate in frame per second

\( \xi_{C_i} \) Probability of transmission of a CTS frame from sub-cluster \( \cap_j C_j \)
due to an RTS originated from \( C_i \) — \( \cup_{j \neq i} C_j \)

\( \Xi \) Latest version of eMAC table that has ever been broadcast

\( \Xi(t) \) Synchronous eMAC table

\( \rho \) Tunable factor in ICP broadcast probability function \( \beta \)

\( \sigma \) Predefined system time slot duration

\( \mathcal{Z}_0 \) Vector of \( N_0, N_1, \ldots, N_N \) maintained by the tagged ICP candidate

\( \mathcal{B}_0 \) Vector of all probability functions \( \beta_0, \beta_1, \ldots, \beta_N \) maintained by the tagged ICP candidate

\( \tau \) Conditional transmission probability

\( \tau_0 \) Decrement factor of contention window size upon a successful transmission in the EIED algorithm

\( \tau_1 \) Increment factor of contention window size upon a frame collision in the EIED algorithm

\( \psi \) Random variable representing the number of time slots needed for the backoff timer to reach zero

ACK Acknowledgment

AFT Asynchronous Frame Transmission

AMACA Advanced Multiple Access Collision Avoidance

AP Access Point

ATIM Announcement Traffic Indication Message
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>Backup Channel</td>
</tr>
<tr>
<td>BEB</td>
<td>Binary Exponential Backoff</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>BI</td>
<td>Beacon Interval</td>
</tr>
<tr>
<td>BNC</td>
<td>Best Next Channel</td>
</tr>
<tr>
<td>BP</td>
<td>Beacon Period</td>
</tr>
<tr>
<td>BSA</td>
<td>Basic Service Area</td>
</tr>
<tr>
<td>BSS</td>
<td>Basic Service Set</td>
</tr>
<tr>
<td>BT</td>
<td>Busy Tone</td>
</tr>
<tr>
<td>BTMA</td>
<td>Busy Tone Multiple Access</td>
</tr>
<tr>
<td>BTP</td>
<td>Begin Transmission Period</td>
</tr>
<tr>
<td>BTMC</td>
<td>Busy Tone Multi-channel</td>
</tr>
<tr>
<td>BTS</td>
<td>Base Station</td>
</tr>
<tr>
<td>CAM-MAC</td>
<td>Cooperative Asynchronous Multi-channel MAC</td>
</tr>
<tr>
<td>CC</td>
<td>Channel Control</td>
</tr>
<tr>
<td>CCAP</td>
<td>Cooperative Collision Avoidance Period</td>
</tr>
<tr>
<td>CDMAC</td>
<td>Cooperative Diversity MAC</td>
</tr>
<tr>
<td>CHMA</td>
<td>Channel Hopping Multiple Access</td>
</tr>
<tr>
<td>CMAC</td>
<td>Cognitive MAC</td>
</tr>
<tr>
<td>CPL</td>
<td>Channel Priority List</td>
</tr>
<tr>
<td>CS</td>
<td>Carrier Sense</td>
</tr>
<tr>
<td>CSCC</td>
<td>Common Spectrum Coordination Channel</td>
</tr>
<tr>
<td>CSCS</td>
<td>Cooperative Stochastic Channel Selection</td>
</tr>
<tr>
<td>CSMA</td>
<td>Carrier Sense Multiple Access</td>
</tr>
<tr>
<td>CST</td>
<td>Carrier Sense Threshold</td>
</tr>
<tr>
<td>CTMAC</td>
<td>Concurrent Transmission MAC</td>
</tr>
<tr>
<td>CTS</td>
<td>Clear-To-Send</td>
</tr>
<tr>
<td>CW</td>
<td>Contention Window</td>
</tr>
<tr>
<td>DBO</td>
<td>Directional Backoff</td>
</tr>
<tr>
<td>DBSMA</td>
<td>Directional Busy Signal Multiple Access</td>
</tr>
<tr>
<td>DBTMA</td>
<td>Dual Busy Tone Multiple Access</td>
</tr>
<tr>
<td>DCA</td>
<td>Dynamic Channel Assignment</td>
</tr>
<tr>
<td>DCBMA</td>
<td>Dual-channel Bi-directional Multiple Access</td>
</tr>
<tr>
<td>DCF</td>
<td>Distributed Coordination Function</td>
</tr>
<tr>
<td>DDMAC</td>
<td>Distance Dependent MAC</td>
</tr>
<tr>
<td>DHN</td>
<td>Double Hop Neighborhood</td>
</tr>
<tr>
<td>DIFS</td>
<td>Distributed Inter-Frame Space</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>DPC</td>
<td>Dynamic Private Channel</td>
</tr>
<tr>
<td>DS</td>
<td>Data Sent</td>
</tr>
<tr>
<td>DSU</td>
<td>Delegate Secondary User</td>
</tr>
<tr>
<td>DSA</td>
<td>Dynamic Spectrum Access</td>
</tr>
<tr>
<td>DTC</td>
<td>Data Transmission Complete</td>
</tr>
<tr>
<td>DTP</td>
<td>Data Transfer Period</td>
</tr>
<tr>
<td>EET</td>
<td>End-To-End Throughput</td>
</tr>
<tr>
<td>EIED</td>
<td>Exponential Increase Exponential Decrease</td>
</tr>
<tr>
<td>FAMA</td>
<td>Floor Acquisition Multiple Access</td>
</tr>
<tr>
<td>FCR</td>
<td>Fast Collision Resolution</td>
</tr>
<tr>
<td>FHSS</td>
<td>Frequency Hopping Spread Spectrum</td>
</tr>
<tr>
<td>GAMA-PS</td>
<td>Group Allocation Multiple Access with Packet Sensing</td>
</tr>
<tr>
<td>HCh</td>
<td>Home Channel</td>
</tr>
<tr>
<td>HCMAC</td>
<td>Hardware-constrained Cognitive MAC</td>
</tr>
<tr>
<td>HOL</td>
<td>Head-Of-Line</td>
</tr>
<tr>
<td>HRMA</td>
<td>Hop Reservation Multiple Access</td>
</tr>
<tr>
<td>ICSMA</td>
<td>Interleaved CSMA</td>
</tr>
<tr>
<td>ICP</td>
<td>Individual Communication Pause</td>
</tr>
<tr>
<td>IS</td>
<td>Invitation Signal</td>
</tr>
<tr>
<td>JMAC</td>
<td>Jamming-based MAC</td>
</tr>
<tr>
<td>LAMM</td>
<td>Load Awareness Multi-channel MAC</td>
</tr>
<tr>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MACA</td>
<td>Multiple Access Collision Avoidance</td>
</tr>
<tr>
<td>MACA-BI</td>
<td>Multiple Access Collision Access - By Invitation</td>
</tr>
<tr>
<td>MACSCC</td>
<td>MAC with a Separate Control Channel</td>
</tr>
<tr>
<td>MANET</td>
<td>Mobile Ad Hoc Network</td>
</tr>
<tr>
<td>MCC</td>
<td>Multi-channel Coordination</td>
</tr>
<tr>
<td>MCLK</td>
<td>Master Clock</td>
</tr>
<tr>
<td>MILD</td>
<td>Multiplicative Increase and Linear Decrease</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>MMAC</td>
<td>Multi-channel MAC</td>
</tr>
<tr>
<td>MSDU</td>
<td>MAC Service Data Unit</td>
</tr>
<tr>
<td>NAV</td>
<td>Network Allocation Vector</td>
</tr>
<tr>
<td>NCTS</td>
<td>Not Clear-To-Send</td>
</tr>
<tr>
<td>NDD</td>
<td>Neighbor Direction Detection</td>
</tr>
<tr>
<td>NTS</td>
<td>Not-To-Send</td>
</tr>
</tbody>
</table>
OC-MAC  Opportunistic Cognitive MAC
ODC  On Demand Channel Switching
OMC-MAC  Opportunistic Multi-channel MAC
OSA  Opportunistic Spectrum Access
OS-MAC  Opportunistic Spectrum MAC
P2MP  Point-To-Multipoint
PAMAS  Power Aware Medium Access Control with Signaling
PCAM  Primary Channel Assignment-based MAC
PCF  Point Coordination Function
PDF  Probability Distribution Function
PGF  Probability Generating Function
PSM  Power Save Mode
PSCS  Predictive Stochastic Channel Selection
PSTMC  Parallel Space – Time Markov chain
PU  Primary User
QoS  Quality of Service
QP  Quiet Period
RBT  Receive Busy Tone
RI  Receiver Initiated
RICHDP  Receiver Initiated Channel Hopping with Dual Polling
RRM  Radio Resource Management
RTS  Ready-To-Send
SCA-MAC  Statistical Channel Allocation MAC
SCIFS  Switching & Carrier Sensing Inter-frame Space
SI  Sender Initiated
SIFS  Short Inter-Frame Space
SSCH  Slotted Seeded Channel Hopping
STA  Station
SU  Secondary User
SUG  Secondary User Group
TBT  Transmit Busy Tone
TC  Transmission Confirmation
TCP  Transmission Control Protocol
TDMA  Time Division Multiple Access
TMMAC  TDMA-based Multi-channel MAC
TR  Transmit Request
UDP  User Datagram Protocol
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VG</td>
<td>Virtual Gateway</td>
</tr>
<tr>
<td>WE</td>
<td>Welfare Enhancement</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
<tr>
<td>WMN</td>
<td>Wireless Mesh Network</td>
</tr>
<tr>
<td>xRDT</td>
<td>Extended Receiver Directed Transmission</td>
</tr>
</tbody>
</table>
Contents

Abstract
Preface 7
List of symbols and abbreviations 11
Contents 23
1 Introduction 25
  1.1 Motivation ................................................................. 25
  1.2 MAC protocol design challenges in ad hoc networks .......... 26
    1.2.1 Overview ......................................................... 26
    1.2.2 MAC entity fundamental components ..................... 27
    1.2.3 Hidden terminal, exposed terminal, and unreachability
         problems: challenges and solutions ......................... 32
    1.2.4 Multi-channel MAC protocols ................................ 35
    1.2.5 Cognitive multi-channel MAC protocols ................. 43
  1.3 Aims and outline of the thesis ..................................... 45
  1.4 Author’s contribution .............................................. 46
2 Performance analysis of contention-based medium access control 47
  2.1 Overview and background .......................................... 47
  2.2 Assumptions ............................................................ 53
  2.3 Non-saturated performance analysis: single-hop perspective 54
    2.3.1 Traffic arrival process ....................................... 55
    2.3.2 Intermediate state transitions .............................. 55
    2.3.3 State transitions in the same and difference spaces ... 57
    2.3.4 Space – time chains state transitions ................... 59
    2.3.5 Throughput analysis ......................................... 64
    2.3.6 Model validation ............................................. 64
    2.3.7 Service time & queuing delay analysis ................... 69
  2.4 Non-saturated performance analysis: multi-hop perspective 75
    2.4.1 Key features of virtual gateway PSTMC model .......... 87
    2.4.2 Modeling the unreachability phenomenon ............... 89
    2.4.3 Throughput analysis ......................................... 92
    2.4.4 Model validation ............................................. 92
    2.4.5 Discussions and model extension .......................... 97
  2.5 Chapter summary ..................................................... 101
3 MAC protocol design for future multi-hop networks 103
3.1  eMAC – a MAC protocol for the next generation ad hoc networks ................................................................. 103
  3.1.1  Protocol architecture and basic concepts ........................................... 107
  3.1.2  Performance evaluation ................................................................ 124
3.2  Chapter summary ............................................................................. 133

4  Multi-channel contention resolution in cognitive wireless networks 135
  4.1  On stochastic channel selection for multi-channel cognitive wireless networks: a medium access control perspective ... 136
    4.1.1  Overview and background ...................................................... 136
    4.1.2  Stochastic channel selection and contention resolution mechanisms .................................................. 139
    4.1.3  Multi-channel unreachable problem ..................................... 144
  4.2  Performance analysis of saturated multi-channel MAC .................... 149
    4.2.1  Overview and background ...................................................... 149
    4.2.2  Multi-channel MAC without primary user consideration ......... 151
    4.2.3  Multi-channel Cognitive MAC with Primary User Consideration ................................................... 159
  4.3  Performance evaluation ................................................................. 164
    4.3.1  Multi-channel MAC without primary user consideration .......... 165
    4.3.2  Cognitive multi-channel MAC with primary user consideration ..................................................... 169
    4.3.3  Analytical model validation .................................................... 174
  4.4  Chapter summary ......................................................................... 178

5  Conclusions and future work ........................................................... 179

References .......................................................................................... 185
Appendices ......................................................................................... 197
1 Introduction

1.1 Motivation

An ad hoc network consists of a group of wireless stations that are capable of enthusiastically figuring a self-organized network, where routes connecting wireless nodes may consist of multiple hops, not necessarily needing to use any pre-installed infrastructure [1]. As a result of their effortless deployment, lots of practical applications have been conceived for ad hoc networks; besides, several attention-grabbing and indeed complex problems arise in such networks due to the nature of the shared wireless medium, limited transmission range of wireless devices, mobility, and battery power limitation.

Since in ad hoc networks stations utilize the same shared radio channel, Medium Access Control (MAC) plays an essential role in coordinating media acquisition among wireless stations so that information gets through from the source to the destination of interest. Even though a wide range of MAC protocols have been broadly studied in the context of wired networks, they cannot be straightforwardly applied to ad hoc networks, which have several distinctive characteristics that well discriminate themselves from their wired counterparts. For instance, wireless channels are not as reliable as their wired counterparts, suffering from path loss, fading, and interference. In addition, as its nature implies, an ad hoc network may be composed of a set of mobile stations that can be in motion stochastically. Consequently, the network topology may experience continuous alteration and cause frequent route ruptures and, as a result, requires supplementary route discovery and fault recovery procedures. An ad hoc network with mobile stations is commonly referred to as a Mobile Ad Hoc Network (MANET) [1]. In ad hoc networks wireless stations are characteristically confined by battery power, meaning that they are unable to afford multifarious and energy intensive functionalities. Last, but not least, ad hoc networks are self-organized, self-controlled, and distributed in nature. This means that there is no centralized coordinator and, therefore, each station can only have an incomplete, deficient, or sometimes skewed view of the network. With imperfect harmonization, collisions could take place, time to time, when more than one station accesses the shared radio channel at the same time. As each wireless station in an ad hoc network can also be a mobile entity, network topology may alter frequently. For that reason, stations may experience dissimilar degrees of media contention and collision. On
the other hand, the concurrent route changes have numerous effects on the interaction between the MAC layer and higher layer(s), as well. Frame losses due to the MAC layer’s unsuccessful contentions will definitely have a negative impact on the performance of the higher layers’ connection control schemes, such as Transmission Control Protocol (TCP) congestion control and routing maintenance, since the wireless station does not know whether an error is due to a collision, a high bit error rate, or an unreachable destination. It has been shown in many articles that multi-hop ad hoc networks perform inadequately with TCP traffic as well as intense User Datagram Protocol (UDP) data streams. As a concluding point, several important issues like energy efficiency, fairness, Quality of Service (QoS) provisioning, and security need to be circumspectly taken into consideration while designing MAC protocols for ad hoc networks. Due to all the aforementioned hurdles posed by these networks, designing a desirable MAC protocol is challenging [1, 2].

Hitherto, a great number of MAC schemes have been proposed in the literature to address a vast variety of issues encountered in ad hoc networks. This chapter aims to provide a comprehensive study of these schemes and to discuss both critical issues and tradeoffs in designing MAC protocols to deliver good performance in ad hoc networks.

1.2 MAC protocol design challenges in ad hoc networks

1.2.1 Overview

In this section, some of the most important MAC protocol design issues are explored, and in each sub-section the existing schemes in the literature are carefully investigated. Related topics are gathered and put into similar sub-sections and, subsequently, the most well-known solutions for the introduced problems are presented.

The key objective of a MAC protocol is to cope with the radio access control and perhaps channel acquisition scheduling among multiple nodes to achieve higher channel utilization. The coordination of medium access should minimize or even eliminate the incidence of frame collisions while taking full advantage of spatial reuse all at once. Collisions come from two aspects in ad hoc networks. They may take place due to simultaneous transmissions by more than one station where their channel accesses interfere with each other. Evidently, the more active
the wireless stations in the range of a transceiver are, the more severe the observed collisions will probably be. Generally, collision occurrence is due to the hidden terminal problem. A hidden terminal is one that can neither sense the transmission of a transmitter nor correctly receive the reservation control frames from its corresponding receiver. So far, a tremendous number of proposals have been suggested in the literature to combat the hidden terminal problem in ad hoc networks, e.g., [1, 2].

In order to achieve higher channel utilization and better performance, the MAC protocol is required to take full advantage of spatial reuse. In this respect, a feasible approach is to reduce the transmission power to allow more concurrent transmissions over the wireless network. However, a smaller transmission range means more transmission hops each packet needs to traverse from the source node to the destination station. This, in turn, leads to a heavier traffic load at each intermediate station and could counteract the advantage of increased spatial reuse. It has been shown that there is a tradeoff between the spatial reuse and the number of hops in order to attain higher aggregate throughput. Indeed, the optimal transmission range depends upon the number of wireless stations, their geographical location, and the speed by which they move through the network and hence is completely difficult to be achieved due to the dynamic and distributed nature of ad hoc networks.

In what follows, fundamental components of a MAC entity specifically designated for ad hoc network applications are discussed. Then, the well-known problems in such systems, namely hidden terminal, exposed terminal, and unreachability problems, and possible solutions to alleviate the impact of these problems in multi-hop ad hoc networks are carefully studied. Finally, not only single channel MAC protocols, but also the most popular multi-channel medium access control schemes are introduced and their channel management mechanisms will be investigated.

1.2.2 MAC entity fundamental components

Basically, a wireless station is unable to detect a frame collision while it is transmitting over the radio channel; rather, it relies on the receiver’s acknowledgment to determine if any collision has taken place during the frame delivery. Obviously, the busy period due to collision is quite long and exorbitant if a long data transmission encounters collision. Thus, minimizing the number of occurred collisions becomes a key issue in ad hoc networks. In this regard, quite a
lot of mechanisms have been proposed to avoid collisions, namely carrier sense, handshake, and backoff mechanisms.

In the case of Carrier Sense (CS), the wireless station concludes that the shared media is busy when the received signal power exceeds a certain threshold, referred to as the Carrier Sense Threshold (CST). If not, the shared radio channel is determined idle. It can be seen that the value of CST determines the approximate sensing range and influences together the collision risk and spatial reuse in multi-hop ad hoc networks. The larger the sensing range, the smaller the possibility that a fresh transmission effort interferes with some ongoing data delivery over the shared radio link. On the other hand, a wider sensing range implies that more stations have to defer from the channel acquisition when only one station is transmitting, which leads to an inefficient spatial reuse and consequently results in significant overall channel utilization degradation.

The IEEE 802.11 Distributed Coordination Function (DCF) is a contention-based MAC protocol and utilizes carrier sense and Binary Exponential Backoff (BEB) scheme. Even though BEB is extensively employed in large numbers of contention-based MAC protocols due to its simplicity, implementation straightforwardness, and reasonable performance, it suffers from both unfairness and inefficiency. In BEB, each station resets its Contention Window (CW) to the minimum value subsequent to a successful transmission, and makes it twice after an unsuccessful packet delivery attempt. As a result, it is possible that a station that has gained the channel and transmitted successfully gains control of the channel in the next contention phase. The worst-case scenario happens when a particular station monopolizes the channel occupation while all other stations are entirely denied channel acquisition. The Multiplicative Increase and Linear Decrease (MILD) algorithm was recommended in the MACAW protocol [3] in order to deal with the large variation of the contention window size and the unfairness problem of BEB. In MILD, the backoff interval is increased by a multiplicative factor 1.5 upon a collision and decreased by one step upon a successful transmission, where step is defined as the transmission time of a Ready-To-Send (RTS) control frame. MILD outperforms relatively for steadily heavy traffic load conditions [1, 2]. On the other hand, the ‘linear decrease’ is sometimes excessively conservative, and it suffers performance degradation when the traffic load is light or the number of contending stations varies piercingly [4]. To overcome the aforementioned problems, the Exponential Increase Exponential Decrease (EIED) backoff algorithm was designed [4, 5]. In the EIED algorithm, the contention window size is decreased by a factor $\tau_D$ upon a successful
transmission, and increased by a factor $\tau_i$ upon a collision. As a result, EIED is not as conservative as the ‘linear decrease’ of MILD. Due to the fact that for every number of competing stations there is a particular optimum contention window size, various investigations concentrated on adaptive contention window issues [6, 7]. By gathering observed collision statistics, the abovementioned methods estimate the number of stations and consequently calculate a new contention window size to arrange the next transmission effort.

In [8], the *Fast Collision Resolution (FCR)* algorithm has been studied. It has the following characteristics: (1) It uses a much smaller initial contention window size in comparison to the IEEE 802.11 BEB; (2) It uses a much larger maximum contention window size in comparison to the IEEE 802.11 BEB; (3) It increases the contention window size when a station is in both the collision and deferral state; (4) It reduces the backoff timer exponentially when a prefixed number of consecutive idle slots has been detected; (5) It assigns a maximum successive frame transmission limit to achieve fair performance. It has been shown in [8] that FCR resolves collisions faster and reduces the idle slots more effectively than the IEEE 802.11 BEB.

Numerous handshake-based approaches between transmitter and receiver stations can be mainly divided into a couple of distinct categories, namely *Sender-Initiated (SI)* and *Receiver-Initiated (RI)* methods. The two-way DATA/ACK and four-way RTS/CTS/DATA/ACK handshake of the IEEE 802.11 MAC protocol are both sender-initiated. Note that CTS stands for *Clear-To-Send* and ACK stands for *acknowledgement*. The sender commences the handshake only when there are pending frames prepared for delivery over the air. The exchange of short RTS and CTS control frames in a four-way negotiation between transmitter and receiver serves as a media reservation that notifies overhearing neighbors to defer their access to the shared radio channel in order to avoid possible collisions. The *Group Allocation Multiple Access with Packet Sensing (GAMA-PS)* incorporates features of contention-based as well as contention-free methods [9]. It divides the time into a series of network activity cycles. Every cycle is divided into two portions dedicated for contention-based and group transmission, respectively. Although the group transmission period is further divided into individual transmission periods, GAMA-PS does not require clock or time synchronization among stations. The stations willing to make a reservation for media acquisition are mandated to perform the RTS/CTS handshake prior to any data delivery over the wireless medium. GAMA-PS organizes associated wireless stations into distinct transmission groups, each one consists of a bunch of
wireless stations that have been allocated a particular transmission period. The stations, associated to a certain transmission group, are aware of all activities performed in their transmission group. Associated members of a transmission group take turns of information transmission, and each one is supposed to send a \textit{Begin Transmission Period (BTP)} control frame ahead of its actual data delivery. The BTP comprises the state of transmission group, position (i.e., rank) of the station within the group, and more importantly the number of group members. A member can transmit up to a fixed length of data, thereby enhancing the channel efficiency. The last member of the transmission group should broadcast a certain and so-called \textit{Transmit Request (TR)} control frame after it finishes its own data frame transmission. The use of TR shortens the maximum length of the contention period by forcing all stations that might contend for group membership to do so. GAMA-PS assumes that there are no hidden terminals. As a result, this scheme may not work well for multi-hop ad hoc networks. When there is not enough traffic in the network, GAMA-PS behaves almost like \textit{Carrier Sense Multiple Access (CSMA)}. However, as the load grows, it starts to mimic \textit{Time Division Multiple Access (TDMA)} and allows associated stations to transmit once per cycle.

In receiver-initiated approaches, the receiver polls its neighbors actively to check if they have any pending MAC frames targeted to it. As a well-known example, \textit{Multiple Access Collision Avoidance by Invitation (MACA-BI)} [10] adopts a three-way handshake, that is, CTS/DATA/ACK, to conduct the channel access where the CTS frame severs as the polling control frame and communication trigger. The receiver needs to get relatively long data frames and has better awareness of the contention status around itself. Additionally, the three-way handshake has less control overhead than the four-way handshake of the IEEE 802.11 MAC protocol, which elucidates the reason for which MACA-BI outperforms the four-way handshake of the IEEE 802.11 when traffic characteristics are stationary or predictable. However, it does not work well in ad hoc networks due to the fact that called parties may have no pending frames for the polling station and the transmission time of polling frames, as a result, is wasted, leading to channel utilization degradation. It is imperative to take into account that in both SI and RI handshakes, acknowledgements for successful transmissions are compulsory due to the unreliable wireless environment.

In an endeavor to attain the benefits of both SI and RI channel access mechanisms, some hybrid channel access methods have already appeared in the literature [11]. A wireless station that puts this method into practice operates
alternately in two different modes, SI or RI. The transmission pair tries to enter into the RI mode when the sender sends the same RTS control frame for more than one half of the times allowed in the IEEE 802.11 MAC protocol and has received no response from the intended receiver. By adaptively sharing the burden of collision-avoidance handshake initiation among those stations that experience different levels of contention, better fairness may be achieved with almost no degradation in throughput. In the multi-hop packet scheduling scheme [12] when the receiver is overloaded, a Negative CTS (NCTS) is used to inform the transmitter of local congestion, and then the transmission pair enters into the RI mode. As soon as congestion is mitigated and backlogged link layer frames have been delivered, the receiver commences a three-way handshake and afterward the transmission pair comes back to the SI mode. In this fashion, the aforementioned scheme effectively keeps upstream nodes from overloading downstream ones. Consequently, End-To-End Throughput (EET) is considerably enhanced by reducing collisions and avoiding dropping frames at the first few hops; furthermore, end-to-end delay is decreased significantly by reducing long queuing delay at forwarding nodes.

Group Transmission (or Batch Transmission [1]) is a further approach to enhance the efficiency of MAC protocols. In this approach, the station is not required to contend for the shared radio channel over again for one or more succeeding frames/fragments after a successful information transaction. This is equivalent to the case where longer DATA frames are employed in the IEEE 802.11 protocol. Since the collision probability may be the same prior to each transmission effort, throughput is improved as the successful transmission period is prolonged. In fact, Group Transmission has already been adopted by the IEEE 802.11 protocol in a fragmentation/defragmentation scheme. Given a fixed channel bit error rate, it is apparent that longer frames are more vulnerable to transmission errors. For that reason, the fragmentation procedure creates smaller data units than the original large DATA frames and therefore can enhance transmission reliability by reducing the frame error probability. It should be noted that each fragment is required to be acknowledged by the receiver, separately. Once a terminal has gained control of the wireless channel, it keeps on sending fragments until all fragments have been successfully received, or an acknowledgement (ACK frame) is not received, or the station is restrained from sending any further fragments due to a maximum retransmission limit. Should fragment-delivery be interrupted due to one of the above reasons, the station will recommence transmission when the next opportunity for transmission arises.
Group Transmission has also been employed in quite many other schemes, such as *Opportunistic Auto Rate (OAR)* \[13\]. In OAR, each station opportunistically sends multiple back-to-back data frames whenever the channel quality is fine and consequently achieves significant throughput improvements over time-varying channels. In spite of its throughput enhancement, Group Transmission itself does not necessarily decrease the potential collision probability. So, the efficiency is still affected by the collisions. Besides, it is destructive for critical communications and time-bounded streams, which have strict end-to-end delay requirements since whenever the station occupies the shared channel it blocks transmissions by other ones. To alleviate this consequence, IEEE 802.11, OAR or FCR also define a maximum period to limit the total duration of continuous transmissions by one station. As a concluding remark, it is worth mentioning that recently a set of innovative concatenation and piggyback mechanisms for ad hoc networks have been proposed to reduce the overhead of the IEEE 802.11 protocol, which also in comparison to the earlier methods overcome to the aforementioned problems \[14\]. Studies show that both of the abovementioned schemes greatly improve the wireless system performance. The idea concerning concatenation is based upon putting more than one data frame into a singular larger data frame in order to reduce the amount of overhead inquired by repeated exchanged RTS/CTS/ACK control frames. The scheme defines few thresholds to control the way the wireless channel is occupied by an ongoing transaction-based on concatenation.

### 1.2.3 Hidden terminal, exposed terminal, and unreachability problems: challenges and solutions

In ad hoc networks, a wireless station that is able to hear the receiver’s messages, but not the frames of the transmitter, is called a hidden terminal. Devoid of proper notification, hidden terminals do not have information about the ongoing transmissions, and as a result, their possible channel acquisition during the period of the other transmissions taking place in their neighborhood annihilates the DATA frames being received at the receiver and degrades the network utilization. Similarly, stations in the radio range of the transmitter but not the receiver are called exposed terminals. When CSMA is used to prevent channel collisions, exposed terminals are prevented from accessing the channel, although such access would not cause potential collisions.
To overcome the hidden terminal and exposed terminal problems, the *out-of-band busy tone* signal has been extensively used in numerous schemes [1, 2]. In *Busy Tone Multiple Access (BTMA)* [15], the *base station* generates a *Busy Tone (BT)* signal to prevent the hidden terminals from accessing the channel when it senses a transmission. The method relies on a centralized infrastructure, which is not obtainable in ad hoc networks. The *Dual Busy Tone Multiple Access (DBTMA)* scheme [16] uses *Transmit Busy Tone (TBT)* at the transmitter side to prevent the exposed terminals from becoming new receivers, and *Receive Busy Tone (RBT)* at the receiver side to prevent the hidden terminals from becoming new transmitters. The busy tone mechanism supplies an effortless solution to the hidden terminal and exposed terminal problems, but it requires additional channels and transceivers. The busy tone channel must be close to the DATA channel and therefore can have analogous channel gain to that of the DATA channel, and there must also be sufficient spectral separation between these two channels to avoid inter-channel interference. On the other hand, the bandwidth requirement of a busy tone signal is negligible, and the decoding is much simpler than that over the DATA channel. A node is only mandated to verify the existence of the busy tone signal at a certain frequency by sensing the received transmission power level.

*Floor Acquisition Multiple Access with Non-persistent Carrier Sensing (FAMA-NCS)* [17] uses long dominating CTS frames to act as a receiver busy tone to prevent any competing transmitter in the receiver’s radio range from media acquisition. In order to avoid facing collisions with an ongoing data transaction, this method necessitates each station to stay silent for the duration of one maximum data frame upon the detection of any interference. Perceptibly, this is not well-organized, especially when the RTS/CTS negotiation process fails or DATA frames are relatively short.

In ad hoc networks, the hidden terminal problem introduces frame collisions while the exposed terminal problem leads to a low spatial reuse ratio. Moreover, the unreachability problem, which is also referred to as the receiver blocking problem [18], deserves serious attention. This problem occurs when the destination of interest is situated either in the *communication* or *interference* range of another transmitting node. In such a case, efforts to set up a link layer connection with the destination station fail either due to frequent collisions taking place between transmitted RTS or DATA frames and unwanted overheard frames originating from an ongoing transmission hidden to the source node. Failure may also occur due to the inability of the destination station to respond by CTS or acknowledgement frame as a consequence of sensing the carrier caused by being
situated in the interference range of a transmitting node, shown in Fig. 1. Here, the former case, caused by communication range, is referred to as the unreachability of *Type I*, while the latter, caused by interference range, is referred to as the unreachability of *Type II* [19 – 21].

![Fig. 1. “C” is an Unreachable Terminal during ‘A’ — ‘B’ data transmission.](image)

When the intended receiver is in the radio range of some other ongoing data transmissions, it is not able to respond to the sender’s RTS control frame according to the carrier sensing specified in the IEEE 802.11 standard. Therefore, the sender attempts to retransmit the RTS frame. As a result, the contention window size becomes larger and larger when the RTS transmission fails, until the sender finally discards the frame waiting for delivery at the head of the MAC layer queue. It has been shown that, the *unreachability problem* results in serious unfairness among flows and severe frame discarding.

To alleviate the unreachability problem, *Advanced Multiple Access Collision Avoidance (AMACA)* [19] adds a few innovative control frames, namely RTSv1, RTSv2, CTSv1, CTSv2, and *DS* (*Data Sent*), to the MAC layer structure in order to improve the media access management process. The DS frame is used to notify the receiver about the start of the data transfer phase. In other words, prior to data transmission, the sender transmits the DS frame to its destination of interest from which it has already received a CTS control frame. As a matter of fact, the DS frame plays a key role in the AMACA protocol; if a terminal receives a DS frame targeted to another terminal, then it has to immediately generate an *Individual Communication Pause (ICP)* frame notifying all its neighbors that it will be unreachable during the time specified in the ICP frame Duration/ID field (which has been copied from the corresponding received DS frame Duration/ID field). As
a result, communicating with this terminal will be postponed until the end of upcoming data transmission. By the means of the aforementioned approach, AMACA prevents those transmission initiation attempts that lead to unsuccessful connection establishments. Thus, the shared wireless channel will be utilized in a more efficient fashion. This means that instead of wireless media exploitation for a predicted ineffective effort, it could have been utilized for another goal, such as delivering data frames to the other neighboring node(s).

AMACA has two distinct operating modes called Normal and Excessive Try while employs a complex timing system and implements two separate memory spaces named the Unreachability Vector and Unreachability Cache to buffer those frames for which the intended destination(s) are unreachable for a specific time duration. When the timer corresponding to the unreachable station reaches zero, the corresponding data frame(s) are fetched and sent at once. It should be noted that during the unreachability period, the source node is able to continue its other data transactions. Note that the Unreachability Vector is completely different from the Network Allocation Vector (NAV). NAV should be used by all terminals in the wireless network while the Unreachability Vector must be employed by a terminal when its desired destination is unreachable for a period of time [19].

1.2.4 Multi-channel MAC protocols

In single channel MAC protocols, all kinds of frames, including RTS/CTS/DATA/ACK as in the IEEE 802.11 media access scheme, are transmitted over the same wireless channel. As a consequence, collisions may occur between any two kinds of frames. One common solution to reduce the number of occurred collisions is to exploit the advantage of multiple channels, and transmit different kinds of frames over separate radio channels. There are several contributions in the literature addressing the problem of coordinated multi-channel operation at the MAC layer. These protocols are typically categorized based on how many radio transceivers they require for operation: single transceiver protocols and multiple transceiver protocols. In the following, the existing contributions in the field of multi-channel MAC protocols are reviewed and, subsequently, the most recent protocols for multi-channel cognitive wireless networks are studied.
**Single transceiver**

In single transceiver MAC protocols, it is assumed that the wireless station is equipped with only one half-duplex transceiver, so it is capable of switching among multiple channels dynamically, while being able to only transmit or receive on exactly one channel at any given time.

In this category, the *Hop Reservation Multiple Access (HRMA)* [22] is a well-studied multi-channel MAC protocol for slow *Frequency Hopping Spread Spectrum (FHSS)* wireless ad hoc networks. In this scheme, wireless stations hop according to a pre-defined hopping pattern. When a station has a data frame to send, it exchanges RTS/CTS control frames with its destination of interest and both remain in the same hop for the entire data transmission. Other stations which are not involved in communication keep on pursuing the original hopping sequence. In *Receiver Initiated Channel Hopping with Dual Polling (RICHDP)* [23], the establishment of a link layer connection is accomplished by the receiver, while technically its functionality is similar to that of HRMA.

The *Channel Hopping Multiple Access (CHMA)* [24] and the *Slotted Seeded Channel Hopping (SSCH)* schemes [25] employ an analogous channel hopping approach with minor variations on the hopping pattern generation. In the abovementioned protocols, the source station follows the destination node’s hopping schedule. If two nodes successfully exchange control frames, they will stay on the current channel to accomplish the data transaction.

In the *Parallel Rendezvous Multi-Channel MAC (McMAC)* [26], each device picks a seed to generate a different pseudorandom hopping sequence. When a station is idle, it follows its “home” hopping sequence. Each device puts its seed in every frame that it sends, so its neighbors eventually learn its hopping sequence. When a station has data to send, it flips a coin and transmits with some probability during each time slot. If it decides to transmit, it will tune into the current channel of the destination station and send an RTS. If the destination node replies with CTS, both devices stop hopping to exchange data frames. Data transaction normally takes place over several time slots. Upon completion of data transmission, both stations return to their original hopping sequence, as if no pause in hopping had happened. SSCH and McMAC are similar, as they allow stations to rendezvous simultaneously on different communication channels; however, there are also slight differences [27]: In SSCH, each station chooses four different hopping sequences and time multiplexes them to form a single hopping sequence. Nodes adjust their hopping sequences over time to the traffic
but are not allowed to deviate from their hopping sequences. In McMAC, each device has an unchangeable hopping sequence. Nevertheless, they are allowed to deviate from their default hopping sequence temporarily. In SSCH, a sender must wait until its current channel overlaps with that of the receiver before it can commence a communication. In contrast, in McMAC the source terminal may temporarily deviate from its sequence to jump to the receiver’s channel to begin data transaction.

The On-Demand Channel Switching (ODC) [28] protocol aims at minimizing the negative impact of switching delay amongst different channels by forcing stations to remain in the current channel as long as traffic conditions on that channel are acceptable. Stations continuously evaluate channel conditions to enable switching decisions. As all channels are equal in ODC, finding intended receivers is more difficult. Besides, ODC performance is not uniform and depends on the traffic pattern [27].

The Multi-channel MAC (MMAC) [29] adopts the Announcement Traffic Indication Messages (ATIM) of the IEEE 802.11 Power Saving Mode (PSM). It utilizes a default control channel that all devices should periodically switch to and synchronize for a predetermined time interval. Within the ATIM window stations with frames to send follow a three-way handshake (ATIM/ATIM-ACK/ATIM-RES) to negotiate an available data channel. Communicating devices may subsequently switch to the agreed data channel and contend for the medium using an RTS/CTS/Data/ACK access mechanism. The Multi-channel Access Protocol (MAP) [30] is basically based on the same idea as MMAC; it splits the time into control periods, during which all devices switch to the common control channel for control signaling, and data periods, during which data transactions are accomplished [27]. By dividing the super-frame interval into an alternating sequence of control and data exchange phases, both MMAC and MAP attempt to overcome the multi-channel hidden terminal problem and the requirement of having two transceivers per station. These protocols, on the other hand, lead to earlier saturation of the control period (e.g., the ATIM window in MMAC [29]); thus, in dense wireless networks, the performance of MMAC and MAP degrades significantly. Furthermore, the access delay per frame in MMAC and MAP is drastically larger than other multi-channel protocols, for which a globally known dedicated common control channel and corresponding control-transceivers are used.

In [31], three asynchronous channels are defined, in each of which the beacon interval is divided into three subintervals. ATIM and ATIM ACK frames are
exchanged in the ATIM window, determining the communication pairs in the data sub-frame. Stations that have completed their transmission in the first subinterval of the first channel switch to the second channel. Furthermore, the stations that have completed their data transactions in the second subinterval of the first channel should switch to the third channel for possible subsequent transmissions. Likewise, the radio stations that have finished their data transmissions on the second channel should either switch to the third or the first channel. The Load Awareness Multi-channel MAC (LAMM) [32] extends the protocol in [31], letting the system to adjust dynamically the number of data channels being used according to the network load.

The Busy Tone Multi-Channel (BTMC) [33] uses a set of \( k \) hash functions \( h_0, h_1, \ldots, h_{k-1} \), each of which returns a channel number from 0 to \( m-1 \) (where \( m \) is the number of available channels) when given the MAC address of a radio station. All stations are assumed to be aware of each other’s MAC addresses and to have the same set of hash functions. The source station uses the first hash function driven by the destination station’s MAC address to get the index number of channel, where the rendezvous is supposed to happen. If this channel is found unavailable for communication, the transmitter will try the next hash function on the list and keep going until they meet each other on one channel. A radio station activates the busy tone of the channel where it is receiving a data frame to inhibit any other station from utilizing that channel. A station waits on an idle channel with the smallest index \( i \), determined by applying its MAC address to hash function \( h_i \). When it receives a control frame intended for another station, it switches to the next idle channel (using the next hash function on the list) and waits there for the entire time needed until its default channel becomes available.

In [34], the TDMA-based Multi-channel MAC (TMMAC) protocol is proposed. TMMAC requires only one half-duplex radio transceiver per station. In addition to explicit frequency negotiation, which is adopted by conventional multi-channel MAC protocols, TMMAC introduces lightweight explicit time negotiation. This two-dimensional negotiation enables TMMAC to exploit the advantage of both multiple channels and TDMA, and achieve aggressive power savings by allowing users that are not involved in any data communication, to go into doze mode. Moreover, TMMAC dynamically adjusts its negotiation window size based on different traffic patterns, which further improves the communication throughput and its energy efficiency.

In [35], the Power Aware Medium Access Control with Signaling (PAMAS) is proposed. The key idea of PAMAS is that all RTS/CTS exchanges should be
conducted over the signaling radio channel while the data transmissions are kept separate over a dedicated data channel. Upon reception of a data frame, the destination node commences to send out a certain busy tone over the signaling channel. Wireless stations are mandated to listen to the signaling channel to deduce when it is advantageous (and optimal) for them to power down their transceivers. Each station makes its own decision on whether to power off. A station powers itself off if it has nothing to transmit and also realizes that its neighbors are willing to deliver something over the shared medium. In addition, it is mandated to power off if at least one of its neighbors is going to transmit a data frame to another station that is also its adjacent neighbor. The authors have developed several rules to determine the length of a power-down state. They also mention briefly some strategies for using the proposed scheme with other protocols, like FAMA [17]. They have also noted that the use of ACK and batch transmission of multiple frames together would enhance the performance of PAMAS; however, the radio transceiver turnaround time, which might not be negligible, was not considered in the design and performance evaluation of the PAMAS scheme.

The Dual Busy Tone Multiple Access (DBTMA) [16] splits a single channel into two sub-channels: a data channel and a control channel. Data frames are delivered over the data channel while control frames (RTS/CTS) are transmitted on the control channel. In addition, a couple of busy tones are employed concurrently: Transmit Busy Tone ($BT_t$), which implies that a wireless station is transmitting over the data channel, and Receive Busy Tone ($BT_r$), which indicates that a mobile station is receiving on the data channel. In DBTMA, no acknowledgment is sent to confirm the successful reception of a transmitted DATA frame, which is evidently deficient for unreliable wireless channels. Moreover, potential collisions between acknowledgments and other frames could considerably degrade the performance. Similar to DBTMA, the MAC with a Separate Control Channel (MACSCC) [36] deploys two separate channels, one as the control channel and the other one as the data channel, taking into account that the data channel is assigned more bandwidth than the control channel. It should be noted that control frames RTS and CTS can be transmitted not only over the control channel but also over the data channel in order to reduce transmission time, as long as the transmitter senses that both channels are idle. MACSCC also uses two distinct network allocation vectors for the data and the control channels, respectively. The use of two NAVs makes it possible for the control channel to
schedule not only the current data transmission but also the next data transmission, thereby reducing the backoff time.

In [37], the Dual Channel Bi-directional Multiple Access (DCBMA) with a couple of dedicated channels, namely control and data channels, and a slightly modified RTS/CTS handshake pattern has been proposed. The control channel is used for RTS and CTS transmission while the data channel is utilized for exchanging data and ACK frames between communicating wireless stations. When the source node has a data frame, at first it senses the control channel. If the channel is sensed to be idle, it sends an RTS frame to the receiving node. Upon intact reception of RTS, the receiver sends an CTS frame back to the source station immediately. The neighbors overhearing the RTS frame defer from accessing the control channel to avoid collision with the CTS frame that is going to be delivered by the receiver. Those neighbors overhearing the CTS frame defer from accessing the data channel in order to avoid interfering during the data frame reception. Up to this point, there is no major difference between the proposed DCBMA and the multi-channel protocols introduced earlier, but the difference becomes apparent when the source station commences to transmit the data frame over the dedicated data channel. In this phase, while the source node starts delivery of data, it also sends a Transmission Confirmation (TC) frame over the control channel to inform the neighboring stations about its frame transmission over the data channel. Stations that overhear TC but have not heard CTS are deemed to be exposed terminals and cannot receive any data frame until the source node finishes the data transmission; however, they still have the ability to transmit over the control channel. When the receiving node captures the data frame successfully, it replies by an ACK frame by the use of a dedicated data channel. In addition to the aforementioned control frames, the authors proposed the use of an extra control frame, namely Not-To-Send (NTS), to alleviate the problems with exposed receiver and intruding nodes.

Multiple Transceiver

In multiple transceiver MAC protocols, it is assumed that the wireless station is equipped with multiple half-duplex transceivers, so it is capable of accessing different channels simultaneously. In fact, this is the key approach to overcome multi-channel hidden- and exposed terminal problems.

In the Dynamic Private Channel (DPC) [38], devices are assumed to be equipped with as many transceivers as the number of channels. Furthermore, one
globally known channel is reserved as the common control channel for negotiation purposes. As a dedicated transceiver is associated with the control channel, the multi-channel hidden terminal problem is eliminated. RTS and reply-to-RTS control frames are used to select a data channel for communication. Once the desired data channel is negotiated, stations exchange CTS/Data/ACK frames through the transceiver associated with the selected data channel. Similar to DPC, the Dynamic Channel Assignment (DCA) [39] defines a pre-determined common control channel. RTS/CTS control frames are exchanged in this channel in order to negotiate a possible deployable data channel for data transmission. The Common Spectrum Coordination Channel (CSCC) [40] is an extension of the DCA protocol, allowing different types of devices to share the available radio spectrum. This is accomplished through negotiation on the common control channel.

In the Primary Channel Assignment-based MAC (PCAM) [41] protocol, each station uses three half-duplex transceivers: Two transceivers, the primary and the secondary transceivers, are used mainly for data communication purposes. The third transceiver is used for transmitting and receiving broadcast messages. The primary transceiver is assigned a fixed data channel, called the “primary channel”, and is used to let other stations contact each other in their primary channels. In contrast, the secondary transceiver is basically used for data transmission and is not assigned any fixed channel. In PCAM, it is explicitly presumed that each radio station is aware of other stations’ primary channels. Technically, this limiting assumption is no more valid in multi-channel mobile networks, where the wireless stations’ neighborhood is changing rather frequently. In such scenarios, the amount of overhead required for constructing each station’s knowledge base about other nodes’ primary channels will be unacceptably high. More importantly, it is not shown how wireless stations are able to be informed about their neighbors’ primary channels without employing a globally known common control channel. Considering the assumptions made in [41], it is almost impossible for a station to inform its neighbors about the chosen primary channel without plentiful broadcast messages transmitted on all available channels. Despite the fact that PCAM is shown to outperform MMAC [29] and DCA [39] when the number of channels is relatively large, such scenarios are rather rare in practical multi-hop cognitive networks.

Finally, the Extended Receiver Directed Transmission (xRDT) [42] defines separate busy tones for different data channels. In this protocol, each station chooses one of the available data channels as its quiescent channel. A wireless
station, when receiving a data frame on its quiescent channel, turns on the corresponding busy tone of that channel to inform potential transmitters about the ongoing frame reception; hence, the multi-channel hidden terminal is resolved.

Furthermore, a radio station, upon returning to its quiescent channel, broadcasts a designated message, called Data Transmission Complete (DTC), to inform any awaiting transmitter to commence transmission of pending frames that are addressed to the same station. Therefore, using this approach xRDT needs to have only one half-duplex data transceiver, in addition to a switchable busy tone interface. This protocol also suffers from the same problem as explained for PCAM. On the other hand, it overcomes the problem caused by hidden terminal and the difference between the transmission and interference range.

Dissimilar to methods that use a common control channel, some medium access schemes do not rely on any control channel. Instead, they are flexible in organizing different channels for RTS/CTS/DATA/ACK to reduce the probability of collisions. Both Interleaved CSMA (ICSMA) [43] and Jamming-based MAC (JMAC) [44, 45] are such protocols, which split the entire bandwidth into a couple of channels and employ one half-duplex transceiver for each channel. ICSMA uses two channels of equal bandwidth. Each station is permitted to commence transmission in either dedicated channel. The transmitter sends RTS and DATA on one channel, and the receiver responds by sending CTS and ACK on the other channel. This scheme supports simultaneous data deliveries between two stations. That is to say, when a mobile station is sending RTS or DATA, or receiving CTS or ACK from the other node, the latter one is also sending the same kind of frames on a different channel to the former one. In JMAC [45], the medium is divided into two channels: the S channel and the R channel. RTS and DATA are transmitted over the S channel, and CTS and ACK are transmitted on the R channel. A transmitter also transmits jamming signals on the S channel while waiting for (or receiving) a CTS/ACK frame on the R channel. For a receiver, while it is waiting for (or receiving) a DATA frame on the S channel, it jams the R channel to prevent neighbors from transmitting RTS frames on the S channel. The jamming signal is the only one that, with sufficient energy, can cause the media to become busy. Since it will stop if the RTS/CTS exchange fails, it resolves the erroneous reservation problem in the IEEE 802.11 protocol. Additionally, it effectively blocks hidden terminals from transmitting.

In the Directional Busy Signal Multiple Access (DBSMA) [46] all the transmissions, receptions, and idle listening tasks are performed in a directional manner. The need to listen in many directions when in an idle state is achieved by
constantly changing the listening direction to cover the whole space. The authors also propose a directional backoff mechanism in which every station maintains independent backoff windows for each direction and show how it yields a better performance. DBSMA has some special features including: 1) While in an idle state, the station sweeps all the directions continuously to listen, i.e., it rotates the listening direction of its directional antenna continuously so as to listen in all directions; 2) After reception of an RTS frame, the receiver sends a unique directional CTS frame to the transmitter using the direction in which it received the RTS. As a result, it does not send any other control frame in any other directions; 3) While in the reception mode, the receiver continuously transmits a busy tone as in the BTMA protocol till the end of the DATA and ACK transmissions. The busy signal is sent along the reception direction. The busy signal is a narrow bandwidth out of band signal; 4) A node that wants to transmit senses the channel in the direction of the receiver (i.e., the next-hop destination), and if the channel is perceived idle, it should try sending a unique directional RTS to the destination. However, before sending the RTS, it sends an Invitation Signal (IS) in the same direction. The invitation signal is a narrow-band signal like the busy tone signal and is sufficiently long so that all idle neighbors can hear it, irrespective of their initial sweeping state, and lock on that direction for reception of the actual RTS message. In addition to the aforementioned features, the authors added some extra enhancements to achieve better performance in multi-hop scenarios, namely Directional Backoff (DBO) and the Neighbor Direction Detection (NDD) scheme. In DBSMA, the authors proposed to use an independent backoff for each direction. Essentially, a wireless station with an antenna beam width of $X$ (in radians) can be considered a set of $2\pi/X$ pseudo-stations, which are active in different directions. These pseudo-stations are able to run their backoff mechanisms autonomously. Alternatively, we may think that each direction behaves as if it were an independent link.

### 1.2.5 Cognitive multi-channel MAC protocols

It is known that cognitive radio is a promising solution to the spectrum scarcity problem in wireless access networks. Cognitive radios are capable of using underutilized frequency bands without interfering with licensed incumbent devices. They scan potential frequency bands, find the spectrum holes, and adjust their parameters, such as frequency, power, and transmission rate, in order to use the spectrum more efficiently. One of the primary goals of cognitive radio
technology is to enable *Opportunistic Spectrum Access (OSA)*. With respect to this fundamental objective, the MAC layer plays a key role in achieving efficient and reliable radio resource allocation, either in centralized or distributed cognitive wireless networks. So far, only a few cognitive MAC protocols have been proposed in the literature, and none of the existing solutions address dynamic channel selection, load balancing, multi-channel contention resolution among *Secondary Users (SUs)*, as well as avoidance of *Primary Users (PUs)* at the same time.

The field of robust MAC protocol design for multi-channel frequency agile wireless networks is an emerging area in cognitive radio technology. As a well-known example, the IEEE 802.22 working group is responsible for standardizing a cognitive MAC protocol to facilitate dynamically accessing UHF/VHF TV bands between 54 and 862 MHz. This standard focuses on constructing consistent and national fixed *Point-to-Multi-Point (P2MP) Wireless Regional Area Networks (WRANs)* [47].

In [48], the *Cognitive MAC (C-MAC)* protocol employs the concept of a dynamic *Rendezvous Channel (RC)* in order to coordinate wireless stations in different channels, multi-channel resource reservation, as well as *Quiet Period (QP)* coordination for incumbent detection. Among all available channels, the RC is set to the most reliable channel. Furthermore, a *Backup Channel (BC)* is used to make the RC robust enough against cases when the incumbents begin accessing the occupied frequency band. The BC is determined by out-of-band measurements carried out by all stations whenever they are not involved with any communication. In C-MAC, each data channel is logically divided into recurring super-frames that begin with a slotted *Beacon Period (BP)* followed by a *Data Transfer Period (DTP)*. During the BP, each device transmits a beacon frame in its designated time slot. Beacon frames contain information about scheduled QPs, spectrum measurements, and multi-channel reservation for data communication. Once the BP is over, stations may switch to other channels for communication.

A full-duplex multi-channel MAC protocol for multi-hop cognitive radio networks is proposed in [49], where each station is equipped with at least two radio transceivers, one for transmitting and the other one for frame reception. Each station selects an unused frequency band as its *Home Channel (HCh)* and tunes its receiving transceiver to that channel. To transmit data frames, one should tune its transmitter to the destination station’s HCh and contend for the shared medium on that channel. The home channel of a radio station is changed when: 1) it becomes unavailable due to a primary user, 2) one or more neighboring station
disagrees with the chosen home channel while there are other channels available for deployment, or 3) the station experiences a high collision rate caused by noise or undetected hidden terminals. Each station is responsible for informing its neighbors about its chosen home channel, using a pre-defined common control channel.

In the Opportunistic Spectrum MAC (OS-MAC) [50], all inter-channel control frames are exchanged via the common control channel, whereas data channels are used for communicating all data frames and all intra-channel control frames. Each data channel always has a Delegate Secondary User (DSU) appointed among users currently using it. All DSUs (one from each data channel) periodically switch to a common control channel to inform each other of the traffic loads experienced on their data channels. Upon learning of the conditions of all data channels, each DSU returns to its original channel and informs all Secondary User Groups (SUGs) currently using that channel about the traffic conditions of all the other channels. Based on this information, each SUG selects and subsequently switches to the “best” channel for data communication until the end of the current super-frame. While DSUs are on the common control channel, all other secondary users continue using their data channels for data communication.

1.3 Aims and outline of the thesis

In this thesis, a few possible approaches to migrating from traditional MAC protocol design techniques to the next generation protocols and algorithms particularly designed for distributed frequency agile ad hoc networks are investigated. All results presented in this thesis have been either published or accepted for publication in peer reviewed journals and conferences. For the sake of clarity and uniformity, the thesis is presented as monograph.

In Chapter 1, the majority of the contents of which has been presented in [150 – 153], the state-of-the-art in medium access control protocol design and the existing challenges in this field are studied. A comprehensive literature review is presented in the area of MAC algorithm design for multi-hop ad hoc networks, as well as frequency agile cognitive networks.

Chapter 2, the results of which have been presented in [154 – 158], presents modern techniques to accurately analyze a non-saturated (i.e., finite load) single-hop and multi-hop ad hoc network. Several approaches to further extend the framework to arbitrarily larger networks with more complex system configurations are introduced. This chapter helps to gain a better understanding of
the most important challenges in ad hoc networks, such as the hidden terminal problem and the unreachability problem.

Chapter 3, the results of which have been presented in [159 – 161], is composed of our contribution in the field of medium access control protocol design in multi-hop ad hoc networks. By introducing this protocol, named eMAC, the unreachability problem is tackled, and through extensive simulations it is shown how this scheme alleviates the negative impact of the receiver blocking problem.

Chapter 4, the results of which have been presented in [162, 163], presents an adaptive stochastic channel selection and multi-channel contention resolution scheme for distributed cognitive ad hoc networks. By incorporating primary user probability of appearance on different data channels, a simple yet effective channel selection technique, as well as a multi-channel contention-resolution mechanism, is proposed. Through extensive simulations, the performance of the presented algorithms is evaluated, showing their superior performance in comparison to the existing protocols in a wide range of considered scenarios.

Eventually, Chapter 5 concludes the thesis. The main results are summarized and several open problems for future research are pointed out.

1.4 Author’s contribution

The thesis is mainly based on five journal papers [154, 158, 159, 161, and 163], four book chapters [150 – 153], and five conference papers [155, 156, 157, 160, and 162]. The author has had the main responsibility for performing the analysis, developing the simulation software, generating the numerical results, and writing all the aforementioned papers. Other authors provided invaluable help, support, comments, and criticism during the postgraduate studies at the Centre for Wireless Communications, University of Oulu. All results and analysis presented in this thesis have been produced by the author.
2 Performance analysis of contention-based medium access control

In this chapter, a set of techniques for precisely analyzing finite load single- and multi-hop ad hoc networks is presented. The focus is mainly drawn on normalized throughput, MAC frame end-to-end delay, as well as frame jitter. Furthermore, the proposed models are validated by comparing numerical results to the simulations results obtained from OPNET™. In the first part of this chapter, a single-hop ad hoc network scenario with a fixed number of contending stations is taken into consideration. An accurate framework based on Markov processes is presented to model the binary exponential backoff sequence when the wireless stations are not saturated (i.e., their transmission queue may become vacant). Subsequently, channel access delay is carefully studied using the proposed Markovian framework, taking into account frame delay due to contention resolution as well as frame delay due to queuing delay in a tagged station’s transmission buffer. Eventually, it will be shown how the aforementioned framework can be extended to model multi-hop scenarios. The extended framework, not only models a multi-hop ad hoc network, but also takes into account the case where the contending stations are non-saturated. Numerous approaches to further generalize the framework to larger networks with more complex system configurations are introduced.

The remainder of this chapter is organized as follows. An overview on specific contributions on the performance analysis of the contention-based 802.11 MAC is given in Section 2.1. Assumptions are then given in Section 2.2 and, subsequently, a Markovian framework is presented in Section 2.3 to model a single-hop non-saturated ad hoc network and, subsequently, MAC frame delay and jitter are studied for such networks using the aforementioned framework. The model is further extended in Section 2.4 to multi-hop systems, and extensive discussions will be provided on possible approaches through which the model can be extended to more complex topologies. Finally, Section 2.5 concludes the chapter.

2.1 Overview and background

The IEEE 802.11 DCF adopts Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) and retransmissions of collided frames are managed according to the binary exponential backoff rule [51]. The DCF defines two
media access techniques to be employed for frame transmission. The default scheme is a two-way handshaking technique called the basic access mechanism. In addition, an optional four-way handshaking technique known as the RTS/CTS mechanism is defined to reduce the hidden terminal problem. Before a station performs any frame transmission, it has to sense the wireless medium. If the shared media is found “idle” for a minimum time equal to the Distributed Inter Frame Space (DIFS), the head-of-line MAC frame, if available, will be transmitted. Otherwise, the station enters into backoff procedure and randomly sets its backoff timer within the range of the contention window. The backoff timer is decremented by one when the channel is sensed “idle” and is frozen when the medium is sensed “busy”. When it reaches zero, the station commences the frame transmission as the next consequence. Upon correct reception of a delivered frame, the receiver is required to send an ACK after a time equal to Short Inter Frame Space (SIFS). If no ACK is received, the sending station assumes either a collision or frame corruption, doubles its current CW, randomly resets its backoff timer, and retransmits the previous frame when the timer reaches zero again.

In [52], a Markov chain-based analysis of the DCF is presented and the saturation throughput in a single-hop network with the assumption of an infinite number of possible retransmissions per frame is derived. In [53], the Markov chain model of [52] was extended to include finite frame retry limits as defined in [51]. An analytical model for non-saturated sources was proposed in [54] with the assumption of a very low traffic load. In [55], a different approach was taken to analyze the performance of the DCF under a statistically varying traffic load. The on-off characteristics of the stations are modeled by a state-dependent single server queue where the service times for the different states were estimated from the saturation throughput obtained in [52]. Moreover, the model in [55] is not very accurate in modeling finite load scenarios since it assumes that a station reaches directly its saturation throughput, which is not reasonable if the active time of a station is comparable with the transitory regime duration. As a whole, in the literature, in order to analyze the DCF under the finite load condition, the MAC layer transmission queue was modeled as an\( M/G/1 \) queuing system (e.g., [56]). However, this assumption degrades the accuracy of analysis due to the fact that only the first moment of MAC layer frame service time is taken into account for different scenarios in order to calculate the probability of the transmission queue being empty. The authors in [57] present a new Markov chain correcting that of [58] by evaluating the channel access probabilities and the station collision
probabilities conditioned upon the channel status. Their analytical model takes into account the type of previous channel state, either idle or busy, in addition to the backoff stage and current backoff counter as in [52]. In [59], the authors propose a new approach to evaluate the throughput/delay performance of IEEE 802.11 DCF based on the elementary conditional probability arguments rather than bi-dimensional Markov chains as in [52] and [53]. In [60, 61], the authors incorporate the notion of so-called *intermediate* states for a tagged station’s backoff stochastic process to take into account the independent media access of hidden stations.

In [62] the authors deal with goodput and link adaptation for IEEE 802.11a wireless LANs. More precisely, they have presented a generic method to analyze the goodput performance of an 802.11a DCF system and derived a closed-form expression of the expected effective goodput. In this case, when a wireless station is ready to transmit a data frame, its expected effective goodput is defined as the ratio of the expected delivered data payload to the expected transmission time, i.e., the expected bandwidth a terminal can actually receive after all the overheads are taken into account, including MAC/PHY overheads, backoff delay, inter-frame intervals, acknowledgment transmission time, and the potential frame retransmission times. In [63], the same authors as in [52] presented an approach to evaluate throughput/delay of 802.11 DCF which relies on elementary conditional probability arguments rather than bi-dimensional Markov chains.

In [64], the authors present an analytical model to evaluate the performance of DCF in imperfect wireless channels from a novel perspective. In this study, they consider the impact of different factors together, including the binary exponential backoff mechanism in 802.11 DCF, various incoming traffic loads, incoming frame size distribution, the queuing system at the MAC layer, and imperfect radio channel condition. In [65], the saturation throughput of IEEE 802.11 DCF in a single-hop wireless network and by the use of queuing theory has been investigated. In [66], the IEEE 802.11 DCF under non-saturation condition has been analyzed with the help of the theory of Markov chains. Actually, the presented investigation is not a highly accurate study but can be assumed to be a reasonable motivation to exploration of non-saturated IEEE 802.11 wireless networks. In this study, the authors assumed that each station can buffer only one frame and there is a constant probability of at least one frame arriving per state; an unusual assumption to make the analysis straightforward. In [67], a comprehensive investigation of the IEEE 802.11 DCF from a new perspective has been presented. The authors’ approach differs fundamentally from
the other works since they do not make the decoupling assumption introduced by [52]. Instead, starting from a Markov chain description that explicitly takes into account the interactions between stations, they try to show that in a large system, namely one with a large number of stations, the Markov chain converges to a typical state. Thus, one can approximate the collision probability seen by any single station by that seen in the typical state. The key assumption has been made by the authors is that the backoff durations are geometrically distributed. In [68], the authors use a discrete Markov chain model to analyze the throughput of the IEEE 802.11 CSMA/CA scheme, taking into consideration the capture phenomenon, which means the frame with the strongest power may capture the receiver even in the presence of other overlapping frames. Moreover, the performance of the IEEE 802.11 CSMA/CA protocol in the presence of path loss, shadowing, and Rayleigh fading is modeled analytically and compared with the model without capture effect. Many other contributions [69 – 73] have partially continued some of the existing works to some extent; however, the lack of sufficient accuracy in the performance study of the contention-based medium access has still been evident.

In [74], a discrete time \( G/G/1 \) queue for modeling nodes in a random access network based on the 802.11 MAC is introduced, and an extensive delay analysis of the random medium access of the 802.11 DCF is conducted. In [75], statistical characteristics of the service time of IEEE 802.11 in saturated wireless networks have been analyzed. The authors derived a closed-form probability generating function for the frame service time of a cluster of 802.11-based terminals for both basic and RTS/CTS media access schemes. In addition, by the use of numerical methods, the obtained generating function has been inverted to derive the probability distribution function of the frame service time. Another work that deals with IEEE 802.11 access delay can be found in [76]. Here, once more, the generating function of the IEEE 802.11 DCF access delay has been investigated. In addition, the authors further inverted the generating function using the Lattice-Poisson numerical method developed by [77].

Multi-hop ad hoc networks are spontaneous systems that consist of wireless stations, either mobile or stationary, devoid of any pre-installed base station or central coordinator. They can be applied in many contexts such as military communication, disaster rescue, and outdoor/indoor activities due to their features of convenience in deployment and flexibility in reconfiguration. In addition to the hidden terminal problem, there are scenarios where the destination is located in the radio range of another transmitting station, so that the efforts to set up
communication with this terminal will fail due to collisions that may occur between transmitted control frames destined to the desired receiving wireless node and unwanted received control or data frames originating from neighboring hidden terminals. In such scenarios, the desired destination is called unreachable for the duration of neighboring hidden nodes’ data transfer. In [19 – 21], practical solutions were proposed to alleviate the aforementioned problem by using extra control frames and by adding more intelligence to the existing 802.11 DCF MAC. When a station is notified about an upcoming data communication due to which it will be unreachable, it is given an opportunity to inform its one-hop neighbors about the forthcoming unreachability. In principle, right after the RTS/CTS negotiation and before commencing the actual DATA transmission phase, the stations that will become unreachable shortly are given the chance to report their imminent unreachability status using a designated broadcast frame called ICP (i.e., individual communication pause).

The authors in [79] proposed a simple Markov chain-based approach to analyze a two-hop wireless system, consisting of only three stations, two of which are hidden with respect to each other. However, their scheme did not consider RTS-to-RTS collisions due to simultaneous RTS transmission by the aforementioned two hidden nodes. The authors in [80] presented an analytical framework for performance study of MAC protocols and their interactions with physical layer parameters in multi-hop ad hoc networks under saturation conditions. On the other hand, in [81 – 86], the capacity of IEEE 802.11 DCF had been analyzed using different methodologies. Both single- and multi-hop wireless systems were considered, while in some of them the impact of physical layer parameters on the overall system performance had been also taken into account. Moreover, recent research papers in the field of multi-hop ad hoc networks analysis have been based on Bianchi’s model [52], which was intended for single-hop saturated networks. The authors in [87] gave an overview of all possible interference topologies in a multi-hop wireless network and computed the aggregate saturation throughput based on the scheme in [52]. Furthermore, [88] and [89] had also derived expressions for similar problems using Bianchi’s time model and renewal theory as the core of their analysis, independently. Although the aforementioned contributions provide computationally efficient means for performance evaluation of CSMA/CA-based networks, their applicability is limited to certain node configurations. While the main contribution of [90] is on throughput modeling and fairness issues in CSMA/CA-based wireless ad hoc networks, in [91, 92], the authors attempted to provide an understanding of the
subtle complexities behind multi-hop 802.11 networks. Specially, in [89] the authors utilized Lorenz curves and the Gini index, originally used in the economics literature to quantify a society’s distribution of wealth to individuals, to evaluate the network’s distribution of flows. On the other hand, they explicitly assumed that a similar expression of conditional transmission probability $T$ [52] as a function of conditional collision probability $P$ [52] could be derived for different system configurations. In contrast, in this thesis it is shown that the aforementioned assumption is not necessarily valid for multi-hop ad hoc networks and, as a result, there is a vital need to have more in depth investigation regarding the interaction of wireless stations with different viewpoints with respect to the whole ad hoc network. Finally, in [93] the impact of RTS frame duration was analyzed.

In this chapter, an analytical framework based on the Parallel Space-Time Markov chain (PSTMC) is presented to simultaneously model the backoff and post-backoff procedures, in addition to the transmission queue status. The structure of PSTMC includes a set of parallel space-time chains; each one represents the backoff procedure of a tagged wireless station when its transmission buffer holds a specific number of frames prepared for delivery. Therefore, the key feature of the proposed model is that for each transmission queue occupancy (i.e., buffer utilization corresponding to the number of awaiting MAC frames in it) the ongoing backoff (or post-backoff) procedure is modeled using a distinct chain of backoff states known as the space-time chain or simply a space. In other words, the overall space-time Markov chain consists of a set of parallel space-time chains, while the number of the aforementioned spaces is exactly equal to the MAC layer transmission queue capacity. Basically, each space represents the tagged station’s backoff procedure for specific queue occupancy in the time domain. This makes the philosophy of our proposed terminology, “Space-Time”, for the presented framework clear; “Space” indicates the modeling of the “Transmission Queue” which is in fact a physical dimension-dependent feature, while “Time” indicates the modeling of the “Backoff Process”, which is essentially a time-dependent aspect. Upon transition between consecutive states, the proposed framework monitors the status of the transmission queue. Hence, during an ongoing backoff procedure, it is possible to jump to another space due to the arrival of a new number of MAC frames.
2.2 Assumptions

A single-hop WLAN is assumed in which all stations are in the radio range of each other constructing a fully connected complete graph and the transmission queue of a wireless station may become empty at some periods during the network lifetime. As defined in the standard [51], a station with a frame ready to transmit typically has to run at least one backoff between two successive transmissions. Thus, after each successful transmission, the contention window size is reset to its minimum value $W_0$ and the station enters into another backoff cycle. If there is a MAC frame, it initiates a fresh backoff procedure; otherwise, if the queue is empty, it enters a so-called post-backoff process. In the latter case, the contention window is again reset to its initial (minimum) value $W_0$ and is decremented according to the DCF access rules. By the end of post-backoff cycle, the station enters an idle state (referred to as the vacancy state) and waits until at least one frame enters the queue. Upon the entrance of at least one frame, while being in the aforementioned state, the station initiates an Asynchronous Frame Transmission (AFT) and without experiencing any further backoff transmits the head-of-line frame over the air.

The main issue of major importance in the proposed framework is the way the backoff procedure is modeled. Although the analysis of the backoff process under saturation condition is achieved by the use of a single-space Markov chain [52, 53], it is not possible to use such a simple framework to model a finite load situation due to the need to take into account queue evacuation during the tagged station activity. A feasible approach to model the “queue status” is to have a singular chain for every possible queue status. This means that different backoff processes are discriminated by different (but analogous in structure) stochastic sub-processes constructing together a multi-space Markov chain, named PSTMC, in order to represent both backoff and transmission queue status simultaneously. Therefore, it may be concluded that for each queue occupation state corresponding to a particular number of buffered MAC frames in it, one shall have a unique space-time chain that models the backoff procedure for that particular queue status, which all together establish a multi-space Markovian framework to model both the backoff and queue status concurrently. Since frame entrance to the MAC layer can occur at any time, there is no guarantee that the queue status will not change during an ongoing backoff procedure. For that reason, throughout the backoff timer decrement, the proposed model monitors the transmission queue continually in order to switch to the next backoff state.
positioned in another space-time chain. In other words, switching between different spaces occurs only due to alteration in the number of buffered frames in the MAC layer queue. In addition, during the post-backoff cycle it is mandatory to verify the transmission queue to see whether any new MAC frame has just buffered or not. Upon at least one new frame entrance, the MAC entity switches from the ongoing post-backoff sequence to a backoff cycle corresponding to the new queue occupation status and continues decreasing the backoff timer accordingly. Consequently, it may be concluded that in the PSTMC framework, each backoff state is reachable by a series of particular backoff (or post-backoff) states placed in the same and/or other spaces.

2.3 Non-saturated performance analysis: single-hop perspective

In the PSTMC model, the queue status is represented by a stochastic process \( q(t) \), representing the number of buffered frames at the time \( t \) for a given station (tagged station). In addition, the stochastic processes \( s(t) \) and \( b(t) \) represent the backoff and post-backoff procedures, respectively [52, 53]. The stochastic process \( s(t) \) symbolizes the backoff stage, while the process \( b(t) \) represents the backoff counter for a given station at time \( t \). Based on the independence condition [52] and the fact that frame collision is considered to be a Bernoulli process, it is possible to model the three-dimensional continuous process \( \{s(t), b(t), q(t)\} \) with a discrete-time Markov chain in which each backoff state is represented by \( B_{i,k,j} \), the stationary distribution of the chain, where

\[
B_{i,k,j} = \lim_{t \to \infty} P\{s(t) = i, b(t) = k, q(t) = j\} \quad \text{for} \quad i \in [0,m], k \in [0,W_i - 1], j \in [0,L]
\]

(\( m \) is the maximum allowed (re)transmissions per frame and \( L \) is the MAC layer queue capacity). Consequently, for every non-negative integer \( j \neq 0 \), the ongoing backoff cycle is represented by a single space-time chain denoted by \( S_j \) where \( j \) indicates that \( j \) MAC frame(s) are currently buffered inside the queue. If the queue capacity of the tagged station is \( L \), exactly \( L \) parallel space-time chains will be required; each one represents the backoff process of the tagged station when its transmission queue contains \( j \) MAC frame(s). Besides, for future purposes, it is worth mentioning that medium busy period due to a successful channel access is denoted by \( T_s \), which can be rewritten as the product form \( K_s \times \sigma \), where \( \sigma \) is the predefined system time slot duration [51]. Also, the medium busy period due to a collision is denoted by \( T_c \) throughout the thesis which is also rewritten as the product form \( K_c \times \sigma \). Furthermore, it is important to note that the time duration considered for a tagged station’s collision is throughout the thesis denoted by \( \Delta_c \).
instead of \( T_e \), since the deferral interval for colliding and non-colliding stations are different according to the IEEE 802.11 specifications [51]. Basically, \( \Delta_c = T_{\text{timeout}} + T_e \) while \( T_s \) and \( T_c \) for the RTS/CTS scheme are \( T_{\text{RTS}} + T_{\text{CTS}} + T_{\text{DATA}} + T_{\text{ACK}} + 3 \times SIFTime + DIFTime + 4 \times \delta \) and \( T_{\text{RTS}} + tEIFTime + \delta \), respectively, where \( \delta \) represents propagation delay.

2.3.1 Traffic arrival process

Due to the nature of traffic generation by the upper layers, the number of existing frames in the transmission queue may be always changing. Therefore, it is inevitable to occasionally switch between different chains, as frame(s) enter or exit the transmission queue. Departures occur upon successful transmission of a frame to the desired destination or upon dropping a frame due to reaching the maximum retry limit for frame (re)transmissions. Moreover, when the transmission queue is full, all incoming traffic will be dropped due to queue overflow. The traffic arrival process at the MAC layer is assumed to be Poisson with the mean rate of \( \lambda \) frames per second. Then, the probability of \( n \) frames arrival during time interval \( t \) is given by \( A_n(t) = e^{-\lambda t}(\lambda t)^n/n! \). Even though the Poisson assumption may not be perfectly realistic, it provides insightful results and allows the model to be tractable. Previously proposed queuing analyze in the literature assume the Poisson arrival pattern specifically due to its independence feature of non-overlapping interarrival times, leading to Markovian property to be valid [52, 53]. On the other hand, it is imperative to determine how argument \( t \) is set in order to model the state transitions between \( L \) parallel spaces. The Poisson traffic pattern makes the time argument determination straightforward.

2.3.2 Intermediate state transitions

Transition between consecutive backoff states occurs with different time durations depending on the corresponding event type taking place in the wireless system. For transitions between any two successive backoff states, one shall traverse through a series of intermediate states in order to reach the next backoff state where the number of these intermediate states relies on the type of event taking place in the wireless network (i.e., either a successful transmission or collision). In the proposed model, all intermediate transitions between two successive backoff states are performed with constant time duration. Basically, upon a successful channel acquisition, the number of intermediate states between
any two consecutive backoff states should be equal to $K_s$, and the transitions through these intermediate states are accomplished with the probability equal to one. On the other hand, being in the same space-time chain during a time slot is conditioned on the fact that for successive transitions no frame is generated by layer three. In contrast, transition to the next state but in a different space-time chain occurs with probability $A_n(t)$, where $n \geq 1$ due to the arrival of $n$ new MAC frames to the queue.

![Fig. 2. Intermediate state transitions between consecutive backoff states in parallel chains.](image)

Different state transitions are illustrated in Fig. 2, in which all states in a single row are situated in the same space representing the backoff process of the tagged station when its transmission queue holds a specific number of frames (i.e., $h, h+1, \ldots$, and $j$ as depicted in Fig. 2). Similarly, each column of states corresponds to the similar states in different space-time chains. Note that in Fig. 2, intermediate states are depicted by dashed circles while the non-negative integer argument $(x)$ represents the remaining busy period of the shared wireless media in time slots. In Fig. 2, transition from $B_{(i,k,h)}$ to the intermediate state located in the same space-time chain with an argument equal to $(x-1)$ shall be performed.
with the probability equal to \( A_0(t) \) (standing for no new frame arrival), whereas the aforementioned transition to the intermediate state located in the chain just below the current space with an argument equal to \((x - 1)\) is done with the probability of \( A_1(t) \) (representing only one new frame arrival), and so on.

In the second step, as illustrated in Fig. 3, transition from state \((n, j_0)\) to state \((n + t, j_t)\) can be accomplished by traversing through different intermediate states due to different combinations of frame arrivals.

In the second step, as illustrated in Fig. 3, transition from state \((n, j_0)\) to state \((n + t, j_t)\) can be accomplished by traversing through different intermediate states due to different combinations of frame arrivals.

\[
\sum_{n_1} \cdots \sum_{n_{k-1}} \sum_{n_t} \left[ e^{-\lambda \cdot \frac{(n_1! \cdot n_2! \cdots n_t)!}{k-j}} \right] = e^{-\lambda \cdot \frac{(n_1! \cdot n_2! \cdots n_t)!}{k-j}}
\]

For simplicity, assume that each intermediate transition is performed during a constant time period equal to one, as illustrated just below the arrows representing the transitions. The probability of each transition has been also shown just above it. The total transition probability between \((n, j_0)\) and \((n + t, j_t)\) can be written as an aggregate probability equal to \( e^{-\lambda k \cdot \frac{k-j}{k-j}} \), where \( j_t = k \) and \( \sum n_i = k - j \). As there might be different combinations of the number of frame arrivals in each step, i.e., \( n_t \), all probable transitions can be merged together to produce a single transition, which is in fact equal to \( A_{k-j}(t) \).

### 2.3.3 State transitions in the same and difference spaces

Let us assume that the current state is \( B_{i,k,j} \), meaning that we are currently situated in the \( j \)th space. Also, suppose that an event has occurred with probability \( P \) and lasts for a time duration equal to \( T \). For a state transition in the same space-time chain caused by the aforementioned event, the probability that there is no frame arrival to the MAC layer queue is given by \( P \times A_0(T) \). On the
other hand, for a transition to a different space caused by the same event, the probability that there are \( n \) frame entrances to the MAC layer queue is given by \( P \times A_n(T) \) with \( j_{\text{new}} = j_{\text{old}} + n \).

Fig. 4. State transitions in the same and between parallel Markov chains during the backoff procedure.

Fig. 4 illustrates possible backoff state transitions in the PSTMC model. These transitions are accomplished when the backoff timer of the tagged station is still non-zero. Therefore, it does not contend for the media. Thus, for all transitions’ probabilities the tagged station should be excluded. If we denote the conditional probability that a given station transmits in a randomly chosen time slot by \( \tau \) [52], the probability that there is at least one transmission in a randomly chosen time slot due to channel access of one or more wireless station except the tagged one by \( P_{tr} \), the probability of a successful transmission by a wireless station except the tagged one by \( P_s \), then for all state transitions between two consecutive backoff states \( P_{tr} \) and \( P_s \) should be used instead of \( P_{tr} \) and \( P_s \), which consider all stations including the tagged station [52]. These probabilities can be written in terms of \( \tau \) and the number of active stations in the wireless network as \( P_{tr} = 1 - (1 - \tau)^n \), \( P_s = n\tau(1 - \tau)^{n-1} \), \( P_{tr}' = 1 - (1 - \tau)^{n+1} \), and \( P_{tr}'P_s' = (n - 1)\tau(1 - \tau)^{n-2} \).

Note that in this thesis, the definition of \( \tau \) is the same as [52] but its mathematical expression is different. Also note that during the period of queue vacancy, the station does not contend for the media, and for that reason it will never initiate any backoff cycle.

Modeling the empty state of the transmission queue can be done by the use of only one state, denoted by \( E \). Also, for all state transitions starting from \( E \), \( P_{tr} \) and \( P_s \) should be respectively used instead of \( P_{tr} \) and \( P_s \). Besides, during an ongoing event, if no new frame enters the queue while being in an empty state, there should be a loop back to state \( E \).
The finite-load study of the IEEE 802.11 DCF should also deal with the analysis of AFT. While being in vacancy state $E$, there might be an arrival of $j$ new MAC frame(s) to the transmission queue. Hence, there should be a state transition to the first backoff stage in space-time chain $S_j$, if, and only if, during the first frame arrival among the abovementioned $j$ new frame(s) the medium has been sensed “busy” [51]; otherwise, there should be a so-called asynchronous frame transmission without experiencing any further backoff cycle. Thus, there should be a state transition to a different set of states denoted by $F_j$. Upon completion of asynchronous frame transmission, and depending on its final outcome (i.e., either successful delivery or collision), there will be a state transition to a proper space-time chain according to the remaining frames in the transmission queue.

2.3.4 Space–time chains state transitions

Fig. 5 illustrates the generic $S_j$ space-time chain ($1 \leq j < L$) in addition to the post-backoff, vacancy, and AFT states in an integrated manner. The vacancy state $E$ is shown next to the post-backoff states $b_i$, while AFT states are shown as $F_i$ with $1 \leq i \leq L$, just above the post-backoff states.
Fig. 5. State transitions related to $S_j$ space-time chain.
While being in states $B_{i,0,j}$, the tagged station transmits a buffered frame, which is received correctly by the desired destination with probability $1 - p$ or collides with other con-current transmission(s) with probability $p$. The same notations for conditional collision and conditional transmission probabilities $p$ and $\tau$ as employed in [52, 53] are used. Conditional collision probability, $p$, is the probability of a collision due to simultaneous media access conditioned to the fact that the tagged station transmits a frame in the next time slot. In single-hop wireless systems, $p$ is equal to $1 - (1 - \tau)^{(n - 1)}$, regardless of the system nature (i.e., saturated or non-saturated), where $n$ is the number of active stations. Upon a collision there will be a transition to the next backoff stage, while upon a correct frame departure there will be a transition to the post-backoff states if, and only if, the MAC control entity has been in $S_0$ and no new frame arrival has taken place during the completed successful transmission. The latter transitions are shown in Fig. 5 as a set of arrows coming out of states $B_{i,0,j}$ (if $j = 0$) and entering either empty state $E$ or post-backoff states $b_i$.

During a post-backoff cycle and upon new frame arrival to the transmission queue there should be a transition to an appropriate space-time chain depending on the number of frame arrivals. This means that each backoff state in the first stage (with $i = 0$) except the one with $k = W_0 - 1$, located in any space-time chain $S_j$ with $1 \leq j \leq L$, should be reachable by a particular post-backoff state with a value of $k$ (i.e., backoff step index) larger exactly by one (see Fig. 5). For instance, during the post-backoff process, if there is a single frame arrival to the transmission queue, a state transition occurs to the next backoff state located in space $S_1$, with the step index, $k$, decreasing by one and the stage index, $i$, being equal to zero as shown by oblique dashed lines in Fig. 5.

While being in vacancy state $E$, three different types of events may occur: a successful frame transmission with probability $P_{tr} \times P_s$ and time duration $T_s$, a collision with probability $P_{tr} \times (1 - P_s)$ and time duration $\Delta$, and finally an idle time slot with probability $(1 - P_{tr})$ and time duration equal to $\sigma$. Note that the aforementioned events exclude the tagged station since it has no awaiting frames inside its transmission queue. If during the abovementioned time intervals, there is no new frame arrival, there should be a loop-back to vacancy state $E$.

As explained earlier, being in vacancy state $E$, if $j$ new MAC frame(s) come into the transmission queue while the first one has arrived during the last “idle” phase of the wireless media, there should be a transition to the $F_j$ AFT state. Transition from AFT states to both backoff and post-backoff processes are also possible as illustrated in Fig. 5. For example, upon a state transition due to a
successful frame delivery coming out of state $F_1$, if no new frame arrival occurs, the MAC entity will initiate a new post-backoff cycle, meaning that the aforementioned transition will be ended by a post-backoff state (and perhaps empty-queue state, $E$) as shown in Fig. 5. On the other hand, consider a state transition due to a successful frame delivery coming out of state $F_1$. In this scenario, if another new frame enters the MAC layer transmission queue, then the MAC entity will initiate a new backoff cycle (not a post-backoff process), meaning that the aforementioned transition will be ended by a backoff state in space $S_1$. In general, it may be concluded that if the current state is AFT state $F_j$ and if the next event leads to a successful frame delivery, while exactly $n$ new frame(s) have entered the transmission queue during the time period of the aforementioned successful transmission, the transition will be ended by a backoff state (with $i = 0$) located in space $S_{j+i-1}$. However, if the next event results in a frame collision, there will be a state transition to a backoff state (with $i = 0$) situated in space $S_{j+i}$.

In Fig. 5, there are a set of transitions to the first backoff stage from states $B_{i,0,h}$ with $0 \leq i \leq m$ and $1 \leq (h \neq j) \leq j + 1$, upon successful frame transaction with the probability of $1 - p$, as illustrated by the left-most single vertical arrow. If a state transition initiated by a successful frame transmission coming out of state $B_{i,0,h}$ ($h \neq j$) is considered while $j - h + 1$ new frame(s) are added to the local transmission queue, the state transition will be directed to the first stage of the backoff process located in space $S_j$. In this case note that since $j - h + 1$ is a non-negative integer, then we will have $1 \leq h \leq j + 1$.

The state transitions from other spaces due to collisions are shown, as well, using horizontal lines at the bottom of Fig. 5. Since by a collision, the number of queued frames is not affected (except for the case in which the frame is dropped due to reaching the maximum retry limit, $m$), there will be a state transition from space(s) $S_h$ to the space $S_j$, for which $1 \leq h \leq j - 1$. Modeling a frame drop due to reaching the maximum retry limit, $m$, is also straightforward. For example, a state transition due to frame dropping, coming out of the $m$th backoff stage of space $S_j$, will be ended by a backoff state with $i = 0$ situated in space $S_j$, if, and only if, no more than one new frame has come into the transmission queue during a busy period caused by a collision.

Finally, Fig. 6 illustrates the space-time chain $S_L$ and all related state transitions while taking queue overflow into account. The steady state balance equations corresponding to the proposed analytical model are presented in Appendix 1.

62
Fig. 6. State transitions related to $S_L$ space-time chain.
2.3.5 Throughput analysis

Using the steady state balance equations and the normalization condition for the proposed PSTMC Markov model, backoff, post-backoff, vacancy, and AFT states’ probabilities, can be derived numerically. Eq. (1) provides the normalization condition, while conditional transmission probability, $\tau$, is given by Eq. (2) as follows.

\[
s_{j=0}^{L} \sum_{m=0}^{l} \sum_{k=0}^{W_{i}-1} B_{l,k,j} + \sum_{l=1}^{L} F_{l} + \sum_{k=1}^{b_{k} - 1} b_{k} + E = 1, \tag{1}
\]

\[
s_{j=0}^{L} \sum_{m=0}^{l} B_{l,d,j} + \sum_{l=1}^{L} F_{l} = \tau, \tag{2}
\]

where $E$ is the probability of being in the vacancy state. Using $\tau$, conditional collision probability $p = 1 - (1 - \tau)(n-1)$ is obtained. Consequently, the transmission probabilities $P_{tr}$ and $P_{t}$ are calculated. Finally, system throughput is given by

\[
S = \frac{P_{tr}P_{t}E[P]}{(1 - P_{tr})\sigma + P_{tr}[P_{t}T_{s} + (1 - P_{t})T_{c}]}, \tag{3}
\]

where $E[P]$ is the expected value of the MAC frame payload size and $\sigma$ is the size of a time slot [51].

2.3.6 Model validation

To validate the proposed framework, the obtained numerical results have been compared with those from simulation results. For all simulations, $\lambda$ is set to 10, 50, and 100 frames per second. The IEEE 802.11 PHY characteristics [78] are considered for all simulations. SIFS and DIFS are set to 16 µs and 56 µs, respectively. The average value of the MAC frame payload size $E[P] = 1024$ bytes, and each radio station’s queue size $L = 30$. The transmission power is set to 20 dBm. The wireless stations in our simulation environment have been configured in a non-fading line-of-sight single-cell hotspot in which the relative distance among any two stations is approximately the same. It should be also mentioned that for each system parameters setup, the achieved results of 30 successive simulation repetitions are averaged.
Fig. 7. Numerical analysis and simulation results for normalized throughput vs. number of contending stations ($\lambda = 10, 50, \text{and} 100, E[P] = 1024, \text{single-hop scenario}$).

Fig. 7 compares the numerical and simulation results for normalized throughput ($\lambda = 10, 50, \text{and} 100, E[P] = 1024$). The acceptable level of error in the numerical results, which are basically based on the analytical framework, shows satisfactory precision of the proposed mathematical framework for modeling such complex systems. In Fig. 7, regarding the numerical results for the analytical model, a divergence between three illustrated curves corresponding to different traffic loads in comparison to the simulation results is observed, while for the proposed model a lower average slope in throughput degradation can be seen when the number of contending stations is increased, which is more realistic based on extensive investigations presented in [52, 53]. As it can be also seen in Fig. 7, there is a difference between numerical results based on the analytical framework and those that have been achieved through simulations executed in OPNET™. In addition, it is worth noting that when the traffic load is increased, the aforementioned difference is decreased consequently.

Fig. 8 shows the normalized throughput for different frame payload sizes (512, 1024, and 2048 Bytes), while the frame arrival rate has been fixed to $\lambda = 10$. As expected, by increasing the payload size, the achieved throughput is increased.
Fig. 8. Numerical analysis for normalized throughput vs. number of contending stations ($\lambda = 10$, $E[P] = 512, 1024$, and 2048 Bytes, single-hop scenario).

Fig. 9 shows the numerical analysis of the normalized throughput using the PSTMC framework for a different number of contending stations (5, 10, 20, and 50), while the frame arrival rate has been fixed to $\lambda = 20$, versus different initial values for the contention window size. At first, it is indeed interesting to see that for each network configuration, there is always a maximum achievable throughput when the initial contention window size is varied from 8 to 1024. This maximum value is also changed and shifted to the larger values for contention window sizes when the number of contending stations is increased. This interesting behavior, in fact, can be observed in saturated systems, as reported also in [52].
Fig. 9. Normalized throughput for numerical analysis vs. initial contention window size ($\lambda = 20, E[P] = 2048$ Bytes, RTS/CTS Mechanism).

Fig. 10 shows the average number of idle time slots per successful packet transmission versus the number of contending stations for different initial contention window sizes (16, 64, and 256), while the frame arrival rate has been fixed to $\lambda = 20$. Apparently, when the number of stations in the network is increased, due to a larger number of occupied slot times for packet transmissions by neighboring stations as well as possible frame collisions, the average number of idle slots is decreased considerably. The number of iterations considered for this set of computer simulations is 100 consecutive simulation repetitions. Furthermore, it can be seen that by increasing the initial contention window size, the average number of idle slots per each number of stations is consequently increased. On the other hand, by increasing the number of stations, all three curves converge to almost the same values. For contention window size values of 16 and 64, when the number of stations reaches 30, no further change is observed in the obtained results. In other words, the average number of idle slots per successful packet transmission becomes independent of the number of associated stations when it goes beyond 30 radio stations. All results shown in Fig. 10 follow the same approach as [52] but of course with different values. The way the network behaves is similar to a saturated system, but on the other hand, as the considered network here is a non-saturated system, the average number of idle slots is more than the case of a saturated ad hoc network.
Fig. 10. Average number of idle slot times per successful packet transmissions vs. number of stations ($\lambda = 20$, $E[P] = 2048$ Bytes).

Fig. 11 shows the average number of transmissions per packet versus the initial contention window size for different numbers of stations (5, 10, 20, and 50), while the frame arrival rate has been fixed to $\lambda = 20$. Obviously, when the initial contention window size is increased, the probability of encountering frame collisions is decreased as the number of possible frame collisions among contending stations is decreased. As a consequence, the average number of (re)transmissions per packet that are required to successfully transmit the packet to the destination of interest is decreased. Since the considered scenario in this thesis is a non-saturated system, the achieved results in Fig. 11 compared to those of [52] are slightly lower. In other words, the average number of (re)transmissions per packet is lower in finite-load non-saturated ad hoc networks as the contending stations are less enthusiastic to contend for the shared wireless medium in order to deliverer buffered MAC frames over the air interface.
2.3.7 Service time & queuing delay analysis

In this sub-section, using the PSTMC framework the IEEE 802.11 DCF frame service time, delay jitter, and queuing delay in a single-hop non-saturated wireless network is analyzed. The PSTMC framework provides the possibility of simultaneous modeling of backoff and post-backoff procedures, in addition to the transmission queue status of a non-saturated 802.11 station. The presented framework facilitates study of the IEEE 802.11 DCF frame service time, i.e., access delay and retransmission delay, plus queuing delay at the same time, when the precise modeling of the binary exponential backoff scheme in the MAC layer is the main issue of concern. The model is also validated by extensive simulations, showing its satisfactory level of accuracy.

Overview and preliminaries

A single-hop ad hoc network with \( n \) stations is assumed while the transmission queue of each station may become vacant at any time during the network lifetime. The traffic arrival process at the MAC layer is assumed to be Poisson with a mean rate of \( \lambda \) frames per second. Then, the probability of \( z \) frames arrival during time interval \( t \) is given by \( A_z(t) = e^{-\lambda t}(\lambda t)^z/z! \). Even though the Poisson assumption
may not be perfectly realistic, it provides insightful results and allows the model to be tractable. Previously proposed queuing analysis in the literature assume Poisson arrival pattern specifically due to its independence feature of non-overlapping interarrival times, leading to the Markovian property to be valid. On the other hand, it is imperative to determine how argument \( t \) is set, in order to model the state transitions between \( L \) parallel spaces. Poisson traffic pattern makes time argument determination straightforward. Before proceeding to the service time analysis, it should be noted that for the Head-Of-Line (HOL) frame, which is experiencing a backoff (or post-backoff) cycle, there is no need to take into account the number of accumulated frames in the transmission queue; thus, for its service time analysis index \( j \) (corresponding to the queue status) can be ignored and, subsequently, \( B_{t,k,j} \) is replaced by \( B_{t,k} \).

**Analysis**

Frame delay is defined as the time interval between two successive successful frame transmissions. The total frame delay \( D \) is divided into two sub-delays: access delay and retransmission delay. Access delay is the time needed for the backoff timer to become zero, while Retransmission delay is the time consumed during the retransmissions due to frame collision. Total frame delay \( D \) is a random variable equal to the sum of \( i.i.d. \) random variables \( X_i \) representing the time period experienced during the (re)transmissions due to collision is modeled as a geometric random variable \( C \) with the mean and variance given below.

\[
P(C = k) = (1 - p)p^k, E[C] = p/(1 - p), Var(C) = p/(1 - p)^2. \tag{4}
\]

The random variable representing frame delay \( D \) is denoted by

\[
D = (X + T_\Delta) + \sum_{i=1}^{C} (X_i + \Delta_c), \tag{5}
\]

where \( X_i \) and \( X \) are \( i.i.d. \) random variables and \( \Delta_c = T\text{timeout} + T_{c\text{RTS}} \). The first term of (5) corresponds to the last frame retransmission leading to a successful frame reception. Taking the expectation value of (5), \( E[D] \) is given by

\[
E[D] = E[X] + T_\Delta + p(1 - p)^{-1}(E[X] + \Delta_c). \tag{6}
\]

The average backoff delay \( E[X] \) depends both on the value of the backoff timer and the time period during which the backoff timer is frozen. Upon being in
state $B_{i,k}$, $k$ slots are required before the timer reaches zero. The aforementioned time period, expressed in slot times and denoted by $\psi$, is given by

$$E[\psi] = \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} kB_{i,k}. \quad (7)$$

If $m$ and $m'$ represent the maximum allowed (re)transmissions per frame and the number of distinct contention window sizes, respectively ($m > m'$), then (7) can be expressed as (8), where $W$ is the minimum size of the contention window, $W = W_{min} = W_0$ and $B_{0,0}$ is given by (9).

$$E[\psi] = \frac{B_{0,0}}{6} \left( \sum_{i=0}^{m} W^2 (4p)^i - \frac{1-p^{m+1}}{1-p} + \sum_{i=m+1}^{m'} W^2 4^{m'} p^i \right) = \frac{W^2 (1-p)(1 - (4p)^{m+1} + 4^{m'}(1 - 4p)p^{m+1}(1 - p^{m-m'})) - (1 - 4p)(1 - p^{m+1})}{6(1-4p)(1-p)} \quad (8)$$

$$B_{0,0} = \frac{2(1-2p)(1-p)}{W(1-p)(1 - (2p)^{m+1}) + (1 - 2p)(1 - p^{m+1}) + 2^m W(1 - 2p)p^{m+1}(1 - p^{m-m'})} \quad (9)$$

Consequently, $E[X]$ can be derived as ($\sigma$ is the pre-defined time slot duration in IEEE 802.11)

$$E[X] = \left( (1 - P_{tr}) \sigma + P_{tr} (P_s T_s + (1 - P_s) T_c) \right) E[\psi]. \quad (10)$$

In (10), $E[X]$ is expressed using $P_{tr}$ and $P_s$ since the modeled station does not contend for the shared medium due to the fact that its backoff counter is still non-zero. To derive frame jitter expression, let us define a dummy random variable $\bar{D}$ as $\bar{D} = D - T_s$. Since $\sigma_2^D = \sigma_2^{\bar{D}}$, instead of the second moment derivation of random variable $D$, the second moment of $\bar{D}$ can be derived, which is given by

$$E[\bar{D}^2] = ((2p - 1)E[X] + 2p \Delta_{e} E[X] + \Delta_{e}^2 p + (1 - p)E[X^2])/(1-p)^2. \quad (11)$$

Since $E[X^2] = \left( (1 - P_{tr}) \sigma + P_{tr} (P_s T_s + (1 - P_s) T_c) \right)^2 E[\psi^2]$, it is only required to derive $E[\psi^2]$, which is given by (12).
\[
E[\psi^2] = \sum_{i=0}^{m} \sum_{k=0}^{W_{i-1}} k^2 B_{i,k} = \\
\frac{12(1-8p)(1-2p)(1-p)}{(1-8p)(1-(2p)^{m'+1})} W(1-p) \left( W^2(1-2p)(1-(8p)^{m'+1}) - (1-8p)(1-(2p)^{m'+1}) \right) + 2^{m'} Wp^{m'+1}(1-2p)(1-8p)(1-p^{m-m'}) (4^{m'} W^2 - 1).
\] (12)

To incorporate the queuing delay into the expression of random variable \( D \), the PSTMC model of the considered non-saturated station is needed to precisely model its transmission queue, in addition to the backoff and post-backoff procedures. Furthermore, the number of accumulated frames ahead of a considered frame in the transmission queue should be taken into account. The number of pre-existing frames can vary between zero and \( L - 1 \). Apparently, when the number of existing frames is \( L \), the incoming frame is simply dropped due to queue overflow. On the other hand, in the case of an empty queue, there is just one possibility: If the system is in the middle of a post-backoff cycle, then the considered frame should wait until the ongoing post-backoff process ends. Conversely, if the post-backoff cycle has already ended, then the frame is delivered without any queuing delay (i.e., asynchronous frame transmission). In the case when there is more than one frame in the transmission queue, the HOL frame is under service; thus, the total delay should have also a term corresponding to the remaining backoff time of the HOL frame. If, upon arrival of the frame, the number of existing frames inside the transmission queue is \( z \), then it is clear that the HOL frame is under service while the rest (i.e., \( z - 1 \)) are waiting to be served subsequently. For each of the \( z - 1 \) frames there will be a total access delay equal to \( (X + T_n) + \sum_{i=1}^{z} \Delta_i \), as derived earlier. This term is denoted by \( D_{\text{wait}} \) as it stands for the total delay for the whole frames waiting for service. Basically, \( D_{\text{wait}} \) is a function of \( z - 1 \), and in fact only when \( z \geq 1 \), it is included in the expression of total delay. The expectation value of \( D_{\text{wait}} \) is given by

\[
E[D_{\text{wait}}] = \sum_{j=2}^{L-1} \sum_{i=0}^{m} \sum_{k=0}^{W_{i-1}} E \left[ \sum_{n=1}^{j-1} D_n \right] B_{i,k,j}.
\] (13)

Note that the term \( E[D_n] \) has been already calculated through (6) to (10). For the HOL frame and depending on its current backoff state, the remaining service time
should be precisely determined. If, exactly before arrival of the considered frame, the HOL frame is assumed to be in backoff state $B_{i,k,z}$ (i.e., space $z$), then the remaining service time can be given by the following expression

$$X_{us} = \left( (1 - P_{tr})\sigma + P_{tr}P_{ts}T_s + (1 - P_{tr})T_c \right) k + (1 - p)T_s + p(\Delta_c + D_{remain}).$$  \hspace{1cm} (14)

Here $X_{us}$ stands for the service time of the HOL frame that is experiencing a backoff cycle. The first term stands for the HOL frame’s ongoing backoff stage. If the ongoing backoff cycle leads to a successful transmission, the service time will last for $T_s$ (the second term). Alternatively, if due to a collision the frame transmission becomes unsuccessful, then it is mandatory to take into account the next backoff stages the HOL frame will experience. This case is represented by the third term in (14). The average number of backoff steps a frame might experience during backoff stage $i$ is

$$E_n[X_i] = W_i^{-1} \sum_{k=0}^{W_i-1} k = (W_i - 1)/2.$$  \hspace{1cm} (15)

In addition, similar to (10), the average backoff time for each backoff stage is given by

$$E[X_i] = ((1 - P_{tr})\sigma + P_{tr}P_{ts}T_s + (1 - P_{tr})T_c)E_n[X_i].$$  \hspace{1cm} (16)

Finally, the expectation of $X_{us}$ is given by (17).

$$E[X_{us}] = \sum_{j=1}^{l-1} \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} B_{i,k,j} \left( (1 - P_{tr})\sigma + P_{tr}P_{ts}T_s + (1 - P_{tr})T_c \right) k + (1 - p)T_s +$$

$$p\left(\Delta_c + (1 - p)(E[X_{i+1}] + T_s) + \right.\sum_{n=1}^{m-i-1} (1 - p)p^n \left( E[X_{n+i+1}] + T_s + \sum_{x=0}^{n-1} E[X_{x+i+1} + \Delta_c] \right) +$$

$$\left. p^{m-i} \sum_{n=0}^{m-i-1} E[X_{n+i+1} + \Delta_c] \right).$$  \hspace{1cm} (17)
It is imperative to consider the case when upon arrival of the considered frame there is no frame inside the transmission queue. The following expression represents the expectation of such a case.

$$E[X_c] = \sum_{k=0}^{W_0-1} \left( (1 - P_{tr}) \sigma + P_{tr} (P'_s T_s + (1 - P'_s) T_c) \right) k b_c, \quad (18)$$

which is related to the current post-backoff step. Considering (13), (17), and (18), the expectation of frame delay can be easily obtained, which incorporates both service time and queuing delay.

**Model validation**

To validate the presented model, the same framework and assumptions are taken into consideration as in Section 2.3.6.

Fig. 12 shows the analytical vs. simulation results for both average total delay ($\lambda = 10, 50, \text{ and } 100, E[P] = 1024$) and frame jitter ($\lambda = 10 \text{ and } E[P] = 1024$). The way that the mean delay value varies with respect to the number of stations is due to the fact that at a certain number of nodes the total offered load to the system reaches the maximum achievable capacity and from that point, the average delay starts increasing exponentially. As it can be seen in Fig. 12, for the case of 50 contending stations when the offered load $\lambda$ per station is 100 frames per second, the difference between numerical and simulation results is almost 100 msec. Such cases have been also reported in [14, 52], especially for the DCF average service time. In fact, when the total offered traffic load to the wireless system reaches the maximum achievable capacity of the entire system, the whole system becomes unstable. For the case of frame jitter, Fig. 12 shows an indeed acceptable level of precision when the number of contending stations is less than 35, while it demonstrates an increment in the difference between the obtained results when the number of stations becomes larger than 35. In fact, the system instability is the main reason for the difference between the numerical and simulation results when the number of stations is larger than a certain value. In other words, the observed instability in the wireless network is reflected in the achieved frame jitter results in the form of estimation error increment.
2.4 Non-saturated performance analysis: multi-hop perspective

The contention-based distributed coordination function has been used in IEEE 802.11 multi-hop ad hoc networks, while so far a comprehensive mathematical analysis of such systems has not been reported in the literature due to complicating factors such as hidden terminal and unreachability problems. In this section, an analytical model based on the Parallel Space-Time Markov chain for such networks is proposed. The proposed scheme is able to model the hidden terminal problem and the unreachability phenomena. Extensive evaluation is conducted to demonstrate the accuracy of the proposed model.

In this section, the framework in Section 2.3 is extended to the multi-hop ad hoc network case. The proposed approach formulates an exact model of the hidden terminal and unreachability phenomena, which is a method to model IEEE 802.11 DCF MAC in multi-hop ad hoc networks. The main idea is to make use of graph partitioning in order to classify wireless stations into distinct membership groups and subsequently in each group use the PSTMC framework to model the medium access behavior of associated stations, taking into account the fact that
their channel acquisition is affected by all neighboring nodes with membership to the same and different groups. It is shown that by the means of the aforementioned approach, one can be directed to a completely localized and independent performance analysis of multi-hop networks in comparison to the earlier proposed methods in the literature. Finally, using extensive simulations, the accuracy of the proposed method is demonstrated.

A single-hop wireless network is defined as a finite set of \( n \) stations, where each station is concurrently situated in the radio range of all other \( n - 1 \) stations. Therefore, a single-hop wireless network may be considered an undirected fully-connected complete graph with \( n \geq 1 \) vertices. Furthermore, the aforementioned graph can be modeled as a fully-connected cluster \( C_n \). In other words, a single-hop wireless network consists of a sole maximal fully-connected cluster, which is called a clique. Similarly, a multi-hop ad hoc network can be easily viewed as a union of all possible maximal fully-connected clusters (or cliques) in which each cluster should have at least an intersection with one of the other clusters in order to provide system connectivity. Note that, throughout the thesis, when a station, say ‘A’, is assumed to be in the radio range of another station, say ‘B’, then the reverse scenario is also assumed to be valid, meaning that station ‘B’ is in the radio range of station ‘A’, as well.

The main idea of the proposed framework is to exploit graph partitioning in order to classify all wireless stations based on their membership to different clusters. Thus, as the first step all possible clusters for the whole graph of the system should be identified and, subsequently, as each station might have multiple memberships to different groups of clusters, all stations in distinct membership-groups are classified according to their association with different sets of clusters. In other words, two stations that have membership to the same set of clusters should be placed in an identical membership-group. From now on and throughout the thesis, a membership-group is referred to as a sub-cluster. Principally, all members of a sub-cluster have exactly the same view of the entire system due to the fact that all of them are situated in the radio range of an analogous set of stations that are outside of that sub-cluster.

It is worth noting that the concept of cluster is equivalent to the concept of collision domain. Basically, the collision domain of a tagged station is defined as the set of all potential transmitting nodes whose transmissions may interfere with the considered station. In other words, a transmission conflict will not be considered a collision unless the tagged station remains in the interference range of the emitting terminal for the whole duration of its data communication. Given
the assumption of bidirectional channels, it may be concluded that all stations, for which the collision domain is the same belong to the same sub-cluster, and their collision domain comprises all stations that are able to interfere on them.

In order to clarify the scenario of interest, consider Fig. 13, which provides a visual illustration of the core concepts employed in the proposed framework. Here, three maximal clusters named $C_1$, $C_2$, and $C_3$ are taken into account. In cluster $C_1$, stations ‘A’ and ‘B’ should be placed in the same membership-group (i.e., sub-cluster), however, stations ‘C’ and ‘D’ cannot be placed in this sub-cluster since in addition to $C_1$ they have membership to clusters $C_2$ and $C_3$, respectively. Similarly, stations ‘E’ and ‘F’ are placed in the same sub-cluster because of membership to only cluster $C_2$. Station ‘C’ has membership to two different clusters: $C_1$ and $C_2$. None of the stations in the network, except ‘C’, has simultaneous membership to both clusters $C_1$ and $C_2$; therefore, the sub-cluster comprising station ‘C’ will have only one associated member which is in fact station ‘C’.

**Fig. 13. An example of a three-hop wireless network.**

In Fig. 14, all possible sub-clusters for an ad hoc network are identified using rectangles with solid-line sides. In addition, the cluster comprising two central sub-clusters is also shown using a rectangle with dotted-line sides.
Fig. 14. A wireless network consisting of three intersecting fully-connected clusters and four non-overlapping sub-clusters (membership-group).

Since all members of a sub-cluster are situated in the radio range of the same set of stations, which is situated outside of the sub-cluster, as far as the backoff process and the impact of other stations on it are concerned, they exhibit similar stochastic behavior. Hence, to deal with a localized approach it is only required to consider the stochastic model of a single station in each sub-cluster while taking into account that the tagged station’s medium access is affected by all its neighbors situated either inside or outside of its sub-cluster. In order to clarify this imperative issue, let us reconsider Fig. 14. It can be verified that the view of each member of a sub-cluster to the entire system is similar to the other members in the same sub-cluster. As an example, let us consider one of the existing sub-clusters in Fig. 14. As can be seen, the considered sub-cluster in Fig. 14 has two members, each of which is in the radio range of three stations that are situated outside the tagged sub-cluster. Since this condition is valid for both members of this sub-cluster, it can be easily inferred that the stochastic behavior of these two stations is affected by exactly the same set of stations: three stations outside the sub-cluster and one station inside the sub-cluster. Therefore, it is only needed to choose a station in each sub-cluster and, subsequently, by applying an extended version of the PSTMC framework it is possible to model its stochastic medium access activities in a totally localized manner. Basically, the PSTMC model of a station should consider different conditional transmission probabilities, \( \tau \), for distinct groups of neighboring stations with association to different adjacent sub-clusters. In addition, due to similarity in stochastic behavior, all members of a tagged station’s sub-cluster should have the same transmission probabilities as the tagged station [80]. Finally, it is assumed that all stations situated in the same sub-cluster are equivalent under the offered traffic characteristics. It is explicitly
assumed that the offered traffic per station follows Poisson distribution with the arrival rate of $\lambda$ frames per second at the MAC layer.

To proceed, without loss of generality, a case shown in Fig. 15 is considered. Each station may have a membership to only one of sub-clusters $C_1 - C_2, C_2 - C_1$ or $C_1 \cap C_2$. As explained earlier, since all members of a sub-cluster see the rest of the network homogeneously, their corresponding stochastic parameters including conditional transmission probability $\tau$ and conditional collision probability $p$ should also be exactly the same [80].

![Fig. 15. A double-hop wireless network depicted as two intersecting fully-connected clusters.](image)

In order to model multi-hop ad hoc networks, it is indispensable to note that channel access is not only performed using an RTS frame, but also through a CTS frame. To explain channel access using a CTS control frame, let us reconsider Fig. 15. In cluster $C_1$, channel access can be accomplished either using an RTS frame transmitted by a station in this cluster or by a CTS frame transmitted by a station with the membership to sub-cluster $C_1 \cap C_2$. The latter case is a result of another channel access performed by a station with membership to $C_2 - C_1$ that has transmitted an RTS frame destined to one of the stations in sub-cluster $C_1 \cap C_2$, stimulating it to respond with a CTS frame. Note that frame transmissions originated by members of sub-cluster $C_1 \cap C_2$ will be heard by all members of clusters $C_1$ and $C_2$, including members of sub-clusters $C_1 - C_2$ and $C_2 - C_1$. Due to the fact that the responsibility of routing data packets between two sub-clusters $C_1 - C_2$ and $C_2 - C_1$ is handled by the members of sub-cluster $C_1 \cap C_2$, such nodes will be referred to as Virtual Gateways throughout the thesis. Finally, there is another important scenario that should be modeled precisely: Since members of sub-clusters $C_1 - C_2$ and $C_2 - C_1$ are hidden to each other, an RTS
frame transmission originated from one of these regions can easily collide with an RTS frame being transmitted by a member of the other sub-cluster destined to a station with membership to $C_1 \cap C_2$. This case has been neglected in for instance [79].

Modeling multi-hop ad hoc networks involves obtaining an analytical model for the so-called virtual gateways, which leads to exploring an appropriate formulation for both hidden terminal and unreachability problems. Using a stochastic model of a virtual gateway (i.e., its PSTMC model), it is possible to derive two imperative conditional probabilities related to any virtual gateway situated in the same sub-cluster: the conditional transmission probability due to an RTS frame in order to initiate a communication, and the conditional transmission probability due to a CTS frame in response to a received RTS frame. In addition, through the stochastic model of a virtual gateway, it is possible to determine when and with which probability it might become unreachable for a certain group of wireless stations.

The way the aforementioned achievements are gained is quite straightforward. As all virtual gateways constructing a sub-cluster are reachable by all members of neighboring sub-clusters, their backoff processes are affected by any event taking place in the adjacent sub-clusters. In Fig. 15 and for a given virtual gateway, as long as the wireless channel of either sub-clusters $C_1 - C_2$ or $C_2 - C_1$ is busy, the virtual gateway’s backoff counter is interdicted from being decremented.

Since the PSTMC framework is used to model the performance of a tagged virtual gateway, its backoff and post-backoff states are represented by $B_{i,k,j}$ and $b_k$, respectively. Here, index $i$ represents the backoff stage ($0 \leq i \leq m$; where $m$ is the maximum allowed retransmissions per frame), index $k$ represents the backoff step ($0 \leq k \leq W_i - 1$; where $W_i$ is the contention window size in the $i$th backoff stage), and index $j$ represents the number of existing frames in the transmission queue ($0 \leq j \leq L$; where $L$ is the transmission queue capacity).

Due to the independence of occurring events in sub-clusters $C_1 - C_2$ and $C_2 - C_1$, the channel state (either ‘idle’ or ‘busy’) for clusters $C_1$ and $C_2$ observed by a virtual gateway is simultaneously modeled by a set of intermediate states, placed between any two consecutive backoff or post-backoff states. The intermediate

---

1 Even though sub-clusters $C_1 - C_2$ and $C_2 - C_1$ are connected to each other through $C_1 \cap C_2$, as long as the shared medium in $C_1 \cap C_2$ is not under utilization, the events taking place in $C_1 - C_2$ and $C_2 - C_1$ are independent, as the members of these two sub-clusters are hidden to each other.
states between successive backoff states \((B_{i,k,j} \text{ and } B_{i,k-1,j})\) and post-backoff states \((b_k \text{ and } b_{k-1})\) are denoted by \(\theta_{i,k,j}(x,y)\) and \(\eta_k(x,y)\), respectively. The non-negative integer arguments \((x \geq 0 \text{ and } y \geq 0)\) represent the remaining channel busy period (in number of time slots) for clusters \(C_1\) and \(C_2\), correspondingly. The virtual gateway’s backoff counter is frozen when at least one of the aforementioned arguments is non-zero. In other words, while having already departed \(B_{i,k,j}\) (or \(b_k\)) in order to proceed to the next state, the virtual gateway’s backoff counter may not enter the state \(B_{i,k-1,j}\) (or \(b_{k-1}\)) given that at least one of the arguments in \(\theta_{i,k,j}(x,y)\) (or \(\eta_k(x,y)\)) is non-zero. The state transition to the next backoff state is accomplished when the arguments in \(\theta_{i,k,j}(x,y)\) are both zero, i.e., \(\theta_{i,k,j}(0,0) = B_{i,k-1,j}\). Similarly, the state transition to the next post-backoff state will be completed if, and only if, the arguments in \(\eta_k(x,y)\) are both zero, i.e., \(\eta_k(0,0) = b_{k-1}\).

While being in \(B_{i,k,j}\) (or \(b_k\)), if no channel access activity is observed for the duration of a time slot \(\sigma\), the backoff process will progress to \(B_{i,k-1,j}\) (or \(b_{k-1}\)). However, following a channel acquisition in at least one of the clusters \(C_1\) or \(C_2\), either resulting in a successful transmission or a frame collision, the state transition is terminated by an appropriate intermediate state, i.e., \(\theta_{i,k,j}(x,y)\) or \(\eta_k(x,y)\). In fact, based on the type of event, arguments \(x\) and \(y\) are determined accordingly.

There are circumstances when one of the arguments in an intermediate state, say \(x\), is zero while the other one, say \(y\), is still non-zero. In such cases, it is possible that due to an event taking place in sub-cluster \(C_1 \rightarrow C_2\) the corresponding argument \((x)\) is getting updated to a non-zero value. Hence, the PSTMC model of a virtual gateway should keep track of all ongoing events in all clusters to which the virtual gateway belongs in order to update the corresponding arguments in intermediate states. As an example, suppose that the current intermediate state is \(\theta_{i,k,j}(x,0)\), meaning that cluster \(C_1\)’s wireless medium will be continuously ‘busy’ for \(x\) successive time slots while the cluster \(C_2\) wireless channel is currently ‘idle’. This might be due to an ongoing data transmission between two stations that are both members of sub-cluster \(C_1 \rightarrow C_2\). Now consider an event taking place in sub-cluster \(C_2 \rightarrow C_1\) leading to a total busy period of \(y \neq 0\) in \(C_2 \rightarrow C_1\), and of course in the whole cluster \(C_2\). Consequently, the next intermediate state will be \(\theta_{i,k,j}(x,y)\) | \((x,y) \neq (0,0)\). Assuming that \(x \geq y\), then it can be easily seen that with a probability equal to one and exactly after \(y\) time slots (or equivalently after \(y \times \sigma\)) the final intermediate state will be \(\theta_{i,k,j}(x-y,0)\). Similarly, the intermediate state \(\theta_{i,k,j}(x,y)\) | \((x,y) = (m,m)\) is continued
by $\theta_{i,k,j}(0,0) = B_{i,k-1,j}$. For instance, the state $\theta_{i,k,j}(K_c - 1, K_c - 1)$ refers to the case when the channel status for both clusters is busy due to a collision that has happened in sub-cluster $C_1 \cap C_2$. Subsequent to the state $\theta_{i,k,j}(K_c - 1, K_c - 1)$, upon conclusion of a single time slot $\sigma$ and with the probability of one, the entire system progresses to the state $\theta_{i,k,j}(K_c - 2, K_c - 2)$. This transition is continued by progressing to the state $\theta_{i,k,j}(K_c - 3, K_c - 3)$ upon conclusion of another single time slot $\sigma$, and again with the probability of one. Finally, the entire system ends up with the state $\theta_{i,k,j}(1,1)$, which itself is followed by the state $\theta_{i,k,j}(0,0) = B_{i,k-1,j}$ with the probability of one. Note that the channel status for both clusters emerges synchronously, and the above-explained process $\theta_{i,k,j}(K_c - 1, K_c - 1) \to B_{i,k-1,j}$ takes exactly $(K_c - 1) \times \sigma$, corresponding to the time period of a collision plus a DIFS, and excluding the first time slot during which the medium was detected ‘busy’ due to the frame collision. Note that $T_c = K_c \times \sigma$, where $T_c$ represents a channel busy period due to frame collision plus a DIFS. Therefore, starting from the state $B_{i,k,j}$ and subsequent to a detected collision in sub-cluster $C_1 \cap C_2$ at the first time slot, the system commences to progress from the state $\theta_{i,k,j}(K_c - 1, K_c - 1)$ to the next backoff state $B_{i,k-1,j}$ after a time period of $(K_c - 1) \times \sigma$.

Fig. 16 illustrates three possible state transitions originating from a tagged virtual gateway’s backoff state $B_{1,k,j}$. Transitions entering dummy states “1” and “2” encircled by thick back lines are continued in Figs. 14 and 15, respectively. In Figs. 13–15, for each state transition the probability of its initiating event along with the corresponding time duration for the initiating event is illustrated. The time duration for each event is required to be determined due to the fact that switching between distinct parallel space-time chains (or simply spaces) necessitates determination of the number of frame arrivals into the transmission queue during the period of each event leading to a state transition.
In Figs. 13–15, below each state transition the number of involving stations in each sub-cluster causing the transition is denoted by an *array* with three *elements* corresponding to $C_1 - C_2$, $C_1 \cap C_2$, and $C_2 - C_1$, respectively. For instance, the array $(\geq 2, 1, 0)$ means that more than one station in $C_1 - C_2$, only one station in $C_1 \cap C_2$, and no station in $C_2 - C_1$ have been involved with a particular state transition of interest. In addition, in some cases one of the elements of an array might be also accompanied with a symbol $d$ or $d'$. The accompanying symbol is only used when the state transition corresponds to an RTS frame delivery; the symbol represents the destination address of the RTS frame which is being transmitted or is to be transmitted later in the corresponding sub-cluster of the labeled element. Basically, these symbols are only used for defining the state transitions’ probabilities. For instance, a certain state transition caused by an RTS frame transmission will have $P\{d \in C_1 \cap C_2\}$ as part of its transition probability in order to represent the probability that the intended receiving station of the RTS frame is a member of sub-cluster $C_1 \cap C_2$. 

---

**Fig. 16. Intermediate state transitions between two consecutive backoff steps.**
Fig. 17. Intermediate state transitions between two consecutive backoff steps.
Fig. 18. Intermediate state transitions between two consecutive backoff steps.

To elucidate the way state flows are traversed in Figs. 13 – 15, consider $B_{i,k,j}$ as the backoff state by the end of a certain time slot. This means that the channel state for both clusters $C_1$ and $C_2$ will be ‘idle’ at the end of the considered time slot. Following the conclusion of the time slot, assume that an RTS frame is transmitted by a member of $C_2 - C_1$ while no channel access is performed in the
two other sub-clusters \( C_1 - C_2 \) and \( C_1 \cap C_2 \). To illustrate the occurrence of this incident visually, we commence by moving from \( B_{i,k,j} \) to dummy state “1”, as shown by a dashed arrow in Fig. 16 accompanied by two labels \((1 - \tau_3)^{n_3-1} \) and \((#,0,#)\). The first label means that none of the virtual gateways in sub-cluster \( C_1 \cap C_2 \) performs channel access: Note that the virtual gateway which is being modeled should be excluded from the others, meaning that instead of \( n_3 \) we have to use \( n_3 - 1 \). In the second label, the # sign indicates that the corresponding element will be determined subsequently in Fig. 17. The state transition is continued from the dummy state “1” while accompanied by an array \((0/d',0,1/d)\), as shown in Fig. 17. Up to the point marked by a flag in Fig. 17, the incident probability is given by \((1 - \tau_1)^{n_1}(1 - \tau_3)^{n_3-1}n_2\tau_2(1 - \tau_2)^{n_2-1}\), which is in fact the multiplication of probabilities along the transition path up to the flagged point. Consequently, assume that the RTS frame originated from the sub-cluster \( C_2 - C_1 \) has been intended to one of the members of sub-cluster \( C_1 \cap C_2 \). To take this possibility into account, it is necessary to multiply \( P(d \in C_1 \cap C_2) \) with the above derived transition probability and traverse along the path with the label \( P(d \in C_1 \cap C_2) \) on top. Now let us proceed to the next time slot and suppose that there is another medium access attempt by one of the members of sub-cluster \( C_2 - C_2 \). As it can be seen from Fig. 17, it is necessary to follow the transition with a label \( n_1\tau_1(1 - \tau_1)^{n_1-1} \) and, consequently, the term \( n_1\tau_1(1 - \tau_1)^{n_1-1} \) should be multiplied by the above aggregate probability. Evidently, the proposed transmission attempt in the second time slot is accomplished using an RTS frame; therefore, it is also mandatory to take into consideration the membership of the intended receiving station for the transmitted RTS frame. Obviously, there are only two possibilities for the receiving station: either membership to sub-cluster \( C_1 - C_2 \) or membership to sub-cluster \( C_1 \cap C_2 \). If the first possibility is assumed, the dashed arrow in Fig. 17 should be followed and, as a consequence, the probability \( P(d' \in C_1 - C_2) \) should be taken into account.

In conclusion, during the first time slot there is a transmission attempt originated from \( C_2 - C_1 \) and intended for a member of \( C_1 \cap C_2 \) and, then, during the second time slot there is another transmission originated from \( C_1 - C_2 \) and intended for a member in the same sub-cluster. Apparently, the second transmission collides with the first one, meaning that the first channel access leads to an unsuccessful medium access (and busy period due to a collision in the whole cluster \( C_2 \)), while the second one does not encounter any problem and leads to a successful data communication taking place in cluster \( C_1 \). If a channel busy period due to a successful medium access is denoted by \( T_2 = K_2 \times \sigma \) and a
channel busy period due to a collision by $T_c = K_c \times \sigma$ (where $\sigma$ is the predefined time slot duration in IEEE 802.11), then it can be stated that at the end of the second time slot the remaining busy period of both clusters $C_1$ and $C_2$ will be $(K_{c-1}) \times \sigma$ and $(K_{c-2}) \times \sigma$, respectively. Note that the second channel access has happened exactly one time slot after the first one and, in fact, this is the main reason why there are $K_c-1$ successive busy time slots for cluster $C_1$ and $K_c-2$ succeeding busy time slots for cluster $C_2$. Based on the concept of intermediate state $\theta_{i,k,j}(x,y)$ given earlier, the channel state of the whole system at the end of the second time slot seen by the tagged virtual gateway can be simply represented by intermediate state $\theta_{i,k,j}(K_{c-1}, K_{c-2})$. Finally, there will be a transition from $\theta_{i,k,j}(K_{c-1}, K_{c-2})$ to $\theta_{i,k,j}(K_c - K_c + 1, 0)$ with the probability of one, which will take $K_c-2$ time slots to be completed.

2.4.1 Key features of virtual gateway PSTMC model

Fig. 17 shows also a scenario in which an RTS frame destined to a station with membership to $C_1 \cap C_2$, and transmitted by a member of sub-cluster $C_1 - C_2 \ (C_2 - C_1)$ collides with an ongoing RTS transmission taking place in sub-cluster $C_2 - C_1 \ (C_1 - C_2)$. In contrast, this scenario has always been ignored in the literature, e.g., in [79, 80]. Here, the length of the RTS frame in terms of the number of time slots is represented by $\tau$. Regarding this case, if there is no channel access in the other sub-cluster for the duration of $\tau \times \sigma$, then the transmitted RTS frame will be received correctly by the destination of interest, which is in fact a virtual gateway with membership to sub-cluster $C_1 \cap C_2$. As a result, the receiver will be able to gain control of both clusters’ channel by sending a CTS frame back to the source station. Since the CTS frame that is being delivered by the virtual gateway, is received successfully not only by the source station, but also by all members of clusters $C_1$ and $C_2$, it consequently updates all network allocation vectors (i.e., NAVs) for the whole network and, as a result, the remaining busy period for both clusters will be $((K_c - \tau) \times \sigma, (K_c - \tau) \times \sigma)$ represented by $\theta_{i,k,j}(K_{c-\tau}, K_{c-\tau})$ intermediate state; thus, the final state transition will be ended by the next backoff state $\theta_{i,k,j}(0,0) = B_{i,k-1,j}$, as shown on the right hand side of Fig. 17.

In multi-hop ad hoc networks, the vacancy state of a virtual gateway’s transmission queue (i.e., when the queue is empty) should be modeled by a set of states $\theta_{0}(x,y)$ representing channel states for both clusters $C_1$ and $C_2$, entirely denoted by $E - plan$, as illustrated in Fig. 19. In this structure, state transitions are
finally terminated at state $\theta_0(0, 0)$, which represents an idle channel state for both clusters $C_1$ and $C_2$ when the local transmission queue is still empty.

Fig. 19. $E -$ and $S_j -$ plan representing state transition from $E$ plan to $S_j$ Space-Time Markov chain.

Basically, transition from $E -$ plan to space-time chains $S_j$ upon arrival of new frames to the transmission queue should also be modeled precisely. Fig. 19 demonstrates the way by which the aforementioned transitions between $E -$ plan and the space-time chain $S_j$ are modeled in a multi-hop non-saturated system. Assume that the current state is $\theta_0(x, y)$ with $(x, y) \neq (0, 0)$, meaning that the transmission queue is empty and the channel state for both clusters $C_1$ and $C_2$ is busy. Suppose that during the period of a state transition in $E -$ plan and before
returning to $\theta_0(0,0), j$ new MAC frame(s) enter the transmission queue. Due to the fact that at least one of the aforementioned clusters observed by the virtual gateway is still busy, according to IEEE 802.11 specifications the virtual gateway is required to wait for both clusters $C_1$ and $C_2$ to become idle before initiating a new backoff procedure. In other words, the entrance of a set of MAC frames into the transmission queue while being in $E$– plan does not impose any allowance for the virtual gateway to initiate a new backoff cycle instantaneously. This means that it is necessary to wait for both clusters to become idle before switching to an appropriate space-time chain. This waiting period is represented by another plan prior to the space-time chain $S_J$ and referred to as $S_J$–plan. In this plan, there is a dummy state denoted by $\theta_J(0,0)$ with the same functionality as state $\theta_0(0,0)$, meaning that the channel state for both monitored clusters is idle. When a state transition in $S_J$–plan is terminated by the dummy state $\theta_J(0,0)$, a new backoff procedure is initiated immediately and the state pointer transits to the first backoff stage of space $S_J$.

2.4.2 Modeling the unreachability phenomenon

Unreachability of a station in sub-cluster $C_i \cap C_j$ is defined as the inability of the members of another sub-cluster $C_j - C_i$ to establish connection with it. This phenomenon takes place during the period when the medium in $C_i$ is busy due to an ongoing intra-cluster data transmission. Based on the definition of unreachability [19 – 21], the following expressions for the probability of unreachability are obtained:

\[
P_{u,C_1} = \sum_{z=1}^{K_s-1} \sum_{j=0}^{L} \theta_j(0,z) + \sum_{j=1}^{L} \sum_{i=0}^{m} \sum_{k=1}^{W_{i-1} K_s-1} \theta_{i,k,j}(0,z) + \sum_{k=1}^{W_0-1} \sum_{z=1}^{K_s-1} \eta_k(0,z) \quad (19)
\]

\[
P_{u,C_2} = \sum_{z=1}^{K_s-1} \sum_{j=0}^{L} \theta_j(z,0) + \sum_{j=1}^{L} \sum_{i=0}^{m} \sum_{k=1}^{W_{i-1} K_s-1} \theta_{i,k,j}(z,0) + \sum_{k=1}^{W_0-1} \sum_{z=1}^{K_s-1} \eta_k(z,0). \quad (20)
\]

Equation (19) represents the probability of unreachability of members in sub-cluster $C_1 \cap C_2$ as seen by members in cluster $C_1 - C_2$, and equation (20) represents the probability of unreachability of members in sub-cluster $C_1 \cap C_2$ as seen by the members in cluster $C_2 - C_1$. In equations (19) and (20), argument $z$ is used to indicate that the wireless channel of cluster $C_2$ and $C_1$ is busy, respectively.

In addition, it is necessary to find the probability that a given virtual gateway
transmits a CTS frame. If the probability of transmission of a CTS frame from sub-cluster $C_1 \cap C_2$ due to an RTS originated from $C_1 - C_2$ is denoted by $\xi_{C_1}$ and the probability of transmission of such a frame due to an RTS originated from $C_2 - C_1$ is denoted by $\xi_{C_2}$, then

$$\xi_{C_1} = \left(\theta_0(0,0) + \sum_{j=1}^{L} \sum_{i=0}^{m} \sum_{k=1}^{w_{j-1}} B_{i,k,j} + \sum_{k=1}^{w_0-1} b_k\right) \times n_1 \tau_1 (1 - \tau_1)^{n_1-1}(1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3} P(d \in C_1 \cap C_2),$$

(21)

$$\xi_{C_2} = \left(\theta_0(0,0) + \sum_{j=1}^{L} \sum_{i=0}^{m} \sum_{k=1}^{w_{j-1}} B_{i,k,j} + \sum_{k=1}^{w_0-1} b_k\right) \times n_2 \tau_2 (1 - \tau_2)^{n_2-1}(1 - \tau_1)^{n_1}(1 - \tau_3)^{n_3} P(d \in C_1 \cap C_2).$$

(22)

In the above equations, $r$ represents the length of an RTS frame in terms of time slot $\sigma$. It is indispensable to determine the state transition probabilities, $P'_{tr}$ and $P'_{s}$, for all stations with different memberships in different sub-clusters. For a tagged station with membership to $C_1 - C_2$, these parameters are given by (for both single- and multi-hop systems, $\tau$ is the same as given in Section 2.1)

$$P'_{tr_{C_1-C_2}} = 1 - (1 - \tau_1)^{n_1-1}\left(1 - (1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3} + \xi_{C_2}\right),$$

(23)

$$P'_{s_{C_1-C_2}} = \frac{1}{1 - (1 - \tau_1)^{n_1-1}\left(1 - (1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3} + \xi_{C_2}\right)} \times \left((n_1 - 1) \tau_1 (1 - \tau_1)^{n_1-2}P(d \in C_1 - C_2) + (1 - P_{u,C_1})P(d \in C_1 \cap C_2)\right) \times$$

$$\left(1 - (1 - \tau_3)^{n_3} + \xi_{C_2}\right) + n_3 \tau_3 (1 - \tau_3)^{n_3-1}(1 - \tau_1)^{n_1-1} + \xi_{C_2}(1 - \tau_1)^{n_1-1} + \xi_{C_2}(1 - \tau_1)^{n_1-1}.$$  

(24)

For a given station with membership to $C_2 - C_1$, $\tau_1$, $n_1$, $\xi_{C_1}$ and $P_{u,C_1}$ should be replaced by $\tau_2$, $n_2$, $\xi_{C_2}$, and $P_{u,C_2}$, respectively, while for a given station with membership in sub-cluster $C_1 \cap C_2$ we have.
\begin{align}
    P_{tr\mathcal{C}_1\cap \mathcal{C}_2} &= 1 - (1 - \tau_1)^{n_1}(1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3 - 1}, \quad (25) \\
    P'_{tr\mathcal{C}_1\cap \mathcal{C}_2} &= \frac{1}{1 - (1 - \tau_1)^{n_1}(1 - (1 - \tau_3)^{n_3} + \xi_{C_2})} \\
    &\times \left( n_1 \tau_1 (1 - \tau_1)^{n_1 - 1} (P(d \in \mathcal{C}_1 - \mathcal{C}_2) + (1 - P_{u,C_1}^d)P(d \in \mathcal{C}_1 \cap \mathcal{C}_2)) \times \right. \\
    &\left. \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_2} \right) \right. \\
    &\left. + n_3 \tau_3 (1 - \tau_3)^{n_3 - 1} (1 - \tau_1)^{n_1} + \xi_{C_2} (1 - \tau_1)^{n_1} \right). \quad (31)
\end{align}

The conditional collision probability for different sub-clusters is then given by

\begin{align}
    p_1 &= 1 - (1 - \tau_1)^{n_1 - 1} \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_2} \right), \quad (27) \\
    p_2 &= 1 - (1 - \tau_2)^{n_2 - 1} \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_1} \right), \quad (28) \\
    p_3 &= 1 - (1 - \tau_3)^{n_3} (1 - \tau_2)^{n_2} (1 - \tau_3)^{n_3 - 1}. \quad (29)
\end{align}

Similarly, for cluster \( \mathcal{C}_1 \) we have the following relations for transmission probabilities

\begin{align}
    P_{tr\mathcal{C}_1\cap \mathcal{C}_2} &= 1 - (1 - \tau_1)^{n_1} \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_2} \right), \quad (30) \\
    P_{2\mathcal{C}_1\cap \mathcal{C}_2} &= \frac{1}{1 - (1 - \tau_1)^{n_1}} \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_2} \right) \times \\
    &\left( n_1 \tau_1 (1 - \tau_1)^{n_1 - 1} (P(d \in \mathcal{C}_1 - \mathcal{C}_2) + (1 - P_{u,C_1}^d)P(d \in \mathcal{C}_1 \cap \mathcal{C}_2)) \times \right. \\
    &\left. \left( 1 - (1 - \tau_3)^{n_3} + \xi_{C_2} \right) \right. \\
    &\left. + n_3 \tau_3 (1 - \tau_3)^{n_3 - 1} (1 - \tau_1)^{n_1} + \xi_{C_2} (1 - \tau_1)^{n_1} \right). \quad (31)
\end{align}

For cluster \( \mathcal{C}_2 \), we should only replace \( \tau_1 \) and \( n_1 \) by \( \tau_2 \) and \( n_2 \), respectively. The aforementioned probabilities for sub-cluster \( \mathcal{C}_1 \cap \mathcal{C}_2 \) are given by

\begin{align}
    P_{tr\mathcal{C}_1\cap \mathcal{C}_2} &= 1 - (1 - \tau_1)^{n_1}(1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3}, \quad (32)
\end{align}
\[ P_{sc_{1},nc_{2}} = (1 - (1 - \tau_1)^{n_1}(1 - \tau_2)^{n_2}(1 - \tau_3)^{n_3})^{-1}(1 - \tau_1)^{n_1-1}(1 - \tau_2)^{n_2-1}(1 - \tau_3)^{n_3-1}(n_1\tau_1(1 - \tau_2)(1 - \tau_3) + n_2\tau_2(1 - \tau_1)(1 - \tau_3) + n_3\tau_3(1 - \tau_1)(1 - \tau_2)). \] (33)

### 2.4.3 Throughput analysis

Based on the previous analysis, system throughput for both clusters \( C_1 \) and \( C_2 \) can be derived using the following equations:

\[ T_{RTS}^s = T_{RTS} + tSIFSTime + \delta + T_{CTS} + tSIFSTime + \delta + T_{DATA} + tSIFSTime + \delta + T_{ACK} + \delta + tDIFSTime, \]
\[ T_{CTS}^s = T_{CTS} + tSIFSTime + \delta + T_{DATA} + tSIFSTime + \delta + T_{ACK} + \delta + tDIFSTime, \]
\[ T_{IFS}^s = T_{RTS} + tEIFSTime + \delta \approx T_{IFS}^c = T_{CTS} + tEIFSTime + \delta = T_c, \]

where \( \delta \) represents propagation delay. \( T_{DATA}, T_{RTS}, T_{CTS}, \) and \( T_{ACK} \) are given in Section 2.1. For cluster \( C_1 \), throughput is then given by:

\[ S_{c_1} = \frac{P_{tr_{c_1}}p_{c_1}E[p]}{(1 - P_{tr_{c_1}})\sigma + P_{tr_{c_1}}(p_{RTS}^s + P_{CTS}^s + (1 - P_{RTS}^c) - P_{CTS}^c)T_c}. \] (34)

\[ P_{tr_{c_1}}p_{RTS}^c = n_1\tau_1(1 - \tau_2)^{n_2-1}(1 - (1 - \tau_3)^{n_3}) + P_{u_{c_1}}P_{d \in C_1}(1 - C_2) + n_3\tau_3(1 - \tau_2)^{n_2-1}(1 - \tau_3)^{n_3-1} \]
\[ P_{tr_{c_1}}p_{CTS}^c = \xi_{c_2}(1 - \tau_2)^{n_2}. \]

For sub-cluster \( C_1 \cap C_2 \), system throughput is given by

\[ S_{c_{1},nc_2} = \frac{P_{tr_{c_{1}}p_{c_{1}}E[p]}(1 - P_{tr_{c_{1}}nc_2})\sigma + P_{tr_{c_{1}}nc_2}(p_{RTS}^c + T_{RTS}^c + (1 - P_{c_{1}nc_2})T_{RTS}^c)}{(1 - P_{tr_{c_{1}}nc_2})\sigma + P_{tr_{c_{1}}nc_2}(p_{RTS}^c + T_{RTS}^c + (1 - P_{c_{1}nc_2})T_{RTS}^c)}. \] (35)

### 2.4.4 Model validation

In this section, the model in Section 2.3.6 is extended to the multi-hop case in order to validate the proposed analytical framework.
Fig. 20 illustrates the numerical and simulation results for normalized throughput ($\lambda = 10, 50, \text{and} 100; E[P] = 1024 \text{and} 2048 \text{bytes}$) with respect to the variation of node population in each sub-cluster $C_1 - C_2$ and $C_2 - C_1$, when the node population of sub-cluster $C_1 \cap C_2$ is set to 5. Also, it is assumed that the probability that a station with membership to sub-cluster $C_1 - C_2$ or $C_2 - C_1$ selects its destination of interest in sub-cluster $C_1 \cap C_2$ is the same as the probability of the case in which it selects its desired destination in its own sub-cluster.
Fig. 20. Numerical analysis and simulation results for normalized throughput with $\lambda = 10$ (Top), 50 (Middle), and 100 (Bottom), $E[P] = 1024$ and 2048 Bytes.
It is worth noting that the simple moving average of the computer simulation results is used to obtain the above illustrated results. Basically, when a particular simulation with a certain number of contending stations is conducted on OPNET™, apparently at the beginning of the simulation the ratio of control traffic to the data traffic is much more than the amount of the above ratio when the whole system reaches a steady state. As the steady state case is being considered, choosing the moving average is wiser due to the fact that it filters out the inaccurate initial results while it relies more on the accurate steady state results achieved at the end of each simulation. At the beginning of each simulation repetition, extra traffic generated by upper layers (i.e., the network layer and above) is comparable to the averaged results obtained at the end of the simulations. Basically, each simulation is left to run long enough to let the system reach a steady state. At the end of each repetition, the channel utilization, which is calculated using the latest results derived from the latest moving average, is used.

As can be verified in Fig. 20, by increasing the frame arrival rate, the achieved normalized throughput per station is increased. The main reason for this outcome is that the obtained results reach the saturation throughput which is expected to be the maximum achievable channel utilization. Note that there is a considerable difference between the channel throughput when $\lambda$ is set to 10 and 50, although this difference reduces significantly compared to the results for $\lambda$ equal to 50 and 100.

It is important to note that for a low number of contending stations, the difference between numerical results based on the mathematical model and the simulation results is much more than the case when the number of contending stations is increased. As it was also predicted by [52, 79], when the number of contending stations per sub-cluster is increased, the accuracy of the model is consequently as the conditional transmission probability, i.e., $\tau$, for each sub-cluster is estimated more accurately. Generally speaking, the parameter $n_i$, i.e., the number of contending stations in sub-cluster $i$, has a great impact on the accuracy of Markov-based models used for random access techniques. Therefore, this is the main reason due to which the curves for the OPNET™ results and the numerical results based on the proposed analytical framework are converging for a large number of contending stations.

For a large number of contending stations (i.e., approximately more than 45 stations per sub-cluster), the numerical results based on the analytical framework and the simulations in the OPNET™ results cross each other. For a larger number of stations, the numerical results for our proposed model remain relatively at the
same values compared to the OPNET™ results. On the other hand, when more accurate analysis without rounding the results is performed, the results converge at a larger number of contending stations, i.e., approximately 50; in this case, the processing time is increased compared to the case when rounding is used for the achieved analytical results.

It should also be noted that to control the processing time for the numerical analysis, the results are truncated, and this truncation of course results in some minor estimation error. Basically, two different types of truncation are used to limit the processing time. The first truncation technique is a basic method through which the number of iterations to achieve the expected operating point for each system setup is limited. Technically, when the difference between two successive probability distributions becomes less than a particular threshold, the iteration process is stopped. In contrast to the aforementioned technique, another approach to limit the processing time of the simulations is employed. This type of truncation is indeed much more efficient when the number of contending stations is quite large. Based on this technique, the probability distributions obtained in the previous operating point with a small number of stations are utilized as the so-called initial value or starting point. Apparently, by this approach the calculation times for a large number of wireless stations becomes significantly shorter than in the case of starting from a constant pre-defined initial value.

Fig. 21 compares normalized throughput versus node population in each sub-cluster $C_1 - C_2$ and $C_2 - C_1$ using the proposed framework for two different destination selection distributions ($\lambda = 100$ and $E[P] = 1024$). Here, for one set of results it is assumed that the probability that a station with membership to sub-cluster $C_1 - C_2$ or $C_2 - C_1$ selects its destination of interest in sub-cluster $C_1 \cap C_2$ is twice as large as the probability of the case in which it selects its desired destination in its own sub-cluster. As could be predicted, due to more inter-sub-cluster collisions the achieved throughput for the case of unequal destination election probabilities is less than the set of results with equal probabilities for selecting the destination between a local- and neighboring sub-cluster.
2.4.5 Discussions and model extension

In this sub-section, it is shown how the proposed framework can be extended to more complex network topologies. So far only double-hop networks have been considered, in which all intermediate states of virtual gateways have only two arguments; each of the arguments represents the channel state of a cluster to which a virtual gateway belongs. In other words, argument $x$ in $\theta_{i,j,k}(x,y)$ represents the channel state of cluster $C_1$: $x = 0$ refers to the case when the shared communication channel of the whole cluster $C_1$ is idle. In addition, $x > 0$ refers to the case when cluster $C_1$’s medium is busy either due to an ongoing communication or a frame collision. Similarly, argument $y$ represents the channel state of cluster $C_2$.

To consider a more general case, let us assume the system configuration shown in Fig. 22. In this setup, there are four overlapping clusters:

Cluster 1: A, B₁, C₁ → Representing argument $x$,
Cluster 2: A, B₂, C₂ → Representing argument $y$,
Cluster 3: A, C₁, C₂ → Representing argument $z$,
Cluster 4: A, B₁, B₂ → Representing argument $v$.

Subsequently, five sub-clusters may be identified, each of which contains only one radio station. For station A, which can be considered a virtual gateway, all
intermediate states have four arguments, each of which represents the channel state of a cluster to which station A belongs.

Fig. 22. A general network topology.

Let us denote conditional transmission probabilities of different sub-clusters by \( \tau_A, \tau_{B_1}, \tau_{B_2}, \tau_{C_1}, \text{ and } \tau_{C_2} \). An intermediate state of station A can be represented as \( \theta_{i,k,j}(x,y,z,v) \). Here, argument \( x \) represents the channel state of the first cluster (i.e., a cluster of A, B_1, and C_1): \( x = 0 \) refers to the case when the shared communication channel of the whole cluster is idle. Moreover, \( x > 0 \) refers to the case when the cluster’s medium is busy either due to an ongoing communication or a frame collision. Similarly, argument \( y \) represents the channel state of the second cluster (i.e., a cluster of A, B_2, and C_2), argument \( z \) represents the channel state of the third cluster (i.e., a cluster of A, C_1, and C_2), and, finally, argument \( v \) represents the channel state of the fourth cluster (i.e., cluster of A, B_1, and B_2).

Now, assume an ongoing data communication between stations C_1 and C_2. Apparently, all clusters to which these two radio stations belong will be busy for the whole data transaction. Thus, \( x = y = z > 0 \). Hence, the whole system is in state \( \theta_{i,k,j}(x,y,z,0) = \theta_{i,k,j}(m,m,m,0) \), where \( m > 0 \). It is indeed worth noting that the transition probability of \( \theta_{i,k,j}(m,m,m,0) \rightarrow \theta_{i,k,j}(m-1,m-1,m-1,0) \) is independent of \( \tau_{B_1} \) and \( \tau_{B_2} \). The probability of such a state transition is equal to one due to the fact that with a probability of one neither station B_1 nor B_2 will commence a frame transmission. Apparently, in this example, it turns out that the fourth argument, \( v \), is redundant that can be eliminated. In other words, when there is an ongoing communication in the third cluster, the fourth argument can be
eliminated as it does not provide any further information, i.e., \( \theta_{i,k,j}(x, y, z, v) = \theta_{i,k,j}(x, y, z) \). Similarly, when there is an ongoing data communication in the first cluster, the second argument may be eliminated, i.e., \( \theta_{i,k,j}(x, y, z, v) = \theta_{i,k,j}(x, z, v) \), and so on.

It is worth determining the initial transition probability from state \( B_{i,k,j} \) to \( B_{i,k,j}(m, m, m) \), with \( m = K_x - 1 \) for the previous network topology when the tagged station is station A. First of all, note that \( \theta_{i,k,j}(m, m, 0), \theta_{i,k,j}(m, m, 0, m), \theta_{i,k,j}(m, 0, m, m) \), and \( \theta_{i,k,j}(0, m, m, m) \) can all be replaced by \( \theta_{i,k,j}(m, m, m) \) for any \( m > 0 \). To give a rough idea, the following few examples are given to show how the initial transition from state \( B_{i,k,j} \) to \( B_{i,k,j}(m, m, m) \) may be accomplished:

1. \( B_{i,k,j} \rightarrow B_{i,k,j}(m, m, m, 1) \) occurs with a probability equal to \( (1 - \tau_{B_1})(1 - \tau_{B_2})(1 - \tau_{C_1})(1 - \tau_{C_2}) \) and the duration of state transition is one time slot, \( \sigma \).

2. \( B_{i,k,j} \rightarrow B_{i,k,j}(K_x - 1, K_x - 1, K_x - 1) \) may occur with a probability equal to

\[
\begin{align*}
\tau_{B_1}(1 - \tau_{B_2})(1 - \tau_{C_1})(1 - \tau_{C_2}) & \left( P\{d_{B_1} = B_2\} + P\{d_{B_1} = C_1\}\right) + \\
\tau_{B_2}(1 - \tau_{B_1})(1 - \tau_{C_1})(1 - \tau_{C_2}) & \left( P\{d_{B_2} = B_1\} + P\{d_{B_2} = C_2\}\right) + \\
\tau_{C_1}(1 - \tau_{B_1})(1 - \tau_{B_2})(1 - \tau_{C_2}) & \left( P\{d_{C_1} = B_1\} + P\{d_{C_1} = C_2\}\right) + \\
\tau_{C_2}(1 - \tau_{B_1})(1 - \tau_{B_2})(1 - \tau_{C_1}) & \left( P\{d_{C_2} = B_2\} + P\{d_{C_2} = C_1\}\right) + \\
\tau_{B_1}\tau_{C_1}(1 - \tau_{B_2})(1 - \tau_{C_2}) & P\{d_{B_1} = B_2\}P\{d_{C_1} = C_2\} + \\
\tau_{B_2}\tau_{C_2}(1 - \tau_{B_1})(1 - \tau_{C_1}) & P\{d_{B_2} = B_1\}P\{d_{C_2} = C_1\} + \\
\tau_{B_1}\tau_{B_2}\tau_{C_1} & P\{d_{B_1} = B_1\}P\{d_{B_2} = B_2\}P\{d_{C_1} = C_2\} + \\
\tau_{C_1}\tau_{C_2}(1 - \tau_{B_1})(1 - \tau_{B_2}) & P\{d_{C_1} = B_1\}P\{d_{C_2} = B_2\},
\end{align*}
\]

which takes \( K_x \times \sigma \) (i.e., \( K_x \) time slots) to be completed. In this example, it is explicitly assumed that simultaneous transmissions \( C_1 \rightarrow B_1 \) and \( C_2 \rightarrow B_2 \) do not interfere with each other, and vice versa. Similarly, it is assumed that two concurrent transmissions \( C_1 \rightarrow C_2 \) and \( B_1 \rightarrow B_2 \) do not interfere with each other, and vice versa. Also note that RTS transmissions corresponding to \( C_1 \rightarrow A, C_2 \rightarrow A, B_1 \rightarrow A, \) or \( B_2 \rightarrow A \) may not be considered in the above transition probability. This is due to the fact that for any of the aforementioned cases, there is still the threat of RTS-to-RTS collision because of possible RTS transmission by another hidden node for the whole duration of the RTS frame (see Fig. 23).
Fig. 23. Station $B_2$ is hidden to $C_1$ and may cause collision during RTS transmission from $C_1$ to $A$.

As another example, consider the system setup illustrated in Fig. 24. Here, there are six overlapping clusters. It is also clear that there are altogether eight sub-clusters, each of which consists of only one station. The above clusters are named as follows:

- Cluster 1: $A_1, C_1, C_2$ → Representing argument $x$,
- Cluster 2: $A_2, C_2, C_3$ → Representing argument $y$,
- Cluster 3: $A_1, B_1, B_2$ → Representing argument $z$,
- Cluster 4: $A_2, B_2, B_3$ → Representing argument $v$,
- Cluster 5: $A_1, A_2, C_2$ → Representing argument $w$,
- Cluster 6: $A_1, A_2, B_2$ → Representing argument $s$. 
Fig. 24. A complex multi-hop network topology.

Now assume the case when stations A₁ and B₁ are conducting an ongoing data transaction. Let us consider station C₂ as the tagged station. Due to the ongoing data communication, the medium of the following clusters is marked as busy: 1, 3, 5, and 6. Let us also assume that our tagged station of interest is chosen to be C₂. As this station belongs to clusters 1, 2, and 5, its intermediate states may be represented by \( \theta_{i,k,j}(x, y, w) \). Here, \( x \) represents the channel state of cluster 1, \( y \) represents the channel state of cluster 2, and \( w \) represents the medium state of cluster 5. Assuming that the ongoing communication is successful, the above intermediate state may be rewritten as \( \theta_{i,k,j}(m, 0, m) \). Now, for station C₂ to transit from \( \theta_{i,k,j}(m, 0, m) \) to \( \theta_{i,k,j}(m-1, 0, m-1) \), station C₃ should not start any transmission. Therefore, the probability of such a state transition is given by \( 1 - \tau_{c₃} \). Here, there is no need to consider station A₂’s transmission probability due to the fact that it is not allowed to perform any kind of medium access (i.e., either by sending an RTS or by returning a CTS frame to for instance station B₃) as its network allocation vector is not zero.

2.5 Chapter summary

The problem of modeling non-saturated contention-based 802.11 ad hoc networks with arbitrary topologies was considered in this chapter. Using Markov processes, a generic framework to model the hidden terminal problem as well as the unreachability problem was presented, leading to better understanding of the underlying challenges in designing efficient medium access control protocols. The mathematical framework was defined in such a way that it provides a clearer view to the unreachability problem and all key factors that cause throughput...
degradation in a multi-hop ad hoc network due to this problem. In addition, the proposed models clearly facilitate demonstrating the impact of the unreachability problem on the performance of multi-hop networks and show how the unreachability problem degrades the achieved throughput and channel capacity by the contending radio stations depending on the deployed network topology.

Furthermore, access delay was carefully studied using the proposed Markovian framework, taking into account frame delay due to contention resolution as well as frame delay due to queuing delay in a tagged station’s transmission buffer. The proposed model was utilized to derive the exact expression of service time and queuing delay. In the following chapter, using the lessons gained in this chapter, the unreachability problem will be tackled and a MAC protocol for single-channel multi-hop ad hoc networks is proposed.
3 MAC protocol design for future multi-hop networks

The IEEE 802.11 standard has been shown to be quite inefficient in multi-hop networks [1, 2, 12, 14, and 16]. Besides the hidden terminal and exposed terminal problems, there is also an unreachability problem [1 and 18], which may result in link/routing failures and unfairness among multiple traffic flows. In this chapter, a MAC protocol, called eMAC, is first proposed. Under the proposed scheme, stations maintain Double Hop Neighborhood (DHN) graphs while exchanging designated eMAC tables to share their knowledge about their neighborhood topology. Using a DHN graph and an adaptive unreachability reporting mechanism, stations are reliably informed about their neighbors’ unreachability status. Hence, they avoid establishing link layer connections with their unreachable neighbors, and, consequently, network resources are not consumed for unsuccessful connection establishment efforts. Furthermore, an adaptive table broadcasting technique is proposed to facilitate topology information dissemination in mobile ad hoc networks. Performance of the proposed schemes is evaluated and compared with earlier known schemes through simulations. The achieved results show performance enhancement due to better handling of the unreachability problem, possible heterogeneous power distribution among contending stations, and node mobility issues.

This chapter is dedicated to the problem of designing a single-channel MAC protocol for mobile ad hoc networks, as will be presented in Section 3.1. The receiver blocking problem is tackled by proposing several mechanisms, explained throughout the chapter. Eventually, concluding remarks are given in Section 3.2.

3.1 eMAC – a MAC protocol for the next generation ad hoc networks

In ad hoc networks the unreachability problem occurs when the destination of interest is situated either in the communication or interference range of another transmitting node. In such a case, efforts to set up a link layer connection with the destination station fail either due to frequent collisions taking place between the transmitted RTS or DATA frames or unwanted overheard frames originated from an ongoing transmission hidden to the source node. Failure may also occur due to the inability of the destination station to respond by CTS or acknowledgement frame as a consequence of sensing the carrier caused as a result of being situated
in the interference range of a transmitting node. Here, the former case, caused by communication range, is referred to as the unreachability of Type I, while the latter, caused by interference range, is referred to as the unreachability of Type II.

The unreachability problem becomes more severe in a multi-hop environment and results in packet dropping, starvation of part of the traffic flows, and possibly unnecessary network layer re-routing [18, 94]. Even though great efforts have been devoted to solving the hidden and exposed terminal problems, limited contributions have been reported in the literature to address the receiver blocking problem, e.g., [18 – 21, 94 – 96]. This problem was first identified by [3] and, subsequently further investigated by [18], [19 – 21], and [95 – 97]. The proposed scheme in [18] attempts to alleviate the impact of receiver unreachability using dual-channel architecture; thus, this results in protocol complexity and dual-channel deployment which may not be always desirable. In fact, both the dual-channel and extra channel approaches have market availability and 802.11 protocol compatibility issues. In [97], on the other hand, although the unreachability problem is tackled without any extra channel, the proposed protocol suffers from a broadcast storm at the MAC layer in certain network topologies when reporting the receiver blocking situation.

The protocol in [97] adds a couple of new control frames to ease reporting of the unreachability situation to solve the receiver blocking problem. When a station is notified about an upcoming data communication due to which it will be unreachable, it is given an opportunity to inform its one-hop neighbors about the forthcoming unreachability. In principle, right after the RTS/CTS negotiation and before commencing the actual DATA transmission phase, the stations, which will shortly become unreachable, are given the chance to report their imminent unreachability status using a designated broadcast frame called ICP (i.e., individual communication pause). Since there is no exception for participation in unreachability reporting, all potential candidates contribute to this phase; thus, in certain network topologies collisions may occur among broadcast ICP frames. Such collisions are caused by unconditional ICP frame broadcasting, which is referred to as an ICP broadcast storm. The protocol in [97] has two distinct operation modes, named Normal and Excessive Try; while employing a complex timing system and implementing two separate memory spaces, called the Unreachability Vector and Unreachability Cache, to buffer those frames for which the receiving node(s) are unreachable.

In this section, the main goal is to solve the broadcast storm problem in [97] by introducing a new technique to prevent unnecessary simultaneous
unreachability reports. As will be explained later, to attain the aforementioned goal, maintenance of a double-hop neighborhood graph by every station is proposed. The DHN graph of each station gives an estimate of its double-hop neighborhood topology. In this chapter, by a one-hop neighbor, we mean an immediate station situated in a considered station’s communication range and vice versa. In addition, a two-hop neighborhood resembles the same concept known in the Network Layer: including two consecutive one-hop neighbors. A station, while reporting its expected unreachability status, indicates the MAC address of the station due to which its unreachability is caused. In this chapter, the station due to which unreachability is caused is called the unreachability cause. This station is the emitting node due to which its neighboring terminals become unreachable. As a result, those stations that need to be informed about the unreachability of their neighbors are only required to receive one report (i.e., ICP frame) to obtain a reasonable picture of their neighbors’ unreachability. Therefore, the number of concurrent unreachability reports can be reduced considerably while the unreachability resolution is accomplished in a more reliable and efficient way. In fact, it is worth noting that the aforementioned approach provides a solution for the case where the unreachable station is situated in communication range of the unreachability cause (i.e., unreachability of Type I).

To address the same problem for those stations situated in interference range of the unreachability cause, but not its communication range (i.e., unreachability of Type II), in this section a simple, yet effective, protocol enhancement called Relayed ICP is proposed. Basically, an out-of-band Busy Tone (BT) is used to facilitate identification of blocking circumstances by unreachable stations of Type II. Such stations, when perceiving a sequence of an overheard BT and an ICP frame of Type I, are informed about their unreachability of Type II. Upon such situations, they become eligible for broadcasting an ICP frame of Type II; in other words, an ICP frame of Type II resembles a relayed ICP frame to further enhance the unreachability resolution.

In ad hoc networks, addressing the unreachability problem becomes even more challenging when stations transmit at different power levels. Although earlier proposed schemes, e.g., [18 and 97], assume a homogeneous network as the basic system model, practical networks are heterogeneous in power capabilities. In a homogeneous network, if station “A” is located in the radio range of station “B”, then station “B” should be located in the radio range of station “A”, as well. In such scenarios, a Friendship relation between “A” and “B” is defined as illustrated in Fig. 25a. On the other hand, in heterogeneous networks,
the aforementioned rule is not valid for all situations. As shown in Fig. 25b, station “B” is resided in the radio range of station “A”, while the reverse situation is not true. In this case, a different notation is used, and “A” is defined as $\overline{X}$ station for “B”, while “B” is defined as an X station for “A”.

![Diagram](image)

**Fig. 25. (a) Friendship scenario, (b) X and $\overline{X}$ scenarios.**

In [98] and [99], two approaches have been proposed to reduce problems encountered in heterogeneous networks. In [98], the authors showed that the performance of the IEEE 802.11 MAC protocol degrades significantly in a network with nodes that transmit at diverse power levels. The primary reason for this degradation is that the high power nodes cannot overhear the exchange of the low power RTS and CTS messages, and, as a result, the high power nodes initiate transmissions while the low power nodes are in communication. This in turn causes an increase in the number of retransmission attempts by low power nodes, thereby increasing the effective traffic load in the network. Broadcasting the CTS message within a neighborhood of a communicating low power node is another scheme proposed in [98]. Nodes that hear the CTS message propagate the message further in an attempt to notify high power nodes in the vicinity of the ongoing communication. This approach, however, will degrade network throughput by about 20% as compared with the legacy IEEE 802.11 MAC scheme, due to an increased amount of overhead generated by propagation of the first CTS message. [99] presents an extension to the aforementioned proposal, recommending reduction of the number of broadcast CTS frames by the use of a
simple timing schedule. Finally, the authors of [100], inspired by [99], applied slight modifications to their earlier work [98] and showed that the channel utilization and throughput can be improved by reduction of overhead incurred due to propagation of CTS frames by the use of some simple modifications.

In this section, while addressing the unreachability problem and related issues, possible heterogeneity in power distributions is also taken into account in order to propose more robust and reliable algorithms, resulting in addition of further features in addition to the presented MAC enhancements. The key contributions of the solution are the following: By incorporating topology-awareness and smarter decision making algorithms into the MAC protocol, the impact of the unreachability problem is reduced, resulting in much more efficient channel utilization and higher transmission capacity. In addition, the impact of stations’ heterogeneous power capabilities on the network topology is taken into account to propose a more robust reporting scheme. Finally, by gathering part of the network topology information at the MAC layer, the overhead due to route discovery and route maintenance is decreased significantly. It is believed that such topology information collected at the MAC layer can be utilized in cross-layer protocol optimization as well.

### 3.1.1 Protocol architecture and basic concepts

#### Preliminaries

In this sub-section, the general architecture of our proposed protocol is explained. Our aim is to introduce a technique to avoid the spread of unnecessary simultaneous unreachability reports, and, therefore, maintenance of the DHN graph by every station is proposed. Each station can be either mobile or stationary. The DHN graph of each station gives an estimate of its double-hop neighborhood topology. Each station is responsible for figuring out the unreachability duration period. This can be accomplished by overhearing RTS/CTS control frames if the unreachable station is situated in the communication range of the unreachability cause (i.e., unreachability of Type I). On the other hand, for those stations situated in the interference range of the unreachability cause, but not its communication range (i.e., unreachability of Type II), this goal may be achieved by overhearing ICP frames of Type I received right after a busy tone of a particular duration. In this case, the unreachable station of Type II broadcasts an ICP frame of Type II.
for which there are two address fields: the first address field carries the MAC address of an unreachable station of Type I from which an ICP frame of Type I has been received, and the second address field carries the MAC address of the unreachability cause. In this thesis, an ICP frame generated by an unreachable station of Type I is referred to as an ICP frame of Type I. Similarly, an ICP frame generated by an unreachable station of Type II is referred to as an ICP frame of Type II. The former has only one address field for carrying MAC address of the unreachability cause, while the latter has two address fields for carrying 1) the MAC address of the unreachable station of Type I from which an ICP frame of Type I has been received and 2) the MAC address of the unreachability cause. For simplicity, throughout the section an ICP frame of Type I is denoted by ICPv1. Similarly, an ICP frame of Type II is denoted by ICPv2.

Note that ICP frames have neither a source address nor a receiver address. In addition, its Duration/ID field is used to indicate the duration of unreachability. As shown in Fig. 26, an announcement of an upcoming unreachability status is either performed right after an overheard RTS and/or CTS frame (Unreachability of Type I), or upon overhearing an ICP frame of Type I received right after a busy tone of a particular duration (Unreachability of Type II). The duration of the busy tone is equal to the length of the DS control frame, as illustrated in Fig. 26.
Fig. 26. Frame exchange in eMAC.

Besides the DHN graph, each station maintains an eMAC table, which keeps track of all its immediate one-hop neighbors. Basically, the eMAC table is generated from the DHN graph. On the other hand, by reception of all one-hop neighbors’ eMAC tables, each station either constructs or updates its local DHN graph, as well. To clarify this issue, consider the network topology illustrated in Fig. 27.

Fig. 27. An example to show how topology information is exchanged using eMAC table broadcasting.
In this configuration, station “A” can receive eMAC tables of all its immediate one-hop neighbors, i.e. stations “B”, “C”, and “F”. Similarly, station “B” is able to obtain tables of stations “A”, “F”, “G”, and “H”. Station “A” is able to easily construct a DHN graph to mimic its double-hop neighborhood topology. In this way, “A” can determine for instance whether stations “E” and “D” are located in their one-hop neighborhood. Now, let us assume that stations “G” and “H” are willing to perform a long-term data exchange using packet fragmentation. In this scenario, “G” is supposed to serve as a source station and “H” is assumed to be the destination station. Apparently, this leads to the unreachability of both “B” and “F”. When station “A” receives an ICP frame sent by either “B” or “F”, it verifies the appended field indicating the unreachability cause (in this example, “G”), and using its DHN graph, it concludes that all one-hop neighbors of stations “G” will be unreachable as well. By this approach, “A” needs to receive only one ICP frame to be informed about the unreachability of “B” and “F”. In other words, there is no need for both “B” and “F” to participate in the ICP frame broadcasting procedure. Note that this example is only about unreachability of Type I, as both “B” and “F” are unreachable stations of Type I. For the case of an unreachable station of Type II, the same approach is followed by station “A”. As in ICP frames of Type II both the MAC address of an unreachable station of Type I, from which an ICP frame of Type I has been received, and the MAC address of the unreachability cause are appended, station “A” concludes that not only “B” and “F”, but also any one-hop neighbor of these two stations will be unreachable.

Control signalling and connection establishment

To alleviate the impact of the receiver blocking problem, two new control frames (i.e., DS and ICP) are defined in [97]. The DS frame is sent by the source station right after reception of the CTS frame, while the ICP frame may be sent by the source station’s one-hop neighbors right after the DS frame and before the actual data communication between the source and the destination nodes.

In eMAC, piggybacking [14] is supported regardless of whether packet fragmentation is accomplished. On the other hand, DS and ICP (both Type I & II) are only employed when the information is delivered through packet fragmentation. When the source and destination stations agree on piggybacking, the More Data bit in the Frame Control field of CTS and DS (if used) is set to one. In addition, the More Fragments bit in the Frame Control field of RTS, CTS, and DS is set to one to imply packet fragmentation during the whole data transaction.
In fact, this is done since all one-hop neighbors of source and destination stations should be informed about the possibility of packet fragmentation for the whole data transmission to consider their participation in the ICP broadcasting phase. Recall that ICP broadcasting is performed only when the neighboring data transmission is planned to be accomplished using packet fragmentation. A station that receives any of the aforementioned frames with the More Fragments bit set is considered as a candidate for ICP frame broadcasting. Note that throughout this chapter, we refer to such stations as ICP candidates. In the data fragmentation scenario, CTS and DS frames’ Duration/ID fields indicate the unreachability periods during which receiver blocking will last; therefore, the same duration value is used in all ICP frames. In eMAC, the Order bit of all ICP frames transmitted by ICP candidates situated in the radio range of the source station should be always set to zero. Besides, the aforementioned bit in ICP frames generated by one-hop neighbors of the destination station is set to zero, if, and only if, the received CTS frame from the destination node has More Fragments = 1 and More Data = 1. If the ICP candidate is situated only in the radio range of the destination station while no piggybacking has been agreed to be conducted, the corresponding Order bit in the ICP frame is set to one.

If a station receives an RTS with More Fragments = 1, it should wait for a DS frame from the same terminal before broadcasting any ICP frame of Type I. In this case, the unreachable station is also supposed to be able to detect the BT generated by the unreachability cause for the whole duration of DS, shown in Fig. 26. Upon reception of the DS frame, the unreachable station of Type I may perform ICP broadcasting during a period of ICPtime. Although the DS frame makes the unreachability announcement possible for the source station’s neighbors, the same problem should be addressed properly for the stations situated in the destination station’s radio range. In fact, CTS serves the same role as DS, but in the receiver side. Thus, when a station receives a CTS frame, for which More Fragments = 1, it first wait for $2 \times SIFS + DS_{time}$, and, subsequently, it may perform ICP broadcasting during a time interval of ICPtime. As an example, in Fig. 26 station “C” is an unreachable station of Type I as it observes both BT and DS frames originated from station “A”. The ICP frame shown as ICPv1 is sent by station “C”, carrying the station “A” MAC address as the unreachability cause.

Following unreachable stations of Type I, it is unreachable stations’ of Type II turn to participate in ICP broadcasting. When a station perceives a sequence of an overheard BT and an ICP frame of Type I (i.e., ICPv1), it becomes eligible for
broadcasting an ICP frame of Type II (i.e., ICPv2), which resembles a relayed ICP frame to further enhance the unreachability resolution. Such stations may also perform ICP broadcasting during a time interval of $ICP_{time}$. In Fig. 26, station “D” is an unreachable station of Type II. The ICP frame shown as ICPv2 is sent by station “D”, carrying the station “C” MAC address as the one from which an ICP frame of Type I has been received, in addition to the station “A” MAC address as the unreachability cause.

**Double-hop neighbourhood graph & eMAC table concepts**

The DHN graph is a time-variant data structure denoted by $G(V(t),E(t))$, where $V(t)$ represents the set of its vertices while $E(t)$ stands for the set of its edges at a given time instant $t$. Based on topology information received from one-hop neighbors in the form of eMAC tables, the DHN is updated accordingly. When the DHN graph is updated, the synchronous eMAC table, denoted by $Ξ(t)$, is generated.

Each station records all its neighbors in its local DHN represented by a set of vertices classified into two different categories. One-hop neighbors are grouped together to form the class $N$ group. Each member of this group is simply referred to as a class $N$ neighbor. In addition, double-hop neighbors are grouped into another set named the class $H$ group. Furthermore, similar to the case of class $N$, each member of a class $H$ group is referred to as a class $H$ neighbor. Principally in the DHN graph, there should be a unique edge connecting each class $N$ neighbor to the local station. Here, by local station, we mean the owner of the DHN graph. On the other hand, class $H$ neighbors are connected to the local station via a couple of edges and obviously through class $N$ neighbors. Fig. 28 illustrates the DHN graph maintained by station “A”, the neighborhood topology of which is shown in Fig. 27.

Fig. 28. Double-hop neighborhood graph maintained by station “A” in the network shown in Fig.24.
Each station maintains two different versions of an eMAC table at any time. One is called the synchronous eMAC table and is denoted by $\Xi(t)$. This table is directly generated from the local DHN, whenever it is updated. The second table is denoted by $\bar{\Xi}$ and represents the latest version of the eMAC table that has ever been broadcast over the air interface. Basically, each station should broadcast its synchronous eMAC table $\Xi(t)$ in a regular fashion; whenever synchronous eMAC table $\Xi(t)$ is broadcast in its corresponding due Beacon Interval (BI), the existing $\bar{\Xi}$ table is simply replaced by $\Xi(t)$. Note that $\bar{\Xi}$ is replaced by synchronous eMAC table $\Xi(t)$ only when $\Xi(t)$ is being broadcast in its due beacon interval, and not when $\Xi(t)$ is updated due to reception of new topology information. Hence, these two tables do not necessarily have the same information at any given time.

Both types of eMAC tables, i.e., $\Xi(t)$ and $\bar{\Xi}$, have the same data structure. Each record of an eMAC table consists of two fields: The first one is a fixed length field that contains a class $N$ neighbor’s MAC address, while the second field consists of a chain of Relationship flags. When an eMAC table contains $k$ records, each record has exactly $k$ relationship flags. These flags determine the one-hop class $N$ neighbors’ relationship status relative to each other. So, if the owner of a local eMAC table is already situated in the radio range of stations $i$ and $j$ and could have received their broadcast tables, then, $j$th flag of the $i$th record in its table, denoted by $m(i,j)$, determines the relationship status of stations $i$ and $j$. The aforementioned flags construct the eMAC matrix, whose elements are determined as follows:

$$
\begin{align*}
&\text{if stations } i \text{ and } j \text{ are friends:} & m(i,j) = m(j,i) = 11 \\
&\text{if stations } i \text{ and } j \text{ are not related:} & m(i,j) = m(j,i) = 00 \\
&\text{if station } i \text{ is } X \text{ for station } j: & m(i,j) = 10 \\
&\text{if stations } i \text{ is } X \text{ for station } j: & m(i,j) = 01
\end{align*}
$$

Diagonal elements of the eMAC matrix are equal to the don’t care value, which is denoted by $\times$. Assuming identical transceivers for all radio stations, for eMAC table broadcasting there is no need to send the entire eMAC matrix (or eMAC table); this is due to the fact that element $m(j,i)$ can be determined using element $m(i,j)$ and vice versa. A typical eMAC table is illustrated in Fig. 29.
When the DHN graph is updated due to reception of new neighborhood topology information, synchronous eMAC table $\Xi(t)$ is consequently regenerated. This means that DHN and $\Xi(t)$ keep the most up-to-date information about the double- and one-hop neighborhood of the local station, respectively. Throughout this chapter, we may refer to the local station as the tagged station.

As stated earlier, the most up-to-date version of the eMAC table should be broadcast in a regular fashion. To determine how frequent and when the eMAC tables are broadcast, the number of beacon intervals that have to elapse before broadcasting the latest version of the eMAC table is specified. When the synchronous eMAC table $\Xi(t)$ is broadcast in its due beacon interval, it is also saved as $\Xi$ to represent the last version of the local eMAC table that has ever been broadcast over the air interface; in addition, the number of beacon intervals that have to elapse before the next table broadcasting is also calculated and included inside the MAC service data unit (MSDU) carrying the contents of the eMAC table. This integer parameter is denoted by $K$ and is recalculated every time the synchronous eMAC table $\Xi(t)$ is broadcast. The way how this parameter is
updated is based on the amount of difference between $\Xi(t)$ and the latest version of table that has ever been broadcast, i.e., $\Xi$. This difference is denoted by $\Delta$ and defines the next $K$ according to the following expression:

$$K_{t+1} = \begin{cases} \max(K_{\min}, K_t - \Delta) & \text{if } \Delta \geq d \\ \min(K_{\max}, K_t + c) & \text{if } \Delta < d \end{cases} \quad (36)$$

where $c$ is a constant value. Typically, it is appropriate to choose $K_{\min} = 10$, $K_{\max} = 100$, and $c = 5$. According to the IEEE 802.11 specifications, the beacon interval is usually set to 0.1 seconds. By setting $K_{\min}$ to 10, the minimum update period becomes 1 second, which is reasonably adequate for relatively highly dynamic mobile ad hoc networks. Similarly, by setting $K_{\max}$ to 100, the maximum update period becomes 10 seconds, which is sensibly sufficient for relatively fixed networks. In each stage, the value of $\Delta$ is calculated using the following expression:

$$\Delta = \Xi \bigtriangleup \Xi(t) = \big(\Xi - \Xi(t)\big) \cup \big(\Xi(t) - \Xi\big). \quad (37)$$

Based on the current value of $\Delta$, whether it is greater or less than a pre-defined threshold $d$, it is decided whether to increase or decrease the value of $K$ for the next table broadcast round. Note that the value of $K$ defines the number of beacon intervals that should elapse before the next eMAC table broadcast occurs.

The way symmetric difference $\Xi \bigtriangleup \Xi(t)$ is calculated is explained using an example illustrated in Fig 30. To map the obtained $\Xi \bigtriangleup \Xi(t)$ to a real value, it is necessary to assign different weights to dissimilar links in $\Xi \bigtriangleup \Xi(t)$. Fig. 30 shows all possible wireless links that may appear in $\Xi \bigtriangleup \Xi(t)$ and their corresponding weights. The two top-most graphs in Fig. 30 resemble $\Xi$ and $\Xi(t)$, respectively. The two differential graphs in the middle of Fig. 30 represent $\Xi - \Xi(t)$ and $\Xi(t) - \Xi$, correspondingly. $\Xi \bigtriangleup \Xi(t)$ is obtained via the union of the aforementioned differential graphs, as shown in the lower part of Fig. 30. The obtained scalar value for $|\Xi \bigtriangleup \Xi(t)|$ will be $12 \times c$. Note that $c$ is exactly the same parameter used for updating the value of $K$. Also, threshold $d$ is defined to be equal to $3 \times c$. 

115
Fig. 30. Symmetric difference calculation example \( \Xi \Delta \Xi(t) \).

Fig. 31 shows the way \( K \) is updated for successive broadcast iterations. Since the beacon interval during which table \( \#i \) was broadcast, three eMAC tables have been received from three different class \( N \) neighbors at time instances \( t_0, t_1, \) and \( t_2 \). Therefore, the DHN graph and consequently \( \Xi(t) \) should have been updated at the aforementioned time instances, as well. Assuming \( K \) to be \( K_i \) when broadcasting table \( \#i \), the next table broadcast occurs after the elapse of \( K_i \) beacon intervals. On the due beacon interval, the station recalculates the value of \( \Delta \) and concludes that it is greater than threshold \( d \); therefore, the value of \( K_i \) is updated consequently (in this case, it is decreased). The new value is denoted by \( K_{i+1} \), and the updated table being delivered over the air is named \( \#i + 1 \). During the second period, there is only one table reception at \( t = t_3 \) which leads to \( \Delta < d \) at the end; as a result, \( K \) is increased for the next broadcast round. The new value of \( K \) is denoted by \( K_{i+2} \). Of course, there are still some updates that should be reflected in
DHN and $\Xi(t)$. The broadcast table is denoted by $#i + 2$. During the third period, no update is received and therefore the table to be broadcast on the next due beacon interval will be the same as table $#i + 2$.

![Fig. 31. eMAC table reception from class $N$ neighbors and $K$ value recalculation procedure based on $\Delta$ during successive broadcast periods.](image)

Broadcasting eMAC tables is accomplished using contention-based media access rules as defined in IEEE 802.11 DCF. If, because of a high traffic load, the station is unable to perform its table transmission on time, it should persistently continue attempting to accomplish it in any subsequent beacon interval. In this case, the next due beacon interval is counted from the actual beacon interval during which the last table broadcast has occurred. Also note that when the eMAC table is broadcast, the updated value of $K$ should also be included inside the MSDU carrying the table.

Since synchronous table $\Xi(t)$ is generated from the DHN graph, it is important to define a set of highly adaptive rules by which it is maintained and updated. By reception of all known class $N$ neighbors’ eMAC tables on a regular basis, DHN is continuously updated, and, consequently, $\Xi(t)$ is generated to reflect the latest topology changes in the local station neighborhood. Fig. 32 illustrates the algorithm based on DHN, and $\Xi(t)$ are updated. The right part of the chart deals with the case when an eMAC table is received from a known class $N$ neighbor that has already been recorded in DHN and $\Xi(t)$. The left part of the chart is dedicated to the scenario in which a table is received from an unknown station or from a class $H$ neighbor. Recall that upon broadcasting an
eMAC table, each station is mandated to put the updated value of \( K \) inside the MSDU carrying the table. Upon reception of an eMAC table from either an unknown station or a class \( H \) neighbor, the appended \( K \) is exploited to load a local counter. This counter is decremented by one whenever a beacon interval is elapsed. If another table is received from the same station before the aforementioned counter reaches zero, DHN and \( \Xi(t) \) are subsequently updated. On the other hand, if the counter reaches zero and no further table is received from the same station, the previous table reception is simply ignored and no modification is made in DHN and \( \Xi(t) \).

To illustrate what is done when a table is received from an unknown station or a class \( H \) neighbor, consider Fig. 33. In this figure, station “B” receives an eMAC table from an unknown station “A”. Consequently, it loads a counter with the value of \( K_i \) included in the received table from “A” and waits, but does not update its DHN or \( \Xi(t) \) until it receives another table from station “A”. The same situation occurs for station “A” with respect to station “B”; station “B” is also an unknown station for “A” and, as a result, “A” loads a counter with the value of \( K_j \) included in the received table from “B”. When station “A” wants to broadcast its table, it does not change the value of \( K_i \) since it has received only one table from “B” but no table from known class \( N \) neighbors. On the other hand, when station “B” wants to rebroadcast its table for the second time, since it has already received two tables from “A”, it updates the value of \( K_j \) to \( K_{j+1} \).
Table-assisted ICP broadcast mechanism

In the previous sub-section, the way the DHN graph and eMAC tables in every station should be maintained in order to keep track of neighborhood information of the station has been explained. Here, it is elucidated how such information may be exploited to alleviate the ICP broadcast storm.

Indeed, the most challenging issue in ICP broadcasting is the way by which the candidates are scheduled in order to accomplish an efficient and reliable unreachability reporting. To prevent an ICP broadcast storm, instead of the 1–persist approach to broadcasting ICP frames [97], each station should sense the channel (i.e., carrier sensing) and only when the medium is found idle broadcast an ICP frame with a probability of $\beta$. For each station, the value of $\beta$ is a function of its neighbors’ geographic distribution and their relationship with each other.

To clarify the core idea of our proposed scheme, at first and for the sake of simplicity, consider a homogenous network. In MANETs with homogenous power capabilities, the relationship between stations can be categorized as either ‘friendship’ or ‘non-related’. The latter corresponds to the case in which stations are situated out of each other’s radio range. When the network allocation vector of the ICP candidate is non-zero, it is not allowed to transmit any ICP frame. Furthermore, the ICP candidate’s probability function $\beta$ should be a function of $N$, i.e., the number of stations that have simultaneous friendship with both the unreachability cause and the ICP candidate (unreachability of Type I), or the number of stations that have simultaneous friendship with any of the unreachable stations of Type I and the ICP candidate (unreachability of Type II). This function, which is strictly descending with respect to $N$, comprises a number of tuneable
factors $a_k$ that should be adjusted so that an objective function $J$ is minimized or maximized. The optimum values for factors $a_k$ give the best value for $\beta$, with which the ICP candidate may broadcast an ICP frame.

A time interval equal to $ICP_{time}$ is dedicated to each type of ICP broadcast phase. Obviously, when the ICP candidate has no neighbors, except the unreachability cause in the case of unreachability of Type I or any of unreachable stations of Type I in the case of unreachability of Type II, it may not broadcast any ICP frame. Thus, if $N'$ denotes the number of the candidate’s neighbors that should be informed by an ICP, then:

$$\beta(N, N') = \beta(N) \times (1 - \delta(N')),$$

where $\delta(N') = 1$ for $N' = 0$ and is zero elsewhere.

If all tuneable factors $a_k$ are denoted as vector $a$, then we have $\beta = \beta(N, a)$. Now consider a tagged ICP candidate. In general, all parameters related to the tagged ICP candidate are defined to have indices equal to “0”, where $a_0$ represents the tagged candidate’s vector of its tuneable factors. Hence, for the tagged candidate we have $\beta_0 = \beta_0(N_0, a_0)$. In the same way, for each of the tagged ICP candidate’s neighbor that is itself an ICP candidate, i.e., station $i$ ($1 \leq i \leq N$), we have $\beta_i = \beta_i(N_i, a_i)$. Subsequently, the following two vectors maintained and utilized by the tagged ICP candidate are defined as $\mathbb{R}_0 = [\beta_0, \beta_1, \beta_2, ..., \beta_N]$ and $\mathbb{S}_0 = [N_0, N_1, N_2, ..., N_N]$. Note that the tagged candidate is able to determine the value of $N_i$ by the use of its local DHN. In addition, due to the fact that each $\beta_i$ is a function of vector $a_i$, the following matrix may be defined to be used in the tagged ICP candidate’s target objective function $A_0 = [a_0, a_1, a_2, ..., a_N]$. The tagged candidate’s objective function, denoted by $J_0$, depends on $\mathbb{R}_0, \mathbb{S}_0$ and $A_0$. This objective function can be also expressed as $J_0 = J_0(\mathbb{R}_0, A_0) = J_0(\mathbb{S}_0, A_0)$. This objective function may be defined in several ways; for instance, one possible definition can be based on the probability of broadcasting an ICP frame by exactly one candidate at a given $ICP_{time}$ interval:

$$J_0 = \sum_{i=0}^{N} \beta_i \prod_{j=0}^{N} (1 - \beta_j).$$

This objective function is maximized as follows:

$$\nabla J_0 = 0 \rightarrow a_0^* = a_0^{{opt}}.$$
By obtaining the optimum values for vector $\mathbf{a}_0$, the tagged ICP candidate may also share its optimum factors with its friends. The friends of a station refer to its one-hop neighbors with which it has a *Friendship* relation. By this feature and after experiencing a finite number of iterations, stations are able to attain the optimum solution for the ICP broadcasting in their neighborhood.

When a wireless network with heterogeneous power capabilities is considered, the probability function is defined as $\beta_i = \beta_i(N_{f,f}, N_{f,x}, N_{x,f}, N_{x,x}, a_i)$, where the first argument is “the number of the unreachability cause’s friends which are also the tagged ICP candidate’s friends”, the second one is “the number of the unreachability cause’s friends that are $X$ for the tagged ICP candidate”, the third argument is “the number of $X$ stations for the unreachability cause which are $X$ for the tagged ICP candidate’s friends”, and finally the fourth is “the number of $X$ stations for the unreachability cause which are $X$ for the tagged candidate”, respectively. Here, attention is focused on the case of unreachability of Type I. Similar mechanisms may be applied for the case of unreachability of Type II.

To make the proposed broadcasting scheme more efficient, it is desirable to investigate scenarios in which it fails due to frame collisions. The first and the most obvious scenario is the case in which ICP frames collide with each other. So far, in this chapter an adaptive location-based scheme to improve the reliability of the broadcasting process has been proposed. Another scenario in which the proposed scheme fails occurs when the station(s) that should be informed about their neighbors’ unreachability state, are also unreachable. This scenario is illustrated in Fig. 34.

**Fig. 34. ICP broadcasting to an unreachable station.**

In this case, station “E” has received the DS frame transmitted by station “A” and will therefore be a candidate for ICP broadcasting. Also assume that “E” knows that station “F” is unreachable due to “D” ↔ “C” communication. In this case, it is important for “E” to know whether “F” is unreachable due to residing in the radio range of either the *source* or *destination* station in “D” ↔ “C” communication. Moreover, it is important for “E” to know whether any piggybacking has been agreed on by stations “D” and “C”. According to the aforementioned expressions, there are three different scenarios as given below:
Scenario A — Station “F” is unreachable since it is situated in the radio range of station “D”, which is itself the receiving end-point in “D” ↔ “C” communication. Also, “D” ↔ “C” data transmission is unidirectional, meaning that station “C” is always the source and station “D” is always the destination. Since “F” has received the CTS frame of station “D”, it has realized that More Fragments = 1 and More Data = 0; therefore, it knows that “D” is only the receiving end-point of this communication and no piggybacking will be done during the data transfer. In this case, station “F” should have set Order bit = 1 in its ICP frame to inform station “E” about the situation. When station “E” becomes unreachable due to “A” ↔ “B” data communication, it should perform ICP broadcasting to “F” since the probability of “F” inability to receive station “E’s” ICP is too low. In other words, “F” should be included in the $N'$ value.

Scenario B — Station “F” is unreachable since it is situated in the radio range of station “D”, which is itself the source station in “D” ↔ “C” communication. “D” ↔ “C” data transmission is unidirectional, and therefore station “C” is always the destination and station “D” is always the source. Since station “F” might have received both the RTS and DS frames of station “D”, it knows that “D” is the transmitting end-point of this communication. In this case, station “F” should have set Order bit = 0 in its ICP frame to inform station “E” about the situation. When station “E” becomes unreachable due to “A” ↔ “B” data communication, station “F” should not be included in the $N'$ value since the probability of “F’s” inability to receive station “E’s” ICP is too high.

Scenario C — Station “F” is unreachable since it is situated in the radio range of station “D”, which is both the transmitting and receiving end-point of “D” ↔ “C” transmission (i.e., piggybacking has been agreed on for this communication). “D” ↔ “C” data transmission is bidirectional. Station “F” knows that “D” is involved in a bidirectional data communication. In this case, station “F” should have set Order bit = 0 in its ICP frame to inform station “E” about the situation. When station “E” becomes unreachable due to “A” ↔ “B” data communication, station “F” should not be included in the $N'$ value since the probability of “F’s” inability to receive station “E’s” ICP is too high.
Fig. 35. ICP broadcasting to an unreachable station.

As a concluding remark and to determine the ascending or descending features of probability function $\beta_0$ with respect to the different types of the unreachability cause’s neighbors, let us consider Fig. 35 while defining an array consisting of three distinct elements. These elements identify different types of the unreachability cause’s neighbors and from left to right have the following meanings (see Fig. 35): The first element corresponds to the relationship status between a tagged ICP candidate and the unreachability cause, which can be either $F$ (Friendship) or $X$. The second element corresponds to the relationship status between a Third Party station and the tagged ICP candidate. This element can be either $F$ or $X$. Finally, the third element corresponds to the relationship status between the Third Party station and the unreachability cause. The third element can be $N$ (Non-related), $F$, $X$, or $\overline{X}$. Using these definitions, it is possible to determine whether the probability function of the tagged ICP candidate is ascending or descending with respect to the Third Party’s relationship with the unreachability cause. According to our investigations, there are nine possible scenarios in which the local probability function should be descending with respect to the Third Party’s relationship with the unreachability cause. These scenarios may be summarized as follow: $(X, \overline{X}, F), (F, \overline{X}, F), (X, F, F), (F, F, F), (X, \overline{X}, X), (F, \overline{X}, X), (X, F, X), (F, F, X)$, and $(F, X, F)$. In addition, there are four scenarios in which the local probability function is ascending with
respect to the abovementioned relationship, i.e., \((X, F, F), (F, F, N), (X, F, \overline{X})\), and \((F, F, \overline{X})\). Other configurations are considered as *don't cares*. In order to reduce the size of information transmitted by each station upon eMAC table broadcasting, all descending scenarios can be denoted by a single integer \(N_-\). In a similar manner, all ascending scenarios are denoted by a single integer \(N_+\). Now the local probability function may be defined as

\[
\beta = \rho \frac{2 - e^{-a_+ N_+}}{2 - e^{-a_- N_-}} (1 - \delta(N')) , 0 < \rho \leq 1.
\]

In the above equation, the \(a_+\) and \(a_-\) coefficients should be precisely tuned using the objective functions introduced earlier, while \(\rho\) is another tunable factor. Although parameter \(\rho\) increases the degree of freedom, it is assumed that it is always constant and is equal to 1 for all conducted simulations.

In single-hop systems, there is no hidden terminal and unreachability problem; therefore, there is no need to use ICP frame broadcasting or to broadcast local tables regularly. Due to the similarity of received tables, the time interval during which tables are broadcast is increased, resulting in considerable overhead reduction. In addition, due to recognition of the system as a single-hop network, the incurred overhead due to DS and ICP frames can be removed. This means that in such scenarios, the proposed scheme reduces to the IEEE 802.11 DCF.

### 3.1.2 Performance evaluation

#### Simulation environment

In order to evaluate the performance of the proposed protocol, extensive simulations for the system throughput, delay jitter, and complexity are provided to compare the achieved results to the IEEE 802.11 DCF [51], multiple access collision avoidance by invitation (MACA−BI) [10], dual busy tone multiple access (DBTMA) [16], advanced multiple access collision avoidance (AMACA) [97], and dual channel multiple access (DUCHA) [18]. The propagation model is the two-ray ground model [1 and 18], and the transmission range of each station is approximately 250m. The channel rate is set to 2 Mbps. In addition, the MAC frame payload is 1500 bytes, while the capture threshold is set to 10 dB.

In the simulation study, the following performance metrics are evaluated.
- Aggregate end-to-end throughput: The sum of data packets delivered to the destinations.
- Normalized control overhead: The ratio of all kinds of control packets to the throughput.
- Transmission efficiency of data packets: The ratio of the aggregate one-hop throughput to the number of the transmitted data packets. This metric reflects the resource wasted by the collided data packets and the discarded data packets due to the overflow of the queue at the intermediate nodes of the path [18].
- Average frame end-to-end delay: The end-to-end delay experienced by data packets.
- Average frame jitter: The frame jitter experienced by data packets.
- Protocol complexity: The average number of handled events per second in the MAC layer to accomplish basic layer functionalities.

**Multi-hop network with simple topology**

To verify the correctness of our protocol, a simple scenario shown in Fig. 36a is taken into consideration, where there is an unreachability problem if the IEEE 802.11 MAC protocol is used. In Fig. 36a, there are two traffic flows originating from station “A” with the same CBR/UDP rate: *Flow I* is from “A” to “B₁” and *Flow II* is from “A” to “B₃”, where “B₁” is the hidden terminal of “B₂” and “B₃”. Moreover, there is another traffic flow originating from station “C” to station “D”, which has a constant CBR/UDP traffic rate of 50 packets per second.
In this simulation study, the performance of eMAC is compared with that of AMACA [97] as well as the DUCHA protocol [18]. Fig. 36b shows the achieved throughput for station “A” with respect to its total offered load in Mbps. Here, the total offered load represents the whole input load due to stations “A”, which is equally distributed between Flow I and Flow II, and the vertical axis represents the one-hop throughput for station “A”. As can be observed in Fig. 36b, due to the so-called ICP broadcast storm AMACA shows the worst performance compared to eMAC and DUCHA. In fact, whenever “B_1” is going to become unreachable due to “C” to “D” transmission, as there is no rule in AMACA to prevent “B_1” and “B_2” from simultaneous ICP frame broadcast, both “B_1” and “B_2” begin transmitting their ICP frames at the same time, and as an apparent consequence they lead to an ICP collision at station “A”. Since the input load at station “A” is equally distributed between Flow I and Flow II, the accumulated packets at station “A’s” transmission queue are representing a random sequence of the packets addressed to “B_1”, and those packets that have been destined for “B_1”. Once “B_1” is unreachable and the head-of-line packet is addressed to this station,
“A” is aggressively attempting to accomplish the planned packet transmission to “B_i”; as a result, those packets that have been addressed to “B_j” and accumulated in the transmission buffer after the HOL packet have to wait until the HOL packet is either delivered successfully to “B_i” or dropped due to reaching the maximum retransmission retry limit. Although AMACA is equipped with the *unreachability vector* and *unreachability cache* [97], it is unable to provide enough bandwidth to *Flow II* due to the aforementioned drawback. In other words, *Flow II* will be starving because of a lack of bandwidth due to the fact that the blocked receiver associated with *Flow I* is unable to reliably inform station “A” about its unreachability status. As shown in Fig. 36b, eMAC outperforms DUCHA. In DUCHA, even though by separating control and data channels reporting an unreachability case becomes much more straightforward, station “A” is still required to contend for the wireless channel to send multiple RTS frames to “B_i” as it does not know exactly when the unreachable station “B_i” is going to be unblocked.

**Multi-hop network with random topology**

In this simulation study, 60 nodes are randomly placed in a 1000m × 300m area. The source of each traffic flow randomly selects one station as the destination, which is at least a certain minimum number of hops away, e.g., 3 hops. It is assumed that there are 20 traffic flows with the same CBR/UDP rate. Furthermore, a pre-computed shortest path with no routing overhead is used. All results are averaged over 50 random simulations.

Fig. 37 shows the aggregate end-to-end throughput in Mbps versus the total offered load. Among all considered schemes, IEEE 802.11 DCF shows the worst performance while DUCHA and the proposed eMAC scheme give the highest results when the offered load varies between 0.5 and 2.5 Mbps. Based on the achieved results, it can be observed that AMACA outperforms DBTMA, but, on the other hand, it is apparently unable to compete with DUCHA as well as eMAC due to the so-called ICP broadcast storm. It is interesting to note that the difference between the performance of IEEE 802.11 DCF and that of DBTMA is indeed considerable. The aggregate end-to-end throughput for all considered protocols decreases when the offered load increases due to the fact that the number of simultaneous contentions for acquiring the shared wireless medium increases.
Fig. 37. Aggregate end-to-end throughput (Mbps) vs. total offered load (Mbps).

Fig. 38 compares the transmission efficiency of DATA packets for different MAC schemes when the total offered load varies. Our protocol shows indeed better performance in comparison to the other schemes; in addition, it maintains relatively stable transmission efficiency of DATA packets for flows with several hop counts, while especially the IEEE 802.11 MAC degrades significantly when a higher amount of traffic is required to be delivered over multiple hops. The main reason for the better performance of eMAC in comparison to the other solutions, e.g., [18], is that in earlier solutions each station is still required to send an RTS frame to the destination of interest to figure out whether the destination is reachable. In [18], although by separating control and data channels reporting an unreachability case becomes easier, the source station is still needed to contend for the wireless channel to send an RTS to an unreachable destination. Hence, such requirements reduce the efficiency of the MAC protocol in spite of deploying multiple channels for different types of traffic.
Fig. 38. Transmission efficiency of DATA packets vs. total offered load (Mbps).

Fig. 39 illustrates normalized control overhead for four different MAC protocols versus the total offered load. The proposed eMAC shows indeed better results in comparison to the other schemes, especially IEEE 802.11 DCF and DBTMA, due to the fact that it avoids unsuccessful link layer connection establishment requests to the unreachable terminals. As can also be seen, the performance of DUCHA and eMAC are close while eMAC shows better performance for a higher amount of offered traffic to the network.
Fig. 40. Average end-to-end frame delay (msec) vs. total offered load (Mbps).

Fig. 41. Average frame jitter (msec) vs. total offered load (Mbps).

Fig. 40 shows the achieved average frame delay in milliseconds while Fig. 41 illustrates the obtained average frame jitter. As far as the average frame delay is concerned, the performance of DUCHA is slightly better than that of eMAC due to the fact that DUCHA uses dual-channel architecture for control and data frame transmissions. In fact, better performance for DUCHA protocol is achieved at the cost of the efficiency of multiple frequency band utilization, which is not always
desirable. The noticeable lower average frame delay for DUCHA and eMAC is due to the fact that in both protocols, link layer connection establishment with unreachable stations is almost avoided, which prevents stations from increasing their backoff contention window size unnecessarily. On the other hand, data transmission to the unreachable terminals can be simply postponed to conduct feasible transmissions to the reachable one-hop neighbors. It is also worth noting that although the average frame delay for DUCHA is lower than that of eMAC, the average frame jitter results of eMAC show better performance for our protocol compared to the DUCHA protocol.

![Graph of Protocol Complexity vs. Total Offered Load](image)

**Fig. 42. Protocol complexity (Events per second) vs. total offered load (Mbps).**

Fig. 42 illustrates the protocol complexity for different schemes versus the total offered load. This performance metric is evaluated in terms of the number of completed events per second for the MAC entity to accomplish its basic link layer functionalities. When this metric is higher, it means that the protocol needs more parallel events to attain a particular goal. Among all considered protocols, eMAC shows higher complexity due to the fact that it tries to solve the unreachability problem using only one channel in comparison to the DUCHA protocol, which uses dual-channel architecture to handle the receiver blocking problem. The higher complexity is because of the need for table construction, table maintenance, table update, and regular table broadcast. On the other hand, it is believed that the complexity of the routing protocol is reduced when eMAC is employed as the MAC protocol for multi-hop ad hoc networks. This is due to the fact that eMAC provides locally a part of the stations’ neighborhood topology information, which
can be easily used by the routing protocol in the network layer; therefore, the complexity of the routing protocol, in addition to the incurred route discovery and route maintenance overhead, will be reduced considerably. It is believed that the presented schemes in this section are not the ultimate solutions and can be improved further in future. For instance, in cases when the difference between the interference range and the communication area is relatively small, the final conclusion regarding the unreachability of Type II for neighboring stations may encounter overestimation. On the other hand, when the difference between the interference range and the communication area is relatively large, the proposed scheme almost leads to the desired conclusion. It is indeed worth evaluating the performance of our proposed scheme when wireless stations are in motion.

Fig. 43. Aggregate end-to-end throughput (Mbps) vs. total offered load (Mbps) for different mobility scenarios.

Fig. 43 shows the aggregate end-to-end throughput in Mbps versus the total offered load for different mobility scenarios. At first, all radio stations are considered to be fixed. In this case, the results are exactly the same as the achieved simulation results shown in Fig. 37. As the second case, it is assumed that wireless stations are in motion but the average speed of each radio station is relatively low. Finally, as the third scenario, it is assumed that wireless stations are in motion and the average speed of the mobile radios is relatively high. The mobility model considered for all simulations is almost similar to the so-called random way point model. In this model, each station randomly chooses a positive
real value from an interval $[0 \text{ sec}, 10 \text{ sec}]$ and stays in its current place until the chosen time interval elapses. Subsequently, the radio station randomly chooses a positive real value from an interval $[0 \text{m/s}, \nu \text{m/s}]$ as well as a geographical point in the simulation area. It then begins moving towards the selected point at the chosen speed. Once the station reaches the target point, it resumes the whole procedure. The average speed of each station is controlled by the parameter $\nu$. For the low speed scenario this parameter is set to $5 \text{ m/s}$, while it is set to $20 \text{ m/s}$ for the high speed scenario. As could be anticipated, when the average speed of radio terminals is increased, the achieved aggregate end-to-end throughput is decreased considerably while eMAC always outperforms DUCHA.

3.2 Chapter summary

The design of a single-channel MAC protocol to address the unreachability problem was considered in the first part of this chapter. The proposed scheme aims to solve the broadcast storm problem by introducing a flexible technique to prevent unnecessary simultaneous unreachability reports. Our protocol facilitates reporting of both the unreachability problem of Type I and Type II using an extra set of control frames as well as an out-of-band busy tone, without the requirement of separating control- and data traffic. Furthermore, an adaptive table broadcasting technique was proposed to facilitate topology information dissemination in mobile ad hoc networks.
4 Multi-channel contention resolution in cognitive wireless networks

Cognitive radio [102 and 103] is among the key solutions to the spectrum scarcity problem in wireless access networks. Cognitive radios are capable of occupying underutilized frequency bands with limited and controllable interference with licensed incumbent devices [104], also referred to as Primary Users (PUs). They scan potential frequency bands, find spectrum holes, and adjust their parameters, such as frequency, power, and transmission rate, in order to use the spectrum more efficiently [105].

One of the primary goals of cognitive radio technology is to enable Opportunistic Spectrum Access (OSA) [106]. With respect to this fundamental objective, the MAC sub-layer plays a key role in achieving efficient and reliable radio resource allocation, either in centralized or distributed cognitive radio networks [107]. So far, only a few cognitive MAC protocols have been proposed in the literature, and none of the existing solutions simultaneously addresses dynamic channel selection, multi-channel contention resolution among Secondary Users (SUs), and avoidance of primary users, e.g., [106 – 108]. Furthermore, the multi-channel hidden terminal [27] and multi-channel unreachability problems are among the most challenging issues in multi-channel ad hoc networks that must be taken into consideration when dealing with the problem of dynamic channel selection without any centralized coordination [109]. The former occurs when a device is operating on a channel and hence cannot listen to any other channel, which is also referred to as the deaf terminal problem [109]. The latter, which is also referred to as the channel conflict problem [109], occurs when a node unintentionally selects an already occupied channel for data transmission, resulting in frame collisions with neighboring data transmissions. Together, these two problems are referred to as the Multi-Channel Coordination (MCC) problem [109].

In this chapter, stochastic multi-channel contention resolution in a distributed cognitive network coexisting with primary users is investigated. In particular, a probabilistic technique for traffic distribution among a set of data channels by incorporating statistical information of primary users’ activities in different channels into the selection process without centralized coordination is proposed. It is shown through simulations that the proposed MAC layer enhancements outperform well-known multi-channel protocols both in the absence and presence of primary users. Furthermore, an analytical framework to study the performance
of the proposed schemes in a multi-channel cognitive network coexisting with primary users is presented.

The remainder of this chapter is organized as follows. The proposed stochastic channel selection and multi-channel contention resolution mechanism are presented in Section 4.1, and complementary discussions are given on how the suggested schemes can be further enhanced to address the so-called multi-channel unreachability problem. A performance analysis of the proposed schemes using a mathematical framework is presented in Section 4.2. Section 4.3 presents the simulation results, and, finally, concluding remarks are given in Section 4.4.

4.1 On stochastic channel selection for multi-channel cognitive wireless networks: a medium access control perspective

4.1.1 Overview and background

In the context of multi-channel MAC protocols, there are several contributions in the literature for conventional wireless ad hoc networks and frequency agile cognitive radio networks, tackling the problem of multi-channel radio resource management.

The Distributed Information Sharing (DISH) is introduced in [109] and applied to a multi-channel scenario to propose the Cooperative Asynchronous Multi-channel MAC (CAM-MAC) in order to tackle the MCC problem, where it is shown to outperform MMAC [29], SSCH [25], and AMCP [110]. In the case of MCC conflict, neighbors of the source and the destination stations send an Invalidation (INV) frame during the so-called Cooperative Collision Avoidance Period (CCAP) to nullify the channel negotiation handshake. This procedure is founded on a simple CSMA-based mechanism where each neighbor schedules to send an INV at a random point. The stations detecting INV frames earlier than their scheduled INVs simply abandon transmission of their INVs.

In [111], a dynamic decentralized multi-channel MAC protocol based on an extended version of slotted ALOHA is proposed, where channels are assigned to flows rather than MAC frames to eliminate collisions. In [112], a simple MAC scheme that can be used in ad hoc networks with asynchronous multi-channel environments is proposed. The Cooperative Diversity Multi-channel MAC (CD-MMAC) protocol is proposed in [113], which leverages the multi-rate capability
of IEEE 802.11b and allows wireless nodes far away from the destination node to transmit at a higher rate by using intermediate nodes as relays.

In [114 – 116] the decentralized cognitive radio-based multi-channel MAC protocols for dynamic access spectrum networks, where each secondary user is equipped with a cognitive radio-based transceiver and multiple channel sensors, is proposed. Moreover, by applying the $M/G/1$ queuing theory, an analytical model to analyze the performance of the proposed MAC protocol is presented. The work in [116] has been first extended in [117] and subsequently in [118] by improving the packet scheduling at the MAC layer. In [119], the *Opportunistic Multi-channel MAC (OMC-MAC)* protocol is proposed and analyzed, which achieves selection diversity gain in a distributed fashion. Following the same approach as proposed in [120], the size of the contention window is also adjusted adaptively based on the estimate of the number of competing stations, which is obtained via using a Sequential Monte Carlo technique. Furthermore, it achieves “resource pooling” and improves the stability of the network. In [121], a game theoretic DSA-driven MAC framework for cognitive radio networks is proposed to utilize the merits of game theoretic DSA techniques, including high spectrum utilization, collision-free channel access for resulting data communication, as well as QoS and fairness support.

In [122], a multi-channel MAC protocol for cognitive mesh networks is proposed. The protocol brings the main idea of cognitive wireless technology into wireless multi-hops and mesh networks. Every host learns the status of channels by listening to the common channel and chooses the best channel that causes the least interference to the rest of network upon its transmission. In [123], a distributed frequency agile MAC extension to the IEEE 802.11s for the next generation wireless mesh networks, taking into account backward compatibility with the legacy IEEE 802.11 and the emerging 802.11s, is proposed. In [124], the *Statistical Channel Allocation MAC (SCA-MAC)* for cognitive radio networks is introduced. This CSMA/CA-based protocol utilizes spectrum usage statistics for making decisions on conducting efficient channel acquisitions with less interference to the rest of the network. In [125], decentralized cognitive MAC protocols that allow secondary users to search independently for spectrum opportunities either without a central coordinator or a dedicated communication channel are proposed.

In [126] and [127], the *Hardware-Constrained Cognitive MAC (HCMAC)* is proposed to conduct efficient spectrum sensing and a spectrum access decision, taking into consideration the hardware constraints, such as the number of
available radios, partial spectrum sensing, and the spectrum aggregation limit. In [128], a slot-based MAC protocol for cognitive radio wireless networks is introduced. A group-based MAC protocol for QoS provisioning in cognitive radio networks is proposed in [129]. In this scheme, secondary users are divided into several non-overlapping groups, and all leftover channels are allocated among groups taking their bandwidth requirements into consideration. Moreover, the allocation of vacant channels can be adjusted dynamically when members join/leave groups or primary users vacate/redeploy the wireless network resources. In [130], an energy efficient distributed multi-channel MAC protocol for cognitive radio networks is presented. In particular, a single rendezvous-based scheme accompanied by sensing and recognition of primary users is developed. In [131], a channel-hopping-based cognitive radio MAC protocol for synchronized wireless networks, enabling the secondary users to utilize opportunistically the unused licensed-spectrum without interfering with the primary users, is studied. In [132], the Distance-Dependent MAC (DDMAC) for cognitive radio networks is proposed. This MAC scheme introduces a suboptimal probabilistic channel assignment algorithm that exploits the dependence between the signal’s attenuation model and the transmission distance while taking into consideration the communicating pairs’ traffic profile.

In [133], the Opportunistic Cognitive MAC (OC-MAC) for cognitive radios to access unoccupied resource opportunistically and coexist with Wireless Local Area Networks (WLANs) is proposed, where by a primary users’ traffic predication model as well as a transmission etiquette technique, fatal damage to licensed users is avoided. In [134], a class of CSMA-based MAC protocols for cognitive radio networks is proposed. Different from conventional schemes, these protocols with an adaptive transmission scheme allow transmissions for secondary users even when the primary users are actively transmitting. In [135], a MAC protocol for multi-channel cognitive radio networks, avoiding the need for a common control channel while solving the multi-channel hidden terminal problem using synchronization, is presented. In [136], based on the CSMA protocol the Concurrent Transmissions MAC (CT-MAC) to identify the possibility of establishing a second link in the presence of the first link in unlicensed bands is presented.

In this section, stochastic channel selection in cognitive wireless networks is studied, and, in particular, a distributed MAC layer resource allocation technique is proposed, which is enabled by a probabilistic channel selection algorithm as well as a multi-channel Binary Exponential Backoff (BEB) mechanism. Each data
channel is tagged by a probability value specifying the probability that a primary user will occupy that data channel during the transmission of an average sized MAC Protocol Data Unit (MPDU), using the lowest possible bit rate. Using these probabilities, a cognitive radio with a frame for transmission runs a probabilistic channel selection procedure to randomly choose one of the data channels. Each station maintains a counter for each data channel, which is randomly loaded with an integer upon initiation of a new channel selection procedure. All counters within a cognitive radio are driven by a common master clock. By each clock beat, the content of the counters is either decremented by one or remains unchanged. The content of a data channel’s counter remains unchanged with the probability of incumbent activity that is associated with the same data channel; otherwise, its content is decremented by one. A channel will be chosen as the possible candidate if its corresponding counter reaches zero before the counter of any other data channel. In this case, the counting down of other data channels’ counters is stopped, and their contents remain unchanged until re-initiation of the stochastic channel selection for the current frame transmission attempt becomes necessary.

The cognitive station also runs a backoff sequence to contend for a known control channel based on IEEE 802.11 DCF access rules, as well as carrier sensing on the selected data channel. Upon primary user appearance or another secondary user activity on the selected data channel while the backoff process on the control channel is progressing, the cognitive station resumes the probabilistic channel selection in order to randomly pick an alternate data channel for the scheduled frame transmission. In any of these cases, the counting down of all data channels’ counters is continued from the point where it was suspended during the previous channel selection attempt. This process is repeated until the backoff counter running on the control channel reaches zero. When that occurs, the currently chosen data channel is negotiated with the destination station.

### 4.1.2 Stochastic channel selection and contention resolution mechanisms

In this section, first the assumptions made throughout this section are discussed, and subsequently the proposed mechanisms are explained in detail.

It is assumed that there is a control channel that is known to all cognitive radios. This channel is used for control signaling and negotiation purposes. In addition, it may not be used for any data transactions. It is further assumed that each cognitive radio has two half-duplex transceivers; one is assumed to be
parked permanently on the control channel while the other one can be dynamically tuned to any communication channel. The transceiver operating on the control channel is used for control signaling and channel negotiation purposes, providing cognitive radios a reasonable picture of multi-channel activities taking place in their vicinity.

The set of all data channels \( \{C_1, C_2, ..., C_M\} \) is denoted by \( \mathcal{M} \), where \( C_i \) represents the \( i \)th data channel. Each data channel is tagged by a given cognitive radio with an estimated a priori probability \( p_i \) that channel \( i \) will be occupied by a primary user during the transmission of an average sized MPDU, plus the corresponding Acknowledgement (ACK) frame, at the lowest possible bit rate. Using probabilities \( p_i \), a cognitive radio runs a probabilistic channel selection mechanism to choose one of the available data channels randomly. It loads the designated counters of data channels with integers selected randomly from the set \( \{1, 2, ..., L_0\} \), where \( L_0 \) is a pre-defined window size. All counters within a cognitive station are driven by a common Master Clock (MCLK). By each clock beat of MCLK, the content of all counters is either decremented by one or remains unchanged. Upon a clock beat the content of data channel \( i \)’s counter is decremented by one with probability \( 1 - p_i \) and remains unchanged with probability \( p_i \), as illustrated in Fig. 44.

Fig. 44. Stochastic channel selection mechanism using a probabilistic reverse counter.

Data channel \( i \) will be chosen if its counter reaches zero before the counter of any other data channel. In such circumstances, the counting down of other data channels’ counters is stopped, and their contents remain unchanged until re-initiation of the probabilistic channel selection for the current frame transmission attempt becomes necessary.

Upon selecting a data channel as the first potential candidate, the cognitive radio initiates a backoff cycle on the control channel, as well as carrier sensing on the selected data channel. Upon primary user appearance or another secondary user activity on the selected data channel while the backoff process on the control
channel is in progress, the cognitive station should resume the stochastic channel selection in order to pick another data channel for the scheduled frame transmission randomly. In either of these cases, the counting down of all data channels’ counters is continued from the point where it has been suspended upon conclusion of the previous channel selection attempt.

Upon observation of a primary user in an already chosen data channel, the channel selection process is resumed excluding the channel in which the incumbent activity was detected. Furthermore, none of the counters for data channels in which a secondary user is currently operating is retrigged for counting down. This prevents the cognitive station from choosing channels that are currently occupied by other secondary users. If there are no data channels left for continuation of the ongoing backoff process on the control channel (i.e., all data channels have been marked as unavailable either due to secondary user or incumbent activity), the backoff sequence will be suspended until at least one data channel becomes available for possible use. As it is not expected that a primary user will likely vacate a data channel during a backoff sequence, we exclude all data channels utilized by incumbents for the entire backoff cycle.

Once secondary user activity is detected on the candidate data channel, the stochastic channel selection process is re-executed to randomly pick another data channel as the next candidate for the planned frame transmission. In addition, the counter associated with the occupied data channel is randomly loaded with an integer selected at random from a larger set \( \{1, 2, \ldots, L_1\} \) with \( L_1 = 2 L_0 \), but inhibited from counting down as long as the corresponding data channel is unavailable due to secondary user activity. So, after \( n \) unsuccessful candidacies of a data channel during the same backoff process, the window size of its pickup set is \( L_n = 2^n L_0 \).

The recursive stochastic channel selection process is repeated until the backoff counter associated with the control channel reaches zero. Subsequently, the currently chosen data channel is negotiated with the destination station using a two-way RTS/CTS handshake.

To illustrate the whole procedure, consider the scenario shown on the left hand side of Fig. 45. At \( t_0 \), the cognitive station randomly loads all data channels’ counters and commences counting down of all counters using its local MCLK, except the counter of data channel 1, as this channel is known to be currently occupied by another secondary user. It turns out that data channel 3 is the mutual candidate for the planned frame transmission, as its counter reaches zero before the counter of any other data channel. The cognitive radio concludes the channel
selection by stopping all counters from being counted down. Henceforth, the cognitive station keeps sensing data channel 3 and lets the backoff sequence on the control channel progress.

At $t_1$, data channel 3 is occupied by another secondary user and as a consequence, the cognitive radio re-executes the channel selection to choose another candidate. It restarts the counting down of all data channels’ counters from the point where they have been stopped upon conclusion of the previous channel selection attempt. It also reloads data channel 3’s counter randomly with an integer taken from the set $\{1, 2, \ldots, L_1\}$, but does not count down the counter. Furthermore, the counter of data channel 1 is inhibited from being counted down, as it is known to be still under utilization by a secondary user. This time, data channel 5 is chosen, as its counter reaches zero before the counter of any other data channel.

At $t_2$, data channel 5 is occupied by a primary user and, as a result, the cognitive radio re-executes the channel selection process to choose another candidate data channel. It restarts the counting down of all data channels’ counters, while excluding data channel 5 for the remainder of the process. As the cognitive station is aware of the availability of data channel 1 at $t_2$, it considers this channel for the current channel selection. In contrast, since it knows that data channel 3 is still under utilization by a secondary user, it does not count down its counter. Eventually, data channel 2 is selected as the next candidate due to the fact that its counter reaches zero before the counter of any other data channel.
At $t_3$, the backoff counter associated with the control channel reaches zero and the cognitive station sends an RTS frame indicating data channel 2 as the candidate for the planned frame transmission. Upon reception of the CTS frame carrying the destination station’s agreement the data frame is transmitted on data channel 2.

There are circumstances when more than one data channel’s counter reaches zero, as shown on the right hand side of Fig. 45. In such cases, one of the data channels for which the counter has reached zero is randomly chosen as the candidate for the scheduled frame transmission. For example, at $t_1$, the channel selection is concluded by data channels’ 2, 4, and 5 counters reaching zero simultaneously. The cognitive radio randomly picks data channel 5; however, at $t_2$ secondary user activity is detected on that channel. Consequently, the channel selection scheme is re-executed and data channel 5’s counter is reloaded with an integer value but inhibited from being counted down. Data channel 3’s counter is also kept suspended as it is still under utilization. It is interesting to note that data channel 1’s counter remains unchanged even though its medium is idle. This is due to the fact that there is at least one data channel for which the content of the counter is zero. Between data channels 2 and 4, the cognitive station randomly picks data channel 4 but since it is found busy, its counter is reloaded with a random integer and restrained from being counted down. Finally, data channel 2 is chosen and its medium is found idle. Henceforth, the backoff process is continued by considering that channel to be the possible candidate.

As a concluding remark, note that when the cognitive station has no a priori knowledge about $p_i$ probabilities, the stochastic channel selection mechanism presumes that the associated probabilities to data channels are equal to zero.

It is worth noting that in the above-considered scenarios it has been assumed that the cognitive radios’ knowledge about the status of all data channels was consistent. In fact, we implicitly assumed that the cognitive station was able to overhear all RTS/CTS frames exchanged by its neighbors to obtain an acceptable picture of all multi-channel activity in its vicinity. However, there are circumstances where the cognitive station is unable to receive the exchanged control frames on the control channel, leading to inconsistency of the collected information. As a result, misguided conclusions could be made by the cognitive radios when re-executing the stochastic channel selection. In what follows, the challenges encountered in practical scenarios and possible solutions to reduce the impact of such problems are explored.
4.1.3 Multi-channel unreachability problem

There are some scenarios where cognitive radios could make misguided conclusions when re-executing the probabilistic channel selection. The difference between the transmission and interference range is the most influential parameter on the efficiency of a MAC protocol. Next, this issue is carefully studied, and a solution to reduce its impact on the performance of the proposed scheme is introduced.

Fig. 46 illustrates a network setup in which station ‘A’ is situated in the interference range of station ‘C’. Assume that station ‘A’ wants to communicate with station ‘B’. Furthermore, consider the availability of five data channels. Fig. 47 illustrates a medium access attempt from station ‘A’, perspective.

![Fig. 46. A typical multi-hop network topology.](image)
Fig. 47. The impact of the difference between the transmission and interference range on stochastic channel selection.

At $t_1$, station ‘A’ chooses data channel 5; however, at $t_2$ and due to incumbent activity, the station re-executes the stochastic channel selection. Eventually, it turns out that data channel 4 is chosen as the candidate channel. Consider the case where station ‘C’ has initiated a frame transmission to station ‘D’ on data channel 4 just before time instant $t_2$. Station ‘A’ tunes its radio transceiver to data channel 4 and detects the interference due to station ‘C’ on that channel. In fact, as station ‘A’ is not in the transmission range of station ‘C’, it could not have overheard the RTS frame transmitted by station ‘C’. As an obvious consequence, station ‘A’ is unaware of the medium status on data channel 4 and makes a misguided selection when re-executing the channel selection process. To deal with this problem, station ‘A’ may follow these steps: 1) Randomly reload the counter of data channel 4 with an integer number taken from $\{1, 2, \ldots, L_1\}$, $L_1 = 2 \times L_0$; 2) Continue the current channel selection sequence for all available data channels; 3) Inhibit the counter of data channel 4 from being counted down for the current channel selection cycle.

In the aforementioned scenario, station ‘A’ has no way to determine for how long data channel 4 should not be involved in any stochastic channel selection. This ambiguity is due to a lack of knowledge about the duration of the ongoing
communication between station ‘C’ and station ‘D’. In fact, using a single wideband transceiver, the cognitive station may ultimately be able to perform carrier sensing on multiple channels simultaneously. Although this might be hard to accomplish in 802.11-based architectures due to relatively wideband individual channels, it is already possible in systems with narrower channels [141 – 143].

Another alternative approach to avoid the above problem is to associate a set of busy tones \( \{b_1, b_2, ..., b_M\} \) to the channel set \( \{C_1, C_2, ..., C_M\} \); the busy tone \( b_i \), which is a known sub-carrier to the entire network, is mapped to the data channel \( C_i \). When the data channel \( C_i \) is occupied, both the source and the destination stations activate the corresponding busy tone \( b_i \) for the entire of data transmission to inform all neighboring stations located in their interference range. By placing all busy tones close enough to each other in the frequency domain it becomes possible to use a single half-duplex transceiver to monitor the status of all available data channels simultaneously. This way, one transceiver is alternatively switched between the channels to conduct carrier sensing and an RTS/CTS handshake on the control channel as well as handle data transmission on any of the data channels, while the other transceiver is always parked on busy-tones to monitor constantly the status of all data channels simultaneously. This transceiver is also used to activate a particular busy-tone when the radio station begins utilizing the corresponding data channel. The main advantage of this approach is its simplicity as well as the possibility to eliminate the channel switching penalty (i.e., channel switching delay) during the backoff procedure; however, wastage of part of the network capacity is the major disadvantage of this method.

As pointed out earlier, the multi-channel unreachability problem is another challenging issue in multi-channel ad hoc networks [159 and 161]. To alleviate the impact of this problem, Fig. 48 suggests a localized solution driven by the source station, called Predictive Stochastic Channel Selection (PSCS).

In PSCS, the source station, in addition to the currently chosen data channel, selects a back-up channel as the second alternative to be provided to the destination station in case the first data channel cannot be deployed due to the unreachability problem. The approach through which the source station chooses the second alternative is straightforward: it further continues the stochastic channel selection to determine which one of the remaining data channels could be the next possible candidate if the current data channel cannot be utilized for any reason. In Fig. 48, at \( t_2 \) the backoff timer on the control channel reaches zero and the source station sends an RTS frame to the destination station, offering data
channel 4 as the candidate channel. It further continues the previous channel selection sequence to determine the next possible candidate. It turns out that by running the PSCS cycle, data channel 2 could be the second possible nominee. Consequently, data channels 4 (with higher priority) and 2 (with lower priority) are advertised to the destination station. By conducting carrier sensing, the destination station finds out that it will be unreachable on data channel 4, and, as a consequence, this channel may not be used for the planned frame reception. It continues the evaluation procedure with data channel 2 and finds this channel appropriate for communication with the source station. A longer inter-frame interval, called Switching & Carrier Sensing Inter Frame Space (SCIFS), is also assigned between the RTS and CTS frame to give the destination station enough time to conduct carrier sensing and any needed channel switching in case the station is not equipped with a wideband transceiver.
Fig. 48. The impact of the unreachability problem (*Left*: predictive stochastic channel selection (PSCS), *Right*: cooperative stochastic channel selection (CSCS)).
Besides PSCS, there is another possible approach to overcome the unreachability problem. Fig. 48 suggests an altruistic solution driven by both the source and the destination station, called Cooperative Stochastic Channel Selection (CSCS). In CSCS, the source station, in addition to advertising the chosen data channel, appends into the RTS frame the current content of all other data channels’ counter for which the medium is idle. The destination station, upon reception of the RTS frame, checks the availability of the advertised data channel. If the suggested data channel is available from the destination station’s point of view, it will be confirmed using the CTS frame; otherwise, the destination station executes a stochastic channel selection sequence using the provided counter values of the specified data channels in the RTS frame. In Fig. 48, the source station appends the content of data channels’ 1, 2, and 5 counters into the RTS frame. The destination station, first, verifies the availability of data channel 4. It turns out that this channel is unavailable due to another secondary user hidden to the source station. Consequently, the destination station runs the stochastic channel selection using the provided counter values of data channels 1, 2, and 5. Eventually, data channel 2 is chosen as the candidate and reported back to the source station. The selection of channel is approved by the source station using another RTS frame, as shown in Fig. 48.

4.2 Performance analysis of saturated multi-channel MAC

In this section, an analytical framework to study the performance of the proposed schemes, both in the absence and presence of primary users, is introduced. First, all data channels are assumed to be free of primary users, i.e., $p_z = 0, \forall C_z \in M$, where $M$ denotes the set of all $M$ data channels $\{C_1, C_2, ..., C_M\}$. Subsequently, the possibility of the incumbents appearing on the available data channels at any given time is incorporated into the framework.

4.2.1 Overview and background

So far, only a few contributions have been reported in the literature to analyze the performance of multi-channel contention-based MAC protocols, either in the absence or presence of primary users.

To analyze the performance of multi-channel contention-based MAC protocols, either in the absence or presence of primary users, numerous contributions have been reported in the literature. In [137], a simple scheme to
describe channel switching by extending IEEE 802.11 MAC for multi-channel networks is proposed. Furthermore, to evaluate the performance of the channel switching scheme a three-dimensional Markov model is presented. In this model, a radio station is allowed to switch dynamically to different channels depending on their traffic condition. When it has a frame to send, it chooses the data channel which has the highest priority (i.e., home channel) in its Channel Priority List (CPL). When the station fails to transmit successfully on this channel, it switches to the most preferable channel in its CPL. The frame is either sent successfully or discarded after trying out all the channels. Moreover, in this model the entire system is considered to have no control channel. In [138], the frameworks presented in [52] and [139] are revised and extended to provide a two-dimensional Markovian analytical model which allows computing an approximation of the MAC layer service time in a multi-channel contention resolution scheme. In [140], a comparative study on the latency performance of DSA protocols based on a dedicated control channel and an embedded dedicated control channel is presented.

In contrast to the existing literature (e.g., [52] and its posteriors, considering a saturated single-channel wireless network), the proposed model in this section takes into account the multi-channel contention resolution and possible secondary users’ communication interruption due to the incumbent appearance. Furthermore, the simplifying assumption of the availability of a CPL in each radio station is not considered. Besides, the scenario where each radio considers the available data channels and selects one of them based on the proposed stochastic channel selection mechanism is assumed.

A single-hop wireless network consisting of $n$ pairs of radio stations, where each pair comprises a transmitter peer as well as a receiver peer, is considered. Each pair of transmitter-receiver entities is referred to as a communication pair. It is assumed that all source stations are operating in saturation mode, i.e., there is always at least one MAC frame in each source station’s transmission queue. A tagged communication pair is randomly chosen, and its backoff process along with the frame transmission on one of the data channels is modeled by a multi-dimensional Markov chain. In contrast to [52], where the control and data frames are exchanged on the same channel, and, as a result, the backoff states can solely represent the whole system behavior, here it is necessary to incorporate the so-called frame delivery states as by the end of each frame transmission, the tagged communication pair should identify whether there is any available data channel, and if not, how long it takes for the occupied channels to become vacant. This
way, it becomes possible to determine how many data channels are going to be occupied, either by the primary users or secondary users, while the tagged communication pair is conducting its data frame transmission on a particular data channel.

In what follows, it is shown that the approach through which a tagged communication pair’s backoff and frame delivery procedures are modeled in a multi-channel network with no primary user differs from the scenario where a multi-channel cognitive network with primary users associated with existing frequency opportunities is considered.

4.2.2 Multi-channel MAC without primary user consideration

In this sub-section, no primary user associated with any of the data channels is considered, i.e., $p_z = 0, \forall C_z \in \mathcal{M}$. The investigation is begun by considering backoff and frame delivery states, and introducing the concept of channel classification based on the remaining busy period of data channels. Next, possible state transitions in the proposed Markov model are discussed. As the last part of this sub-section, statistical characterization of the service time in a saturated multi-channel network is studied.

A) Channel classification

In a multi-channel wireless network, data channels may have different remaining busy periods at any given time. In fact, data channels with similar remaining busy periods may be classified into the same groups; however, as it is assumed that all data channels are identical in terms of physical specifications, it is wiser to keep track only of the number of data channels with similar remaining busy periods. As will be shown, using either classification methods the same analysis results are obtained, though, the latter leads to a much easier approach to analyze the stochastic behavior of the entire system.

As pointed out earlier, a tagged communication pair is chosen and its backoff process along with the frame transmission on one of the data channels is modeled by a multi-dimensional Markov chain. For the tagged communication pair’s backoff procedure, along with the stochastic processes $s(t)$ and $b(t)$ representing the backoff stage and step, respectively \cite{52}, a set of processes $S_z(t), \forall z \in \{0,1,2, ..., t_{\text{max}}\}$ are defined, where $S_z(t)$ represents the number of data channels with $z$ remaining busy time slots at the given time $t$ and $t_{\text{max}}$ denotes the
maximum possible number of remaining busy time slots in a data channel. Assuming the size of a time slot to be $\sigma$, $t_{\text{max}}$ represents the number of time slots that a data transmission may last on any data channel, i.e., $t_{\text{max}} = \frac{(T_{\text{DATA}} + t_{\text{SFSTime}} + T_{\text{ACK}})}{\sigma}$ . The stationary probability distribution of each backoff state is given by:

$$\lim_{t \to \infty} P[s(t) = i, b(t) = k, S_0(t) = s_0, S_1(t) = s_1, \ldots, S_{t_{\text{max}}}(t) = s_{t_{\text{max}}}] = \theta_{i,k}(s_0, s_1, \ldots, s_{t_{\text{max}}})$$  \hspace{1cm} (41)$$

Similar to the backoff procedure, a set of processes $S_z(t), \forall z \in \{0,1,2, \ldots, t_{\text{max}}\}$ is employed for the tagged communication pair’s frame delivery phase. During this phase, it is also necessary to keep track of the remaining busy time slots in the data channel that is being occupied by the tagged communication pair. This is accomplished by introducing another stochastic process $G(t)$. The stationary probability distribution of each frame delivery state is given by:

$$\lim_{t \to \infty} P[G(t) = g, S_0(t) = s_0, S_1(t) = s_1, \ldots, S_{t_{\text{max}}}(t) = s_{t_{\text{max}}}] = \theta_s(g, s_0, s_1, \ldots, s_{t_{\text{max}}})$$  \hspace{1cm} (42)$$

In addition to $(s_0, s_1, \ldots, s_{t_{\text{max}}})$, representing the number of data channels with all possible remaining busy periods ranging from 0 to $t_{\text{max}}$, the vector $(b_1, b_2, \ldots, b_M)$ is defined to denote the remaining busy period of all data channels $\{C_1, C_2, \ldots, C_M\}$, where $b_z (\in \{0,1, \ldots, t_{\text{max}}\})$ denotes channel $C_z$’s remaining busy period in the number of time slots. Throughout this section, “the remaining busy time slots” and “the remaining busy period” are used interchangeably. In addition, either form is referred to as the multi-channel system channel status.

Note that in the case of primary user absence, there is no need to keep track of the control channel busy status. In fact, as long as the control channel is busy either due to an RTS/CTS handshake or a frame collision, none of the communication pairs is allowed to begin a new RTS/CTS handshake on the control channel to initiate a channel negotiation. Depending on the length of an ongoing event on the control channel, either part of or the whole remaining busy period of the occupied data channels is elapsed, while the unoccupied data channels remain unutilized for the entire duration of that particular event. By the end of an RTS/CTS handshake, which is assumed to take $\Delta_s = \frac{(t_{\text{SFSTime}} + T_{\text{RTS}} + t_{\text{SFSTime}} + T_{\text{CTS}})}{\sigma}$ slots to complete, one of the idle data channels is occupied, resulting in its remaining busy period
becoming \( K_e - \Delta_s = (T_{DATA} + tSIFSTime + T_{ACK})/\sigma = t_{\text{max}} \). In general, by the end of an RTS/CTS handshake, the remaining busy period of a data channel, which is presumed to be \( e \) slots, will be \( \max(0, e - \Delta_s) \). Similarly, by the end of a frame collision, which is assumed to last for \( K_c = (T_{RTS} + tEIFSTime)/\sigma \) slots, the remaining busy period of a data channel, which is presumed to be \( e \) slots, will be max(0, e - K_c).

To grasp a better perception, let us assume that \( K_e = 10, K_c = 4, \text{ and } \Delta_s = 3 \) slots (i.e., \( t_{\text{max}} = 7 \)). Suppose when the tagged communication pair is operating at the \( i \)th backoff stage and \( k \)th backoff step, the channel status (i.e., remaining busy period) of a multi-channel system with 5 orthogonal data channels is as follows: data channel \( C_1 \) with 0, data channel \( C_2 \) with 5, data channel \( C_3 \) with 1, data channel \( C_4 \) with 0, and data channel \( C_5 \) with 3 remaining busy time slots, i.e., \( \langle \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \rangle = \langle 0,5,1,0,3 \rangle \). This means that \( \delta_0 = 2, \delta_1 = 1, \delta_2 = 0, \delta_3 = 1, \delta_4 = 0, \delta_5 = 1, \delta_6 = 0, \delta_7 = 0 \). Hence, the tagged communication pair’s current backoff state can be identified as \( \theta_{i,k}(2,1,0,1,0,1,0,0) \). Note that in every multi-channel network and at any backoff or frame delivery state, it is clear that \( \sum_{h=0}^{t_{\text{max}}} \delta_h = M \).

Now consider the case when a successful RTS/CTS handshake is conducted on the control channel and the data channel \( C_1 \) is selected for carrying out the data frame transmission. By the end of the RTS/CTS exchange, the channel status of the multi-channel system will become \( \langle \theta_1, \theta_2, \theta_3, \theta_4, \theta_5 \rangle = \langle 7,2,0,0,0 \rangle \). This means that the backoff state to which the tagged communication pair transits will be \( \theta_{i,k-1}(3,0,1,0,0,0,0,1) \). In other words, not only the backoff stage and step, but also the channel status of the entire wireless network \( \langle \delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7 \rangle \) is updated. On the other hand, if a frame collision is considered on the control channel, the channel status of the system will change from \( \langle 0,5,1,0,3 \rangle \) to \( \langle 0,1,0,0,0 \rangle \). This means that the backoff state to which the tagged communication pair transits is going to be \( \theta_{i,k-1}(4,1,0,0,0,0,0,0) \). If nothing happens during a time slot on the control channel, the eventual channel status of the system will change from \( \langle 0,5,1,0,3 \rangle \) to \( \langle 0,4,0,0,2 \rangle \). This means that the backoff state to which the tagged communication pair transits will be \( \theta_{i,k-1}(3,0,1,0,1,0,0,0) \).

Once the tagged communication pair’s backoff counter reaches zero (i.e., backoff state \( \theta_0 \)), it will select one of the available idle data channels and proceed with the frame delivery state \( \theta_s \) if we have \( \delta_0 \neq 0 \). During the frame delivery phase, the tagged communication pair conducts a data frame transmission on the
chosen data channel while keeping on tracking the channel status of the entire wireless network.

To summarize, it turns out that for a simple multi-channel MAC protocol there is no need to define an extra stochastic process to maintain the control channel’s latest timing update; however, in the subsequent sub-section it will be shown that for a cognitive multi-channel MAC it is necessary to keep track of the control channel’s timing (i.e., remaining busy period) due to the fact that primary users may occupy any of the data channels at any given time, regardless of the status of the control channel.

**B) Markov model state transition**

In this sub-section, possible state transitions within and between backoff and frame delivery procedures are considered and a systematic approach to represent them is introduced. Each state transition has a starting point, i.e., the state where it begins, as well as an ending point, i.e., the state where it terminates. The entire wireless network channel status at the starting point of a state transition is denoted by \( (\delta_0, \delta_1, \delta_2, \ldots, \delta_{t_{\text{max}}}) \), and assuming the tagged communication pair transits from \( \theta_{i',k'} \) to \( \theta_{i,k} \), the backoff state at the starting point of state transition is denoted by \( \theta_{i',k'}(\delta_0, \delta_1, \delta_2, \ldots, \delta_{t_{\text{max}}}) \), while the backoff state at the ending point of state transition is denoted by \( \theta_{i,k}(\delta_0, \delta_1, \delta_2, \ldots, \delta_{t_{\text{max}}}) \). Note that in every multi-channel network with no primary user and for any state transition, it is clear that \( \sum_{h=0}^{t_{\text{max}}} \delta_h = \sum_{h=0}^{t_{\text{max}}} s_h = M \).

Before discussing possible state transitions in the Markov model, first, it is necessary to define a set of probabilities that will be used for calculation of state transitions’ probability. The probability that all data channels are occupied and the tagged communication pair is blocked as a result of lack of radio resources (i.e., blocking probability) is given as follows:

\[
\mathcal{P}_{\text{block}} = \sum_{i=0}^{m} \sum_{k=0}^{W_{i-1}} \sum_{(s_1, \ldots, s_{t_{\text{max}}})} \theta_{i,k}(0, s_1, \ldots, s_{t_{\text{max}}}).
\]  

(43)

In (43), the interior summation is taken over all possible combinations for \( (s_1, \ldots, s_{t_{\text{max}}}) \), while the condition \( \sum_{h=1}^{t_{\text{max}}} s_h = M \) (i.e., \( s_0 = 0 \)) indicates that there is no unoccupied data channel left for the contending communication pairs (i.e., blocking probability). The conditional transmission probability that a communication pair transmits in a randomly chosen slot, given the channel status
of the entire network $\mathcal{S} = (s_0, s_1, ..., s_{t_{\text{max}}})$, symbolically represented by $[s_h]_{h=0}^{t_{\text{max}}}$, is expressed as follows:

$$
\tau([s_h]_{h=0}^{t_{\text{max}}}) = \begin{cases} 
\frac{1}{\pi(s)} \sum_{i=0}^{m} \theta_{i,0}(s_0, s_1, ..., s_{t_{\text{max}}}) & \text{for } s_0 \neq 0 \\
0 & \text{for } s_0 = 0
\end{cases}. \quad (44)
$$

The probability that a communication pair is utilizing one of the data channels, given the channel status $(s_0, s_1, ..., s_{t_{\text{max}}})$, is obtained as follows:

$$
\eta([s_h]_{h=0}^{t_{\text{max}}}) = \frac{1}{\pi(s)} \sum_{g=1}^{t_{\text{max}}} \theta_{g}(g, s_0, s_1, ..., s_{t_{\text{max}}}). \quad (45)
$$

The conditional deferral probability that a communication pair due to an ongoing backoff process with the channel status $(s_0, s_1, ..., s_{t_{\text{max}}})$ is given as follows:

$$
\bar{\tau}([s_h]_{h=0}^{t_{\text{max}}}) = \begin{cases} 
\frac{1}{\pi(s)} \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} \theta_{i,k}(s_0, s_1, ..., s_{t_{\text{max}}}) & \text{for } s_0 \neq 0 \\
\frac{1}{\pi(s)} \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} \theta_{i,k}(s_0, s_1, ..., s_{t_{\text{max}}}) & \text{for } s_0 = 0
\end{cases}. \quad (46)
$$

where $\pi(s) \triangleq \sum_{h=0}^{t_{\text{max}}} \theta_{h}(g, s) + \sum_{i=0}^{m} \sum_{k=0}^{W_i-1} \theta_{i,k}(s)$.

Using the aforementioned probabilities, the following conditional probabilities are defined that are widely used to express the state transitions’ probability. The probability that none of the contending communication pairs transmit on the control channel during a randomly chosen slot, given the channel status $\mathcal{S} = [s_h]_{h=0}^{t_{\text{max}}}$, is derived as follows:

$$
\mathcal{P}_n(n, [s_h]_{h=0}^{t_{\text{max}}}) = \left( \frac{\bar{\tau}(s)}{\tau(s) + \bar{\tau}(s)} \right)^n \cdot \frac{\tau(s)}{\tau(s) + \bar{\tau}(s)}. \quad (47)
$$

Furthermore, the probability that only one of the contending communication pairs transmits on the control channel during a randomly chosen slot, given the channel status $\mathcal{S} = [s_h]_{h=0}^{t_{\text{max}}}$, is expressed by:

$$
\mathcal{P}_1(n, [s_h]_{h=0}^{t_{\text{max}}}) = \left( n - \sum_{h=1}^{t_{\text{max}}} s_h \right) \cdot \frac{\tau(s)}{\tau(s) + \bar{\tau}(s)} \cdot \frac{\bar{\tau}(s)}{\tau(s) + \bar{\tau}(s)}. \quad (48)
$$
The probability of a collision on the control channel among the contending communication pairs, given the channel status of the multi-channel system, is obtained as follows:

\[ P_{\text{coll}}(n, [s_h]_{h=0}^{t_{\text{max}}}) = 1 - P_{\text{tr}}(n, [s_h]_{h=0}^{t_{\text{max}}}) - \mathcal{P}_s(n, [s_h]_{h=0}^{t_{\text{max}}}). \] (49)

Finally, the probability that at least one of the contending communication pairs transmits on the control channel during a randomly chosen slot with the channel status \([s_0, s_1, ..., s_{t_{\text{max}}}]\) is given as follows:

\[ P_{\text{tr}}(n, [s_h]_{h=0}^{t_{\text{max}}}) = 1 - P_{\text{tr}}(n, [s_h]_{h=0}^{t_{\text{max}}}). \] (50)

In Appendix 2, possible state transitions in the Markov model, in addition to their corresponding transition probabilities, are presented.

**A) Statistical characterization of service time**

In this sub-section, using the proposed Markov model and by following the same approach as [144], a closed-form **Probability Generating Function (PGF)** for the MAC frame service time of a saturated multi-channel wireless network is derived. The obtained PGF can be inverted by numerical methods, providing the **Probability Distribution Function (PDF)** of the service time.

To begin the investigation, first the probability of observing an empty slot on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking, i.e., at least one data channel is available to be utilized by the contending communication pairs, is derived. This probability, denoted by \(p_{L\text{-nb}}\), is given as follows:

\[ p_{L\text{-nb}} = \sum_{\mathcal{S}} \pi(\mathcal{S}) \cdot P_{\text{tr}}(n - 1, \mathcal{S}). \] (51)

Next, the probability of observing a successful channel acquisition on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking, is derived. This probability, denoted by \(p_{S\text{-nb}}\), is given as follows:

\[ p_{S\text{-nb}} = \sum_{\mathcal{S}} \pi(\mathcal{S}) \cdot P_s(n - 1, \mathcal{S}). \] (52)
Similarly, the probability of observing a frame collision on the control channel, given that the tagged communication pair is pursuing its backoff process while there is no blocking, denoted by $p_{c,nb}$, is derived as follows:

$$p_{c,nb} = \sum_{s} \pi(s) \cdot \mathcal{P}_{coll}(n-1,s).$$

(53)

The probability of observing an empty slot on the control channel, given that the tagged communication pair is pursuing its backoff process and all contending communication pairs are blocked due to a lack of radio resources while the minimum remaining busy period in the entire system is $b$, denoted by $p_{l,b}$, is given as follows:

$$p_{l,b} = \sum_{s : b = \min(n \neq 0|s_b \neq 0)} \pi(s) \cdot \mathcal{P}_{nt}(n-1,s).$$

(54)

The service time $d$ is defined as the time taken by the MAC layer control entity to successfully deliver a MAC frame over the radio channel [144], and is given as follows:

$$d = T_s + \sum_{j=0}^{n_c} T_{c,j} + T_B(n_c),$$

(55)

where $T_s$ and $T_{c,j}$ denote the random variables representing the time taken by a successful frame transmission, including an RTS/CTS handshake on the control channel, a DATA/ACK exchange on a randomly chosen data channel, and an RTS frame collision on the control channel, respectively. Moreover, $n_c$ denotes the random variable representing the number of collisions the MAC frame undergoes before being successfully received by the destination station, taking $T_{c,j}$ at each retransmission attempt. The random variable $T_B(n_c)$ is the contribution to the delay due to the backoff procedure. As the successful transmission time depends on the length of the MAC frame, it is assumed that the successful transmission times are independent identically distributed random variables with probability mass function $p_T(k)$. As in a multi-channel network RTS/CTS handshake is always used prior to any actual data transmission, collisions can only occur on RTS frames, so that the collision time $T_{c,j}$ is always constant.

Under stationary assumptions, it is assumed that the probability for a transmitted frame to collide at a given time slot is a constant $p$, independent of the transmission history [52] [144]. The probability of collision given that the tagged
communication pair begins transmitting an RTS frame on the control channel is given as follows:

\[ p = \sum_{x, x_0 > 0} \pi(s). \left(1 - P_{\text{ntr}}(n - 1, s)\right). \]  \hspace{1cm} (56)

Similar to [144], it is assumed that the number of collisions \( n_c \) a MAC frame undergoes before successful reception is a geometric random variable with probability mass function \( p_{n_c}(r) = (1 - p)p^r \). The moment generating function for the service time \( d \) is then given by:

\[
G_d(x) = G_T(x). \sum_{r=0}^{\infty} (1 - p)p^r . E \left[Z_{\Sigma j=0}^{\sum \tau_c} | n_c = r \right]. E \left[Z^\tau_c | n_c = r \right]
= (1 - p). G_T(x). \sum_{r=0}^{\infty} p^r . E \left[Z^\tau_c | n_c = r \right]. E \left[Z^\tau_c | n_c = r \right]
= (1 - p). G_T(x). \sum_{r=0}^{\infty} \left( \epsilon . G_T(x) \right)^r . E \left[Z^\tau_c | n_c = r \right]. \]  \hspace{1cm} (57)

where \( G_T(x) = E[Z^X] \) and \( G_T_c(z) = z^T_c \) with \( T_c = K_c \times \sigma \). The overall backoff time, given that the MAC frame experiences \( r \) collisions before success, denoted by \( T_B(r) \), is given as follows:

\[ T_B(r) = \sum_{i=0}^{r} \sum_{j=0}^{x_i} t_{i,j} \]  \hspace{1cm} (58)

where \( t_{i,j} \) is the random variable representing the duration of a slot time, defined as the time between two successive decrements of the backoff counter [52]. Moreover, \( x_i \) is the random variable representing the number of slot times spent in backoff stage \( s(i) \), where the backoff stage \( s(i) \) after \( i \) collisions is defined as follows [144]:

\[ s(i) = \begin{cases} i, & \text{if } i < m; \\ m, & \text{if } i \geq m; \end{cases} \]

Taking into account that random variables \( \{t_{i,j}\} \) are independent, the moment generating function for the backoff time \( T_B(r) \) is given by:

\[
G_{T_B}(z) = E \left[Z_{\Sigma j=0}^{\epsilon_0, j + \Sigma j=0}^{\tau_x, 1, j} + \ldots + \Sigma j=0}^{\tau_y, r, j} \right] = \prod_{i=0}^{r} G_{\epsilon_i}(z). G_{\tau_x}(z) \ldots G_{\tau_y}(z). \]  \hspace{1cm} (59)
where $G_{y_1}(z)$ can be expressed as:

$$G_{y_1}(z) = E \left[ z^{\sum_{j=0}^{\infty} \tau_{i,j}} \right] = E \left[ z^{\tau_{i,0} + \sum_{j=1}^{\infty} \tau_{i,j}} \right] = E \left[ z^{\sum_{j=1}^{\infty} \tau_{i,j}} \right] = G_{s_1}(G_{e}(z)). \quad (60)$$

In (60), $\tau_{i,0} = 0$ for any $i$ as when the backoff counter is loaded with 0, the station immediately begins transmitting an RTS frame on the radio channel. Subsequently, the following expression for $G_{r_0}(z)$ can be derived:

$$G_{r_0}(z) = \prod_{i=0}^{r} G_{s_1}(G_{e}(z)). \quad (61)$$

During a slot time, four different events are possible: an idle slot time due to transmission deferral of all contending communication pairs, a busy period due to a successful RTS/CTS handshake, a busy period due to a frame collision, and an idle slot time due to blocking of all radio resources. Hence, the following expression for $G_{e}(z)$ can be perceived:

$$G_{e}(z) = p_{l;nb} z^\sigma + p_{S;nb} z^{A_1} + p_{C;nb} z^{K_1} + \sum_{b=1}^{t_{max}} p_{l;b} z^{B_{\sigma}}, \quad (62)$$

where $p_{l;nb}, p_{S;nb}, p_{C;nb}$, and $p_{l;b}$ are given by (51) to (54), respectively. Furthermore, the probability generating function of random variable $x_i$ is given by [144]:

$$G_{x_i}(z) = \sum_{k=0}^{W_{i}-1} \frac{z^k}{W_i} = \frac{1 - z^{W_{i}2^{s(i)}}}{(1 - z) W_0 2^{s(i)}}, \quad (63)$$

where $W_i$ is the contention window size in backoff stage $s(i)$ and is defined as $W_i = W_0 \times 2^{s(i)}$.

4.2.3 Multi-channel Cognitive MAC with Primary User Consideration

In this sub-section, the possibility of the incumbents to occupy any of $M$ data channels irrespective of whether they are being utilized by the secondary users is incorporated into the analytical framework proposed in the previous section. For a multi-channel cognitive network, when specifying the status of a data channel, it is imperative to determine not only the channel remaining busy period, but also the type of radio station, either primary or secondary, utilizing the wireless medium. In what follows, first the Markov model state space is introduced, and,
subsequently, possible state transitions within the proposed framework are
discussed.

A) Markov model state space

For a tagged communication pair’s backoff procedure, along with the stochastic
processes \( s(t) \) and \( b(t) \) representing the backoff stage and backoff step, the
processes \( S_z(t), \forall z \in \{0,1,2,\ldots,t_{\text{max}}\} \) are defined, where \( S_0(t) \) corresponds to
the number of idle data channels, while process \( S_z(t), z \neq 0 \) corresponds to the
number of data channels with \( z \) remaining busy time slots occupied by the
secondary users. The maximum number of remaining busy time slots in a data
channel due to the communication pair’s transmission is \( t_{\text{max}}^l = K_z \). The
processes \( U_z(t), \forall z \in \{1,2,\ldots,T_{\text{max}}\} \) are also defined, where \( U_z(t) \) represents the
number of data channels with \( z \) remaining busy time slots occupied by the
primary users at the given time \( t \). Moreover, \( T_{\text{max}} \) denotes the maximum possible
number of remaining busy time slots in a data channel due to an incumbent
transmission. Finally, the process \( E(t) \) is defined to represent the remaining busy
period of the control channel, ranging from 0 to \( \Delta_e = (T_{\text{DIFS}} + \tau_{\text{RTS}} + t_{\text{SIFS}} + T_{\text{CTS}})/\sigma \) in the case of a successful
RTS/CTS handshake or from 0 to \( K_e = (T_{\text{RTS}} + T_{\text{EIFS}})/\sigma \) in the case of a
frame collision.

The stationary probability distribution of each backoff state is given by:

\[
\theta_{l,k}(e, s_0, s_1, s_2, \ldots, s_{t_{\text{max}}^l}, u_1, u_2, \ldots, u_{t_{\text{max}}^r}) = \lim_{t \to \infty} \mathbb{P}\{ s(t) = i, b(t) = k, E(t) = e, S_0(t) = s_0, S_1(t) = s_1, \ldots, S_{t_{\text{max}}^l}(t) = s_{t_{\text{max}}^l}, U_1(t) = u_1, U_2(t) = u_2, \ldots, U_{t_{\text{max}}^r}(t) = u_{t_{\text{max}}^r} \}. \tag{64}
\]

Similarly, the stationary probability distribution of each frame delivery state is
given by:

\[
\theta_{e} (g, e, s_0, s_1, s_2, \ldots, s_{t_{\text{max}}^r}, u_1, u_2, \ldots, u_{t_{\text{max}}^r}) = \lim_{t \to \infty} \mathbb{P}\{ s(t) = i, b(t) = k, G(t) = g, E(t) = e, S_0(t) = s_0, S_1(t) = s_1, \ldots, S_{t_{\text{max}}^r}(t) = s_{t_{\text{max}}^r}, U_1(t) = u_1, U_2(t) = u_2, \ldots, U_{t_{\text{max}}^r}(t) = u_{t_{\text{max}}^r} \}. \tag{65}
\]
A) Markov model state transitions

When a cognitive network is assumed, \((\bar{e}, \delta_0, \delta_1, \delta_2, ..., \delta_{t_{\text{max}}}^e, u_1, u_2, ..., u_{t_{\text{max}}}^e)\) denotes the entire wireless network channel status at the starting point of a state transition, where \(\delta_0, \delta_1, \delta_2, ..., \delta_{t_{\text{max}}}^e\) represent the number of data channels with all possible remaining busy periods ranging from 0 to \(t^e_{\text{max}}\) that are either idle (i.e., \(\delta_0\)) or occupied by the secondary users (i.e., \(\delta_e\)). Moreover, \(u_1, u_2, ..., u_{t_{\text{max}}}^e\) represent the number of data channels with all possible remaining busy periods ranging from 1 to \(T_{\text{max}}\) that are occupied by the primary users. Note that in every multi-channel cognitive network and for any state transition, it is clear that
\[
\sum \delta_k + \sum u_k = M.
\]

Before discussing possible state transitions in the Markov model, first, it is necessary to define a set of probabilities that will be used for calculation of state transitions’ probability. The probability that all data channels are occupied and the tagged communication pair is blocked as a result of a lack of radio resources (i.e., blocking probability) is given as follows:

\[
P_{\text{block}}(e) = \sum_{l=0}^{m} \sum_{k=0}^{W_l-1} \sum_{\delta_0, \delta_1, \delta_2, ..., \delta_{t_{\text{max}}}^e, u_1, u_2, ..., u_{t_{\text{max}}}^e} \theta_{l,k}(e, 0, \delta_1, \delta_2, ..., \delta_{t_{\text{max}}}^e, u_1, u_2, ..., u_{t_{\text{max}}}^e). \tag{66}
\]

The conditional transmission probability that a communication pair transmits in a randomly chosen slot, given the channel status of the entire network \(e, s = (\delta_0, \delta_1, ..., \delta_{t_{\text{max}}}^e)\), represented by \((e, s, [u_l]_{l=0}^{t_{\text{max}}}^e)\), is derived as follows:

\[
\tau(e, [u_l]_{l=0}^{t_{\text{max}}}^e) = \begin{cases} 
\frac{1}{\pi(s, u)} \sum_{l=0}^{m} \theta_{l,0}(0, s, u) & \text{for } e = 0 \land s_0 \neq 0, \\
0 & \text{for } e \neq 0 \land s_0 = 0.
\end{cases} \tag{67}
\]

The probability that a communication pair is utilizing one of the data channels, given the channel status of the multi-channel network, is expressed as follows:

\[
\eta(e, [s_h]_{h=0}^{t_{\text{max}}^e}, [u_h]_{h=0}^{T_{\text{max}}}) = \frac{1}{\pi(s, u)} \sum_{g=1}^{t_{\text{max}}^e} \theta_{s}(g, e, s, u). \tag{68}
\]
The probability that a communication pair is in transmission deferral due to an ongoing backoff process, given the channel status of the multi-channel system, is obtained as follows:

\[
\tilde{r}(e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) = \\
\begin{cases} \\
\frac{1}{\pi(s,u)} \sum_{i=0}^{m} \sum_{k=1}^{W_i - 1} \theta_{i,k}(e,s,u) & \text{for } e = 0 \land s_0 \neq 0 \\
\frac{1}{\pi(s,u)} \sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} \theta_{i,k}(e,s,u) & \text{for } e \neq 0 \land s_0 = 0 \\
\end{cases}
\] (69)

where \(\pi(s,u) = \sum_{i=0}^{t_{\text{max}}} \theta_i(s,u) + \sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} \theta_{i,k}(s,u)\). In addition, the conditional probabilities \(P_{\text{nt}}\), \(P_{\text{tr}}\), \(P_{\text{coll}}\), and \(P_{\text{tr}}\) can be derived as follows:

\[
P_{\text{nt}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) = \left(\frac{\tilde{r}(s)}{\tau(s) + \tilde{r}(s)}\right)^{n - \sum_{h=1}^{t_{\text{max}}} s_h},
\] (70)

\[
P_{\text{tr}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) = \\
\left(1 - \sum_{h=1}^{t_{\text{max}}} \delta_h\right) \cdot \frac{\tau(s)}{\tau(s) + \tilde{r}(s)} \left(\frac{\tilde{r}(s)}{\tau(s) + \tilde{r}(s)}\right)^{n - \sum_{h=1}^{t_{\text{max}}} s_h - 1}.
\] (71)

\[
P_{\text{coll}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) = \\
1 - P_{\text{nt}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) - P_{\text{tr}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}).
\] (72)

\[
P_{\text{tr}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}) = 1 - P_{\text{nt}}(n,e,[s_h]_{h=0}^{t_{\text{max}}},[u_h]_{h=0}^{t_{\text{max}}}).
\] (73)

In Appendix 2, possible state transitions in the Markov model, in addition to their corresponding transition probabilities, are presented.

A) Modeling a stochastic channel selection mechanism

To incorporate a stochastic channel selection mechanism into the Markov process, it is necessary to derive the probability that a particular counter reaches zero before any other counter given the initial contents of all counters corresponding to unoccupied data channels. To achieve this goal, first a particular counter, e.g.,
counter $z$, loaded with the initial value $a_z$ is considered. The probability that this counter reaches zero exactly at $b$ master clock beats is given by:

$$q^\xi_z(a_z) = \binom{b-1}{a_z-1} (1 - p_z)^{a_z} p_z^{b-a_z},$$  \hspace{1cm} (74)$$

where $p_z$ represents the incumbent appearance probability on data channel $C_z$. Furthermore, the probability that the aforementioned counter does not reach zero by $b$ master clock beats is given by:

$$\bar{q}^\xi_z(a_z) = \sum_{l=0}^{a_z-1} \binom{b}{l} (1 - p_z)^l p_z^{b-l}. \hspace{1cm} (75)$$

Using $q^\xi_z(a_z)$ and $\bar{q}^\xi_z(a_z)$, the probability that counter $z$ wins the stochastic channel selection exactly at master clock beats given the initial contents of all counters corresponding to unoccupied data channels is given by:

$$Q^\xi([a_j]) = q^\xi_1(a_1) \cdot q^\xi_2(a_2) \cdots q^\xi_2(a_2) = q^\xi_z(a_z). \prod_{i=1}^{a_z} \bar{q}^\xi_z(a_i). \hspace{1cm} (76)$$

where $a = [a_1 a_2 ...]$ is the vector of the initial values of unoccupied data channels’ counters. Finally, the probability that counter $z$ wins the stochastic channel selection given the initial contents of all counters corresponding to unoccupied data channels is given by:

$$p^\xi_z([a_j]) = \sum_{l=1}^{\infty} Q^\xi([a_j]) \cdot \left( 1 - \sum_{m=1}^{l-1} Q^\xi([a_j]) \right). \hspace{1cm} (77)$$

Considering the fact that the initial content of unoccupied data channels’ counters is uniformly drawn from $[1, 2, ..., L_0]$, the probability that channel $C_{g'}$ is chosen among all unutilized channels when the stochastic channel selection is executed, is given by:

$$p^g_{cs}([C_j]) = \sum_{[a_z]} \sum_{[a_x] \in [1, L_0]} \frac{1}{L_0} p^g_{cs}([a_z]), \hspace{1cm} (78)$$

where the summation is taken over all possible combinations of the initial values $[a_z]$ ranging from 1 to $L_0$ for each data channel. Note that in the previous sub-sections we implicitly supposed that the probabilities of primary user
appearance on different data channels are the same as we explicitly assumed that the data channels are all identical. In contrast, here we derived the probability that a particular data channel’s counter wins the channel selection process, given that the probabilities of primary user appearance on different data channels are not the same.

4.3 Performance evaluation

To evaluate the proposed schemes, the stochastic channel selection and multi-channel contention resolution mechanisms have been coded on top of existing IEEE 802.11 models. In this simulation study, the terrain is a $1000 \times 1000 \, m^2$ area and 150 radio stations form 150 non-disjoint flows randomly (i.e., each node is the source of one flow and the destination of another flow) [27] [25]. Each node has a single data packet queue (instead of per-neighbor queues, such as used by [27] and [24], which bypass head-of-line (HOL) blocking and yield higher throughput and lower delay) [25]. For all simulations and similar to [18], a pre-computed shortest path with no routing overhead is used.

In this simulation study, the following performance metrics are evaluated.

- Aggregate end-to-end throughput: The total end-to-end throughput of the entire network.
- Aggregate frame end-to-end delay: The end-to-end delay experienced by data frames.
- Data channel conflict rate: The frame collisions on data channels per second for all radios [25].
- Packet delivery ratio: The number of data frames successfully received by destinations normalized by the number of data frames sent by source stations [25].

There is one control channel and data channels, each of which has a bandwidth of 1 Mbps. The transmission range of each station is 250m and the capture threshold is set to 10 dB. The physical (PHY) and MAC layers’ parameters, i.e., PLCP, SIFS, and retry limit, are the same as in IEEE 802.11. Furthermore, the beacon interval is set to 100 msec. Each source generates data frames with 2,048 Bytes payload according to the Poisson point process unless otherwise mentioned. Each simulation is terminated when a total of 500,000 data frames are sent over the network. All results are averaged over 30 randomly generated networks.
4.3.1 Multi-channel MAC without primary user consideration

First, the case when our proposed scheme is deployed as a simple multi-channel MAC protocol is taken into consideration. In this scenario, there is no primary user, and therefore all probabilities associated with the data channels are set to zero. The performance of the proposed protocol is compared with MMAC [29] and DCA [39]. For MMAC, the specified values in [29] of 20 $msec$ for the ATIM window and 80 $msec$ for the rest of the beacon interval are used.

In Fig. 49, the aggregate end-to-end throughput for multi-channel stochastic channel selection, abbreviated by SCS, is compared with that of DCA [39] and MMAC [29], when there are 6 available data channels with no associated primary users. As the traffic generation rate per flow is increased, the aggregate throughput is increased for all three multi-channel schemes. By an approaching traffic generation rate per flow of 9 $Kbps$, MMAC enters saturation. In contrast, saturation is approached for DCA and SCS approximately at 12 $Kbps$ and 13 $Kbps$, respectively. Indeed, this can be simply anticipated by noting the fact that MMAC has a relatively shorter control phase (i.e., ATIM window) available for channel negotiation purposes. In contrast, for DCA and SCS the control channel is readily available for such purpose at any time. The higher aggregate end-to-end throughput of SCS is due to its smarter approach to accessing existing data channels. In the saturated control channel based on SCS, more successful channel negotiations in the form of RTS/CTS frame exchange can be accommodated, leading to more concurrent data transmissions and, consequently, higher aggregate end-to-end throughput. A station adopting SCS will not start exchanging the RTS/CTS with the destination station unless it is assured that the intended data channel is free of any secondary and primary users’ activity. By further increasing the total offered load, all three protocols begin experiencing performance degradation due to a significant increase in the number of collisions among contending stations.
Fig. 49. Aggregate end-to-end throughput (Mbps) vs. traffic rate per flow (Kbps): No primary user, 6 data channels.

Fig. 50 shows the aggregate throughput for SCS, DCA, and MMAC when there are 13 data channels with no associated primary users. It is worth noting that the traffic generation rate per flow at which SCS enters saturation is now shifted to 20 Kbps, while the saturation is approached for MMAC and DCA approximately at 15 Kbps and 18 Kbps, respectively. Furthermore, the maximum achievable aggregate end-to-end throughput for SCS is now at 3.15 Mbps (compared to the maximum aggregate throughput of 2.6 Mbps when there were 6 data channels with no associated primary users).
Figs. 51 and 52 show the average end-to-end frame delay for SCS, DCA, and MMAC when there are 6 and 13 available data channels with no associated primary users. Among these three schemes, SCS shows lower average frame delay for both scenarios. This is due to the fact that in SCS, before commencing the RTS/CTS negotiation on the control channel, a radio station is free to select another candidate when the currently chosen data channel is utilized either by another secondary user or an incumbent; thus, the ongoing backoff sequence on the control channel is not suspended. In both DCA and MMAC, channel negotiation on the control channel and the connection establishment on the agreed data channel are conducted independently; hence, they require carrying out a separate backoff cycle on the data channel to avoid possible frame collision due to hidden terminals, leading to larger average access delay compared to the proposed SCS scheme.
Fig. 51. Average frame end-to-end delay (msec) vs. traffic rate per flow (Kbps): No primary user, 6 data channels.

Fig. 52. Average frame end-to-end delay (msec) vs. traffic rate per flow (Kbps): No primary user, 13 data channels.
4.3.2 Cognitive multi-channel MAC with primary user consideration

In the second set of simulations, a cognitive network coexisting with primary users is considered. To take into account the possibility of primary users’ operation in each data channel, an incumbent is associated with each data channel. Furthermore, in the simulation environment time is assumed to be slotted for each data channel, and the length of each slot is set to 1 msec. When data channel \( i \) is free of the associated primary user, it may only be occupied by the incumbent at the beginning of each time slot with probability \( p_i \). Once it has entered the data channel, the primary user continues utilizing the data channel for a random period with uniform distribution defined on the interval \([5\text{msec}, 20\text{msec}]\). In all simulations, it is assumed that cognitive stations have access to all data channels’ status; therefore, ideal primary user detection is always assumed. In addition, it is assumed that there are 10 data channels with associated primary user arrival probabilities of 0.1, 0.1, 0.2, 0.2, 0.4, 0.4, 0.6, 0.6, 0.8, and 0.8, respectively. In addition, the channel status of each frequency opportunity is assumed to be available to all cognitive stations.

In Figs. 53 and 54, the performance of SCS is compared to two other techniques, when there are ten data channels, each of which has an associated probability representing the primary user appearance on that channel. Here, two heuristic schemes, namely the Best Next Channel (BNC) and random-based channel selection, are considered. According to BNC, when the current candidate channel is occupied by a primary user, the cognitive radio tries to choose an available data channel with the lowest associated probability. Apparently, this approach leads to the worst performance as it increases the level of congestion on data channels with relatively low probabilities of primary user appearance. In contrast, in the random-based approach the cognitive radio tries to choose a data channel randomly among all the available channels. As shown in Fig. 53, there is a considerable difference between the performance of SCS and these two other channel selection schemes. The worst performance of BNC is due to the fact that it carries out channel selection in the greediest possible way. Furthermore, though the random-based channel selection is less greedy compared to the BNC, it does not take into account the probability of primary user appearance on the frequency opportunities when conducting load balancing among available data channels. In addition, even though there is a significant difference between the average end-to-end frame delay of SCS without coexistence with the primary users and that of
the SCS coexisting with the primary users, SCS still outperforms both BNC and random-based channel selection schemes.

Fig. 53. Aggregate frame end-to-end throughput (Mbps) vs. traffic rate per flow (Kbps): With primary user, 10 data channels.

Fig. 54. Average frame end-to-end delay (msec) vs. traffic rate per flow (Kbps): With primary user, 10 data channels.
Fig. 55 illustrates the data channel conflict rate versus the traffic generation rate per flow for the SCS, BNC, and random-based channel selection mechanism. As could be predicted, SCS shows the lowest conflict rate, while BNC demonstrates the worst performance. The conflict rate for the proposed SCS begins from 50 at a 2.5 Kbps traffic generation rate per flow and continues increasing up to almost 400 at the traffic rate of 50 Kbps traffic rate per flow. In fact, the lower data channel conflict rate of SCS gives an insightful observation on why SCS achieves higher aggregate throughput, as shown in Fig. 53. Encountering less data channel conflict, SCS conducts smarter channel allocation among contending radio stations, leading to higher aggregate throughput and lower average end-to-end frame delay.

![Graph showing data channel conflict rate vs. traffic rate per flow for SCS, BNC, and random-based channel selection.](image)

Fig. 55. Data channel conflict rate vs. traffic rate per flow (Kbps): With primary user, 10 data channels.

Fig. 56 shows the packet delivery ratio versus the traffic generation rate per flow for SCS, BNC, and random-based channel selection schemes. As the proposed SCS encounters a smaller data channel conflict rate than the two other mechanisms, it gains significant performance for a wide range of incoming traffic generation per flow. In fact, the difference between the performance of SCS and that of the other two schemes is considerable within the range of an 8 to 24 Kbps traffic rate per flow.
Fig. 56. Packet delivery ratio vs. traffic rate per flow (Kbps): With primary user, 10 data channels.

Fig. 57 illustrates the data channel conflict rate versus the data payload size in bytes for SCS, BNC, and random-based channel selection schemes. Interestingly, for all protocols the channel conflict rate exhibits a bell shape, similar to the phenomenon that is reported and discussed in [109]; a larger payload size offsets protocol overhead more effectively and, thus, leads to higher throughput, while longer data packets keep nodes on data channels longer, and, hence, fewer nodes are able to initiate new communication on the control channel, which reduces the possibility of channel conflicts.
Fig. 57. Data channel conflict rate vs. data payload size (Bytes): With primary user, 10 data channels.

Fig. 58 demonstrates the impact of per data channel a priori probability estimation error on the multi-channel MAC protocol, overall performance. In this figure, for example, 20% error refers to the case where the estimated probability for data channel $i$ is a uniform random value taken from the interval $[\max(0, p_i - \frac{1}{5}p_{ave}), \min(1, p_i + \frac{1}{5}p_{ave})]$, and in all simulations $p_{ave} = 0.42$. As it can be observed from Fig. 58, by increasing the estimation error the performance of our proposed scheme degrades noticeably. Interestingly, the performance degradation for the case of 20% estimation error is almost twice of that of the 10% case. Similarly, performance degradation for the case of 50% estimation error is roughly twice of that of the 20% case.
Fig. 58. Aggregate throughput (Mbps) for different estimation error rates in a priori probabilities $p_i$ vs. traffic rate per flow (Kbps): With primary user, 10 data channels.

4.3.3 Analytical model validation

In this sub-section, the proposed analytical framework is validated by comparing the numerical results obtained through the coded mathematical model with the simulation results. This sub-section is begun by introducing a performance metric similar to that of [52], while taking into account the channel status of the entire network as well as primary user appearance on data channels.

First, the case of a simple multi-channel MAC without primary user appearance on any data channel is considered. The normalized channel utilization $\mathcal{R}(n)$ of a data channel is defined as the ratio of channel timing used for successful data frame transmission to the total channel timing consumed for all possible incidents that may occur during a slot time. The channel timing consumed for a successful data communication can be derived as follows:

$$\mathcal{N}(n) \triangleq \sum_{\Delta = (\sigma_0, \ldots, \sigma_{t_{\text{max}}})}^{\Delta_{\text{t_{\text{max}}}} \geq 1} \sigma_{t_{\text{max}}} p_s(n, \Delta) p_{e_{t_{\text{max}}}}(\sigma_0) \quad (79)$$
where the above summation is taken over all possible combinations of channel status for the multi-channel system. Furthermore, the total channel timing consumed for all events that may occur during a slot time can be expressed as follows:

$$\mathcal{D}(n) = \sum_{s = (s_0, \ldots, s_{\text{max}})} \pi(s) \left( \mathcal{P}_s(n, s) (1 - p_{cs}(s_0)) \right) \sigma \Delta_s +$$

$$+ \mathcal{P}_s(n, s) p_{cs}(s_0) \sigma K_s + \mathcal{P}_{\text{rtr}}(n, s) \sigma + \mathcal{P}_{\text{coll}}(n, s) \sigma K_c. \quad (80)$$

In (80), the first term in summation represents the incident when a successful RTS/CTS handshake is conducted on the control channel. This RTS/CTS handshake is not intended for the tagged data channel, denoted by the probability \((1 - p_{cs}(s_0))\). In fact, \(p_{cs}(s_0)\) denotes the probability that the tagged data channel, which is presumed to be part of \(s_0\), is selected by the MAC layer built-in channel selection mechanism. Since throughout such a handshake no medium access is allowed for the rest of contending communication pairs, the channel timing for the tagged data channel is wasted for \(\sigma \Delta_s\). The second term represents the case when a successful RTS/CTS handshake is conducted on the control channel that is particularly intended for the tagged data channel. Obviously, this case results in a channel deployment, which lasts for a period of \(\sigma K_s\). The third term refers to the event corresponding to an idle channel timing, which lasts for \(\sigma\). Finally, the last term corresponds to a collision on the control channel, wasting channel timing of the tagged channel for \(\sigma K_c\). The normalized channel utilization is then defined as:

$$\mathcal{R}(n) = \frac{\mathcal{N}(n)}{\mathcal{D}(n)}. \quad (81)$$

When a cognitive multi-channel MAC is considered, both \(\mathcal{N}(n)\) and \(\mathcal{D}(n)\) should be revised accordingly. The channel timing consumed for a successful data communication is given as follows:

$$\mathcal{N}(n) = \sum_{(s, u)} \pi(s, u) \mathcal{P}_s(n, 0, s, u) p_{cs}(s_0) \sigma K_s (1 - p_{inc}) K_s. \quad (82)$$
where \((1 - p_{inc})^n\) denotes the probability that no incumbent appears during the data transmission on the tagged data channel. The total channel timing consumed for all events that may occur during a slot time is given as follows:

\[
\mathcal{D}(n) \triangleq \sum_{(s,u)} \pi(s,u) \left( \mathcal{P}_s(n,0,s,u) \cdot (1 - p_{cs}(s_0)) \cdot \sigma \cdot \Delta_s + \mathcal{P}_s(n,0,s,u) \cdot p_{cs}(s_0) \cdot \sigma \cdot K_c \cdot (1 - p_{inc}) + \mathcal{P}_{ntr}(n,0,s,u) \cdot \sigma \cdot K_c \cdot (1 - p_{inc}) + \mathcal{P}_{call}(n,0,s,u) \cdot p_{cs}(s_0) \cdot p_{inc} \cdot \sum_{z=0}^{K_c-1} \left( (T_{max} + \sigma \cdot z) \cdot (1 - p_{inc})^z \right) \right) + \mathcal{P}_{call}(n,0,s,u) \cdot (1 - p_{cs}(s_0)) \cdot \sigma \cdot K_c + \mathcal{P}_s(n,0,s,u) \cdot p_{cs}(s_0) \cdot p_{inc} \cdot \sum_{z=0}^{K_c-1} \left( (T_{max} + \sigma \cdot z) \cdot (1 - p_{inc})^z \right) \right). \tag{83}
\]

In (83), the first term in summation represents the incident when a successful RTS/CTS handshake is conducted on the control channel. This RTS/CTS handshake is not intended for the tagged data channel, denoted by the probability \((1 - p_{cs}(s_0))\). If the primary user interferes with this handshake, the channel timing is wasted for \(\sigma \cdot \Delta_s\). The second term in (83) is exactly the same as the term in (82). The third term refers to an event corresponding to an idle time slot with no interference due to incumbent, which lasts for \(\sigma\). The fourth term refers to an idle time slot where no secondary user performs medium access but, a primary user enters the tagged data channel, resulting in channel timing wastage of \(T_{max}\).

The fifth and sixth terms represent a frame collision on the control channel when the tagged data channel is chosen by at least one of the colliding communication pairs. In this case, for the incident to last for only \(\sigma \cdot K_c\), no primary user should enter the tagged channel throughout the frame collision (i.e., fifth term). In contrast, depending on the time instance incumbent occupies the channel when the primary user enters the tagged data channel during frame collision, the channel timing is determined. In other words, if with the probability of \(p_{inc} \cdot (1 - p_{inc})^z\) the incumbent enters the tagged data channel after \(\sigma \cdot z\) time slots, the channel timing of the tagged data channel will be \((T_{max} + \sigma \cdot z)\).

The seventh term corresponds to the case when the tagged data channel is chosen by none of the colliding radios. In this case, whether the primary user
enters the chosen channels or not, the channel timing of the tagged data channel is consumed for \( \sigma \cdot K_c \).

Finally, the last term represents the case when an incumbent enters the tagged data channel while it is chosen by a communication pair that has conducted a successful RTS/CTS handshake on the control channel. When the incumbent enters the tagged data channel after \( t \) time slots with the probability of \( p_{inc} \), the channel timing of the tagged data channel becomes \( T_{max} + \sigma \cdot z \).

Fig. 59 illustrates per data channel normalized utilization factor \( \mathcal{R}(n) \) versus the number of contending communication pairs \( n \), and for different probabilities of incumbent appearance \( p_{inc} \). By increasing the number of radio stations, the normalized channel utilization is decreased; similar behavior is also observed in [52], and is well investigated in [152]. Furthermore, as it could be predicted, by increasing \( p_{inc} \), the normalized channel utilization \( \mathcal{R}(n) \) is decreased considerably. Interestingly, the performance degradation for the case of \( p_{inc} = 0.2 \) is almost twice of that of \( p_{inc} = 0.1 \). Similarly, the performance degradation for the case of \( p_{inc} = 0.5 \) is roughly twice and a half of that of the \( p_{inc} = 0.2 \) case. Fig. 59 also compares the simulation results with the obtained numerical results based on the proposed analytical framework. For a larger number of contending communication pairs the difference between numerical and simulation results is decreased roughly down to zero when \( n \geq 45 \). For instance, the lowest amount of difference for the case of \( p_{inc} = 0.5 \) is observed when \( n = 50 \), though, it should be noted that the processing time to achieve numerical results with satisfactory precision is increased significantly when \( p_{inc} \neq 0 \) and \( n \) is larger than 25. It is interesting to note that normalized channel utilization exhibits a bell shape for \( p_{inc} = 0 \) and \( p_{inc} = 0.1 \), though, this behavior does not exist in other scenarios. In addition, the difference between simulation and numerical results around the peak value of \( \mathcal{R}(n) \) for \( p_{inc} = 0 \) and \( p_{inc} = 0.1 \) is larger than that of \( n < 15 \).
4.4 Chapter summary

In this chapter a stochastic channel selection and load balancing mechanism for distributed multi-channel cognitive networks was introduced. The proposed scheme incorporates the probability of incumbent user appearance in each frequency opportunity obtained through distributed measurements into the channel selection process, resulting in smarter load balancing, higher achievable channel throughput, and lower access delay.

To continue the work presented in this chapter, it would be interesting to incorporate smarter algorithms based on advanced machine learning [145] into the proposed multi-channel MAC enhancements presented in this chapter. In addition, distributed learning and reasoning algorithms [146 and 147] can be applied in the proposed cognitive MAC protocol to enhance the performance of the entire wireless network.
5 Conclusions and future work

In this thesis, the problem of medium access control in multi-hop ad hoc networks was considered. An analytical framework, through which a closer observation of the problem of medium access control can be achieved, was proposed. The majority of performance evaluations for the contention-based IEEE 802.11 DCF that have been previously reported in the literature are based on saturation analysis, while only a few models under a finite load condition have adopted queuing systems with general service time distributions. However, such a queuing model only considers the first moment of frame service time to derive the probability of the transmission queue being vacant. The IEEE 802.11 DCF was modeled using the so-called Parallel Space-Time Markov chain (i.e., PSTMC), in which frame arrivals were tracked by monitoring the transmission queue during transitions between successive states of the space-time Markov chain. The proposed framework has been shown to provide the possibility of modeling the contention phase, the backoff and post-backoff procedures, and the transmission queue status at the same time. The suitability of the proposed model became more obvious when applied to multi-hop scenarios; it has been shown that the PSTMC framework has the potential of being utilized for performance analysis of multi-hop networks with any arbitrary network topology. With the help of a multi-hop extension of the aforementioned framework, the hidden terminal and unreachability phenomena could be easily formulated. In this way, graph partitioning has been employed in order to classify wireless stations into distinct membership groups, and, subsequently, in each group PSTMC was applied to model medium access behavior of associated stations. Normalized throughput, MAC frame end-to-end delay, and frame jitter were carefully studied throughout Chapter 2 and the proposed models were validated by comparing numerical results to the simulations results obtained from OPNET™.

As a possible future research direction, the PSTMC model could be extended so as to model a multi-channel 802.11-based MAC. How to extend the framework to model a multi-channel contention resolution scheme is quite an interesting issue that may lead to better understanding of the challenges in multi-channel medium access control. As an obvious consequence, this will definitely lead us to propose more robust and more flexible multi-channel contention resolution techniques. Another interesting topic for future work is to extend the multi-hop PSTMC framework to consider the case where the interference range is different from the radio transmission range: How should this difference be incorporated
into the framework and how does it affect the complexity of the whole system? Moreover, it is indeed interesting to combine the multi-channel model and multi-hop PSTMC framework to obtain one of the most comprehensive tools in analyzing non-saturated multi-hop ad hoc networks. Finally, it is important to consider node mobility in a PSTMC mathematical framework. One approach to take to consider the mobility of radio stations is to propose a supplementary queuing system in parallel to the PSTMC model to keep track of the number of stations associating with each sub-cluster.

As pointed out earlier, the above systematic approach assisted us in grasping thorough understanding on the nature of the unreachability problem, which was the main issue of concern throughout the thesis. In Chapter 3, the eMAC protocol was proposed; a single-channel MAC protocol for multi-hop ad hoc networks. Under eMAC, stations maintain double-hop neighborhood graphs while exchanging designated eMAC tables to share their knowledge about their neighborhood topology. Furthermore, an adaptive table broadcasting technique was proposed to facilitate topology information dissemination in mobile ad hoc networks. Performance of the proposed schemes was carefully evaluated and compared with earlier schemes through simulations. Our results show performance enhancement due to better handling of unreachability, possible heterogeneous power distributions among contending stations, and mobility issues. According to the simulation results, it can be observed that eMAC outperforms some of the well-known MAC protocols such as DUCHA, DBTMA, AMACA, MACA-BI, and IEEE 802.11 DCF. Aggregate end-to-end throughput, end-to-end frame delay, frame jitter, transmission efficiency, and protocol complexity were the key performance evaluation metrics taken into consideration. Based on the obtained numerical results, eMAC maintains relatively stable transmission efficiency of DATA packets for flows with several hop counts, while the IEEE 802.11 MAC in particular degrades significantly when a higher amount of traffic is required to be delivered over multiple hops. The main reason for better performance of eMAC in comparison to the other solutions is that in earlier solutions each station was still required to send an RTS frame to the destination of interest to figure out whether the destination is reachable. In DUCHA, although by separating control and data channels reporting an unreachability case becomes easier, the source station is still needed to contend for the wireless channel to send an RTS to an unreachable destination. Hence, such requirements reduce the efficiency of the MAC protocol in spite of deploying multiple channels for different types of traffic. On the other hand, as far as average frame delay is
concerned, the performance of DUCHA is slightly better than that of eMAC due to the fact that DUCHA uses dual-channel architecture for control and data frame transmissions. In fact, better performance for DUCHA protocol is achieved at the cost of efficiency for multiple frequency bands’ utilization, which is not always desirable. The noticeable lower average frame delay for DUCHA and eMAC is due to the fact that in both protocols, link layer connection establishment with unreachable stations is almost avoided, which prevents stations from increasing their backoff contention window size unnecessarily. Moreover, data transmission to the unreachable terminals can be simply postponed to conduct feasible transmissions to the reachable one-hop neighbors. It is also worth noting that although the average frame delay for DUCHA is lower than that of eMAC, the average frame jitter results of eMAC show better performance for our protocol compared to the DUCHA protocol. Among all considered protocols in Chapter 3, eMAC was shown to have higher complexity due to the fact that it tries to solve the unreachability problem using only one channel in comparison to the DUCHA protocol, which uses dual-channel architecture to handle the receiver blocking problem. The higher complexity is because of the need for table construction, table maintenance, table update, and regular table broadcast. On the other hand, it is believed that the complexity of the routing protocol could have been reduced if eMAC had been employed as the MAC protocol for multi-hop ad hoc networks. This is due to the fact that eMAC provides locally part of the stations’ neighborhood topology information, which can be easily used by the routing protocol at the network layer; therefore, the complexity of the routing protocol, in addition to the incurred route discovery and route maintenance overhead, will be reduced considerably.

The presented schemes in this thesis are not the ultimate solutions and can be improved further in future. For instance, in cases when the difference between the interference range and the communication area is relatively small, the final conclusion regarding unreachability of Type II for neighboring stations may encounter overestimation. On the other hand, when the difference between the interference range and the communication area is relatively large, the proposed scheme almost leads to the desired conclusion. Furthermore, the proposed eMAC table broadcasting mechanism can be improved by incorporating more flexible and robust techniques to handle the scenarios where there are large numbers of mobile stations. In such cases, the proposed schemes in this thesis may not work properly, leading to rather a high level of overhead compared to simpler protocols, e.g., DUCHA. In this respect, DUCHA seems to be more favorable as it does not
need to be accompanied by any table broadcasting mechanism. However, 
DUCHA still suffers from the same problem as pointed out earlier: Any source 
station still needs to contend for the wireless channel to send an RTS to an 
unreachable destination. For highly dynamic systems, extensive and more 
detailed simulations are required to be conducted in order to observe all the pros 
and cons of eMAC and other well-known solutions. Finally, it is obvious that 
more advanced optimization techniques can be employed to improve the 
performance of ICP frame broadcasting. How to combine topology information 
and optimization techniques to achieve better performance in ICP frame 
broadcasting is indeed an interesting topic for future investigations. Besides, 
complexity study of any potential solution in this respect deserves careful 
attention. How can possible solutions by which optimization techniques are 
combined with topology information be implemented in practical platforms? 
What are the impacts of such algorithms on the system complexity? In possible 
future work, this part of the thesis can be continued in many ways, for instance, 
how can the basic idea of the proposed eMAC protocol be used by frequency 
agile 802.11 MAC, or how to tackle the so-called unreachability problem in 
multi-channel cognitive 802.11 systems using the same approach that has been 
proposed in eMAC?

Chapter 4 tackles a real problem in multi-channel systems and opens up 
several interesting problems and issues that can be further studied in future. Using 
comprehensible illustrations, it has been shown in this chapter that the so-called 
multi-channel unreachability problem can easily cause considerable performance 
degradation in multi-channel cognitive ad hoc networks, though only a few 
solutions to address this problem were proposed. Studying the robustness of the 
proposed solutions in Chapter 4 in the presence of estimation error is an indeed 
interesting topic for future work. Extending the introduced stochastic channel 
selection technique to multi-channel systems without any control channel is 
another important topic that deserves more attention and serious work. How does 
combining the stochastic channel selection with the multi-channel backoff 
algorithm brings a better performance for the multi-channel MAC protocol? As 
security issues in cognitive networks are receiving considerable attention and in 
fact, it deserve serious concern due to the high potential of several types of 
attacks initiated by misbehaving radio stations, applying the solutions reported in 
[148 and 149] in order to propose more secure schemes to improve the robustness 
of the stochastic channel selection against radio stations’ misbehavior is definitely 
interesting.
The proposed practical schemes in this thesis can be re-used in other emerging applications, too. For instance, the proposed eMAC protocol can be deployed with some modifications by future femtocell access points when establishing a backup wireless backbone among them, either for control signaling or for exchanging part of their cell incoming traffic in emergency cases. In such scenarios, it is obvious that the unreachability (i.e., receiver blocking) problem will serve as the main source of low achievable capacity; thus, eMAC may play an important role in supplying higher performance and lower traffic latency. Furthermore, the introduced stochastic channel selection technique can be modified in order to be utilized by the next generation 3G LTE systems in which network entities, either mobile users or the central base station, are able to mimic cognitive radio functionality. In such scenarios, the base station may conduct channel allocation to mobile users based on the proposed stochastic channel selection in which a mobile user’s traffic type, the level of interference in each channel, the probability of interruption by the licensed (i.e., primary users), level of congestion per channel, etc. are altogether incorporated into the so-called count-down probabilities associated with each channel.

Finally, there is a fundamental question regarding the re-usability of the proposed schemes in this thesis for the future wireless Internet. As the future Internet architecture will be entirely information-centric, rather than endpoint-centric, it is imperative to investigate how the current user-centric architecture of the wireless networks can be modified to support the envisioned highly flexible information-centric structure. With this regard, the key questions that should be properly answered are as follows: How the currently deployed routing protocols in mobile ad hoc networks should be changed to eliminate the requirement for address-based traffic forwarding within multi-hop wireless ad hoc networks? How the security issues for information-centric based data sharing should be addressed in wireless ad hoc networks, which are well-known to be indeed vulnerable to the wide variety of security attacks? Is it necessary to re-design the current MAC protocols in wireless ad hoc networks to support the information-centric traffic forwarding and how the current radio resource allocation techniques are influenced by the envisioned future system architecture? As a prospective future direction to continue the current work in this thesis, one may focus on addressing the aforementioned fundamental questions, which is anticipated to bring considerable attention both from academic communities and industry.


78. (2005) IEEE 802.11g, Part 11: Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) Specifications: Further High-Speed Physical Layer Extension in the 2.4GHz Band, supplement to IEEE 802.11 Standard.


139. Kwak BJ, Song NO, & Miller LE (2005) Performance Analysis of Exponential
IEEE International Symposium on Personal, Indoor, and Mobile Radio
Communications, 1: 1–5.
143. Yu Z, Hoyos S & Sadler BM (2008) Mixed-Signal Parallel Compressed Sensing and
Reception for Cognitive Radio. Proc. IEEE International Conference on Acoustics,
Technology Assessment and Research Challenges. Chapter in Cognitive Networks:
Towards Self-Aware Networks, Edited by Q. H. Mahmoud, UK, John Wiley & Sons,
1: 97–120.
Chapter in Cognitive Networks: Towards Self-Aware Networks, Edited by Q. H.
Mahmoud, UK, John Wiley & Sons, 1: 203–222.
and Reasoning in Cognitive Networks: Methods and Design Decisions. Chapter in
Cognitive Networks: Towards Self-Aware Networks, Edited by Q. H. Mahmoud, UK,
Chapter in Cognitive Networks: Towards Self-Aware Networks, Edited by Q. H.
Networks: Towards Self-Aware Networks, Edited by Q. H. Mahmoud, UK, John
Wiley & Sons, 1: 293–314.
Area Networks. Chapter in Medium Access Control in Wireless Networks, Edited by
Access Control Protocols. Chapter in Medium Access Control in Wireless Networks,
Chapter in Medium Access Control in Wireless Networks, Edited by H. Wu and Y.


Appendix 1 Parallel space – time Markov chain balance equations

In this appendix, the Markovian balance characteristic equations describing the steady state behavior of the proposed analytical model in Chapter 2 are presented. For the single-hop scenario, if the probability of being in vacancy state, \( Pr(E) \), according to the explanations provided earlier, it is straightforward to derive the following Markovian Balance equations:

\[
E = Pr\{E\} = \frac{1}{W_0} \left( (1 - p)A_0(T_o) \sum_{i=0}^{m} B_{i,0,1} + pA_0(\Delta)B_{m,0,1} + b_1W_0 \left[ (1 - P'_{tr})A_0(\sigma) + P'_{tr}\left(P'_{s}A_0(T_o) + (1 - P_2)A_0(\Delta)\right) \right] \right) \left( 1 - \frac{\left[ (1 - P'_{tr})A_0(\sigma) + P'_{tr}\left(P'_{s}A_0(T_o) + (1 - P_2)A_0(\Delta)\right) \right] - W_0^{-1}(1 - p)(1 - P'_{tr})A_1(\sigma)A_0(T_o)}{W_0^{-1}(1 - p)(1 - P'_{tr})A_1(\sigma)A_0(T_o)} \right)^{-1}. \]  

(A.1)
\[ B_{0,k,j} = \]
\[ \frac{1}{W_0} \left( (1 - p) \sum_{n=0}^{j} \sum_{i=0}^{m} A_n(T_x)B_{i,0,(j+1-n)} + p \sum_{n=0}^{j} A_n(\Delta_c)B_{m,0,(j+1-n)} + \right. \]
\[ P_{tr} \left( P'_s A_j(T_x) + (1 - P'_s)A_j(\Delta) \right) E + (1 - p) \sum_{n=0}^{j} A_n(T_x)F_{(j+1-n)} + \]
\[ p \sum_{n=0}^{j-1} A_n(\Delta_c)F_{(j-n)} \] 
\[ \left. + \left[ (1 - P'_s)A_j(\sigma) + P'_s \left( P'_s A_j(T_x) + (1 - P'_s)A_j(\Delta) \right) \right] b_{(k+1)} \right) \] 
\[ (A.2) \]

for \( 0 \leq k < W_0 - 1, 1 < j < L. \)

\[ B_{0,k,j} = \]
\[ \frac{1}{W_0} \left( (1 - p) \sum_{n=0}^{j} \sum_{i=0}^{m} A_n(T_x)B_{i,0,(j+1-n)} + p \sum_{n=0}^{j} A_n(\Delta_c)B_{m,0,(j+1-n)} + \right. \]
\[ P_{tr} \left( P'_s A_j(T_x) + (1 - P'_s)A_j(\Delta) \right) E + (1 - p) \sum_{n=0}^{j} A_n(T_x)F_{(j+1-n)} + \]
\[ p \sum_{n=0}^{j-1} A_n(\Delta_c)F_{(j-n)} \] 
\[ \left. + \left[ (1 - P'_s)A_j(\sigma) + P'_s \left( P'_s A_j(T_x) + (1 - P'_s)A_j(\Delta) \right) \right] b_{(k+1)} \right) \] 
\[ (A.3) \]

for \( k = W_0 - 1, 1 < j < L. \)
\[ B_{i,k,j} = \]
\[ \frac{1}{W_i} p \sum_{n=0}^{j-1} A_n(\Delta_c) B_{(i-1),0,(j-n)} + \]
\[ \left[ (1 - P_{tr}') \sum_{n=0}^{j-1} A_n(\sigma) B_{i,(k+1),(j-n)} + \right. \]
\[ P_{tr}' \left( p_s \sum_{n=0}^{j-1} A_n(T_s) B_{i,(k+1),(j-n)} + (1 - p_s) \sum_{n=0}^{j-1} A_n(\Delta) B_{i,(k+1),(j-n)} \right), \]
\[ \text{(A.4)} \]

for \( 0 \leq k < W_i - 1, \quad 1 < j < L. \)

\[ B_{i,k,j} = \frac{1}{W_i} p \sum_{n=0}^{j-1} A_n(\Delta_c) B_{(i-1),0,(j-n)}, \]  
\[ \text{(A.5)} \]

for \( k = W_i - 1, 1 < j < L. \)

\[ B_{i,k,j} = \]
\[ \frac{1}{W_i} p A_0(\Delta_c) B_{(i-1),0,1} + \]
\[ \left[ (1 - P_{tr}') A_0(\sigma) B_{i,(k+1),1} + \right. \]
\[ P_{tr}' \left( p_s A_0(T_s) B_{i,(k+1),1} + (1 - p_s) A_0(\Delta) B_{i,(k+1),1} \right), \]
\[ \text{(A.6)} \]

for \( 0 \leq k < W_i - 1, j = 1. \)

\[ B_{i,k,j} = \frac{1}{W_i} p A_0(\Delta_c) B_{(i-1),0,1}, \]  
\[ \text{(A.7)} \]

for \( k = W_i - 1, j = 1. \)
\[ B_{l,k,j} = \begin{align*}
\frac{1}{W_l}^p \sum_{z=0}^{L-1} \sum_{n=0}^{L-1} A_{(n+z)}(\Delta_c)B_{(l-1),0,(l-n)} + \\
\left[ (1 - P_{tr}') \sum_{z=0}^{L-1} \sum_{n=0}^{L-1} A_{(n+z)}(\sigma)B_{l,(k+1),(l-n)} + \\
P_{tr}' \left( P_s \sum_{z=0}^{L-1} \sum_{n=0}^{L-1} A_{(n+z)}(T_s)B_{l,(k+1),(l-n)} + \\
(1 - P_s) \sum_{z=0}^{L-1} \sum_{n=0}^{L-1} A_{(n+z)}(\Delta)B_{(k+1),(l-n)} \right) \right].
\end{align*} \tag{A.8} \]

for \(0 \leq k = W_l - 1, j = L.\)

\[ B_{0,k,j} = \begin{align*}
\frac{1}{W_0}^p \sum_{n=0}^{L-1} A_{(n+z)}(\Delta_c)B_{(l-1),0,(l-n)}
\end{align*} \tag{A.9} \]

for \(0 \leq k = W_l - 1, j = L.\)

\[ B_{0,k,j} = \begin{align*}
\frac{1}{W_0} \left[ (1 - p) \sum_{n=0}^{m} A_{n}(T_s)B_{0,(2-n)} + p \sum_{n=0}^{m} A_{n}(\Delta_c)B_{0,0,(2-n)} + \\
P_{tr}'(P_sA_1(T_s) + (1 - P_s)A_1(\Delta))E + (1 - p) \sum_{n=0}^{m} A_{n}(T_s)F_{(2-n)} + \\
pA_0(\Delta_c)F_1 \right] + \\
\left[ (1 - P_{tr}')A_0(\sigma)B_{0,(k+1),1} + \\
P_{tr}'(P_sA_0(T_s)B_{0,(k+1),1} + (1 - P_s)A_0(\Delta)B_{0,(k+1),1}) + \\
[(1 - P_{tr}')A_1(\sigma)b_{(k+1)} + P_{tr}'(P_sA_1(T_s)b_{(k+1)} + (1 - P_s)A_1(\Delta)b_{(k+1)})] \right].
\end{align*} \tag{A.10} \]

for \(0 \leq k < W_0 - 1, j = 1.\)
\[
B_{0,k,j} = \frac{1}{W_0} \left[(1 - p) \sum_{n=0}^{1} \sum_{i=0}^{m_1} A_n(T_s)B_{i,0,(2-n)} + p \sum_{n=0}^{1} A_n(\Delta)B_{m,0,(2-n)} + pA_0(\Delta)F_1 \right]
\]

for \( k = W_0 - 1, j = 1 \).

\[
B_{0,k,j} = \frac{1}{W_0} \left[(1 - p) \sum_{n=1}^{\infty} \sum_{i=0}^{m} A_{n+2}(T_s)B_{i,0,(L-n)} + p \sum_{n=1}^{\infty} A_{n+2}(\Delta)B_{m,0,(L-n)} + \right.
\]

\[
pA_0(\Delta)F_1 \left] \right.,
\]

\[
\frac{1}{W_0} \left[(1 - p) \sum_{n=0}^{\infty} \sum_{z=1}^{L-1} \sum_{n=0}^{m} A_{n+2}(T_s)B_{z,0,(l-n)} + p \sum_{n=0}^{\infty} A_{n+2}(\Delta)B_{m,0,(l-n)} + \right.
\]

\[
\frac{1}{W_0} \left[(1 - p) \sum_{n=0}^{\infty} \sum_{z=1}^{L-1} \sum_{n=0}^{m} A_{n+2}(T_s)F_{l-n} + p \sum_{n=0}^{\infty} A_{n+2}(\Delta)F_{l-n} \right],
\]

for \( k = W_0 - 1, j = L \).
\[ B_{0,k,j} = \frac{1}{W_0} \left( (1 - p) \sum_{z=1}^{\infty} \sum_{n=0}^{L-1} \sum_{l=0}^{m} A_{n+z}(T_s) B_{l,0,(l-n)} + \right. \\
\left. p \sum_{z=1}^{\infty} \sum_{n=0}^{L-1} A_{n+z}(\Delta r) B_{m,0,(l-n)} + \right. \\
\left. P'_{tr} \left( P'_{s} \sum_{z=0}^{\infty} A_{L+z}(T_s) + (1 - P'_{s}) \sum_{z=0}^{\infty} A_{(L+z)}(\Delta) \right) E + \right. \\
\left. (1 - p) \sum_{z=1}^{\infty} \sum_{n=0}^{L-1} A_{n+z}(T_s) F_{l-n} + p \sum_{z=1}^{\infty} \sum_{n=0}^{L-1} A_{n+z}(\Delta r) F_{l-n} \right] + \\
\left[ (1 - P'_{tr}) \sum_{z=0}^{\infty} \sum_{n=0}^{L-1} A_{n+z}(\sigma) B_{0,(k+1),(l-n)} + \right. \\
\left. P'_{tr} \left( P'_{s} \sum_{z=0}^{\infty} A_{L+z}(T_s) B_{0,(k+1),(l-n)} + \right. \\
\left. (1 - P'_{s}) \sum_{z=0}^{\infty} A_{(L+z)}(\Delta) B_{0,(k+1),(l-n)} \right] + \\
\left[ (1 - P'_{tr}) \sum_{z=0}^{\infty} A_{(L+z)}(\sigma) + \right. \\
\left. P'_{tr} \left( P'_{s} \sum_{z=0}^{\infty} A_{L+z}(T_s) + (1 - P'_{s}) \sum_{z=0}^{\infty} A_{(L+z)}(\Delta) \right) b_{(k+1)} \right], \tag{A.13} \]

for \( 0 \leq k < W_0 - 1, j = L. \)

\[ b_k = \frac{1}{W_0} \left[ (1 - p) \sum_{l=0}^{m} A_0(T_s) B_{l,0.1} + p A_0(\Delta r) B_{m,0.1} + (1 - p) A_0(T_s) F_1 \right] + \tag{A.14} \\
\left. b_{(k+1)} \left[ (1 - P_{tr}) A_0(\sigma) + P_{tr} \left( P'_{s} A_0(T_s) + (1 - P'_{s}) A_0(\Delta) \right) \right]. \right] 

\]

for \( 1 \leq k < W_0 - 1. \)
\[ b_k = \frac{1}{W_0} \left[ (1 - p) \sum_{t=0}^{m} A_0(T_{s})B_{t,0,1} + pA_0(\Delta_0)B_{m,0,1} + (1 - p)A_0(T_{s})F_{1} \right], \quad (A.15) \]

for \( k = W_0 - 1. \)
Appendix 2 Simple and cognitive MAC protocols
Markov model balance equations

In this appendix, the Markovian balance characteristic equations describing the steady state behavior of the proposed analytical models in Chapter 4 are presented.

Possible state transitions for the simple multi-channel MAC protocol analytical model are summarized in Tables I to III. Table I represents possible state transitions during an idle slot on the control channel. Table II summarizes possible state transitions when a successful RTS/CTS handshake is conducted on the control channel. Finally, Table III tabulates all possible transitions when a frame collision occurs on the control channel.

Table 1. Idle slot on the control channel (Simple Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>( P_{\text{ctrl}}(n-1,[x_h]_{h=0}^{\infty}) )</td>
<td>Available resource: ( x_0 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [1,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 )</td>
</tr>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>1</td>
<td>No available resource: ( x_0 = 0 ), ( x_0 + x_1 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [0,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 )</td>
</tr>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>( P_{\text{ctrl}}(n,[x_h]_{h=0}^{\infty}) )</td>
<td>Available resource: ( x_0 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [1,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 ), ( g = g' - 1 )</td>
</tr>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>1</td>
<td>No available resource: ( x_0 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [0,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 ), ( g = g' - 1 )</td>
</tr>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>( \frac{1}{W_0}P_{\text{ctrl}}(n,[x_h]_{h=0}^{\infty}) )</td>
<td>Available resource: ( x_0 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [1,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 )</td>
</tr>
<tr>
<td>( \theta_{\theta,1}([{x_h}]<em>{h=0}^{\infty}) ) to ( \theta</em>{\theta,2}([{x_h}]_{h=0}^{\infty}) )</td>
<td>1</td>
<td>No available resource: ( x_0 = 0 ), ( x_0 + x_1 = x_0 ) if ( h \in [0,t_{\text{max}}-1] ), ( t_{\text{max}} = 0 )</td>
</tr>
</tbody>
</table>
Table 2. Successful RTS/CTS handshake on the control channel (Simple Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category to Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
</table>
| $\theta_{k+1}([x_{h_{\text{max}}}]_{n=0})$ to $\theta_{k}([x_{h_{\text{max}}}]_{n=0})$ | $p_k(n-1,[x_{h_{\text{max}}}]_{n=0})$ | Available resource: $a_0 = 0$ and $g = t_{\text{max}}$  
$\sum_{h=0}^{\Delta_h} a_h - 1 = a_0$  
$a_{h+\Delta_h} = a_h$ if $h \in [1,t_{\text{max}} - \Delta]$  
$a_0 = 0$ if $h \in [t_{\text{max}} - \Delta + 1,t_{\text{max}} - 1]$  
$t_{\text{max}} = 1$ |
| $\theta_{k}([x_{h_{\text{max}}}]_{n=0})$ to $\theta_{k}(g,[x_{h_{\text{max}}}]_{n=0})$ | $p_{\text{diff}}(n-1,[x_{h_{\text{max}}}]_{n=0})$ | Available resource: $a_0 = 0$ and $g = t_{\text{max}}$  
$\sum_{h=0}^{\Delta_h} a_h - 1 = a_0$  
$a_{h+\Delta_h} = a_h$ if $h \in [1,t_{\text{max}} - \Delta]$  
$a_0 = 0$ if $h \in [t_{\text{max}} - \Delta + 1,t_{\text{max}} - 1]$  
$t_{\text{max}} = 1$ |
| $\theta_{k}(g^*,[x_{h_{\text{max}}}]_{n=0})$ to $\theta_{k}(g,[x_{h_{\text{max}}}]_{n=0})$ | $p_k(n,[x_{h_{\text{max}}}]_{n=0})$ | Available resource: $a_0 = 0$  
$\sum_{h=0}^{\Delta_h} a_h - 1 = a_0$  
$a_{h+\Delta_h} = a_h$ if $h \in [1,t_{\text{max}} - \Delta]$  
$a_0 = 0$ if $h \in [t_{\text{max}} - \Delta + 1,t_{\text{max}} - 1]$  
$\max = 1$  
g = $\max(0,g^* - \Delta_h)$ |
| $\theta_{k}(g^*,[x_{h_{\text{max}}}]_{n=0})$ to $\theta_{k}(g,[x_{h_{\text{max}}}]_{n=0})$ | $\frac{1}{W_0} p_k(n,[x_{h_{\text{max}}}]_{n=0})$ | Available resource: $a_0 = 0$  
$\sum_{h=0}^{\Delta_h} a_h - 1 = a_0$  
$a_{h+\Delta_h} = a_h$ if $h \in [1,t_{\text{max}} - \Delta]$  
$a_0 = 0$ if $h \in [t_{\text{max}} - \Delta + 1,t_{\text{max}} - 1]$  
$t_{\text{max}} = 1$ |
To derive the Markovian balance equations for the multi-channel cognitive MAC protocol analytical model, it is assumed that in each transition $v = v_1 + v_2 + \cdots + v_{t_{\text{max}}}$ data channels are going to be occupied by the incumbents, where $v_z$ denotes the number of data channels with $z$ remaining busy time slots taken by the secondary users before the state transition that will be occupied by the incumbents ($0 \leq v_z \leq \delta_z$). In addition, it is assumed that $p_{\text{inc}}$ denotes the probability that an incumbent occupies a data channel, given that the channel is free of primary users during the previous time slot.

Possible state transitions for the multi-channel cognitive MAC protocol analytical model are summarized in Tables IV to VII. Table IV represents possible state transitions during an idle slot on the control channel. Table V summarizes possible state transitions when a successful RTS/CTS handshake is conducted on the control channel. Moreover, Table VI tabulates all possible transitions when a frame collision occurs on the control channel. Finally, Table VII shows possible transitions when there is a busy time slot on the control channel.

---

**Table 3. A frame collision on the control channel (Simple Multi-Channel MAC).**

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{2\to1}(x_{h_{k=0}}^{\text{max}})$</td>
<td>(P_{\text{coll}}(h-1,x_{h_{k=0}}^{\text{max}}))</td>
<td>Available resource: $\delta_2 = 0$ $\sum_{h \in D} \delta_h = \delta_2$ $\delta_{h\to c} = \delta_2$ if $h \in [1,t_{\text{max}} - K_c]$ $\delta_2 = 0$ if $h \in [t_{\text{max}} - K_c + 1,t_{\text{max}}]$ $\delta_{\text{max}} = 0$</td>
</tr>
<tr>
<td>$\theta_{2\to1}(x_{h_{k=0}}^{\text{max}})$</td>
<td>$\frac{1}{W_{i}} \cdot P_{\text{coll}}(n-1,x_{h_{k=0}}^{\text{max}})$</td>
<td>Available resource: $\delta_2 = 0$ $\sum_{h \in D} \delta_h = \delta_2$ $\delta_{h\to c} = \delta_2$ if $h \in [1,t_{\text{max}} - K_c]$ $\delta_2 = 0$ if $h \in [t_{\text{max}} - K_c + 1,t_{\text{max}}]$ $\delta_{\text{max}} = 0$</td>
</tr>
<tr>
<td>$\theta_{2\to1}(x_{h_{k=0}}^{\text{max}})$</td>
<td>$\frac{1}{W_{i}} \cdot P_{\text{coll}}(n,x_{h_{k=0}}^{\text{max}})$</td>
<td>Available resource: $\delta_2 = 0$ $\sum_{h \in D} \delta_h = \delta_2$ $\delta_{h\to c} = \delta_2$ if $h \in [1,t_{\text{max}} - K_c]$ $\delta_2 = 0$ if $h \in [t_{\text{max}} - K_c + 1,t_{\text{max}}]$ $\delta_{\text{max}} = 0$</td>
</tr>
<tr>
<td>$\theta_{2\to1}(x_{h_{k=0}}^{\text{max}})$</td>
<td>$\frac{1}{W_{i}} \cdot P_{\text{coll}}(n,x_{h_{k=0}}^{\text{max}})$</td>
<td>Available resource: $\delta_2 = 0$ $\sum_{h \in D} \delta_h = \delta_2$ $\delta_{h\to c} = \delta_2$ if $h \in [1,t_{\text{max}} - K_c]$ $\delta_2 = 0$ if $h \in [t_{\text{max}} - K_c + 1,t_{\text{max}}]$ $\delta_{\text{max}} = 0$</td>
</tr>
</tbody>
</table>
Table 4. Idle slot on the control channel (Cognitive Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{l,k} (e',</td>
<td>[s_h, h_{max}]</td>
<td>\rightarrow \theta_{l,k} (e, [s_h, h_{max}])</td>
</tr>
</tbody>
</table>
| No available resource: \( \delta_a = 0, e' = 0, e = 0 \)
| \( \delta_a - u_1 + u_4 = s_0 \)
| \( \delta_{h+1} - v_{h+1} = s_h \) if \( h \in [1, t_{max} - 1] \)
| \( u_{h+1} = u_h \) if \( h \in [1, t_{max} - 1] \)
| \( u_{max} = v \) |
| Available resource: \( \delta_a = 0, e' = 0, e = 0 \)
| \( \delta_a - v_2 + \delta_1 - v_4 + u_4 = s_0 \)
| \( \delta_{h+1} - v_{h+1} = s_h \) if \( h \in [1, t_{max} - 1] \)
| \( u_{h+1} = u_h \) if \( h \in [1, t_{max} - 1] \)
| \( u_{max} = v \) |

No available resource & no conflict with incumbent for the communication pair: \( \delta_a = 0, e' = 0, e = 0 \)
| Available resource & no conflict with incumbent for the communication pair: \( \delta_a = 0, e' = 0, e = 0 \)
| No available resource & conflict with incumbent for the communication pair: \( \delta_a = 0, e' = 0, e = 0 \)

207
Table 5. Successful RTS/CTS handshake on the control channel (Cognitive Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{k} e', g' \in [\Delta_{k}^{\max}, [\nu_{k}], [\nu_{k}^{\max}]$</td>
<td>$\frac{1}{W_0} (\frac{\delta_{0}}{\nu_{0}} - 1) (\frac{\delta_{1}}{\nu_{1}} - 1) \cdots (\frac{\delta_{u_{i}}}{\nu_{u_{i}}} - 1)$</td>
<td>Available resource &amp; conflict with incumbent for the communication pair: $\delta_{0} \neq 0, \delta_{e} = 0, \delta_{0} - \nu_{0} = 0, \delta_{1} - \nu_{1} + \nu_{k} = \delta_{0}$ $\forall h \in [1, T_{\max} - 1]$ $u_{i} = u_{h}$ if $h \in [1, T_{\max} - 1]$ $u_{T_{\max}} = v$</td>
</tr>
<tr>
<td>to $\theta_{k+1} e', g' \in [\Delta_{k+1}^{\max}, [\nu_{k}], [\nu_{k}^{\max}]$</td>
<td>$P_{\text{inc}}^{s}(1 - P_{\text{inc}})\text{Var}<em>{\Delta</em>{k}}^{\max} \Delta_{k} - \nu_{i}$</td>
<td>$\forall h \in [1, T_{\max} - 1]$ $u_{i} = u_{h}$ if $h \in [1, T_{\max} - 1]$ $u_{T_{\max}} = v$</td>
</tr>
<tr>
<td>$\theta_{k+1} e', g' \in [\Delta_{k+1}^{\max}, [\nu_{k}], [\nu_{k}^{\max}]$</td>
<td>$\frac{1}{W_0} (\frac{\delta_{0}}{\nu_{0}} - 1) (\frac{\delta_{1}}{\nu_{1}} - 1) \cdots (\frac{\delta_{u_{i}}}{\nu_{u_{i}}} - 1)$</td>
<td>Available resource: $\delta_{0} \neq 0, \delta_{e} = 0, \delta_{0} - \nu_{0} = 0, \delta_{1} - \nu_{1} + \nu_{k} = \delta_{0}$ $\forall h \in [1, T_{\max} - 1]$ $u_{i} = u_{h}$ if $h \in [1, T_{\max} - 1]$ $u_{T_{\max}} = v$</td>
</tr>
<tr>
<td>to $\theta_{k} e', g' \in [\Delta_{k}^{\max}, [\nu_{k}], [\nu_{k}^{\max}]$</td>
<td>$P_{\text{inc}}^{s}(1 - P_{\text{inc}})\text{Var}<em>{\Delta</em>{k}}^{\max} \Delta_{k} - \nu_{i}$</td>
<td>$\forall h \in [1, T_{\max} - 1]$ $u_{i} = u_{h}$ if $h \in [1, T_{\max} - 1]$ $u_{T_{\max}} = v$</td>
</tr>
<tr>
<td>$\theta_{k} e, g' \in [\Delta_{k}^{\max}, [\nu_{k}], [\nu_{k}^{\max}]$</td>
<td>$\frac{1}{W_0} (\frac{\delta_{0}}{\nu_{0}} - 1) (\frac{\delta_{1}}{\nu_{1}} - 1) \cdots (\frac{\delta_{u_{i}}}{\nu_{u_{i}}} - 1)$</td>
<td>Available resource &amp; no conflict with incumbent neither within $\Delta_{0}$ nor for the communication pair: $\delta_{0} \neq 0, \delta_{e} = 0, \delta_{0} - \nu_{0} = 0, \delta_{1} - \nu_{1} + \nu_{k} = \delta_{0}$ $\forall h \in [1, T_{\max} - 1]$ $u_{i} = u_{h}$ if $h \in [1, T_{\max} - 1]$ $u_{T_{\max}} = v$</td>
</tr>
<tr>
<td>Transition Category</td>
<td>Transition Probability</td>
<td>Notes</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>( \theta_A ) ((c', g', [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) ) to ( \theta_A ) ((c, g, [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) )</td>
<td>( \frac{1}{W_{h}} \left( \frac{s_{0} - 1}{v_0} \right) \left( \frac{s_{1} - 1}{v_1} \right) \left( \frac{1}{v_{T_{max}}^{h}} \right) \left( 1 - p_{in} \right)^{v_{T_{max}}^{h} s_0} )</td>
<td>Available resource &amp; conflict with incumbent within ( s_0 ) but not for the communication pair: ( s_0 \neq 0, c' = 0, c = \Delta_0, g = g' - 1 ) ( s_0 - v_0 + s_1 - v_1 + g_1 = s_0 ) ( s_{0, h, 1} - v_{0, h} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \sigma_{T_{max}} = 0 ) ( s_{0, h, 1} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \tau_{T_{max}} = 0 ) ( \tau_{T_{max}} = v )</td>
</tr>
<tr>
<td>( \theta_A ) ((c', g', [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) ) to ( \theta_A ) ((c, g, [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) )</td>
<td>( \frac{1}{W_{h}} \left( \frac{s_{0} - 1}{v_0} \right) \left( \frac{s_{1} - 1}{v_1} \right) \left( \frac{1}{v_{T_{max}}^{h}} \right) \left( 1 - p_{in} \right)^{v_{T_{max}}^{h} s_0} )</td>
<td>Available resource &amp; no conflict with incumbent within ( s_0 ) but for the communication pair: ( s_0 \neq 0, c' = 0, c = \Delta_0 ) ( s_0 - v_0 - 1 + s_1 - v_1 + g_1 = s_0 ) ( s_{0, h, 1} - v_{0, h} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \sigma_{T_{max}} = 1 ) ( s_{0, h, 1} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \tau_{T_{max}} = 0 ) ( \tau_{T_{max}} = v )</td>
</tr>
<tr>
<td>( \theta_A ) ((c', g', [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) ) to ( \theta_A ) ((c, g, [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) )</td>
<td>( \frac{1}{W_{h}} \left( \frac{s_{0} - 1}{v_0} \right) \left( \frac{s_{1} - 1}{v_1} \right) \left( \frac{1}{v_{T_{max}}^{h}} \right) \left( 1 - p_{in} \right)^{v_{T_{max}}^{h} s_0} )</td>
<td>Available resource &amp; conflict with incumbent both within ( s_0 ) and for the communication pair: ( s_0 \neq 0, c' = 0, c = \Delta_0 ) ( s_0 - v_0 + s_1 - v_1 + g_1 = s_0 ) ( s_{0, h, 1} - v_{0, h} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \sigma_{T_{max}} = 0 ) ( s_{0, h, 1} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \tau_{T_{max}} = 0 ) ( \tau_{T_{max}} = v )</td>
</tr>
<tr>
<td>( \theta_A ) ((c', g', [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) ) to ( \theta_A ) ((c, g, [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) )</td>
<td>( \frac{1}{W_{h}} \left( \frac{s_{0} - 1}{v_0} \right) \left( \frac{s_{1} - 1}{v_1} \right) \left( \frac{1}{v_{T_{max}}^{h}} \right) \left( 1 - p_{in} \right)^{v_{T_{max}}^{h} s_0} )</td>
<td>Available resource &amp; conflict with incumbent for the communication pair: ( s_0 \neq 0, c' = 0, c = \Delta_0 ) ( s_0 - v_0 + s_1 - v_1 + g_1 = s_0 ) ( s_{0, h, 1} - v_{0, h} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \sigma_{T_{max}} = 0 ) ( s_{0, h, 1} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \tau_{T_{max}} = 0 ) ( \tau_{T_{max}} = v )</td>
</tr>
<tr>
<td>( \theta_A ) ((c', g', [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) ) to ( \theta_A ) ((c, g, [s_{1}^{T_{max}}], [u]<em>{n=0}^{T</em>{max}}) )</td>
<td>( \frac{1}{W_{h}} \left( \frac{s_{0} - 1}{v_0} \right) \left( \frac{s_{1} - 1}{v_1} \right) \left( \frac{1}{v_{T_{max}}^{h}} \right) \left( 1 - p_{in} \right)^{v_{T_{max}}^{h} s_0} )</td>
<td>Available resource &amp; no conflict with incumbent for the communication pair: ( s_0 \neq 0, c' = 0, c = \Delta_0 ) ( s_0 - v_0 - 1 + s_1 - v_1 + g_1 = s_0 ) ( s_{0, h, 1} - v_{0, h} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \sigma_{T_{max}} = 1 ) ( s_{0, h, 1} = s_0 ) if ( h \in [1, T_{max} - 1] ) ( \tau_{T_{max}} = 0 ) ( \tau_{T_{max}} = v )</td>
</tr>
</tbody>
</table>
Table 6. A frame collision on the control channel (Cognitive Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 \left( \epsilon', \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$ to $\theta_1 \left( \epsilon, \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$</td>
<td>$p_{\text{incumbent}} \cdot \left(1 - p_{\text{trans}} \right) \cdot \left( 1 - p_{\text{trans}} \right)^{d_h - 1} \cdot \frac{1}{W_h} \cdot \left( u_h \right)^{d_h - 1}$</td>
<td>Available resource: $\mathbf{a}_h \neq 0, \epsilon = 0, \epsilon = K_r$ $\mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h = \mathbf{a}_h$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$</td>
</tr>
<tr>
<td>$\theta_1 \left( \epsilon, \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$ to $\theta_1 \left( \epsilon', \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$</td>
<td>$p_{\text{incumbent}} \cdot \left(1 - p_{\text{trans}} \right) \cdot \left( 1 - p_{\text{trans}} \right)^{d_h - 1} \cdot \frac{1}{W_h} \cdot \left( u_h \right)^{d_h - 1}$</td>
<td>Available resource &amp; no conflict with incumbent for the communication pair: $\mathbf{a}_h \neq 0, \epsilon = 0, \epsilon = K_r, \epsilon = \epsilon' - 1$ $\mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h = \mathbf{a}_h$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$</td>
</tr>
<tr>
<td>$\theta_1 \left( \epsilon, \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$ to $\theta_1 \left( \epsilon', \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$</td>
<td>$p_{\text{incumbent}} \cdot \left(1 - p_{\text{trans}} \right) \cdot \left( 1 - p_{\text{trans}} \right)^{d_h - 1} \cdot \frac{1}{W_h} \cdot \left( u_h \right)^{d_h - 1}$</td>
<td>Available resource &amp; conflict with incumbent: $\mathbf{a}_h \neq 0, \epsilon = 0, \epsilon = K_r, \epsilon = \epsilon' - 1$ $\mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h = \mathbf{a}_h$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$</td>
</tr>
<tr>
<td>$\theta_1 \left( \epsilon', \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$ to $\theta_1 \left( \epsilon, \left[ \mathbf{a}<em>{h,h}^{\text{trans}} \right], { \mathbf{a}</em>{h,h}^{\text{trans}} } \right)$</td>
<td>$p_{\text{incumbent}} \cdot \left(1 - p_{\text{trans}} \right) \cdot \left( 1 - p_{\text{trans}} \right)^{d_h - 1} \cdot \frac{1}{W_h} \cdot \left( u_h \right)^{d_h - 1}$</td>
<td>Available resource &amp; no conflict with incumbent for the communication pair: $\mathbf{a}_h \neq 0, \epsilon = 0, \epsilon = K_r, \epsilon = \epsilon' - 1$ $\mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h - \mathbf{a}_h + \mathbf{a}_h = \mathbf{a}_h$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$ $\mathbf{a}_h + \mathbf{a}_h = \mathbf{a}<em>h$ if $h \in [1, T</em>{\text{max}} - 1]$ $\mathbf{a}_h^{\text{trans}} = 0$</td>
</tr>
</tbody>
</table>
Table 7. A busy time slot on the control channel (Cognitive Multi-Channel MAC).

<table>
<thead>
<tr>
<th>Transition Category</th>
<th>Transition Probability</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{lk}(e', [\delta_h]<em>{h=0}^{\max}, [u_h]</em>{h=0}^{\max})$</td>
<td>$\frac{3}{2} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$</td>
<td>$e = e' - 1$</td>
</tr>
<tr>
<td>to $\theta_{lk}(e, [\delta_h]<em>{h=0}^{\max}, [u_h]</em>{h=0}^{\max})$</td>
<td>$p_{\text{inc}} \cdot (1 - p_{\text{inc}}) \frac{2}{3} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$</td>
<td>$\delta_h - v_0 + \delta_1 v_1 + u_1 = \delta_0$</td>
</tr>
</tbody>
</table>

| $\theta_{l}(e', g', [\delta_h]_{h=0}^{\max}, [u_h]_{h=0}^{\max})$ | $\frac{1}{\nu_0} \left( \frac{\delta_0}{v_0} \right) \cdots \left( \frac{\delta_{g-1}}{v_{g-1}} \right) \frac{2}{3} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$ | $e = e' - 1, g = g' - 1$ |
| to $\theta_{l}(e, g, [\delta_h]_{h=0}^{\max}, [u_h]_{h=0}^{\max})$ | $p_{\text{inc}} \cdot (1 - p_{\text{inc}}) \frac{2}{3} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$ | $\delta_h - v_0 + \delta_1 v_1 + u_1 = \delta_0$ |

$\delta_{l_{\max}} = 0$ if $h \in [1, T_{\text{max}} - 1]$ $\nu_{l_{\max}} = \nu$

$\delta_{l_{\max}} = 0$ if $h \in [1, T_{\text{max}} - 1]$ $\nu_{l_{\max}} = \nu$

Conflict with incumbent:

$\theta_{l}(e', g', [\delta_h]_{h=0}^{\max}, [u_h]_{h=0}^{\max})$ | $\frac{1}{\nu_0} \left( \frac{\delta_0}{v_0} \right) \cdots \left( \frac{\delta_{g-1}}{v_{g-1}} \right) \frac{2}{3} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$ | $e = e' - 1$ |

$\theta_{l}(e, g, [\delta_h]_{h=0}^{\max}, [u_h]_{h=0}^{\max})$ | $p_{\text{inc}} \cdot (1 - p_{\text{inc}}) \frac{2}{3} \delta_0 \left( \frac{\delta_1}{v_1} \right) \cdots \left( \frac{\delta_{l_{\max}}}{v_{l_{\max}}} \right)$ | $\delta_h - v_0 + \delta_1 v_1 + u_1 = \delta_0$ |

$\delta_{l_{\max}} = 0$ if $h \in [1, T_{\text{max}} - 1]$ $\nu_{l_{\max}} = \nu$
<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>320.</td>
<td>Komulainen, Mikko (2009)</td>
<td>Bandwidth enhanced antennas for mobile terminals and multilayer ceramic packages</td>
</tr>
<tr>
<td>322.</td>
<td>Liedes, Toni (2009)</td>
<td>Improving the performance of the semi-active tuned mass damper</td>
</tr>
<tr>
<td>323.</td>
<td>Marina Tyunina &amp; Orest Vendik (Eds.) (2009)</td>
<td>Proceedings of the 16th International Student Seminar “Microwave and optical applications of novel phenomena and technologies”, June 8–9, Oulu, Finland</td>
</tr>
<tr>
<td>324.</td>
<td>Belt, Pekka (2009)</td>
<td>Improving verification and validation activities in ICT companies—product development management approach</td>
</tr>
<tr>
<td>327.</td>
<td>Härkönen, Janne (2009)</td>
<td>Improving product development process through verification and validation</td>
</tr>
<tr>
<td>332.</td>
<td>Risikko, Tanja (2009)</td>
<td>Safety, health and productivity of cold work. A management model, implementation and effects</td>
</tr>
</tbody>
</table>
Kaveh Ghaboosi

INTELLIGENT MEDIUM ACCESS CONTROL FOR THE FUTURE WIRELESS NETWORKS