Université Hassan 1er
Centre d'Études Doctorales en Sciences et Techniques & Sciences Médicales

Faculté des Sciences et Techniques
Settat

THÈSE DE DOCTORAT

Pour l'obtention de grade de Docteur en Sciences et Techniques

Formation Doctorale: Mathématiques, Informatique et Applications

Spécialité: Mathématiques Appliquées

Sous le thème

Performance Evaluation and Enhancement of Restricted Access Wi-Fi IoT Networks: Stochastic Approaches

Présentée par :

Hamid TARAMIT

Soutenue le: *16 mars 2023*

A la Faculté des Sciences et Techniques de Settat devant le jury composé de :

Pr. Abderrahim Marzouk	**PES**	**Faculté des Sciences et Techniques, Settat**	**Président**
Pr. Mohamed Hanini	**PH**	**Faculté des Sciences et Techniques, Settat**	**Rapporteur**
Pr. Said El Kafhali	**PH**	**Faculté des Sciences et Techniques, Settat**	**Rapporteur**
Pr. Boujemâa Achchab	**PES**	**Ecole Nationale des Sciences Appliquées, Berrechid**	**Examinateur**
Pr. Abdelghani Ben Tahar	**PES**	**Faculté des Sciences et Techniques, Settat**	**Examinateur**
Pr. Luis Orozco Barbosa	**Professeur**	**Université de Castilla-La Mancha, Espagne**	**Co-Directeur de thèse**
Pr. Abdelkrim Haqiq	**PES**	**Faculté des Sciences et Techniques, Settat**	**Directeur de thèse**

Année Universitaire: 2022/2023

To all my family members, professors, and friends
who supported me during this journey.

Acknowledgments

This dissertation summarizes the research works conducted within the framework of a Cotutelle PhD program between the Computer, Networks, Mobility and Modeling (IR2M) laboratory, Hassan First University, Settat, Morocco, and the Albacete Research Institute of Informatics (I3A), University of Castilla-La Mancha, Albacete, Spain.

I would like to express my sincere gratitude to my supervisors, Dr. Abdelkrim Haqiq and Dr. Luis Orozco Barbosa, for their invaluable guidance, support, and mentorship throughout my PhD journey. Their expert knowledge and insightful feedback have been invaluable in shaping my research ideas and improving the quality of my work. I am grateful for their patience, kindness, and availability, which have been a constant source of motivation and encouragement. I feel privileged to have had the opportunity to work under their mentorship, and I will always cherish the invaluable lessons I learned from them. Sincere appreciation to Dr. Luis Orozco Barbosa for his invaluable personal assistance during my stay in Spain. He was exceptionally hospitable and went beyond his duties as a supervisor to ensure that I was comfortable and had everything I needed. Thank you, Dr. Abdelkrim Haqiq and Dr. Luis Orozco Barbosa, for your immeasurable contribution to my academic and personal growth.

I would like to express my appreciation to Dr. Javier Gomez and Dr. José Jaime Camacho-Escoto, for their contributions to our collaborative work. Their expertise, insights, and dedication have been invaluable in shaping the direction and scope of our research. Working with them has been a pleasure, and I am grateful for their unwavering commitment to producing high-quality research.

I would like to express my gratitude for the unwavering support and friendship provided by my colleagues in both the IR2M laboratory at Hassan First University and the I3A institute at the University of Castilla-La Mancha. I am honored to have had the opportunity to work alongside such talented and inspiring individuals.

I am grateful for the support and kindness provided by the administrative staff of the Hassan First University, specifically, Dr. Ahmed Fahli and Dr. Bouchaib Bencharki, and the administrative staff of the University of Castilla-La Mancha, specifically, Julia Corredor Cañadas, María Blanca Caminero Herráez and Ana Raquel Sicilia Fernandez. I am also grateful to the technical staff of I3A, particularly Raúl Galindo Moreno, Vicente López Camacho and José

Valentín García Gómez, who have always been available to provide invaluable assistance when needed.

Special thanks to the Ministry of Science and Innovation of Spain for supporting my work through the research projects RTI2018-098156-B-C52 and PID2021-123627OB-C52. I also deeply appreciate the financial support provided by the Erasmus+ mobility program KA107 and the mobility grant from Hassan First University.

I would like to express my sincere appreciation to all members of my thesis committee who have generously dedicated their time, expertise, and valuable insights to review my work and provide invaluable feedback. I sincerely thank Dr. Ahmed Karmouch, Dr. Said El Kafhali, and Dr. Mohamed Hanini for serving as preliminary examiners of my thesis manuscript. I am grateful for their time and effort in carefully reviewing my work and providing constructive criticism and thoughtful suggestions to improve the quality of my manuscript. I wish to express my sincere gratitude to Dr. Boujemâa Achchab and Dr. Abdelghani Ben Tahar for their invaluable contributions as examiners. Furthermore, I am grateful to Dr. Abderrahim Marzouk for honoring me by chairing of my thesis committee.

My deep and sincere gratitude goes to all my family and friends for their unwavering support throughout my academic journey. Their encouragement, love, and understanding have been my source of strength and motivation.

Hamid Taramit
Settat, February 2023

Abstract

Wireless communication technologies face the challenge of designing the mechanisms enabling the efficient operation of dense and large-scale Internet of Things (IoT) scenarios. Existing low-power IoT communication technologies suffer from short transmission ranges and low data rates. Accordingly, the IEEE Task Group ah (TGah) has developed a new standard called IEEE 802.11ah (also known as Wi-Fi HaLow) that offers a trade-off between throughput and range coverage. Significant changes have been introduced in the standard aiming to improve the performance in high contention traffic scenarios. Novel channel access mechanisms, such as the Restricted Access Window (RAW), should enable the deployment of network platforms consisting of up to 8191 devices over a range of up to 1 km. Under the RAW mechanism, the Access Point (AP) periodically restricts channel access to a limited number of stations during each RAW slot, which reduces collisions in dense scenarios. Although the standard defines the principles of operation of the RAW mechanism, it does not specify how to configure its various parameters. The design of efficient RAW configuration schemes will significantly benefit from the development of mathematical models. Such models should enable the performance evaluation and identification of the shortcomings of the potential configuration mechanisms. Throughout the thesis, we consider dense IEEE 802.11ah-based IoT networks and contribute to modeling, evaluating, and enhancing the network's performance using stochastic approaches. We develop several mathematical models for different network operation scenarios. We start with a system consisting of a saturated traffic load operating over an ideal channel to explore the maximum gain of IEEE 802.11ah-based networks. Then, we examine the case of non-ideal channel conditions with transmission errors and propose an efficient RAW configuration scheme. Additionally, we develop a model to evaluate the network performance of a more realistic scenario of IoT networks deployed in dense urban environments. In this model, we consider a Rayleigh fading channel condition in the presence of the capture effect. We then establish resource-efficient schemes to reconfigure the RAW slots by eliminating the extra time wasted during the RAW periods. This scheme greatly improves the network performance. Moreover, we develop an analytical model for unsaturated IEEE 802.11ah networks operating under a Rayleigh-fading channel with capture enabled. We also propose an optimal channel allocation scheme to ensure complete data delivery. Finally, we develop a Markov Decision Process (MDP) model for multiclass traffic routing control within a relay in an IEEE 802.11ah network. We consider a multichannel relay with forwarding service priority in terms of the transmission

data rate. Then, we derive a deterministic stationary optimal policy for the proposed routing control mechanism, which remarkably reduces the latency of packets within the relay. The developed mathematical models are meticulously validated through extensive campaigns of simulations derived via a discrete-event simulator developed in MATLAB. Although we focus on the IEEE 802.11ah standard, the developed models apply to all alternative standards that integrate periodic channel reservations (e.g., .11aa, .11ad, and .11ax). Our proposals are practical for enhancing the throughput, improving energy efficiency, and optimizing the channel resources of dense IoT-enabled Wi-Fi Networks.

Keywords: Stochastic Modeling, Stochastic Processes, Markov Decision Processes, Dense Wireless Networks, IEEE 802.11ah, Wi-Fi, MAC Protocols, Internet of Things, Rayleigh Channels, Performance Evaluation, Resource Allocation.

Résumé

Les technologies des communications sans fil sont confrontées au défi de concevoir des mécanismes permettant un fonctionnement efficace des scénarios d'Internet des objets (IoT) denses et à grande échelle. Les technologies de communication IoT existantes à faible puissance souffrent de courtes distances de transmission et de faibles débits de données. Par conséquent, IEEE Task Group ah (TGah) a développé une nouvelle norme appelée IEEE 802.11ah (également connue sous le nom de Wi-Fi HaLow) qui offre un compromis entre le débit et la distance de transmission. Des changements significatifs ont été introduits dans cette norme afin d'améliorer les performances dans les scénarios de trafic à haute compétition. De nouveaux mécanismes d'accès aux canaux, tels que le Restricted Access Window (RAW), devraient permettre la mise en place de plates-formes réseau comprenant jusqu'à 8191 appareils sur une portée allant jusqu'à 1 km. Dans le cadre du mécanisme RAW, le point d'accès (AP) restreint périodiquement l'accès au canal à un nombre limité de stations pendant chaque slot RAW, ce qui réduit les collisions dans les scénarios denses. Bien que la norme explique les principes de fonctionnement du mécanisme RAW, elle ne précise pas comment configurer ses différents paramètres. La conception de schémas de configuration efficaces pour le RAW bénéficiera grandement du développement de modèles mathématiques. De tels modèles devraient permettre d'évaluer les performances et d'identifier les lacunes des mécanismes de configuration potentiels. Tout au long de la thèse, nous considérons des réseaux IoT denses basés sur IEEE 802.11ah et contribuons à la modélisation, l'évaluation et l'amélioration des performances du réseau en utilisant des approches stochastiques. Nous développons plusieurs modèles mathématiques pour différents scénarios de fonctionnement du réseau. Nous commençons par un système constitué d'une charge de trafic saturée fonctionnant sur un canal idéal pour explorer le gain maximal des réseaux basés sur IEEE 802.11ah. Ensuite, nous examinons le cas de conditions de canal non idéales avec des erreurs de transmission et proposons un schéma de configuration RAW efficace. En outre, nous développons un modèle pour évaluer la performance du réseau d'un scénario plus réaliste de réseaux IoT déployés dans des environnements urbains denses. Dans ce modèle, nous considérons une condition de canal à évanouissement de Rayleigh en présence de l'effet de capture. Nous établissons ensuite des schémas efficaces en termes de ressources pour reconfigurer les créneaux RAW en éliminant le temps supplémentaire gaspillé durant les périodes de RAW. Ce schéma améliore considérablement les performances du réseau. De plus, nous développons un modèle analy-

tique pour les réseaux IEEE 802.11ah non saturés fonctionnant sous un canal à évanouissement de Rayleigh avec capture activée. Nous proposons également un schéma d'allocation optimale des canaux pour garantir la livraison complète des données. Finalement, nous développons un modèle de processus de décision Markovien (MDP) pour le contrôle du routage du trafic multiclasse au sein d'un relais dans un réseau IEEE 802.11ah. Nous considérons un relais multicanal avec une priorité de service de routage en termes de débit de transmission. Ensuite, nous dérivons une politique optimale déterministe stationnaire pour le mécanisme de contrôle de routage proposé, qui réduit remarquablement la latence des paquets dans le relais. Les modèles mathématiques développés sont méticuleusement validés par plusieurs simulations dérivées via un simulateur à événements discrets développé avec MATLAB. Bien que nous nous concentrions sur la norme IEEE 802.11ah, les modèles développés s'appliquent à toutes les normes alternatives qui intègrent des réservations périodiques de canaux (par exemple, .11aa, .11ad et .11ax). Nos propositions sont pratiques pour augmenter le débit, améliorer l'efficacité énergétique et optimiser les ressources en canaux dans les réseaux Wi-Fi denses de l'IoT.

Mots clés: Modélisation stochastique, Processus stochastiques, Processus de Décision Markovien, Réseaux sans fil denses, IEEE 802.11ah, Wi-Fi, Protocoles MAC, Internet des objets, Canaux de Rayleigh, Évaluation des performances, Allocation des ressources.

Resumen

Las tecnologías de comunicación inalámbrica se enfrentan al reto de diseñar los mecanismos que permitan el funcionamiento eficiente de escenarios densos y a gran escala del Internet de las cosas (IoT). Las tecnologías de comunicación del IoT de baja potencia existentes adolecen de un corto alcance de transmisión y de una baja velocidad de datos. Por ello, el Grupo de Trabajo ah (TGah) del IEEE ha desarrollado un nuevo estándar denominado IEEE 802.11ah (también conocido como Wi-Fi HaLow) que ofrece un equilibrio entre rendimiento y cobertura de alcance. Se han introducido cambios significativos en la norma con el objetivo de mejorar el rendimiento en escenarios de tráfico de alta contención. Los nuevos mecanismos de acceso al canal, como la Ventana de Acceso Restringido (RAW), deberán permitir el despliegue de plataformas de red formadas por hasta 8191 dispositivos en un rango de hasta 1 km. Con el mecanismo RAW, el punto de acceso (AP) restringe periódicamente el acceso al canal a un número limitado de estaciones durante cada ranura RAW, lo que reduce las colisiones en escenarios densos. Aunque la norma define los principios de funcionamiento del mecanismo RAW, no especifica cómo configurar sus distintos parámetros. El diseño de esquemas de configuración eficaces para RAW se beneficiarán significativamente del desarrollo de modelos matemáticos. Dichos modelos permitirán la evaluación del rendimiento y la identificación de las deficiencias de los posibles mecanismos de configuración. A lo largo de la tesis, se aborda el estudio de redes IoT densas operando bajo la norma IEEE 802.11a abordando el desarrollo de modelos estocásticos con el fin de evaluar y mejorar las prestaciones de este tipo de redes. En esta tesis, se desarrollan varios modelos matemáticos considerando diferentes escenarios de funcionamiento de la red. Primeramente se considera un sistema operando bajo un régimen de carga de tráfico saturado y un canal ideal con el objetivo de explorar la ganancia máxima de las redes basadas en IEEE 802.11ah. A continuación, se examina el caso de condiciones de canal no ideales con errores de transmisión y se propone un esquema de configuración RAW eficiente. Además, se desarrolla un modelo para evaluar el rendimiento de la red en un escenario más realista. En este modelo, se asume una condición de canal con desvanecimiento Rayleigh en presencia del efecto de captura. A continuación, se definen esquemas eficientes en cuanto a recursos para reconfigurar las ranuras temporales del RAW eliminando el tiempo extra que se pierde durante los periodos RAW. Este esquema mejora enormemente el rendimiento de la red. Además, se desarrolla un modelo analítico para las redes IEEE 802.11ah no saturadas que operan bajo un canal con desvanecimiento Rayleigh con captura habilitada. También se propone un esquema

óptimo de asignación de canales para garantizar la entrega completa de datos. Finalmente, desarrollamos un modelo de Proceso de Decisión de Markov (MDP) para el control de enrutamiento de tráfico multiclase dentro de un relé en una red IEEE 802.11ah. Consideramos un relé multicanal con prioridad de servicio de reenvío en función de la velocidad de transmisión de datos. A continuación, derivamos una política óptima estacionaria determinista para el mecanismo de control de enrutamiento propuesto, que reduce notablemente la latencia de los paquetes dentro del relé. Los modelos matemáticos desarrollados se validan meticulosamente mediante extensas campañas de simulaciones derivadas a través de un simulador de eventos discretos desarrollado en MATLAB. Aunque la tesis se centra en el estándar IEEE 802.11ah, los modelos y la metodología de evaluación son aplicables a todos aquellos estándares que definen reservas periódicas de canales, como por ejemplo los estándares, .11aa, .11ad y .11ax. Los resultados de la tesis pueden ser utilizados para guiar el diseño de productos comerciales con el fin de mejorar las prestaciones, así como la eficiencia energética con el fin de optimizar el uso de los recursos de canal de las redes IoT IEEE 802.11ah.

Palabras clave: Modelado estocástico, Procesos estocásticos, Procesos de Decisión de Markov, Redes inalámbricas densas, IEEE 802.11ah, Wi-Fi, Protocolos MAC, Internet de las cosas, Canales Rayleigh, Evaluación del rendimiento, Asignación de recursos.

Contents

List of Tables

List of Figures

List of Algorithms

List of Abbreviations

A | B | C | D | E | F | G | H | I | L | M | N | O | P | Q | R | S | T | U | V | W

A

AC Access Category. 96, 146

ACCP Average Conditional Capture Probability. 99, 100, 127, 136, 147, 148

ACK Acknowledgment. 39, 55, 95, 115, 126, 146

ACS Authentication Control Slot. 25

AID Association Identification. 21, 25, 26

AoI Age of Information. 141–143, 160, 184, 185, 187, 189, 191, 193

AP Access Point. 1, 2, 5, 7, 20–28, 36–38, 61, 63, 65, 68, 70, 92–97, 99, 100, 103, 108–110, 113, 114, 117, 120, 160, 162, 163, 184, 187, 191

B

BER Bit Error Rate. 86–88

BI Beacon Interval. 2

BLE Bluetooth Low Energy. 21, 22

C

CAC Centralized Authentication Control. 25

CBAP Contention-Based Access Period. 36, 63

CDF Cumulative Distribution Function. 11

CGR Channel Gain Ratio. 133, 136

CSMA Carrier Sense Multiple Access. 45, 93

CTMC Continuous-Time Markov Chain. 17

CU Channel Usage. 5, 184, 187, 191

CW Contention Window. 39–41, 95, 109, 117, 127, 146

D

DAC Distributed Authentication Control. 25

DCF Distributed Coordination Function. 2, 24, 26, 29, 30, 37, 39, 40, 45, 52, 61, 63, 65, 81–86, 89, 92, 93, 152, 183, 187, 191

DIFS DCF Inter-frame Spacing. 39, 55, 78, 95, 124, 149

DTMC Discrete-Time Markov Chain. 13–15, 42

E

EDCA Enhanced Distributed Channel Access. 2, 26, 29–33, 36, 37, 61, 95, 109, 110, 126, 140, 146

F

FD Full-Duplex. 21

FIFO First-In-First-Out. 164

G

GS-DCF Group-Synchronized Distributed Coordination Function. 30, 31

H

HCCA Hybrid coordination function Controlled Channel Access. 36, 63

I

IDC International Data Corporation. 61

IID Independent and Identically Distributed. 8, 9, 11, 12, 48–50, 105–107, 150

IoT Internet of Things. 1–4, 19, 21–24, 27, 28, 30, 32–35, 37, 60–62, 90, 92, 96, 122, 123, 137, 138, 140, 141, 160–162, 183–193

IP Internet Protocol. 30

ISM Industrial, Scientific and Medical. 22

L

LACA Load-Aware Channel Allocation. 141, 142, 144, 146, 150, 151, 153, 157, 159, 160, 184, 187, 191

LOS line of sight. 96

LP Linear Programming. 170

LPWAN Low-Power Wide Area Networks. 1, 21, 22, 61

LT Lookup Table. 170

LTE Long-Term Evolution. 21, 23

M

MAC Medium Access Control. 20, 21, 24, 28–30, 34, 37, 38, 45, 63, 93, 123, 140

MC Markov Chain. 3, 39

MCCA Mesh coordination function Controlled Channel Access. 36, 63

MCS Modulation and Coding Scheme. 24

MDP Markov Decision Process. 5, 17, 18, 160, 163, 171, 181, 182, 184, 187, 191

MST Mean Sojourn Time. 163, 176–178, 180–182

N

NB-IoT Narrowband Internet of Things. 21, 22

NLOS non-line-of-sight. 92, 96, 143

O

OFDM Orthogonal Frequency-Division Multiplexing. 24

P

PDF Probability Density Function. 99, 100

PDR Packet Delivery Ratio. 5, 151, 154–156, 159, 184, 187, 191

PER Packet Error Rate. 86

PHY Physical. 24

PI Policy Iteration. 170

PLCP Physical Layer Convergence Protocol. 39, 96, 126, 146

PMF Probability Mass Function. 9

Q

QoS Quality of Service. 33, 123

QTP Quiet Time Period. 37

R

RAW Restricted Access Window. 2–5, 23, 26–34, 37–40, 43–45, 47–59, 61–63, 65–82, 84–89, 92–98, 101–104, 106–121, 123, 124, 126–128, 130–138, 140–144, 146, 150–156, 158–160, 162, 183–188, 190–192

RPS RAW Parameter Set. 28, 29

S

SG Stochastic Geometry. 185, 188, 192

SI Sector Interval. 27, 162

SIFS Short Inter-frame Space. 39, 55, 95, 146

SIR Signal-to-Interference Ratio. 110, 147

SP Service Period. 36, 63

sub-RAW sub-Restricted Access Window. 62, 68–73, 75, 89

T

TDMA Time Division Multiple Access. 20, 65

TGah Task Group ah. 2, 21, 22, 26, 27, 61, 123, 139, 162

TIM Traffic Indication Map. 23

TWT Target Wake Time. 23

TXOP Transmission Opportunity. 36, 39, 63, 95, 124, 126, 146

U

UAV Unmanned Aerial Vehicle. 141, 185, 189, 193

UWB Ultra Wide Band. 20

V

VI Value Iteration. 170, 178

VoIP Voice over Internet Protocol. 111, 133, 152

W

Wi-SUN Wireless Smart Utility Network. 21

WLAN Wireless Local Area Network. 36, 45

WPAN Wireless Personal Area Networks. 1, 21, 22, 61

General Introduction

Introduction

The world's lifestyle is becoming increasingly digitized thanks to the fast growth and development of communication technologies and smart devices. The Internet of Things (IoT) has evolved in this sense to organize the new smart life and provide reliable connectivity to smart devices [1]. Several challenges appeared, such as offering connectivity to ultra-dense networks, maximizing the lifetime of battery-based devices, enhancing the network throughput, optimizing the network latency, etc.

As IoT incorporates an increasing number of devices, it is vital to develop wireless communication technologies that meet the requirements of IoT applications, such as large-scale connectivity, low power consumption, and reliable throughput [2]. New wireless technologies are expected to ensure the reliable interconnection of a massive number of IoT devices, such as sensors, smartphones, and home appliances. Two types of low-power IoT communication technologies are currently available: Wireless Personal Area Networks (WPAN) and Low-Power Wide Area Networks (LPWAN) technologies [3]. WPAN technologies (e.g., Zig-Bee and Bluetooth) provide medium data rates (i.e., up to a few hundred kilobits per second) at a short range (i.e., tens of meters), whereas LPWAN technologies (e.g., LoRa and SigFox) provide long-range communications (i.e., up to tens of kilometers) at low data rate (i.e., up to a few kilobits per second). Due to the short transmission range of WPAN and the insufficient throughput of both WPAN and LPWAN, the use cases of such technologies in IoT scenarios are minimal. As such, it is still required to have an IoT communication technology offering higher data rates over large transmission ranges.

Wi-Fi technology could be an adequate solution to IoT applications, as it has established great success since its invention and has become one of the most widely utilized wireless technologies worldwide. However, it only applies to small-scale networks and cannot address IoT applications due to long-distance transmissions and tremendous targeted devices. Turning a large-scale network into several small-scale networks by deploying many Access Points (APs) is costly and raises issues of Overlapping Basic Service Sets (OBBS) interference. Therefore, to make use of Wi-Fi technology and fill the gap between WPAN and LPWAN technologies, IEEE

Task Group ah (TGah) has developed a new standard called IEEE 802.11ah (also known as Wi-Fi HaLow) [4]. It aims to offer a trade-off between throughput and range coverage, where a single 802.11ah AP can provide connectivity to up to 8191 devices over a communication range of up to 1 km.

The IEEE 802.11ah Wi-Fi standard was officially released in 2016, presenting several amendments to the legacy Wi-Fi technology. Being the first Wi-Fi solution to address IoT applications, it operates in unlicensed sub-1Ghz frequency bands and provides data rates ranging from 150 Kbps to 78 Mbps [3]. To improve the network performance in high contention scenarios, significant changes have been proposed in legacy IEEE 802.11 standards, the most important of which is the new channel access mechanism, called Restricted Access Window (RAW). This mechanism aims to address the high channel contention of IoT networks and increase throughput and energy efficiency by splitting the channel time into periods and allocating each period to a given group of stations. During each Beacon Interval (BI), the AP divides the channel time into RAW periods and restricts channel access to a given group of stations during each period. To alleviate the contention even more, each RAW is decomposed into one or more periods, named RAW slots. The stations within the RAW are distributed among the RAW slots and the channel is limited to only the assigned stations during each RAW slot. During each RAW slot, the stations contend for channel access using the Distributed Coordination Function (DCF) or Enhanced Distributed Channel Access (EDCA) protocols. Although the standard explains how the RAW mechanism works, it does not specify how to configure its parameters. The design of the configuration of such a mechanism will significantly benefit from the development of mathematical models. Such models should enable the performance evaluation and identification of the impact of involving parameters. Additionally, closed-form analytical results allow for performing practical configurations to enhance the overall network performance effectively.

Research Objectives

We address in this thesis the mathematical modeling, performance evaluation, and optimization of dense Wi-Fi based IoT networks. We consider the prominent IEEE 802.11ah standard, which is the first Wi-Fi technology aiming to address IoT applications. We aim to examine the novelty in this standard, especially the restricted channel access, develop adequate mathematical models, evaluate the network performance, and propose solutions to tune up the RAW configurations and enhance the network performance. The main questions we address in this thesis are as follows:

- Many works in the literature have addressed the modeling and performance evaluation of Wi-Fi networks. Are those models applicable to restricted channel access based networks? If not, how to model time-restricted channel access contention?

- How to model and evaluate the channel access in a RAW-based Wi-Fi network? Moreover, how to tune up RAW parameters to improve the network performance?
- As Wi-Fi HaLow technology addresses large-scale IoT applications, what is the appropriate channel model, and how to integrate the power fading in such scenarios?
- Considering a non-saturated IoT network, how to make an appropriate channel reservation for periodic channel access in order to maximize the gain of the RAW mechanism and optimize the network resource allocation?

Research Contributions

This thesis aims to develop accurate mathematical models to evaluate and enhance the performance of IEEE 802.11ah-based networks. This standard adopts a new channel access mechanism based on periodic channel reservation and allows only a group of stations to access the channel during a given time interval. Markov Chain (MC) based models are commonly used to represent legacy IEEE 802.11 networks, providing accurate stationary results. However, those models are valid only for the case of permanent channel access contention and do not apply to time restricted channel access. The existing analytical frameworks for networks are mainly based on Markov chains, where the first and main model was proposed by Bianchi [5].

Therefore, we develop new mathematical models using stochastic approaches based on renewal theory, where we deploy arrival processes, counting processes, and renewal processes. The developed analytical frameworks allow us to accurately model the behavior of a time-constrained network with restricted channel access. Then, evaluate the network performance and propose enhancements through different scheduling and configuration schemes. Developing accurate models for IEEE 802.11ah-based networks is an essential step towards proposing efficient RAW configuration schemes, especially since the standard does not mention how to perform this configuration and many parameters are involved in this procedure. Additionally, the main applications of this technology are densely deployed and large-scale networks that periodically access the channel using the RAW mechanism. Therefore, RAW configuration critically affects the overall network performance and needs to be dynamically adjusted based on the network conditions. We address different network scenarios for IoT applications, develop appropriate mathematical models to evaluate network performance, and propose adaptive enhancements using stochastic approaches. Besides the analytical frameworks, We also develop discrete event simulations using MATLAB for every network scenario considered, and we meticulously validate all the developed mathematical models.

Thesis Structure

The structure of this thesis is organized as follows:

- Chapter 1 presents the stochastic tools deployed to develop analytical approaches for different dense wireless communication scenarios. We mainly focus on the stochastic processes related to Renewal theory and Markov chains.
- Chapter 2 presents the state of the art for the Wi-Fi technology under consideration, IEEE 802.11ah, and the associated analytical models. We depict an overview of this technology and its potential to meet the requirements of dense IoT applications, particularly the new channel access mechanism, RAW. Then, we present various analytical works that established the modeling and improvement of IEEE 802.11ah networks. In addition, we highlight the shortcomings of existing mathematical models.
- Chapter 3 develops a new mathematical model for time-bounded access in Wi-Fi networks. We deploy arrival processes to model the instants of initiated transmissions and counting processes to track these transmissions over time. With this approach, we evaluate the performance of restricted access in IEEE 802.11ah. We also present a Markov chain model that provides stationary results, and we show the convergence of our model results (time-constrained) to stationary (time-independent) results.
- Chapter 4 develops an analytical framework for designing and optimizing an IEEE 802.11ah RAW configuration mechanism. We propose a grouping scheme that divides the channel time into RAW slots and assigns stations. Then, we evaluate the gain of this scheme and propose a configuration algorithm to maximize the network performance. We also analyzed the impact of different parameters on the network performance and showed the effectiveness of our proposals in enhancing the network throughput. Additionally, we examine the impact of a non-ideal channel with transmission errors by integrating our proposals into this scenario.
- Chapter 5 develops an accurate mathematical model to evaluate the network performance in a more realistic scenario of IoT networks deployed in dense urban environments. In this model, We consider a Rayleigh fading channel condition in the presence of the capture effect. We model channel access under capture awareness and extend the accuracy of the model in Chapter 3 by deploying counting processes to track both transmissions and idle slots during RAW slots under the constraints of the Rayleigh channel, the capture effect, and the spatial distribution of stations. Therefore, we evaluate the performance of a single RAW slot and a RAW composed of multiple RAW slots in terms of throughput and channel wastage. We examine the impact of different parameters on the overall performance, including access time, the number of stations, and the capture threshold. We also investigate the impact of the capture effect within a RAW-based network.

- Chapter 6 develops efficient schemes to manage IEEE 802.11ah networks operating in Rayleigh channels with capture effect. First, We establish the analytical modeling of the considered scenario. Then, We propose an algorithm that selects the best choices for a RAW slot duration. This approach avoids channel wastage and maximizes the performance of the RAW slot in terms of throughput and energy efficiency. Then, we contribute to designing a RAW configuration algorithm that provides a new configuration that optimizes channel allocation resources and significantly improves network performance in terms of throughput and energy efficiency.

- Chapter 7 develops an analytical model based on a two-level renewal process for unsaturated IEEE 802.11ah networks operating under a Rayleigh fading channel with capture enabled. Based on this model, we propose an algorithm for optimal channel allocation for stations. This algorithm considers the number of stations and their spatial distribution around the AP, Payload size, and capture threshold. To prove the effectiveness of the proposed algorithm, we evaluate the RAW slot performance in terms of Packet Delivery Ratio (PDR) and Channel Usage (CU). Compared to a random pre-allocation of the RAW slot period, the proposed algorithm achieves full delivery of all packets while significantly improving the CU.

- Chapter 8 addresses the routing control of multiclass traffic through a relay in an IEEE 802.11ah network. We model the forwarding mechanism as a queuing system with three M/M/1 queues and three types of traffic, where the packets in each queue are transmitted with a different data rate under a different channel. We propose a routing control mechanism for packets with lower priority to be forwarded through queues with higher data rates without interfering with the performance of the prioritized traffic. We formulate the proposed routing control under an Markov Decision Process (MDP) framework and derive a deterministic stationary optimal policy minimizing the system discounted cost. We establish the characterization of the optimal policy and the service priority in the system. Therefore, We evaluate the Mean Sojourn Time (MST) for non-prioritized packets when the optimal routing policy is applied and prove the effectiveness of our proposed routing control. Finally, We present some numerical results to analyze the optimal policy and the impact of the routing costs. In addition, we show that our proposal remarkably reduces the latency of packets within the relay.

The manuscript ends with a conclusion summarizing the main contributions and results of the thesis and proposes several perspectives for future work.

CHAPTER 1

Stochastic Tools for Modeling Wireless Communications

Contents

1.1 Introduction

Wireless communications have been one of the fastest growing and most impactful technologies over the last two decades. Wireless local area networks are currently replacing wired networks in many homes, businesses, and campuses. They allowed the building of new applications such as wireless sensor networks, smart homes and appliances, and remote telemedicine [6]. A wireless communication network can be seen as a collection of devices distributed in a given area. Each device can be either a transmitter or a receiver at a given time, depending on the network topology. The shared resources of the wireless medium present a challenge to establishing reliable connectivity. Hence, different adaptive access protocols are deployed to manage resource utilization depending on the network scenario. Since medium access protocols and other network setup parameters are crucial to the overall network performance, it is essential to determine the best configuration and tune up the network parameters for every use case. However,

such tasks are not obvious as it is challenging to characterize wireless network behavior and assess its performance in large-scale scenarios [7]. Deploying field trials for testing purposes is very costly, and running simulations is time-consuming due to the high system complexity. Thus these two options are not practical for several use cases, especially a network with dynamic conditions which requires real-time adjustments. Hereafter, an adequate alternative is to evaluate the network performance analytically using adaptive mathematical approaches. Analytical frameworks provide closed-form results of the network performance in terms of its parameters and hence allow the investigation of the impact of each parameter and optimize the network configuration for different performance metrics such as throughput, energy efficiency, delay, etc. Therefore, analytical evaluations of the network performance are not only limited to theoretical studies but also significantly necessary for practical purposes such as optimal configuration, real-time resource management, and development [8].

Stochastic processes present an excellent tool to theoretically model a wireless network behavior. They describe the evolution of a system that evolves randomly in time. When the system is observed at discrete times, we obtain discrete-time stochastic processes, whereas if the system is observed continuously in time, we obtain continuous-time stochastic processes. A wireless communication network is evolving randomly in time, and many aspects can be viewed as stochastic processes, including the state of the AP or a single station. The system state can be represented by many aspects, such as traffic load and channel usage, the spatial distribution and mobility of devices, and the events occupying the wireless medium. Therefore, stochastic modeling is an excellent approach to studying such a system effectively. However, it is also challenging due to the complexity and random behavior of real network systems. A stochastic model providing closed-form results allows many practical tasks, including evaluating the system performance, studying the impact of several parameters, and optimizing the system in terms of a given parameter or a tradeoff of several parameters. Once a stochastic framework is validated via simulations, it becomes an effective tool allowing the optimization and enhancement of a system in a real-time matter. Mathematical modeling can go further than that and can be an excellent tool for developing effective decision-making systems directly implemented within an agent device (Station, AP, Server,...). The development of several communication technologies was and still is growing in parallel with appropriate mathematical models. As many network systems have stochastic behavior, stochastic processes can significantly participate in the enhancement and development of such systems. Hereafter, we present in this chapter the main tools of stochastic processes that we deploy throughout the contributions of this thesis.

1.2 Stochastic Processes

Definition 1.1 (Stochastic process)**.** *A stochastic process is an infinite collection of random variables defined on a common probability model. These random variables are indexed by an integer or a real number usually interpreted as time. If they are indexed by positive integers,*

the process $X_1, X_2, ...$, is called a ***discrete-time stochastic process****. Whereas, if they are indexed by a positive real number t, the process is noted $\{X_t; t \geq 0\}$ and is called a* ***continuous-time stochastic process****.*

Note that stochastic processes are also referred to as random processes, however, it has become to commonly use the word random for random variables and the word stochastic for stochastic processes.

A discrete-time stochastic process assigns a sequence of numbers to every outcome $\omega \in \Omega$:

$$\omega \mapsto \{X_1(\omega), X_2(\omega), ...\},$$

while a continuous-time stochastic process assigns a function defined on the half-line $[0, \infty)$:

$$\omega \mapsto X_t(\omega)$$

The value of the stochastic process on a particular outcome ω is called a sample path of the process. When studying stochastic processes, one may choose to abstract to a probability model in which the outcomes are the sample paths, and events are sets of sample paths.

The stochastic processes can be used to model repeated experiments, arrivals or departures (of customers, orders, signals, packets, etc.), and random walks (over a line, in a plane, in a 3D space).

1.3 Bernoulli Processes

Definition 1.2 (Bernoulli process)**.** *A Bernoulli process is a sequence $Z_1, Z_2, ...$, of Independent and Identically Distributed (IID) binary random variables, each follows a Bernoulli distribution with the same parameter p, i.e., $p = \Pr\{Z_i = 1\}$ and $1 - p = \Pr\{Z_i = 0\}$. The independence here is understood in the sense that for every integer n, the random variables $Z_1, Z_2, ..., Z_n$ are IID.*

A Bernoulli process is usually seen as an evolving discrete-time stochastic process describing the arrival of customers. The event $\{Z_i = 1\}$ represents the arrival of a customer at time i, and $\{Z_i = 0$ represents no arrival at time i. Hence, at most one arrival can occur at each integer time.

It is hereby interesting to study the interarrival times in this process. The first interarrival

time X_1 represents the number of time units required for the first customer to arrive. We have

$$X_1 = \begin{cases} 1 & \text{if } Z_1 = 1 \\ 2 & \text{if } Z_1 = 0, Z_2 = 1 \\ 3 & \text{if } Z_1 = 0, Z_2 = 0, Z_3 = 1 \\ \vdots & \\ k & \text{if } Z_1 = 0, Z_2 = 0, ..., Z_{k-1} = 0, Z_k = 1 \end{cases} \tag{1.1}$$

We can see that X_1 has the geometric Probability Mass Function (PMF) as follows:

$$\Pr\{X_1 = k\} = p(1-p)^{k-1}, \quad k \geq 1 \tag{1.2}$$

Similarly, we can define the interarrival times $X_2, X_3, ...$, where X_i is the interarrival time between the $(i-1)^{\text{th}}$ and the i^{th} arrivals. Each X_i has the same geometric PMF (1.2), and is statistically independent of $X_1, X_2, ..., X_{i-1}$. Therefore, the Bernoulli process can be characterized by the interarrival times $X_1, X_2, ...$, which form an IID sequence of geometric random variables. We present below an example of a sample path of the Bernoulli process in terms of the arrivals and the interarrival times:

$$\begin{aligned} \{Z_i\} &: 0,1,1,0,0,1,0,1,0,0,0,1,1,... \\ \{X_i\} &: 2,1,3,2,4,1,... \end{aligned} \tag{1.3}$$

Additionally, an important result of a Bernoulli process is the aggregated number of arrivals up to a certain integer time. Let $S_n = \sum_{i=0}^{n} Z_i$ be the random variable representing the aggregated number of arrivals up to and including time instant n. S_n is a sum of n IID Bernoulli random variables. Thus, it has a binomial distribution as follows:

$$\Pr\{S_n = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \geq 1 \tag{1.4}$$

1.4 Poisson Processes

A Poisson process is a widely used stochastic process for modeling the times at which arrivals enter a system [9]. It can be seen as the continuous-time version of the Bernoulli process. However, the arrivals for a Poisson process may occur at arbitrary positive times, and the probability of an arrival at any particular instant is 0. This means that there is no very clean way of describing a Poisson process in terms of the probability of an arrival at any given instant. It is more convenient to define a Poisson process in terms of the sequence of interarrival times $X_1, X_2, ...$, which are defined to be IID. We first define arrival processes as follows:

Definition 1.3 (Arrival process)**.** *An arrival process is a sequence of increasing random variables* $0 < A_1 < A_2 < \cdots$*, where* $A_i < A_{i+1}$ *means that* $A_{i+1} - A_i$ *is a positive random variable, i.e.,* $\Pr\{A_{i+1} - A_i < 0\} = 0$*.*

The random variables $0 < A_1 < A_2 < \cdots$ are called arrival epochs and represent the successive times at which some random repeating phenomenon occurs, such as arrival times or departure times of customers in a system.

The arrival process $\{A_i; i \geq 1\}$ can also be specified by two additional alternatives:

1) The sequence of interarrival times $\{X_i; i \geq 1\}$ as follows:

$$X_1 = A_1 \quad \text{and} \quad X_i = A_i - A_{i-1} \ \text{ for } i \geq 2 \tag{1.5}$$

Similarly, each arrival epoch A_i can be specified in terms of interarrival times as follows:

$$A_i = \sum_{j=1}^{i} X_j \tag{1.6}$$

Henceforth, the joint distribution of the random variables $X_1, X_2, ..., X_k$ for all $k > 1$ is sufficient to specify the arrival process $\{A_i; i \geq 1\}$.

2) The counting process $\{N(t); t > 0\}$, where for each $t > 0$ the random variable $N(t)$ is the aggregate number of arrivals up to and including time t. Note that the $\Pr\{N(0) = 0\} = 1$, which means that arrivals are considered only at strictly positive times.

Lemma 1.1. *The counting process* $\{N(t); t > 0\}$ *for any arrival process has the property that* $N(t_2) \geq N(t_1)$ *for all* $t_2 \geq t_1$ *(i.e.,* $N(t_2) - N(t_1)$ *is a non-negative random variable for* $t_2 \geq t_1 > 0$*).*

Theorem 1.1. *For any given integer* $n \geq 1$ *and time* $t > 0$*, the* n^{th} *arrival epoch* A_n *and the counting random variable* $N(t)$ *are related as follows:*

$$\{A_n \leq t\} = \{N(t) \geq n\} \tag{1.7}$$

Proof. The event $\{A_n \leq t\}$ means that the n^{th} arrival occurs at some epoch $\tau \leq t$. Hence, $N(\tau) = n$

Using Lemma 1.1, we have

$$\tau \leq t \Rightarrow n = N(\tau) \leq N(t)$$

Thus, $\{N(t) \geq n\}$.

Reciprocally, the event $\{N(t) \geq n\}$ means that there exist an integer $m \geq n$ such that $N(t) = m$. This implies that $\{A_m \leq t\}$. Since $m \geq n \Rightarrow A_m \geq A_n$, we have $\{A_n \leq t\}$. □

Corollary 1.1. *An alternative form of Theorem 1.1 comes from taking the complement of both sides of (1.7). We have*

$$\{A_n > t\} = \{N(t) < n\} \tag{1.8}$$

We hereby define the Poisson processes as follows:

Definition 1.4 (Poisson process)**.** *A Poisson process is an arrival process in which the interarrival times have an exponential Cumulative Distribution Function (CDF), i.e., for a given real $\lambda > 0$, each random variable X_i has the density*

$$f_X(x) = \lambda \exp(-\lambda x), \quad x \geq 0 \tag{1.9}$$

The parameter λ is called the rate of the process, and it represents the number of occurred events in a time unit. Thus, for a given time interval of size t, the expected number of arrivals is λt.

1.5 Renewal Processes

Definition 1.5 (Renewal process)**.** *A renewal process is an arrival process for which the sequence of interarrival times is a sequence of positive IID random variables.*

As an arrival process, a renewal process can be specified in three forms:

1) The arrival epochs $A_1, A_2, ...,$

2) The interarrival times $X_1, X_2, ...,$

3) The counting process $\{N(t); t > 0\}$.

These processes are called renewal processes due to the fact that the process probabilistically starts over at each arrival epoch A_n. This behavior exists in many systems such as wireless networks, where the event occurs on random instants at which the system state is probabilistically equivalent to the starting state. The renewal epochs allow distinguishing the study of both the long-term behavior of the process and the behavior within each renewal period.

A renewal process can be seen as a never-ending sequence of random variables $X_1, X_2, ...$ However, the observations of sample outputs may stop after a certain trial n, where n depends on the sample values of $X_1, X_2, ..., X_n$. Such a scenario occurs in several applications. For example, repeating an experiment until certain hypotheses are satisfied, where the number of trials depends on the aggregated outcomes.

The first study of such situations where the number of trials depends on the process outcomes was made by the statistician Abraham Wald, introducing the field of sequential analysis [10]. One of the main results he presented is Wald's equality, which is useful in studying the expected value $\mathrm{E}[A_n]$ of the renewal process at the n^{th} trial.

The rule for stopping trials is specified by a positive integer-valued random variable L, called the stopping trial or stopping time.

Definition 1.6 (Stopping trial)**.** *A stopping trial (or stopping time) L for a sequence of random variables $X_1, X_2, \ldots$ is a positive integer-valued random variable such that for each $n \geq 1$, the indicator random variable $\mathbb{I}_{\{L=n\}}$ is a function of $\{X_1, X_2, \ldots, X_n\}$.*

Stopping times are commonly applied to the study of random processes, particularly the analysis of arrivals to a system or random walks and martingales. Hereafter, for a given stopping trial L, one important challenge is to evaluate the sum $A_L = \sum_{i=1}^{L} X_i$, representing the occurring instant of the L^{th} event (or arrival instant of the L^{th} customer). Wald's equality presented below establishes the relation between the expected value $\mathrm{E}[A_L]$ and $\mathrm{E}[L]$. Thus, it allows finding $\mathrm{E}[A_L]$ if $\mathrm{E}[L]$ can be determined and vice versa.

Theorem 1.2 (Wald's equality)**.** *Let $\{X_n; n \geq 1\}$ be a sequence of IID random variables, each of mean $\overline{X}$. If L is a stopping trial for $\{X_n; n \geq 1\}$ and $\mathrm{E}[L] < \infty$, then the sum $A_L = X_1 + X_2 + \ldots + X_L$ at the stopping trial L satisfies*

$$\mathrm{E}[A_L] = \overline{X}\,\mathrm{E}[L]\,. \tag{1.10}$$

Proof. We have

$$\begin{aligned} A_L &= \sum_{n=1}^{L} X_n \\ &= \sum_{n=1}^{\infty} X_n \mathbb{I}_{\{n \leq L\}} \end{aligned}$$

The event $\{n \leq L\}$ is the complement of $\{L < n\} = \{L = 1\} \cup \{L = 2\} \cup \cdots \cup \{L = n-1\}$. Additionally, the events $\{L = 1\}, \{L = 2\}, \ldots, \{L = n-1\}$ are determined by the random variables $X_1, X_2, \ldots, X_{n-1}$, and hence they are independent of X_n. Therefore, $\{L < n\}$ is independent of X_n.

Since $\{n \leq L\} = \{L < n\}^c$, it follows that $\{n \leq L\}$ is independent of X_n. Therefore, we have

$$\mathrm{E}\left[X_n \mathbb{I}_{\{n \leq L\}}\right] = \mathrm{E}[X_n]\,\mathrm{E}\left[\mathbb{I}_{\{n \leq L\}}\right]$$

Thus,

$$\begin{aligned} \mathrm{E}[A_L] &= \mathrm{E}\left[\sum_{n=1}^{\infty} X_n \mathbb{I}_{\{n\leq L\}}\right] \\ &= \sum_{n=1}^{\infty} \mathrm{E}\left[X_n \mathbb{I}_{\{n\leq L\}}\right] \\ &= \sum_{n=1}^{\infty} \mathrm{E}[X_n]\,\mathrm{E}\left[\mathbb{I}_{\{n\leq L\}}\right] \\ &= \mathrm{E}[X_n] \sum_{n=1}^{\infty} \mathrm{E}\left[\mathbb{I}_{\{n\leq L\}}\right] \end{aligned} \tag{1.11}$$

The interchange of the expected value in (1.11) is valid because we have $\mathrm{E}[L] < \infty$.

Since L is a positive integer-valued random variable, we have

$$\sum_{n=1}^{\infty} \mathrm{E}\left[\mathbb{I}_{\{n\leq L\}}\right] = \sum_{n=1}^{\infty} \Pr\{n \leq L\} = \mathrm{E}[L]$$

Finally, we have

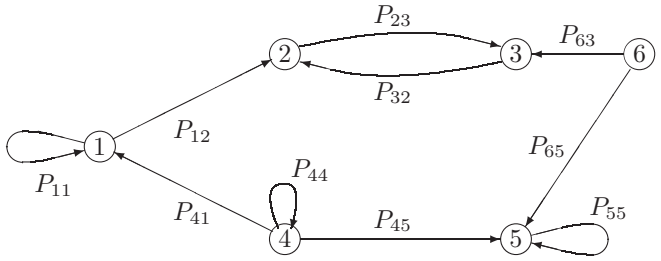
$$\mathrm{E}[A_L] = \mathrm{E}[X_n]\,\mathrm{E}[L] = \overline{X}\,\mathrm{E}[L]\,.$$

□

1.6 Markov Chains

The counting process $\{N(t); t > 0\}$ described above changes its value at discrete instants defined by the arrival epochs, however, $N(t)$ is defined for all reals $t > 0$. Conversely, a Markov chain is a stochastic process at integer values of time $n = 0, 1, \ldots$ At each integer time $n \geq 0$, there is a random variable X_n, called the state at time n. The Markov chain is then represented by the process $\{X_n; n \geq 0\}$. This process is referred to as a discrete-time process and can represent a continuous-time process $\{X(t); t \geq 0\}$ in a way that $X_n = X(t)$ for $n \leq t < n+1$. However, since the process changes its state only at discrete times, it is more practical to view it only at those instants. The set of possible values for each random variable X_n is a countable set $\mathbb{S}$, which can be either finite or infinite, depending on the application.

Definition 1.7 (Discrete-Time Markov Chain)**.** *A Discrete-Time Markov Chain (DTMC) is a discrete-time process $\{X_n, n \geq 0\}$, for which the sample values for each random variable $X_n, n \geq 1$ lie in a countable set $\mathbb{S}$ and depend on the past only through the most recent random variable X_{n-1}. More specifically, the process satisfies the following Markov property:*



(a) Graphical representation

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{16} \\ P_{21} & P_{22} & \cdots & P_{26} \\ \vdots & \vdots & \ddots & \vdots \\ P_{61} & P_{62} & \cdots & P_{66} \end{bmatrix}$$

(b) Matrix representation

Figure 1.1: Representation of a six-state Markov chain.

$\forall n \in \mathbb{N}$, $\forall i_0, ..., i_{n+1} \in \mathcal{S}$,

$$\Pr\{X_{n+1} = i_{n+1} \mid X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = \Pr\{X_{n+1} = i_{n+1} \mid X_n = i_n\} \tag{1.12}$$

The initial state X_0 has an arbitrary probability distribution, and the random variable X_n is called the state of the Markov chain at time n. The set $\mathcal{S}$ is called the state space of the Markov chain.

The quantities $\Pr\{X_{n+1} = i_{n+1} \mid X_n = i_n\}$ are called the *transition probabilities*, and they are in general depending on i, j, n. We note

$$P_{ij}(n) = \Pr\{X_{n+1} = i_{n+1} \mid X_n = i_n\} \tag{1.13}$$

The Markov chain $\{X_n, n \geq 0\}$ is *homogeneous* if the transition probabilities depend only on i and j (not n). That is
$\forall (i, j) \in \mathcal{S}^2$

$$\Pr\{X_{n+1} = j \mid X_n = i\} = \Pr\{X_1 = j \mid X_0 = i\} \tag{1.14}$$

We denote by $P_{ij} = \Pr\{X_1 = j \mid X_0 = i\}$ the probability to go from state i to state j in one step.

The matrix P with elements $\{P_{ij}\}_{i,j \in \mathcal{S}}$ is called the *transition matrix*, and it satisfies the following:

$$\begin{aligned} P_{ij} &\geq 0 \quad \forall i, j \\ \sum_j P_{ij} &= 1 \quad \forall i \end{aligned} \tag{1.15}$$

A DTMC is usually described by a transition graph, as the example in Figure 1.1a. However, if the state space $\mathcal{S}$ is finite, the DTMC can also be represented by the transition matrix, as in Figure 1.1b. Note that in the graphical representation, the directed arc from state i to state j is included only if $P_{ij} > 0$.

The *n-step transition probability* of a Markov chain, noted $P_{ij}^{(n)}$, is the probability that it

goes from state i to state j in n transitions:

$$P_{ij}^{(n)} = P\left(X_n = j \mid X_0 = i\right) \tag{1.16}$$

and the associated *n-step transition matrix* is

$$P^{(n)} = \left\{P_{ij}^{(n)}\right\} \quad \left(P^{(1)} = P\right). \tag{1.17}$$

Theorem 1.3 (Chapman–Kolmogorov equations)**.** *The n-step transition probabilities verify the following system of Chapman–Kolmogorov equations:*

$$P_{ij}^{(m+n)} = \sum_{k\in\mathcal{S}} P_{ik}^{(m)} P_{kj}^{(n)} \tag{1.18}$$

In terms of transition matrices, we have

$$P^{(m+n)} = P^{(m)} P^{(n)} \tag{1.19}$$

and in particular

$$P^{(n)} = P^{(n-1)} P \tag{1.20}$$

so that,

$$P^{(n)} = P^n \quad \text{for } n \geq 1. \tag{1.21}$$

Proof of Theorem 1.3.

$$\begin{aligned}
\Pr\{X_{m+n} = j \mid X_0 = i\} &= \sum_{k\in\mathcal{S}} \Pr\{X_{m+n} = j, X_m = k \mid X_0 = i\} \\
&= \sum_{k\in\mathcal{S}} \Pr\{X_{m+n} = j \mid X_m = k, X_0 = i\} P\left(X_m = k \in \mathcal{S} \mid X_0 = i\right\} \\
&= \sum_{k\in\mathcal{S}} \Pr\{X_{m+n} = j \mid X_m = k\} \Pr\{X_m = k \mid X_0 = i\} \\
&= \sum_{k\in\mathcal{S}} P_{kj}^n p_{ik}^m
\end{aligned}$$

□

The states of a DTMC are classified depending on their accessibility. A state i is called an *absorbing state* if the process can never leave this state after entering it. i.e., $P_{ij} = 0$ for every $j \in \mathcal{S}$.

Definition 1.8 (Irreducible Markov chain)**.** *A Markov chain is said to be irreducible if all states can be reached from any other state in a finite number of transitions. That is,*

$$\forall i, j \quad \exists n \geq 1 \text{ such that } P_{ij}^n > 0$$

Definition 1.9 (Aperiodic Markov chain)**.** *The period of a state i, denoted $d(i)$, is the greatest common divisor (gcd) of those values of n for which $P_{ii}^n > 0$. If $d(i) = 1$, the state is aperiodic, and if $d(i) \geq 2$ or more, the state is periodic. A Markov chain is said to be aperiodic if all of its states are aperiodic.*

Definition 1.10 (Recurrent Markov chain)**.** *For any state i, we define*

$$f_{ii} = \Pr\{X_n = i, \text{ for some } n \geq 1 \mid X_0 = i\}.$$

State i is recurrent if $f_{ii} = 1$, and it is transient if $f_{ii} < 1$.

A Markov chain is said to be recurrent if all of its states are recurrent.

Let f_{ii}^n be the probability that the first return to state i occurs in n transitions:

$$f_{ii}^n = \Pr\{X_1 \neq i, X_2 \neq i, \ldots, X_{n-1} \neq i, X_n = i \mid X_0 = i\}$$

Let f_{ii} be the probability to return to state i in a finite time. We have

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^n$$

Let M_i be the average time to return to i:

$$M_i = \sum_{n=1}^{\infty} n f_{ii}^n$$

State i is:

- transient if $f_{ii} < 1$
- recurrent if $f_{ii} = 1$. Adittionaly, it is:
  - recurrent null if $M_i = \infty$
  - positive recurrent if $M_i < \infty$

Definition 1.11 (Ergodic Markov chain)**.** *A finite-state Markov chain is ergodic if it is irreducible, aperiodic, and positive recurrent.*

Theorem 1.4. *Let P be the transition matrix of an ergodic finite-state Markov chain. Then there is a unique steady-state vector (or a steady-state distribution) π, that is positive and satisfies*

$$\lim_{n\to\infty} P_{ij}^n = \pi_j \quad \forall i, j \in \mathcal{S} \tag{1.22}$$

Proof. Refer to [9]. □

Definition 1.12 (Continuous-time Markov chain). *The stochastic process* $\{X_t, t \geq 0\}$ *taking values in a state space* $\mathcal{S}$ *is a Continuous-Time Markov Chain (CTMC) if:*

$$\Pr\{X_{t+s} = j \mid F_t, X_t = i\} = \Pr\{X_{t+s} = j \mid X_t = i\} \quad \forall i, j \in \mathcal{S}; \forall t \in IR^+, \forall s \in IR_+^*$$

where $F_t = \sigma(X_{s'}; s' \leq t)$ *represents the history of the process* $\{X_t, t \geq 0\}$ *until the instant* t.

The process $\{X_t, t \geq 0\}$ *also verifies the weak Markov property. for any sequence of times* $0 \leq s_0 < s_1 < \ldots s_n < s$ *and any set of states* $i_0, \ldots, i_n, i, j \in S$ *the following property holds*

$$\Pr\{X_{t+s} = j \mid X_s = i, X_{s_n} = i_n, \ldots, X_{s_0} = i_0\} = \Pr\{X_{t+s} = j \mid X_s = i\}. \tag{1.23}$$

1.7 Markov Decision Processes

An MDP is a Markov process with feedback control. That is, a controller takes an action a_t when the Markov process is in the state x_t. This action is fed back to the Markov process and controls its transition matrix. The transition probabilities of the Markov process are then conditioned to the actions taken by the controller. The objective of the controller is to decide a sequence of actions over a time horizon to minimize a cumulative cost function associated with the expected value of the trajectory of the Markov process [11]. MDPs construct a suitable stochastic optimization models in wireless telecommunication networks, providing effective dynamic scheduling and control schemes in various scenarios.

An MDP model consists of the following elements:

1. The state space E, with $x_t \in E$ is the state of the controlled Markov chain at time t.
2. The action space A, with $A_{x_t} \in A$ is the set of admissible action at state x_t. $a_t \in A_{x_t}$ is the action chosen at time t.
3. The transition probabilities. $\Pr\{y \mid x, a\}$ denotes the conditional probability that the system moves to state y when the action a is applied to state x.
4. The instant cost $c(x, a)$ that occurs when the action a is applied to state x.

Definition 1.13 (Markov Decision Process). *Let* $\{x_n, n \geq 0\}$ *be an* E*-valued stochastic process.* $\{x_n, n \geq 0\}$ *is an MDP if at each decision instant n, we have:*

- *The instant cost* $c(x_n, a_n)$ *depends only of the state* x_n *and the action* a_n.
- *The transition probability* $\Pr\{x_{n+1} \mid x_n, a_n\}$ *depends only on the current state and action, i.e.,*

$$\Pr\{x_{n+1} \mid H_n, x_n, a_n\} = \Pr\{x_{n+1} \mid x_n, a_n\} \tag{1.24}$$

with H_n being the history of states and actions up to instant $n-1$.

Additionally, for every $x \in E$

$$\sum_{y \in E} \Pr\{y \mid x_n, a_n\} = 1 \tag{1.25}$$

Definition 1.14 (Decision rule)**.** *The decision rule at time n is a conditional probability law $\pi_n(. \mid H_n)$ defined on A.*

Definition 1.15 (Admissible policy)**.** *An admissible policy π is a sequence of decision rules, i.e.,*

$$\pi = \{\pi_0, \pi_1, \ldots, \pi_n, \ldots\} \tag{1.26}$$

We denote by $\mathcal{P}^a$ the set of admissible policies.

Definition 1.16 (Markovian policy)**.** *An admissible policy is Markovian or memoryless if:*

$$\pi_n(. \mid H_n) = \pi_n(. \mid x_n) \qquad \text{for every } \; n \geq 0 \tag{1.27}$$

We denote by $\mathcal{P}^m$ the set of Markovian policies.

Definition 1.17 (Stationary policy)**.** *An admissible policy is stationary if it is time-independent, i.e., $\pi_n(. \mid H_n)$ does not depend on n.*

We denote by $\mathcal{P}^s$ the set of stationary policies.

Definition 1.18 (Deterministic policy)**.** *An admissible policy is deterministic if the probability measure $\pi_n(. \mid H_n)$ is focused at a single point. That is, a single action a is taken whenever the system is at state x.*

We denote by $\mathcal{P}^d$ the set of deterministic policies.

We have the following classification:

$$\mathcal{P}^d \subset \mathcal{P}^s \subset \mathcal{P}^m \subset \mathcal{P}^a \tag{1.28}$$

The objective of an MDP model is to derive an optimal policy π that minimizes the expected cumulative cost incurred by using policy π up to time N or in an infinite horizon, expressed respectively as follows:

$$J_N^{\delta}(x, \pi) = \mathrm{E}_{\pi}\left[\sum_{n=0}^{N-1} \beta^n c(x_n, a_n) \mid x_0 = x\right], \tag{1.29}$$

$$J^{\delta}(x, \pi) = \mathrm{E}_{\pi}\left[\sum_{n=0}^{+\infty} \beta^n c(x_n, a_n) \mid x_0 = x\right], \tag{1.30}$$

where E_{π} is the expected value under the policy π, and $0 < \beta < 1$ is the discount factor. x denotes the initial state of the system.

1.8 Conclusion

This chapter introduced some main stochastic tools we will deploy in the upcoming chapters. We presented different stochastic processes characterizing a system with random events. The occurrence of such events can be represented by the Bernoulli process in the case of a system is observed in discrete times and the interarrival time of events is constant. Furthermore, the arrivals in a dynamic and continuous-time system can be characterized by a Poisson process in the case of exponentially distributed interarrival times. However, if the time factor is irrelevant, one could characterize the system state using Markov chains and derive the stationary distribution of the system state. The development of a suitable stochastic model requires careful attention to the wireless communication system scenario, its constraints, and developing a mathematical model under appropriate definitions of the considered stochastic framework. Although a stochastic model can be mathematically well defined, it is always important to validate the analytical results through extensive simulations to prove the accuracy of the proposed model in characterizing and evaluating the network scenario conditions. The fast growth of wireless technologies raises the challenge of meeting the requirements of dense and large-scale scenarios, especially in the case of IoT applications. An IoT brings into consideration several constraints on the behavior of the network setup and devices' conditions. Hereafter, many typical mathematical models become not applicable, and it is hence necessary to develop new appropriate analytical frameworks. For example, dense IoT applications encounter a crucial impediment to sharing resources among all connected devices. Thus, to address this issue, most recent communication technologies propose to limit resource access to a small set of devices at a time. That is, reforming a series of small-scale networks, each operating for a limited time. This solution yields reliable connectivity, however, it presents a complex challenge to represent by a mathematical model. Therefore, traditional stochastic frameworks that are based on stationary results are not applicable to such cases. Hence, we contribute in the next chapters to new stochastic approaches to accurately model and evaluate the new wireless communication technologies addressed to dense IoT applications. However, we first present an overview of the new communication technology IEEE 802.11ah that we consider in this thesis, and the related works carried out in the sense of modeling and evaluating the performance of IEEE 802.11ah-based IoT networks.

CHAPTER 2

Overview of the IEEE 802.11ah Standard and Related Research

Contents

2.1 Introduction

Traditional Wi-Fi technologies have established a huge success to become one of the most widely deployed wireless communication technologies in the world. They are originally designed to provide high throughput for small-scale networks where an AP provides connectivity to a few dozen stations over a small range of tens of meters. Nevertheless, many amendments and variants have been proposed to support several applications and their requirements. For example, Time Division Multiple Access (TDMA)-based Medium Access Control (MAC) protocols for long-range networks [12] and for combining Wi-Fi and Ultra Wide Band (UWB)

technologies [13], and Full-Duplex (FD)-based MAC protocols increasing the symmetry between uplink and downlink traffic loads [14]. Other MAC protocols have been proposed to establish the coexistence of Wi-Fi and Long-Term Evolution (LTE) systems [15], and for mesh networks with multi-beam antennas [16].

However, the emergence of the IoT has changed the vision of communication and connectivity. It brings a technological revolution aiming to link anything, anywhere at any time. It hence changes our contemporary life by embedding smart devices and technologies to fully or partially automate all things in the environment [1]. Therefore, higher performance criteria are required to offer sustainable wireless connectivity for IoT applications, including long-distance transmission, large-scale and reliable connectivity, good throughput, and low power consumption [2].

Henceforth, TGah developed the IEEE 802.11ah standard to the first enhanced and optimized Wi-Fi technology to meet the requirements of IoT applications and fill the gap between the current existing IoT communication technologies [4]. The IEEE 802.11ah standard was officially released in 2016, with many remarkable amendments to the legacy Wi-Fi. Although IEEE 802.11ah inherits most of the legacy IEEE 802.11 features, several amendments have been proposed in this new standard to support the general requirements of large-scale dense IoT networks. The novel hierarchical Association Identification (AID) structure allows a single IEEE 802.11ah AP to associate up to 8191 stations. Furthermore, the IEEE 802.11ah standard operates in the unlicensed sub-1GHz frequency band and provides connectivity over a range of up to 1 km, with data rates ranging from 150 Kbps to 78 Mbps.

2.2 Comparison to Existing IoT Communication Technologies

Although IEEE 802.11ah overcomes legacy IEEE 802.11 technology in many aspects to become an optimized Wi-Fi technology for IoT applications, it still has to compete with other existing IoT communication technologies. Thus, this section highlights the novelties and advantages of the IEEE 802.11ah standard compared to existing competitors.

Existing low-power IoT communication technologies can be categorized into two classes: (1) WPAN [17, 18] and (2) LPWAN [19]. Table 2.1 summarizes the characteristics of these technologies. WPAN technologies, such as ZigBee and Bluetooth Low Energy (BLE), provide medium throughput (i.e., up to a few hundred kilobits per second) at a short-range (i.e., tens of meters). On the other hand, LPWAN technologies (e.g., LoRa, SigFox, Narrowband Internet of Things (NB-IoT), eMTC and Wireless Smart Utility Network (Wi-SUN)) provide long-range communications up to tens of kilometers at low throughput rates (few hundred bits to a few megabits per second) [20]. For WPAN technologies, Zig-Bee is based on IEEE 802.15.4 and

Table 2.1: Comparison of IEEE 802.11ah and the existing IoT communication technologies.

Category	Technology	Frequency	Chanel width	Topology	Range	Date rate	Advantages	Disadvantages
WPAN	BLE	2.4 GHz (ISM)	2 MHz	star	30 m	1 - 2 Mbps	medium data rate	short distance
	Zigbee (IEEE 802.15.4)	2.4/Sub GHz (ISM)	5 MHZ	mesh	100 m	10 - 250 Kbps	large coverage	low data rate
LPWAN	LoRa	Sub GHz (ISM)	125/500 KHz	star	20 Km	300 - 500 Kbps	long distance, high mobility	low data rate
	SigFox	Sub GHz (ISM)	100 KHz	star	50 Km	100 bps	long distance	low data rate
	NB-IoT (5G)	Sub GHz (licensed)	180 KHz	cellular	15 Km	250 bps	long distance	low data rate
	eMTC (5G)	Sub GHz (licensed)	1.4 MHz	cellular	N.A.	1 Mbps	long distance, high mobility, low latency, high reliability	low data rate
	Wi-SUN (IEEE 802.15.4)	Sub GHz (ISM)	200 KHz - 1.2 MHz	mesh	1000 m	50 Kbps - 2.4 Mbps	medium date rate	medium distance
	IEEE 802.11ah	Sub GHz (ISM)	1–16 MHz	star	1000 m	150 Kbps - 78 Mbps	high date rate	medium distance

uses a mesh topology to provide connectivity to a large number of devices in a wide coverage area, while BLE reduces energy consumption. Regarding LPWAN technologies, NB-IoT and eMTC are developed as 5G technologies for IoT, operating in licensed frequency bands, while the others operate in Industrial, Scientific and Medical (ISM) bands. SigFox provides the longest transmission range, while LoRa and eMTC support high mobility.

Due to the short range of WPAN and the very low throughput of both WPAN and LPWAN, the use of these two technologies exhibits severe limitations to some relevant IoT scenarios. Therefore, there is still a need for a low-power IoT communication technology capable of offering higher transmission rates over long transmission ranges. Therefore, to meet the IoT requirements and fill the gap between WPAN and LPWAN, TGah has developed a new standard called IEEE 802.11ah (also known as Wi-Fi HaLow) [4], as an LPWAN technology that offers a trade-off between throughput and range coverage. IEEE 802.11ah systems operate in the unlicensed sub-GHz frequency bands below 1 GHz. A single AP may provide connectivity to a maximum of 8191 low-power devices (nodes) at data rates ranging from 150 Kbps to 78 Mbps over a transmission range of up to 1 Km [20]. Figure 2.1 presents the position of the IEEE 802.11ah standard among the existing WPAN and LPWAN technologies.

The main competitors of IEEE 802.11ah in IoT applications are IEEE 802.15.4 and 5G cellular technologies, as seen in Table 2.1. However, the studies in [21] and [22] established that IEEE 802.11ah outperforms IEEE 802.15.4 in terms of association time, coverage range, throughput, and latency. Additionally, while comparing IEEE 802.11ah With BLE, it is found in [21] that IEEE 802.11ah yields higher throughput and lower latency than BLE. The 5G cellular networks are also competing to lead IoT communications, and Ericsson expects that Several IoT applications will be deploying cellular networks in the future [23]. Even though NB-IoT and eMTC introduce high data rates and large coverage to IoT applications, it is very unlikely to ignore the other LPWAN technologies due to their remarkable features, as highlighted in Table 2.1. Moreover, in [24], IEEE 802.11ah is introduced to increase the capability of coping with massive access attempts in 5G massive Machine Type Communications (mMTC) networks. It is

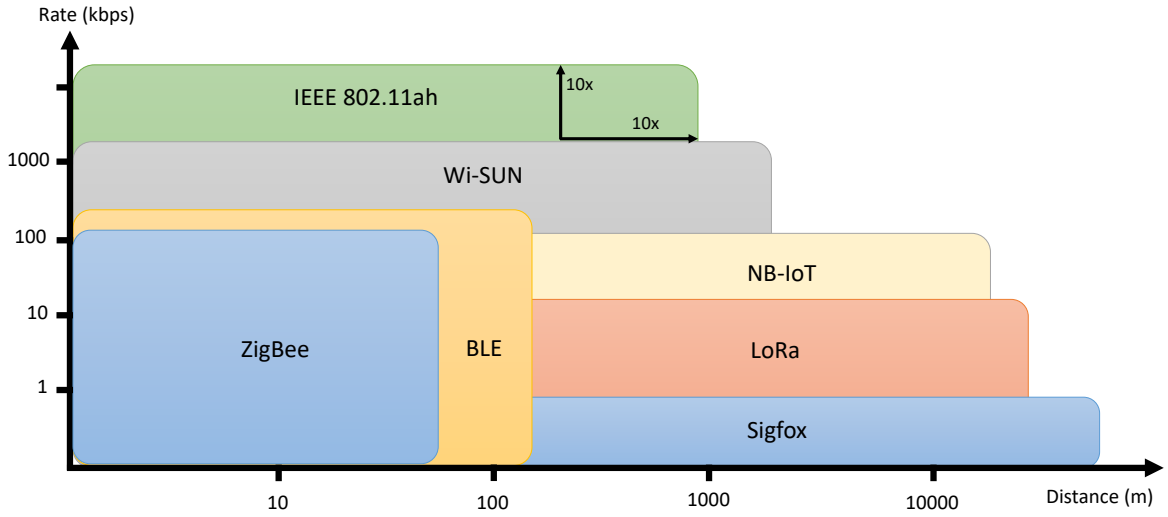


Figure 2.1: Position of the IEEE 802.11ah standard among the existing WPAN and LPWAN technologies.

found that it significantly enhances the access delay besides the advantages of being a low-cost, high-access capacity, and handy deployment technology.

In order to be able to provide the capabilities mentioned above, the IEEE 802.11ah introduces novel mechanisms such as RAW, Traffic Indication Map (TIM) segmentation, and Target Wake Time (TWT), allowing better management of dense scenarios with reliable connectivity and low power consumption for battery-constrained devices. Such features make IEEE 802.11ah an interesting and promising technology for several IoT use cases [25] as we list in the following:

- Smart metering and monitoring: IEEE 802.11ah can offer reliable connectivity to thousands of sensors distributed in a wide area, with the periodic transmission of short packets to the AP. This includes smart meters (gas, water, and power consumption), smart grids, environmental and agricultural monitoring (temperature, humidity, wind speed, water level, pollution, animal condition, forest fire detection, etc.), automation of industrial process (petroleum refinement, iron and steel, pharmacy), indoor healthcare/fitness system (blood pressure, heart rate, weight), elderly care system (fall detection, pill bottle monitor).

- Backhaul aggregation: Since IEEE 802.15.4g wireless devices are battery-constrained and limited to very low transmission range and data rates, IEEE 802.11ah can be implemented as a backhaul connection between those devices and remote servers with long transmission range and higher data rates. Additionally, IEEE 802.11ah provides better channel access than IEEE 802.15.4g under similar scenarios [26].

- Extended range hotspot and cellular offloading: Although legacy Wi-Fi such as IEEE 802.11n/ac provides similar or higher data rates than mobile networks such as LTE, it is

not practical to offload traffic in outdoor scenarios due to the short transmission range. Alternatively, IEEE 802.11ah can take this role and extend the hotspot range thanks to its long transmission range and high throughput.

2.3 IEEE 802.11ah Features

Many enhancements are proposed in the IEEE 802.11ah amendment in order to construct an adequate Wi-Fi technology to address the major requirements of IoT applications. This section highlights some of the main features of the IEEE 802.11ah standard in the Physical (PHY) and MAC layers.

2.3.1 PHY Layer Overview

The IEEE 802.11ah inherited its PHY layer from the IEEE 802.11ac standard and defined an Orthogonal Frequency-Division Multiplexing (OFDM) PHY in the sub-1GHz bands, with channels 10 times narrower than those in IEEE 802.11ac. Among the supported 1, 2, 4, 8, 16 MHz channel bandwidths, only 1 and 2 MHz channels are mandatory. Using sub-1GHz frequency bands (e.g., 902-928 MHz in North America, 863-868 MHz in Europe, and 755-787 MHz in China) along with narrow bandwidth allows the IEEE 802.11ah to improve coverage range up to 1 km, and to significantly lower the energy consumption than traditional Wi-Fi technologies, which operate in the 2.4 and 5 GHz bands. For each channel width, IEEE 802.11ah deploys different sets of Modulation and Coding Schemes (MCSs) with different properties and transmission rates from the IEEE 802.11ac standard, as presented in Table 2.2. MCS9 is not available in 2 MHz as it is the same case for 20 MHz in IEEE 802.11ac. In addition, the IEEE 802.11ah defines a new MCS10 for 1 MHz channels as a 2x repetition mode of MCS0, providing a longer transmission range with better reliability. It has been found in [27] that the transmission range exceeds 1 km when MCS0 is deployed in an outdoor scenario with a transmission power higher than 200 mW and 1 MHz channel.

2.3.2 MAC Layer Overview

2.3.2.1 Fast Association and Authentication

During the association stage with the AP, stations use the DCF mechanism for channel access. Thus, linking more than 8000 stations may take a very long time due to extensive collisions of the association and authentication messages. To overcome this issue, IEEE 802.11ah proposes the following two effective mechanisms for reducing the association delay by limiting the number of stations that can send simultaneous authentication requests:

Table 2.2: IEEE 802.11ah MCSs for a single spatial stream and normal OFDM symbol.

MCS Index	Modulation	Coding rate	Data rate (Mbps)				
			1 MHz	2 MHz	4 MHz	8 MHz	16 MHz
0	BPSK	1/2	0.3	0.65	1.35	2.925	5.850
1	QPSK	1/2	0.6	1.30	2.70	5.850	11.700
2	QPSK	3/4	0.9	1.95	4.05	8.775	17.550
3	16-QAM	1/2	1.2	2.60	5.40	11.700	23.400
4	16-QAM	3/4	1.8	3.90	8.10	17.550	35.100
5	64-QAM	2/3	2.4	5.20	10.80	23.400	46.800
6	64-QAM	3/4	2.7	5.85	12.15	26.325	52.650
7	64-QAM	5/6	3.0	6.50	13.50	29.250	58.500
8	256-QAM	3/4	3.6	7.80	16.20	35.100	70.200
9	256-QAM	5/6	4.0	–	18.00	39.000	78.00
10	BPSK	1/2 with 2x repetition	0.15	–	–	–	–

BPSK: Binary Phase-Shift Keying.
QPSK: Quadrature Phase-shift Keying.
QAM: Quadrature Amplitude Modulation.

1. Centralized Authentication Control (CAC) [28]: The AP sets a threshold and broadcasts it to all stations through the beacon frame. On the other hand, each station generates a random value from $[0, 1022]$ after its initialization and competes to send its authentication request to the AP if its chosen random value does not exceed the threshold from the beacon frame. Otherwise, it waits until it receives the next beacon frame and checks the new threshold. Note that the threshold should be dynamically configured by AP in order to accelerate the association process.

2. Distributed Authentication Control (DAC) [29]: The AP divides each beacon interval into several sub-intervals with equal duration, called Authentication Control Slots (ACSs). Each station selects a BI and an ACS to send its authentication request. If it does not succeed, it selects again the next i^{th} BI and i^{th} ACS for a new attempt, where the values of i and j are selected following the truncated binary exponential backoff mechanism.

2.3.2.2 Hierarchical Station Organization

Since the legacy IEEE 802.11 standard can only associate up to 2007 stations due to limited size of AID field (1-2007), the new IEEE 802.11ah amendment needs to overcome such a limitation to support large-scale networks. Hence, a new hierarchical mechanism is proposed to organize stations by 13-bit AIDs following a four-level structure [4], as depicted in Figure 2.2. The new AID structure is organized in a way that the stations are divided into up to four pages,

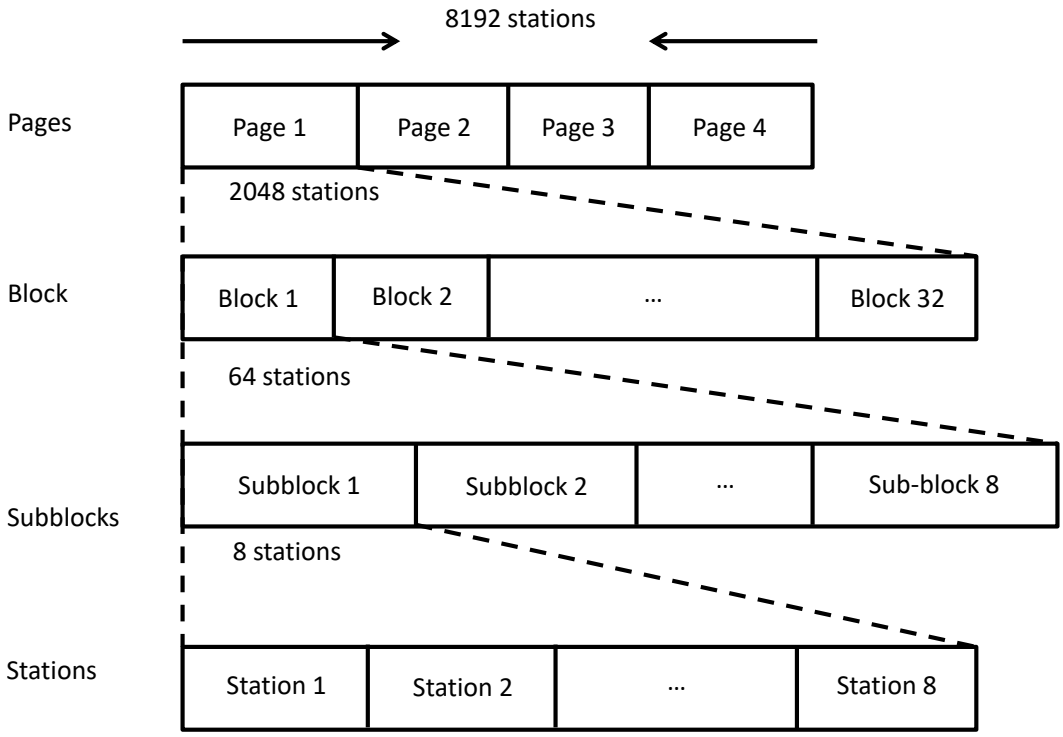


Figure 2.2: IEEE 802.11ah AID hierarchy

each containing up to 32 blocks. Every block consists of 8 sub-blocks of 8 stations each. The numbers of pages and blocks are variable and can be configured by network operators. Hence, a 13-bit AID is assigned to every associated station, which is decomposed of 2-bit pages, 5-bit blocks, 3-bit sub-blocks, and 3-bit stations. Therefore, a single IEEE 802.11ah AP provides connectivity to up to $2^{13} - 1 = 8191$ associated stations.

The four-level AID structure allows grouping stations with the same characteristics and effectively reducing overhead when referring to stations. So an AID is assigned according to the station type, the power management mode, the location, the station's traffic, etc [25]. When a station changes its characteristics, the AID may be reassigned following a dynamic AID assignment handshake between the AP and the station. Additionally, AID-based grouping of stations that are located near each other enhances the gain of channel access using the RAW mechanism by restricting channel access to stations that are in the listening range of each other, hence avoiding the overlapping transmissions of hidden nodes.

2.3.2.3 Medium Access

As an amendment to legacy IEEE 802.11 technology, IEEE 802.11ah provides medium access to stations through the protocols of DCF and EDCA. However, it will not be practical to allow all stations to contend for channel access in such a case of IEEE 802.11ah-based networks with thousands of associated stations. The collision chains will drastically degrade the network performance and drain out the devices' batteries. To solve this problem, TGah introduces the

new RAW mechanism for wireless medium access. The RAW mechanism divides each BI into sub-intervals, called RAWs, and each RAW consists of several periods, called RAW slots. The stations are first assigned to RAWs, then distributed among the available RAW slots following a round-robin rule. To this end, only the stations in each RAW slot are allowed to contend for channel access during the duration of this RAW slot. Since this mechanism is essentially considered in the upcoming chapters, we present it in more detail in Section 2.4

2.3.2.4 Group Sectorization

Due to the long-range coverage of an IEEE 802.11ah AP and a large number of associated stations, the problem of hidden nodes is very crucial for IEEE 802.11ah-based networks. To address this issue, TGah developed a group sectorization mechanism to organize stations. This mechanism combines space- and time-division multiplexing, where the AP divides the coverage area into several geographical sectors, each containing a subset of stations. Each station obtains the index of the sector it belongs to during the association procedure. All stations in the same sector listen to each other transmissions. Additionally, the AP splits the channel time into Sector Intervals (SIs) and transmits a beamformed beacon at the beginning of each SI with information about the sector that is allowed for channel access during this SI. Group sectorization can be seen as a simplified RAW version based only on location-aware grouping. It improves the throughput by eliminating the hidden nodes problem and the energy efficiency as each station can be in sleep mode during SIs other than the one it belongs to. Nevertheless, the two methods of RAW and group sectorization combined together will significantly improve the flexibility of the network. During our studies in the following chapters, we omit the hidden nodes problem by assuming that the stations assigned to a given RAW are from the same sector and they can hear each other signals.

2.3.2.5 Relay Forwarding Mechanism

To support large-scale IoT applications with efficient data rates, IEEE 802.11ah deploys relays to extend the transmission range between the AP and stations. A relay forwards frames between its associated stations and the AP, called root AP. Hence, it extends the distance between stations and the AP, especially in scenarios with the presence of obstructions. Relays are also contributing to saving energy consumption for stations. Since the relay is close to its associated stations, the latter can transmit their frames with lower power and higher data rates, shortening the transmission time and hence the awake time of stations. A relay logically consists of two components: a relay AP that is associated with the stations and a relay station associated with the root AP. During downlink transmissions, the root AP transmits the packets to the relay station, which forwards them to the relay AP. The latter transmits the packets to the addressed stations. The operation is reversed for uplink transmissions. Although the multi-hop concept has already been introduced in the mesh network of the IEEE 802.11s standard [30],

the mesh forwarding protocols produce a heavy routing framework which is not practical for battery-constrained IoT devices. Therefore, IEEE 802.11ah introduces this concept into IoT centralized networks using the relay forwarding mechanism, limited to a two-hop link between the root AP and the stations.

2.4 The IEEE 802.11ah RAW Mechanism

Since connecting thousands of stations using basic IEEE 802.11 MAC mechanisms is impossible, the IEEE 802.11ah standard introduces a new MAC mechanism named RAW. The RAW mechanism aims to alleviate collisions and improve the performance of dense IoT networks where a significant number of devices are competing for channel access simultaneously. It can be viewed as a combination of deterministic and stochastic media access strategies. Figure 2.3 depicts a schematic representation of the RAW mechanism functionality. Within each beacon interval, the AP allocates one or more RAWs, where each RAW is allocated to a group of stations. Only its designated stations are allowed to contend for channel access during a RAW period. In contrast, all stations can access the channel in the shared channel time. The AP is responsible for allocating RAW periods within the beacon interval and assigning a group of stations to each RAW. Thus, it broadcasts this information through the preceding beacon frame using the RAW Parameter Set (RPS) element. The RPS element indicates the stations assigned to RAWs, the configuration of RAWs, and the starting time of each RAW.

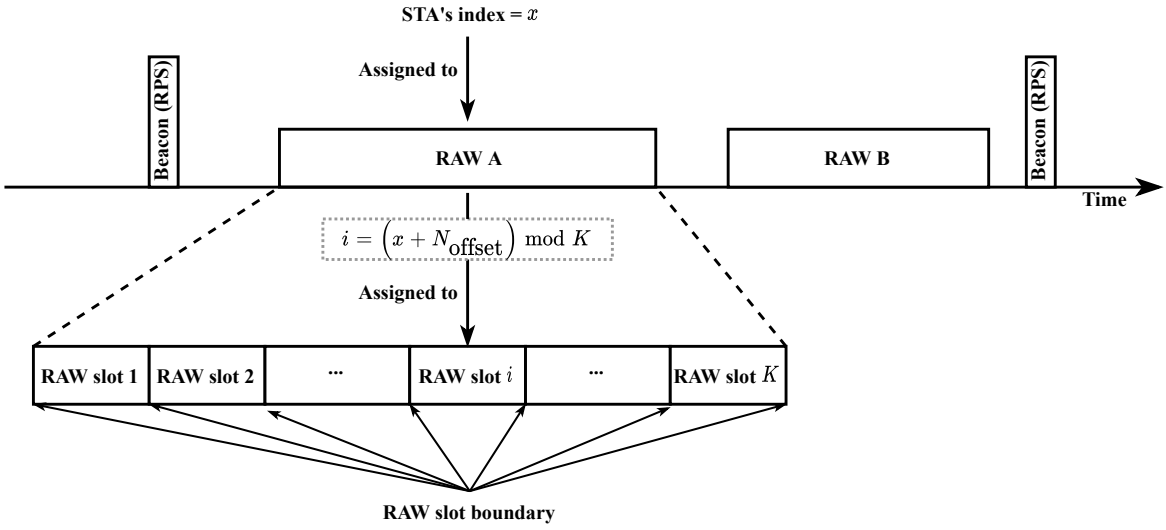


Figure 2.3: Schematic representation of the RAW mechanism.

To reduce the contention even more, each RAW can be divided into one or more short periods, referred to as RAW slots. The stations assigned to a RAW are also split across the RAW slots using a round-robin assignment policy. Each station is allowed to contend for channel access only during its designated RAW slot and prohibited from contending for medium access

during the other RAW slots of the RAW. The number of RAW slots within each RAW and their duration are also included in the RPS element. A station determines the index of the RAW slot i_{slot} in which it is allowed to begin competing for medium access based on the following mapping function [4]:

$$i_{slot} = (x + N_{\text{offset}}) \mod K, \tag{2.1}$$

where x is the position index of the station among others in the RAW, N_{offset} represents the offset value in the mapping function, and K is the number of RAW slots in the RAW.

During each RAW slot, the assigned stations use DCF or EDCA mechanism to access the channel. To make that compatible with the RAW mechanism, each station maintains two backoff function states to manage medium access inside and outside its assigned RAW slot [4]. The first backoff function is used outside RAW slots during the shared channel time, where all stations contend for channel access, while the second is used inside. At the start of each RAW period, every station stores and suspends the state of its first backoff function, then later, it restores and resumes the backoff timer at the end of the RAW period. The second backoff function is used inside RAW periods, where each station initiates the backoff timer at the beginning of its designated RAW slot, then terminates and discards the backoff timer state at the end of the RAW slot.

2.5 Research on IEEE 802.11ah-based Networks

Since we mainly focus on channel access and network resources, we mostly consider in this section the works related to MAC layer in the IEEE 802.11ah standard. As detailed above, IEEE 802.11ah deploys the RAW mechanism, where the channel timeline is composed into RAW periods, and hence the performance of the entire network is proportional to the performance of these periods.

Since the IEEE 802.11ah standard does not provide any scheme or strategy for configuring the RAW parameters, adequate optimal RAW configurations are henceforth required to achieve the highest gain of this mechanism. Hereafter, several works focus on modeling and evaluating RAW-based networks in order to tune up the impacting parameters.

2.5.1 Performance Evaluation

The research community has been interested in investigating the advantages and challenges of the new IEEE 802.11ah standard. To highlight the potentials of the IEEE 802.11ah standard, several works evaluated its performance by studying the impact of several parameters related to the MAC layer based on the RAW mechanism.

Recent performance studies have concluded that the IEEE 802.11ah standard outperforms other prominent alternatives, e.g., the IEEE 802.15.4, in terms of association time, throughput, delay, and coverage range [22]. The amendments introduced into the IEEE 802.11ah standard make this new technology very relevant to the development of IoT network infrastructure. Several studies have already examined the key features presented in the standard, aiming to illustrate the advantages and challenges in the design of the proposed mechanisms [20, 31, 27, 25, 32]. Domazetović et al. [33] examined the performance of IEEE 802.11ah systems in terms of an achievable range. Their results proved that IEEE 802.11ah systems are very promising solutions for different IoT applications. Additionally, the analysis in [20] showed that the implementation of an IEEE 802.11ah-based IoT infrastructure may prove less costly than a typical scenario based on the legacy IEEE 802.11 technology.

Zhao et al. [34] evaluated the energy efficiency of the IEEE 802.11ah MAC layer based on a system-level simulator. They showed that segmenting the channel into RAW slots improves the overall energy efficiency of the network. Tian et al. [35] studied the implementation of the RAW mechanism in the ns-3 simulator, which reveals the effectiveness of the RAW in boosting throughput and latency and minimizing power consumption in dense scenarios. Their findings emphasize the necessity to tune up the RAW parameters and develop smart RAW configuration schemes in order to maximize the network performance under dynamic conditions. Similar efforts have been reported in [36] and [37], where the authors proved the importance of the RAW mechanism in high contention scenarios by evaluating the network performance in terms of throughput and energy efficiency. Their results show that the RAW mechanism improves the overall performance in dense networks. They also have pointed out the need for tuning up the RAW parameters in order to make the best use of the network resources, i.e., the maximum achievable performance. Šljivo et al. [38] carried out a performance evaluation of the RAW mechanism using simulations based on an implementation of the IEEE 802.11ah standard in the ns-3 network simulator. Their results showed that the IEEE 802.11ah network could support up to 20 streaming Internet Protocol (IP)-cameras operating at a data rate of 160 Kbps, and ten 255-Kbps IP-cameras over a range of 200 m.

2.5.2 Mathematical Models

During the RAW period, the assigned stations compete for channel access using DCF or EDCA protocols. Hence, the developed RAW models are mostly based on the backoff process with consideration of the RAW characteristics.

In [39, 40], the authors established the performance of the Group-Synchronized Distributed Coordination Function (GS-DCF), which is the same idea in the RAW mechanism. They developed a mean value analysis approach to model the contention within a RAW slot and evaluate the saturated throughput for both cases where the RAW slot boundary-crossing is enabled or disabled. They also studied the handover between RAW slots and showed that the

throughput faces a fluctuating behavior in the case of a non-cross slot boundary. They also highlighted that GS-DCF is a promising solution capable of significantly improving the throughput in dense networks. Chang et al. [41] proposed a load-balanced grouping algorithm that distributes sensors among groups according to their traffic demands. However, they assumed that the number of groups is fixed and given, while this is a crucial parameter that requires to be optimized. The handover between RAW slots is another issue to be addressed in the configuration and operation of the RAW mechanism. In the context of the IEEE 802.11ah standard, the handover between RAW slots refers to the fact that the stations associated with a RAW group may cross the boundary between two consecutive RAW slots. Zheng et al. [40] presented an analytical study of the RAW operation considering the cross-slot boundary condition. Qutab-ud-din et al. [42] extended this work and evaluated several holding schemes to prevent the crossing of the boundary slot. Nawaz et al. [43] extended the work in [40] with a model where the RAW frame is divided into two sub-frames, with a different RAW slot duration in each sub-frame. However, they did not provide details on the length of the RAW slots and assumed an infinite retransmission limit scenario as in [5]. Both of these features are highly relevant to evaluate the benefits of the RAW scheme. Raeesi et al. [37] established an analytical model to evaluate the throughput and energy efficiency of a RAW-based network and extended it to the case of multi-access points in [36]. However, their model assumes known collision and error probabilities which are not obvious to determine in such a case of RAW-based channel access.

The proposed models in these works above ignore the limited duration of the RAW slot and hence provide results that are valid only for an infinite RAW slot, which is not the case for the IEEE 802.11ah standard. Ali et al. [44, 45] proposed an analytical model to evaluate the performance of the RAW scheme taking into account the time constraint. However, they assumed that the beacon interval is fully occupied with one RAW, which is divided uniformly into several RAW slots. Besides, they considered pre-configured RAW slots where each RAW slot duration and the number of participating stations are already defined. In [46, 47], the authors presented a mathematical model for the contention within the RAW using a four-dimensional Markov chain framework. The Markov chain is depicted by the number of contending stations and the total number of occurred virtual slots in three forms: idle, busy with the transmission, and busy with collision. However, the model shows higher complexity with no closed-form results as there are no explicit definitions of state probabilities at the absorbing states, which is the main objective of the proposed model. However, the proposed model estimates throughput and energy consumption for a RAW-based network with the feature of boundary-crossing. Their results show that enabling boundary-crossing increases the throughput but with the cost of more energy consumption, as the stations in each group will awake at the beginning of their allocated RAW slot period and may have to wait for a transmission from the previous group to be completed. In [48], the authors proposed a mathematical model for EDCA within RAW without retries, i.e., a station gets only one attempt to transmit its packet. Although such a scenario would have very limited applications, it is an easy to calculate model compared to the normal EDCA with retries as the authors depicted while comparing it with the model in [47]. A

similar model with no retries is presented in [49] with an additional assumption that a RAW slot can host at most only one transmission attempt, which is an even more limited scenario for real network applications. This approach is not applicable in the case of non-saturated traffic, as the channel access probabilities change when a station leaves the contention. Other mathematical models for EDCA within RAW with no retransmissions are presented in [48] where a station gets only one attempt to transmit its packet, and in [49] where a RAW slot can host at most only one transmission attempt. In [50, 51, 52], Wang et al. carried out an analysis based on the probability theory [50, 51] and the Markov chain [52] to derive the energy efficiency within a RAW. They then proposed a channel access window algorithm to enhance the energy efficiency of the up-link medium access mechanism.

Since IoT applications mainly consist of sensors, it is more common to have networks with unsaturated traffic in real scenarios. The stations require channel access periodically when they have new data to transmit, where the packet arrival is usually considered to be following a Bernoulli or Poisson distribution.

The authors in [53] developed a Markov chain model to obtain the distribution of the time needed for an arbitrary station to transmit a packet successfully. They assume that each station has one packet at most to transmit during the RAW slot period. Their model also examines the distribution of the time for all the stations sharing a given RAW slot. However, their model application is very limited since they focus on the performance of a target RAW slot as opposed to the overall RAW structure. The same model is deployed in [54] to evaluate the energy consumption within the RAW under the assumption that every station has only one packet to transmit in each beacon interval. Their results show that dividing the RAW into more RAW slots yields better energy efficiency but decreases the latency. This model is based on the absorbing states of the Markov chain, which are defined by the constraint of the total duration of events reaching the RAW slot limit. Thus, it is complex and challenging to define closed-form results to evaluate the network. The authors in [55] considered a scenario of sensors with heterogeneous traffic stations and established a regression-based analytical model to accurately estimate the contention success probability. Tian et al. [56, 57, 58] developed a surrogate model for RAW-based channel access with heterogeneous stations in order to depict more realistic scenarios. A surrogate model is an efficient mathematical representation of a complex system. It is based on supervised learning (e.g., Kriging or neural networks), and is effective for tasks with large inputs [59]. An accurate model can be trained with very few sampled data points by entering realistic simulation results into the surrogate modeling toolbox. The proposed model allows evaluating the throughput and energy consumption given specific network conditions [56]. Additionally, it allows to accurately predict RAW performance under a given RAW configuration in heterogeneous networks [57]. The extended version of this model allows for determining the optimal RAW configuration for both throughput and energy efficiency in real time through multi-objective optimization.

Considering unsaturated traffic conditions with packet arrival following a Bernoulli dis-

tribution, Ometov et al. [60] developed a Markov chain model to evaluate the network performance in terms of throughput, delay, and power consumption. In addition, a Poisson arrival distribution and channel impairments are considered in [44], where the authors proposed a Markov chain based model to evaluate the network throughput. The results show that a higher arrival rate of packets significantly degrades the network performance. Furthermore, Ali et al. [61, 45] established a Markov chain and M/G/1 queuing model to evaluate the RAW performance with EDCA on differentiated Quality of Service (QoS). Their findings reveal that the RAW mechanism is suitable for low load IoT applications and supports the coexistence of heterogeneous traffic. They also show that the RAW scheme supports a large number of QoS and non-QoS nodes in unsaturated conditions. However, increasing the number of nodes and packet arrival rate results in having some non-QoS nodes unable to transmit all of their packets during the RAW slot period. Nevertheless, the proposed models in [60, 44, 61, 45] are based on Markov chains which provide steady-state results. These models may not necessarily be accurate as the channel access using the RAW mechanism is constrained by the short durations of the RAW slot periods.

2.5.3 RAW Configuration

The configuration of the RAW mechanism depends on several parameters, such as assigned stations and their locations, the payload size, the channel model, and the capture threshold. Therefore, an accurate mathematical model with closed-form results is the best solution for such a task. It permits predicting the RAW performance given the input values of all related parameters and hence deriving the best configuration optimizing the entire network performance. Several researchers have proposed analytical models to study and evaluate the IEEE 802.11ah standard [25, 3]. Due to the time-limited contention introduced in the RAW mechanism, analytical models based on steady-state results such as Bianchi's [5] are not applicable for this standard. Although several modified and enhanced models have been proposed [62], they are still all based on stationary results. The limited duration of a RAW slot prevents the contention from reaching the stationary state [63]. Therefore, new innovative approaches have been required to model and evaluate the IEEE 802.11ah standard.

Although the studies mentioned above model and evaluate the RAW mechanism and propose station grouping algorithms to optimize RAW parameters, they still have some shortcomings. Some models have a limited application due to the assumption that each station has only one packet to send during the RAW. Others do not consider critical RAW parameters, such as the number of groups, the number of stations in each group, and the duration of a RAW slot allocated to each group in the grouping scheme. They further assume that some of these parameters are fixed and given inputs. Some other studies do not take into account the channel waste time during the backoff slots and holding periods of the RAW slots.

Differently from the works mentioned above, we developed models based on a new ap-

proach in our previous works [63, 64, 65]. We deployed the renewal theory in which a counting process tracks transmissions up to the end of the RAW slot [63]. This approach allows us to predict the number of occurred events within the RAW slot period in terms of time and hence evaluate the gain of the time-limited contention. In [64], we proposed an analytical framework based on renewal theory to evaluate the RAW throughput and then optimize the RAW configuration to maximize the throughput in both cases of an ideal channel and a channel with transmission errors. We further developed an accurate model for an IEEE 802.11ah-based network, operating under a Rayleigh-fading channel with capture effect enabled at the AP [65]. These models allowed us to evaluate the network performance and study the impact of different parameters in both RAW configuration and network conditions.

2.6 Conclusion

We presented in this chapter an overview of the IEEE 802.11ah standard and the related research on this new wireless communication technology. Being the first Wi-Fi technology addressed to IoT applications, the IEEE 802.11ah gained considerable attraction from researchers. Its key features are proposed at the level of the MAC layer, mainly the new channel access mechanism, RAW. However, the standard describes how the RAW mechanism works without providing a concrete scheme or strategy for its configuration. Hence, many works established proposals to evaluate the IEEE 802.11ah-based networks and propose schemes and algorithms to tune up its configuration. Nevertheless, the time-limited contention for channel access in such networks presents a significant challenge, as the contention ends before the system reaches a state that can be represented by stationary results provided by approaches such as Markov chains. Although some works introduced new Markov models based on adsorbing states, the randomness of the system makes such a solution very complex and impractical for real-time usage. Henceforth, we propose in the next chapter a new stochastic frameworks to model to model a network with time restricted access. This analytical framework provides closed-form and practical results that we validate throughout simulations and convergence to stationary state results.

CHAPTER 3

Modeling and Evaluation of Time-limited Channel Access

Contents

3.1 Introduction

One of the main challenges of modern wireless technologies is the wide range and dense networks, especially for IoT applications. These scenarios raise hidden node issues due to large

distances between stations and degrade the efficiency of the legacy random channel access methods because of the high contention. Therefore, the new IEEE 802.11 standards integrate mechanisms for deterministic channel access to meet such requirements. This solution allows the scheduling of channel time into periods, where each period is reserved for a small-scale network with random access. First, the AP can easily eliminate the hidden nodes problem by assigning to each period only the stations that are located in the same region where they are all listening to each other's transmissions. Second, the deterministic channel allocation mechanisms limit channel access to only a small group of stations, decreasing collisions and enhancing network performance.

The periodic channel reservation feature is introduced in several standards under different names and operating schemes [66]. We list some of them as follows:

- IEEE 802.11s: The IEEE 802.11s standard defining the Wi-Fi mesh technology introduces a new channel access mechanism called Mesh coordination function Controlled Channel Access (MCCA) [67]. This mechanism is optional and is used with the mandatory EDCA mechanism. MCCA permits a station (called owner) to reserve time intervals during which only this station is allowed to transmit its data to another station (called responder) [68]. During these time intervals, the neighboring stations of both the owner and the responder are forbidden from accessing the channel.

- IEEE 802.11aa: To improve real-time audio/video streaming, the IEEE 802.11aa standard proposes to enhance the functionality of the Hybrid coordination function Controlled Channel Access (HCCA) mechanism, proposed in the IEEE 802.11e standard. The HCCA allows the AP to allocate a Transmission Opportunity (TXOP) for a single station. However, because of the dense deployment of APs, the networks overlap, and two APs may allocate the same TXOP for stations, which results in collisions. Therefore, the IEEE 802.11aa standard introduces a mechanism called HCCA TXOP Negotiation [69]. This mechanism allows the APs to negotiate the reserved time intervals for TXOPs and defer from each other.

- IEEE 802.11ad/ay: The IEEE 802.11ad and IEEE 802.11ay standards, operating on 60 GHz millimeter-wave (mmWave), are very promising Wireless Local Area Network (WLAN) technologies for ultra-high-speed communications [70]. Operating on a 60 GHz band attenuates the signal strength and shortens the coverage of the AP. To enhance channel resources utilization, an IEEE 802.11ad/ay allows a tagged station to operate as a central coordinator. This station splits the channel time into BIs and divides each BI into Contention-Based Access Periods (CBAPs) and Service Periods (SPs). The scheduling information is broadcasted via beacons. During SPs, only assigned stations are allowed to access the channel, whereas all stations can randomly access the wireless medium during CBAPs.

- IEEE 802.11ah: Aiming to offer connectivity to over 8000 stations within a range of up to 1 km, the IEEE 802.11ah introduces the RAW mechanism, explained previously in Chapter 2. The RAW allows to limit medium access in such a dense scenario, where the AP splits the channel time into RAW periods and assigns a set of stations to each RAW [71].

- IEEE 802.11ax: The IEEE 802.11ax standard, marketed as Wi-Fi 6, introduces the Quiet Time Period (QTP) mechanism that allows a given station to request a time interval from the AP, during which only this station is allowed to use the wireless medium [72, 73]. The QTP mechanism can also be implemented at the level of APs in order to eliminate interference between overlapping networks. In this case, a centralized controller is deployed to manage and operate the QTP mechanism for APs, where the channel time is split into periods, each allocated to a single network.

Among all the standards mentioned above, we consider in our study the IEEE 802.11ah standard, which is the main candidate for IoT applications, constructing ultra-dense networks. Since the IEEE 802.11ah MAC is based only on the legacy DCF and EDCA protocols [25], a large number of connected devices will have a negative impact on the access mechanism performance. To address this problem, the RAW mechanism proposed in the IEEE 802.11ah limits the concurrent channel access to a reduced number of stations at a time. The scheduling process takes place at the beginning of each Beacon Interval (BI). The AP may allocate one or more RAWs within a BI, and each RAW is divided into several RAW slots. The stations may only attempt to access the channel within a pre-determined RAW slot period. The periodic channel reservation introduced in the RAW mechanism is one of the promising solutions for preserving energy in IoT networks, where most connected devices are battery-operated

The existing analytical models to study the RAW mechanism are based on different approaches, such as probability theory [41, 40, 42] and Markov chains [45, 47]. However, these works still have some shortcomings such as the assumption of only one packet to send during the RAW, and the non-consideration of the channel waste time during the backoff slots and holding periods of the RAW slots.

As IoT devices are mostly battery-operated, saving their energy to extend their lifetime is a topic of current interest, especially within an IEEE 802.11ah network where the standard brings more suitable features for IoT applications [4]. In [57], the authors proposed a sampling algorithm to optimize RAW parameters in the IEEE 802.11ah standard. It is based on a multi-objective optimization model addressing both throughput and energy efficiency within the RAW. The results are interesting, yet not confirmed by simulations. Kai et al. [74] studied the energy efficiency within the RAW based on an integer nonlinear programming model. They also proposed a traffic grouping scheme to optimize the energy efficiency within RAW slots while maintaining fairness among stations. However, their approach is based on steady-state results, which do not apply to the short contention time interval of the RAW slot.

Differently from the works mentioned above, we propose an approach based on renewal theory, which accurately models the finite RAW slot duration. Although many aspects of wireless networks have renewal process behaviors, very limited works integrate the renewal theory in their analytical models [75, 76]. However, the time constraint for contention is not present in these works as they consider only the renewal cycle periods, which leads to steady-state results.

As the duration of a RAW slot is very short, the stations will not have enough time to reach a steady state. Hereafter, we construct a counting process to track the transmissions within the RAW slot up to its end. We shall prove that the RAW slot duration critically impacts its performance in terms of successful transmissions and energy efficiency. Hence, our model considers the time-limited channel access of the RAW slot and proves effective as it yields implicit closed-form results, which are suitable for real-time networks.

We shall validate our developed model and associated results through two perspectives:

1. Accuracy: To prove the accuracy of our analytical findings, we develop a discrete-event simulator on MATLAB. We henceforth display the results of 10000 simulations to validate our analytical results.

2. Convergence: Since we are modeling and evaluating the performance of time-limited channel access, our analytical results should converge to the stationary state results when the contention time (i.e., RAW slot duration) is long enough.

3.2 System Model

The RAW mechanism combines both deterministic and random approaches for channel access as described in Chapter 2. The channel time is divided into BIs, and the AP allocates one or more RAW periods within each BI. The RAW period in its turn, may be split into one or more RAW slots. Finally, only the designated stations are allowed to compete for channel access during their allocated RAW slot. Henceforth, we focus on modeling and evaluating only one RAW slot with disabled boundary crossing.

We consider a RAW slot with duration T_S and g contending stations under saturated conditions. Only the g stations get access to the channel during the time interval T_S.

Our objective is to model, evaluate and analyze the contention during the uplink MAC operation: the transmission process of packets from the stations to the Access Point (AP). Since our main interest is to model the time-limited contention of the RAW mechanism, we assume: 1) ideal channel conditions and no capture effect: transmission errors are only due to packet collisions, 2) packets have a constant and fixed size, and 3) the stations operate under saturated conditions, i.e., stations are always ready to transmit.

The stations access the channel within a RAW slot using the DCF mechanism. That is to say, whenever there is a new packet, a backoff process is invoked. The station first senses if the channel is idle for DCF Inter-frame Spacing (DIFS) duration given by T_{DIFS}. The station then chooses a random backoff time counter uniformly from $[0, W_j - 1]$, where W_j is the current Contention Window (CW) size in time slots. The size of CW is initially set to W_0 and doubled after each unsuccessful transmission until it reaches a maximum value $W_m = 2^m \cdot W_0$. It is reset to W_0 whenever a packet is successfully delivered or is dropped. Note that a station is allowed to transmit only at the beginning of each slot time, and each packet can be sent in $m+1$ attempts before it is dropped. When the backoff counter reaches zero, the station obtains a TXOP to transmit its packet. For simplicity and to maintain fairness among stations, we assume that only one packet is transmitted within a TXOP. Hence, the duration of one TXOP is given by

$$\varphi = T_{\text{DATA}} + \text{SIFS} + T_{\text{ACK}}\,, \tag{3.1}$$

where Short Inter-frame Space (SIFS) is the time interval defined in the standard [4]. T_{ACK} and T_{DATA} are the transmission times of an Acknowledgment (ACK) frame and data frame defined, respectively. The data frame transmission time is defined as:

$$T_{\text{DATA}} = T_{\text{PLCP}} + \frac{\text{Payload} + \text{MacHeader}}{\text{DataRate}}\,, \tag{3.2}$$

where T_{PLCP} is the duration of the Physical Layer Convergence Protocol (PLCP) header.

According to the standard [4], there are two cases regarding whether an ongoing transmission is allowed to cross the RAW slot boundary. In order to provide a fair resource allocation to the contending stations, we consider only the non-crossing case of the RAW slot boundary. That is, any station in a group cannot cross its assigned RAW slot boundary during the transmission of a packet.

3.3 Markov Chain Model

This section presents a MC model for channel access contention of a given group of stations. Although this approach is not considering the limited time constraint, it provides stationary state results, which provide the limits of time-limited results.

Most existing analytical models of DCF-based networks are based on MCs. The idea was started by Bianchi [5] who proposed the first model, considering an infinite retry limit for packet transmission. Several modified models are proposed later, which integrate other assumptions to cover more realistic scenarios. Therefore, we propose to study the contention within the RAW slot using a MC approach in this section.

3.3.1 Model

Consider a group of g stations, competing for channel access in saturated traffic conditions. That is, every station has always an available packet to transmit. The stations contend for channel access using the DCF mechanism explained in the previous section. Thus, each station transmits its packet after a random backoff time.

We model the backoff procedure using a bi-dimensional Markov chain. This model is based on the one introduced in [5], but with the assumption of finite retry limits. We may also refer to the findings of this section as Bianchi's results because we shall mainly be based on his model [5]. We first aim to obtain the probability τ that a station will transmit a packet. Afterward, we shall derive the other channel access probabilities and hence the performance of an infinite RAW slot.

Theorem 3.1. *Let $b(t)$ and $s(t)$ be the stochastic processes representing the backoff time counter and the backoff stage, respectively, for a given station at time t.*

The bi-dimensional process $\{s(t), b(t)\}_{\{t\in\mathbb{N}\}}$ is a discrete-time Markov chain with a unique stationary distribution.

Proof. Let $b(t)$ be the stochastic process representing the backoff time counter for a given station at time t. Since the channel time is decomposed into slots and the system keeps the same state during each slot, we adopt an integer time scale where t corresponds to the beginning of the t^{th} slot time. Note that the adopted discrete time scale for the process $b(t)$ is not related to the system time. That is because the duration of a given slot is not constant and depends on its state. A randomly chosen slot can be either idle when no station is transmitting, busy with successful transmission when a single station is transmitting, or busy with a collision when at least two stations are transmitting simultaneously. Hence, the time between two consecutive backoff times counter decrements is variable and depend on the state of the slot that separates them.

The value of the backoff time counter is randomly chosen from a doubled CW size after every failed transmission attempt, hence it depends on the transmission attempts (history) of the packet. Therefore, the stochastic process $b(t)$ is not Markovian.

Let $s(t)$ be the stochastic process representing the backoff stage $(0,...,m)$ of the station at slot time t. Since all stations contend for channel access using the same configuration of the DCF mechanism (Same initial CW and transmission attempts), a transmitted packet of a given station collides with a constant and independent probability p. The probability p is referred to as conditional collision probability, depending on the number of stations g, which is fixed.

Hereafter, we construct the bi-dimensional process $\{s(t), b(t)\}_{\{t\in\mathbb{N}\}}$, where $b(t)$ decrements by one unit at the beginning of each slot as long as $s(t)$ keeps the same value, and

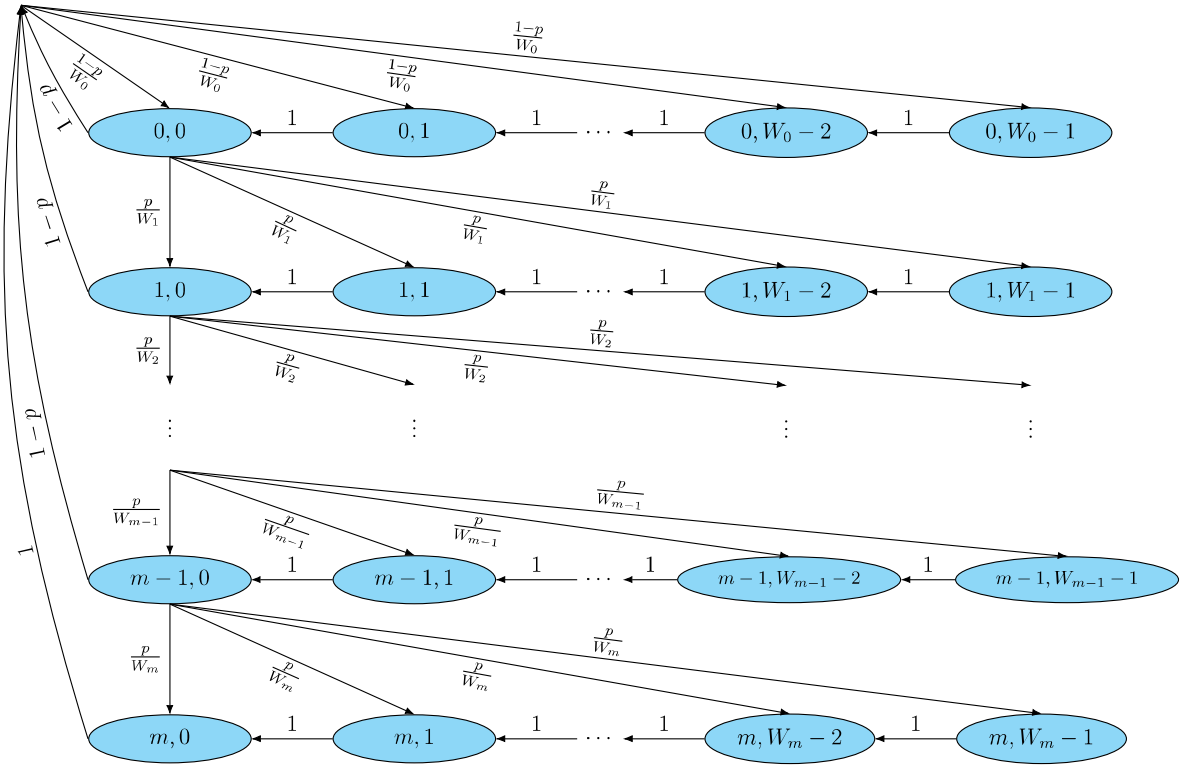


Figure 3.1: Markov chain model for the backoff window scheme.

chooses uniformly a new random value from $[0, W_j - 1]$ at the same time that $s(t)$ changes its value to j. The transitions of this process depend only on the independent probability p and the fixed sizes of the CW: $W_0, \ldots W_m$.

Therefore, the bi-dimensional process $\{s(t), b(t)\}_{\{t \in \mathbb{N}\}}$ is a discrete-time Markov chain as depicted in Figure 3.1.

We note the transition probabilities as follows:

$$\Pr\{j_1, k_1 \mid j_0, k_0\} = \Pr\{s(t+1) = j_1, b(t+1) = k_1 \mid s(t) = j_0, b(t) = k_0\}. \tag{3.3}$$

The only non-null transition probabilities are the following:

$$\begin{cases} \Pr\{j, k \mid j, k+1\} = 1 & \text{for } 0 \le j \le m \text{ and } 0 \le k \le W_j - 2 \quad (3.4) \\ \Pr\{0, k \mid j, 0\} = \dfrac{1-p}{W_0} & \text{for } 0 \le j \le m-1 \text{ and } 0 \le k \le W_0 - 1 \quad (3.5) \\ \Pr\{j, k \mid j-1, 0\} = \dfrac{p}{W_j} & \text{for } 1 \le j \le m \text{ and } 0 \le k \le W_j - 1 \quad (3.6) \\ \Pr\{0, k \mid m, 0\} = \dfrac{1}{W_0} & \text{for } 0 \le k \le W_0 - 1 \quad (3.7) \end{cases}$$

The transition probabilities in (3.4) models the fact that the backoff time counter decrements at the beginning of each slot time. The transition probabilities in (3.5) models the fact

that after a successful packet transmission with probability $1-p$, the backoff time counter starts for a new packet with a random number in the stage 0 with probability $1/W_0$. The transition probabilities in (3.6) models the fact that when an unsuccessful transmission occurs with probability p at stage $j-1$, the backoff stage increases and the backoff time counter chooses a new value uniformly in the range $[0,W_j]$ with probability $1/W_j$. The transition probabilities in (3.7) models the fact that when the backoff stage reaches the maximum value m, the station drops the current packet and resets its backoff state for a new packet.

Let $b_{j,k}$ be the stationary distribution of the DTMC $\{s(t),b(t)\}$ at state $(j,k) \in [0,m] \times [0,W_j-1]$, defined as follows:

$$b_{j,k} = \lim_{t\to\infty} \Pr\{s(t)=j, b(t)=k\} \tag{3.8}$$

The balance equations are expressed as

$$\begin{cases} b_{j,0} = p^j \cdot b_{0,0}, & \text{for } 1 \le j \le m; \\ b_{j,k} = \frac{W_j-k}{W_j} \cdot b_{j,0}, & \text{for } 1 \le j \le m \text{ and } 1 \le k \le W_j-1; \\ b_{0,k} = \frac{W_0-k}{W_0}\left[b_{m,0} + (1-p)\sum_{j=0}^{m-1} b_{j,0}\right], & \text{for } 1 \le k \le W_0-1. \end{cases} \tag{3.9}$$

with the normalization condition

$$\sum_{j=0}^{m} \sum_{k=0}^{W_j-1} b_{j,k} = 1 \tag{3.10}$$

□

The relations in (3.9) shows that all the values $b_{j,k}$ are expressed in function of $b_{0,0}$ and the probability p. Hence, by simplifying the normalization condition in (3.10), we obtain

$$b_{0,0} = \frac{2(1-2p)(1-p^{m+1})}{W_0(1-(2p)^{m+1})(1-p)+(1-2p)(1-p^{m+1})} \tag{3.11}$$

Therefore, the probability τ that a station will transmit a packet in a randomly chosen slot is given by

$$\tau = \sum_{j=0}^{m} b_{j,0} = \frac{1-p^{m+1}}{1-p} \cdot b_{0,0}\,. \tag{3.12}$$

The probability p that the transmitted packet encounters a collision is the probability that at least one of the $g-1$ remaining stations transmits in the same time slot. knowing that each station transmits with probability τ, the conditional collision probability p is expressed as follows

$$p = 1-(1-\tau)^{g-1}\,. \tag{3.13}$$

The probabilities τ and p can be obtained by solving the nonlinear system formed by (3.12) and (3.13) using numerical methods.

Let P_{tr}^{g} be the probability that at least one station transmits a packet in the considered slot. Since g stations are contending to access the channel, and each one transmits with probability τ, we have

$$P_{tr}^{g} = 1 - (1-\tau)^{g} . \tag{3.14}$$

The probability P_{s}^{g} for a packet to be transmitted successfully is given by the probability that exactly one station transmits, conditioned on the fact that at least one station transmits, i.e.,

$$P_{s}^{g} = \frac{g\tau(1-\tau)^{g-1}}{P_{tr}^{g}} . \tag{3.15}$$

3.3.2 Performance Evaluation

Evaluating the RAW slot performance based on only channel access probabilities will provide results that are suitable for an infinite RAW slot without any channel time constraint. Thus, stationary results are essential to validate time-constrained results, as the latter should converge to the first.

3.3.2.1 Normalized Throughput

The throughput S corresponding to a tagged group of g stations is the average quantity of bits transmitted successfully by the contending stations. We have

$$S = \frac{P_{tr}^{g} \cdot P_{s}^{g} \cdot T_{DATA}}{(1-P_{tr}^{g})\sigma + P_{tr}^{g} \cdot P_{s}^{g} \cdot \mathrm{E}[B_s] + P_{tr}^{g}(1-P_{s}^{g})\,\mathrm{E}[B_c]} , \tag{3.16}$$

where σ is the duration of an empty slot time. $\mathrm{E}[B_s]$ and $\mathrm{E}[B_c]$ are the average time periods for a successful and an unsuccessful transmission, respectively. For convenience, we assume $\mathrm{E}[B_s] = \mathrm{E}[B_c] = \phi + DIFS$.

3.3.2.2 Energy Efficiency

Denote by eb_{tx} and eb_{l} the energy per bit consumed by one station when transmitting and listening to medium, respectively. Hence, during a randomly chosen busy slot, the transmitting station consumes $e_{tx} = (\phi + DIFS) \cdot DataRate \cdot eb_{tx}$. A station listening to a busy slot consumes $e_{l} = (\phi + DIFS) \cdot DataRate \cdot eb_{l}$. During an idle slot, each station wastes $e_{idle} = \sigma \cdot DataRate \cdot eb_{l}$.

During a random slot containing a successful transmission, one station transmits while the remaining stations listen to a busy medium. Hence the energy consumed is given by:

$$E_s^{st} = e_{tx} + (g-1) \cdot e_l \,. \tag{3.17}$$

During a random slot containing a collision, at least two stations transmit while the remaining stations listen to a busy medium. The energy wasted in this case is expressed as follows:

$$E_c^{st} = g \cdot (\tau \cdot e_{tx} + (1-\tau) \cdot e_l) \,. \tag{3.18}$$

The energy consumed during an idle slot is given by the total listening energy wasted by all stations. We have

$$E_{idle}^{st} = g \cdot e_{idle} \,. \tag{3.19}$$

We henceforth define energy efficiency by the ratio of energy invested in the successful transmission of packets among the total wasted energy in a randomly chosen slot. Thus, we define the stationary energy efficiency for a group of g stations as follows:

$$E_{eff} = \frac{P_s^g \cdot E_s}{(1-P_{tr}^g)E_{idle} + P_{tr}^g \cdot P_s^g \cdot E_s + P_{tr}^g(1-P_s^g)E_c} \,. \tag{3.20}$$

3.4 New Accurate Approach for Time-Limited Contention

According to the standard in [4], the number of RAW slots within the RAW and their durations are defined in two scenarios: 1) the RAW consists of up to 8 RAW slots, the length of each is up to 246.14 ms, or 2) the RAW consists of up to 64 RAW slots, the duration of each is up to 31.1 ms. Such a short duration is insufficient for the network to converge to a stationary state. Hence, a different approach is required to study the contention for a time-limited scenario, as in the RAW slot. Therefore, we propose a new approach based on renewal theory that accurately models the contention under a time constraint. The renewal theory is defined to be the study of renewal processes, which mainly addresses complicated processes that have randomly occurring instants (generically called renewals or arrivals) at which the system returns to a state probabilistically equivalent to the starting state [9]. The time interval between two consecutive renewal epochs constructs a renewal period, which is also referred to as the renewal cycle.

3.4.1 Renewal Theory Based Modeling

Despite the fact that many aspects of wireless networks have renewal process behaviors, there is limited use of this approach. Nevertheless, some interesting studies integrate the renewal the-

ory into their analytical models. In [77], the authors introduced a three-level renewal process to model the contention access period of the IEEE 802.15.4 MAC protocol. However, their model is limited to the classical slotted non-persistent Carrier Sense Multiple Access (CSMA) defined in the IEEE 802.15.4 standard, which is different from the DCF MAC protocol. Zhang [75] implemented the renewal theory to analyze non-saturated IEEE 802.11 DCF networks by modeling the behavior of each station in the network as a Markov renewal process, which renews when the station enters the first backoff stage. The developed approach is proved to be more explicit than the well-known Bianchi's approach, especially for non-saturated networks. In [76], Khairy et al. developed an analytical model based on the renewal theory to study the performance of multi-channel MAC protocols in IEEE 802.11 WLANs with coexisting single-channel legacy users. Their concept is based on modeling the transmissions of single-channel and multi-channel users with and without channel bonding as a two-level renewal process. However, in both [75], and [76], the time factor, which is a critical constraint in the RAW mechanism, is not present in their framework. Nevertheless, the renewal process in these studies is deployed only considering the renewal cycle periods, which represent the interarrival times. In our approach, we introduce the counting process, which provides impressive, accurate results, especially when the contention time is a critical factor in the network.

Hereafter, we provide a mathematical model for evaluating the performance within a RAW slot period. Unlike the models mentioned above, our proposed framework is based on the renewal theory, which yields implicit closed-form results and is suitable for accurately modeling time-constrained channel access. Rather than the slotted channel time assumed in [44, 45, 77], we consider random continuous-time channel access. We develop a counting process to track the transmissions over the channel up to a given time instant, which is simplified into a closed-form function. Therefore, we evaluate the performance of time-limited contention within the RAW slot period. The analytical results of our model are more suitable for a real-time network than other complex approaches, as in [47].

3.4.2 Medium Access Probabilities

We first need to derive channel access probabilities, which will be used later to develop the proposed stochastic framework. Assuming that each station in the RAW slot attempts to transmit in a given slot with the same probability τ, we have the following proposition.

Proposition 3.1. *The probability τ that a station will transmit a packet in an idle slot is given by*

$$\tau = \frac{2p\left((2p-1)(m(p-1)+p-2)p^m+2\right)-2}{W_0(p-1)^2\left(2^{m+1}p^{m+1}-1\right)+2p\left((2p-1)(m(p-1)+p-2)p^m+2\right)-2}. \tag{3.21}$$

Proof. Consider a tagged station, and let A and B be the random variable denoting the number of attempts and backoffs experienced by one packet. By adopting the approach of mean value

analysis as in [40], the probability τ is defined as follows

$$\tau = \frac{\mathrm{E}[A]}{\mathrm{E}[A] + \mathrm{E}[B]} \tag{3.22}$$

where E[.] is the expected value function.

Let p be the probability that the transmitted packet encounters a collision. As we consider here a finite number of retransmissions limited at m, both A and B follow a truncated geometric distribution with rate p, but with different sample spaces Ω_A and Ω_B respectively. We have

$$\Omega_A = \{1, 2, ..., m+1\},$$

and

$$\Omega_B = \left\{ \frac{W_0}{2}, \frac{2W_0}{2}, ..., \frac{2^m W_0}{2} \right\}.$$

Hence,

$$\begin{aligned} \mathrm{E}[A] &= \frac{(1-p)\sum_{k=0}^{m}(k+1)p^k}{1-p^{m+1}} \\ &= \frac{(1-p)\left(-mp^{m+1} - 2p^{m+1} + mp^{m+2} + p^{m+2} + 1\right)}{(p-1)^2(1-p^{m+1})}, \end{aligned} \tag{3.23}$$

and

$$\begin{aligned} \mathrm{E}[B] &= \frac{(1-p)\sum_{k=0}^{m}\left(2^k W_0\right)p^k}{2(1-p^{m+1})} \\ &= \frac{W_0(1-p)\left(2^{m+1}p^{m+1} - 1\right)}{2(2p-1)(1-p^{m+1})}. \end{aligned} \tag{3.24}$$

Therefore, by substituting (3.23) and (3.24) in (3.22) we obtain the result.

□

The probability p that the transmitted packet encounters a collision is the probability that at least one of the $g-1$ remaining stations transmits in the same time slot. Hence,

$$p = 1 - (1-\tau)^{g-1}. \tag{3.25}$$

The nonlinear system formed by (3.21) and (3.25) can be solved numerically (e.g., using Newton's method [78]) to obtain the probabilities τ and p.

Since each station transmits with probability τ, the probability P_{idle}^{g} that a given random slot is idle is defined as follows

$$P_{idle}^{g} = (1-\tau)^{g}. \tag{3.26}$$

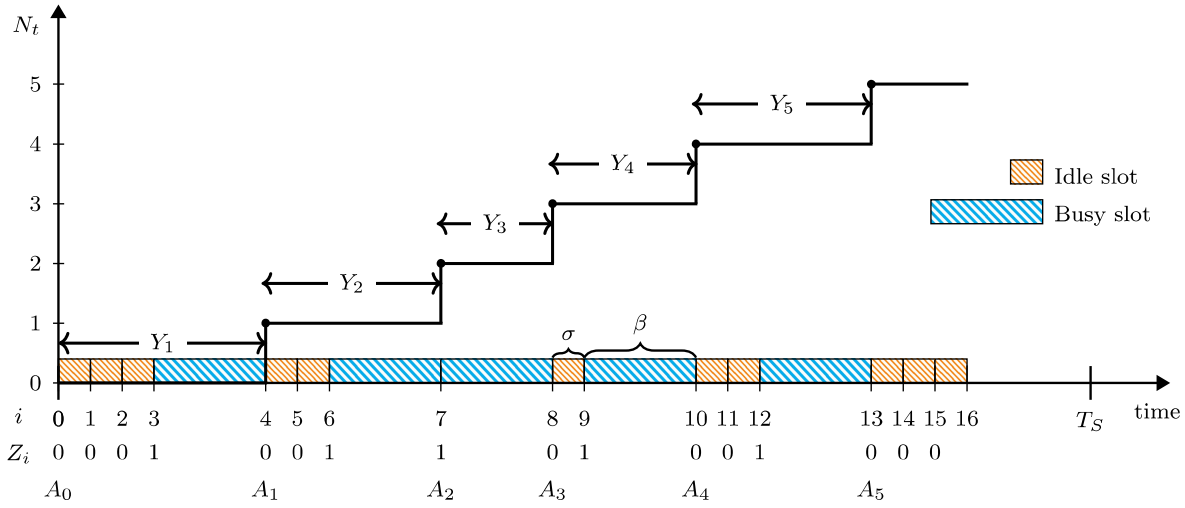


Figure 3.2: Illustration of the stochastic model: the sample values of Z_i and the arrival epochs $\{A_k\}_{k\geq 0}$ are shown below the time instants. The sample function of the counting process $\{N_t\}_{t\geq 0}$ is the step function illustrated with a unit step at each arrival epoch.

The probability P_s^g for a packet to be transmitted successfully is given by the probability that exactly one station transmits, i.e.,

$$P_s^g = g\tau(1-\tau)^{g-1}. \tag{3.27}$$

The probability P_c^g that a transmitted packet encounters a collision is given by the probability that at least two stations are transmitting at the same time. Thus we have

$$P_c^g = \sum_{k=2}^{g} \binom{k}{g} \tau^k (1-\tau)^{g-k}. \tag{3.28}$$

We can easily verify that $P_{idle}^g + P_s^g + P_c^g = 1$, which proves the three possible states of a given random slot.

3.4.3 Time-Limited Contention Model

While the stations within the RAW slot start contending to access the medium, each slot can be seen in a binary way where it is either idle or busy. A busy slot may contain a successful transmission if only one station is transmitting or a collision otherwise.

We present in Figure 3.2 an illustration of our proposed stochastic model that we describe afterwards. Let $0 = t_0 < t_1 < t_2 < \cdots$ be the instants in time constructing slots, where for $i = 1,2,...$, the length of slot i is $t_i - t_{i-1}$. Let Z_{i-1}, $i = 1,2,...$ be a random variable representing the type of slot i, where $\{Z_i = 1\}$ indicates a busy slot and $\{Z_i = 0\}$ otherwise. We also refer to the events $\{Z_i = 1\}$ and $\{Z_i = 0\}$ as success and failure of the i^{th} trial, respectively.

Denoting by σ the length of an idle slot and $\beta = \varphi + DIFS$ the duration of a busy slot, we can define the length of the slot i as follows:

$$t_i - t_{i-1} = \begin{cases} \beta, & \text{if } \{Z_{i-1} = 1\} \\ \sigma, & \text{if } \{Z_{i-1} = 0\} \end{cases} \tag{3.29}$$

Since the medium is busy with probability $(1 - P_{idle})$ and idle otherwise, we have for $i = 0, 1, ...$, $\Pr\{Z_i = 1\} = 1 - P_{idle}$ and $\Pr\{Z_i = 0\} = P_{idle}$. Hence, $Z_0, Z_1, ...$ form a sequence of IID binary random variables, which construct a Bernoulli process $\{Z_i\}_{i \geq 0}$.

We intend to track transmissions that occurred during the limited-time interval within the RAW slot. To achieve that, we propose to model the contention in the RAW slot within the context of the renewal theory in order to construct a counting process to track the transmissions. To this end, the arrivals will be presented by busy slots. Assuming the slot type (idle or busy) to be characterized by the Bernoulli process $\{Z_i\}_{i \geq 0}$, it is not enough to rely only on this process as it is evolving in discrete time with events without considering the time occupied by the events. In other words, tracking the transmissions with the aggregated number of successes associated to the Bernoulli process $\{Z_i\}_{i \geq 0}$ is not sufficient as it is not possible to know the number of events that happened within the RAW slot duration T_S. That is because the length of the slot is changing depending on its type as defined in (3.29).

Therefore, we shall construct a continuous-time framework taking into account the RAW slot duration T_S. Such a framework is based on a counting process to track the busy slots up to T_S, where the interarrivals are random time intervals constructed by the total duration of a busy slot and the preceding idle slots. That is, the interarrival time of the n^{th} busy slot is given by the time between the $(n-1)^{\text{th}}$ and n^{th} successes of the of the Bernoulli process $\{Z_i\}_{i \geq 0}$, respectively.

Despite that this modeling is seen to be a mixture of discrete-time processes and continuous-time processes, we shall accurately define the targeted counting process within the standards and definitions of the renewal theory.

Let N_t be the number of busy slots up to and including time t, which is illustrated by the step function in Figure 3.2. Our objective is to find the expected number of busy slots at the end of the RAW slot, which is the time instant T_S. To proceed with this objective, we define a renewal process framework in which N_t will be defined as a counting process for the busy slots up to time t. Hereafter, we consider each complete busy slot as a new arrival to the system.

Theorem 3.2. *Let $A_1, A_2, ...$ be random variables representing the arrival epochs of transmissions. The sequence $\{A_k, k \in \mathbb{N}\}$ is a renewal process, and $\{N_t, t \in \mathbb{R}^+\}$ is its associated counting process.*

Proof. Let X_1 be the random variable representing the number of failures before the first suc-

cess, and $X_i, i \geq 2$ the random variable representing the number of failures between the $(i-1)^{th}$ success and the i^{th} success. The random variables $\{X_i\}_{i\geq 1}$ are IID following the geometric distribution on $\mathbb{N}$ as follows

$$\Pr\{X_i = k\} = (1 - P_{idle})P_{idle}^k, \quad k \geq 0.$$

Let $Y_1, Y_2, \ldots$ be the interarrival times as depicted in Figure 3.2. We have

$$Y_i = \sigma X_i + \beta, \quad i \geq 1 \tag{3.30}$$

A new arrival represents a new complete busy slot. Hence, the arrival epochs are defined by the time instants indicating the end of busy slots. In other words, A_k is the time instant of the k^{th} arrival, which represents the time instant at the end of the k^{th} busy slot. Hence, we have:

$$A_0 = 0, \quad A_k = A_{k-1} + Y_k, \quad k \geq 1 \tag{3.31}$$

where $A_0 = 0$ represents zero arrival at time 0.

Since $X_1, X_2, \ldots$ are IID random variables, then the interarrival times $Y_1, Y_2, \ldots$ are also IID random variables positively distributed ($\Pr\{Y_i > 0\} = 1$). Hence, $\{A_k\}_{k\geq 0}$, is a well-defined renewal process.

Hereafter, the number of arrivals (busy slots) N_t at time t is the counting process associated to $\{A_k\}_{k\geq 1}$. We have

$$N_t = \arg\max_{k\in\mathbb{N}} \{A_k \leq t\} \tag{3.32}$$

□

We can now proceed to find the expected number of arrivals (busy slots) $\mathrm{E}[N_{T_S}]$ at the end of the RAW slot T_S. Note that the holding period is considered here as the last counted transmission should be ended before the RAW slot boundary. Otherwise, it will not be considered by the counting process at the end of the RAW slot.

Proposition 3.2. *Consider a RAW slot of duration T_S allocated for g stations. The expected number of busy slots that occurred during this RAW slot is given by*

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{T_S - k\cdot\beta}{\sigma} \right\rfloor} \binom{j+k-1}{j} (1 - P_{idle})^k P_{idle}^j \tag{3.33}$$

Proof. We have

$$N_{T_S} = \arg\max_{k\in\mathbb{N}} \{A_k \leq T_S\}$$

Since $\{A_k; k \geq 0\}$ is an increasing process, N_{T_S} can also be defined as the size of the set $\{k \geq 1 : A_k \leq T_S\}$. Hence,

$$\begin{aligned}
\mathrm{E}[N_{T_S}] &= \mathrm{E}\left[\arg\max_{k\in\mathbb{N}}\{A_k \leq T_S\}\right] \\
&= \mathrm{E}[\#\{k \geq 1 : A_k \leq T_S\}] \\
&= \mathrm{E}\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{A_k \leq T_S\}}\right] \\
&= \sum_{k=1}^{\infty} \Pr\{A_k \leq T_S\}
\end{aligned}$$

Hereafter, the maximum number of successes can be achieved when all trials of the process $\{Z_i\}_{i\geq 0}$ come out with successes until the time limit T_S, hence, within the RAW slot interval $[0, T_S]$, the maximum number of busy slots that can be occurred is $\left\lfloor \frac{T_S}{\beta} \right\rfloor$. Thus, for $k > \left\lfloor \frac{T_S}{\beta} \right\rfloor$, we have $\Pr\{A_k \leq T_S\} = 0$. Hence,

$$\begin{aligned}
\mathrm{E}[N_{T_S}] &= \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \Pr\{A_k \leq T_S\} \\
&= \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \Pr\left\{\sum_{i=1}^{k} Y_i \leq T_S\right\} \\
&= \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \Pr\left\{k\cdot\beta + \sigma\sum_{i=1}^{k} X_i \leq T_S\right\} \\
&= \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \Pr\left\{\sum_{i=1}^{k} X_i \leq \frac{T_S - k\cdot\beta}{\sigma}\right\} \\
&= \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{T_S - k\cdot\beta}{\sigma} \right\rfloor} \Pr\left\{\sum_{i=1}^{k} X_i = j\right\}
\end{aligned}$$

$\sum_{i=1}^{k} X_i$ is sum of k IID random variables geometrically distributed with parameter $(1 - P_{idle})$ and hence, $\sum_{i=1}^{k} X_i$ is following the negative binomial distribution with parameters k and $(1 - P_{idle})$. Finally, we obtain

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{T_S - k\cdot\beta}{\sigma} \right\rfloor} \binom{j+k-1}{j} (1 - P_{idle})^k P_{idle}^j$$

□

Now that we have obtained the expected number of busy slots within the RAW slot time

interval $[0, T_S]$, we can distinguish the average number of busy slots containing successful transmissions and those containing collisions.

The average number of slots used for successful transmissions $\mathrm{E}\left[N^s_{T_S}\right]$ is the portion associated with the probability of success defined in (3.27) conditioned to the probability that the slots are busy. Hence,

$$\mathrm{E}\left[N^s_{T_S}\right] = \frac{\mathrm{E}\left[N_{T_S}\right] \cdot P_s}{1 - P_{idle}}. \tag{3.34}$$

Similarly, the average number of slots containing collisions is given by

$$\mathrm{E}\left[N^c_{T_S}\right] = \frac{\mathrm{E}\left[N_{T_S}\right] \cdot P_c}{1 - P_{idle}}. \tag{3.35}$$

Hereafter, the average number of idle slots during the RAW slot is defined as follows

$$\mathrm{E}\left[N^i_{T_S}\right] = \left\lfloor \frac{T_s - \mathrm{E}\left[N_{T_S}\right] \cdot \beta}{\sigma} \right\rfloor. \tag{3.36}$$

3.5 Performance Evaluation

After deriving the average number of busy slots used for transmissions and collisions, and idle slots, we can now evaluate the performance of a time-limited contention within the RAW slot in terms of throughput and energy efficiency.

3.5.1 Throughput

Now that we obtained the expected number of busy slots within the RAW slot time interval $[0, T_S]$, we can distinguish the average number of busy slots containing successful transmissions by multiplying $\mathrm{E}\left[N_{T_S}\right]$ by the probability of success defined in (3.27). In addition, multiplying the result by T_{DATA} yields the portion of time used for successful transmissions of data frames during the RAW slot duration T_S. Therefore, we define the normalized throughput within the RAW slot as follows:

$$S_{\mathrm{RS}} = \frac{\mathrm{E}\left[N_{T_S}\right] \cdot P_s \cdot T_{\mathrm{DATA}}}{T_S} \tag{3.37}$$

3.5.2 Energy Efficiency

Let e_{tx}, e_l, and e_{idle} be the energy consumed by one station when transmitting, listening to a busy slot, and listening to an idle slot, respectively. When a station has successfully transmitted

its packet, the rest of the stations have to listen to the busy slot. Hence, the energy consumed by successful transmissions during the RAW slot is given by

$$E_s^{T_S} = \mathrm{E}\left[N_{T_S}^s\right] \cdot (e_{tx} + (g-1) \cdot e_l)\,. \tag{3.38}$$

To compute the energy wasted during collisions, we need to distinguish the stations participating in the collision and the stations only listening to the busy slot. Thus, the energy consumed by collisions during the RAW slot is defined as follows

$$E_c^{T_S} = \mathrm{E}\left[N_{T_S}^c\right] \cdot g \cdot (\tau \cdot e_{tx} + (1-\tau) \cdot e_l)\,. \tag{3.39}$$

The energy consumed during idle slots within the RAW slot is the total listening energy wasted by all stations. We have

$$E_{idle}^{T_S} = \mathrm{E}\left[N_{T_S}^i\right] \cdot g \cdot e_{idle}\,. \tag{3.40}$$

The energy efficiency within a RAW slot represents the ratio of energy used for successful transmissions to the total energy wasted during the RAW slot time. Hence, we define the RAW slot energy efficiency as follows

$$E_{eff}^{T_S} = \frac{E_s^{T_S}}{E_s^{T_S} + E_c^{T_S} + E_{idle}^{T_S}}\,. \tag{3.41}$$

3.6 Numerical Results and Analysis

This section analyzes and evaluates our findings and proposals in the previous sections using MATLAB software. The numerical parameters are presented in Table 3.1. We also validate our model and results by simulation results obtained using a discrete-event simulator developed on MATLAB. The simulator mimics the DCF mechanism with consideration of all analytical assumptions. Every single simulation yields results of contention up to the length of the RAW slot. The depicted simulation results are obtained by averaging the results of 10000 RAW-slot simulations. As seen in the figures, the results obtained from our model closely match the simulation results, which validates our model.

3.6.1 Throughput

We first evaluate the performance within one RAW slot to validate our model. Figure 3.3 presents the average number of transmissions that occurred within the RAW slot, which we

Table 3.1: System parameters.

Parameter	**Value**	**Parameter**	**Value**
Data rate	1 Mbps	σ	52 µs
Payload	512 bits	CW_{min}	16
MacHeader	272 bits	CW_{max}	1024
SIFS	160 µs	m	6
DIFS	264 µs	e_{tx}	2 mJ
ACK	112 bits	e_l	1.3 mJ
T_{PLCP}	80 µs	e_{idle}	0.04 mJ

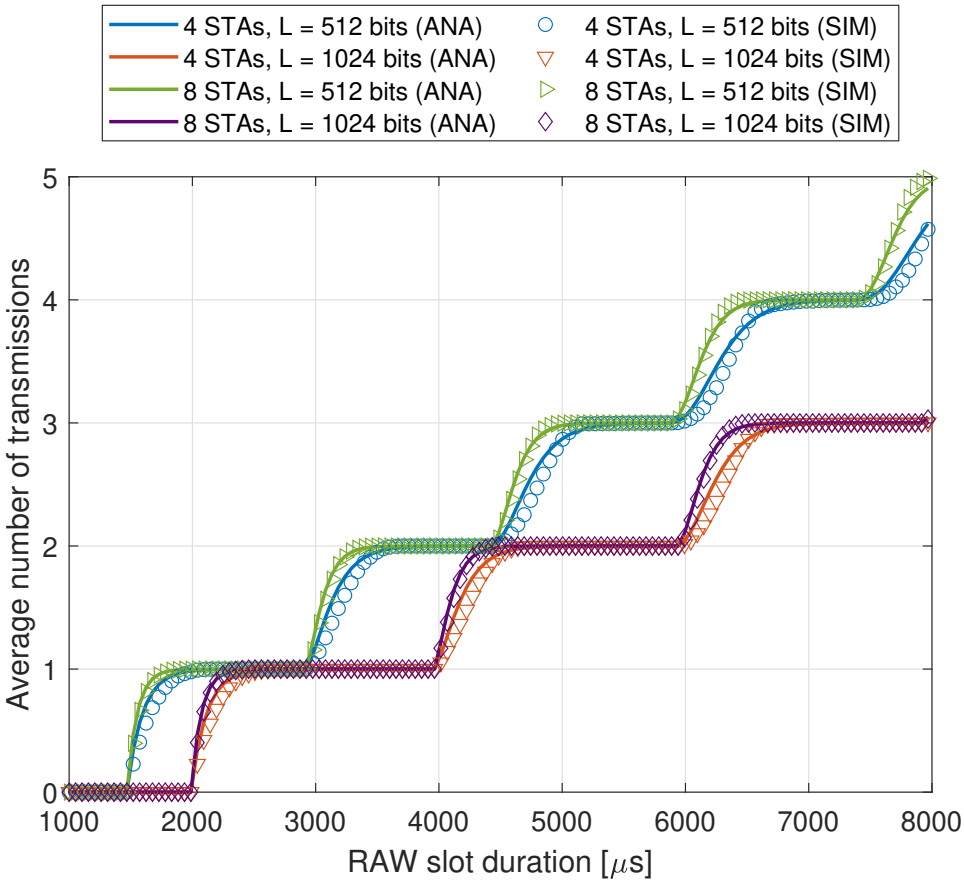


Figure 3.3: Average number of transmissions within the RAW slot in terms of its duration T_S.

have derived in our model with (3.33). We consider four different scenarios for two sets of stations allocated to the RAW slot and two different payloads. The simulations yield excellent accuracy for this metric, which confirms the results obtained in our analytical framework. We observe that the graph follows a similar path for each scenario where the expected number of transmissions stays constant for some time while T_S is increasing. This behavior is due to the fact that the additional time added to T_S is not enough to complete an additional transmission. The period during which $\mathrm{E}[N_{T_S}]$ is constant depends on the payload size, as we can see in the figure. That is because a 1024-bits payload requires a longer transmission time than a 512-bits payload.

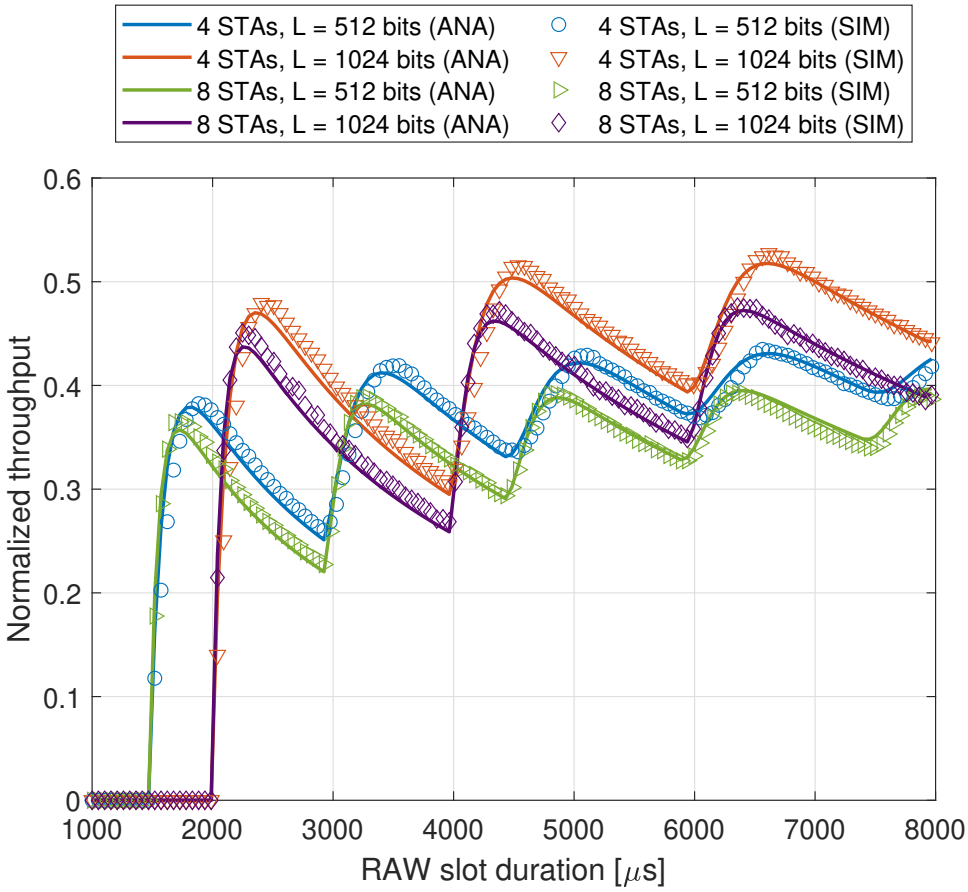


Figure 3.4: Normalized throughput of the RAW slot in terms of its duration T_S.

With the same context of the four scenarios, we depict the corresponding normalized throughput of the RAW slot in Figure 3.4. The simulations show good accuracy, which validates the analytical throughput defined in (3.37). We observe that for every scenario, the normalized throughput of the RAW slot always has the same shape where it increases and repeatedly decreases while T_S is increasing. Considering the results from Figure 3.3, we can see that the throughput decreases when the average number of transmissions is not increasing. As the normalized throughput presents the time ratio used to transmit data, the lack of transmissions when T_S is increasing explains the throughput degradation. In other words, a longer duration of the RAW slot without additional transmissions yields more wasted time and hence lower throughput. Considering the stations' factor, four stations reach a higher throughput than eight stations due to lower contention. The payload size has a critical impact on the performance as well. A small payload size causes higher contention among stations, which reduces the overall normalized throughput.

In the well-known Bianchi's model [5], the throughput is derived from the steady state of the network, which is reached after the stations are contending for a long time. Such a result does not consider the time-limited contention presented in the RAW mechanism. However, a time-limited throughput should converge to Bianchi's result when the system is running for a long enough time. Hereafter, to validate the normalized throughput of the RAW slot defined by (3.37), we depict in Figure 3.5 a comparison of the normalized throughput between our

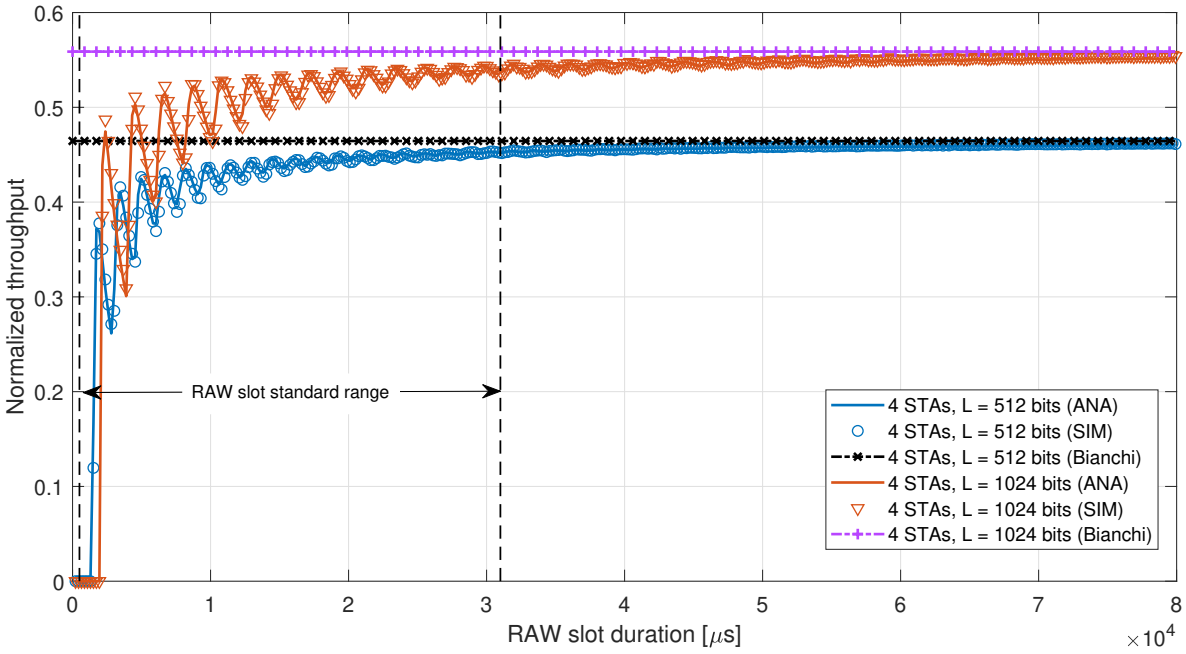


Figure 3.5: Throughput convergence in terms of time.

approach, Bianchi's model, and the simulation. We consider two different payloads: 512 bits and 1024 bits. We observe satisfactory accuracy with the simulation for our proposal, whereas Bianchi's result is constant on the same value as it does not depend on time. Besides, we see that when the RAW slot duration is long enough, our result converges to Bianchi's, which gives additional validation to our proposal.

According to the standard in [4], the number of RAW slots within the RAW and their durations are defined in two scenarios: 1) the RAW consists of up to 8 RAW slots, the duration of each is up to 246.14 ms, or 2) the RAW consists of up to 64 RAW slots, the duration of each is up to 31.1 ms. In both scenarios, the minimum duration of a RAW slot is 0.5 ms. As we consider the second scenario in our analysis, we present the RAW slot standard duration range of this case in Figure 3.5. The results prove that the RAW slot duration is a mandatory constraint to consider when modeling this mechanism and that other approaches based on Bianchi's model or the steady-state of a Markov chain cannot provide accurate results, especially when considering the standard measures defined in the amendment [4].

Besides, we observe that the throughput fluctuates when T_S increases. To understand this behavior, we present in Figure 3.6 the simulation results for the usage ratio of the different events that occupy the channel. a RAW slot interval can be seen as a reunion of five different durations: 1) the time used to transmit data frames, which represents the throughput, 2) the overhead caused by ACK frames and inter-frame spaces DIFS and SIFS, 3) the time consumed by collisions, 4) the time used by Backoff slots, and 5) the time wasted within the holding period. The reason behind the fluctuating behavior of the throughput is the partial usage of the holding period. When we increase T_S, the additional time is wasted unless it is sufficient for one additional transmission. The additional wasted time is significant compared to a very

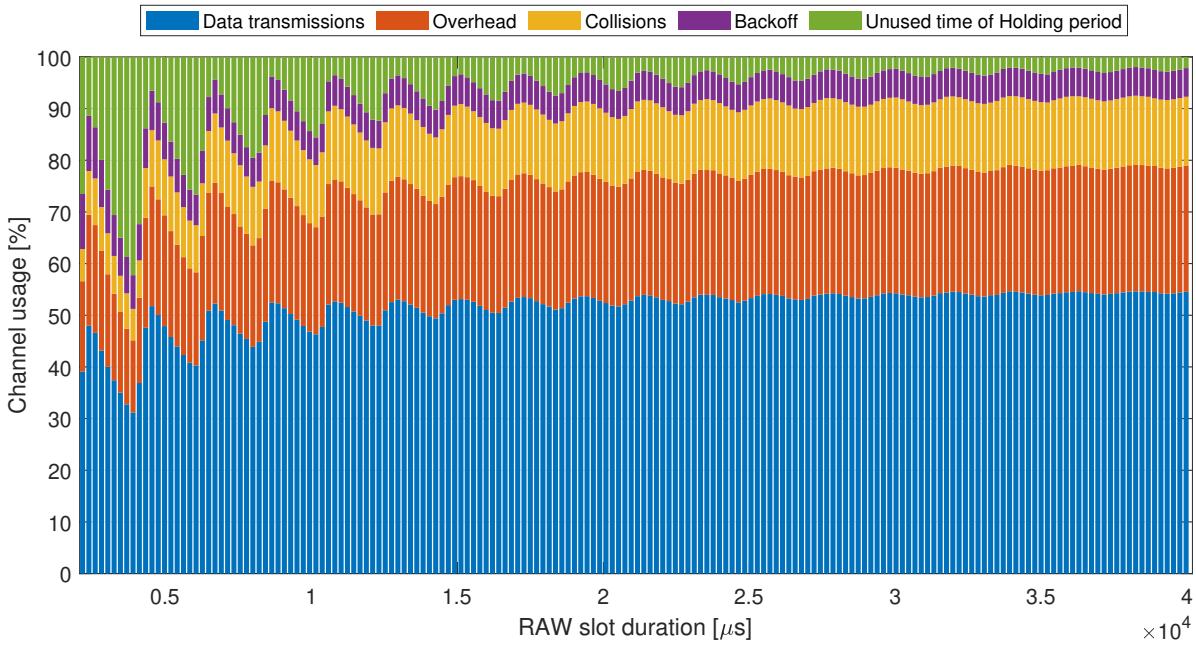


Figure 3.6: Usage ratio of different events occupying the channel in terms of the length of a RAW slot allocated for four stations with a 1024-bits payload.

short duration T_S, which causes a critical drop in the throughput. However, when T_S becomes larger, the additional time is more likely to be partially occupied with an additional transmission as the stations are contending randomly for a more extended period. Hence, the throughput degradation is smaller. Eventually, for a long enough T_S, the average number of transmissions becomes proportional to T_S, which yields a constant throughput.

3.6.2 Energy Efficiency

We consider a RAW slot of five stations with a 512 bits payload, contending to access the medium for duration T_S. The simulations yield excellent accuracy for this metric, which confirms the results obtained in our analytical framework.

Figure 3.7 presents the average number of slots occupied by successful transmissions, collisions, and idle listening during T_S. While T_S is increasing, both the number of slots containing successful transmissions and the number of slots containing collisions are either constant or increasing. That is because more contention time can contribute to additional transmissions only if it is sufficient enough. On the other hand, the average number of idle slots keeps fluctuating as the RAW slot duration increases. When the time added to the current T_S is not enough to complete an additional transmission, it becomes occupied by idle slots and hence an increase in the number of idle slots within the RAW slot. In contrast, the inverse happens when the additional time allows an extra transmission.

Figure 3.8 presents the energy consumption of the stations during the events of transmissions (E_s), collisions (E_c), and idle listening (E_{idle}). We observe the same behavior as the

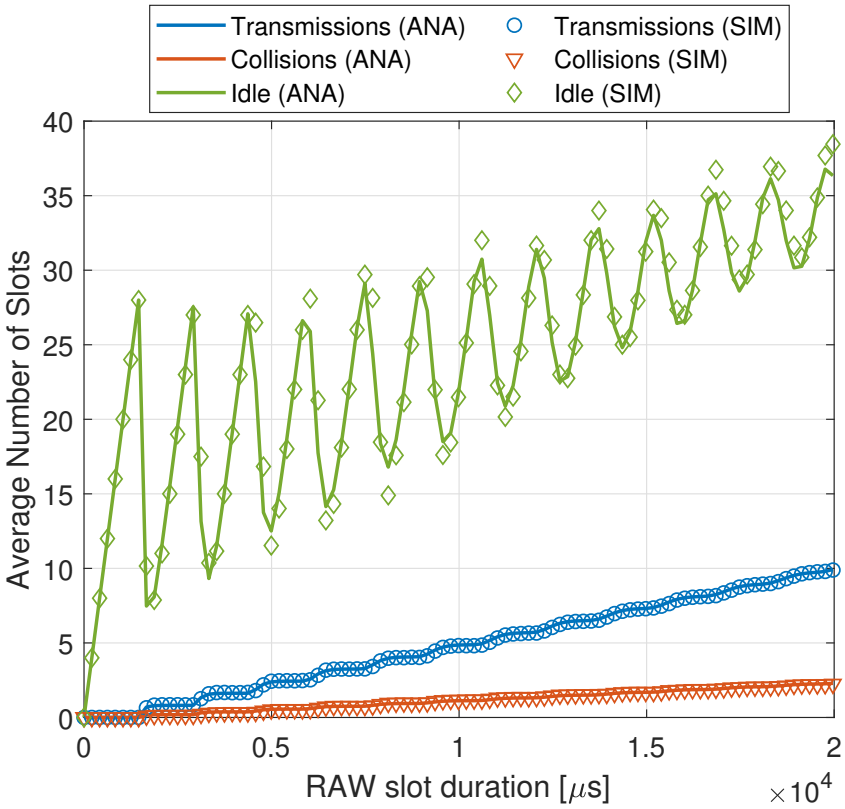


Figure 3.7: Average number of slots occupied by the events in terms of the RAW slot duration.

average number of slots. The energy consumption of both successful transmissions and collisions keeps increasing when the additional time is long enough for additional transmissions. In this case, the energy wasted by idle slots decreases as the RAW slot interval becomes less occupied by idle listening to the medium. Otherwise, when the additional time to the RAW slot is not exploitable by transmissions, the energy consumed by successful transmissions and collisions remains constant while the energy wasted in idle listening increases.

To analyze the good usage of the RAW slot interval, we depict in Figure 3.9 the energy efficiency in terms of the RAW slot duration T_S. The latter represents the energy ratio used for the successful transmission of packets over the total energy wasted during the RAW slot. The energy efficiency of the RAW slot has a fluctuating behavior where it repeatedly increases and decreases while T_S is increasing. Considering the results from Figure 3.7, We can see that the energy efficiency decreases when the average number of transmissions is not increasing. As the energy efficiency presents the energy ratio used to transmit data, the lack of transmissions while T_S is increasing explains the degradation of energy efficiency of the RAW slot. In other words, a longer duration of the RAW slot without additional transmissions yields more wasted time and hence, a waste of energy. As fewer stations present lower contention, we see that five stations make better use of the RAW slot period than the case of ten stations. Hence, a RAW slot with five stations yields higher energy efficiency for every duration. When T_S increases, the impact of the holding period T_H becomes ineffective as the ratio T_H/T_S tends to zero. Hence, the energy efficiency loses the fluctuating behavior and becomes more stable. Eventually, the

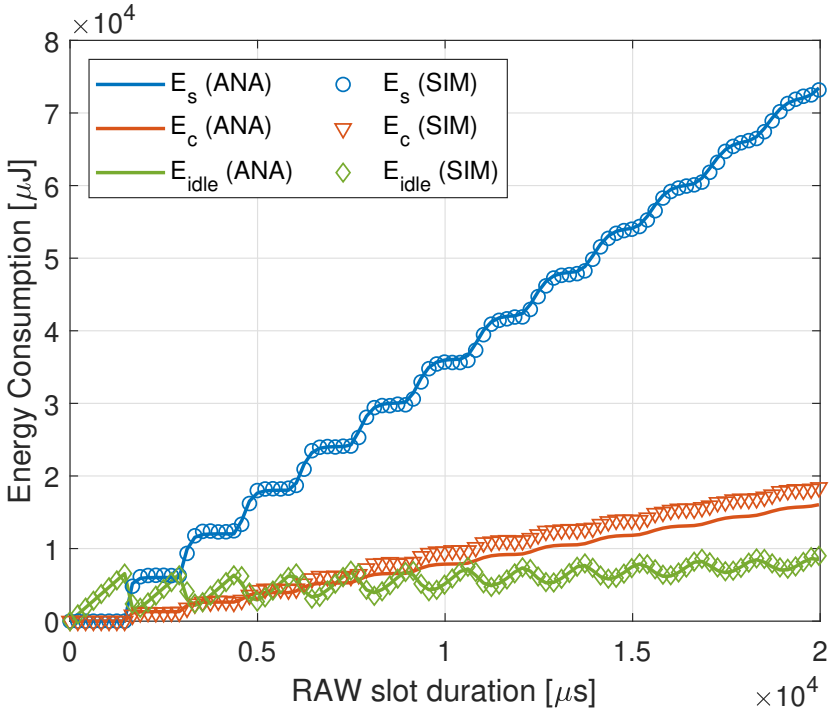


Figure 3.8: Energy consumption of transmissions, collisions, and idle listening in terms of the RAW slot duration.

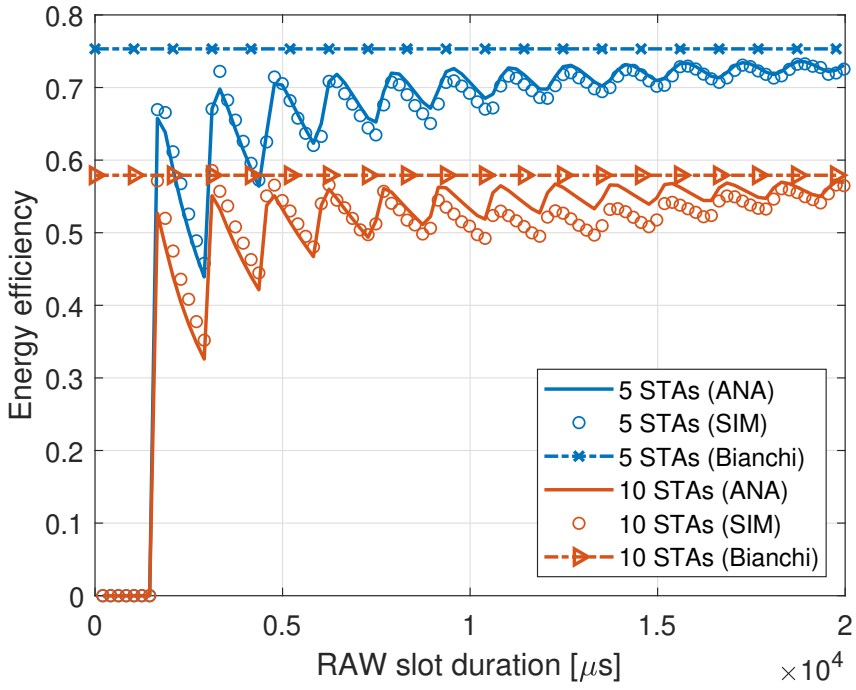


Figure 3.9: Energy efficiency in terms of the RAW slot duration.

RAW slot energy efficiency converges to the maximum achievable value, which can be reached in the case of an infinite RAW slot, representing the stationary state result given by Bianchi's approach.

3.7 Conclusion

In this chapter, we have established a new analytical framework to model, analyze, and evaluate the time-limited contention which is integrated into several new IEEE 802.11 standards addressing large-scale and dense scenarios. We particularly focus on the RAW mechanism of the IEEE 802.11ah standard, where only the assigned stations are allowed to compete for channel access during the RAW slot period. We first present a Markov Chain model for a given group of stations which provide stationary results valid for an infinite RAW slot. We then developed a new mathematical framework based on renewal theory, which is accurately characterizing the network behavior during a time-limited contention. We proposed a novel way to observe the events within the wireless medium in terms of time, where a counting process is deployed to track the number of transmissions within the RAW slot period. Our approach allows deriving the expected number of different events that occurred during the time-limited period of contention. Hereafter, we evaluate the RAW slot performance in terms of throughput and energy efficiency. We also evaluate the same metrics for the case of an infinite RAW slot using the Markov Chain model. Additionally, we developed a discrete-time event simulator with MATLAB that mimics the considered channel access during the time-limited period. Our analytical findings are meticulously validated via extensive simulations and the convergence to the stationary state results representing the case of an infinite RAW slot. We have shown that the length of the RAW slot period is critically impacting its performance, especially when it is very short. The mathematical model developed in this work can be applied to other IEEE standards integrating periodic channel reservations. Additionally, we extend this analytical framework in the next chapter to design efficient grouping schemes and parameters' configuration for the RAW mechanism.

CHAPTER 4

Efficient Configuration Framework of the IEEE 802.11ah RAW Mechanism

Contents

4.1 Introduction

The emergence of the IoT has changed the perspective of wireless communications since the number of interconnected devices has increased exponentially in recent years. According to the

International Data Corporation (IDC) [79], there will be 55.7 billion connected devices by 2025, 75% of which will be connected to IoT platforms, generating 73.1 zettabytes of data. Consequently, it is necessary to provide connectivity to this large number of energy-constrained smart devices while maximizing throughput and transmission range. As stated in Chapter 2, existing WPAN and LPWAN technologies suffer from short range coverage of WPAN and very low throughput, respectively. However, TGah has developed a new IEEE 802.11ah standard (also known as Wi-Fi HaLow) [4], which offers a trade-off between throughput and range coverage. IEEE 802.11ah systems operate in the unlicensed sub-GHz frequency bands below 1 GHz. A single AP may provide connectivity to a maximum of 8192 low-power devices (nodes) at data rates ranging from 150 Kbps to 78 Mbps over a transmission range of up to 1 km [20]. A network model for IEEE 802.11ah standard is illustrated in Figure 4.1.

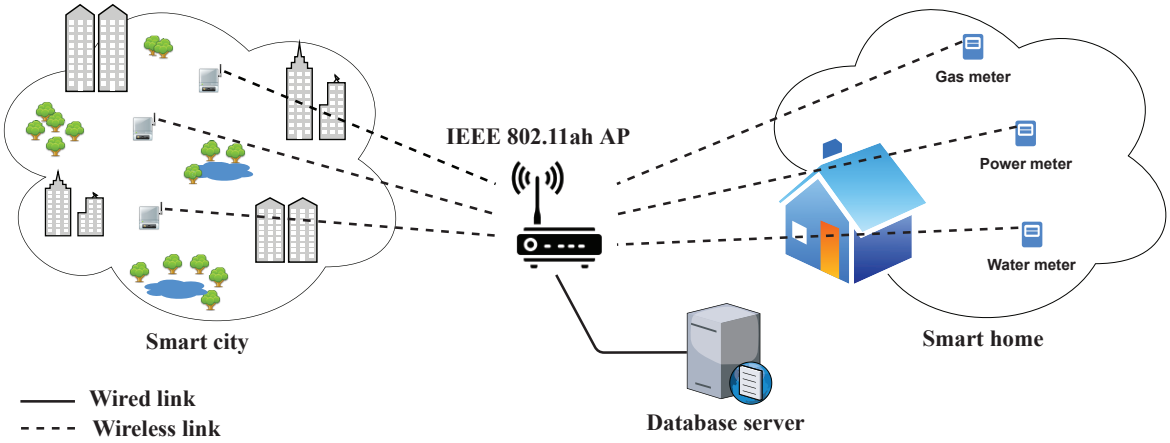


Figure 4.1: IEEE 802.11ah network model.

The IoT is expected to deploy edge computing to reduce data transfer delay and bandwidth consumption [80]. However, the number of edge nodes is estimated by billions of devices, which presents a challenge to offer good scalability and stringent latency, especially with the exponential growth of IoT devices [80, 81]. One possible solution to handle the massive contention presented in edge computing is to limit the number of nodes accessing the network resources at the same time [82]. Such an idea of putting devices in groups with periodic reservations is one significant amendment presented in the IEEE 802.11ah standard [4], which eventually makes this standard an excellent candidate to answer major challenges in IoT and edge computing applications.

To address the high contention rates that may be generated in a scenario consisting of hundreds and even thousands of nodes making use of the DCF and EDCA protocols, the IEEE 802.11ah standard introduces a group-based contention mechanism called Restricted Access Window (RAW) [4]. The RAW mechanism aims to reduce the collision probability and enhance scalability: a feature of particular interest in dense IoT networks. According to the RAW mechanism, the stations may only attempt to access a channel within a predetermined time (RAW) period, as described in Chapter 2 All these features, higher throughput, longer coverage range, and scalability make IEEE 802.11ah systems an excellent candidate to deploy IoT multimedia

(i.e., audio/video) services. Moreover, a multitude of services that require the interconnection of a large number of nodes are implemented in more extensive areas, such as smart city or smart grid applications [31]. Some studies have already shown that the key features included in IEEE 802.11ah may be very efficient in the deployment of IoT smart city applications [25].

To manage medium access inside and outside the RAW, each station maintains two backoff function states [4]. The first backoff function state is used outside the RAW, and the second is used inside the RAW. The first backoff function state is stored and suspended at the start of each RAW, and then restored and resumed at the end of RAW. The second backoff function state is used inside the RAW by the participating stations; however, it is reset and disregarded at the end of the RAW. In this work, we focus only on the medium access inside the RAW, where the stations use just the second backoff function. Therefore, we omit the impact of the two states of the backoff function.

Among the main open issues, the IEEE 802.11ah standard does not specify the details of the RAW grouping strategy. Although numerous studies have already been reported in the literature, there is no general consensus on the best strategy. In this chapter, our main goal is to contribute to the definition of an efficient grouping strategy. In Chapter 3, we developed an accurate mathematical model to evaluate performance within a RAW slot period. The analytical framework is based on the deployment of a counting process to track transmission during the RAW slot period, resulting in implicit closed-form results suitable for a real-time RAW configuration algorithm. Based on those results, we propose a RAW grouping scheme that splits the RAW into two Sub-Restricted Access Windows (sub-RAWs). Each Sub-RAW is divided into groups. The stations are then assigned to the different groups. Therefore, we propose a RAW configuration algorithm that provides the best parameters for the grouping scheme to maximize the overall performance of the RAW under saturated throughput. Furthermore, we validate and analyze our proposals with numerical results and simulations. Finally, we examine the impact of non-ideal channel conditions on the throughput within the RAW slot and the overall RAW optimal configuration.

The main contributions of this work are summarized as follows:

- We propose a RAW grouping scheme to distribute stations and channel access time within the RAW into several and different groups. We then evaluate the performance of our proposed grouping scheme in terms of the overall normalized throughput of the RAW, which we derive from the performance within each RAW slot.
- We define a RAW configuration algorithm, which tunes up the parameters of our proposed RAW grouping scheme, to alleviate channel contention and maximize the normalized throughput within the RAW. We further explore the impact of channel error conditions on RAW slot performance and overall optimal RAW configuration.

The remainder of this chapter is organized as follows. Section 4.2 defines the system

model based on the principles of operation of the RAW mechanism and sets the main problem treated in this work. Section 4.3 presents the contention within the RAW slot and derives the corresponding performance. In Section 4.4, we propose a grouping scheme for the RAW mechanism. Then Section 4.5 derives the overall associated performance of the RAW. We further introduce a RAW configuration algorithm to tune-up the RAW parameters maximizing the normalized throughput in Section 4.6. We validate and analyze our proposals with numerical results and simulations in Section 4.7. We examine in Section 4.8 the impact of the non-ideal channel with transmission errors on the performance of the RAW slot and the overall optimal RAW configuration. Finally, we conclude our work in Section 4.9.

4.2 System Model

We consider an IEEE 802.11ah network consisting of one AP ensuring the interconnection of a large number of nodes (stations). Our main goal is to model and evaluate the performance of the uplink direction of the MAC protocol: the transmission process of packets from the stations to the Access Point (AP). Since our main interest is to evaluate the RAW mechanism, we assume: 1) ideal channel conditions and no capture effect: transmission errors are only due to packet collisions; 2) packets have a constant and fixed size; and 3) the stations operate under saturated conditions, i.e., the stations are always ready to transmit. However, we adapt our findings to the case of non-ideal channel in Section 4.8. As indicated in [83, 84], a network operating below saturation traffic conditions will be able to deliver all traffic while keeping the delay within reasonable limitations. Thus, the saturation traffic allows us to determine the upper bounds of the available capacity within the RAW, and hence explore the achievements of the RAW mechanism in comparison with the legacy DCF.

To address the high traffic within dense networks, a reliable solution is to limit access to the available resources, such as the RAW slot in IEEE 802.11ah, which is an interval of time when only a limited number of stations can contend to access the medium. The same concept is repeated in other IEEE standards with different names. Therefore, the analytical framework we present below can be applied to all standards that implement the periodic channel reservation mechanism, from which we can cite: MCCA in IEEE 802.11s, HCCA TXOP negotiation in IEEE 802.11aa, CBAP and SP in IEEE 802.11ad/ay, Trigger-based multi-user scheduled access in IEEE 802.11ax, and coordinated schedule in IEEE 802.11be [66].

The AP has complete knowledge of all active stations on the network before generating the grouping schedule at the beginning of each Beacon Interval (BI). Then, the grouping parameters are broadcasted to all connected stations within a beacon frame. The grouping procedure is carried out at two levels [4].

At the first level, the AP distributes the stations into groups and splits the BI into RAWs, where each RAW r contains N_R^r stations and has a duration T_R^r. This grouping could be used to

Table 4.1: Model main parameters.

Parameter	Description
A_k	Renewal process representing the time instant at the end of the k^{th} busy slot
$\mathrm{E}[N_{T_S}]$	Expected number of busy slots within the tagged RAW slot
$\mathrm{E}\left[N_{T_{S_i}}\right]$	Expected number of busy slots within a RAW slot of sub-RAW i
g	Number of stations within the tagged RAW slot
g_i	Number of stations within the i^{th} RAW slot
K_{op}	Optimal number of groups that maximizes the throughput within the RAW
K^r	Number of groups within the r^{th} RAW
$\mathcal{K}_{\mathcal{R}}$	Grouping range
L	Payload size
N_R	Number of stations in the RAW
N_R^r	Number of stations in the r^{th} RAW within the BI
N_t	Number of busy slots up to and including time t
p	Probability that the transmitted packet encounters a collision
$P_s^{g_i}$	Probability for a packet to be transmitted successfully within the i^{th} RAW slot
P_{tr}^{g}	Probability that at least one station transmits a packet within the tagged RAW slot
$P_{tr}^{g_i}$	Probability that at least one station transmits a packet within the i^{th} RAW slot
RAW^{op}	Optimal RAW configuration
$S(K)$	Normalized throughput within a RAW composed of K RAW slots
$S(K_{op})$	Maximum achievable throughput within the RAW
S_{RS}	Normalized throughput within the tagged RAW slot
T_{ACK}	Duration an ACK frame
T_{DATA}	Duration of a data frame
T_{DIFS}	Duration of a DIFS interval
T_{F_i}	Duration of free access period within a RAW slot
T_H	Duration of holding period within a RAW slot
T_R	RAW duration
T_R^r	Duration of the r^{th} RAW within the BI
T_S	Duration of the tagged RAW slot
T_{S_i}	Duration the i^{th} RAW slot
W_0	Minimum value of contention window
W_j	Value of contention window at stage j
W_m	Maximum value of contention window
X_i	Random variable representing the number of failures between the $(i-1)^{\text{th}}$ success and the i^{th} success of the Bernoulli process $\{Z_i\}_{i\geq 0}$
Y_k	Interarrival time between the $(i-1)^{\text{th}}$ and the i^{th} arrivals
Z_i	Random variable representing the type of the $(i+1)^{\text{th}}$ slot
β	Length of a busy slot
σ	Length of an idle slot
τ	Probability that a station will transmit a packet in an idle slot
φ	Duration of one TXOP

mitigate the hidden-node problem by dividing the AP's coverage into sectors and allocating the stations in each sector to a different RAW.

At the second level, the grouping is applied within the RAW. The AP assigns the N_R^r stations in each RAW r to K^r groups, and the RAW length T_R^r is divided accordingly. Each group is associated with a given RAW slot, i.e., each RAW slot provides support to g_i stations and whose duration is given by T_{S_i}. The total number of stations and the RAW length are given by $\sum_{i=1}^{K^r} g_i = N_R^r$ and $\sum_{i=1}^{K^r} T_{S_i} = T_R^r$, respectively. In this way, the groups within the RAW access the channel consecutively, i.e., only the g_i stations that have the right to access the channel within their corresponding RAW slot T_{S_i}.

In the rest of this chapter, we focus on the second-level grouping inside one tagged RAW. We consider one RAW of a duration of T_R and N_R stations. Our main objective is twofold: (1) to propose a grouping scheme to properly distribute the stations and the RAW length across groups and (2) to propose a RAW configuration algorithm maximizing the overall RAW throughput. Table 4.1 summarizes our model's main parameters.

The configuration of the RAW mechanism is carried out through an allocation process and a stochastic medium access mechanism, i.e., the DCF mechanism. Hence, to evaluate the performance of the RAW mechanism, we first start by evaluating the performance of one RAW slot. We then derive the performance of a RAW that is composed of several RAW slots.

4.3 RAW Slot Analysis

We consider a tagged RAW slot i with duration T_{S_i} and g_i contending stations. Note that only this group of g_i stations gets access to the channel within the time interval T_{S_i}.

The handover between two RAW slots is important factor to consider [4]. In the case where crossing the RAW slot boundary is allowed, an ongoing transmission can cross the RAW slot boundary, and the stations in the next RAW slot should not start accessing the channel until the ongoing transmission is completed. Whereas, in the case that RAW slot boundary crossing is not allowed, if a station has an ongoing transmission that may cross the RAW slot boundary, it should hold its transmission and wait until its next allocated RAW slot. Allowing boundary-crossing within a group in a saturated scenario will cause extra waste of energy for stations in the next group, as these stations will wake up at the beginning of the RAW slot and they will not be able to access the channel until the last transmission from the previous group is completed. The same impact goes to the subsequent groups. In addition, boundary-crossing in the last group will increase the total duration of the RAW, which will cause waste in terms of energy for stations contending outside the RAW. Moreover, boundary-crossing may cause difficulties in the implementation of the network, mainly when TDMA is used between two or more APs. For these reasons and in order to provide a fair resource allocation to the contending stations

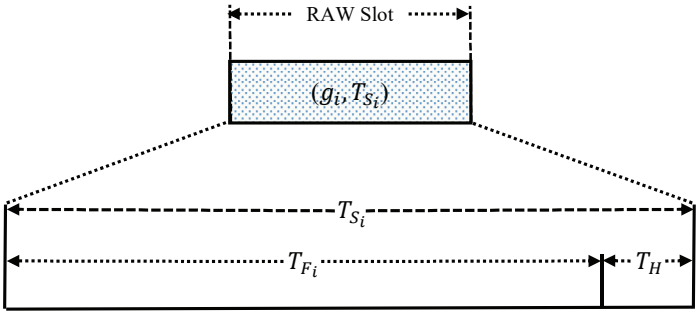


Figure 4.2: Structure of the RAW slot.

within the RAW, we consider only the non-crossing case of the RAW slot boundary. That is, any station in a group cannot cross its assigned RAW slot boundary during the transmission of a packet. To guarantee that a packet transmission will not cross the RAW slot boundary, we define a holding period T_H at the end of the RAW slot. Hereafter, as illustrated in Figure 4.2, the length T_{S_i} of the RAW slot is split into two periods: the free access period T_{F_i} during which stations within the group can contend to access the channel and the holding period T_H during which stations can no longer initiate a packet transmission. Hence, T_H should be set to the time required to complete one successful packet transmission. We then define $T_H = DIFS + \varphi$.

Therefore, within a tagged group i, the g_i stations can start their transmission only during the free access period given by $T_{F_i} = T_{S_i} - T_H$.

Note that the free access period T_{F_i} should be strictly positive, i.e., the RAW slot should be at least sufficient for one transmission. Additionally, we assume that all packets have the same fixed length.

As proven in Chapter 3-Proposition 3.1, the probability τ that a station will transmit a packet in an idle slot is given by

$$\tau = \frac{2p\left((2p-1)(m(p-1)+p-2)p^m+2\right)-2}{W_0(p-1)^2\left(2^{m+1}p^{m+1}-1\right)+2p\left((2p-1)(m(p-1)+p-2)p^m+2\right)-2} \tag{4.1}$$

The probability p that the transmitted packet encounters a collision is the probability that at least one of the $g_i - 1$ remaining stations transmits in the same time slot. Knowing that each station transmits with probability τ, the conditional collision probability p is expressed as follows

$$p = 1 - (1-\tau)^{g_i - 1} \tag{4.2}$$

The nonlinear system formed by (4.1) and (4.2) can be solved numerically (e.g., using Newton's method [78]) to obtain the probabilities τ and p.

Let $P_{tr}^{g_i}$ be the probability that at least one station transmits a packet in the considered slot. Since g_i stations are contending to access the channel, and each one transmits with probability

τ, we have

$$P_{tr}^{g_i} = 1 - (1-\tau)^{g_i} \tag{4.3}$$

The probability $P_s^{g_i}$ for a packet to be transmitted successfully is given by the probability that exactly one station transmits, conditioned on the fact that at least one station transmits, i.e.,

$$P_s^{g_i} = \frac{g_i \tau (1-\tau)^{g_i - 1}}{P_{tr}^{g_i}} \tag{4.4}$$

To derive the throughput of the RAW slot period, we track the transmissions that occurred up to T_S. While the stations within the RAW slot start contending to access the medium, each slot can be seen in a binary way where it is either idle or busy. A busy slot may contain a successful transmission if only one station is transmitting or a collision otherwise. Hereafter, following the proven result in Chapter 3-Proposition 3.2, we have the expected number of busy slots occurred during a RAW slot of duration T_S allocated for g stations given by

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\left\lfloor \frac{T_S}{\beta} \right\rfloor} \sum_{j=0}^{\left\lfloor \frac{T_S - k\cdot\beta}{\sigma} \right\rfloor} \binom{j+k-1}{j} P_{tr}^k (1-P_{tr})^j \tag{4.5}$$

Now that we obtained the expected number of busy slots within the RAW slot time interval $[0, T_S]$, we can distinguish the average number of busy slots containing successful transmissions by multiplying $\mathrm{E}[N_{T_S}]$ by the probability of success defined in (4.4). In addition, multiplying the result by T_{DATA} yields the portion of time used for successful transmissions of data frames during the RAW slot duration T_S. Therefore, we define the normalized throughput within the RAW slot as follows:

$$S_{\mathrm{RS}} = \frac{\mathrm{E}[N_{T_S}] \cdot P_s \cdot T_{\mathrm{DATA}}}{T_S} \tag{4.6}$$

Additionally, the energy efficiency within a RAW slot representing the ratio of energy used for successful transmissions to the total energy wasted during the RAW slot time, is defined as follows:

$$E_{eff}^{T_S} = \frac{E_s^{T_S}}{E_s^{T_S} + E_c^{T_S} + E_{idle}^{T_S}}, \tag{4.7}$$

where $E_s^{T_S}$, $E_c^{T_S}$, and $E_{idle}^{T_S}$ are respectively the energy consumptions by successful transmissions, collisions, and idle listening, and are given in Chapter 3 by (3.38), (3.39), and (3.40), respectively.

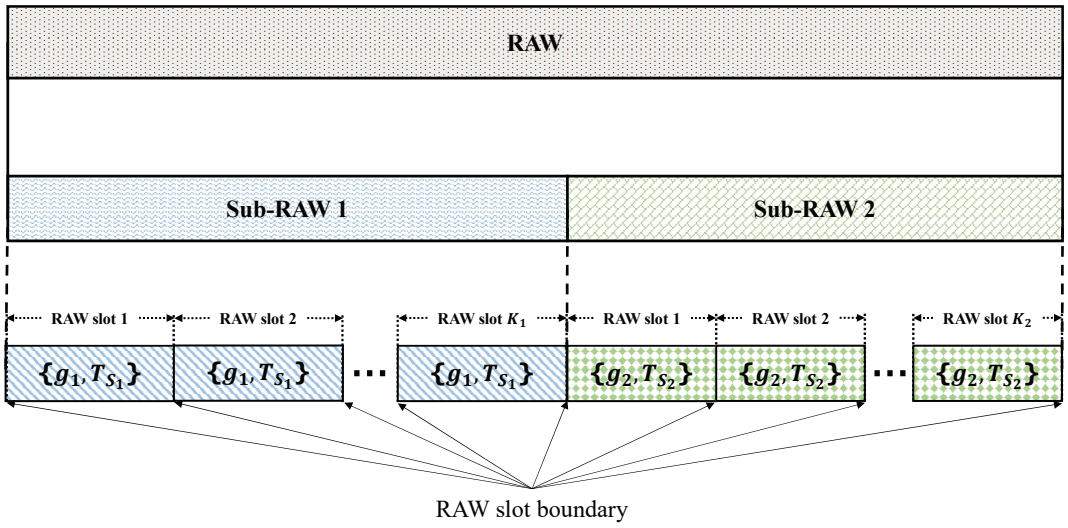


Figure 4.3: RAW grouping strategy.

4.4 RAW Grouping Scheme

According to the standard in [4], the AP should broadcast the grouping configuration information to all stations at the beginning of each beacon interval. Each RAW consists of $K \leq 64$ RAW slots, and each station should only wake up during its designated RAW slot. However, determining the number of groups within a RAW, the number of stations in each group, and the duration of the RAW slot assigned to each group is not an obvious task.

To alleviate the channel contention process while keeping fairness among stations in each RAW slot, we propose a RAW grouping scheme where the AP distributes one tagged RAW into K RAW slots, K being an integer between 1 and 64.

A RAW is defined by two parameters, N_R: the number of stations assigned to the RAW, and T_R: the time length of the RAW. As the IEEE 802.11ah aims to support thousands of nodes, it is mandatory to alleviate the contention of N_R stations within the limited channel access time interval T_R allocated to them. Note that in order to create K groups, both N_R and T_R have to be divided by K. Here, we may have to deal with the case where the contending stations N_R may not be divisible by K. We then need to define a grouping scheme adapted to any possible value of K while guaranteeing the best RAW utilization rate.

Suppose a fixed value for the number of groups K in the RAW. To guarantee that the resulting number of stations g_i assigned to each group i after splitting the total N_R stations within the RAW is always an integer, we create two Sub-RAWs within the RAW as an extension of the model in [43]. We split stations among several groups within two types of RAW slots, each type providing service to a different number of assigned stations and different duration. Hereafter, the RAW can be seen as two Sub-RAWs, where each one contains a uniform station grouping, as depicted in Figure 4.3.

For a fixed number of stations N_R, we have

$$N_R = (N_R \bmod K) \cdot \left(\left\lfloor \frac{N_R}{K} \right\rfloor + 1 \right) + (K - (N_R \bmod K)) \cdot \left\lfloor \frac{N_R}{K} \right\rfloor \tag{4.8}$$

Hence, we can divide the RAW into two Sub-RAWs: 1) Sub-RAW 1: it consists of $K_1 = N_R \bmod K$ RAW slots, each containing $g_1 = \left\lfloor \frac{N_R}{K} \right\rfloor + 1$ stations. And 2) Sub-RAW 2: it consists of $K_2 = K - (N_R \bmod K)$ RAW slots, each being allocated for $g_2 = \left\lfloor \frac{N_R}{K} \right\rfloor$ stations.

We now need to allocate a channel access duration to each RAW slot. Let T_{S_1} and T_{S_2} be the duration of a RAW slot in Sub-RAW 1 and Sub-RAW 2, respectively. Suppose that $K_1 \neq 0$, i.e., N_R is not divisible by K (the case of $K_1 = 0$ is discussed in Remark 4.1). For fairness purposes among stations in both Sub-RAW 1 and Sub-RAW 2, we define T_{S_1} and T_{S_2} according to the ratio of stations in each Sub-RAW. Hence, we define

$$T_{S_1} = \frac{g_1}{N_R} T_R \tag{4.9}$$

and

$$T_{S_2} = \frac{g_2}{N_R} T_R \tag{4.10}$$

Remark 4.1. *If N_R is divisible by K, then $K_1 = 0$. In this case Sub-RAW 1 will not be created, and all stations will be assigned uniformly to the K_2 RAW slots of Sub-RAW 2, where $K_2 = K$, $g_2 = \frac{N_R}{K}$ and $T_{S_2} = \frac{T_R}{K}$. In other words, the grouping within the RAW is created by splitting uniformly the RAW duration T_R and the contending stations N_R among K groups.*

To sum up, Algorithm 4.1 defines our proposed grouping scheme. After defining all grouping variables for a given number of groups K, we represent the RAW configuration by the parameters of the two Sub-RAWs: (K_1, g_1, T_{S_1}) and (K_2, g_2, T_{S_2}). We note

$$\text{RAW} \triangleq \{(K_1, g_1, T_{S_1}), (K_2, g_2, T_{S_2})\} \tag{4.11}$$

4.5 RAW Performance

After we have defined the RAW grouping scheme and the performance of each RAW slot, we can now derive the performance of an entire RAW with several RAW slots assigned following the proposed grouping scheme.

Algorithm 4.1 RAW grouping scheme.

Input: N_R, T_R, K
Output: $\mathcal{R}_1$, $\mathcal{R}_2$

1: $K_1 \leftarrow N_R \bmod K$
2: $K_2 \leftarrow K - K_1$
3: $g_1 \leftarrow \left\lfloor \frac{N_R}{K} \right\rfloor + 1$
4: $g_2 \leftarrow \left\lfloor \frac{N_R}{K} \right\rfloor$
5: $T_{S_1} \leftarrow \frac{g_1}{N_R} \cdot T_R$
6: $T_{S_2} \leftarrow \frac{g_2}{N_R} \cdot T_R$
7: **Define:** $\mathcal{R}_1 \triangleq \{K_1, g_1, T_{S_1}\}$
8: **Define:** $\mathcal{R}_2 \triangleq \{K_2, g_2, T_{S_2}\}$

4.5.1 Throughput

The normalized throughput within a RAW composed of K RAW slots is defined by the ratio of time used for successful transmission within the RAW duration T_R. Considering RAW definition (4.11), we have

$$S(K) = \sum_{i=1}^{2} \frac{K_i \cdot \mathrm{E}\left[N_{T_{S_i}}\right] \cdot P_s^{g_i} \cdot T_{\mathrm{DATA}}}{T_R} \tag{4.12}$$

where $\mathrm{E}\left[N_{T_{S_i}}\right]$ is the expected number of busy slots within a RAW slot of Sub-RAW i.

4.5.2 RAW Energy Efficiency

The energy efficiency of a RAW composed of K RAW slots is defined by the ratio of energy used for successful transmissions during the RAW duration T_R. Considering the RAW definition (4.11), we have

$$E_{eff}(K) = \sum_{i=1}^{2} \frac{K_i \cdot E_s^{T_{S_i}}}{E_{total}}\,, \tag{4.13}$$

where $E_s^{T_{S_i}}$ is the energy consumed by successful transmission width a RAW slot of Sub-RAW i. E_{total} is the total energy wasted during the RAW and it is given by

$$E_{total} = K_1 \left(E_s^{T_{S_1}} + E_c^{T_{S_1}} + E_i^{T_{S_1}}\right) + K_2 \left(E_s^{T_{S_2}} + E_c^{T_{S_2}} + E_i^{T_{S_2}}\right), \tag{4.14}$$

where $E_c^{T_{S_i}}$ and $E_{idle}^{T_{S_i}}$ are the energy wasted during collisions and idle slots, respectively, within a RAW slot of Sub-RAW i.

In an IEEE 802.11ah network, the AP defines the RAW configuration at the beginning of each beacon interval. The implementation of this task is central to the success of the re-

source allocation mechanisms introduced by the IEEE 802.11ah standard. Since the main goal is to define mechanisms to get the most from network resources, the relevance of our findings comes from the fact that from the average number of transmissions within a RAW slot, we can deduce the RAW throughput $S(K)$. The closed-form expression (4.5) answers the needs of real-time RAW configuration and proves the effectiveness of our approach compared to the high complexity presented by other approaches such as [40] and [47].

4.6 Tuning-up RAW Configuration

In Section 4.4, we proposed a RAW grouping scheme and defined the RAW configuration for a given grouping K. However, choosing the number of groups K providing the best performance results is not a straightforward task. In this section, we provide an algorithm to determine the value of K, yielding the best performance results. More specifically, we shall obtain a RAW configuration allowing us to maximize the overall RAW throughput.

Consider one RAW of length T_R, containing N_R contending stations. We seek to optimize the RAW and derive the optimal values of RAW parameters in (4.11) that yields the best performance for the RAW in terms of throughput. Note that the channel access time intervals allocated to RAW slots in both Sub-RAWs should be sufficient at least for one transmission, as described in Section 4.3. Moreover, the number of groups K should not exceed $\min\{N_R, 64\}$.

Hence, we define the grouping set by

$$\mathcal{K}_{\mathcal{R}} = \{1, \cdots, \min\{N_R, 64\}\} \tag{4.15}$$

Hereafter, to optimize the RAW, we seek the optimal number of groups K_{op} that yields the best performance within the RAW period. Denoting by $Perf(K)$ the RAW performance when composed of K RAW slots, which represents either throughput or energy efficiency. Thus, the optimal RAW grouping is given by:

$$K_{op} = \underset{K \in \mathcal{K}_{\mathcal{R}}}{\arg\max}\, Perf(K) \tag{4.16}$$

and the corresponding maximum performance is given by:

$$Perf(K_{op}) = \max_{K \in \mathcal{K}_{\mathcal{R}}} Perf(K) \tag{4.17}$$

To find the optimal values in (4.16) and (4.17) and apply the optimal station grouping, we propose an optimal RAW configuration algorithm, see Algorithm 4.2. The required inputs are the two elements that characterize the RAW: duration of the RAW T_R and the number of stations N_R in this RAW. The algorithm is composed of two main phases: Phase 1 is dedicated

Algorithm 4.2 Optimal RAW configuration.

Require: Duration of the RAW: T_R
Require: Number of stations in the RAW: N_R
1: $\mathcal{K}_{\mathcal{R}} \leftarrow \{1, \cdots, \min\{N_R, 64\}\}$
2: $K_{op} \leftarrow \arg\max_{K \in \mathcal{K}_{\mathcal{R}}} Perf(K)$
3: $Perf_{op} \leftarrow Perf(K_{op})$
4: $[\mathcal{R}_1, \mathcal{R}_2] \leftarrow RAW\_Grouping\_Scheme(N_R, T_R, K_{op})$
5: **Define:**

$$\begin{aligned} \text{RAW}_{\text{op}} &\triangleq \{\mathcal{R}_1, \mathcal{R}_2\} \\ &\triangleq \left\{ \left(K_1^{op}, g_1^{op}, T_{S_1}^{op}\right), \left(K_2^{op}, g_2^{op}, T_{S_2}^{op}\right) \right\} \end{aligned}$$

to finding the optimal number of groups K_{op} within the RAW and the corresponding throughput, i.e., the optimums in (4.16) and (4.17). And Phase 2 is dedicated to configuring the RAW grouping according to the optimum K_{op} obtained in Phase 1. Thus, to define the optimal RAW configuration, we call Algorithm 4.1 with entries N_R, T_R and K_{op} in order to define the grouping scheme corresponding to K_{op}. We obtain the structure of Sub-RAW 1 and Sub-RAW 2, and hence the optimal RAW configuration is defined as

$$\text{RAW}^{op} \triangleq \left\{ \left(K_1^{op}, g_1^{op}, T_{S_1}^{op}\right), \left(K_2^{op}, g_2^{op}, T_{S_2}^{op}\right) \right\} \tag{4.18}$$

Hereafter, we perform the optimal grouping assignment, which creates Sub-RAW 1 and Sub-RAW 2 for the optimal total number of groups K_{op} following the grouping scheme procedure described above.

4.7 Evaluation and Analysis

In this section, we analyze and evaluate the performance of our proposed grouping scheme and RAW configuration algorithm in terms of the throughput by using the MATLAB software. The numerical values of the system parameters are presented in Table 4.2. To validate our model and results, we also provide simulation results obtained using a simulator developed with MATLAB. As seen in the figures, the results obtained from our model closely match the simulation results, which validates our model.

4.7.1 Impact of Grouping

This section evaluates and examines the impact of grouping stations and dividing the RAW period into several RAW slots. We deploy our proposed grouping scheme and investigate its impact on alleviating the channel contention and increasing the network performance.

Table 4.2: Network setup parameters.

Parameter	Value	Parameter	Value
Data rate	1 Mbps	σ	52 µs
Payload	512 bits	CW_{min}	16
MacHeader	272 bits	CW_{max}	1024
SIFS	160 µs	m	6
DIFS	264 µs	e_{tx}	2 mJ
ACK	112 bits	e_l	1.3 mJ
T_{PLCP}	80 µs	e_{idle}	0.04 mJ

4.7.1.1 Throughput

We consider a RAW consisting of a given K RAW slots configured following our proposed grouping scheme and evaluate the overall throughput, representing the portion of time used to successfully deliver data frames. Figure 4.4 depicts the normalized throughput in terms of stations participating in a $T_R = 100$ ms RAW using three different values of K: 15, 20 and 25 groups. For a better understanding of the behavior of RAW throughput, we present in Table 4.3 detailed data results for $K = 15$. More stations within the RAW may cause an increase or decrease in the overall RAW throughput. The throughput degradation is due to the higher contention within each group, which eventually decreases the throughput within each RAW slot, as in the table when N_R increases from 61 to 67 stations. However, as defined in our proposed grouping scheme, the duration of the RAW slot in each Sub-RAW is allocated proportionally to the participating stations within the Sub-RAW. Hence, more stations within the RAW (68 to 75 stations in this example) can also cause a smaller duration for the majority of RAW slots, which eliminates the unnecessary additional time and eventually increases the throughput.

As can be seen from the figure, choosing the best grouping for the RAW is not apparent. We notice that increasing the number of groups does not always provide better performance results in all system configurations. Hereafter, the optimal configuration depends on both the number of stations and the RAW duration. Among the three groupings depicted in the figure, we observe that 15 groups are the best choice for 50 stations, whereas $K = 20$ and $K = 25$ are optimal for 66 and 80 stations, respectively.

Configuring the RAW into more groups may also decrease the RAW performance due to the overhead introduced by the protocol header and the holding period. As we observe in the figure, dividing the RAW into 25 groups reports worse performance results for more than 70 stations. It is evident that there is a trade-off between the channel resources used by the overhead and the number of backoff slots wasted during the conflict resolution.

In Figure 4.5, we present the same three groupings as in Figure 4.4, but in terms of RAW duration while we fix the stations within the RAW at $N_R = 100$ stations. We observe the same

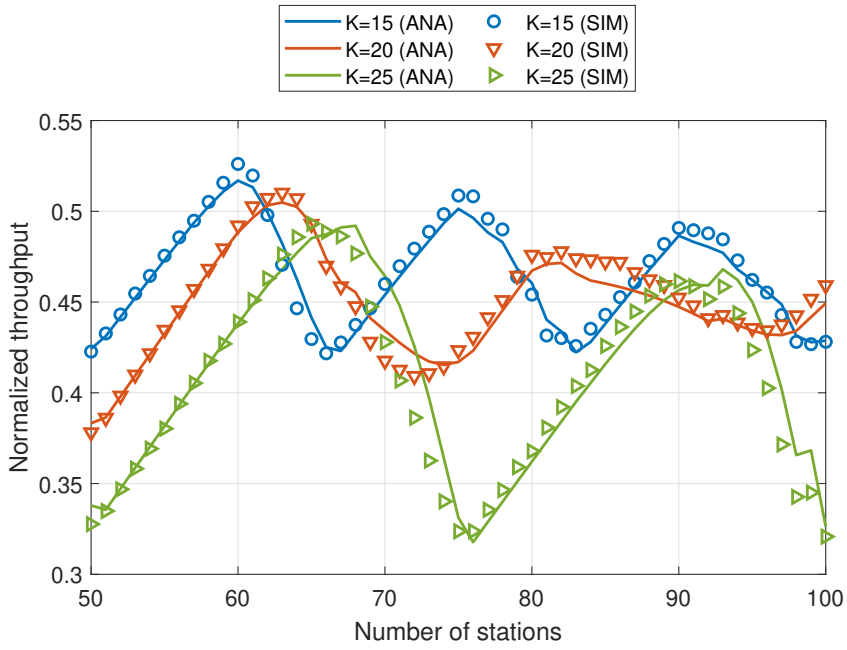


Figure 4.4: Impact of the RAW grouping on the normalized throughput in terms of the stations within a RAW of length $T_R = 100$ ms with a 1024-bits payload.

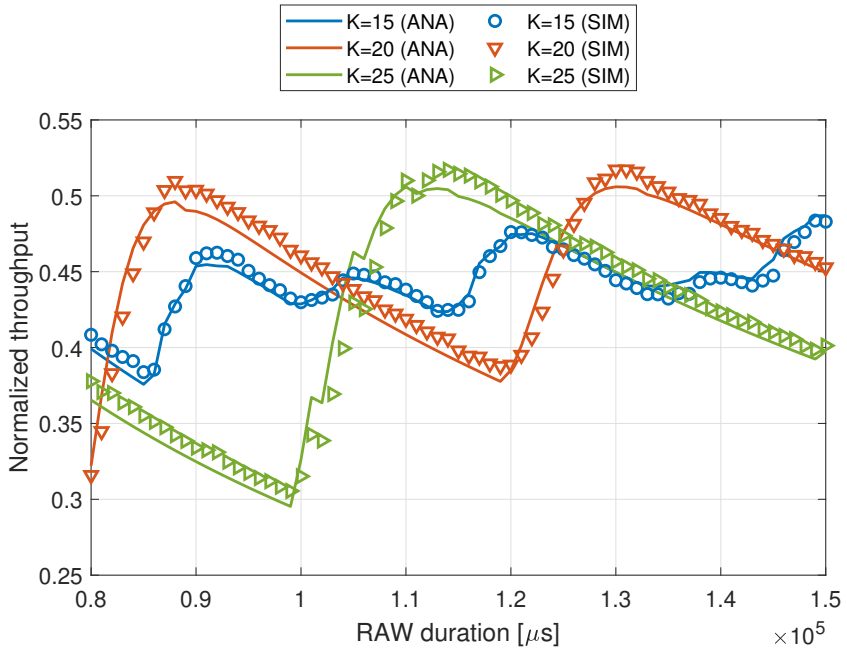


Figure 4.5: Impact of the RAW grouping on the normalized throughput in terms of the RAW duration with $N_R = 100$ stations and $L = 1024$ bits.

behavior of the throughput going up and down repeatedly. This is explained by the usage of the duration of each RAW slot. Increasing the RAW duration might not benefit the RAW when the transmissions are still the same. The additional time will be wasted, which affects the throughput negatively.

We can conclude that the grouping of a RAW is a very sensitive task as it is easily affected by the RAW duration and the number of stations participating in the RAW. Hence, it is mandatory to configure the RAW optimally in order to achieve its best usage for each different

Table 4.3: RAW configuration and associated performance for different number of stations within a RAW of length $T_R = 100$ ms composed of $K = 15$ RAW slots with a 1024-bits payload.

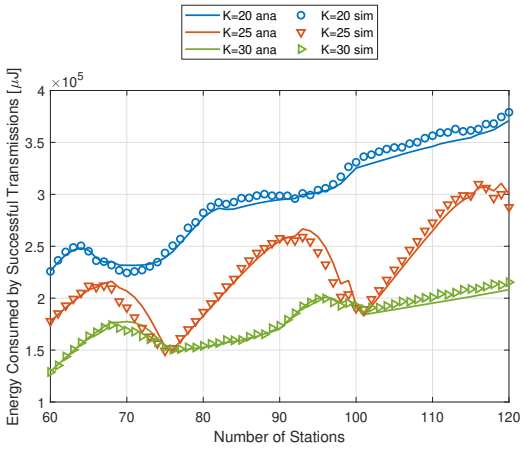
N_R	sub-RAW 1					sub-RAW 2					RAW throughput
	K_1	g_1	T_{S_1} (µs)	$\mathrm{E}\left[N_{T_{S_1}}\right]$	S_1	K_2	g_2	T_{S_2} (µs)	$\mathrm{E}\left[N_{T_{S_2}}\right]$	S_2	
60	**-**	**-**	**-**	**-**	**-**	**15**	**4**	**6666.67**	**2.9470**	**0.3847**	**0.5169**
61	1	5	8196.72	3.3673	0.3435	14	4	6557.38	2.9040	0.3854	0.5133
62	2	5	8064.52	3.0748	0.3188	13	4	6451.61	2.8318	0.3820	0.4996
63	3	5	7936.51	3.0000	0.3161	12	4	6349.21	2.7168	0.3724	0.4824
64	4	5	7812.50	3.0000	0.3211	11	4	6250.00	2.5487	0.3549	0.4627
65	5	5	7692.31	3.0000	0.3261	10	4	6153.85	2.3334	0.3300	0.4414
66	6	5	7575.76	3.0000	0.3311	9	4	6060.61	2.1174	0.3040	0.4251
67	**7**	**5**	**7462.69**	**2.9999**	**0.3361**	**8**	**4**	**5970.15**	**2.0000**	**0.2915**	**0.4231**
68	8	5	7352.94	2.9999	0.3411	7	4	5882.35	2.0000	0.2959	0.4334
69	9	5	7246.38	2.9997	0.3461	6	4	5797.10	2.0000	0.3002	0.4437
70	10	5	7142.86	2.9993	0.3511	5	4	5714.29	2.0000	0.3046	0.4539
71	11	5	7042.25	2.9983	0.3560	4	4	5633.80	2.0000	0.3089	0.4641
72	12	5	6944.44	2.9963	0.3608	3	4	5555.56	2.0000	0.3133	0.4742
73	13	5	6849.32	2.9919	0.3653	2	4	5479.45	1.9999	0.3176	0.4838
74	14	5	6756.76	2.9881	0.3698	1	4	5405.41	1.9999	0.3220	0.4934
75	**-**	**-**	**-**	**-**	**-**	**15**	**5**	**6666.67**	**2.9749**	**0.3731**	**0.5014**

scenario.

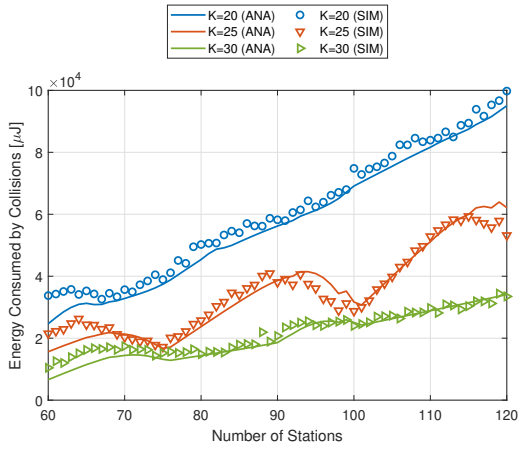
Therefore, finding the overall best grouping size is not a straightforward task. From the above results, we realize that the overhead in each RAW slot may have an impact on the performance of the grouping scheme. Since the amount of overhead is closely related to the size of the payload, we should therefore explore the impact of the payload size on the performance of our proposed grouping scheme.

4.7.1.2 Energy Consumption

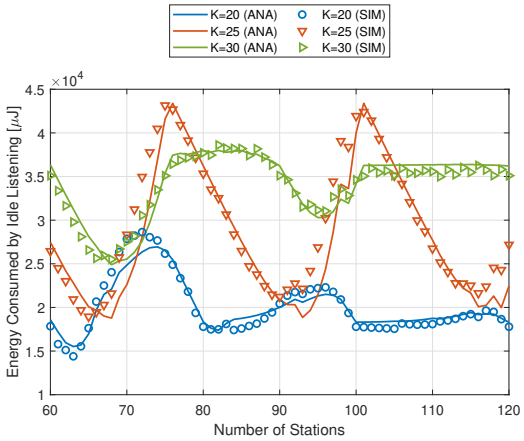
We now examine the impact of our proposed grouping scheme in alleviating the channel contention and decreasing the energy consumption within the RAW. Figure 4.6 depicts the RAW performance in terms of stations participating in a $T_R = 100$ ms RAW using three different values of K: 20, 25 and 30 groups. We present in Figure 4.6a, Figure 4.6b and Figure 4.6c the energy consumption due to successful transmissions, collisions and idle listening, respectively. More stations within the RAW may cause in an increasing or decreasing of energy consumption. The increasing of energy consumption is due to the higher contention within each group. However, as defined in our proposed grouping scheme, the duration of the RAW slot in each Sub-RAW is allocated proportionally to the participating stations within the Sub-RAW. Hence, more stations within the RAW can also cause in smaller duration for the majority of RAW slots, which eliminates the unnecessary additional time and eventually decreases the overall energy



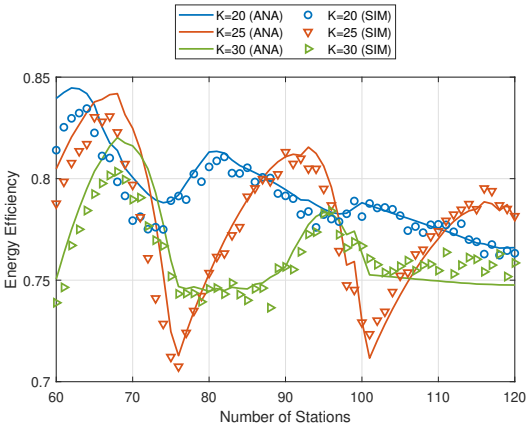
(a) Energy consumption of successful transmissions.



(b) Energy consumption of collisions.



(c) Energy consumption of idle listening.



(d) Energy Efficiency.

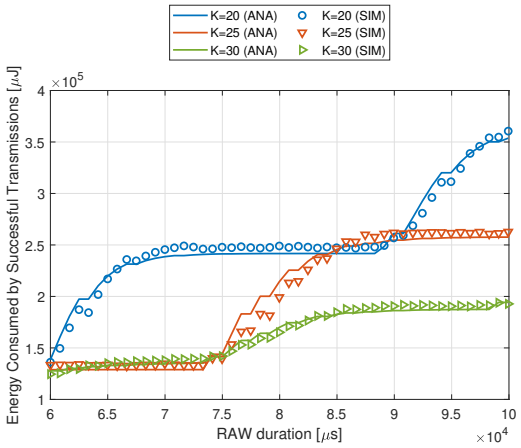
Figure 4.6: Impact of RAW grouping on the energy efficiency in terms of stations within a RAW of length $T_R = 100\,\text{ms}$ with a 1024-bits payload.

consumption within the RAW.

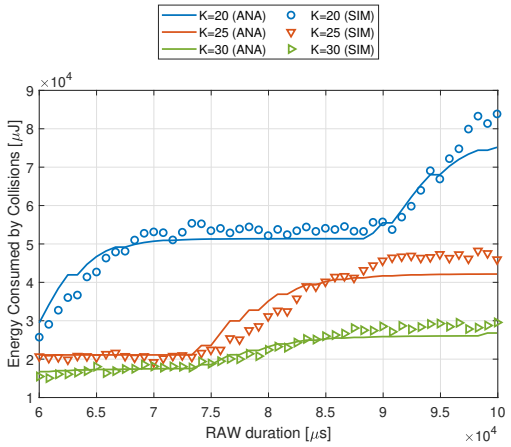
Such behavior is directly affecting the RAW energy efficiency as displayed in Figure 4.6, where we notice that choosing the best grouping for the RAW is not apparent. Increasing the number of groups does not always provide better performance results in all system configurations. Hereafter, the optimal configuration depends on both the number of stations and the RAW duration.

In Figure 4.7, we present the same three groupings as in Figure 4.6, but in terms of RAW duration while we fix the stations within the RAW at $N_R = 100$ stations. We observe the same behavior of the energy consumption and energy efficiency going up and down repeatedly. This is explained by the usage of the duration of each RAW slot. Increasing the RAW duration might not benefit the RAW when the transmissions are still the same. The additional time will be wasted, which affects the energy efficiency negatively.

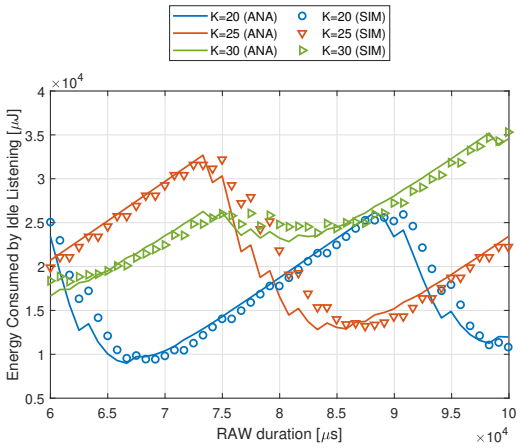
We can conclude that the grouping of a RAW is a very sensitive task as it is easily affected



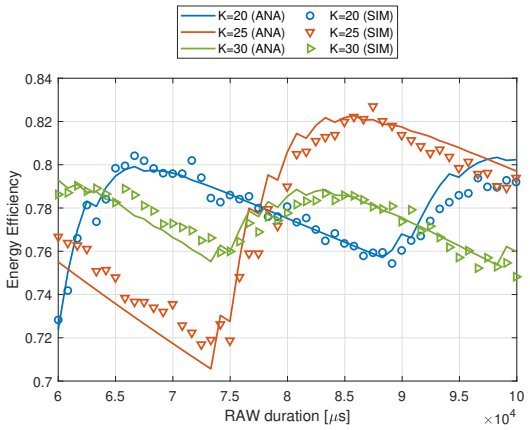
(a) Energy consumption of successful transmissions.



(b) Energy consumption of collisions.



(c) Energy consumption of idle listening.



(d) Energy Efficiency.

Figure 4.7: Impact of RAW grouping on energy efficiency in terms of RAW duration with $N_R = 100$ stations and $L = 512$ bits.

by RAW duration and number of stations participating in the RAW. Hence, it is mandatory to configure the RAW optimally in order to achieve its best usage for each different scenario.

4.7.2 RAW Grouping Performance

We examine the RAW performance for all possible groupings. We consider a RAW of length 350 ms and 70 contending stations. Figure 4.8 depicts the normalized RAW throughput for all possible groupings defined by our proposed grouping scheme and four different payload sizes: 256, 512, and 1024 bits.

As can be seen in Figure 4.8, a different optimal throughput is reached for every payload size. As the number of groups increases, the grouping scheme proves effective in reducing the collisions. However, as the number of groups is further increased, the normalized throughput decreases. This is mainly due to the holding period, which is not totally used for packet transmission. Recall that in order to be transmitted, a packet transmission must start at the latest at

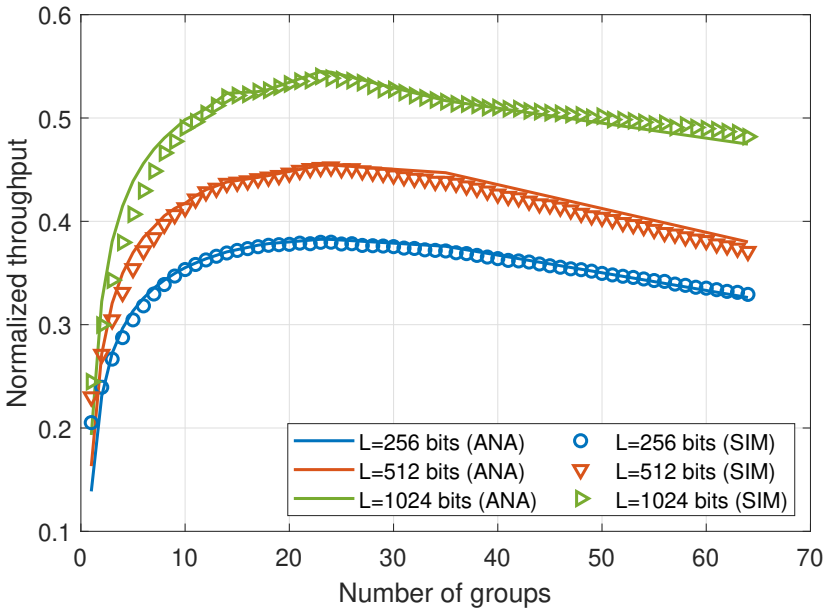


Figure 4.8: Normalized throughput versus number of groups for different payloads within a RAW of duration $T_R = 350$ ms, containing $N_R = 70$ stations.

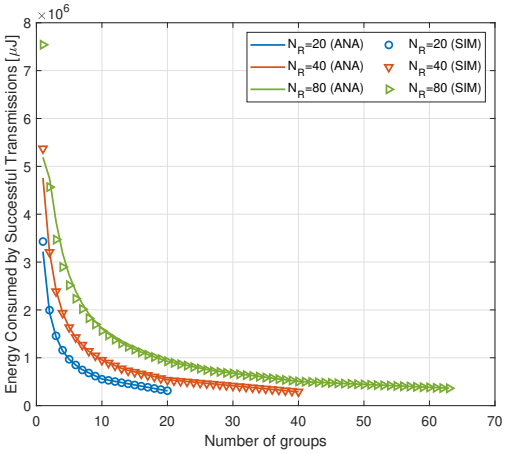
the beginning of the holding period. In the case that packet transmission is on its way at the time when the holding period starts, the ongoing packet transmission should end before the end of the holding period. In this case, once the packet transmission is completed, the time elapsed between the end of the transmission and the end of the holding period is wasted. It is clear that in the case of a longer payload, it will result in a higher waste of the channel resources, i.e., a longer idle period, and hence faster degradation of the throughput as more groups are employed.

From the analysis of the depicted results, we may conclude:

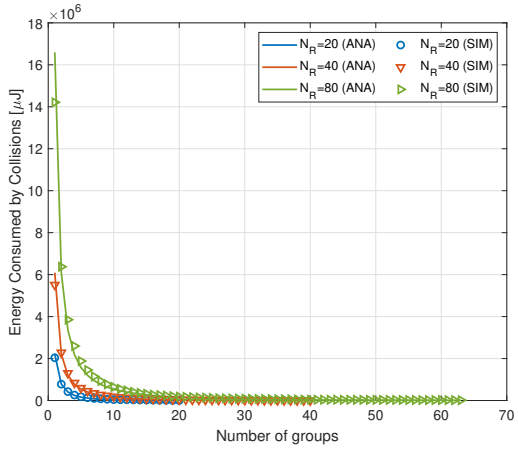
- **Maximum throughout:** Deploying a longer payload increases the maximum throughput and reduces the overhead introduced by the headers and the various time periods used by the medium access mechanism, e.g., DIFS. In the case of a longer payload, once a station initiates the transmission of its packet, it occupies the channel for a longer period of time. In contrast, transmissions of a small payload are significantly suffering from the headers' overhead. It is worth mentioning that such headers problem is one of the recent research topics, where it is required to find solutions to shorten the headers.

- **Throughput for large grouping:** We observe that for longer payload size, the RAW throughput decreases sharply when the number of groups increases. This is due to the fact that longer payloads result in a higher value of the holding period T_H. Since this period is included in each RAW slot as a measure to avoid the crossing of boundaries between two consecutive RAW slots, having more groups increases the wasted time in the total duration T_R of the RAW.

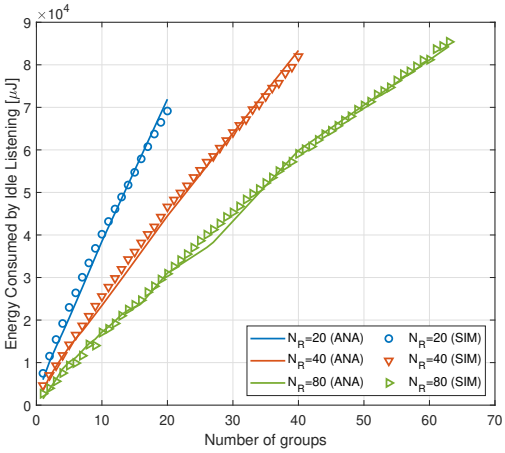
Hereafter, we examine the RAW performance over all possible grouping states in terms



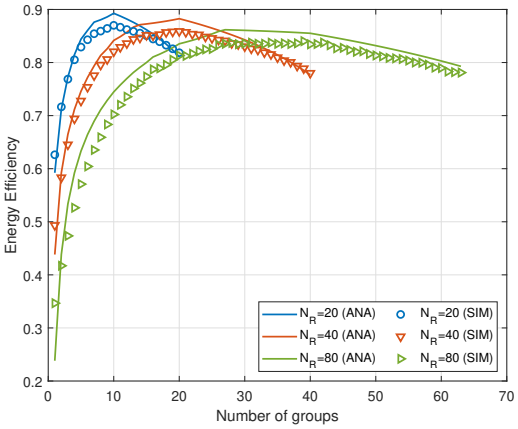
(a) Energy consumption of successful transmissions.



(b) Energy consumption of collisions.



(c) Energy consumption of idle listening.



(d) Energy Efficiency.

Figure 4.9: Energy consumption of different grouping within a 300 ms RAW for different number of stations with a 512-bits payload.

of energy consumption and efficiency. In Figure 4.9, we consider a RAW of duration $T_R = 300\,\text{ms}$ for $N_R = 20, 40, 80$ nodes and a payload size of 512 bits. More stations within the RAW yields higher energy consumption from transmissions and collisions, where the lowest number of nodes suffers from higher idle listening energy consumption. That is because of the lack of idle slots within a dense RAW, due to high contention. However, the best and worst energy efficiency is obtained for RAW consisting of the lowest and highest number of stations, respectively. Accordingly, a low number of stations achieve the maximum energy efficiency in a low number of groups. Similar to the previous results, the grouping scheme proves effective up to a certain value where the energy efficiency decreases gradually to values comparable to the ones obtained when fewer groups are used. While high collision probability is the main cause of poor performance in the case of few groups, the need to introduce a holding period in each RAW slot heavily penalizes the performance of the grouping scheme when the number of groups increases.

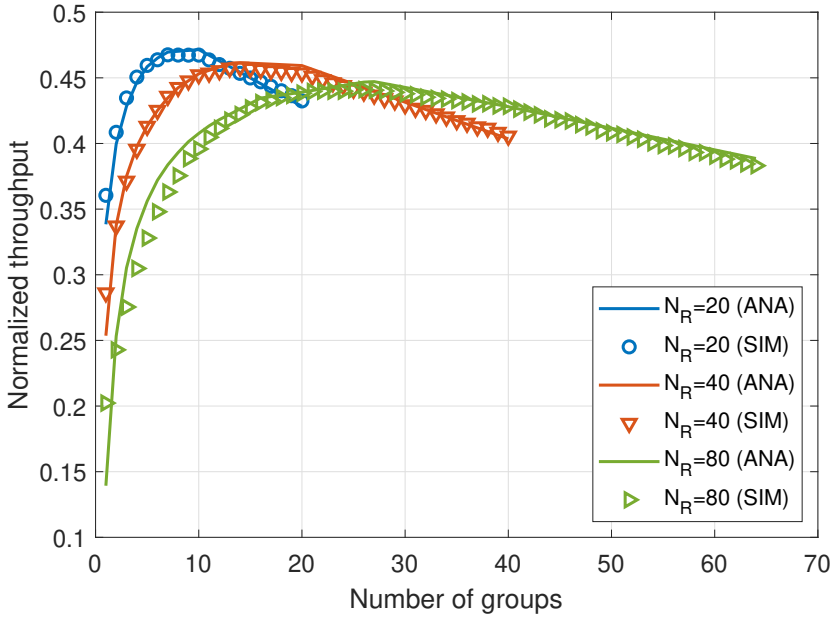


Figure 4.10: Normalized throughput in terms of the number of groups within a RAW of length $T_R = 300$ ms for different number of stations with a 512-bits payload.

4.7.3 Impact of Number of Stations and RAW Length

As stated in the previous section, our proposed RAW grouping scheme requires two prior inputs: the number of stations assigned to the RAW N_R, and the duration of the RAW T_R. Hereafter, we examine the impact of these two parameters on the RAW performance.

To analyze the impact of the number of stations assigned to the RAW, we examine the performance of a RAW of duration $T_R = 300$ ms for $N_R = 20, 40, 80$ nodes and a payload size of 512 bits. Figure 4.10 shows the normalized throughput vs. the number of groups for the three different values of N_R.

As can be seen in the figure, the best and worst results are obtained for networks consisting of the lowest and highest number of nodes, respectively. Accordingly, a low number of stations achieve the maximum throughput in a low number of groups. Similar to the previous results, the grouping scheme proves effective up to a certain value where the throughput decreases gradually to values comparable to the ones obtained when fewer groups are used. While high collision probability is the main cause of poor performance in the case of fewer groups, the need to introduce a holding period in each RAW slot heavily penalizes the performance of the grouping scheme when the number of groups increases.

To examine the impact of RAW duration, we consider a tagged RAW with 64 stations, and we explore the corresponding performance for different values of T_R: 240, 360, and 540 ms. As can be seen in Figure 4.11, a larger value of T_R turns out a higher throughput. This is because increasing the RAW length results in having more channel access time for each group, and hence a higher achievable throughput. Furthermore, the throughput decreases faster as the number of

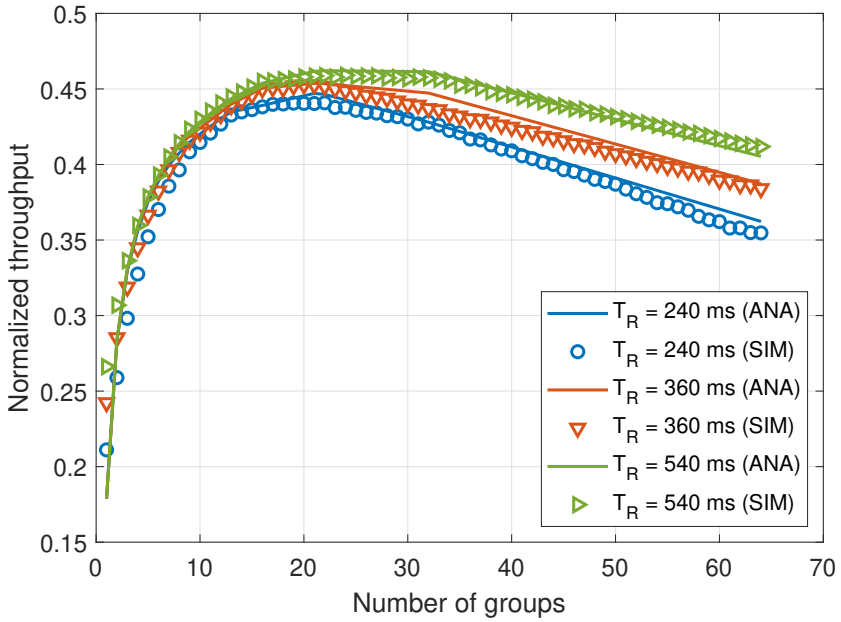


Figure 4.11: Normalized throughput in terms of the number of groups within a RAW of $N_R = 64$ stations for different RAW lengths with a 512-bits payload.

groups exceeds the optimal grouping K_{op} for a shorter RAW. This is mainly caused by the channel waste time added in each RAW slot, which is more significant in shorter RAW slots.

4.7.4 Proposed RAW Grouping Algorithm versus Conventional DCF

To evaluate the effectiveness of our proposed RAW configuration algorithm, we compare its performance with the results obtained when the legacy DCF is preferred.

4.7.4.1 Throughput

We first evaluate the RAW performance in terms of throughput. Figure 4.12 compares the performance of our proposed RAW configuration algorithm with the legacy DCF mechanism in terms of the number of stations participating in the RAW. We fix $T_R = 500$ ms and consider three payload sizes. As can be seen from the figure, the DCF throughput decreases very fast down to zero as the number of contending stations increases. However, in the case the RAW is configured by our proposed algorithm, the reported throughput is much better and decreases steadily as the number of stations increases. Additionally, the performance of our proposed algorithm is more stable over the dynamic variation of the RAW size, where it adapts every situation to the highest possible performance.

To analyze the relationship between the optimal grouping and the stations within the RAW, we present in Figure 4.13 the optimal grouping K_{op} associated with the maximum throughput depicted in Figure 4.12. One could assume that an optimal choice for saturated

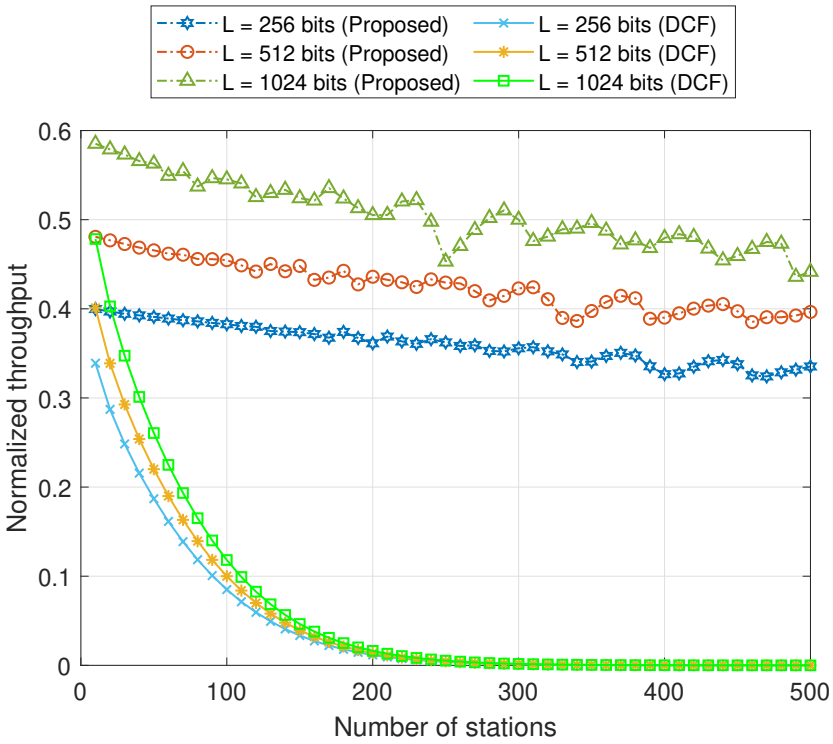


Figure 4.12: Throughput comparison of the proposed optimal RAW configuration with conventional DCF in terms of contending stations within an $T_R = 500$ ms RAW.

traffic would be one station per group. However, we observe that assigning every single station to a different group, i.e., $K_{op} = N_R$, is not optimal. This is because of the channel time wasted during the first backoff stage. Hence, it is better to allocate at least two stations for each RAW slot, which means K_{op} is always less or equal to $\left\lfloor \frac{N_R}{2} \right\rfloor$. On the other hand, a minimal value of K_{op} may lead to a higher number of collisions in each group, and hence a lower throughput. Hereafter, the optimal grouping K_{op} is very dependent on the total number of stations N_R allocated within the RAW.

Figure 4.14 shows the normalized throughput in terms of RAW duration for the proposed grouping scheme and the DCF mechanism with $N_R = 100$ stations participating in the RAW. As can be seen from the figure, the RAW duration does not have a significant impact on the performance of the DCF mechanism. However, our proposed algorithm reports much better results than the legacy DCF. This result is explained by the fact that our proposed algorithm profits from a higher value of the RAW length by dividing stations into more RAW slots and by allocating more time to each RAW slot. With such a grouping, each RAW slot contains fewer stations with a longer free access period and hence yields a higher RAW utilization rate, which eventually increases the overall throughput within the RAW.

We present in Figure 4.15 the optimal grouping K_{op} associated to the maximum throughput depicted in Figure 4.14. We see that for a short RAW, the optimal grouping K_{op} is lower. This is because a higher value will lead to RAW slots with a duration shorter than the length of one packet transmission, and hence a null throughput. Although a higher duration of T_S allows

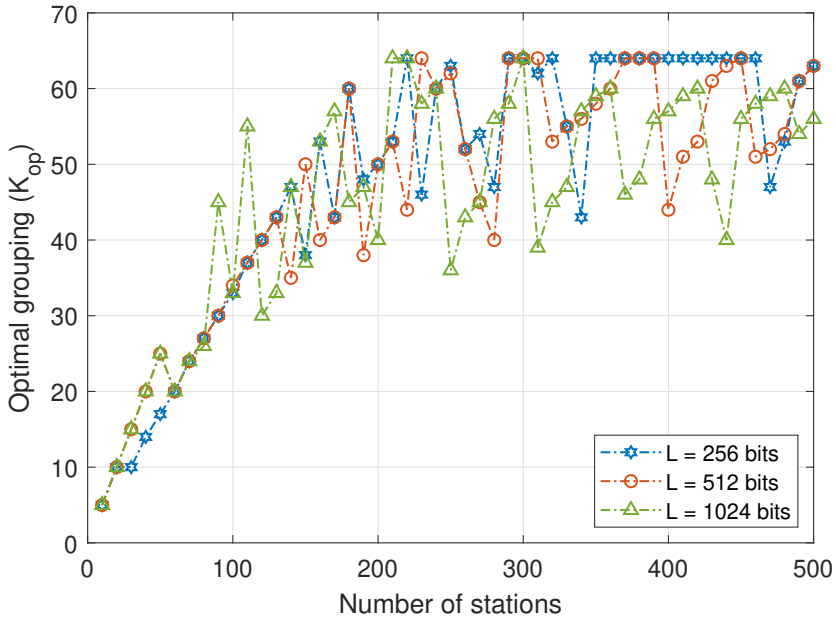


Figure 4.13: Optimal grouping in terms throughput and the number of contending stations within an $T_R = 310$ ms RAW.

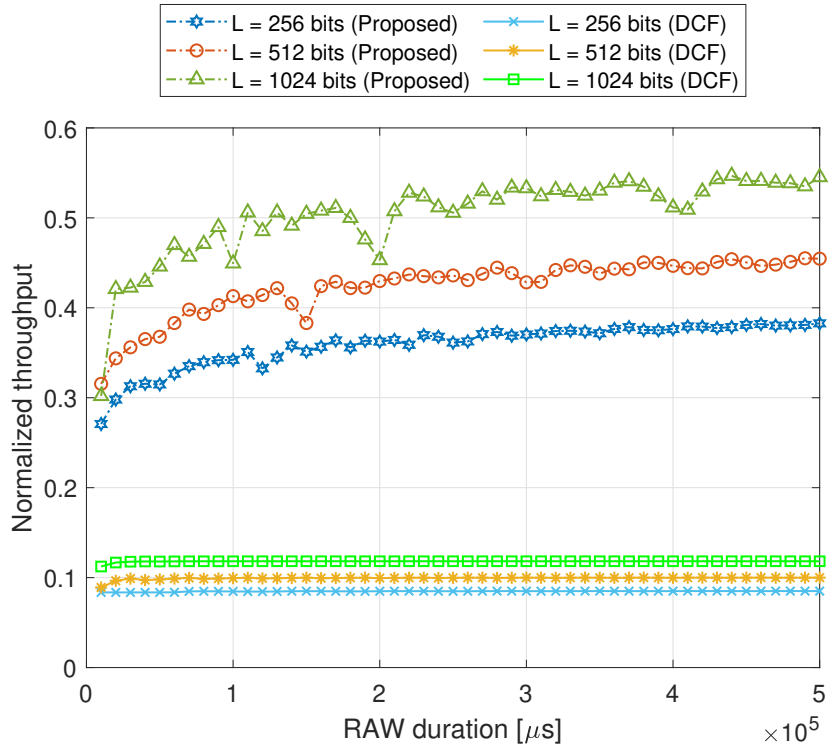


Figure 4.14: Throughput comparison of the proposed optimal RAW configuration with conventional DCF in terms of the RAW duration with $N_R = 100$ stations.

creating more groups, K_{op} is eventually limited to the value of half number of stations (50 stations in this example). That is for the same reason mentioned previously, which is the channel time wasted during the first backoff stage when there is one station per group.

The results obtained from the comparisons in Figure 4.12 and Figure 4.14 validate the performance of our proposed algorithm over the legacy DCF in increasing the network throughput.

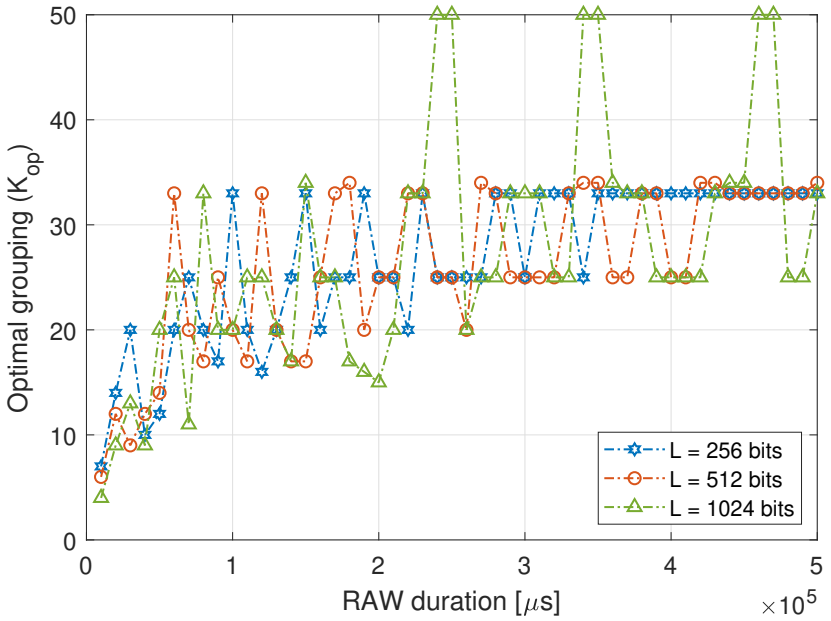


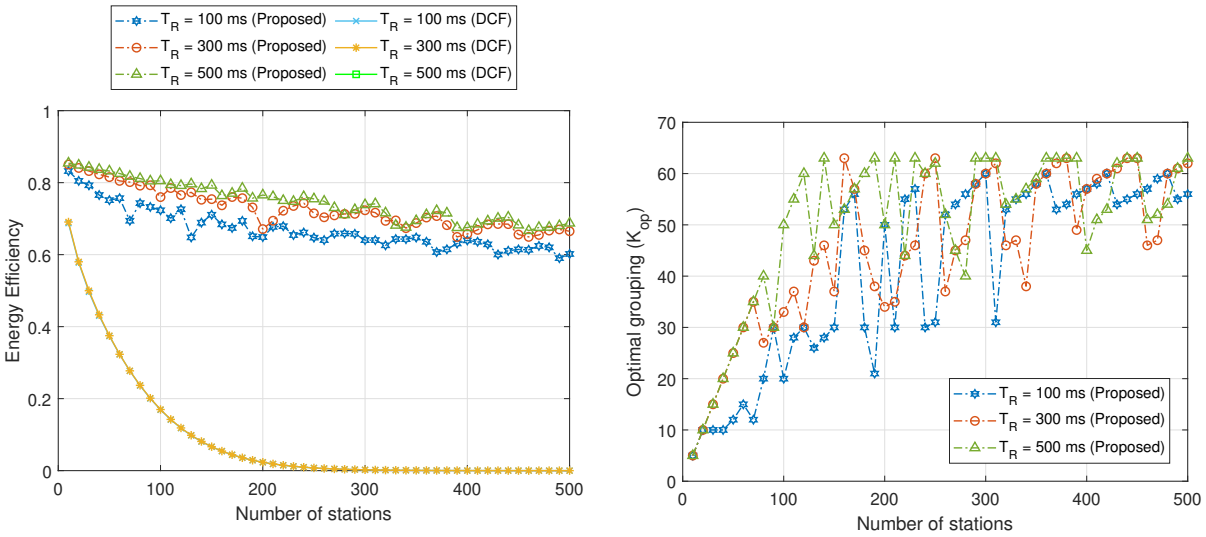
Figure 4.15: Optimal grouping in terms throughput and the RAW duration with $N_R = 100$ stations.

This is due to the reliability of our proposal in deriving for every RAW the best configuration according to our proposed grouping scheme, which maximizes the overall normalized throughput. Consequently, the combination of our proposed grouping scheme that sets the RAW parameters and our proposed algorithm to optimally configure those parameters proves effective in maximizing the overall RAW throughput in highly dense scenarios.

4.7.4.2 Energy Efficiency

We now explore the improvement of our proposal in terms of energy efficiency compared to DCF protocol. Figure 4.16 compares the performance of our proposed RAW configuration algorithm with the legacy DCF mechanism in terms of the number of stations participating in the RAW. We consider three different RAW lengths: 100 ms, 300 ms and 500 ms. As seen from Figure 4.16a, the energy efficiency provided by DCF decreases very fast down to zero as the number of contending stations increases. However, in the case the RAW is configured by our proposed algorithm, the reported energy efficiency is much better and decreases steadily as the number of stations increases. We can also see that a RAW with longer length yields better energy efficiency when the optimal configuration is applied. Whereas, it make no difference for DCF. Additionally, the performance of our proposed algorithm is more stable over the dynamic variation of the RAW size, where it adapts every situation to the highest possible performance.

To analyze the relationship between the optimal grouping and the stations within the RAW, we present in Figure 4.16b the optimal grouping K_{op} associated with the energy efficiency depicted in Figure 4.16a. One could assume that an optimal choice for saturated traffic would be station per group. However, We observe that assigning every single station to a differ-



(a) Energy efficiency.

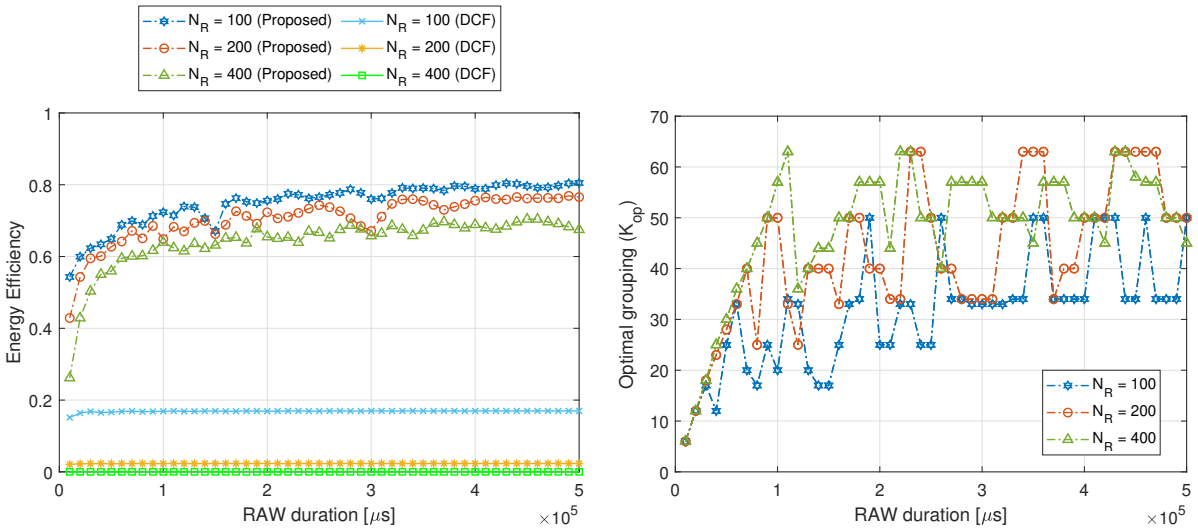
(b) Optimal grouping.

Figure 4.16: Energy efficiency comparison of the proposed optimal RAW configuration with conventional DCF in terms of contending stations.

ent group, i.e., $K_{op} = N_R$, is not optimal. This is because of the channel time wasted during the first backoff stage. Hence, it is better to allocate at least two stations for each RAW slot, which means K_{op} is always less or equal to $\left\lfloor \frac{N_R}{2} \right\rfloor$. On the other hand, a minimal value of K_{op} may lead to a higher number of collisions in each group, and hence, more waste of energy. Hereafter, the optimal grouping K_{op} is very dependent on the total number of stations N_R allocated within the RAW.

Figure 4.17 shows the performance comparison between our proposed optimal configurations and and the DCF mechanism in terms of RAW duration. We consider three different numbers of contending stations participating in the RAW: 100, 200 and 400. As seen from Figure 4.17a, the RAW duration does not have a significant impact on the performance of the DCF mechanism. However, our proposed algorithm reports much better results than the legacy DCF. This result is explained by the fact that our proposed algorithm profits from a higher value of RAW length by dividing stations into more RAW slots and by allocating more access time interval to each RAW slot. With such grouping, each RAW slot contains fewer stations that have more free access period, and hence, yields more successful transmissions, which eventually increases the overall energy efficiency within the RAW.

We present in Figure 4.17b the optimal grouping K_{op} associated to the energy efficiency depicted in Figure 4.17a. We see that for a short RAW, the optimal grouping K_{op} is lower. This is because a higher value will lead to RAW slots with duration shorter than the length of one packet transmission, and hence, no transmissions. Although a higher duration of T_S allows creating more groups, K_{op} is eventually limited to the value of half number of stations. That is for the same reason mentioned previously, which is the channel time wasted during the first backoff stage when there is only one station per group.



(a) Energy efficiency.

(b) Optimal grouping.

Figure 4.17: Energy efficiency comparison of the proposed optimal RAW configuration with conventional DCF in terms of RAW duration.

The results obtained from the comparisons in Figure 4.16 and Figure 4.17 validate the performance of our proposed algorithm over the legacy DCF in increasing the network energy efficiency. Our proposal derives for every RAW the best configuration that maximizes the overall energy efficiency. Consequently, the combination of our proposed grouping scheme that sets the RAW parameters and our proposed algorithm to optimally configure those parameters proves effective on maximizing the overall RAW energy efficiency in highly dense scenarios.

4.8 Non-ideal Channel Scenario

In this section, we analyze and evaluate the impact of non-ideal channel conditions. Since we assumed ideal channel conditions throughout our study, we now consider the impact of Bit Error Rate (BER) on the performance of the time-limited contention within the RAW slot [85]. We consider the same operating conditions and analysis as in Section 4.3. With an additional presence of BER, the difference from the previous analytical model lies in the collision probability defined in (4.2) and the successful transmission probability defined in (4.4).

In this context, the Packet Error Rate (PER) is defined by

$$PER = 1 - (1 - BER)^{\ell} \tag{4.19}$$

where ℓ is the packet length in bits.

The conditional collision probability in (4.2) is now the probability that a packet is col-

lided or received in error and is defined by

$$p = 1 - (1-\tau)^{g_i - 1} \cdot (1 - PER) \tag{4.20}$$

The probability for a packet to be successfully transmitted in (4.4) is now conditioned to the fact that the packet is not corrupted and it is given by

$$P_s^{g_i} = \frac{g_i \tau (1-\tau)^{g_i - 1}}{P_{tr}^{g_i}} \cdot (1 - PER) \tag{4.21}$$

To analyze the impact of BER on the RAW slot performance, we consider a RAW slot of duration $T_S = 20\,\text{ms}$ with 4 stations. We depict in Figure 4.18 the throughput in terms of BER for three different payloads: 256, 512, and 1024 bits. We observe a good accuracy of the analytical results compared to the simulations. The throughput for each payload decreases with a different speed when the BER increases. A very low BER has almost no impact on the payload size as smaller payload yields lower throughput due to the high contention.

However, the throughput begins to rapidly decrease as BER passes a certain value for each different payload: $BER > 3 \cdot 10^{-5}$ for $L = 256\,\text{bits}$, $BER > 2 \cdot 10^{-5}$ for $L = 512\,\text{bits}$, and $BER > 10^{-5}$ for $L = 1024\,\text{bits}$. This can be explained by the that becomes higher for a longer payload, as can be seen in (4.19). For the same reason, the throughput for a longer payload decreases faster when BER increases. That eventually causes the shortest payload to have the higher throughput at some values of BER, as observed in Figure 4.18: the payload $L = 256\,\text{bits}$ yields higher throughput than $L = 512\,\text{bits}$ and $L = 1024\,\text{bits}$ when $BER > 7 \cdot 10^{-4}$, and the payload $L = 512\,\text{bits}$ yields higher throughput than $L = 1024\,\text{bits}$ when $BER > 3.4 \cdot 10^{-4}$. Eventually, the throughput becomes null for every payload size when $BER > 10^{-2}$ due to the effect of high transmission errors.

We further analyze the impact of BER on the optimal grouping and the associated maximum throughput within the RAW. Figure 4.19 presents the results of our proposed optimal grouping algorithm (Algorithm 4.2) for different stations within a RAW of duration $T_R = 100\,\text{ms}$. We consider the results for two cases of an ideal channel and a non-ideal channel with $BER = 0.5 \cdot 10^{-3}$. Table 4.4 summarizes the data for important variations. We observe that the BER has a direct impact on decreasing the maximum achievable throughput.

In the scenario of a non-ideal channel, the optimal grouping K_{op} that provides the maximum throughput is either the same or higher than the case of ideal channel conditions. That is due to the fact that the corrupted transmissions are seen as collisions, and hence more groups might alleviate the contention and provide a higher overall RAW throughput. Whereas, a lower number of groups when considering non-ideal channel conditions cannot be optimal as it leads to longer RAW slots, and hence more wasted time due to transmission errors.

In both cases, the maximum throughput fluctuates when N_R increases. That can be ex-

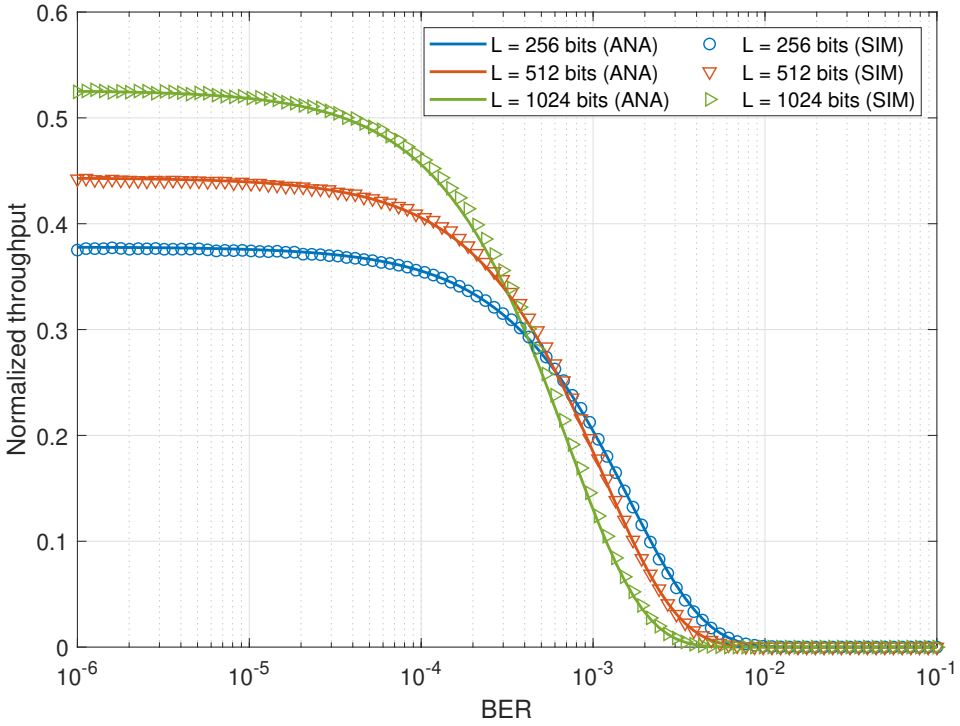
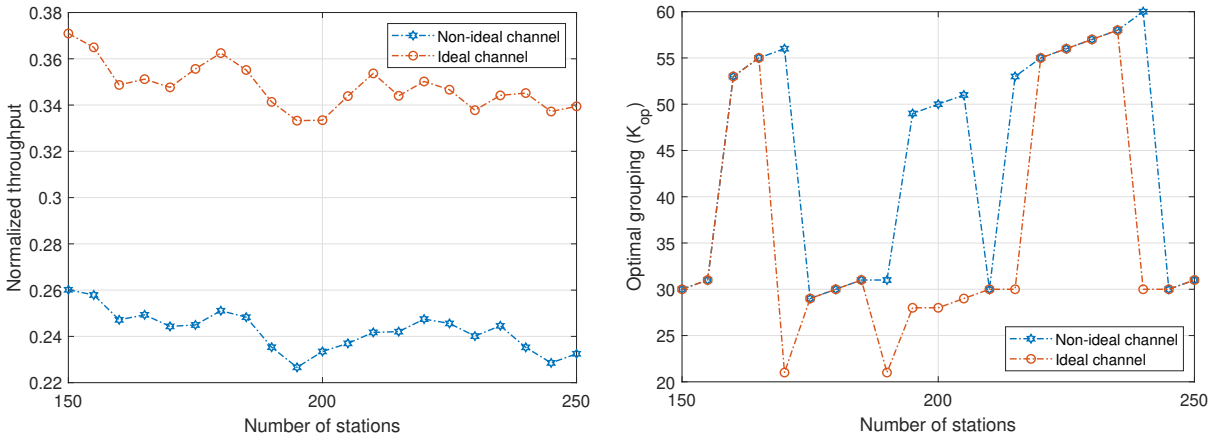


Figure 4.18: Normalized throughput in terms of the BER within a RAW slot of duration $T_S =$ 20 ms with 4 stations for different payloads.



(a) Maximum achievable throughput.

(b) Optimal grouping.

Figure 4.19: Impact of the non-ideal channel with $BER = 0.5 \cdot 10^{-3}$ on the maximum achievable throughput and the optimal grouping within a RAW of duration $T_R = 100$ ms.

plained by the fact that more stations within the RAW may increase or decrease the overall RAW throughput. The throughput degradation is due to the higher contention within each group, which eventually decreases the throughput within each RAW slot. However, as the duration of the RAW slot is allocated proportionally to the number of participating stations, more stations within the RAW can also lead to shorter RAW slots (see (4.9) and (4.10)), which eliminates the unnecessary additional time and eventually increases the throughput. Consequently, our proposed algorithm takes into account all different factors: the number of stations, the RAW duration, and the BER, in order to provide an optimal configuration that maximizes the overall

Table 4.4: Impact of the BER on the optimal grouping and the maximum achievable throughput within a RAW of duration $T_R = 100\,\text{ms}$.

Number of stations	Ideal channel		Channel with $BER = 0.5 \cdot 10^{-3}$	
	Optimal grouping	Maximum throughput	Optimal grouping	Maximum throughput
160	53	0.3487	53	0.2472
165	55	0.3512	55	0.2493
170	21	0.3476	56	0.2443
175	29	0.3556	29	0.2449
180	30	0.3625	30	0.2511
185	31	0.3552	31	0.2483
190	21	0.3414	31	0.2353
195	28	0.3332	49	0.2266
200	28	0.3335	50	0.2335
205	29	0.3439	51	0.2370
210	30	0.3537	30	0.2417
215	30	0.3440	53	0.2421
220	55	0.3502	55	0.2475

RAW throughput.

4.9 Conclusion

In this chapter, we have established an analytical framework for designing and optimizing a grouping scheme for the RAW mechanism of the IEEE 802.11ah standard. We proposed a grouping scheme that splits the RAW into two Sub-RAWs, which allows us to distribute the stations into groups of similar size within each Sub-RAW. Based on the framework developed in the first part, we evaluate the RAW performance and then propose a configuration algorithm to maximize the network throughput with the provided number of nodes and RAW length. Finally, we analyzed our proposals through numerical results, which are also accurately validated via discrete-event simulations that we developed on MATLAB software. We have established the importance of station grouping in increasing the network throughput and energy efficiency. We also carried out an analysis and examination of the impact of different variables on the overall RAW performance. Our proposed RAW configuration algorithm is shown to achieve higher throughput and energy efficiency compared to the conventional mechanism DCF. We further examined the impact of a non-ideal channel with transmission errors on both time-limited contention within the RAW slot and the overall RAW optimal grouping and configuration. Our

proposal can be applied to several new standards that address dense Wi-Fi networks using periodic channel reservations. A more realistic scenario of a dense IoT network consists of devices distributed in a heavily built urban environment. We address such a scenario in the next chapter, where transmitted signals are subject to the effect of fading, and the feature of capture effect is essential to improve the success rate of transmitted data.

CHAPTER 5

Accurate Analytical Model and Evaluation of Wi-Fi HaLow Based IoT Networks under a Rayleigh-Fading Channel with Capture

Contents

5.1 Introduction

The IEEE 802.11ah standard, marketed as Wi-Fi HaLow, comes with significant improvements over the legacy Wi-Fi technology, making it an excellent Wi-Fi solution for IoT applications. It is shown that the IEEE 802.11ah standard provides better performance than the other alternatives in terms of data rate, coverage range, association time, and delay [22]. The amendments in the IEEE 802.11ah standard make this new technology prominent in developing reliable IoT network infrastructures. Numerous studies have previously been conducted on the standard's key features to show the potentials and challenges in designing the new proposed mechanisms [25].

The time-limited contention within short periods of RAW slots presents a new challenge to derive an accurate evaluation of the overall performance of a single RAW slot or a RAW consisting of several RAW slots. The legacy analytical models based on steady-state results as Bianchi's [5] are not applicable in this case. The channel access contention within a RAW slot is terminated at the end of its period, which prevents the system from reaching the stationary state [63]. Henceforth, the RAW mechanism is critically affected by the time constraint presented by the lengths of RAW slots. Since the standard did not mention how to configure RAWs within beacon intervals and RAW slots within RAW, an accurate evaluation of the performance of a given configuration is necessary in order to derive efficient scheduling and configuration algorithms. However, we do not focus on RAW configuration strategies in this work. Instead, we shall develop an accurate analytical framework to evaluate the contention within a RAW slot period and a RAW consisting of several RAW slots. We assume that the RAW periods within beacon intervals are allocated in a way that eliminates the hidden nodes problem so that all stations assigned to a given RAW can listen to each other.

In real-life applications, the capture effect highly affects the network performance, and the signal of a transmitted packet is subject to fading effects over the channel, background noise, and potential interference from concurrent transmissions. An IEEE 802.11ah based IoT network in an urban environment will be affected by many factors. Considering that the coverage is up to 1 km, and Non-line-of-sight (NLOS) conditions with respect to the AP are not met for all stations. Besides, a transmitted signal in such an environment will be subject to reflections, resulting in multipath fading at the receiver. Such a scenario is mainly represented by a Rayleigh fading channel [86]. Several works carried out an analytical framework to evaluate the performance of a network operating under a Rayleigh fading channel. In [87, 88], the authors presented a Markov chain model to evaluate the throughput of a DCF-based network. They considered a non-ideal channel with capture effects in a Rayleigh fading environment. Their study was carried out within saturated [88], and unsaturated [87] scenarios. Similar work was reported in [89], where the authors proposed a 3-D Markov Chain based model to evaluate a homogenous 802.11 network operating with the DCF mechanism under a Rayleigh fading channel that incorporates transmission error and capture.

Therefore, the capture effect feature imports significant enhancements to the network performance by enhancing channel usage during RAW periods. That is, when a collision occurs, the receiver can still capture the data frame from one station when its received signal exceeds that of other interfering signals by a certain threshold. Furthermore, the received signal depends on the channel model, which characterizes the wireless medium and depends on several parameters, such as the distance between the transmitter and receiver, the line of sight, and the environment (obstacles and surroundings) [90].

The analytical frameworks presented in these works have some shortcomings in the capture effect modeling. First, they do not consider the geometric distribution of stations around the AP. Second, they assume that all stations are power-controlled so that the mean received power of interfering packets is the same. Third, they compute the capture probability at the level of slot states so that when a slot contains a collision, one packet from one of all stations can still be captured. However, that may not be accurate, as only a few stations will participate in a collision at a given slot. In this work, we take these shortcomings into account besides the time-limited contention presented in the RAW mechanism. We also consider the geometric distribution of stations around the AP and the same transmission power for all stations so that the mean received powers of interfering packets are different and depend on the distance from the AP to the source stations. Furthermore, we compute the capture probability at the level of a transmitted packet as a subevent of the packet collision. Hence, such a result is extended to derive a slot that contains collision with a capture.

During the contention in a wireless network, the timeline of a station or the wireless medium is discretely occupied by a limited number of events (e.g., successful transmissions, collisions, idle slots) occurring repeatedly. Each event occurs periodically, creating renewal cycles. Renewal theory is an adequate mathematical approach that studies similar behavior of periodic events. This approach is dedicated to stochastic processes that occur at random instants at which the system returns to a state probabilistically equivalent to the starting state [9]. The authors in [77] introduced a three-level renewal process framework, modeling the periodic channel access in the IEEE 802.15.4 standard. However, their model does not apply to the DCF MAC protocol as medium access in IEEE 802.15.4 is based on a non-persistent CSMA mechanism. In [75], Zhang presented a renewal process to model the backoff procedure of a station within a non-saturated IEEE 802.11 network. The introduced method is more precise and practical than Bianchis's approach, particularly for non-saturated scenarios. In [76], the authors established a renewal theory based model to evaluate the IEEE 802.11 multichannel MAC protocol y deploying a two-level renewal process. Nevertheless, the time-limited contention factor introduced in the RAW mechanism is not considered in [75, 76]. The models introduced in these works are based on renewal cycles that only provide accurate results in a long enough contention time.

A renewal process can also be associated with a counting process that tracks events in time. The authors in [91] introduced the simple counting processes and their extension to gen-

eralized counting processes under the presence of environmental stochasticity, where the generalization corresponds to the case in which the process transitions depend on the environmental conditions. Henceforth, in this work, we develop a counting process that tracks transmission in the medium up to a given time. Although the different events occupying the channel yield a discrete timeline, the counting process is defined in a continuous timeline, which presents a challenge to adapt such a process to a wireless medium. This process can be seen as a generalized counting process defined in [91], where the process transitions depend on the outcome of wireless medium slots. The developed counting process provides accurate modeling for the limited channel access introduced in the RAW mechanism. To the best of our knowledge, the counting process is not deployed before in the literature to model time-restricted medium access contention.

The contributions of this work are summarized as follows:

- We model the channel access within a RAW slot under a Rayleigh fading channel with capture, considering the geometric distribution of stations around the AP.
- We develop a renewal theory based framework to model the contention within the RAW slot by deploying a counting process to track transmissions within the RAW slot period. We then evaluate the RAW slot throughput representing the channel usage ratio in delivering data frames during the RAW slot period.
- We derive the throughput of a RAW consisting of several RAW slots, representing the channel usage ratio to deliver data frames during the overall RAW duration. We also evaluate the RAW performance in a no-capture channel scenario and derive the capture ratio in a given RAW configuration.
- We present a meticulous validation of the analytical findings through extensive simulations obtained via a discrete-event simulator developed with MATLAB software. We henceforth discuss and analyze several parameters affecting the performance of a RAW slot and a RAW consisting of several RAW slots.

The remainder of this chapter is organized as follows: Section 5.2 describes the considered scenario and the channel model. We further model the channel access mechanism in the presence of capture effect and power attenuation under a Rayleigh fading channel. Then, Section 5.3 presents a renewal theory based model to evaluate the RAW slot throughput, representing the channel usage ratio to deliver data frames during the RAW slot period. We henceforth derive the throughput of a RAW comprising several RAW slots and highlight the capture ratio in the RAW period. In Section 5.4, we validate our analytical findings via simulations and examine the impact of several parameters on the performance of a RAW slot and a RAW consisting of several RAW slots. Finally, we conclude our work in Section 5.5.

5.2 System Model

This section expounds on the network system and the channel access model. First, Section 5.2.1 introduces the network scenario considered. Then, Section 5.2.2 describes the Rayleigh fading channel model. Thereafter, Section 5.2.3 presents the analytical model for channel access under fading and capture awareness. Finally, Section 5.2.4 yields the different states of a channel slot.

5.2.1 Scenario

We consider an IEEE 802.11ah-based network consisting of stations randomly distributed over a given area around one AP within a radius ρ, as illustrated in Figure 5.1. We assume that one RAW occupies the entire beacon interval and that all stations are allocated to this RAW. Note that this RAW may contain one or more RAW slots. Our goal is to model the contention within a RAW slot and then evaluate and analyze the performance of a single RAW slot and a RAW consisting of several RAW slots. We consider a channel with capture effect enabled, taking into account the distances of stations from the AP and the fading of transmissions throughout the channel. The AP assigns a RAW to a set of stations with packets to deliver. Thus, we aim to examine the gain of the RAW in the worst scenario, and therefore we consider a network with saturated traffic.

During the RAW slot period, only the assigned stations are allowed to compete for channel access using the EDCA protocol [4]. The station first checks if the channel is unoccupied for a DIFS time. Then, it selects a random backoff time counter uniformly from $[0, W_j - 1]$, $j \in [0, m]$, where W_j is the size of the current CW, and m denotes the number of backoff stages. The CW is initiated with the value W_0, and it doubles after each collision until it achieves the maximum size $W_m = 2^m \cdot W_0$. When the transmitted packet is delivered successfully or dropped, the CW resets to W_0. Note that the transmission attempt of a given packet can only be initiated at the beginning of the time slot, and the packet is dropped after m failed retransmissions. The station gets a TXOP to transmit packets when its backoff time counter reaches zero. For the convenience of analysis and to keep fair resources allocation for stations, we assume that a TXOP allows only the transmission of one packet, and its duration is given by:

$$T_{\text{TXOP}} = T_{\text{DATA}} + \text{SIFS} + T_{\text{ACK}}\,, \tag{5.1}$$

where T_{ACK} is the transmission time of an ACK frame. SIFS is the time interval between the transmission of a data frame and its corresponding ACK, defined in the standard. T_{DATA} is the transmission time of a data frame defined as follows:

$$T_{\text{DATA}} = T_{\text{PLCP}} + \frac{\text{Payload} + \text{MacHeader}}{\text{DataRate}}\,, \tag{5.2}$$

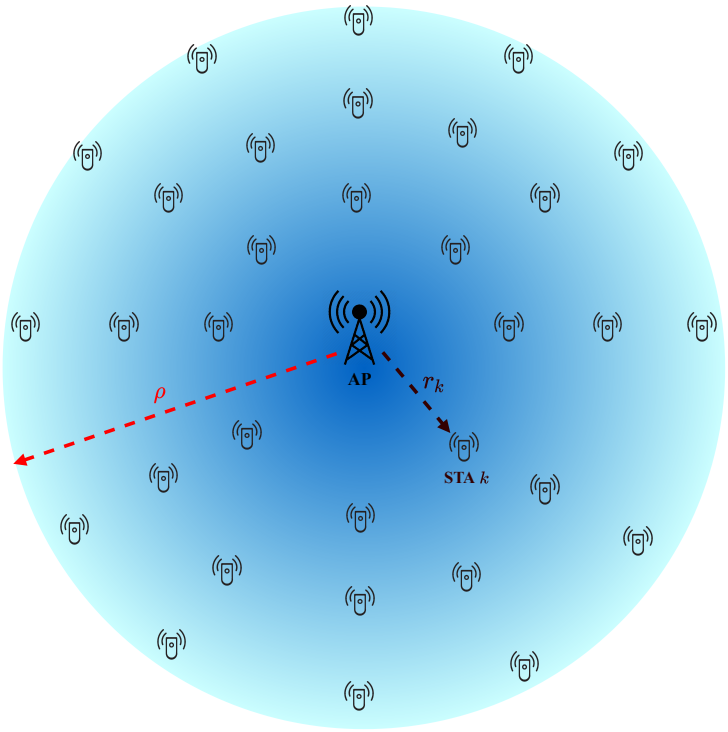


Figure 5.1: Illustration of the network scenario.

where T_{PLCP} is the duration of the PLCP header.

For the convenience of analysis, we assume that all stations assigned to the RAW belong to the same Access Category (AC) and carry similar fixed-size payloads.

5.2.2 Channel Model

The capture effect imports significant improvements to the network performance by decoding one packet from a collision of several stations. Thus, when two stations or more transmit simultaneously, and a collision occurs, it is still possible for the AP to capture the data frame from one station if the power of its received signal exceeds that of interfering stations by a certain threshold. However, the received power of a transmitted packet depends on the channel model representing the behavior and changes of transmitted signals throughout the channel. The channel model characterizes the wireless medium and depends on several parameters, such as the distance between the transmitter and receiver, the line of sight, and the environment (obstacles and surroundings) [90]. Therefore, the signal of a transmitted packet is subject to fading effects over the channel, background noise, and potential interference from concurrent transmissions. We omit in this work the impact of background noise and focus only on the impact of interference and fading effects. The IEEE 802.11ah standard mainly addresses long-distance IoT use cases, where there is mostly a lack of Line of sight (LOS) between stations and the AP. Henceforth, we consider a Rayleigh fading channel model representing a heavily built-up urban environment where transmissions are subject to multipath fading and NLOS propagation in the wireless channel [92].

During simultaneous transmission of $n+1$ stations, the packet of a tagged transmitting station k is assumed to be captured by the AP if and only if its instantaneous power ω_k exceeds the instantaneous joint interference power $\omega_{int} = \sum_{i=1}^{n} \omega_k$ by at least a threshold z. This threshold reflects the capture performance of the AP, and it is referred to as the capture threshold. Furthermore, with the consideration of a Rayleigh fading channel model, the instantaneous power of a received packet from station k is exponentially distributed as follows [6]:

$$f_{\omega_k}(x) = \frac{1}{\omega_{0k}} e^{-\frac{x}{\omega_{0k}}}, \quad x \geq 0 \tag{5.3}$$

where ω_{0k} is the local mean power of the transmitted packet at the receiver, and it is defined as follows:

$$\omega_{0k} = A \cdot r_k^{-\alpha} \cdot \omega_T \,. \tag{5.4}$$

where

- ω_T is the transmitted signal power.
- $A \cdot r_k^{-\alpha}$ represents the deterministic path-loss law, where r_k is the distance between the station k and the AP.
- α is the path-loss exponent, and it depends on the propagation environment. For the convenience of analysis, we assume $\alpha = 4$ for the rest of this chapter.
- A is a dimensionless constant in the path-loss law.

We assume that the constants A and ω_T are the same for all transmitted packets.

5.2.3 Capture Aware Channel Access

Although there are works in the literature that model the channel access under the capture effect feature [88, 87, 89], they derive the packet capture probability as a subevent of a slot containing a collision. However, such a result is not exactly accurate since capturing a given packet is a subevent of the collision of the same packet. Henceforth, in our model, we track the packet transmitted from a tagged station and derive its capture probability conditioned to its collision. We then extend this result to obtain the probability of having a captured packet in a randomly chosen slot.

Consider a RAW slot allocated to N_S stations. In order to consider the capture effect in the channel access analysis, we assume $N_S \geq 2$. The case of $N_S = 1$ presents no collisions and hence no capture. Thus, it belongs to the scenario of a no-capture channel, which we discuss later in Section 5.3.2.

We assume that all stations inside the RAW slot attempt to transmit packets with the same probability τ. Let p denote the probability that the transmission attempt of a packet fails to be successfully delivered. Denote by A and B two random variables representing the attempts and backoffs experienced by a single transmitted packet, respectively. Using the mean value analysis method as in [40], we define the probability τ as follows:

$$\tau = \frac{\mathrm{E}[A]}{\mathrm{E}[A] + \mathrm{E}[B]} \tag{5.5}$$

where E[.] is the mathematical expectation function.

Given that each packet can be retransmitted up to m times, both A and B follow a truncated geometric distribution with the rate p and different supports Ω_A and Ω_B, respectively. Thus, we have

$$\Omega_A = \{1, 2, \ldots, m+1\},$$

and

$$\Omega_B = \left\{ \frac{W_0}{2}, \sum_{k=0}^{1} \frac{2^k W_0}{2}, \ldots, \sum_{k=0}^{m} \frac{2^m W_0}{2} \right\}.$$

Hence,

$$\mathrm{E}[A] = \sum_{k=0}^{m} (k+1) \cdot \frac{(1-p)p^k}{1-p^{m+1}} = \frac{1 + (m+1)p^{m+2} - (m+2)p^{m+1}}{(1-p)(1-p^{m+1})}, \tag{5.6}$$

and

$$\mathrm{E}[B] = \sum_{k=0}^{m} \frac{2^k W_0}{2} \cdot \frac{(1-p)p^k}{1-p^{m+1}} = \frac{W_0 \left(1 + \left(1 - 2^{m+2}(1-p) - 2p\right) p^{m+1}\right)}{2(1-2p)(1-p^{m+1})}. \tag{5.7}$$

Thus, we substitute (5.6) and (5.7) in (5.5) to obtain the probability τ for a station to initiate the transmission of a packet in a given slot expressed as follows:

$$\tau = \frac{2(2p-1)\left(1 + (m+1)p^{m+2} - (m+2)p^{m+1}\right)}{W_0(1-p)\left((2^{m+2}-1)p^{m+1} - (2^{m+2}-2)p^{m+2} - 1\right) + 2(2p-1)\left(1 + (m+1)p^{m+2} - (m+2)p^{m+1}\right)}. \tag{5.8}$$

A transmitted packet from a tagged station will encounter a collision if at least one of the remaining $N_S - 1$ stations starts transmitting in the same time slot. Thus, the packet collision probability is defined as follows:

$$p_{col}^{p} = 1 - (1-\tau)^{N_S - 1}. \tag{5.9}$$

Since we consider a channel with the capture effect feature enabled, a packet involved in a collision may be captured and delivered successfully. Let p_{cap}^{p} be the probability of capturing a

transmitted packet after encountering a collision. The capture of a transmitted packet is a subset event of the collision events; a tagged packet fails to be delivered if it experiences a collision and is not captured. Hence, the probability p for failed delivery of a transmitted packet is given by:

$$p = p_{col}^{p} \cdot (1 - p_{cap}^{p}) . \tag{5.10}$$

A prior definition for the probability p_{cap}^{p} to capture a transmitted packet after encountering a collision is given by:

$$p_{cap}^{p} = \frac{\sum_{n=1}^{N_S-1} R_n \cdot p_{cap}(z, n+1)}{p_{col}^{p}} , \tag{5.11}$$

where $p_{cap}(z, n+1)$ is the Average Conditional Capture Probability (ACCP) for $n+1$ interfering packets. This probability is defined by the probability of signal-to-interference ratio $\gamma = \frac{\omega_0}{\omega_{int}}$ exceeding the threshold z. We have

$$p_{cap}(z, n+1) = \Pr\{\gamma > z | n+1\} . \tag{5.12}$$

R_n is the probability that the transmitted packet encounters n interfering packets, and it is given by:

$$R_n = \binom{N_S - 1}{n} \tau^n (1-\tau)^{N_S-1-n} . \tag{5.13}$$

We now proceed to determine the expression of the ACCP defined in (5.12). This probability depends on the received powers of the captured signal and the interfering packets. Furthermore, the received power of a given packet depends on the distance of the source station from the AP.

All N_S stations are randomly distributed around the AP in a circular area of radius ρ, as illustrated in Figure 5.1. Hence, the Probability Density Function (PDF) of the distance between a station and the AP is given by:

$$h(r) = \frac{2r}{\rho^2}, \qquad 0 \le r \le \rho . \tag{5.14}$$

Assume that the AP receives $n+1$ packets simultaneously. Among these packets, we tag one packet transmitted by a station located at a distance r_0 from the AP.

Let $\omega_{00}, \omega_{01}, \ldots, \omega_{0n}$ be the local mean powers of the tagged packet and the n interfering packets, respectively. The conditional capture probability for the tagged packet can be expressed

as follows:

$$
\begin{aligned}
p_{cap}(z,r_0,r_1,\ldots,r_n) &= \Pr\{\gamma > z | r_0,r_1,\ldots,r_n\} \\
&= \Pr\left\{ \omega_0 > z\cdot\sum_{i=1}^{n}\omega_k \Big| r_0,r_1,\ldots,r_n \right\} \\
&= \int_0^\infty f_{\omega_1}(x_1)\left(\int_0^\infty f_{\omega_2}(x_2)\cdots\int_0^\infty\left(f_{\omega_n}(x_n)\cdot\int_{z\cdot\sum_{i=1}^{n}\omega_k}^\infty f_{\omega_0}(x_0)\,dx_0\right)dx_n\cdots dx_2\right)dx_1 \\
&= \int_0^\infty f_{\omega_1}(x_1)\left(\int_0^\infty f_{\omega_2}(x_2)\cdots\int_0^\infty\left(f_{\omega_n}(x_n)\cdot e^{-\frac{z\cdot\sum_{i=1}^{n}\omega_k}{\omega_{00}}}\right)dx_n\cdots dx_2\right)dx_1 \\
&= \frac{1}{1+z\cdot\frac{\omega_{0n}}{\omega_{00}}}\cdot\int_0^\infty f_{\omega_1}(x_1)\left(\int_0^\infty f_{\omega_2}(x_2)\cdots\int_0^\infty\left(f_{\omega_{n-1}}(x_{n-1})\cdot e^{-\frac{z\cdot\sum_{i=1}^{n-1}\omega_k}{\omega_{00}}}\right)dx_{n-1}\cdots dx_2\right)dx_1 \\
&\vdots \\
&= \prod_{k=1}^{n}\frac{1}{1+z\cdot\frac{\omega_{0k}}{\omega_{00}}} \\
&= \prod_{k=1}^{n}\frac{1}{1+z\cdot\left(\frac{r_k}{r_0}\right)^{-4}}.
\end{aligned}
\tag{5.15}
$$

Since all stations are distributed following the PDF $h(.)$ defined (5.14), all factors in the product in (5.15) are statistically equal. Consequently, the average conditional capture probability for the tagged packet when encountering n interfering packets can be defined as follows:

$$
\begin{aligned}
p_{cap}(z,r_0,n) &= \left(\int_0^\rho \frac{h(r)}{1+z\cdot\left(\frac{r}{r_0}\right)^{-4}}\,dr\right)^n \\
&= \left(1-\frac{r_0^2\sqrt{z}}{\rho^2}\cdot\arctan\left(\frac{\rho^2}{r_0^2\sqrt{z}}\right)\right)^n
\end{aligned}
\tag{5.16}
$$

Given that the distance r_0 is also randomly distributed following the PDF $h(.)$, the ACCP to capture one packet from $n+1$ interfering packets can be derived by integrating $p_{cap}(z,r_0,n)$ over all possible values of the distance r_0 from the AP. Henceforth, the ACCP $p_{cap}(z,n+1)$ is expressed as follows:

$$
\begin{aligned}
p_{cap}(z,n+1) &= \int_0^\rho p_{cap}(z,r_0,n)\cdot h(r_0)\,dr_0 \\
&= \int_0^\rho \left(1-\frac{r_0^2\sqrt{z}}{\rho^2}\cdot\arctan\left(\frac{\rho^2}{r_0^2\sqrt{z}}\right)\right)^n\cdot\frac{2r_0}{\rho^2}\,dr_0.
\end{aligned}
\tag{5.17}
$$

Consequently, by substituting (5.13) and (5.17) in (5.11), we derive the expression of the

probability p_{cap}^{p} as follows:

$$p_{cap}^{p} = \frac{1}{p_{col}^{p}} \sum_{n=1}^{N_S-1} \binom{N_S-1}{n} \tau^n (1-\tau)^{N_S-1-n} \cdot \int_0^{\rho} \left(1 - \frac{r_0^2\sqrt{z}}{\rho^2} \cdot \arctan\left(\frac{\rho^2}{r_0^2\sqrt{z}}\right)\right)^n \cdot \frac{2r_0}{\rho^2} dr_0 . \tag{5.18}$$

Note that the integral in this expression can only be calculated by numerical integration.

Moreover, the values of probabilities τ, p, and p_{cap}^{p} can be obtained by resolving numerically the non-linear system defined by (5.8), (5.10) and (5.18).

5.2.4 Channel Slot State

During the RAW slot period, all assigned stations contend for channel access, and each station starts transmitting with probability τ. Hence, the probability P_i that a randomly chosen slot is idle is given by:

$$P_i = (1-\tau)^{N_S} . \tag{5.19}$$

The probability P_s of a packet being transmitted and successfully delivered without interference is the probability of having exactly one station that transmits, conditioned on the current slot to be not empty. We have

$$P_s = \frac{N_S \cdot \tau \cdot (1-\tau)^{N_S-1}}{1-P_i} . \tag{5.20}$$

Let P_{cap} be the probability of delivering a packet successfully after being captured in a collision. Since only one packet can be captured from each collision, P_{cap} is given by the probability that one of the N_S stations transmits and its packet is captured, conditioned to the fact that the current slot contains a collision. Hence, we have

$$P_{cap} = \frac{N_S \cdot \tau \cdot p_{cap}^{p}}{(1-P_i)(1-P_s)} . \tag{5.21}$$

Henceforth, during the RAW slot period, the channel comprises several slots that we classify into four types. As illustrated in Figure 5.2, a randomly chosen slot can be in only one of the following four states:

- Idle: When there are no ongoing transmissions in the slot. That is, all stations are listening to the medium while counting down their backoff counter. A randomly chosen slot is idle with probability P_i.
- Success: When only one station transmits and its packet is delivered successfully. Given that this state is a subset of busy slots, a randomly chosen slot contains a successful single transmission with probability $(1-P_i)P_s$.

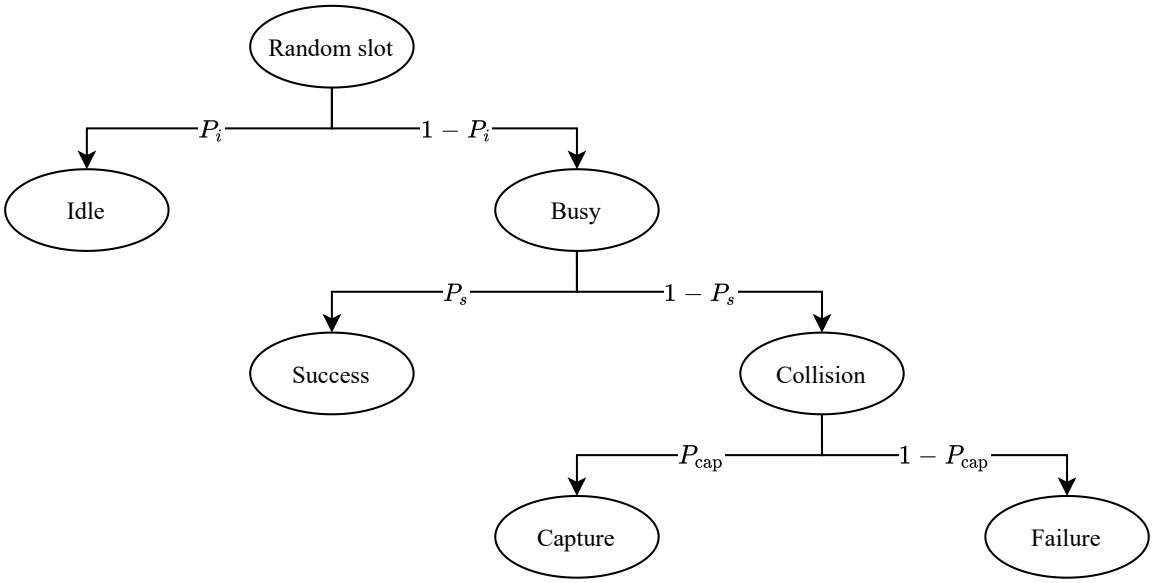


Figure 5.2: Different states of a randomly chosen slot.

- Capture: When two or more stations transmit and the packet of one station is captured successfully from the collision. This type of slot is a subset of slots containing collisions, which are a subset of busy slots. Thus, a randomly chosen slot contains a collision with a capture with probability $(1-P_i)(1-P_s)P_{cap}$.
- Failure: When two or more stations transmit and no packet is captured. Given that this state is a subset of slots containing collisions, the probability that a randomly chosen slot contains a collision and no packet is captured is given by $(1-P_i)(1-P_s)(1-P_{cap})$.

5.3 Analytical Model and Performance Evaluation

Since each station is allowed to contend for channel access only during its designated RAW slot, we start by modeling and evaluating the contention within one RAW slot. We consider a tagged RAW slot of duration T_S allocated for N_S stations. Note that only the N_S stations can compete for channel access during the period T_S.

The handover between two consecutive RAW slots is another factor to be considered. There are two options to deal with the transmission of a packet at the end of the RAW period according to the standard [4]. In the first case, the transmission of the last packet in a given RAW is not allowed to overrun the slot boundary. In the second case, the transmission of the last packet in a given RAW slot may cross the slot boundary. In this latter case, the stations in a RAW slot will not start contending for channel access until the ongoing transmissions from the previous RAW slot are completed.

Enabling RAW slot boundary-crossing will result in more energy loss for stations in the following RAW slot. That is because the stations will awake at the beginning of their allocated

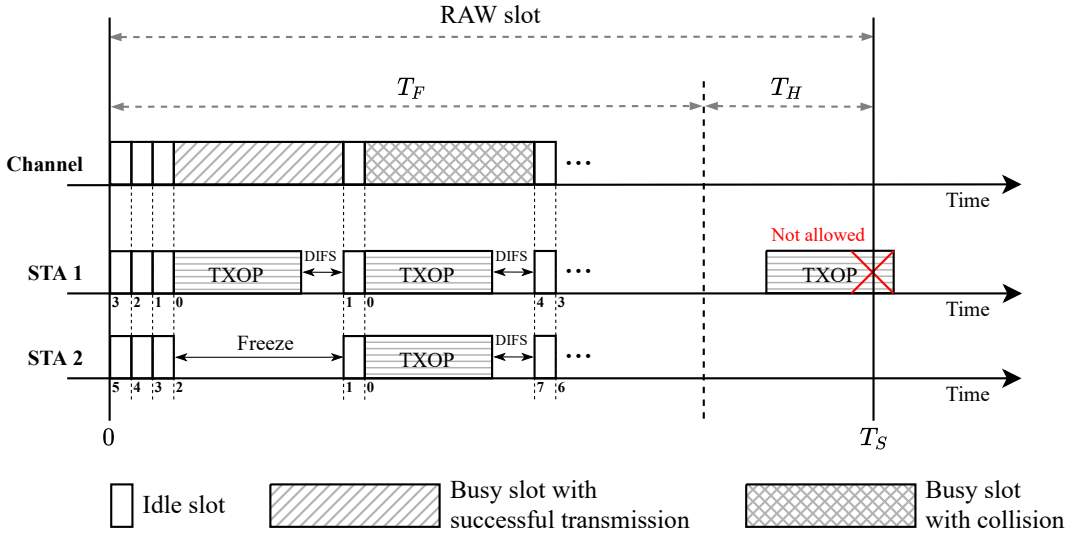


Figure 5.3: Structure of RAW slot.

RAW slot and will not be able to contend for channel access until the last transmission from the preceding RAW slot is complete. Additionally, boundary-crossing in the final RAW slot will lengthen the overall time of the RAW, resulting in energy waste for stations competing outside the RAW. Thus, for these reasons and to guard fair resource allocation for all assigned stations in the RAW, we consider the RAW slot boundary-crossing to be disabled. Hence, stations are not allowed to cross the boundary of their allocated RAW slot with transmissions of their packets. To ensure this condition, a holding time T_H is defined at the end of the RAW slot, where stations are not allowed to initiate transmissions.

Hereafter, as depicted in Figure 5.3, the RAW slot period T_S is decomposed into a free access period T_F during which stations contend for channel access, and a holding period T_H during which stations are not to initiate a packet transmission. The holding period can only host ongoing transmissions initiated before the free access period boundary.

The remaining of this section is structured as follows. Section 5.3.1 introduces the stochastic model for the contention within the RAW slot period. Then, Sections 5.3.2 and 5.3.3 derive the performance metrics in a single RAW slot and several RAW slots, respectively.

5.3.1 Stochastic Model

We aim to evaluate the RAW slot in terms of normalized throughput, which represents the portion of the RAW slot period used to successfully deliver data frames to the AP, whether through single transmissions or captured packets. We consider a tagged RAW slot with duration T_S allocated for N_S stations. To derive the RAW slot throughput, we first need to determine the slots that contain successful delivery of data frames. That includes both successful single transmissions and successful delivery of captured packets.

In Chapter 3, we developed a renewal theory based model to derive the slots used for transmissions during the RAW slot period. We extend that model in this work with the additional assumptions of channel fading and capture effect. Additionally, we enhance the accuracy of this new model by distinguishing the time exploited by transmissions during the holding period. Once the RAW slot interval begins, the assigned stations start to contend for channel access up to the instant T_F, which defines the free access period within the RAW slot. At the time T_F, all stations suspend their backoff timers, and only ongoing transmissions are allowed during the holding period T_H. Hereafter, we observe the channel timeline during the RAW slot period in a binary state, where a randomly chosen slot can be idle or busy.

The stochastic model we propose is illustrated in Figure 5.4. We aim to evaluate the throughput within the RAW slot, which represents the channel usage ratio during its duration. To do so, we propose to deploy a renewal theory approach to model the contention of stations during the RAW slot period by constructing a counting process to keep track of occurred transmissions up to the end of the RAW slot interval T_S.

For every $i \in \mathbb{N}$, denote the type of the $(i+1)$th slot by the random variable Z_i, where the events $\{Z_i = 1\}$ and $\{Z_i = 0\}$ represents a busy and idle slot, respectively. A randomly chosen slot is idle with probability P_i and busy with probability $(1-P_i)$. Thus, we have $\Pr\{Z_i = 1\} = 1 - P_i$, and $\Pr\{Z_i = 0\} = P_i$. Therefore, the sequence $\{Z_i\}_{i\geq 0}$ forms a Bernoulli process, with success and failure of the ith trial given by the events $\{Z_i = 1\}$ and $\{Z_i = 0\}$, respectively.

Let i be the counter of slots during the RAW slot timeline, and $\{t_i\}_{i\geq 0}$ be an increasing sequence representing the beginning instants of slots with $t_0 = 0$. Let $\beta = T_{\text{TXOP}} + \text{DIFS}$ be the length of a busy slot and σ be the duration of an idle slot. Then, we can define the duration of the i^{th} slot as follows:

$$t_i - t_{i-1} = \begin{cases} \beta, & \text{if } \{Z_{i-1} = 1\}, \\ \sigma, & \text{if } \{Z_{i-1} = 0\}. \end{cases} \tag{5.22}$$

The holding period T_H should be sufficient to host one complete transmission, and hence we set $T_H = \beta$.

Note that in the case of a system where events have the same duration (i.e., $\sigma = \beta$), the number of busy slots during the RAW slot will be given by the aggregated number of successful trials of the Bernoulli process $\{Z_i\}_{i\geq 0}$ up to time T_S. However, that is not applicable as the number of events during the RAW slot duration T_S is unknown. That is because the slot length defined in (5.22) varies according to the outcome of the process $\{Z_i\}_{i\geq 0}$.

Therefore, we need a continuous-time model that considers the binary length of slots within the RAW slot period. For that, we shall construct a counting process that tracks busy slots continuously up to the length T_S of the RAW slot duration. Let busy slots represent the arrivals. An interarrival is a random time interval given by the length of a busy slot and the preceding idle slots. That is the total length of events associated with a successful outcome

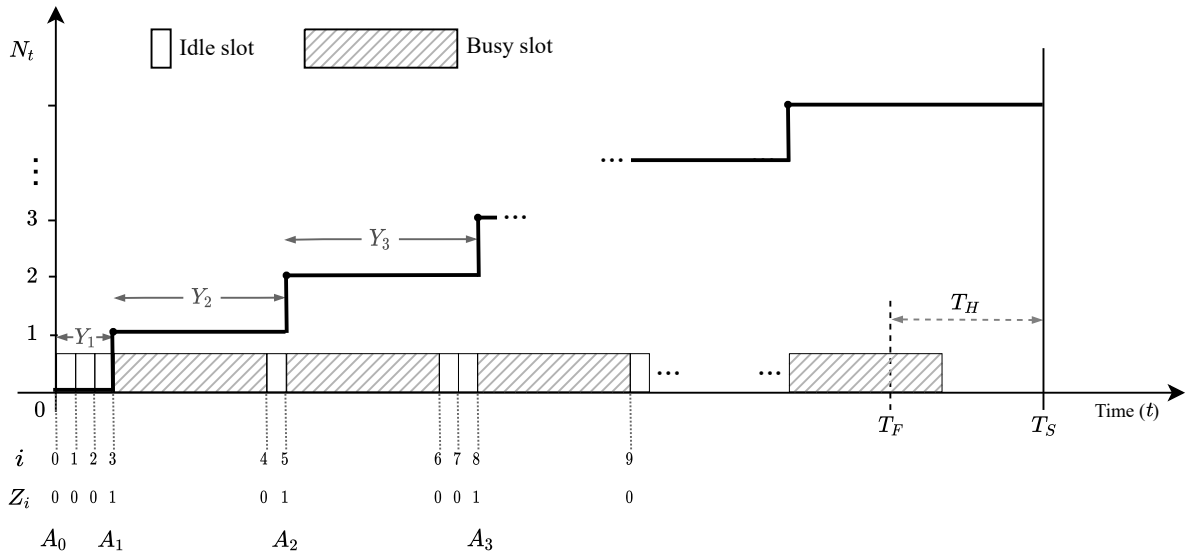


Figure 5.4: Illustration of the stochastic model.

of the Bernoulli process $\{Z_i\}_{i\geq 0}$ and the preceding aggregated failure outcomes. Although our framework is based on both discrete-time and continuous-time processes, we shall accurately derive the intended counting process within the context and norms of renewal theory.

Let N_t denote the number of busy slots that occurred up to the instant t, as depicted by the step function in Figure 5.4. We shall develop a renewal theory framework in which $\{N_t\}_{t\geq 0}$ represents a counting process that tracks busy slots up to and including time t.

For every $i \geq 1$, let X_i be a random variable representing the number of consecutive idle slots before the i^{th} success of the process $\{Z_i\}_{i\geq 0}$. Hence, $\{X_i\}_{i\geq 1}$ construct a sequence of IID random variables following a geometric distribution on $\mathbb{N}$. We have

$$\Pr\{X_i = k\} = (1-P_i)P_i^k\,, \quad k \geq 0\,. \tag{5.23}$$

Let $\{Y_i\}_{i\geq 1}$ be a sequence of random variables representing the interarrival times of busy slots as illustrated in Figure 5.4. Thus, we have

$$Y_1 = \sigma X_1, \quad Y_i = \sigma X_i + \beta, \quad i \geq 2 \tag{5.24}$$

Let $A_1, A_2, \ldots$ denote the arrival epochs of busy slots, where a new arrival is counted whenever the process $\{Z_i\}_{i\geq 0}$ gains a successful trial. Hence, we have

$$A_0 = 0, \quad A_k = A_{k-1} + Y_k, \quad k \geq 1\,, \tag{5.25}$$

where $A_0 = 0$ means zero arrival at time 0.

The interarrival periods $Y_1, Y_2, \ldots$ are positively distributed random variables, i.e., $\Pr\{Y_i >$

$0\} = 1$ and IID following the fact that $X_1, X_2, \ldots$ are IID random variables. Therefore, $\{A_k\}_{k\geq 0}$ is a well-defined renewal process and $\{N_t\}_{t\geq 0}$ is its associated counting process defined as follows:

$$N_t = \arg\max_{k\in\mathbb{N}} \{A_k \leq t\} \tag{5.26}$$

We now proceed to derive the expected number of busy slots $\mathrm{E}[N_{T_S}]$ within the RAW slot period T_S. The stations stop counting down their backoff counters at the end of the free access period T_F, and the holding period T_H is reserved for only the ongoing transmissions initiated before the boundary of T_F. Hence, we define the maximum number of transmissions that can occur during the RAW slot period as follows:

$$\Gamma_b = \left\lfloor \frac{T_F}{\beta} \right\rfloor + \mathbb{1}_{\left\{T_F > \left\lfloor \frac{T_F}{\beta} \right\rfloor \cdot \beta + \sigma\right\}} \tag{5.27}$$

Proposition 5.1. *The expected number of busy slots within a RAW slot of duration T_S allocated for N_S stations is given by:*

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\Gamma_b} \sum_{j=0}^{\left\lfloor \frac{T_F-(k-1)\cdot\beta}{\sigma} \right\rfloor} \binom{j+k-1}{j} (1-P_i)^k P_i^j . \tag{5.28}$$

Proof. We have by definition $N_{T_S} = \arg\max_{k\in\mathbb{N}}\{A_k \leq T_F\}$. Since $\{A_k\}_{k\geq 0}$ is an increasing process, we can define N_{T_S} as the cardinal of the set $\{k \geq 1 : A_k \leq T_F\}$. Thus,

$$\begin{aligned}
\mathrm{E}[N_{T_S}] &= \mathrm{E}\left[\arg\max_{k\in\mathbb{N}}\{A_k \leq T_F\}\right] \\
&= \mathrm{E}[\#\{k \geq 1 : A_k \leq T_F\}] \\
&= \mathrm{E}\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{A_k \leq T_F\}}\right] \\
&= \sum_{k=1}^{\infty} \Pr\{A_k \leq T_F\}
\end{aligned}$$

Since the number of transmissions within the RAW slot is limited by Γ_b, $\Pr\{A_k \leq T_F\} = 0$

for $k > \Gamma_b$. Henceforth,

$$\begin{aligned}
\mathrm{E}[N_{T_S}] &= \sum_{k=1}^{\Gamma_b} \Pr\{A_k \leq T_F\} \\
&= \sum_{k=1}^{\Gamma_b} \Pr\left\{\sum_{i=1}^{k} Y_i \leq T_F\right\} \\
&= \sum_{k=1}^{\Gamma_b} \Pr\left\{(k-1)\cdot\beta + \sigma\sum_{i=1}^{k} X_i \leq T_F\right\} \\
&= \sum_{k=1}^{\Gamma_b} \Pr\left\{\sum_{i=1}^{k} X_i \leq \frac{T_F-(k-1)\cdot\beta}{\sigma}\right\} \\
&= \sum_{k=1}^{\Gamma_b} \sum_{j=0}^{\left\lfloor\frac{T_F-(k-1)\cdot\beta}{\sigma}\right\rfloor} \Pr\left\{\sum_{i=1}^{k} X_i = j\right\}
\end{aligned}$$

$\sum_{i=1}^{k} X_i$ forms a sum of k IID random variables following a geometric distribution with parameter $(1-P_i)$. Hence, $\sum_{i=1}^{k} X_i$ is a random variable that follows the negative binomial distribution with parameters k and $(1-P_i)$. Thus, we have

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\Gamma_b} \sum_{j=0}^{\left\lfloor\frac{T_F-(k-1)\cdot\beta}{\sigma}\right\rfloor} \binom{j+k-1}{j} (1-P_i)^k P_i^j .$$

□

Proposition 5.2. *The expected number of idle slots within a RAW slot of duration T_S allocated for N_S stations is given by:*

$$\mathrm{E}[I_{T_S}] = \frac{P_i}{1-P_i} \cdot \mathrm{E}[N_{T_S}] . \tag{5.29}$$

Proof. Let I_{T_S} be a random variable representing the number of idle slots within the RAW slot period T_S. Note that idle slots occur only during the free access period T_F as all stations terminate their backoff counters at the boundary of T_F.

Thus, I_{T_S} represents the aggregated number of idle slots preceding all transmissions that occurred during the RAW slot. Since we have N_{T_S} busy slots within the RAW slot and each busy slot k is preceded by X_k idle slots, we have $I_{T_S} = \sum_{k=1}^{N_{T_S}} X_k$.

We have $\{X_k; k \geq 1\}$ is a sequence of IID random variables with same mean $\overline{X} = \mathrm{E}[X_k], \forall k \geq 1$. Additionally, $\mathrm{E}[N_{T_S}] < \infty$ as proved in Proposition 5.1.

Therefore, according to Wald's equality [9], we have

$$\mathrm{E}[I_{T_S}] = \mathrm{E}\left[\sum_{k=1}^{N_{T_S}} X_k\right] = \overline{X} \cdot \mathrm{E}[N_{T_S}] . \tag{5.30}$$

In addition, for every $1 \leq k \leq N_{TS}$, X_k follows a geometric distribution on $\mathbb{N}$ with rate $1-P_i$.

Thus,

$$\overline{X} = \frac{P_i}{1-P_i} \tag{5.31}$$

By substituting (5.31) in (5.30), we obtain the result. □

After deriving the expected number of idle and busy slots that occupy the free access period of the RAW slot, we can now study the channel usage during the holding period. Thus, we define the holding period usage ratio, representing the portion of time from the holding period that is exploited by transmissions, as follows:

$$U_h = \frac{E[I_{TS}] \cdot \sigma + \mathrm{E}[N_{TS}] \cdot \beta - T_F}{T_H}. \tag{5.32}$$

5.3.2 Evaluating a Single RAW Slot

Since busy slots represent three types of slots, as mentioned previously, we can derive from $\mathrm{E}[N_{TS}]$ three average numbers of slots containing successful single transmissions, collisions with capture, and collisions without capture.

As stated previously, a randomly chosen slot accommodates a successful single transmission (i.e., is in the state "Success") with probability $(1-P_i)P_s$. Therefore, knowing that a random slot is busy with probability $(1-P_i)$, we define the expected number of slots containing successful single transmissions during the RAW slot period as follows:

$$A^s_{TS} = \mathrm{E}[N_{TS}] \cdot P_s. \tag{5.33}$$

A randomly chosen slot contains a collision with capture when one of the interfering packets is successfully captured and delivered to the AP. Given that such an event happens with probability $(1-P_i)(1-P_s)P_{cap}$ and that it is a subset event of non-idle slots, we define the average number of slots in the state "Capture" within the RAW slot as follows:

$$A^{cap}_{TS} = \mathrm{E}[N_{TS}] \cdot (1-P_s) \cdot P_{cap}. \tag{5.34}$$

A randomly chosen slot contains a collision with no capture when none of the interfering packets is captured from the collision. That is, all transmitted packets faced a failed delivery in the considered slot. Since such an event is a subset of busy slots and a random slot is in this state with probability $(1-P_i)(1-P_s)(1-P_{cap})$, we have the average number of slots in the state

"Failure" within the RAW slot given by:

$$A_{T_S}^{f} = \mathrm{E}[N_{T_S}] \cdot (1 - P_s) \cdot (1 - P_{cap}) . \tag{5.35}$$

A data frame delivery can be successful when only one station is transmitting, or two or more stations are transmitting simultaneously and the AP successfully captures one packet following the collision. We define the RAW slot throughput as the ratio of time used to successfully transmit data frames during the period T_S. That is the time occupied by the successful delivery of data frames during successful single transmissions and during the successful delivery of captured packets from collisions. Therefore, the RAW slot throughput is expressed as follows:

$$Th_S = \frac{\left(A_{T_S}^{s} + A_{T_S}^{cap}\right) \cdot T_{\mathrm{DATA}}}{T_S} . \tag{5.36}$$

To determine the gain from the capture feature in the network, we need to compare the network performance with the case of the no-capture channel. There are several differences between a channel with capture and a no-capture channel. Starting from the EDCA mechanism for channel access, in which a collided packet can still be captured in the first scenario, and hence the source station resets its CW to CW_0. In contrast, there is no such behavior in the case of a no-capture channel.

Henceforth, in the case of a no-capture channel, $p_{cap}^{p} = 0$ and $p = p_{col}^{p}$. Therefore, τ and p can be obtained by solving numerically the non-linear system defined by (5.8) and (5.9).

As for the channel slots, a randomly chosen slot has three states: idle, success, and collision, where the states "Idle" and "Success" are the same as defined in the previous scenario. Additionally, a chosen random slot can be in a state of collision when it hosts more than one transmission, which happens with probability $(1 - P_s)(1 - P_{cap})$.

Hereafter, the RAW slot throughput in the case of a no-capture channel is given by:

$$Th_S^{NC} = \frac{A_{T_S}^{s} \cdot T_{\mathrm{DATA}}}{T_S} . \tag{5.37}$$

5.3.3 Evaluating Several RAW Slots

Consider a RAW uniformly divided into K RAW slots, each of duration $T_S = T_R/K$. According to the stations' assignment defined in Chapter 2-(2.1), the RAW structure will end up with $K_1 = K - (N_R \mod K)$ RAW slots, each one allocated to $N_S^1 = \lfloor N_R/K \rfloor$ and $K_2 = K - K_1$ RAW slots, each allocated to $N_S^2 = N_S^1 + 1$. Note that if the total number of stations in the RAW N_R is divisible by the number of RAW slots K, we will have a uniform allocation of both RAW slots'

duration and assigned stations. Thus, the RAW will consist of K similar RAW slots of duration $T_S = T_R/K$ with $N_S = N_R/K$ assigned stations.

The RAW throughput is defined by the portion of the time used to deliver data frames during the RAW period T_R successfully. Hence, the normalized throughput of a RAW with K RAW slots is expressed as follows:

$$Th_R(K) = \frac{\left(K_1\left(A^s_{T_S^1} + A^{cap}_{T_S^1}\right) + K_2\left(A^s_{T_S^2} + A^{cap}_{T_S^2}\right)\right) \cdot T_{\text{DATA}}}{T_R}. \tag{5.38}$$

The normalized throughput of a RAW consisting of K RAW slots in the case of a no-capture channel is given by:

$$Th^{NC}_R(K) = \frac{\left(K_1 \cdot A^s_{T_S^1} + K_2 \cdot A^s_{T_S^2}\right) \cdot T_{\text{DATA}}}{T_R}. \tag{5.39}$$

To derive the added value of the capture feature within the RAW, we compare the RAW throughput for a channel with capture to the no-capture channel. Thus, we define the RAW capture ratio for a RAW consisting of K RAW slots as the portion of the RAW throughput gained through the capture effect, and it is given by:

$$G_{cap}(K) = \frac{Th_R(K) - Th^{NC}_R(K)}{Th_R(K)}, \tag{5.40}$$

where $Th_R(K)$ is the RAW throughput in the case of a channel with capture, defined in (5.38). The definition in (5.40) means that $G_{cap}(K)$ percent of the RAW throughput $Th_R(K)$ is gained thanks to the capture effect, where packets are successfully delivered after being captured from collisions.

5.4 Results and Discussion

This section evaluates and analyzes our findings through numerical results and simulations using MATLAB software. It also presents a validation for our mathematical framework and the obtained analytical results through extensive simulations derived via a discrete-event simulator we developed with MATLAB. The simulator mimics the channel access using the EDCA mechanism, considering the random distribution of stations around the AP, the power attenuation of received packets under a Rayleigh fading channel, and the capture of a packet in case of collision if its Signal-to-Interference Ratio (SIR) exceeds the capture threshold z. A single simulation of a RAW slot yields the contention results from its beginning until the end of its duration. The final results represent the average of 10,000 simulations. We observe in the figures below a close match of the results obtained from our model and the simulations, which

Table 5.1: Parameters

Parameter	**Value**
Data rate	1.95 Mbps
MacHeader	272 bits
T_{ACK}	1000 µs
T_{PLCP}	80 µs
σ	52 µs
SIFS	160 µs
DIFS	264 µs
CW_{min}	8
CW_{max}	16
m	1
Payload	160 bytes
ρ	100 m

validates the accuracy of our mathematical framework.

We shall first study the performance of one RAW slot for different parameters: its duration T_S, the number of designated stations N_S, and the capture threshold z. We then examine a RAW consisting of several RAW slots by studying the impact of the grouping on the overall RAW performance under the presence of the capture effect. We additionally investigate the RAW capture ratio under different RAW configurations. We consider sensor stations of category video or Voice over Internet Protocol (VoIP), allowing one retransmission attempt of a collided packet according to the standard [4]. Besides the parameters we put in the study, the remaining parameters considered in both the analytical framework and simulations are presented in Table 5.1.

5.4.1 RAW Slot

5.4.1.1 RAW Slot Period

We evaluate in this section the RAW slot throughput as a function of its duration T_S, which represents the portion of time used for the successful transmission of data frames. We consider a RAW slot allocated to 10 stations, distributed uniformly within the radius $\rho = 100$ m, and a channel with capture threshold $z = 8$ dB.

Figure 5.5 depicts the expected number of busy and idle slots in terms of the duration of the RAW slot period. The figure also depicts the holding period usage ratio U_h, representing the portion of T_H exploited by transmissions. These three results start with the value zero and yield

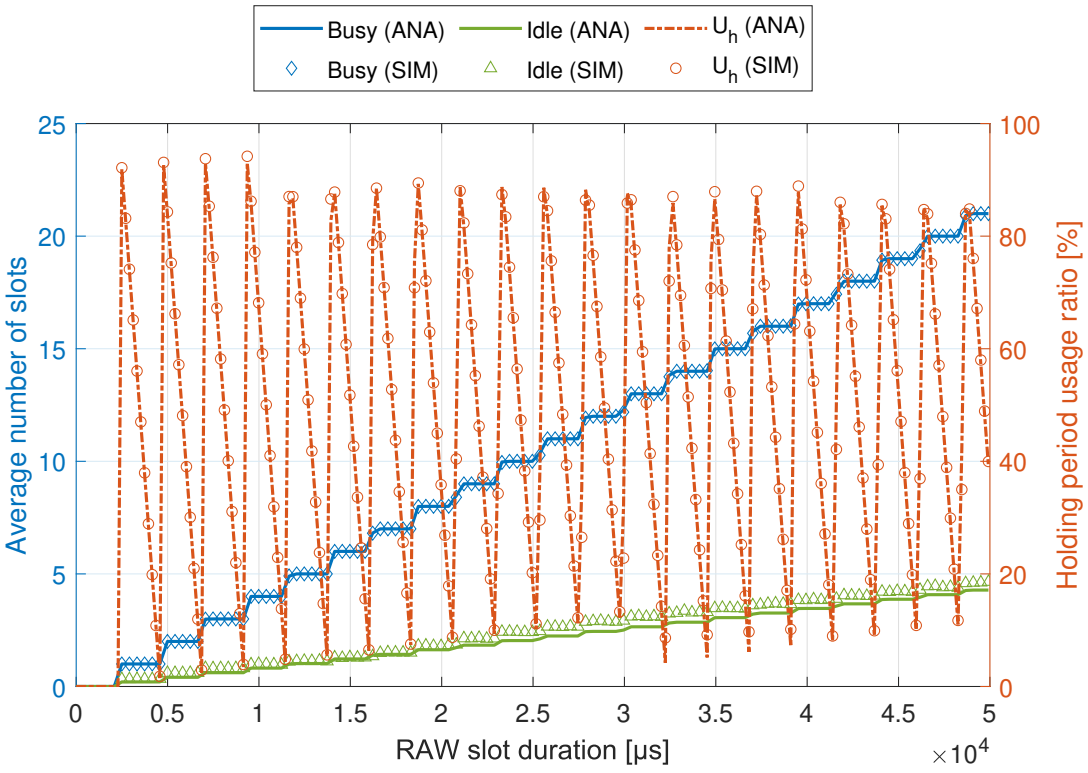


Figure 5.5: Average number of busy and idle slots, and the holding period usage ratio within the RAW slot interval, with channel capture threshold $z = 8\,\text{dB}$.

positive values after a given value of T_S. That corresponds to when T_S is long enough to host at least one transmission. The average number of busy slots $\mathrm{E}[N_{T_S}]$ is almost a step function that increases by one unit whenever T_S is long enough to host an additional complete transmission. Similar behavior is shown by the average number of idle slots $\mathrm{E}[I_{T_S}]$ since idle slots precede each busy slot. U_h has a fluctuating behavior as T_S increases. When the additional time in the RAW slot interval is not enough to host additional transmissions, $\mathrm{E}[N_{T_S}]$ remains constant, and the holding period becomes unused. That causes a fast drop in the value of U_h. When RAW slot duration T_S increases enough to allow additional transmissions, $\mathrm{E}[N_{T_S}]$ and $\mathrm{E}[I_{T_S}]$ increase while U_h jumps to the maximum possible usage ratio. We observe that the variation interval of U_h decreases as T_S increases. That is because the holding period T_H becomes less significant within the RAW slot interval and hence less affecting in cases used or wasted. Eventually, as T_S tends to infinity, the usage of the holding period becomes completely unaffecting, and the value of U_h converges to 50%.

Figure 5.6 represents the average number of slots in each state of success, capture and failure. Since all three states are a subset of the busy slot state, we observe that the average number of slots in each state has the same behavior as $\mathrm{E}[N_{T_S}]$ presented in Figure 5.5. That is explained by the same aforementioned phenomena of the RAW slot period, which allows additional transmissions only when long enough. Note that in the presented scenario, the average number of slots with a successful single transmission $A^s_{T_S}$ is larger than the other two types due to the low contention of the considered ten stations.

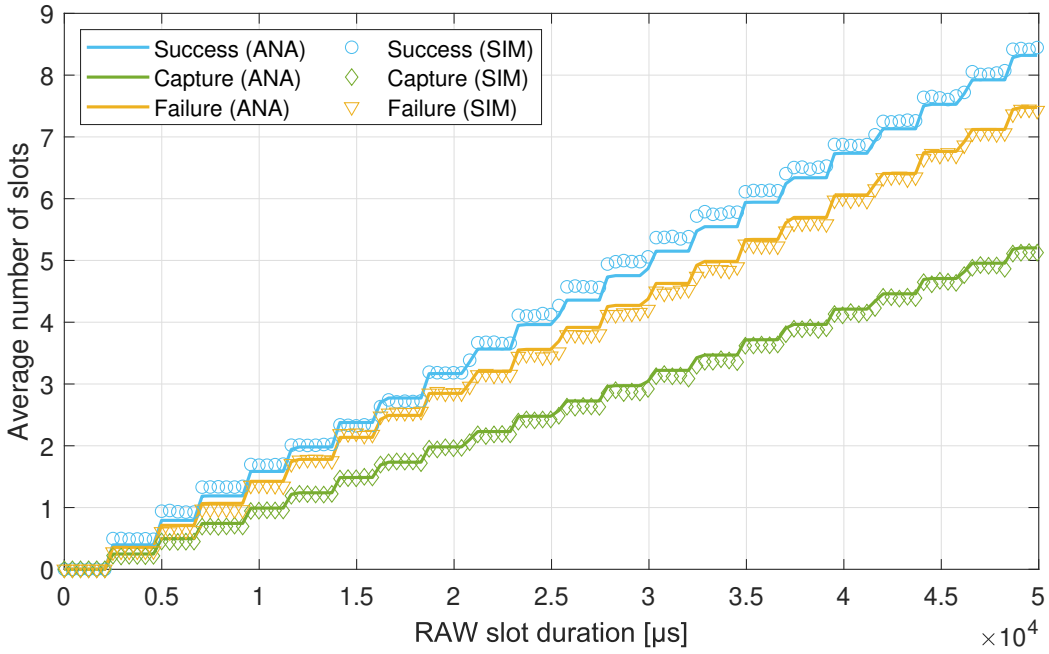


Figure 5.6: Average number of slots with successful single transmission, captured packet, and failure, within the RAW slot interval, with channel capture threshold $z = 8\,\text{dB}$.

Figure 5.7 presents the RAW slot throughput in terms of its allocated period T_S. We consider two channel scenarios: a no-capture channel and a channel with capture for different capture threshold values z: 2 dB, 4 dB, 8 dB, and 16 dB. The RAW slot throughput represents the time used to successfully transmit data frames to the AP during the RAW slot period T_S. In both channel scenarios, the RAW slot throughput has a fluctuating behavior due to the usage of the holding period when T_S increases, as depicted in Figure 5.5. When the RAW sot duration T_S increases and the extended period is not enough to host another transmission, the additional time in T_S is wasted, and the holding period usage ratio U_h decreases, which results in a drop in the RAW slot throughput. On the other hand, as long as T_S can allow an additional transmission, the RAW slot throughput jumps to a higher value. When T_S is long enough, the presence or absence of one transmission does not affect the total transmissions in the RAW slot period, and hence, the fluctuating behavior slightly disappears. Eventually, the RAW slot throughput converges to the case of an infinite RAW slot. Regarding the capture feature, we observe that a channel with capture yields higher throughput for the RAW due to the supplementary delivered packets captured from collisions. However, the capture threshold has an inverse impact on the throughput. A lower capture threshold means that the AP can capture a packet from a collision even if its received power shortly exceeds the joint power of interfering packets. This effect results in more captured packets from collisions and hence higher throughput. Nevertheless, increasing the capture threshold lowers the possibility of capturing packets, which reduces the RAW slot throughput. Ultimately, when the capture threshold tends to infinity, the capture of packets becomes unfeasible, and hence the RAW slot throughput converges to the throughput provided by a channel without capture.

Our developed analytical framework assumes that every station transmits its packet in a

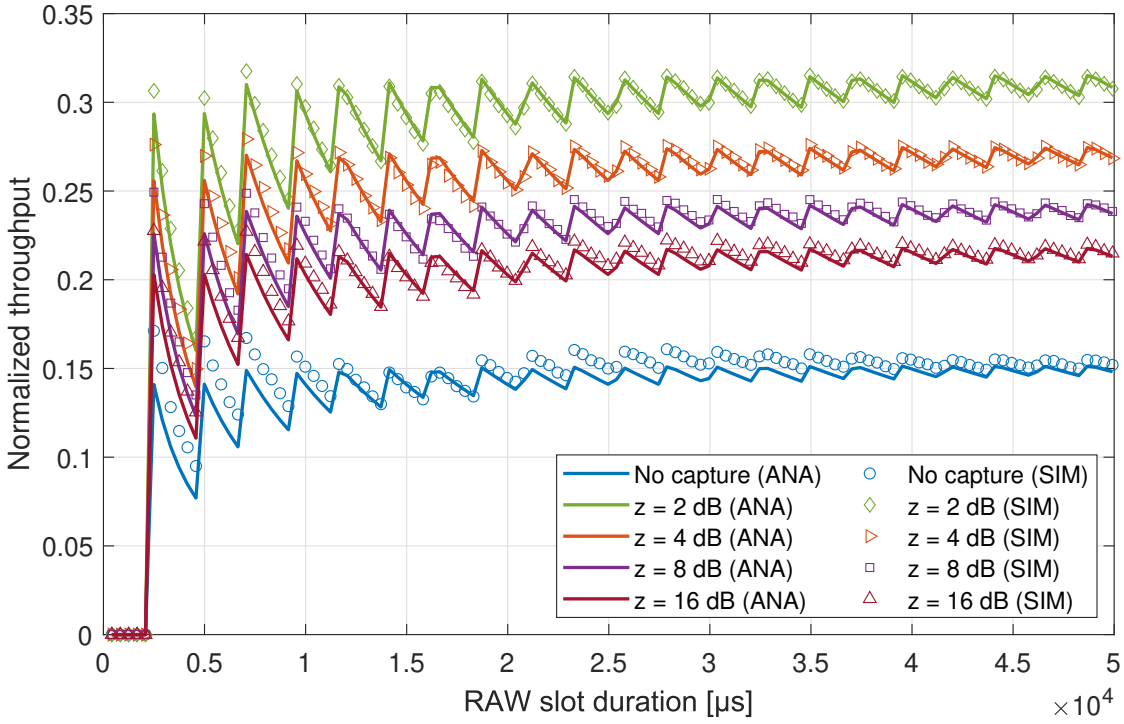


Figure 5.7: Impact of the capture threshold on the RAW slot throughput.

randomly chosen slot with the same probability τ, defined in (5.8). The latter depends on the attempts and idle slots experienced by the station. The slight differences between the analytical and simulation results depicted in Figures 5.5–5.7 are due to the channel access probabilities that depend on τ and the number of contending stations. The analytical throughput in Figure 5.7 yields slightly lower values than the simulation, especially for a very short RAW slot. That is because the model assumes that every station goes through all possible transmission attempts and backoffs, and hence more potential collisions. Whereas in the simulation, the contention is just starting and the backoff counters of most of the stations are still in their first contention window. However, the analytical throughput always keeps the same trends as the simulation, which preserves accuracy for deriving optimal configuration that maximizes the throughput.

5.4.1.2 Contending Stations

We consider a group of stations distributed uniformly around the AP within the radius $\rho = 100\,\text{m}$. We evaluate the throughput of a RAW slot with length $T_S = 20\,\text{ms}$ in terms of the designated stations.

Figure 5.8 presents the average number of the three types of busy slots during the RAW slot period regarding the number of contending stations. When more stations are assigned to the RAW slot, the contention increases and yields more collisions. This lowers the number of slots with a successful single transmission and raises slots containing a capture and slots with failure. Furthermore, the average number of slots $A_{T_S}^{cap}$ containing a collision with capture has a downward concave shape behavior. Given that more stations bring on more collisions, the AP

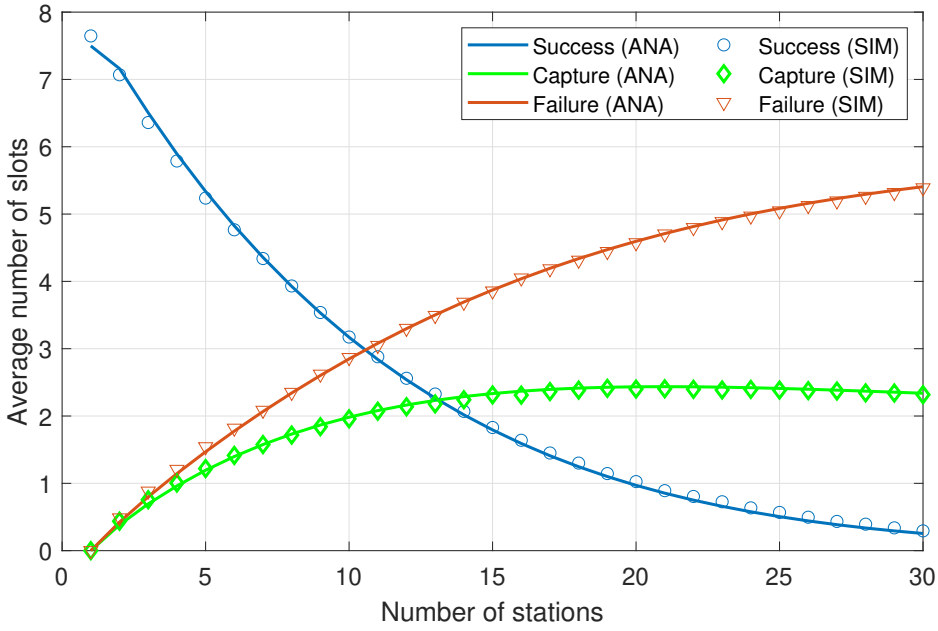


Figure 5.8: Average number of slots with successful single transmissions, captured packets, and failures within a RAW slot interval of length $T_S = 20\,\text{ms}$ and channel capture threshold $z = 8\,\text{dB}$.

captures more packets from these collisions, which increment the value of $A_{T_S}^{cap}$. Eventually, a higher number of stations will be involved in collisions, which lowers the chances of capturing a packet due to the higher joint received power of interfering packets. That results in limiting the increment rate of $A_{T_S}^{cap}$ and ultimately diminishes its value when a large number of stations are assigned to the RAW slot.

Figure 5.9 shows the channel usage ratio during the RAW slot period of length $T_S = 20\,\text{ms}$ with capture enabled at threshold $z = 8\,\text{dB}$. When more stations participate in the RAW slot, the contention arises, and the channel becomes more occupied by transmissions, diminishing the idle time. Moreover, the time consumed by collisions (slots with capture and failure) increases, whereas the time used for successful single transmissions rapidly degrades due to the immense contention generated by more stations in such a short period of the RAW slot. However, the capture feature takes advantage of the rise of collisions, and more packets are captured when the RAW slot becomes denser. As observed in the figure, the capture feature in the channel allows to successfully deliver packets even when typical single transmissions are not possible (e.g., when $N_S = 30$). The time used to transmit ACK frames follows the trend of successful transmissions that occurred during both single transmissions and captured packets. The figure also shows a slightly better usage of the holding period for fewer stations. That is, fewer stations can initiate transmissions occupying a more significant part of the holding period. Note that such a result is also related to other factors, such as the RAW slot duration and payload size. The usage ratio of the holding period is mainly affected by the RAW slot duration, as discussed in Figure 5.5.

Hereafter, we examine in Figure 5.10 the impact of the capture threshold z. We first consider a RAW slot with a fixed length $T_S = 20\,\text{ms}$. Then, we explore its throughput in terms of the designated stations for a no-capture channel and a channel with capture for different capture

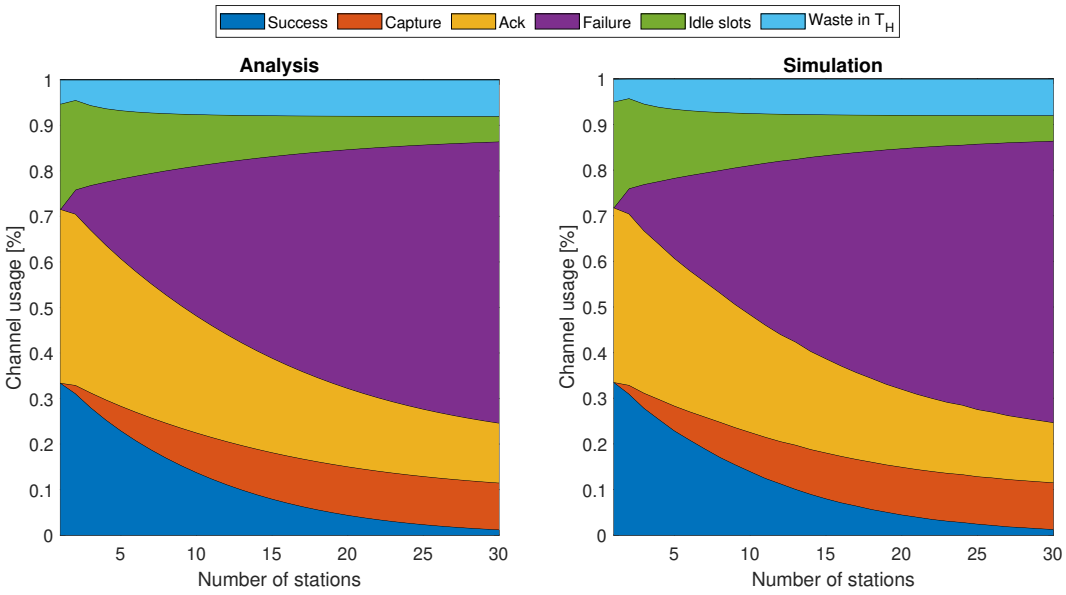


Figure 5.9: Channel usage ratio by different events occupying the RAW slot interval with capture threshold $z = 8\,\text{dB}$, in terms of contending stations.

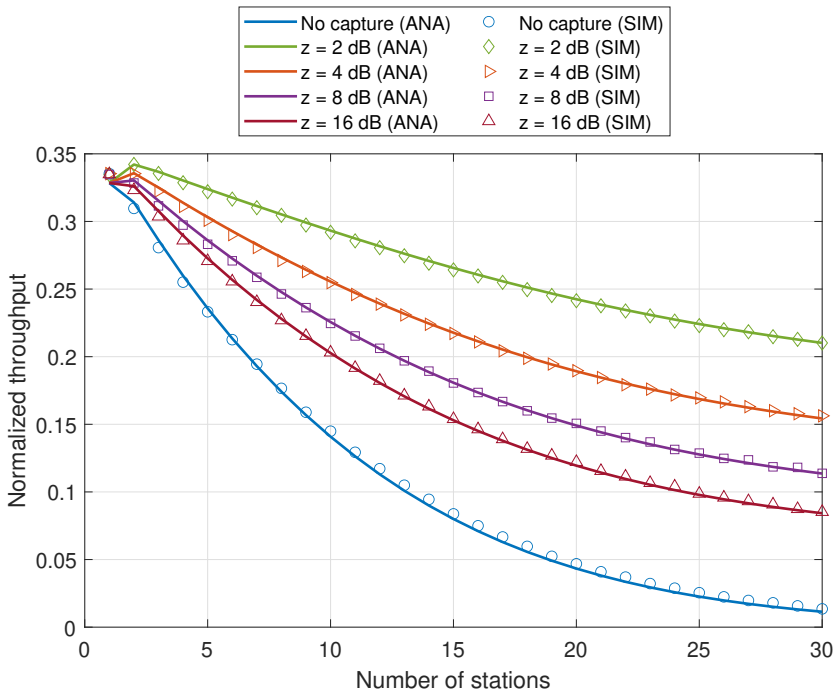


Figure 5.10: Impact of the capture threshold on the RAW slot throughput.

thresholds: $z = 2\,\text{dB}, 4\,\text{dB}, 8\,\text{dB}, 16\,\text{dB}$. A channel with a smaller capture threshold provides higher throughput during the RAW slot period. Conversely, a higher capture threshold value lessens the captured packets and henceforth degrades the RAW slot throughput. Ultimately, when the capture threshold z rises, the RAW slot throughput converges to the result provided by a normal channel without capture.

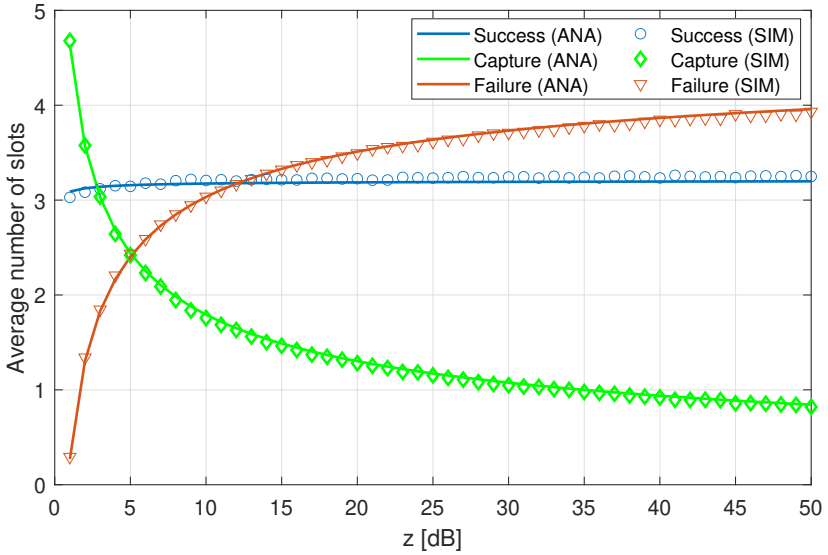


Figure 5.11: Average number of slots with successful single transmissions, captured packets, and failures, in terms of the capture threshold, within a RAW slot of duration $T_S = 20$ ms and $N = 10$ assigned stations.

5.4.1.3 Capture Threshold

We consider a RAW slot of length $T_S = 20$ ms allocated for $N = 10$ stations, and we investigate the impact of the capture threshold z on the RAW slot throughput.

Figure 5.11 depicts the average number of each of the three states of busy slots during the RAW slot period. The average number of slots with a successful single transmission shows a slightly increasing variation for short values of the capture threshold z, which is explained by the higher values achieved by $A_{T_S}^{cap}$ at the same period of z. When a packet is captured, its source station resets its CW to CW_0, which increases the station's chances of delivering another packet without colliding. As the capture threshold z increases, it becomes unlikely for a received power of one packet to exceed the joint interfering powers by the threshold z. That translates into a decrease in the average number of slots containing collision with capture $A_{T_S}^{cap}$ and an increase in the average number of slots containing failure (collision with no capture) $A_{T_S}^{f}$.

Hereafter, we explore the impact of both the capture threshold and the number of assigned stations in the RAW slot. Figure 5.12 depicts the normalized throughput within a RAW slot of duration $T_S = 25$ ms in terms of the capture threshold for different sets of stations. As long as z gets a higher value, it becomes less likely for the AP to capture packets from collisions, as seen in Figure 5.11. For that reason, increasing the capture threshold value lowers the ability to capture packets from collisions and degrades the RAW slot throughput. As a result, it eliminates the capture feature and converges to the scenario of a no-capture channel. From the perspective of stations assigned to the RAW slot, we observe a higher capture gain for a denser RAW slot. That is due to high contention, which presents more collisions and hence more captured packets. Hereafter, we see that ten stations operating within a channel with a capture threshold $z \leq 6$ dB yield higher throughput than five stations operating within a no-capture channel. Similarly,

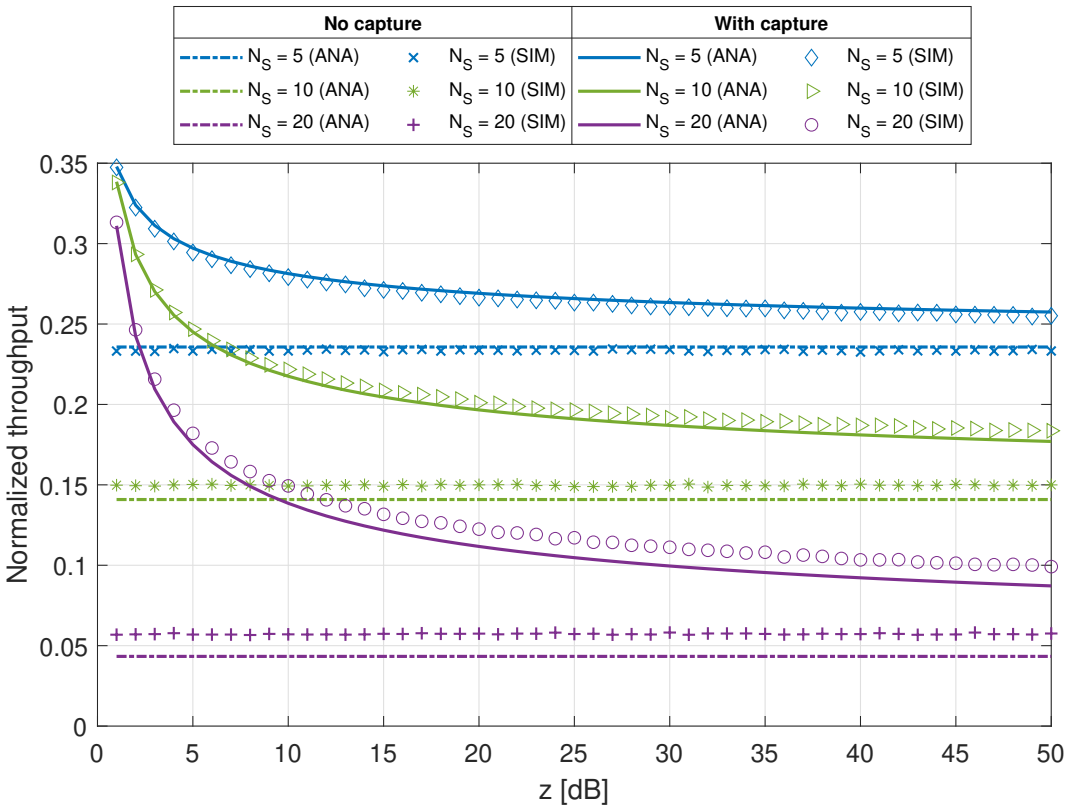


Figure 5.12: Impact of the capture threshold and the number of stations assigned to the RAW slot.

we observe that 20 stations operating within a channel with a capture threshold $z \leq 2\,\text{dB}$ and $z \leq 9\,\text{dB}$ provide better throughput than five stations and ten stations, respectively, operating within a no-capture channel. Such results are significant to consider in developing efficient scheduling algorithms for configuring the RAW mechanism.

The observed difference between the analytical and simulation results for 10 and 20 stations is due to the same reason explained in the previous subsection. The evaluation of a short RAW slot of duration $T_S = 25\,\text{ms}$ when allocated to a higher number of stations predicts more collisions. Whereas, in the simulation, the backoff counters of most of the stations are still on the first contention window W_0 and it is less likely for collisions to occur. However, we observe a higher accuracy of the throughput for five stations, which is a more practical scenario for such a short RAW slot. Additionally, we can see a higher accuracy of the throughput for 10 and 20 stations in the case of a channel with capture enabled, especially for a small threshold that allows more captured packets and eliminates the misprediction of higher collisions in the no-capture channel case.

5.4.2 RAW Performance

The IEEE 802.11ah standard introduced the grouping within RAW to reduce collisions among allocated stations to the RAW. However, the standard does not specify the algorithm to perform

such grouping. Henceforth, this section examines the impact of grouping the nodes as evenly as possible across the available RAW slots. As shown in the previous section, short RAW slots yield a fluctuating throughput behavior. From such behavior, we can deduce that it is not obvious to make an optimal choice about the number of RAW slots that maximizes the RAW gain.

Figure 5.13 shows the normalized throughput for a $T_R = 500\,\text{ms}$ RAW in terms of the number of RAW slots in a RAW for four different numbers of stations, $N_R = 64, 140, 300, 600$. As seen in the figure, subdividing the RAW into more RAW slots increases the normalized throughput up to a certain point from which the throughput presents light fluctuations around a saturation point. This behavior is explained by the fact that by subdividing the RAW into more groups, the RAW slot duration T_S becomes shorter, and as a result, the T_H/T_S ratio increases. Since T_H may be exploited or not, the RAW throughput keeps decreasing and increasing as the number of groups in the RAW increases. Note that such behavior does not appear for a small number of groups because the RAW slots in this range are long enough not to be affected by the use or non-use of the holding period. Furthermore, the overall RAW throughput achieves a maximum value at a given number of RAW slots, as depicted in the figure, where K_{op}^N is the optimal number of slots when N stations are assigned to the RAW. In the case of RAWs consisting of $N_R = 140, 300, 600$, a higher number of slots than the use for $N_R = 64$ is required before reaching the maximum achievable throughput. As depicted in the figure, an increase in the number of slots beyond the optimal number of slots may cause a drop of up to 5% on the maximum achievable throughput for a network consisting of $N_R = 140$. We also observe that a RAW with fewer assigned stations has a sharper increase in normalized throughput when the RAW is subdivided into more RAW slots. That is due to lighter contention in each RAW slot after allocating more RAW slots in the RAW period. However, a dense RAW is slightly affected by the grouping due to the high contention that is still present in each RAW slot.

Figure 5.14 depicts the RAW capture ratio, representing the ratio of throughput gained from captured packets during the RAW period $T_R = 500\,\text{ms}$. We observe that when the RAW is composed of one RAW slot, i.e., all stations are contending for channel access simultaneously, 100% of the RAW throughput is derived from the capture of packets from collisions. That means there will be no successful transmissions within the RAW duration in such conditions with a no-capture channel. Allocating more RAW slots within the RAW leads to fewer stations in each RAW slot, decreasing collisions and allowing successful single transmissions. Thus, the percentage of captured packets among the total delivered ones becomes less, and hence the RAW capture ratio decreases. Eventually, the capture ratio becomes null when collisions are eliminated within the RAW, which coincides with the case of one station per RAW slot. Note that this case may not be possible if the number of stations is greater than 64 or when the length of each RAW slot is not sufficient to host at least one transmission. Furthermore, we observe a slight difference between the analytical results and simulations in both Figures 5.13 and 5.14 for the cases of 300 and 600 stations. That is due to the fact that each RAW slot in these cases

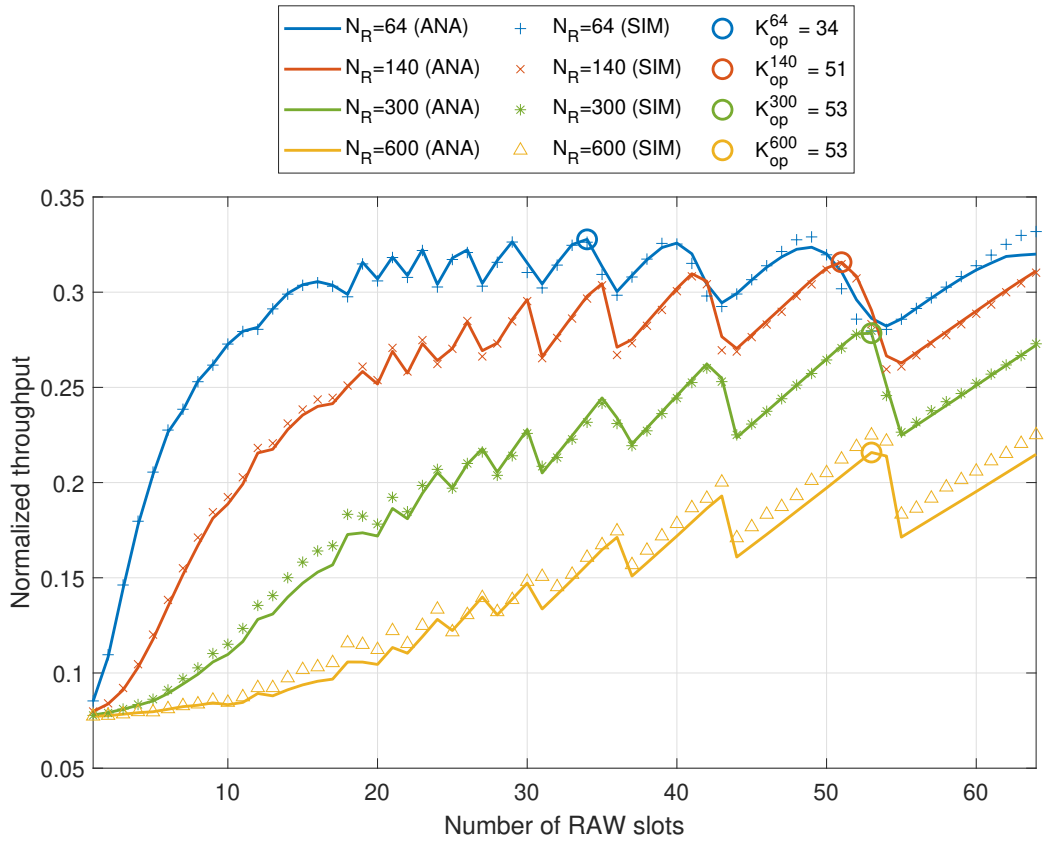


Figure 5.13: RAW throughput in terms of allocated RAW slots for different sets of assigned stations.

contains a higher number of stations, and hence more disturbance to the throughput due to the prediction of more collisions as discussed in Figures 5.7 and 5.12.

Our results show that our modeling framework allows us to evaluate the performance of the IEEE 802.11ah RAW mechanism by taking into account key system parameters, namely, the number of nodes, number of groups, channel bit rate, and packet length. Furthermore, our methodology also includes modeling the wireless channel and the network layout, i.e., placement of the nodes with respect to the AP. Our resulting modeling framework contributes to the development of evaluation and optimization solutions of IEEE 802.11ah networks: an area of great interest to the research and development of dense wireless systems [93].

5.5 Conclusion

In this chapter we presented an accurate mathematical framework to model and evaluate the performance of an IEEE 802.11ah-based network with the presence of the capture effect in the channel. We first modeled channel access under capture awareness, considering a Raleigh-fading channel where the capture model considers the stations' distance from the AP and the power attenuation of the received packets. We further defined the different states of a randomly chosen slot within the RAW slot period. Since the RAW mechanism presents a time-limited

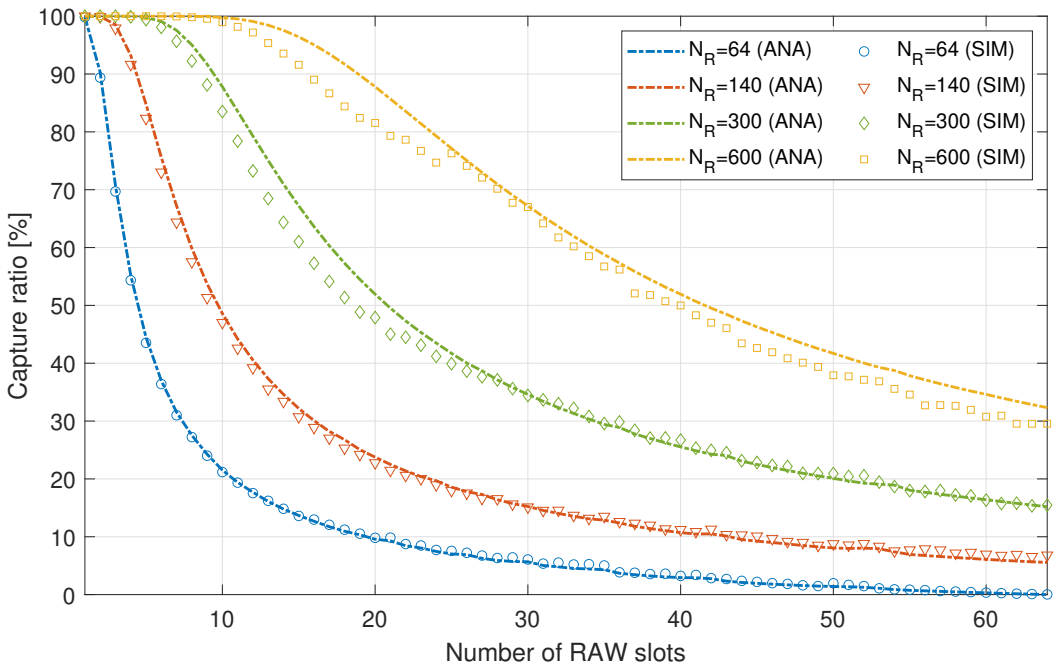


Figure 5.14: Capture gain in terms of allocated RAW slots for different sets of assigned stations to the RAW.

contention for channel access, we then developed a counting process to track transmissions up to the end of the contention time interval. Hereafter, we evaluated the performance of a single RAW slot and a RAW comprising several RAW slots in terms of throughput, representing the ratio of time used to successfully deliver data frames during the total contention period. The derived analytical results are meticulously validated through extensive simulations obtained via a simulator we developed with MATLAB software. We studied the impact of different parameters on the overall performance of the time-limited contention, including the contention time, the number of stations, and the capture threshold. We also examined the capture ratio within the RAW, representing the portion of RAW throughput gained via the capture effect in the channel. It is shown that a very dense RAW can only deliver data through captured packets. However, the capture ratio degrades when the RAW comprises a more significant number of RAW slots due to less contention in each RAW slot. Besides the RAW duration and the number of assigned stations, several other parameters affect the overall RAW throughput when considering a channel with capture as the capture model, the capture threshold, stations' locations, and the power attenuation. Henceforth, it is critical to consider all these parameters to derive efficient scheduling algorithms for the RAW mechanism. The proposed analytical model and associated results present an excellent basis for proposing practical scheduling algorithms to configure the RAW mechanism under a fading channel and a capture effect at the receiver. Henceforth, we develop in the next chapter efficient algorithms to optimize the RAW slot duration and to configure a RAW consisting of several RAW slots. These algorithms significantly improve the throughput and enhance the energy efficiency of the network.

CHAPTER 6

Resource and Energy-Efficient Configuration of IEEE 802.11ah Networks under Rayleigh Channels

Contents

6.1 Introduction

The immense growth of devices of the IoT makes traditional channel access protocols ineffective and unable to provide reliable interconnection between different nodes [94]. The low-power and long-range Wi-Fi standard IEEE 802.11ah is a promising candidate to meet the high traffic

load of IoT applications and provide excellent QoS [4]. IEEE TGah introduces a new MAC protocol, called RAW [4]. The RAW mechanism limits concurrent channel access by dividing the channel into RAW slots and assigning a small group of stations to each. The scheduling process is carried out at the start of each Beacon Interval (BI) [4]. The Access Point (AP) may allocate one or more RAWs within a BI. Only designated stations are allowed to access the channel during their allocated RAW slot. However, the IEEE 802.11ah standard does not mention how to configure the RAW parameters. Numerous studies have been released; nevertheless, there is no general consensus on the most effective RAW scheduling strategy [95].

Evaluating the RAW mechanism presents the challenge of time-limited contention [63], which makes analytical models based on stationary results such as Bianchi's [5] inapplicable. However, an accurate model for time-limited contention should provide results that converge to the stationary case where the contention lasts long enough time. Nevertheless, many works proposed analytical frameworks to model and evaluate the IEEE 802.11ah standard [25, 3]. A Markov chain based model is proposed in [47], where the Markov process counts the number of active stations and slots that occurred in the idle, collision, and success states. A similar model is deployed in [54] evaluating the energy consumption for RAW-based channel access while considering one packet at most per station in each BI. This model is effective but highly complex due to the absorbing states of the Markov chain defined by the condition of the duration of occurring events reaching the RAW slot limit.

Alternatively, we established different models based on renewal theory approaches in our previous works [63, 64, 65]. We modeled time-limited contention within a renewal theory framework and implemented a counting process to track transmissions that occurred during the RAW slot period [63, 64]. Such an approach allows distinguishing the different events that occurred during the contention and hence evaluating the gain of a RAW slot and a RAW composed of several RAW slots. We also developed an analytical framework to model and evaluate the network in the case of a Rayleigh channel with the capture effect enabled [65].

We observe from our previous studies [63, 64] that the RAW slot performance is very sensitive to its duration, resulting in fluctuating behaviors. Increasing the allocated time may enhance or degrade the performance, spending on the number of assigned stations. Therefore, in this work, we aim to efficiently configure the RAW mechanism under a Rayleigh-fading channel with the capture effect, which is a realistic scenario for dense IoT networks deployed in urban environments [65].

Our contributions in this chapter are listed as follows:

- We develop an algorithm that identifies the best lengths for the RAW slot period by considering the number of assigned stations and the maximum allowed duration. It allows for a smart selection of optimal RAW slot lengths and avoids durations that degrade its performance.

- We propose a resource-efficient scheme to optimally reconfigure the RAW in order to maximize its gain in terms of throughput and energy efficiency. This scheme eliminates the extra time wasted in the RAW period, which improves channel usage.

The remainder of this chapter is organized as follows. Section 6.2 introduces the system model considered. Then Section 6.3 presents the proposed algorithm to efficiently allocate RAW slots. Section 6.4 evaluates the throughput and energy efficiency of a RAW slot. Therefore, Section 6.5 established the proposed configuration scheme for the RAW. Section 6.6 presents some results with discussions. Finally, Section 6.7 concludes our work.

6.2 System Model

We consider an IEEE 802.11ah-based wireless network with saturated traffic conditions operating on a Rayleigh fading channel with the capture effect enabled in the AP. The list of parameters used in our analytical framework is presented in Table 6.1. The AP broadcasts the configuration of the RAW within the beacon frame at the beginning of each beacon interval (BI) [4]. The RAW is configured in terms of its duration, assigned stations, and the number of RAW slots. The stations are distributed among the RAW slots following a round-robin rule [65]. The channel access is then limited during each RAW slot period to only its designated stations.

The last ongoing transmission within the RAW period can either be allowed to cross the RAW slot boundary or not [4]. Allowing the boundary-crossing results in a waste of energy for stations in the subsequent RAW slot, as stations will be awake without being able to compete for channel access until the transmission from the previous RAW slot is completed. Similarly, subsequent RAW slots are subject to the same impact. Henceforth, for fairness reasons among all stations in the RAW, we consider the case of disabled boundary-crossing of the RAW slot. The channel timeline during the RAW slot period is illustrated in Figure 6.1. In order to prevent transmissions from crossing the RAW slot boundary, the RAW slot duration is split into two parts: A free access period T_F during which designated stations can compete to access the channel, and a holding period T_H during which the stations stop the contention and only transmissions initiated before T_F are allowed to be completed. Hence, T_H should be set to the duration of a complete transmission. Thus, we define $T_H = DIFS + T_{\text{TXOP}}$, with DIFS being the duration of a DIFS time. T_{TXOP} is the duration of the TXOP that a station obtains after having access to the channel.

6.3 Adaptive RAW slot allocation

The RAW mechanism allows the AP to reserve the channel to a specific number of stations for a given period. Hereafter, maximizing the gain of the RAW period and avoiding channel

Table 6.1: List of the model parameters

Parameter	**Description**
DIFS	the duration of a DIFS interval
eb_{tx}, eb_l	the energy per bit for a station when transmitting and listening to the medium
e_{tx}, e_l, e_{idle}	the energy consumed by a station when transmitting a packet, listening to a busy slot, and listening to an idle slot
L	the payload size
m	the retry limit
N	the number of stations in a RAW slot
N_R	the number of stations within a RAW
K	the number of RAW slots within a RAW
p	the probability of packet failed delivery
p^p_{cap}	the packet capture probability
p^p_{col}	the packet collision probability
P_{cap}	the probability that a randomly chosen slot contains a captured packet
P_i	the probability that a randomly chosen slot is idle
T_{PLCP}	the duration of the PLCP header
P_s	the probability that a randomly chosen slot contains a successful transmission
$p_{cap}(z, n+1)$	the ACCP for $n+1$ interfering packets
r_i	the distance of the station i from the AP
R_n	the probability that a transmitted packet encounters n interfering packets
SIFS	the duration of a SIFS interval
T_{ACK}	the transmission time of an ACK frame
T_{DATA}	the transmission time of a data frame
T_F	the free access period in a RAW slot
T_H	the holding period in a RAW slot
T_R	the duration of a RAW
T_S	the duration of a pre-allocated RAW slot
$T_S^{\max}$	the maximum length of a RAW slot
T_{TXOP}	the duration of one TXOP
W_0	the minimum value of the CW
X	the number of idle slots before a transmission
z	the capture threshold
β	the length of a busy slot
ρ	the range of the AP
σ	the length of an idle slot
τ	the probability that a station initiates a transmission

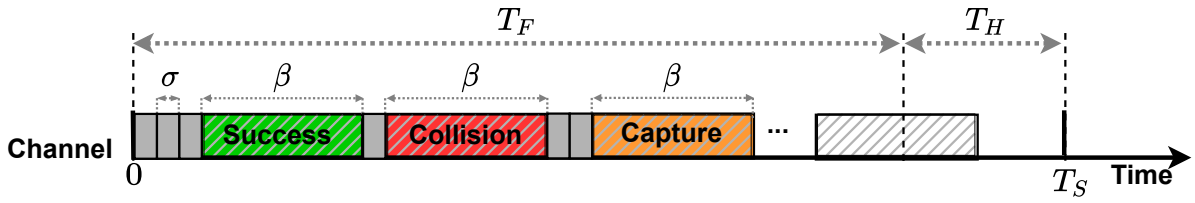


Figure 6.1: Channel timeline during the RAW slot period.

waste time is essential to maximize throughput and minimize energy consumption. As shown in our previous work [65], the performance of a RAW slot is very sensitive to its duration, which produces a fluctuating behavior. This is due to the holding period that might be utilized by transmission or wasted. Therefore, in this paper, we propose a fast scheme to identify only the RAW slot lengths that enhance its performance and skip the cases that decrease its gain.

Let $T_S^{\max}$ be the maximum duration of a RAW slot allocated to N stations. We aim to derive the set $V(N, T_S^{\max})$, which contains all lengths of the RAW slot that maximize its gain within the interval $[0, T_S^{\max}]$. According to the standard [4], there are two scenarios for the RAW slot limits: (1) If the RAW consists of up to 8 RAW slots, then $T_S^{\max} = 246.14\,\text{ms}$, whereas (2) in the case of a RAW consisting of up to 64 RAW slots, then each is limited to $T_S^{\max} = 31.1\,\text{ms}$.

We first need to evaluate the performance of a pre-configured RAW slot in terms of its duration and assigned stations. We consider a tagged RAW slot of length T_S, reserved for N stations. Only the N stations are allowed to compete for channel access during the period T_S.

6.3.1 Channel Access

During the RAW slot interval, assigned stations compete for channel access using the EDCA protocol [95]. When the backoff counter state reaches zero, the station obtains a TXOP to initiate the transmission. We assume that only one packet can be transmitted during TXOP, which has the length $T_{\text{TXOP}} = T_{\text{DATA}} + \text{SIFS} + T_{\text{ACK}}$. The lengths of the transmission of a data frame and an ACK frame, respectively, are given by

$$T_{\text{DATA}} = T_{\text{PLCP}} + \frac{\text{Payload} + \text{MacHeader}}{\text{DataRate}}, \tag{6.1}$$

$$T_{\text{ACK}} = T_{\text{PLCP}} + \frac{\text{ACK}}{\text{DataRate}}, \tag{6.2}$$

where T_{PLCP} is the duration of the PLCP header.

Within the RAW slot, every station attempts to transmit its packet in a random virtual slot with probability τ expressed as follows [65]:

$$\tau_k = \frac{2(2p-1)\left(1+(m+1)p^{m+2}-(m+2)p^{m+1}\right)}{W_0(1-p)\left(\left(2^{m+2}-1\right)p^{m+1}-\left(2^{m+2}-2\right)p^{m+2}-1\right)+2(2p-1)\left(1+(m+1)p^{m+2}-(m+2)p^{m+1}\right)}, \tag{6.3}$$

where W_0 is the minimum value of the CW size, and m is the retransmission limit. p is the probability that a transmitted packet faces a failed delivery given by

$$p = p_{col}^{p} \cdot (1 - p_{cap}^{p}), \tag{6.4}$$

where p_{col}^{p} and p_{cap}^{p} are, respectively, the probability that a transmitted packet faces a collision and the probability that a collided packet is captured.

We have

$$p_{col}^{p} = 1 - (1 - \tau)^{N-1}, \tag{6.5}$$

and

$$p_{cap}^{p} = \frac{\sum_{n=1}^{N-1} R_n \cdot p_{cap}(z, n+1)}{p_{col}^{p}}, \tag{6.6}$$

where R_n is the probability that a transmitted packet encounters n interfering packets, given by:

$$R_n = \binom{N-1}{n} \tau^n (1 - \tau)^{N-1-n}. \tag{6.7}$$

$p_{cap}(z, n+1)$ denotes the ACCP for $n+1$ colliding packets and is expressed as follows [65]:

$$p_{cap}(z, n+1) = \int_0^{\rho} \left(1 - \frac{r_0^2 \sqrt{z}}{\rho^2} \cdot \arctan\left(\frac{\rho^2}{r_0^2 \sqrt{z}}\right)\right)^n \cdot \frac{2 r_0}{\rho^2} \, dr_0, \tag{6.8}$$

with ρ being the radius of the circular area around the AP where the stations are randomly distributed.

Hereafter, the probabilities τ, p, and p_{cap}^{p} and they can be obtained by resolving numerically (e.g., using Newton's method [78]) the non-linear system defined by (6.3), (6.4) and (6.6).

During the RAW slot period, a virtual slot can be in one of the following four states:

- Idle with probability:

$$P_i = (1 - \tau)^N. \tag{6.9}$$

- Contain a successfully delivered packet with probability:

$$P_s = N \tau (1 - \tau)^{N-1}. \tag{6.10}$$

- Contain a captured packet from a collision with probability:

$$P_{cap} = N \cdot \tau \cdot p_{cap}^{p}. \tag{6.11}$$

- Contain a collision with no capture (i.e., failed delivery), with probability: $P_f = 1 - P_i - P_s - P_{cap}$.

6.3.2 Contention Results

Given the scenario without boundary-crossing transmissions within RAW slots, all stations cease counting down their backoff at the instant T_F. However, if a transmission is initiated before T_F, it can continue within the holding period T_H. Therefore, the maximum number of transmissions that can be initiated and completed within the RAW slot period is given by

$$\Gamma_b = \left\lfloor \frac{T_F}{\beta} \right\rfloor + \mathbb{1}_{\left\{T_F > \left\lfloor \frac{T_F}{\beta} \right\rfloor \cdot \beta + \sigma\right\}}, \tag{6.12}$$

where σ and $\beta = T_{\text{TXOP}} + DIFS$ are the lengths of an idle and a busy slot, respectively.

Hereafter, the expected number of transmissions that occurred during the RAW slot period is expressed as follows [65]:

$$\mathrm{E}[N_{T_S}] = \sum_{k=1}^{\Gamma_b} \sum_{j=0}^{\left\lfloor \frac{T_F-(k-1)\cdot\beta}{\sigma} \right\rfloor} \binom{j+k-1}{j} (1-P_i)^k P_i^j . \tag{6.13}$$

Additionally, the expected number of idle slots within the RAW slot is given by:

$$\mathrm{E}[I_{T_S}] = \frac{P_i}{1-P_i} \cdot \mathrm{E}[N_{T_S}] . \tag{6.14}$$

Since a RAW slot duration is usually very short (up to 31.1 ms or 246.14 ms), as mentioned above, its performance is mainly affected by the usage or waste of channel resources during the holding period T_H. Thus, we can define the holding period's usage ratio as its portion occupied by transmissions. Therefore, for a RAW slot T_S allocated for N stations, the usage ratio for the holding period is expressed as follows:

$$U_h(T_S) = \frac{E[I_{T_S}] \cdot \sigma + \mathrm{E}[N_{T_S}] \cdot \beta - T_F}{T_H} . \tag{6.15}$$

6.3.3 Adaptive RAW Slot Allocation

Increasing the duration of the RAW slot is not necessarily beneficial. When the added time is not enough to contain an additional transmission, it will waste resources and degrade the gain of the RAW slot. That is directly associated with the usage ratio of the holding period.

We present in Algorithm 6.1 our proposal to derive the adaptive lengths of the RAW slot, which maximizes its performance. The algorithm takes as input the number of designated stations N and the radius ρ of the circular area where they are located, the maximum allowed length for the RAW slot $T_S^{\max}$, the payload size L, and the capture threshold z.

Algorithm 6.1 Adaptive RAW slot allocation

Input: N: Number of stations.

$T_S^{\max}$: Maximum length of the RAW slot.

L: Payload size.

ρ: Range of stations distribution.

z: Capture threshold.

Output: $V(N, T_S^{\max})$: Adaptive lengths for the RAW slot.

1: Derive τ from the system defined by (6.3), (6.4) and (6.6).
2: Calculate β for the payload size L.
3: $\Delta \leftarrow \frac{P_i}{1-P_i} \cdot \sigma + \beta$
4: $\Delta_0 \leftarrow \Delta - \Delta \mod \sigma$
5: $\Delta_1 \leftarrow \Delta_1 + \sigma$
6: Compute $U_h(\Delta_0)$ and $U_h(\Delta_1)$
7: **while** $U_h(\Delta_0) <= U_h(\Delta_1)$ **do**
8: $\quad U_h(\Delta_0) \leftarrow U_h(\Delta_1)$
9: $\quad \Delta_1 \leftarrow \Delta_1 + \sigma$
10: $\quad$ Compute $U_h(\Delta_1)$
11: **end while**
12: $\Delta_0 \leftarrow \Delta_1 - \sigma$
13: $V(N, T_S^{\max}) \leftarrow \left\{ \Delta_0 + k\Delta : k \leq \frac{T_S^{\max} - \Delta_0}{\Delta} \right\}$
14: **return** $V(N, T_S^{\max})$

First, we derive the transmission probability τ by numerically solving the non-linear system defined by (6.3), (6.4) and (6.6). Then we calculate the duration of a busy slot $\beta = T_{\text{TXOP}} + DIFS$ in terms of the considered payload L, using T_{DATA} in (6.1) which is included in the definition of T_{TXOP}. Hereafter, we derive the average duration required to successfully complete one transmission. That is, the length of both the transmission and the preceding idle slots during the backoff countdown. Hence, the average time required to obtain a complete transmission is defined as follows:

$$\Delta = \mathrm{E}[X] \cdot \sigma + \beta \,, \tag{6.16}$$

where $\mathrm{E}[X]$ is the expected number of idle slots before a busy slot.

A random virtual slot within the channel timeline is idle with probability P_i and busy when at least one station transmits with probability $1 - P_i$. Hence, with X being the random variable representing the number of idle slots that occurred before a transmission is initiated, we have X is geometrically distributed on $\mathbb{N}$ with parameter $1 - Pi$ as follows:

$$\Pr\{X = j\} = (1 - P_i) P_i^j \,, \quad j \geq 0 \,. \tag{6.17}$$

Thus, $\mathrm{E}[X]$ is given by

$$\mathrm{E}[X] = \frac{P_i}{1-P_i}. \tag{6.18}$$

Therefore, we have

$$\Delta = \frac{P_i}{1-P_i} \cdot \sigma + \beta\,, \tag{6.19}$$

The duration Δ can also be seen as the time between two consecutive transmissions. However, the first transmission requires more time since it occurs by the first attempt of at least one station. That is, the transmission attempt probability τ may not be accurate yet, as it is based on the $m+1$ transmission attempts, while the first transmission occurs through only the first attempt. Noting by Δ_0 the time for the first transmission, we initiate its value by Δ, then synchronize it with the channel timeline in terms of idle slots (lines 3-4 in the algorithm). Therefore, we evaluate the holding period usage ratio $U_h(.)$ for a RAW slot of duration Δ_0, and for larger lengths with a step of one idle slot σ until reaching a peak for $U_h(.)$ (lines 5-11). Δ_0 is then updated by the latest value maximizing the holding period usage ratio, i.e., $\Delta_1 - \sigma$. This loop takes very few iterations with a maximum of W_0 iterations. Finally, considering the length of the first transmission Δ_0 and the time required for each additional transmission Δ, we define the set of adaptive lengths for the RAW slot as follows:

$$V(N, T_S^{\max}) = \left\{ \Delta_0 + k\Delta : k \leq \frac{T_S^{\max} - \Delta_0}{\Delta} \right\} \tag{6.20}$$

6.4 RAW Slot Performance

We define the throughput of a RAW slot by the portion of its period that is used to successfully transmit data frames. The expected number of transmissions that occurred during the RAW slot period T_S is given by $\mathrm{E}[N_{T_S}]$ defined in (6.13), and a non-idle slot can contain a successful transmission and a captured packet with probability P_s and P_{cap}, respectively. Thus, we express the RAW slot throughput as follows:

$$Th_S(N, T_S) = \frac{\mathrm{E}[N_{T_S}] \cdot T_{\mathrm{DATA}}}{T_S} \cdot \frac{P_s + P_{cap}}{1-P_i}. \tag{6.21}$$

Let eb_{tx} and eb_l be the energy per bit consumed by a single station when transmitting and listening to the medium, respectively. Therefore, during a busy slot, the transmitting station consumes $e_{tx} = \beta \cdot DataRate \cdot eb_{tx}$. A station listening to a busy slot consumes $e_l = \beta \cdot DataRate \cdot eb_l$. During an idle slot, each station wastes $e_{idle} = \sigma \cdot DataRate \cdot eb_l$.

When a random virtual slot contains a successful transmission, one station transmits while the remaining listens to a busy slot. Thus, during the RAW slot period, the energy consumed by

successful transmissions is given by

$$E_s^{T_S} = \mathrm{E}[N_{T_S}] \cdot \frac{P_s}{1-P_i} \cdot (e_{tx} + (N-1) \cdot e_l) \,. \tag{6.22}$$

During a collision, at least two stations transmit with probability τ while the remaining listen to the medium. A collision may result in a captured packet with probability P_{cap} or a complete failed delivery for all colliding packets with probability P_f. Therefore, the energy consumed by collisions with capture and collisions with failure within a RAW slot of length T_S is given by $E_{cap}^{T_S}$ and $E_c^{T_S}$, respectively, expressed as follows:

$$E_c^{T_S} = \frac{\mathrm{E}[N_{T_S}] \cdot P_c}{1-P_i} \cdot \sum_{j=2}^{N} \binom{N}{j} \tau^j (1-\tau)^{N-j} (j \cdot e_{tx} + (N-j) \cdot e_l) \,, \tag{6.23}$$

$$E_{cap}^{T_S} = \frac{\mathrm{E}[N_{T_S}] \cdot P_{cap}}{1-P_i} \cdot \sum_{j=2}^{N} \binom{N}{j} \tau^j (1-\tau)^{N-j} (j \cdot e_{tx} + (N-j) \cdot e_l) \,. \tag{6.24}$$

In addition to the energy consumed during transmissions, the remaining time in the RAW slot is occupied by idle slots and the unused part of the holding period. Hence, the energy consumed during idle listening to the medium is expressed as follows:

$$E_i^{T_S} = (\mathrm{E}[I_{T_S}] + (1 - U_h(T_S)) \cdot T_H) \cdot N \cdot e_{idle} \,. \tag{6.25}$$

Therefore, we define the energy efficiency performance of a RAW slot as the ratio of useful energy consumed by successful transmissions (through single transmission or capture from interference) to the total energy consumption during the period T_S. We have

$$E_{eff}^{T_S} = \frac{E_s^{T_S} + E_{cap}^{T_S}}{E_s^{T_S} + E_{cap}^{T_S} + E_c^{T_S} + E_i^{T_S}} \,. \tag{6.26}$$

6.5 Resource-Efficient RAW Configuration

The RAW consists of one or more RAW slots and is limited to 64 groups according to the standard [4]. We consider a RAW of length T_R with N_R assigned stations. When the AP decides to configure the RAW to be decomposed into K RAW slots, each will have a duration of $T_S = T_R/K$. The stations will be assigned to the RAW slot following a round-robin rule. Eventually, the RAW structure will consist of $K_1 = K - (N_R \mod K)$ RAW slots with $N_1 = \lfloor N_R/K \rfloor$ stations and $K_2 = K - K_1$ RAW slots allocated to $N_2 = N_1 + 1$ stations. We note this RAW configuration by $\{(K_1, N_1, T_S), (K_2, N_2, T_S)\}$.

Due to the fluctuating behavior of each RAW slot performance, the overall RAW per-

Algorithm 6.2 Resource-efficient RAW configuration

Input: N_R: Number of stations
T_R: Length of the RAW
K: Number of RAW slots
L: Payload size.
ρ: Range of stations distribution.
z: Capture threshold.

Output: RAW_Config: RAW configuration

1: $K_1 \leftarrow K - (N_R \mod K)$
2: $K_2 \leftarrow K - K1$
3: $N_1 \leftarrow \lfloor N_R/K \rfloor$
4: $N_2 \leftarrow N_1 + 1$
5: $T_S \leftarrow T_R/K$
6: $V1 \leftarrow V(N_1, T_S)$ {use Algorithm 6.1 for N_1, T_S, L, ρ, z}
7: $T_S^1 \leftarrow V1(end)$
8: $V2 \leftarrow V(N_2, T_S)$ {use Algorithm 6.1 for N_2, T_S, L, ρ, z}
9: $T_S^2 \leftarrow V2(end)$
10: **return** $RAW\_Config = \{(K_1, N_1, T_S^1), (K_2, N_2, T_S^2)\}$

formance will be affected accordingly. Hence, we deploy our proposal in Algorithm 6.1 to optimize channel resources in terms of the RAW duration. We present in Algorithm 6.2 our proposed scheme to derive an efficient RAW configuration. We aim to alter the original RAW configuration to minimize the channel allocation and maximize the network performance during the RAW period. Therefore, we optimize each RAW slot's duration by replacing it with the last element of the set of adaptive lengths obtained by Algorithm 6.1. Thus, it is more likely to shorten the RAW slot duration and enhance its performance.

We first define the standard RAW configuration $\{(K_1, N_1, T_S), (K_2, N_2, T_S)\}$ explained above (lines 1-5). Since the RAW contains two types of RAW slots with a different number of assigned stations, we optimize the allocated duration for each type. Hence, using Algorithm 6.1 we obtain K_1 RAW slots of duration T_S^1 which corresponds to the last element of $V(N_1, T_S)$ (lines 6-7), and K_2 RAW slots with length T_S^2 corresponding to the last element of $V(N_2, T_S)$ (lines 8-9). We therefore obtain the new resource-efficient RAW configuration noted as $RAW\_Config = \{(K_1, N_1, T_S^1), (K_2, N_2, T_S^2)\}$.

To study the effectiveness of our proposed configuration scheme, we evaluate the RAW performance in terms of throughput and energy efficiency. We define the RAW throughput by the ratio of time used to successfully deliver data frames within its allocated period T_R. Hence, the normalized throughput for a RAW consisting of K RAW slots is given by:

$$Th_R(K) = \frac{T_{\text{DATA}}}{T_R} \sum_{j=1}^{2} \left(K_j \cdot \mathrm{E}\left[N_{T_S^j} \right] \cdot \frac{P_s^{N_j} + P_{cap}^{N_j}}{1 - P_i^{N_j}} \right), \tag{6.27}$$

where $P_i^{N_j}$, $P_s^{N_j}$ and $P_{cap}^{N_j}$ are the probabilities defined in (6.9), (6.10) and (6.11), respectively, for a RAW slot with N_j stations.

The energy efficiency of a RAW consisting of K RAW slots is defined by the portion of the energy used for successful transmissions during the RAW period T_R. We have

$$E_{eff}(K) = \sum_{j=1}^{2} \frac{K_j(E_s^{Ts_j} + E_{cap}^{Ts_j})}{E_{total}}, \tag{6.28}$$

where E_{total} is the total energy consumption during the RAW, and it is given by

$$E_{total} = \sum_{j=1}^{2} K_j \left(E_s^{Ts_j} + E_{cap}^{Ts_j} + E_c^{Ts_j} + E_i^{Ts_j} \right). \tag{6.29}$$

The configuration scheme in Algorithm 6.2 redefines the lengths of RAW slots, and hence the new overall RAW duration is reduced to $T_R^* = K_1 \cdot T_S^1 + K_2 \cdot T_S^2$. We, therefore, evaluate the Channel Gain Ratio (CGR) , representing the percentage of time eliminated from the original RAW duration T_R. This reduction allows for optimizing channel resources by removing unused time within the RAW period. We define the CGR as follows:

$$CGR(N_R, T_R, K) = \frac{T_R - T_R^*}{T_R}. \tag{6.30}$$

6.6 Numerical Results and Discussion

This section analyzes and evaluates our proposals through numerical results. We validate our analytical findings via simulations using a discrete-event simulator we developed with MATLAB software. The simulator mimics the channel access using the RAW mechanism, where each simulation of a single RAW slot provides the contention results of assigned stations during the RAW slot period T_S. The final results are averaged over 1000 simulations. We consider sensor stations of category video or VoIP, which allows two transmission attempts for every packet according to the standard [4]. Table 6.2 lists the network configuration parameters.

6.6.1 Single RAW Slot

We consider a RAW slot allocated to N stations with a maximum length of $T_S^{\max} = 15\,\text{ms}$. We examine the RAW slot performance and the effectiveness of the adaptive selection of optimal lengths proposed in Algorithm 6.1. "ANA" and "SIM", respectively, represent the results obtained analytically and by simulations."Proposed" depicts the results obtained by Algorithm 6.1, and "Bianchi" represents the stationary (time-independent) results obtained by Bianchi's ap-

Table 6.2: Parameters

Parameter	**Value**	**Parameter**	**Value**
Data rate	1.95 Mbps	CW_{min}	8
MACHeader	272 bits	m	1
ACK	112 bits	Payload	512 bits
T_{PLCP}	80 µs	ρ	100 m
σ	52 µs	e_{tx}	2 mJ
SIFS	160 µs	e_{l}	1.3 mJ
DIFS	264 µs	e_{idle}	0.04 mJ

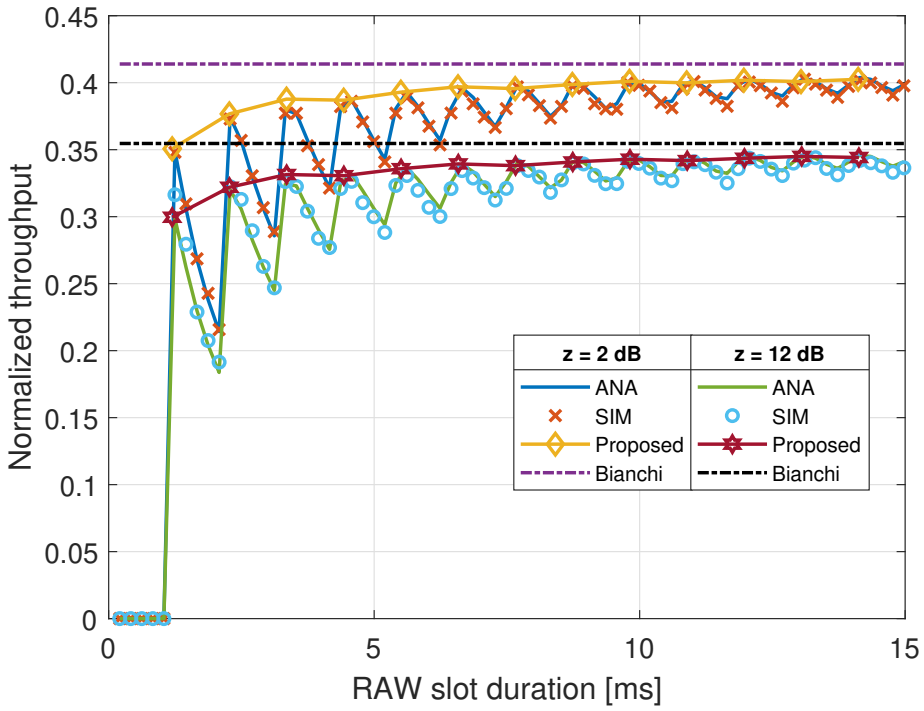


Figure 6.2: Normalized throughput of a RAW slot allocated to 5 stations

proach [5]. We present in Figure 6.2 the normalized throughput of this RAW slot while increasing its duration, with two scenarios of capture threshold z (2 dB and 12 dB). The results show a fluctuating behavior of throughput, which is related to the usage ratio of the holding period. When T_S increases, it enhances the performance in case the added time allows an additional complete transmission. Otherwise, it becomes a waste of channel time and hence degrades the RAW slot performance. Increasing the capture threshold z lowers the ability to capture packets [65]. Thus, the scenario of 2 dB capture threshold provides higher throughput than the case of $z = 12$ dB. In both scenarios, our proposed approach in Algorithm 6.1 is dynamically adapting to the channel conditions, and it effectively selects adaptive lengths for the RAW slot. It determines the times that maximize the usage ratio of the holding period, which proves effective in matching the best cases of the RAW slot throughput.

Figure 6.3 presents the energy efficiency of a RAW slot with two cases of five and ten assigned stations, operating under a Rayleigh channel with capture threshold $z = 4\,\text{dB}$. The energy efficiency follows the same behavior as the throughput due to the impact of the holding period, as mentioned above. It also converges to Bianchi's result, which further validates our model. The case of ten stations yields lower performance than the case of five stations. That is due to the higher contention followed by more collisions within the RAW slot period, which increases energy consumption and decreases the number of successfully delivered packets. Since the RAW slot performance depends on the usage of the holding period and Algorithm 6.1 is based on this constraint, we see that our proposed scheme provides the best cases for the RAW slot durations, maximizing the energy efficiency throughout the time interval $[0, T_S^{\max}]$.

The analytical results depicted in Figure 6.2 and Figure 6.3 show a good match with the simulation results, which validates the accuracy of our analytical model. Since the performances of a RAW slot are subject to its duration, an accurate evaluation of such performances shall converge to the case of an infinite RAW slot, which can be represented by the stationary results given by Bianchi's model [5]. Hereafter, we see that both throughput in Figure 6.2 and energy efficiency in Figure 6.3 converge in time to the stationary results, which presents an additional validation for our model.

The results of Algorithm 6.1 are practical to make optimal decisions for channel allocation using the RAW mechanism. As shown in Figure 6.2 and Figure 6.3, the last maximal performance is not essentially corresponding to the maximum length of the RAW slot. Thus, better performance can be achieved within a lower period. Therefore, a better configuration for the entire RAW can be based on this approach, such as our scheme proposed in Algorithm 6.2.

6.6.2 Several RAW Slots

To analyze the effectiveness of the proposed resource-efficient RAW configuration scheme in Algorithm 6.2, we evaluate the performance of a RAW consisting of several RAW slots and compare the results of our proposed configuration with the default configuration given by the standard. We represent by "Standard" the results for default RAW configuration given by the standard. That is the uniform division of the RAW into RAW slots and round-robin assignment of stations to the RAW slots. "Proposed" depicts the results obtained after applying the configuration given by Algorithm 6.2.

Figure 6.4 presents the throughput RAW of length 80 ms, consisting of $K = 20$ RAW slots. We evaluate its performance when allocated to different numbers of stations (from 20 to 200) operating under a Rayleigh channel with the capture threshold $z = 4\,\text{dB}$. For both configurations, the throughput increases with the number of stations up to a certain extent, where it starts to degrade. This effect is due to the capture effect, where more stations result in more collisions that end up with capture. However, when the number of stations in each RAW slot increases,

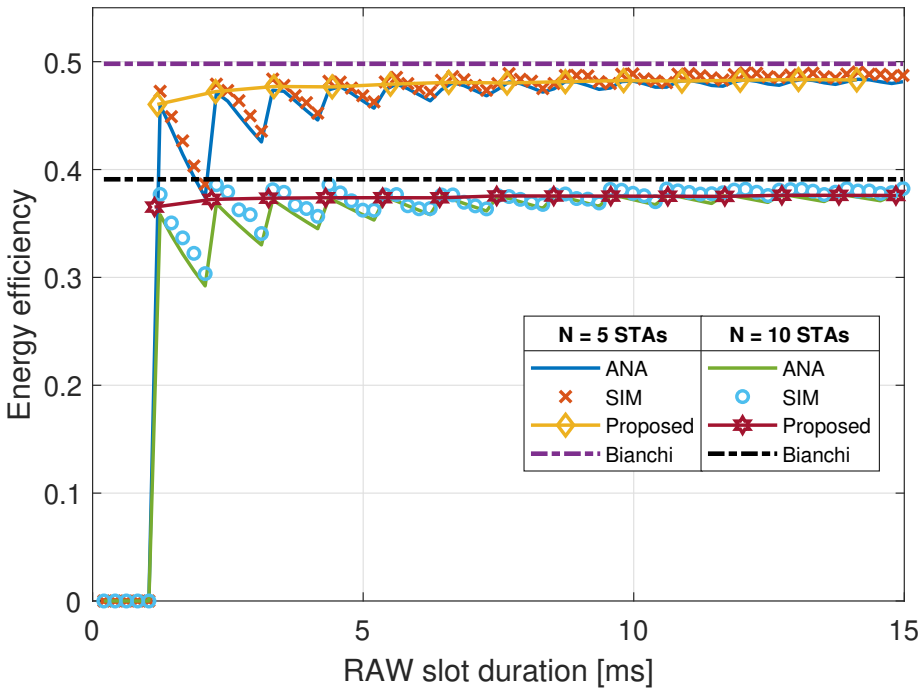


Figure 6.3: Energy efficiency of a RAW slot allocated to N stations with capture threshold $z = 4\,\text{dB}$

a collision may contain a large number of interfering stations which results in failed delivery with no capture. Comparing the two configurations, we observe that our proposal provides higher throughput than the standard for all different numbers of assigned stations. The RAW throughput has increased by 16% and 21% for 100 and 200 stations, respectively. Additionally, the channel allocation resource is optimized by a CGR reaching 16% and 19% for 100 and 200 stations, respectively. This shows the effectiveness of the proposed scheme, not only in enhancing the RAW performance but also in providing significant CGR , which represents the channel wastage eliminated from the RAW period.

Figure 6.5 presents the energy efficiency in terms of capture threshold for a RAW of length 150 ms, consisting of 50 RAW slots. Considering 200 assigned stations, we observe that our proposed configuration scheme yields higher performance than the standard for all capture conditions, reaching more than 8% improvement. The RAW performance decreases when the capture threshold increases due to the fact that a higher value of z lowers the value of the ACCP in (6.8), which decreases the chances of capturing packets from collisions. Eventually, the RAW performance converges in terms of z to the case of a no-capture channel with ideal conditions.

6.7 Conclusion

This chapter considers an IEEE 802.11ah-based network operating under a Rayleigh fading channel with the capture effect. We presented efficient schemes that maximize the gain of the

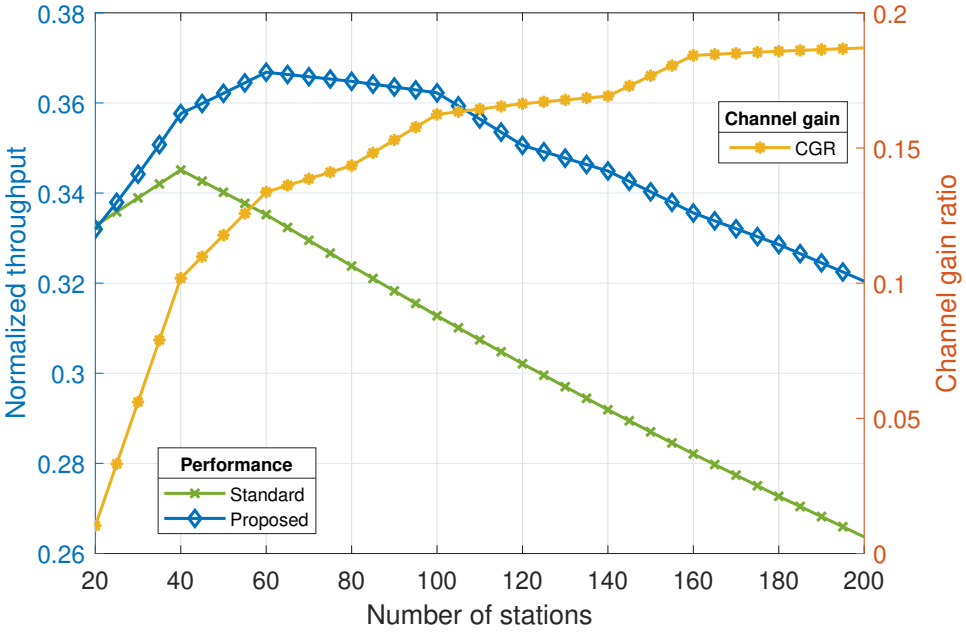


Figure 6.4: Normalized throughput in terms of assigned stations for an 80 ms RAW consisting of 20 RAW slots, with capture threshold $z = 4$ dB.

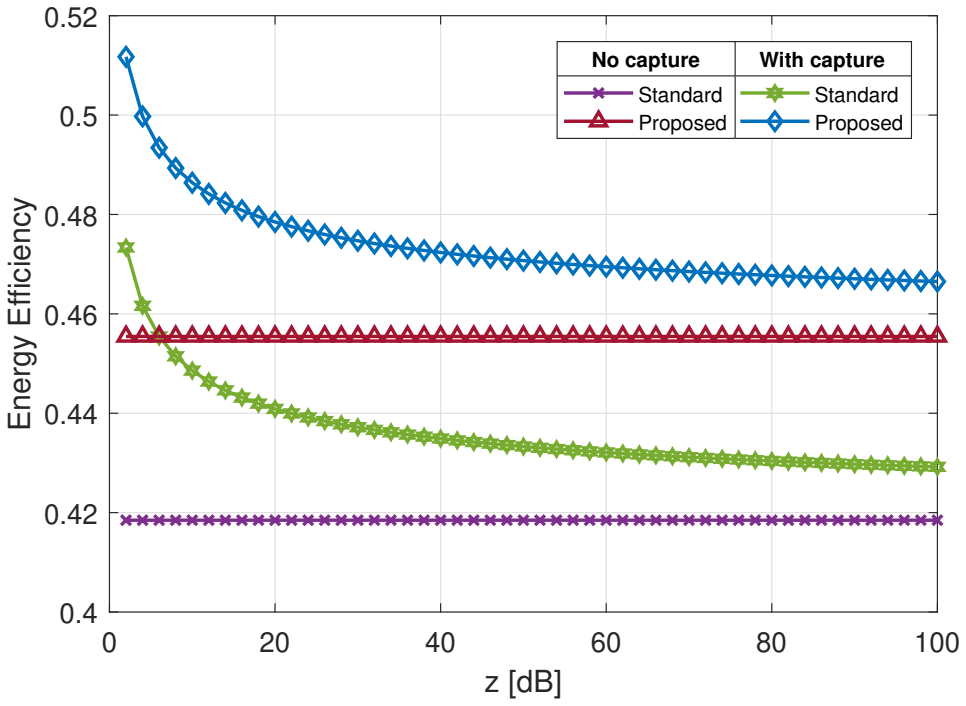


Figure 6.5: Energy efficiency in terms of capture threshold for an 150 ms RAW with 200 stations, consisting of 50 RAW slots.

RAW mechanism, the new channel access protocol introduced in IEEE 802.11ah. We considered a Rayleigh channel with capture effect, which is a realistic scenario for dense IoT networks deployed in urban environments. We first established the contention results within a RAW slot period and evaluated its performance. Then, we proposed an algorithm that selects the best choices for a RAW slot duration up to the maximum allowed value. This approach avoids channel waste caused by the misuse of the holding period and maximizes the RAW slot per-

formance in terms of throughput and energy efficiency. Based on this result, we proposed a resource-efficient RAW configuration scheme, maximizing the performance of a RAW consisting of several RAW slots. This scheme alters the channel time allocated to each RAW slot by replacing it with an efficient duration that is lower than the original allocated length. The new configuration eliminates the extra time wasted during the holding period, resulting in a higher performance of the RAW in terms of both throughput and energy efficiency. It also improves the channel gain ratio, representing the portion of wasted time that is removed from the original RAW duration. We established an improvement in RAW performance by up to 21% for throughput and up to 8% for energy efficiency. Furthermore, channel allocation resources were improved with a channel gain ratio of up to 19%. The algorithms proposed in this work are practical for managing large-scale networks in urban environments with Rayleigh channels.

In addition to being practical for an AP to provide effective configurations for the RAW mechanism, our proposals can also be deployed as an additional layer to optimize existing channel allocation scheduling schemes for IEEE 802.11ah and other alternative standards that implement periodic channel reservations.

Such scheduling schemes as our proposals in this chapter are effective in enhancing the network performance in terms of making the best use of available resources. However, new IoT applications, such as remote sensing and control, are prioritizing the freshness of received data more than system resources. Therefore, we address this kind of time-sensitive applications in the next chapter and develop a mathematical model to evaluate channel allocation. We then propose an efficient scheme to perform adaptive channel allocation to stations based on traffic load and network setup.

CHAPTER 7

Load-Aware Channel Allocation for Unsaturated IEEE 802.11ah-Based Networks

Contents

7.1 Introduction

Wi-Fi technology is one of the most widely deployed and used Internet access technologies in the world. However, the exponential growth of wireless devices tends to overcome this technology in its traditional form, as it is designed for small networks over short ranges. Therefore, the IEEE TGah developed an enhanced Wi-Fi technology, named IEEE 802.11ah standard [4], and marketed it as Wi-Fi HaLow. The IEEE 802.11ah standard comes with several amendments, providing connectivity to thousands of devices within a range of up to 1 km [3]. One

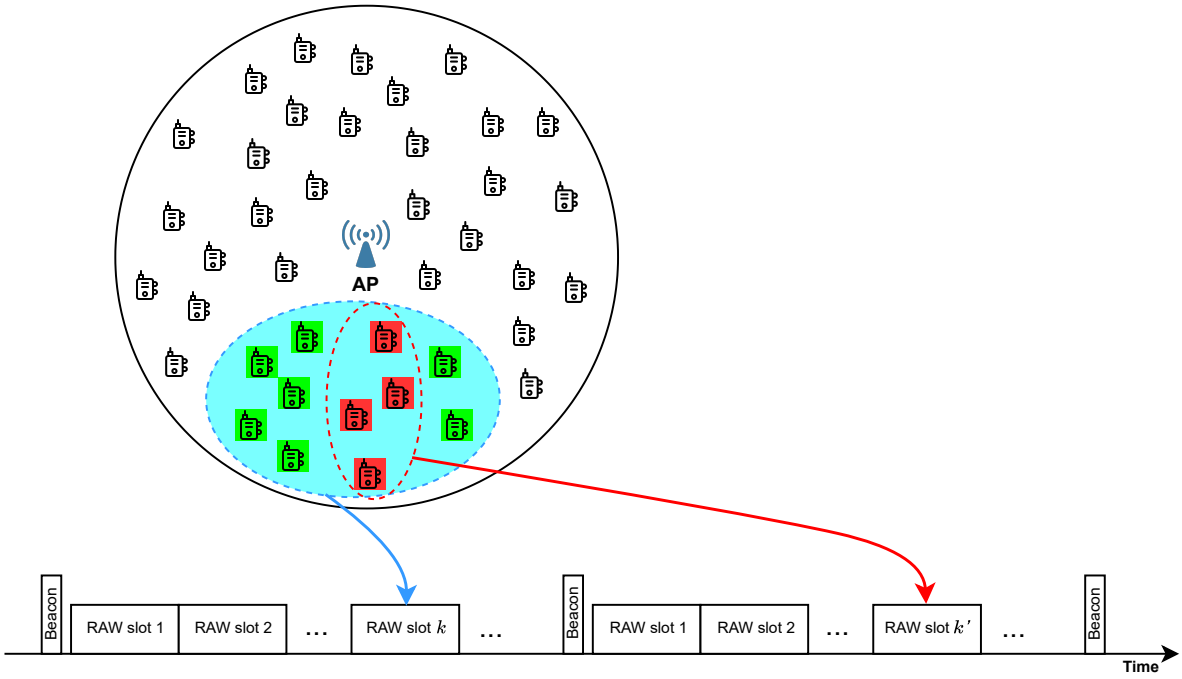


Figure 7.1: RAW-based channel access in the IEEE 802.11ah standard.

of its main features is a new channel access mechanism based on periodic channel reservations, called RAW. Only a small group of devices can contend for channel access during a given period called RAW slot, using the EDCA protocol.

The IEEE 802.11ah standard presents an excellent solution for large-scale IoT applications, consisting of dense networks with energy-constrained devices [95]. Its new RAW-based MAC support allows this standard to meet the requirements of IoT applications such as energy efficiency, coverage range, and scalability. Figure 7.1 presents an illustration of RAW-based channel access in the IEEE 802.11ah standard. The stations are distributed around the AP, and only a group of stations are allowed to contend for channel access during their allocated RAW slot. Suppose a station does not transmit its traffic load during the allocated RAW slot. In that case, it will have to wait until the next allocated RAW slot to try again within a new contention, hence increasing the age of the generated packet. This is the case depicted in the figure where the stations that are located within the highlighted area are assigned to the RAW slot k. This RAW slot period resulted in a successful data delivery for the green stations, whereas the red stations were unable to transmit their data. The latter will have to wait for access to the RAW slot k' in the next Beacon Interval (BI). Since the AP broadcasts the RAW configuration through the beacon frame, additional stations could be assigned to the same RAW slot k'. However, the standard does not provide schemes or strategies for configuring the RAW mechanism, dividing stations into groups and the channel timeline into RAW slots.

In order to derive efficient RAW configuration schemes, current works focus on distributing stations among pre-allocated RAW slots [3]. This approach enhances the network performance by eliminating hiding nodes and decreasing collisions in each RAW slot. Such approaches enhance channel usage during the RAW slot period but do not guarantee the successful

delivery of all packets carried by the designated stations. For time-sensitive applications, delivering updates as they are generated is more crucial to optimizing system resources [96, 97]. Therefore, instead of packet delay, the latency metric that gains priority in such applications is the Age of Information (AoI), which measures the freshness of information at the receiver [98]. That is the time since the last received update was generated. Such a metric is crucial in many systems, such as Unmanned Aerial Vehicle (UAV) networks, where exchanging the latest position, speed, and control information is vital and can contribute to significant performance enhancements [99].

The study of AoI is usually carried out using queueing theory, where the objective is to evaluate the average age and the peak age of information [98]. The authors in [97] considered a multiserver queuing system, where updates are generated following a renewal process. The system contains a controller that routes arriving packets to servers and decides the service distribution time and the service order of packets in each server. The authors carried out the age-delay and age-delay variance tradeoff and showed that the average age achieves its minimum when the packet delay tends to infinity. In [99], the authors considered a UAV-assisted wireless powered IoT system and established an optimization problem to minimize the average AoI of the data collected from sensor nodes on the ground by jointly optimizing the UAV's trajectory, energy transfer and data collection time for each sensor node. They showed that the AoI increases when the altitude of the UAV rises and that the average AoI increases linearly with the size of data frames.

Configuring the RAW mechanism based only on distributing stations among pre-allocated RAW slots does not guarantee timely communication of data [100], and may not be reliable for time-sensitive remote sensing and control applications. If the RAW slot is not long enough to allow full delivery of all stations' data, some stations will have to compete for channel access in subsequent RAW slots, which increases the AoI of data at the AP. However, if the RAW slot duration is longer than the time required to transmit the stations' traffic load, the extra time will result in channel waste. Therefore, to reduce the AoI at the AP upon successful delivery and prevent channel waste, we propose to allocate adaptive load-aware RAW slots, ensuring the successful delivery of all packets carried by designated stations.

Differently from configuring the RAW mechanism in terms of grouping stations and dividing the RAW duration into RAW slots, as we established in Chapter 4, we address in this chapter a load-aware channel allocation for the lengths of RAW slots. We assume that stations access the channel only through the RAW mechanism. That is, the channel time consists of successive RAW slots periods, and hence the performance of the entire network is proportional to the performance of these periods. Therefore, we consider a single RAW slot allocated to a group of stations. We evaluate the contention performance of the designated stations under a Rayleigh channel with capture, considering the stations' traffic load and their spatial distribution. Then, we develop a Load-Aware Channel Allocation (LACA) scheme for the RAW slot duration based on a two-level renewal process model. A load-aware RAW slot allows the

successful delivery of all packets of the assigned stations without extra waste of channel time. Such an approach is very beneficial for time-sensitive applications that require receiving data as they are generated. As illustrated in Figure 7.1, a load-aware allocation of the RAW slot k shall serve all assigned stations and avoid having red stations that need to wait for a subsequent allocated RAW slot in order to compete again for channel access.

In this work, we target time-sensitive applications where the RAW slot duration should meet the requirements of its designated stations. That is, to be long enough to deliver packets up to the desired ratio and not too long to avoid wasting channel time. Such an adaptive RAW slot allocation requires predicting the duration needed to serve the traffic load of the designated stations. Since a station may not have available data to transmit or might be in a sleep state, each group may have a different traffic load that requires an adaptive RAW slot duration, ensuring the successful delivery of all packets. To the best of our knowledge, estimating the RAW size in terms of the dynamic conditions of the network is not yet addressed in the literature [95]. Therefore, we develop an analytical framework based on a two-level renewal process to model the contention within one group of stations. We assume that each station has enough buffer for only one packet. That is the case of sensor stations where the previous data is disregarded when new data is generated. Considering the spatial distribution of stations operating under a Rayleigh-fading channel with capture effect enabled in the AP, we derive a LACA for a given group of stations, representing an efficient length of the RAW slot that ensures successful delivery of all packets without extra waste of channel time.

The contributions of this chapter are listed as follows:

- We develop an analytical model for channel access contention of a group of stations operating under a Rayleigh-fading channel. Then, we develop a two-level renewal process based model for the channel timeline.
- Based on the analytical framework, we propose a load-aware channel allocation algorithm for the duration of RAW slots, considering the spatial distribution of stations and the capture effect in the AP.
- We validate our analytical framework via extensive simulations and prove the effectiveness of the proposed algorithm through the metrics of packet delivery ratio and channel usage. We show that our proposal provides an adequate solution for channel allocation under Rayleigh fading conditions and capture effect, making it practical to ensure data delivery from the first contention attempt, which minimizes the AoI at the AP.

The remainder of this chapter is organized as follows. Section 7.2 explains the considered scenario and the addressed problem. Then Section 7.3 presents the proposed analytical model and the LACA algorithm. Section 7.4 validates our analytical findings via simulations and analyzes the impact of different parameters. Finally, Section 7.5 concludes our work.

7.2 Problem Statement

We consider an IEEE 802.11ah-based network where stations are distributed around the Access Point (AP). The stations periodically obtain access to the channel following the configuration of the RAW mechanism. Such a configuration is based on two primary operations: (1) divide stations into groups, and (2) allocate a certain amount of channel time to each group, named RAW slots. Knowing that the number of groups within each RAW interval cannot exceed 64, the standard suggests distributing stations following a round-robin rule. To the best of our knowledge, existing research works focus on evaluating and studying the assignment of stations into pre-defined RAW slots. Such approaches improve the gain of the RAW mechanism but do not guarantee that all stations will be able to get a chance to transmit their packets during their allocated RAW slot. Therefore, we propose in this work to define the duration of the RAW slot after knowing the conditions of its designated stations, including their traffic load and locations. Note that this procedure can be applied repeatedly to every group of stations and perform an adequate channel allocation to every RAW slot.

The channel time reserved for a given RAW slot highly depends on the number and state of the designated stations. As proven in our previous works [63, 64, 65], a slight change in the length of a RAW slot can either enhance or degrade the overall performance of the RAW slot. This is due to the holding period, which might be long enough to host an additional transmission or otherwise be wasted. Henceforth, the RAW slot throughput results in a fluctuating behavior when its duration varies. The fluctuation behavior is more significant for a short RAW slot, where the holding period represents a large part of the RAW slot interval. In contrast, it becomes less significant when the length of the RAW slot is longer.

Therefore, defining the channel duration allocated to each RAW slot is critical, especially in an unsaturated traffic condition. Thus, the duration of the RAW slot shall be efficiently allocated according to the traffic load of the designated stations. Note that the overall RAW duration is simply the total length of its composed RAW slots. Hence, in the rest of this chapter, we focus on evaluating only one RAW slot allocated to a group of stations. We consider a Rayleigh fading channel that models a dense urban environment where transmissions are subject to multipath fading and NLOS propagation in the wireless medium [92]. We assume that the buffer of a station can store only one packet. That is, a newly generated packet replaces the old one. This is practical in the case of remote sensing applications, where old packets are not informative [97] and only packets with fresh updates are considered by the AP for a better AoI performance [100].

We henceforth consider a group of N_S stations to be assigned to a RAW slot, where each station has a single packet to transmit. We aim to prevent having red stations in Figure 7.1 and to have all stations served in the first allocated RAW slot. Therefore, we shall accurately derive the RAW slot duration that is long enough to allow the successful delivery of all packets from

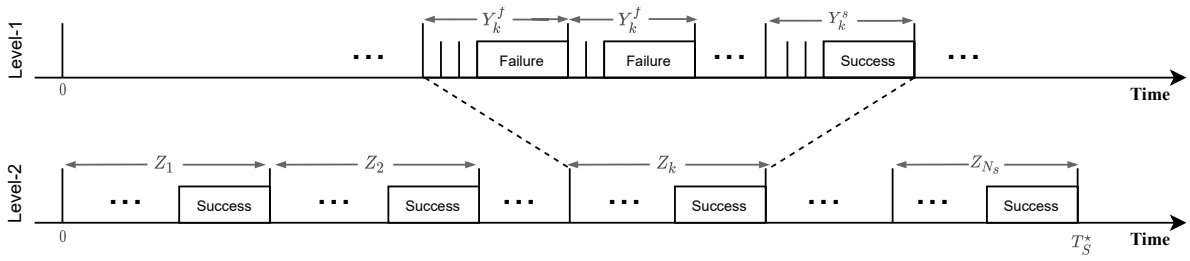


Figure 7.2: Illustration of the proposed two-level renewal process.

the N_S stations without extra channel time waste. We refer to this result as LACA time for the RAW slot period, denoted by T_S^*.

7.3 Proposed Model

During the contention of N_S stations, the channel timeline is occupied by idle and busy slots. Each busy slot occurs after a sequence of idle slots, where the stations count down their backoff counters. A good mathematical approach to model this behavior is the renewal theory [9], where stochastic processes update at random instants, during which the state of the system becomes probabilistically equivalent to the starting state. This approach effectively derives closed-form solutions for time-dependent events by implementing the counting process associated with the renewal process [65]. To derive the LACA time for a group of stations, we seek the average time required to transmit all packets successfully. Therefore, we model the channel timeline using a two-level renewal process, as illustrated in Figure 7.2. The list of the main model parameters is presented in Table 7.1.

The first-level process represents the time between consecutive transmission attempts, whether they result in a successful or failed packet delivery, whereas the second-level process depicts the interarrival times of successful transmissions. The second-level renewal cycle is defined by the period between two successive successful transmissions. We denote the level-2 renewal process by $\{Z_k\}_{k\geq 1}$, where Z_1 is the arrival instant of the first successful transmission and $Z_k, k \geq 2$ is the period between the $(k-1)$-th and the k-th successful transmissions.

For each process Z_k, we associate a first-level process $\{Y_k\}_{k\geq 1}$, representing the inter-arrival times of busy slots within the RAW slot channel time. The process $\{Y_k\}_{k\geq 1}$ has a binary outcome, where $Y_k = Y_k^f$ if the busy slot is a failure and $Y_k = Y_k^s$ if the busy slot contains a successful delivery of a packet. Note that a successful transmission can occur when a single station is transmitting or when the AP captures a packet during a collision.

Since each station has only one packet to transmit, the number of contending stations decrements by one after each successful packet transmission. That is, $N_S - k + 1$ stations compete for channel access during the period Z_k.

Table 7.1: List of the main model parameters

Parameter	**Description**
$\mathcal{C}_\ell^n$	the set of all n-combinations of the distances of stations interfering with the station STA_ℓ
$\left\lvert\mathcal{C}_\ell^n\right\rvert$	the cardinality of the set $\mathcal{C}_\ell^n$
$\text{CU}(T_S)$	the CU for a RAW slot of duration T_S
DIFS	the duration of a DIFS interval
L	the payload size
m	the retry limit
N_S	the number of stations in the RAW slot
p_k	the packet failed-delivery probability during the renewal cycle Z_k
$p_{k,cap}^p$	the packet capture probability during the renewal cycle Z_k
$p_{k,col}^p$	the packet collision probability during the renewal cycle Z_k
$P_{k,cap}$	the probability that a randomly chosen slot within Z_k contains a captured packet
$P_{k,i}$	the probability that a randomly chosen slot within Z_k is idle
$P_{k,s}$	the probability that a randomly chosen slot within Z_k contains a successful transmission
T_{PLCP}	the duration of the PLCP header
$\text{PDR}(T_S)$	the PDR for a RAW slot of duration T_S
$p_{cap}(z,n+1)$	the ACCP for $n+1$ interfering packets
r_i	the distance of the station i from the AP
$R_{k,n}$	the probability that a transmitted packet encounters n interfering packets
SIFS	the duration of a SIFS interval
T_{ACK}	the transmission time of an ACK frame
T_{DATA}	the transmission time of a data frame
T_S	the duration of a pre-allocated RAW slot
T_S^*	the LACA time
T_{TXOP}	the duration of one TXOP
W_0	the minimum value of the CW
X_k	the number of idle slots before a busy slot
Y_k	the level-2 renewal process (see Figure 7.2)
z	the capture threshold
Z_k	the level-1 renewal process (see Figure 7.2)
β	the length of a busy slot
γ	the signal-to-interference ratio
σ	the length of an idle slot
τ_k	the probability that a station initiate a transmission within Z_k
ω_ℓ	the instantaneous power of a packet received from a tagged station STA_ℓ
ω_{int}	the instantaneous joint power of interfering packets

Hereafter, we define the LACA time for the RAW slot T_S^* as the time required to deliver all packets from the N_S stations. Therefore, T_S^* coincides with the end instant of the period Z_{N_S}. Denoting by $\overline{Z_k}$, the average length of the renewal cycle Z_k, we define the LACA time for the RAW slot as follows:

$$T_S^* = \sum_{k=1}^{N_S} \overline{Z_k} \tag{7.1}$$

7.3.1 Channel Access

After each successful transmission, the corresponding station goes to sleep mode and leaves the contention. Thus, the level of contention decreases, resulting in different channel access probabilities within each cycle Z_k. Hence, to determine the average lengths of the renewal cycles of the process $\{Z_k\}_{k\geq 1}$, we need to evaluate the channel access within each cycle.

The renewal cycle Z_k contains $N_S - k + 1$ stations, competing for channel access using the EDCA protocol [4]. Each station obtains a TXOP to initiate the transmission of its packets when its backoff time counter reaches zero [65]. We assume that a TXOP allows transmitting only one single packet, and its duration is given by:

$$T_{\text{TXOP}} = T_{\text{DATA}} + \text{SIFS} + T_{\text{ACK}}\,, \tag{7.2}$$

where T_{ACK} is the transmission time of an ACK frame. SIFS is the time interval between the transmission of a data frame and its corresponding ACK, defined in the standard [4]. Denoting by L the payload size, the transmission time of a data frame is defined as follows:

$$T_{\text{DATA}} = T_{\text{PLCP}} + \frac{L + \text{MacHeader}}{\text{DataRate}}\,, \tag{7.3}$$

where T_{PLCP} is the duration of the PLCP header.

For the convenience of analysis, we assume that all the N_S stations belong to the same AC and carry similar fixed-size payloads. Thus, within the renewal cycle Z_k, each station starts to transmit its packet in a randomly chosen slot with probability τ_k expressed as follows [65]:

$$\begin{aligned}\tau_k = &\left[2(2p_k-1)\left(1+(m+1)p_k^{m+2}-(m+2)p_k^{m+1}\right)\right] \Big/ \\ &\left[W_0(1-p_k)\left(\left(2^{m+2}-1\right)p_k^{m+1}-\left(2^{m+2}-2\right)p_k^{m+2}-1\right)\right. \\ &\left.+2(2p_k-1)\left(1+(m+1)p_k^{m+2}-(m+2)p_k^{m+1}\right)\right]\,,\end{aligned} \tag{7.4}$$

where p_k is the probability that the transmitted packet fails to be delivered successfully after encountering a collision and not being captured. W_0 is the starting value of the CW, and m is the retry limit.

The packet collision probability $p_{k,col}^p$ is the probability that at least one of the remaining

$N_S - k$ stations transmits in the same slot. Since each station transmits with probability τ_k, we have

$$p^p_{k,col} = 1 - (1 - \tau_k)^{N_S - k}. \tag{7.5}$$

A transmitted packet fails to be delivered when it encounters a collision and not being captured. Hence, the probability p_k of failed delivery of a transmitted packet is given by:

$$p_k = p^p_{k,col} \cdot (1 - p^p_{k,cap}), \tag{7.6}$$

where $p^p_{k,cap}$ is the probability of capturing a transmitted packet after encountering a collision. A prior definition of $p^p_{k,cap}$ is expressed as follows [65]:

$$p^p_{k,cap} = \frac{\sum_{n=1}^{N_S-k} R_{k,n} \cdot p_{cap}(z, n+1)}{p^p_{k,col}}, \tag{7.7}$$

where $R_{k,n}$ is the probability that the transmitted packet encounters n interfering packets and is given by:

$$R_{k,n} = \binom{N_S - k}{n} \tau^n (1 - \tau)^{N_S - k - n}. \tag{7.8}$$

$p_{cap}(z, n+1)$ is the Average Conditional Capture Probability (ACCP) for $n+1$ interfering packets. This probability depends on the received powers of the colliding packets. Denote by ω_ℓ the instantaneous power of a packet received from a tagged station STA_ℓ, located at a distance r_ℓ from the AP. Under the conditions of a Rayleigh fading channel, ω_ℓ is a random variable exponentially distributed as follows [6]:

$$f_{\omega_\ell}(x) = \frac{1}{\omega_{0\ell}} e^{-\frac{x}{\omega_{0\ell}}}, \quad x \geq 0 \tag{7.9}$$

where $\omega_{0\ell}$ is the local mean power of the transmitted packet at the receiver and is defined as:

$$\omega_{0\ell} = A \cdot r_\ell^{-\alpha} \cdot \omega_T, \tag{7.10}$$

where ω_T is the transmitted signal power, A is a dimensionless constant in the path-loss law, and α is the path-loss exponent. For the convenience of analysis, we assume that ω_T is constant for all stations.

Hereafter, the ACCP is defined by the probability that the SIR $\gamma = \frac{\omega_\ell}{\omega_{\text{int}}}$ exceeds a given threshold z, where $\omega_{\text{int}} = \sum_{i=1}^{n} \omega_k$ is the instantaneous joint power of interfering packets, and z represents the capture performance of the AP, and it is called the capture threshold. Hence, we have

$$p_{cap}(z, n+1) = \Pr\{\gamma > z | n+1\}. \tag{7.11}$$

The instantaneous power of a received packet at the AP depends on the location of the

source stations. Hence, the evaluation SIR γ requires the distances of colliding stations from the AP. Let $r_1, r_2, ..., r_n$ be the distances from the AP of the n interfering stations with STA_ℓ, respectively. The conditional capture probability for the packet of STA_ℓ can be expressed as follows:

$$\begin{aligned} p_{cap}(z, r_\ell, r_1, ..., r_n) &= \Pr\{\gamma > z | r_\ell, r_1, \ldots, r_n\} \\ &= \Pr\left\{ \omega_\ell > z \cdot \sum_{i=1}^{n} \omega_k | r_\ell, r_1, \ldots, r_n \right\} \end{aligned} \tag{7.12}$$

Therefore, by implementing the instantaneous power distribution of the received packets defined in (7.9), we obtain [65]:

$$p_{cap}(z, r_\ell, r_1, ..., r_n) = \prod_{i=1}^{n} \frac{1}{1 + z \cdot \left(\frac{r_i}{r_\ell}\right)^{-\alpha}}. \tag{7.13}$$

To derive the ACCPs for a colliding packet of STA_ℓ when encountering n interfering packets, we need to average over all possible combinations of potential interfering stations. Note by $\mathcal{C}_\ell^n$ the set of all n-combinations of the distances of the $N_S - k$ stations that are contending with the tagged stations STA_ℓ. Thus, the ACCPs for a packet of STA_ℓ when encountering n interfering packets is defined as follows:

$$p_{cap}(z, r_\ell, n) = \frac{1}{|\mathcal{C}_\ell^n|} \cdot \sum_{\{r_1, ..., r_n\} \in \mathcal{C}_\ell^n} p_{cap}(z, r_\ell, r_1, ..., r_n). \tag{7.14}$$

To obtain the ACCP for $n+1$ interfering packets within the period Z_k, we take the average of the ACCPs for all stations in contention. Denoting by $r_1, r_2, ..., r_{N_S-k+1}$ the distances from the AP for the $N_S - k + 1$ contending stations, we have the following:

$$p_{cap}(z, n+1) = \frac{1}{N_S - k + 1} \cdot \sum_{\ell=1}^{N_S-k+1} p_{cap}(z, r_\ell, n) \tag{7.15}$$

Consequently, substituting (7.8) and (7.15) in (7.7), we derive the expression of the probability $p_{k,cap}^p$. We have:

$$\begin{aligned} p_{k,cap}^p = \frac{1}{p_{k,col}^p} \sum_{n=1}^{N_S-k} \Bigg[\binom{N_S - k}{n} \tau_k^n (1 - \tau_k)^{N_S-k-n} \cdot \frac{1}{N_S - k + 1} \\ \cdot \sum_{\ell=1}^{N_S-k+1} \frac{1}{|\mathcal{C}_\ell^n|} \cdot \sum_{\{r_1, ..., r_n\} \in \mathcal{C}_\ell^n} \prod_{i=1}^{n} \frac{1}{1 + z \cdot \left(\frac{r_i}{r_\ell}\right)^{-\alpha}} \Bigg] \end{aligned} \tag{7.16}$$

Therefore, the values of probabilities τ_k, p_k, and $p_{k,cap}^p$ can be obtained by solving nu-

merically (e.g., using Newton's method [78]) the nonlinear system defined by (7.4), (7.6) and (7.16).

7.3.2 Load-Aware Channel Allocation

The probability $P_{k,i}$ for a randomly chosen slot within the period Z_k to be idle is given by the following:

$$P_{k,i} = (1-\tau_k)^{N_S-k+1}\,. \tag{7.17}$$

The probability $P_{k,s}$ that a randomly chosen slot within Z_k contains a successful transmission of a single packet without interference is given by:

$$P_{k,s} = \frac{(N_S-k+1)\cdot\tau_k\cdot(1-\tau_k)^{N_S-k}}{1-P_{k,i}}\,. \tag{7.18}$$

Let $P_{k,cap}$ be the probability that a randomly chosen slot within Z_k contains a successful packet delivery by capture. That is, when a collision occurs and a packet from one of the colliding stations is captured successfully. We have

$$P_{k,cap} = (N_S-k+1)\cdot\tau_k\cdot p^p_{k,cap}\,. \tag{7.19}$$

The renewal cycle Z_k is constructed by the cycles of the process $\{Y_k\}_{k\geq 1}$, each has an average length of $\overline{Y_k} = \sigma\cdot\overline{X_k}+\beta$, where X_k is the number of idle slots before at least one of the stations starts transmitting. σ is the length of an idle slot. $\beta = T_{\text{TXOP}} + \text{DIFS}$ is the duration of a busy slot, where DIFS is the duration of a DIFS time.

The random variable X_k follows a geometric distribution on $\mathbb{N}$ with parameter $(1-P_{k,i})$. We have

$$\Pr\{X_k = j\} = (1-P_{k,i})P^j_{k,i}\,,\quad j\geq 0\,. \tag{7.20}$$

Hence, the average number of idle slots before the occurrence of a busy slot is given by:

$$\overline{X_k} = \frac{P_{k,i}}{1-P_{k,i}}\,, \tag{7.21}$$

and the average length of a renewal cycle of the process $\{Y_k\}_{k\geq 1}$ is given by:

$$\overline{Y_k} = \sigma\cdot\frac{P_{k,i}}{1-P_{k,i}}+\beta \tag{7.22}$$

Let W_k be a random variable representing the number of renewal cycles of the process

$\{Y_k\}_{k\geq 1}$ within the period Z_k. Therefore, we have

$$Z_k = \sum_{n=1}^{W_k} Y_n \tag{7.23}$$

We have $\{Y_k\}_{k\geq 1}$ is a sequence of IID random variables with the same mean $\overline{Y_k}$ Additionally, the cycle Z_k ends after a period of $\{Y_k\}_{k\geq 1}$ of type Y_k^s and a packet can be successfully delivered through a single transmission or capture from a collision. Therefore, W_k is geometrically distributed on $\mathbb{N}^*$ with probability $P_{k,s} + P_{k,cap}$, and we have

$$\overline{W_k} = \frac{1}{P_{k,s} + P_{k,cap}} \tag{7.24}$$

Hereafter, $\overline{W_k} < \infty$ and according to Wald's equality [9], we have the average length of the cycle Z_k expressed as follows:

$$\begin{aligned} \overline{Z_k} &= \mathrm{E}\left[\sum_{n=1}^{W_k} Y_n\right] \\ &= \overline{W_k} \cdot \overline{Y_k} \\ &= \frac{1}{P_{k,s} + P_{k,cap}} \cdot \left(\sigma \cdot \frac{P_{k,i}}{1 - P_{k,i}} + \beta\right) \end{aligned} \tag{7.25}$$

Finally, by substituting (7.25) in (7.1), we obtain the LACA time for the RAW slot, allowing the successful delivery of all packets of the N_S station expressed as follows:

$$T_S^* = \sum_{k=1}^{N_S} \frac{1}{P_{k,s} + P_{k,cap}} \cdot \left(\sigma \cdot \frac{P_{k,i}}{1 - P_{k,i}} + \beta\right) \tag{7.26}$$

We summarize the proposed procedure for obtaining the LACA time T_S^* in Algorithm 7.1. The input parameters are the number of stations, their locations, the payload size, and the capture threshold. Since a station leaves the contention after delivering its packet, the channel access is evaluated in every renewal cycle of the process $\{Z_k\}_{k\geq 1}$ for $k = 1, \ldots, N_S$. Therefore, by evaluating the average length of renewal cycles of the process $\{Y_k\}_{k\geq 1}$ within each cycle Z_k, we obtain the average length of the cycle Z_k. Finally, the LACA time for the RAW slot in the considered scenario is derived as the aggregated length of renewal cycles of the process $\{Z_k\}_{k\geq 1}$.

Since the main objective of the RAW mechanism is to limit the number of stations accessing the channel simultaneously in order to decrease collisions, the number of stations N_S in the input of Algorithm 7.1 is supposed to be small, which lowers the computation complexity. Additionally, it is very likely to have several groups of stations with the same traffic characteristics, and the algorithm is then executed only once to derive the common LACA time T^* for

Algorithm 7.1 Load-Aware Channel Allocation

Input: N_S: Number of stations.

$r_1, ..., r_{N_S}$: Distances of stations from the AP.

z: Capture threshold.

L: Payload size.

Output: T_S^*: LACA time for the RAW slot.

1: **for** $k = N_S$ to 1 **do**

2: Derive τ_k, p_k and $p_{k,cap}^p$ using the non-linear system defined by (7.4), (7.6) and (7.16).

3: $P_{k,i} \leftarrow (1-\tau_k)^{N_S-k+1}$

4: $P_{k,s} \leftarrow \dfrac{(N_S-k+1)\cdot\tau_k\cdot(1-\tau_k)^{N_S-k}}{1-P_{k,i}}$

5: $P_{k,cap} \leftarrow (N_S-k+1)\cdot\tau_k\cdot p_{k,cap}^p$

6: **end for**

7: $T_S^* \leftarrow \sum_{k=1}^{N_S} \dfrac{1}{P_{k,s}+P_{k,cap}} \cdot \left(\sigma\cdot\dfrac{P_{k,i}}{1-P_{k,i}}+\beta\right)$

8: **return** T_S^*

all RAW slots allocated to these groups.

The LACA time for a given group of stations represents the channel time required to deliver all these stations' traffic load. We henceforth propose to investigate the impact of different duration on the network performance in terms of PDR and channel usage. Consider a pre-defined RAW slot of duration T_S allocated to N_S stations. We define the PDR for this RAW slot by the portion of delivered packets among the total stations' traffic load. Since the process $\{Z_k\}_{k\geq 1}$ starts a new cycle after each successful transmission, the number of delivered packets is given by the number of occurred renewal cycles of $\{Z_k\}_{k\geq 1}$. Thus, we have:

$$\mathrm{PDR}(T_S) = \begin{cases} \dfrac{\sum_{k=1}^{n^*}\overline{Z_k}}{T_S} + \dfrac{T_S - \sum_{k=1}^{n^*}\overline{Z_k}}{\overline{Z_{n^*+1}}} & \text{, if } T_S < T_S^* \\ 1 & \text{, otherwise} \end{cases} \tag{7.27}$$

where

$$n^* = \underset{n\in\mathbb{N}^*}{\arg\max}\left\{\sum_{k=1}^{n}\overline{Z_k} < T_S\right\}. \tag{7.28}$$

The channel usage during the RAW slot period is defined as the portion of the time used to successfully deliver data during T_S. Using the previous result of PDR, the number of delivered packets can be expressed by $\mathrm{PDR}(T_S)\cdot N_S$. Therefore, we define the channel usage for the considered RAW slot as follows:

$$\mathrm{CU}(T_S) = \frac{\mathrm{PDR}(T_S)\cdot N_S\cdot\beta}{T_S} \tag{7.29}$$

Table 7.2: Parameters

Parameter	**Value**
Data rate	1.95 Mbps
MacHeader	272 bits
T_{ACK}	1000 µs
T_{PLCP}	80 µs
σ	52 µs
SIFS	160 µs
DIFS	264 µs
CW_{min}	8
CW_{max}	16
m	1
α	4

7.4 Numerical Results and Discussion

This section validates our proposed model and analytical findings via extensive simulations obtained by a discrete-event simulator we developed with MATLAB. The simulator mimics the DCF mechanism, considering the location of stations, the channel fading, and the capture effect. The depicted simulation results represent the average outcome of 1000 simulations. We also evaluate and analyze the performance of a RAW slot in terms of different parameters and scenarios. The slight differences between the analytical and simulation results depicted below are due to the analytical assumption of the same transmission probability τ_k of stations in each renewal cycle of the process $\{Z_k\}_{k\geq 1}$. In the simulator, the first transmission attempt of each station is initiated after its backoff counter reaches zero during the first backoff stage. Whereas, τ_k considers the transmission attempts initiated from all backoff stages. Additionally, if a station successfully transmits its packet, it leaves the contention, and a new renewal cycle of the process $\{Z_k\}_{k\geq 1}$ starts. During this cycle, the remaining stations continue to contend for channel access with their current backoff counter states. This effect is not considered analytically as τ_k represents the transmission attempt probability for $N_S - k + 1$ contending stations without any inheritance of backoff counter states from the previous contention. We define the spatial distribution of N_S contending stations by an interval $[\rho_1, \rho_2]$, where the distances $r_1, r_2, ..., r_{N_S}$ of the stations from the AP are defined by N_S evenly spaced points in this interval. The latter is also referred to as the stations' area, with a range $\rho_2 - \rho_1$. We consider sensor stations transmitting video or VoIP data, with one retransmission attempt of a collided packet according to the standard [4]. The system setup parameters are presented in Table 7.2.

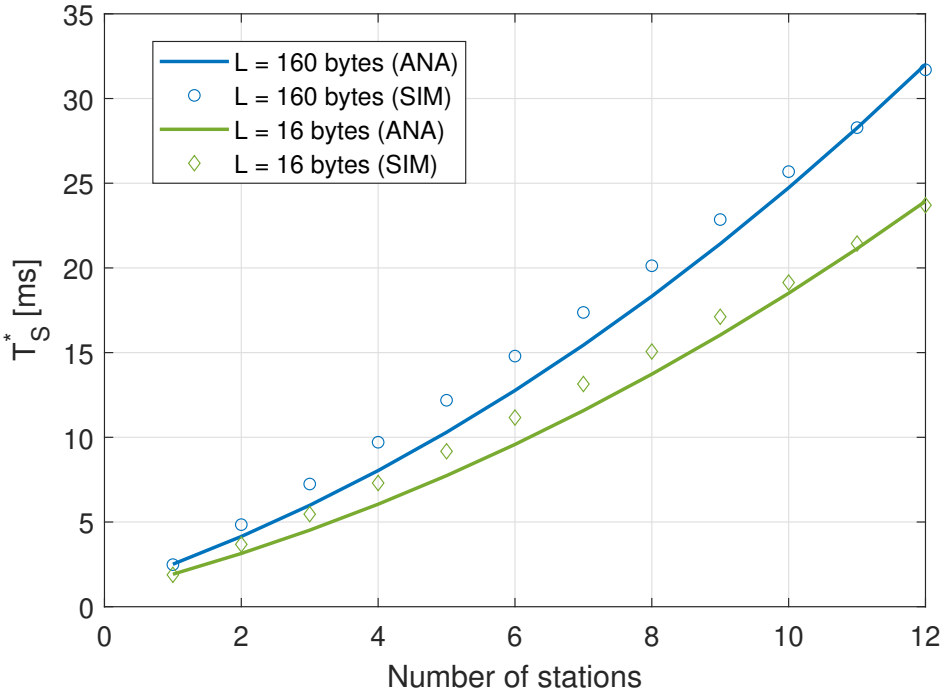


Figure 7.3: LACA time for stations located in the area of range $[1,10]$, with a capture threshold $z = 4\,\text{dB}$.

7.4.1 Load-Aware Channel Allocation

This section validates and analyzes our proposal, Algorithm 7.1. Figure 7.3 presents the LACA time for a RAW slot in terms of the number of assigned stations located in $[1,10]$. We observe an excellent match with the simulations, which validates the closed-form analytical result of T_S^* in (7.26). More stations in a group result in a higher traffic load, which requires more time to ensure the delivery of all packets. Moreover, extra stations in the RAW slot escalate the contention, which results in more collisions, and hence expands the time T_S^* needed to transmit all packets successfully. The figure shows the results for two payload sizes, $L = 16\,\text{Bytes}$ and $L = 160\,\text{Bytes}$. The case of stations with larger payloads requires more service time. However, the difference in T_S^* values between the two payloads is not very significant due to the considerable ACK frame length $T_{\text{ACK}} = 1\,\text{ms}$ that is introduced in the IEEE 802.11ah standard.

To analyze the impact of both payload size and the capture effect, we present in Figure 7.4 the results for T_S^* in terms of the capture threshold z for the two different payloads. We first observe that the simulations validate the analytical results. As mentioned previously, a more extended payload requires more channel access time to deliver the packets of all stations in the group successfully. Furthermore, a more significant value of the capture threshold z lowers the capture probability, as it becomes more unlikely for the received power of a given packet to exceed the joint received powers of the interfering packets by the threshold z. Therefore, the length of the first-level renewal cycle Y_k becomes longer, affecting the renewal cycles Z_k, and hence the value of T_S^* for the RAW slot increases. On the other hand, a lower value of

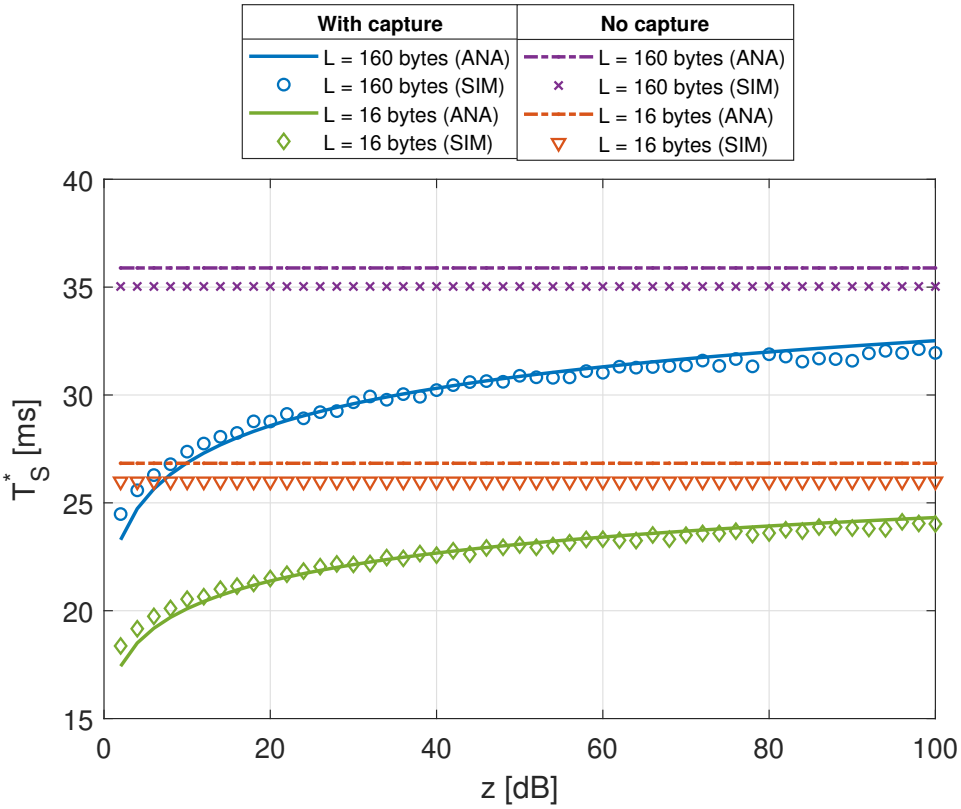


Figure 7.4: LACA time in terms of the capture threshold for stations located in the area of range $[1, 10]$.

the capture threshold z allows to capture packets more often, and hence making good usage of channel time. This allows for delivering all packets faster in a shorter period. Whereas a higher value of z has an inverse impact, resulting in a larger T_S^* and converging to the case of a no-capture channel.

7.4.2 RAW slot performance

We now investigate how to set up the RAW slot duration T_S according to the design objective and performance requirements of the end application. Towards this end, two performance metrics are of interest. First, we evaluate the PDR of a pre-allocated RAW slot, representing the portion of successfully delivered packets among the total packets of all stations assigned to the RAW slot. Let us consider a RAW slot with pre-defined duration $T_S = 10\,\text{ms}$, and evaluate the PDR in terms of the number of its designated stations, as depicted in Figure 7.5. By definition, if the RAW slot duration is less than T_S^*, the time allocated to the stations will not be sufficient to deliver all their traffic load (packets). Otherwise, when the length of the RAW slot is greater than T_S^*, all packets will be delivered before the RAW slot ends, and the remaining channel time will be wasted. Furthermore, the larger payload $L = 160\,\text{Bytes}$ results in a lower PDR than the shorter payload of $L = 16\,\text{Bytes}$. However, the difference is not significant due to the higher overhead of the ACK frame and the inter-frame spaces defined in the standard [4].

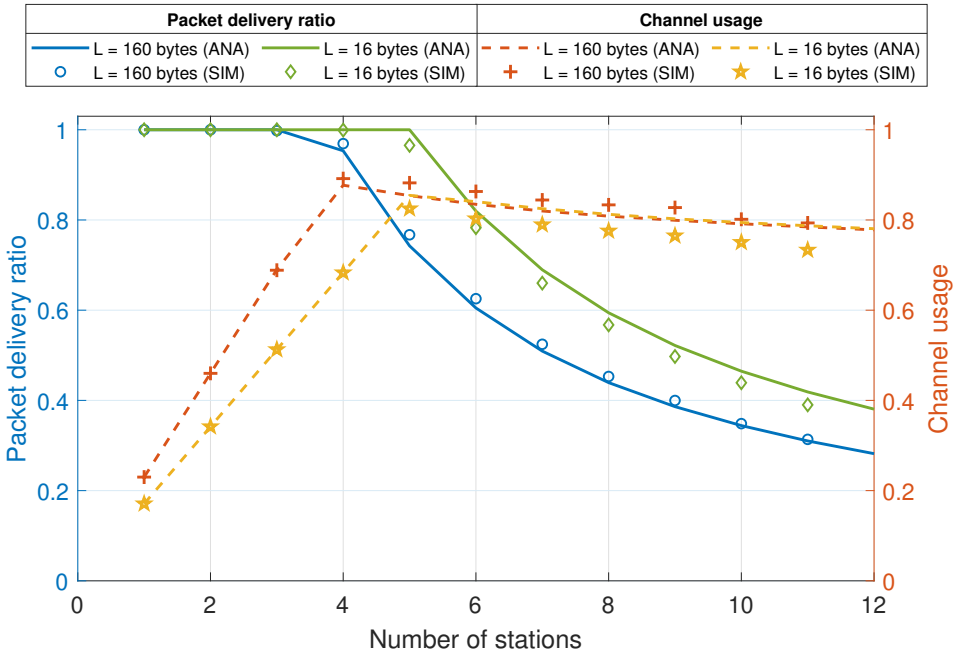


Figure 7.5: PDR and channel usage for stations in the area of range $[1, 10]$, assigned to a 10 ms RAW slot, with capture threshold $z = 4$ dB.

Figure 7.5 also depicts the channel usage representing the portion of the RAW slot duration that is used to transmit packets successfully. The channel usage starts with its lowest value when only one station is assigned to the RAW slot. That is because only a small portion of the RAW slot is used to transmit the station's packet and the remaining time in the RAW slot is wasted. The channel usage keeps increasing as long as the assigned stations yield a 100% PDR, for the same reason for the case of one station. However, this trend changes when the PDR starts to degrade due to the inadequacy of the 10 ms RAW slot to host the entire traffic load of stations. Hereafter, the extra contention introduced by more stations increases the collision probability, which expands the time needed to achieve a successful transmission, i.e., extends the length of the renewal cycle Z_k. Eventually, the RAW slot T_S becomes occupied by more idle slots, failed transmissions, and less successful packet deliveries, which diminishes the channel usage.

We also observe in the results that the gain of a pre-defined RAW slot can be maximized when an adequate number of stations is assigned to it. The gain of the RAW slot can be presented by a trade-off between PDR and channel usage. In the case of the scenario considered in Figure 7.5, we see that four stations for $L = 160$ bytes and five stations for $L = 60$ bytes are good choices for this RAW slot of 10 ms, as they provide higher combined values for PDR and channel usage. These results can also be seen in Figure 7.3, where T_S^* is valued at about 10 ms for four and five stations with a payload size of 160 and $L = 60$ bytes, respectively. In such cases, the RAW slot is efficiently allocated, providing two main gains: (1) maximize the channel usage and avoid waste of channel time; (2) ensure successful delivery of all stations' packets and guarantee that no station will have to wait for another allocated RAW slot in the

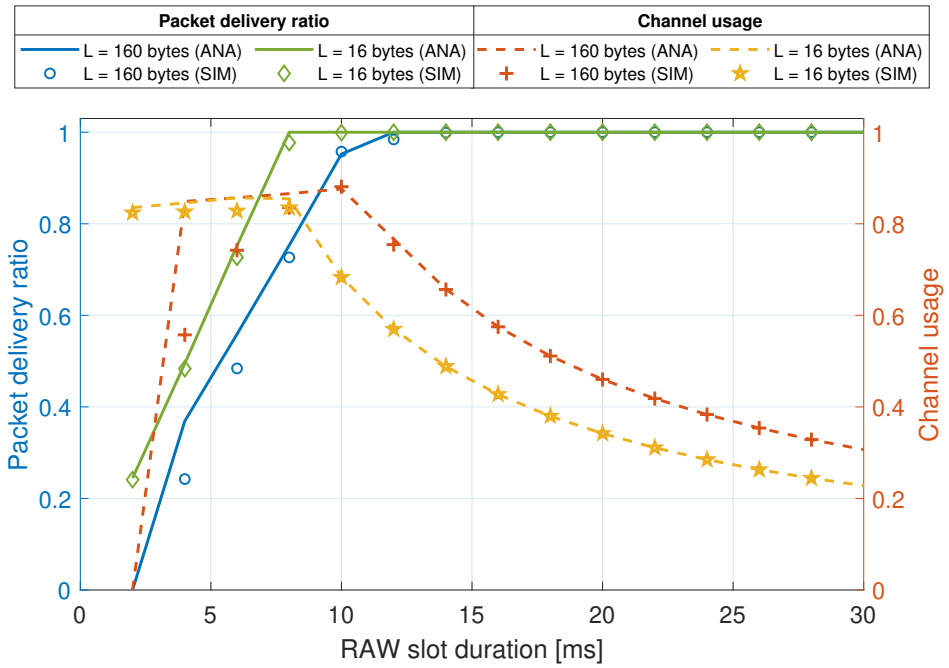


Figure 7.6: Performance of a pre-defined RAW slot in terms of its duration, with 4 assigned stations and capture threshold $z = 4\,\text{dB}$.

future to try again to transmit the same packet, which eventually enhances the AoI of received data at the AP.

To analyze the impact of allocating a different duration than T_S^* to a group of stations, we depict in Figure 7.6 the performance of a RAW slot allocated to four stations in terms of its duration. As the RAW slot length increases, more channel time becomes available to transmit packets, which enhances the PDR. Nevertheless, a very short period may not be enough to host at least one transmission and results in a null performance, as the case here for the length of 2 ms with a 160 Bytes payload. The channel usage keeps an almost stable performance up to the duration corresponding to T_S^*, where it starts decreasing. As seen above in Figure 7.3, T_S^* for four stations with a 160 Bytes payload is around 10 ms, which is the adaptive length observed in Figure 7.6, maximizing the PDR and enhancing the channel usage up to more than 80%.

Furthermore, we present in Figure 7.7 simulation results for the percentage of different events occupying the RAW slot period when allocated to four stations with a 160 Bytes payload. As the length of 2 ms is not enough to host at least one transmission, it results in a complete waste of channel time. A more extended RAW slot duration allows transmitting more packets, however, the waste of channel time still appears when the additional time is insufficient to host an extra transmission, and hence wasted. Nevertheless, in this case, the efficient choice corresponding to 10 ms yields the best results with the highest percentage of successful packet delivery through single transmissions and the capture of packets from collisions. A longer value for the RAW slot provides a complete delivery of the stations' traffic load but with a significant waste of channel time since each station leaves the contention after successfully delivering its packet. Henceforth, an adaptive allocation of the RAW slot as our proposal is essential to

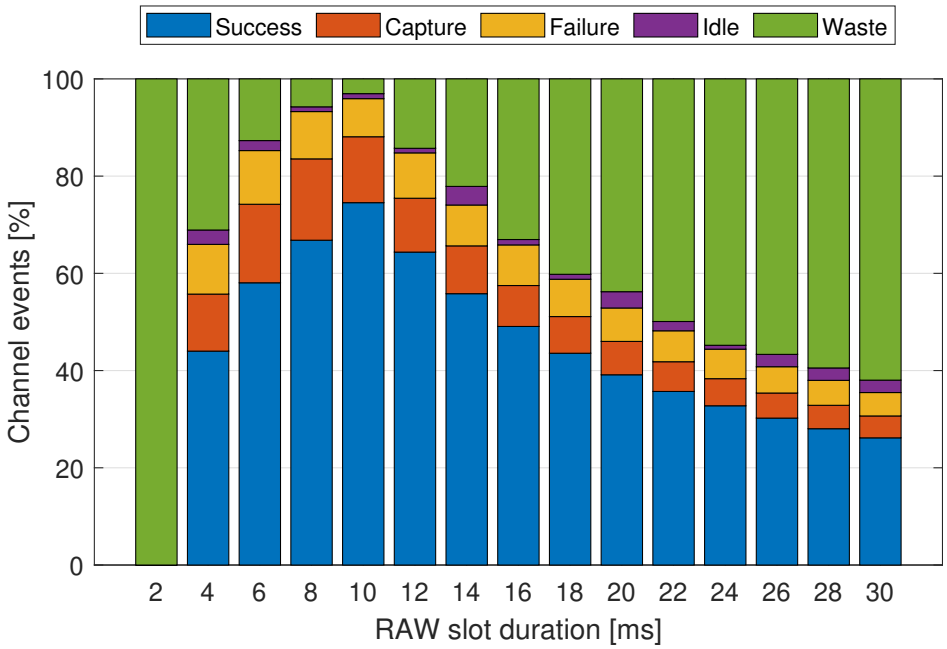


Figure 7.7: Channel events occupying a RAW slot allocated to 4 stations with a 160 Bytes payload and capture threshold $z = 4$ dB.

ensure complete packets' delivery on the one hand, and save resources by preventing the waste of channel time on the other hand.

7.4.3 Spatial Distribution of Stations

To further investigate the impact of the capture effect, we present in Figure 7.8 the LACA time for 15 stations distributed in an area of range 10 m. We evaluate T_S^* for this group of stations in terms of the distance of their area from the AP, under two different capture thresholds ($z = 2$ dB and $z = 10$ dB). The lower threshold value 2 dB increases the ACCP, providing more captured packets than the case of a 10 dB threshold. That is, eliminating the slots containing a failure transmission and eventually shortening T_S^* for the stations. We observe that the capture is more efficient when the stations are closer to the AP, where T_S^* yields minimal values. Even when the stations keep the same interdistance within a range of 10 m, the value of T_S^* increases when the stations' area is located far from the AP, and almost stabilizes when the stations are in a distance of 40 m or more. This behavior is due to the Rayleigh-fading channel considered in this scenario, where the capture of packets depends on the received power at the AP, which is distributed exponentially in terms of the stations' distance from the AP, as expressed in [65]. Such distribution of received power yields higher power attenuation when the station is located at a larger distance. Additionally, the distance between stations becomes ineffective when they are located in an area far from the AP.

Hereafter, we analyze the impact of the stations' area range in Figure 7.9. Starting from

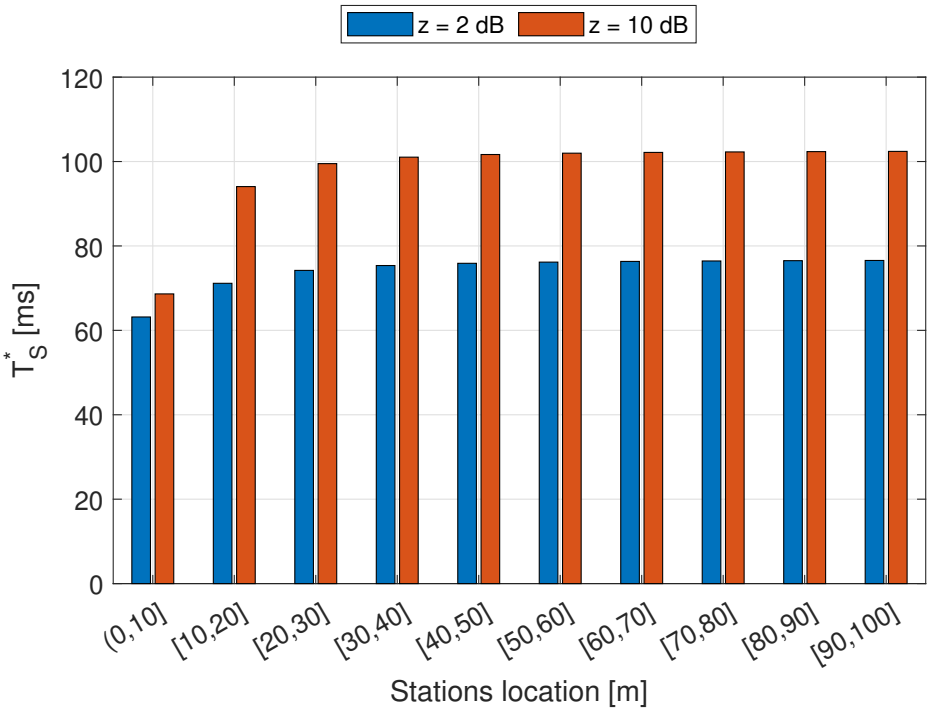


Figure 7.8: Impact of the distance of the location area from the AP for 15 stations.

the distance of 500 m, we increase the area range for 15 stations by a step of 50 m in both directions, and we present the results of T_S^* for two capture thresholds of 2 dB and 10 dB. As long as the stations are distributed over a wider area, the packets sent by stations closer to the AP are more likely to be captured when encountering a collision. Hence, as seen in the figure, the entire group of stations requires a shorter T_S^* to deliver its traffic load. That is explained by the fact that a lower ratio of the distances of two colliding stations (i.e., r_i/r_ℓ in (7.13)) increases the conditional capture probability. Henceforth, the best case is when the stations are distributed in a range of 1000 m. Although all stations have the same chances of encountering a collision, the closest ones to the AP will mostly have their packets captured during a collision due to their higher received power. We additionally observe that the case of a 10 dB capture threshold is more affected by the stations' area range than the case of a 2 dB threshold. That is because the 2 dB threshold tolerates more packets' capturing than the 10 dB threshold even if the stations are close to each other, as the signal-to-interference ratio is more likely to exceed the lower threshold [65]. These results show that the best strategy to profit from the capture effect at the AP is to assemble stations located in a broader area. This will provide shorter RAW slot periods to serve the stations. However, the hidden node problem may appear if the stations are out of the sensing range of each other, e.g., when two stations are located on opposite sides of the AP. This issue can be solved by performing sectorization before configuring the RAW and choosing stations from the same sector where all stations can hear each other.

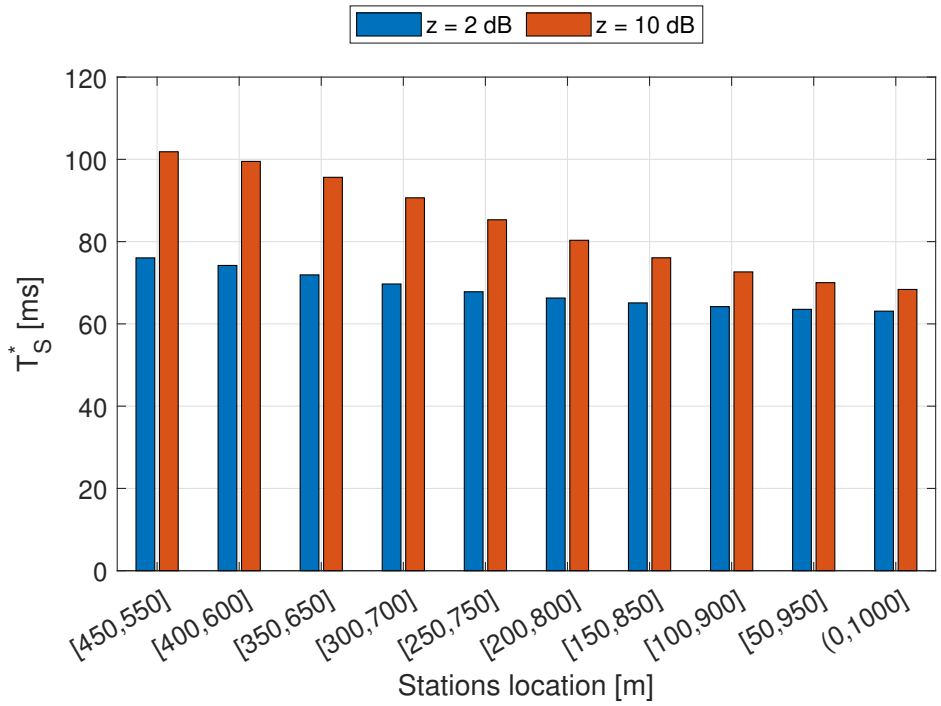


Figure 7.9: Impact of the location area range for 15 stations.

7.5 Conclusion

This chapter considered an IEEE 802.11ah-based network where stations are distributed around the AP and operate under a Rayleigh-fading channel with capture enabled. We develop an analytical model to evaluate the channel access performance and propose an algorithm to derive a LACA time for a group of stations assigned to a RAW slot. Our framework is based on a two-level renewal process that models the events within the channel when allocated to a group of stations, where each station has one packet to transmit. The first-level renewal cycle resets after every transmission attempt within the channel, while the second-level renewal cycle ends after every successful transmission, which can occur either when only one single station transmits or when the AP captures a packet from a collision. To highlight the effectiveness of the proposed LACA algorithm, we evaluate the PDR and channel usage for a pre-allocated RAW slot in terms of designated stations, payload size, and capture threshold. We further investigated the impact of the spatial distribution of stations in terms of both the area range and the distance from the AP.

Our analytical findings have been validated via extensive simulations obtained by a discrete-event simulator we developed with MATLAB. While most scheduling schemes for dense networks are based on assigning stations into groups with pre-allocated channel time, we showed in this work that allocating the channel periods for groups of stations is crucial. This operation is based on the state of contending stations (i.e., traffic load and locations) and can ensure successful delivery of all packets or a PDR and channel usage up to a certain extent. This approach of load-aware RAW slots is practical for remote sensing applications to mini-

mize the AoI at the AP and enhance the freshness of received data. With appropriate RAW slot allocation, all designated stations can transmit their packets without further attempts in subsequent RAW slots. Compared to a random pre-allocation of the RAW slot period, the proposed algorithm yields complete delivery of all packets while improving the channel usage up to more than 80%. The analytical framework developed in this paper provides an adaptive and load-aware channel allocation scheme based on the dynamic conditions of the network. This framework can be implemented in a global configuration scheme for the RAW mechanism in a heterogeneous network. One could consider different types of stations with different payload sizes and rates of generating packets. The proposed LACA algorithm can be improved towards dynamic inputs such as the arrival rate of packets and the mobility of stations. Additionally, our proposals are practical for alternative communication technologies addressing dense scenarios with the integration of periodic channel reservations.

To support large-scale IoT applications with efficient data rates, IEEE 802.11ah deploys relays to extend the transmission range between the AP and stations. A relay for the root AP is a station with buffered data, which is delivered through allocated RAW periods. We develop in the next chapter an analytical model for the forwarding mechanism in the relay, and propose a routing control policy to enhance the latency using an MDP framework.

CHAPTER 8

Dynamic Routing Control of Multiclass Traffic Through Relay-Based IEEE 802.11ah Networks

Contents

8.1 Introduction

The IoT applications are becoming mandatory to organize sensors and actuators in a complete IoT system. Several use cases evolve along with IoT smart devices, such as healthcare, smart agriculture, smart industry, home automation, etc [1, 101]. However, IoT communication technologies face several challenges in offering reliable connectivity, especially the large number of devices, their energy constraints as battery-operated devices and the broad deployment of

sensors. Therefore, these large scale scenarios present crucial challenges to offer high scalability management of such complex networks [102]. The IEEE 802.11.ah standard is one of the promising technologies to address IoT requirements. It is the first Wi-Fi solution for IoT application aiming to connect thousands of devices data rates ranging from 150 Kbps to 78 Mbps over a transmission range of up to 1 km [3]. A single AP provides connectivity to up to 8191 nodes using new channel access mechanisms. The AP periodically allocates the channel to a set of stations for a limited period using the new channel access mechanism, RAW.

Due to the long-range coverage of an IEEE 802.11ah AP and the large number of associated stations, the problem of hidden nodes is very crucial for IEEE 802.11ah-based networks [95]. To address this issue, TGah developed a group sectorization mechanism to divide the coverage area into several geographical sectors, each containing a subset of stations that listen to each other transmissions. Additionally, the AP splits the channel time into SIs and allocates each SI to a single sector. Group sectorization can be seen as a simplified RAW version based only on location-aware grouping. It improves the throughput by eliminating the hidden nodes problem and the energy efficiency as each station can be in sleep mode during SI other than the one it belongs to.

To support large-scale applications with efficient data rates, IEEE 802.11ah deploys one or more relays in each sector to extend the transmission range between the AP and stations [4]. A relay forwards frames between its associated stations and the AP, called root AP. Hence, it extends the distance between stations and the AP, especially in scenarios with the presence of obstructions. Relays are also contributing to saving energy consumption for stations. Since the relay is close to its associated stations, the latter can transmit their frames with lower power and higher data rates, shortening the transmission time and hence the awake time of stations.

We consider in this work a heterogeneous IoT network with multiclass traffic forwarded through relays. It is crucial to prioritize traffic in dense and large-scale IoT applications. The time-sensitive sensors, for example, are required to deliver the updates as soon as they are generated, and optimizing system resources is less considered [96, 97]. Therefore, instead of packet delay, the latency metric that gains priority in such applications is the Age of Information (AoI), which measures the freshness of information at the receiver [98]. Hence, we consider a relay forwarding three different types of generated packets. The relay consists of three infinite buffers, each transmits with a different data rate under a different channel. Therefore, we represent the relay as a queuing system with three parallel M/M/1 queues, noted Q_1, Q_2 and Q_3. We assume that Q_1 has the highest data rate than Q_2, in which packets are transmitted with a data rate higher than Q_3. The traffic received at the relay arrives following different Poisson process rates. Although each type of the arriving packets should be processed through its assigned queue, we propose to transmit non-prioritized packets with a higher data rates without interfering with the performance of prioritized traffic.

The contributions of this chapter are listed as follows:

- We model the relay forwarding mechanism as a queuing system with three M/M/1 queues and three types of packets with different service priorities. Then, we propose routing control mechanism for the non-prioritized packets through queues with higher data rates.
- We formulate the proposed routing control under a MDP framework, where we aim to derive an optimal policy that minimizes the system discounted cost.
- We derive the equivalent discrete-time MDP problem and structural properties of the optimal routing policy. Then, we establish the characterization of the optimal policy and the service priority in the system.
- We evaluate the Mean Sojourn Time (MST) for non-prioritized packets when the optimal routing policy is applied and prove the effectiveness of our proposed routing control.
- We establish some numerical results to analyze the optimal policy and the impact of the routing costs. Additionally, we show that our proposal remarkably reduces the latency of packets within the relay.

The remaining of this chapter is organized as follows. Section 8.2 models the relay forwarding mechanism as a queuing system. Section 8.3 Proposes the routing control of arriving packets and formulates it under an MDP framework. Then, in Section 8.4 we derive the discrete-time equivalence of the formulated MDP problem. Section 8.5 proposes properties to characterize the optimal policy and the service priority. Then, evaluate the MST of non-prioritized packets. Therefore, Section 8.6 presents numerical results and discussion of the optimal policy, the routing costs and the MST. Finally, Section 8.7 concludes our work.

8.2 Problem Description

We consider a relay-based IEEE 802.11ah network, where each relay is associated with a group of stations. A relay logically consists of two components: a relay AP that is associated with the stations and a relay station associated with the root AP. During downlink transmissions, the root AP transmits the packets to the relay station, which forwards them to the relay AP. The latter transmits the packets to the addressed stations. The operation is reversed for uplink transmissions. We focus on the uplink transmissions where the relay collects data with different priorities and forward them to the root AP using different servers. Therefore, we model the relay as a queuing system consisting of three parallel M/M/1 queues noted Q_1, Q_2 and Q_3. Each queue has infinite capacity. The transmission of packets in each queue is carried out with a different data rate under a different channel. The transmission data rate of Q_1 is higher than Q_2, which is in turn higher than Q_3. Hereafter, the server of each queue represents the data rate and associated operating channel. Hence, we consider that the service time in each queue $Q_i, i = 1,2,3$ follows an exponential distribution of rate μ_i, where $\mu_1 > \mu_2 > \mu_3$. That is, the

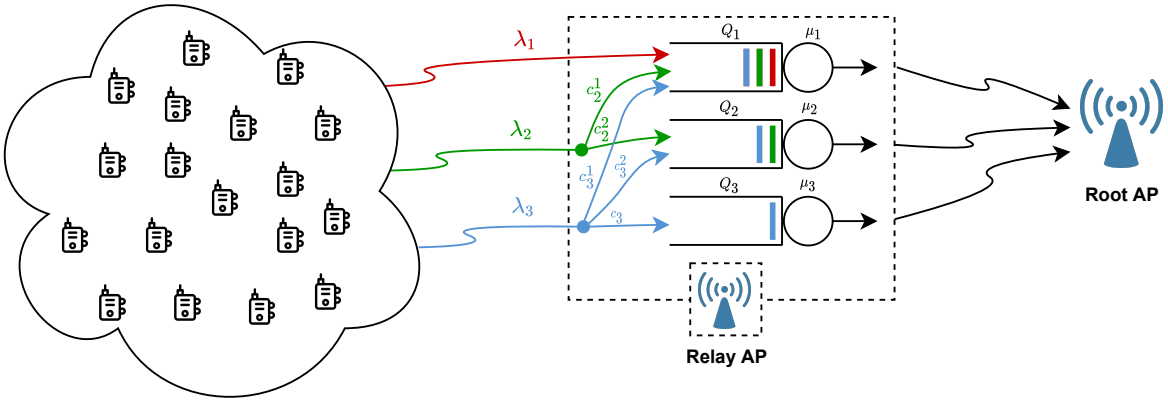


Figure 8.1: Illustration of the system model.

service time in Q_1 is shorter than that in Q_2, which is shorter than the service time in Q_3. Figure 8.1 illustrates the considered system model.

The relay receives three types of packets. Type-1 packets arrive following a Poisson process of rate λ_1 and enter to Q_1. Type-2 and type-3 packets arrive following a Poisson process of rate λ_2 and λ_3, respectively. Each type of packets is normally forwarded through its associated queue. However, since the service time in Q_1 is faster than Q_2, which is in turn faster than Q_3, it is important to avoid resource waste in the fast servers and use these servers for other packets with lower priority. Hereafter, we proposed to deploy a routing controller for the arriving type-2 and type-3 packets and forward them to the appropriate queue. Type-2 packets can be routed to either Q_1 or to Q_2, whereas type-3 packets can be routed to one of the three queues (Q_1, Q_2 or Q_3). The routing decision depends on the current state of the queuing system.

Hereafter, Q_1 and Q_2 contain different types of packets. However, the service in these queues does not follow the First-In-First-Out (FIFO) service discipline. In Q_1, type-1 packets have priority over type-2 packets, which have priority over type-3 packets. In Q_2, type-2 packets have priority type-3 packets. Within Q_1 or Q_2, the packets are served depending on their priority and not their arrival time. Upon the arrival of a prioritized packet, it is automatically put in front of a non-prioritized packet.

8.3 Problem Formulation

Let x_t^{ij} be the random variable representing the number of type-j packets in the queue Q_i at time t, where $(i,j) \in \{(1,1),(1,2),(1,3),(2,2),(2,3)\}$. Hence, we characterize the state of Q_1 at time t by the three-dimensional random vector $x_t^1 = (x_t^{11}, x_t^{12}, x_t^{13})$. Similarly, the state of Q_2 at time t is denoted by the two-dimensional random vector $x_t^2 = (x_t^{22}, x_t^{23})$. The state of the queue Q_3 at time t is represented by the random variable x_t^3 denoting the number of type-3 packets in Q_3 at time t. Note that the total number of packets in a queue also includes the packet in service.

Hereafter, we define the queuing system state by stochastic process $\{x_t\}_{t\geq 0}$, where $x_t = (x_t^1, x_t^2, x_t^3)$. x_t is a six-dimensional stochastic process representing the number of packets in the queuing system. The system state changes when a new packet arrives to the system or when a packet is served and then quits the system. Each type of packet arrives following a Poisson process. The interarrival times are exponentially distributed. Additionally, the inter-departure times of packets from the queuing system are exponentially distributed. Therefore, with the fact that the exponential distribution is memoryless, we conclude that the system state $\{x_t\}_{t\geq 0}$ is a Markov process with state space $E = \mathbb{N}^6$.

If at time t a type-2 (respectively, type-3) packet arrives in the system, an action a_t^2 (respectively, a_t^3) is taken. $a_t^2 = i$, $i = 1,2$ if the controller sends the arrived type-2 packet to the queue Q_i. $a_t^3 = j$, $j = 1,2,3$ if the controller sends the arrived type-3 packet to the queue Q_j. Therefore, we associate to each state $x_t \in E$ a set of admissible actions $D = \{1,2\} \times \{1,2,3\}$. Hence, and admissible action at time t is given by $a_t = (a_t^2, a_t^3)$.

An admissible policy is a mapping from states to actions. It allows the controller to decide the appropriate action for each arriving packet depending on the system state at the arrival instant. Henceforth, denoting the action space by $A = D^E$, an admissible policy is an A-valued stochastic process $\pi = \{a_t, t \in \mathbb{R}^+\}$.

Let t_n be the instant of the n^{th} transition of the system state, and x_n be the state of the system after this transition. We denote by $\tau_n = t_n - t_{n-1}$ the n^{th} sojourn time, and a_n the action taken at time t_n. Let x_t and a_t denote the state and the action at the last transition before the instant t. We have $(x_t, a_t) = (x_n, a_n)$ for $t_n \leq t < t_{n+1}$.

Hereafter, for each instant t, an instant cost $c(x_t, a_t)$ is incurred and it is given by:

$$c(x_t, a_t) = c_2^1(x_t^{11} + x_t^{12}) + c_2^2 x_t^{22} + c_3^1(x_t^{11} + x_t^{12} + x_t^{13}) + c_3^2(x_t^{22} + x_t^{23}) + c_3 x_t^3, \tag{8.1}$$

where $c_i^j, i = 2,3, j = 1,2$ is the routing cost of a type-i packet to Q_j. c_3 is the routing cost of a type-3 packet to Q_3. We assume that all routing costs are positive real numbers.

Our objective is to find an optimal policy π minimizing the following discounted cost:

$$\limsup_{T\to\infty} E_\pi \left[\int_0^T e^{-\delta t} c\,(x_t, a_t)\, dt \mid x_0 = x \right] \tag{8.2}$$

where E_π is the expected value under the policy π, and $\delta > 0$ is the discount rate. x denotes the initial state of the system.

Remark 8.1.

1. *To avoid the scenario where the policy that sends arriving packets to the faster server is optimal, we assume that* $c_2^1 > c_2^2$ *and* $c_3^1 > c_3^2 > c_3$.
2. *To ensure the stability of the queuing system, we assume that* $\mu_1 > \lambda_1 + \lambda_2 + \lambda_3$, $\mu_2 > \lambda_2 + \lambda_3$ *and* $\mu_3 > \lambda_3$.

An admissible policy π that is time-independent and decides a fixed action $a = \pi_i$ whenever the system is in state x is called a Deterministic Stationary Policy (DSP) [103]. Let $\mathcal{P}^d$ denote the set of all DSP's.

Let $A(x)$ is the set of actions available for the controller when the system is in state x. Since the state space E is countable and $A(x)$ is countable for each $x \in E$, then according to [104] (Theorem 1, p. 1229), there exists an optimal policy for (8.2). Furthermore, we have $c(x,a) \geq 0$ for every $x \in E$ and $a \in A(x)$, then according to [105] (Theorem 7.2.11, p. 294), there exists a deterministic stationary optimal policy for (8.2).

Therefore, we focus on DSPs and define the δ-discounted cost associated to the problem (8.2) as follows:

$$J^{\delta}(x) = \min_{\pi \in \mathcal{P}^d} \mathrm{E}_{\pi} \left[\int_0^{+\infty} e^{-\delta t} c(x_t, a_t)\, dt \mid x_0 = x \right] \tag{8.3}$$

where the objective function is the δ-discounted expected cost for the infinite horizon under a given policy $\pi \in \mathcal{P}^d$ and an initial system state $x \in E$, noted as follows:

$$J^{\delta}(x,\pi) = \mathrm{E}_{\pi} \left[\int_0^{+\infty} e^{-\delta t} c(x_t, a_t)\, dt \mid x_0 = x \right] \tag{8.4}$$

The expression (8.3) can further be simplified as a limit of the expected cost over a finite horizon. Let $0 = t_0 < t_1 < \ldots < t_n < \ldots$ be the random instants in time denoting transition epochs of the system state. The δ-discounted expected cost over the time-horizon $[0, t_N)$, under the policy $\pi \in \mathcal{P}^d$ is expressed as follows:

$$J_N^{\delta}(x,\pi) = \mathrm{E}_{\pi} \left[\int_0^{t_N} e^{-\delta t} c(x_t, a_t)\, dt \mid x_0 = x \right] \tag{8.5}$$

Let

$$J_N^{\delta}(x) = \min_{\pi \in \mathcal{P}^d} J_N^{\delta}(x,\pi), \quad N = 0, 1, \ldots \tag{8.6}$$

and

$$J_{\infty}^{\delta}(x) = \lim_{N \to \infty} J_N^{\delta}(x). \tag{8.7}$$

Following the same procedure as in [106], one can prove that the minimum cost in (8.3) for a every initial state x can be expressed alternatively as follows:

$$J^{\delta}(x) = J_{\infty}^{\delta}(x). \tag{8.8}$$

8.4 Problem Discretization

This section converts the original continuous-time problem (8.3) into its discrete-time equivalent using the standard procedure of uniformization [107, 108, 109].

The inter-epoch intervals are seen to be Independent and Identically Distributed (IID) random variables distributed as follows:

$$\Pr\{t_{k+1} - t_k > t\} = e^{-t\nu} \qquad k = 0, 1, 2, \ldots$$

where $\nu = \lambda_1 + \lambda_2 + \lambda_3 + \mu_1 + \mu_2 + \mu_3$.

Proposition 8.1. *Let $(x_n, a_n) = (x_t, a_t)$ for $t_n \leq t < t_{n+1}$. For every policy $\pi \in \mathcal{P}^d$ and every initial state $x \in E$, we have*

(i)

$$J^{\delta}(x,\pi) = \frac{1}{\delta+\nu}\,\mathrm{E}_{\pi}\left[\sum_{n=0}^{+\infty} \beta^n c(x_n, a_n) \mid x_0 = x\right] \tag{8.9}$$

where $\beta = \dfrac{\nu}{\delta+\nu}$.

(ii) Furthermore, for every $N \in \mathbb{N}$ we have

$$J_N^{\delta}(x,\pi) = \frac{1}{\delta+\nu}\,\mathrm{E}_{\pi}\left[\sum_{n=0}^{N-1} \beta^n c(x_n, a_n) \mid x_0 = x\right] \tag{8.10}$$

Proof. Both assertions of the proposition can be proved following the same procedure. Hence, we restrict our proof to only the assertion *(i)*.

Consider a system state x_t at time t with the corresponding action a_t. For a given policy $\pi \in \mathcal{P}^d$ and an initial state $x \in E$, we have $J^{\delta}(x,\pi)$ expressed as follows:

$$\begin{aligned}
\mathrm{E}_{\pi}\left[\int_0^{+\infty} e^{-\delta t} c(x_t, a_t)\, dt \mid x_0 = x\right] &= \mathrm{E}_{\pi}\left[\sum_{n=0}^{+\infty} \int_{t_n}^{t_{n+1}} e^{-\delta t} c(x_t, a_t)\, dt \mid x_0 = x\right] && (8.11)\\
&= \sum_{n=0}^{+\infty} \mathrm{E}_{\pi}\left[\int_{t_n}^{t_{n+1}} e^{-\delta t} c(x_n, a_n)\, dt \mid x_0 = x\right] && (8.12)\\
&= \sum_{n=0}^{+\infty} \mathrm{E}_{\pi}\left[\int_{t_n}^{t_{n+1}} e^{-\delta t}\, dt\right] \mathrm{E}_{\pi}\left[c(x_n, a_n) \mid x_0 = x\right] && (8.13)
\end{aligned}$$

The equality (8.13) comes from the fact that the stochastic processes $\{x_t, t \in \mathbb{R}^+\}$ and $\{a_t, t \in \mathbb{R}^+\}$ representing the system states and control actions are independent of the sojourn times $\{\tau_n, n \in \mathbb{N}^*\}$. That is because the policy π is deterministic, i.e., it does not depend on τ_n.

Since the random variables τ_n are IID, we have:

$$\begin{aligned}
\mathrm{E}_\pi\left[\int_{t_n}^{t_{n+1}} e^{-\delta t}\right] &= \mathrm{E}_\pi\left[e^{-\delta t_n}\int_0^{\tau_{n+1}} e^{-\delta t}\,dt\right] && (8.14)\\
&= \frac{1}{\delta}\mathrm{E}_\pi\left[e^{-\delta t_n}\right]\left(1-\mathrm{E}_\pi\left[e^{-\delta\tau_{n+1}}\right]\right) && (8.15)\\
&= \frac{1}{\delta}\left(\frac{\nu}{\delta+\nu}\right)^n\left(1-\frac{\nu}{\delta+\nu}\right) && (8.16)\\
&= \left(\frac{1}{\delta+\nu}\right)\left(\frac{\nu}{\delta+\nu}\right)^n && (8.17)
\end{aligned}$$

The result in (8.16) comes from the fact that for any $n \geq 0$, the random variable τ_{n+1} is exponentially distributed with the rate ν. That is,

$$\mathrm{E}_\pi\left[e^{-\delta\tau_{n+1}}\right] = \frac{\nu}{\delta+\nu}$$

Additionally, we have

$$\mathrm{E}_\pi\left[e^{-\delta t_n}\right] = \mathrm{E}_\pi\left[e^{-\delta(\Sigma_{k=1}^n t_k - t_{k-1})}\right] = \prod_{k=1}^n \mathrm{E}_\pi\left[e^{-\delta\tau_k}\right] = \left(\frac{\nu}{\delta+\nu}\right)^n$$

Denote $\beta = \frac{\nu}{\delta+\nu}$. Therefore, by substituting (8.17) in (8.13), we obtain the result as follows:

$$\begin{aligned}
\mathrm{E}_\pi\left[\int_0^{+\infty} e^{-\delta t}c(x_t,a_t)\,dt \mid x_0 = x\right] &= \sum_{n=0}^{+\infty}\mathrm{E}_\pi\left[\left(\frac{1}{\delta+\nu}\right)\beta^n\right]\mathrm{E}_\pi\left[c(x_n,a_n)\mid x_0=x\right] && (8.18)\\
&= \frac{1}{\delta+\nu}\mathrm{E}_\pi\left[\sum_{n=0}^{+\infty}\beta^n c(x_n,a_n)\mid x_0=x\right] && (8.19)
\end{aligned}$$

□

Hereafter, for a given policy $\pi \in \mathcal{P}^d$, we can define the β-discounted cost incurred by the N-step discrete time system as follows:

$$V_N^\beta(x,\pi) = \mathrm{E}_\pi\left[\sum_{n=0}^{N-1}\beta^n c(x_n,a_n)\mid x_0=x\right] \tag{8.20}$$

Additionally, the β-discounted cost for the infinite horizon is defined as:

$$V^\beta(x,\pi) = \mathrm{E}_\pi\left[\sum_{n=0}^{+\infty}\beta^n c(x_n,a_n)\mid x_0=x\right] \tag{8.21}$$

From the assertions of Proposition 8.1, we have

$$J_N^{\delta}(x,\pi) = \frac{1}{\delta+\nu} V_N^{\beta}(x,\pi) \tag{8.22}$$

and

$$J^{\delta}(x,\pi) = \frac{1}{\delta+\nu} V^{\beta}(x,\pi) \tag{8.23}$$

For a given initial system state $x \in E$, we denote by $V_N^{\beta}(x)$ and $V^{\beta}(x)$ the optimal β-discounted cost in N-step discrete system times and infinite horizon,respectively. We have:

$$V_N^{\beta}(x) = \min_{\pi \in \mathcal{P}^d} V_N^{\beta}(x,\pi) \tag{8.24}$$

and

$$V^{\beta}(x) = \min_{\pi \in \mathcal{P}^d} V^{\beta}(x,\pi) \tag{8.25}$$

Let

$$V_{\infty}^{\beta}(x) = \lim_{N \to \infty} V_N^{\beta}(x) \tag{8.26}$$

Similar to the previous section, we have

$$V_{\infty}^{\beta}(x) = V^{\beta}(x) \tag{8.27}$$

for every initial state $x \in E$.

Finally, we can now derive the equivalence in the sense of optimal discounted cost between the original optimization problem (8.3) and its discrete-time formulation. From (8.8) we have

$$J^{\delta}(x) = \lim_{N \to \infty} \min_{\pi \in \mathcal{P}^d} J_N^{\beta}(x,\pi) \tag{8.28}$$

Hence, using (8.22), (8.24) and (8.27), we obtain

$$J^{\delta}(x) = \frac{1}{\delta+\nu} V^{\beta}(x) \tag{8.29}$$

Hereafter, we restrict our attention to the discrete-time β-discounted cost problem defined in (8.25). We assert from Walrand [110] (Prop. 8.5.3, p. 275) that $V^{\beta}(x)$ is the unique bounded solution of the following dynamic programming equation:

$$V^{\beta}(x) = \min_{a \in A(x)} \left\{ C(x,a) + \beta \sum_{y \in E} \Pr\{y \mid x,a\} V^{\beta}(y) \right\}, \tag{8.30}$$

where $\Pr\{y \mid x,a\}$ is the conditional probability that the system moves to state y when the action a is applied to state x.

As proven in Section 8.3, our problem admits an optimal stationary policy which is a sequence of actions that solve the dynamic programming equation (8.30). This equation is also referred to as Bellman's optimality equation and can be solved using algorithms such as Value Iteration (VI), Policy Iteration (PI), and Linear Programming (LP). In Section 8.6, we use the VI algorithm because of its theoretical simplicity and ease of implementation. It provides a Lookup Table (LT) containing the optimal stationary policy (i.e., optimal actions for every system state) and the corresponding minimum expected total costs.

Let's now express the transition probabilities in (8.30). For each control action $a = (a^2, a^3)$, we define the transition probability function $\Pr\{. \mid ., a\}$ on $E \times E$ by:

$$\begin{aligned}\Pr\{y \mid x, a\} = \frac{1}{\nu}\Big(&\lambda_1 \mathbb{1}_{\{y=A_{11}x\}} + \lambda_2 \mathbb{1}_{\{a^2=1\}}\mathbb{1}_{\{y=A_{12}x\}} + \lambda_2 \mathbb{1}_{\{a^2=2\}}\mathbb{1}_{\{y=A_{22}x\}} \\ &+ \lambda_3 \mathbb{1}_{\{a^3=1\}}\mathbb{1}_{\{y=A_{13}x\}} + \lambda_3 \mathbb{1}_{\{a^3=2\}}\mathbb{1}_{\{y=A_{23}x\}} + \lambda_3 \mathbb{1}_{\{a^3=3\}}\mathbb{1}_{\{y=A_3x\}} \\ &+ \mu_1\Big(\mathbb{1}_{\{y=D_{11}x\}} + \mathbb{1}_{\{y=D_{12}x, x^{11}=0\}} + \mathbb{1}_{\{y=D_{13}x, x^{11}=0, x^{12}=0\}}\Big) \\ &+ \mu_2\Big(\mathbb{1}_{\{y=D_{22}x\}} + \mathbb{1}_{\{y=D_{23}x, x^{22}=0\}}\Big) + \mu_3 \mathbb{1}_{\{y=D_3x\}}\Big)\end{aligned} \tag{8.31}$$

where for any state $x = (x^{11}, x^{12}, x^{13}, x^{22}, x^{23}, x^3) \in E$:

$$\begin{aligned}A_{11}x &= (x^{11}+1, x^{12}, x^{13}, x^{22}, x^{23}, x^3) \\ A_{12}x &= (x^{11}, x^{12}+1, x^{13}, x^{22}, x^{23}, x^3) \\ A_{13}x &= (x^{11}, x^{12}, x^{13}+1, x^{22}, x^{23}, x^3) \\ A_{22}x &= (x^{11}, x^{12}, x^{13}, x^{22}+1, x^{23}, x^3) \\ A_{23}x &= (x^{11}, x^{12}, x^{13}, x^{22}, x^{23}+1, x^3) \\ A_3x &= (x^{11}, x^{12}, x^{13}, x^{22}, x^{23}, x^3+1) \\ D_{11}x &= ((x^{11}-1)^+, x^{12}, x^{13}, x^{22}, x^{23}, x^3) \\ D_{12}x &= (x^{11}, (x^{12}-1)^+, x^{13}, x^{22}, x^{23}, x^3) \\ D_{13}x &= (x^{11}, x^{12}, (x^{13}-1)^+, x^{22}, x^{23}, x^3) \\ D_{22}x &= (x^{11}, x^{12}, x^{13}, (x^{22}-1)^+, x^{23}, x^3) \\ D_{23}x &= (x^{11}, x^{12}, x^{13}, x^{22}, (x^{23}-1)^+, x^3) \\ D_3x &= (x^{11}, x^{12}, x^{13}, x^{22}, x^{23}, (x^3-1)^+)\end{aligned}$$

with $x^+ = \max\{x, 0\}$.

Therefore, we can characterize V^β by the dynamic optimality equation $V^\beta = TV^\beta$, with

T being the operator defined on the set $\mathcal{F}(E,\mathbb{R})$ of real-valued functions on E as follows:

$$\begin{aligned} TV^{\beta}(x) = {} & c_1^1(x^{11}+x^{12}) + c_1^2 x^{22} + c_2^1(x^{11}+x^{12}+x^{13}) + c_2^2(x^{22}+x^{23}) + c_3 x^3 \\ & + \frac{1}{\nu(\delta+\nu)}\Big[\mu_1\left(V^{\beta}(D_{11}x) + V^{\beta}(D_{12}x)\mathbb{1}_{\{x^{11}=0\}} + V^{\beta}(D_{13}x)\mathbb{1}_{\{x^{11}=0,x^{12}=0\}}\right) \\ & + \mu_2\left(V^{\beta}(D_{22}x) + V^{\beta}(D_{23}x)\mathbb{1}_{\{x^{22}=0\}}\right) + \mu_3 V^{\beta}(D_3 x) \\ & + \lambda_1 V^{\beta}(A_{11}x) + \lambda_2 \min\{V^{\beta}(A_{12}x), V^{\beta}(A_{22}x)\} \\ & + \lambda_3 \min\{V^{\beta}(A_{13}x), V^{\beta}(A_{23}x), V^{\beta}(A_3 x)\}\Big], \end{aligned} \tag{8.32}$$

for every $V^{\beta} \in \mathcal{F}(E,\mathbb{R})$.

8.5 Optimal Policy and Discounted Cost

8.5.1 Structural Properties

Although the optimal policy is derived from solving Bellman's dynamic programming equation (8.30), it is very useful to have structural properties characterizing the MDP model. These properties allow determining optimal actions without extra computations of solving the dynamic programming equation. We shall show that the optimal cost function $V_N^{\beta}(x)$ defined in (8.24) satisfies the following properties:

Proposition 8.2. *$\forall x \in E$, $\forall N \geq 0$, we have:*

(P1) $V_N^{\beta}(x)$ *is increasing in* $x^{11}, x^{12}, x^{13}, x^{22}, x^{23}, x^{23}$ *and* x^3.

(P2) $V_N^{\beta}(A_{12}x) - V_N^{\beta}(A_{22}x)$ *is increasing in* x^{11}, x^{12} *and* x^{13}*, decreasing in* x^{22} *and* x^{23}*, and constant in* x^3.

(P3) $V_N^{\beta}(A_{13}x) - V_N^{\beta}(A_{23}x)$ *is increasing in* x^{11}, x^{12} *and* x^{13}*, decreasing in* x^{22} *and* x^{23}*, and constant in* x^3.

(P4) $V_N^{\beta}(A_{13}x) - V_N^{\beta}(A_3 x)$ *is increasing in* x^{11}, x^{12} *and* x^{13}*, decreasing in* x^3*, and constant in* x^{22} *and* x^{23}.

(P5) $V_N^{\beta}(A_{23}x) - V_N^{\beta}(A_3 x)$ *is increasing in* x^{22} *and* x^{23}*, decreasing in* x^3*, and constant in* x^{11}, x^{12} *and* x^{13}.

Let H be the set of functions V_N^β defined in (8.24) verifying the properties (P1), (P2), (P3), (P4) and (P5), such that

$$V_N^\beta \in H \Longrightarrow V_{N+1}^\beta := TV_N^\beta \in H \tag{8.33}$$

where T is the operator defined in (8.32).

Assuming further that H contains the function that is identically zero, and that it is closed for pointwise limits (i.e., the pointwise limit of the sequence of functions V_N of H is a function of H). Hence, we have $V^\beta \in H$.

Proof. See Appendix. □

Following these results, the properties of V_N^β are also verified by the optimal discounted cost V^β.

8.5.2 Characterization of the Optimal Policies

This section presents interpretations of the structural properties above for the optimal discounted cost V^β.

(P1) represents the monotonicity of the discounted cost in terms of the system load, where an extra arrival of any type of packets results in a higher cost.

(P2) characterizes the sending regions of type-2 packets that arrive to Q_1 and Q_2. There is a monotonic switching curve between these two regions. If it is optimal to send the arriving type-2 packet to Q_1 in state x (i.e., $V^\beta(A_{12}x) \leq V^\beta(A_{22}x)$), then the same decision remains optimal when there is an additional packet in the queue Q_2 (i.e., in the states $A_{22}x$ and $A_{23}x$).

Indeed, the property (P2) implies that

$$\begin{cases} V^\beta(A_{12}A_{22}x) - V^\beta(A_{22}A_{22}x) & \leq & V^\beta(A_{12}x) - V^\beta(A_{22}x) & < & 0 \\ V^\beta(A_{12}A_{23}x) - V^\beta(A_{22}A_{23}x) & \leq & V^\beta(A_{12}x) - V^\beta(A_{22}x) & < & 0 \end{cases}$$

$$\Longrightarrow \begin{cases} V^\beta(A_{12}A_{22}x) & < & V^\beta(A_{22}A_{22}x) \\ V^\beta(A_{12}A_{23}x) & < & V^\beta(A_{22}A_{23}x) \end{cases}$$

Similarly, (P2) indicates that if it is optimal to send arriving type-2 packet to Q_2 when the system state is x (i.e., $V^\beta(A_{22}x) < V^\beta(A_{12}x)$), then the same decision remains optimal when there is an additional packet in Q_1 (i.e., in states $A_{11}x, A_{12}x$ and $A_{13}x$).

The properties (P3), (P4) and (P5) characterize the sending regions of the arriving type-3 packets to the queues Q_1, Q_2 and Q_3. There are switching surfaces between these three regions. The projection of the surface that separates every two regions is a monotonic curve with respect to the two types of packets associated with these regions, and a constant curve with respect to the other type of packets.

If it is optimal to send the arriving type-3 packet to Q_1 when the system state is x (i.e., $V^\beta(A_{13}x) \leq \min\{V^\beta(A_{23}x), V^\beta(A_3x)\}$, then the same decision remains optimal when there is an additional packet in the queues Q_2 and Q_3 (i.e., in states A_{22}, A_{23} and A_3).

When a type-3 packet arrives to the system in state x, the control action can be expressed as follows:

$$a^3 = \begin{cases} 1, & \text{if } V^\beta(A_{13}x) \leq \min\{V^\beta(A_{23}x), V^\beta(A_3x)\} \\ 2, & \text{if } V^\beta(A_{23}x) \leq \min\{V^\beta(A_{13}x), V^\beta(A_3x)\} \\ 3, & \text{if } V^\beta(A_3x) \leq \min\{V^\beta(A_{13}x), V^\beta(A_{23}x)\} \end{cases} \tag{8.34}$$

Furthermore, the property (P3) implies that

$$\begin{cases} V^\beta(A_{13}A_{22}x) - V^\beta(A_{23}A_{22}x) \leq V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \\ V^\beta(A_{13}A_{23}x) - V^\beta(A_{23}A_{23}x) \leq V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \\ V^\beta(A_{13}A_3x) - V^\beta(A_{23}A_3x) = V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \end{cases}$$

and the property (P4) implies that

$$\begin{cases} V^\beta(A_{13}A_{22}x) - V^\beta(A_3A_{22}x) = V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \\ V^\beta(A_{13}A_{23}x) - V^\beta(A_3A_{23}x) = V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \\ V^\beta(A_{13}A_3x) - V^\beta(A_3A_3x) \leq V^\beta(A_{13}x) - V^\beta(A_{23}x) < 0 \end{cases}$$

Hence

$$\begin{cases} V^\beta(A_{13}A_{22}x) < \min\{V^\beta(A_{23}A_{22}x), V^\beta(A_3A_{22}x)\} \\ V^\beta(A_{13}A_{23}x) < \min\{V^\beta(A_{23}A_{23}x), V^\beta(A_3A_{23}x)\} \\ V^\beta(A_{13}A_3x) < \min\{V^\beta(A_{23}A_3x), V^\beta(A_3A_3x)\} \end{cases}$$

Similarly, using the properties (P3) and (P5), we can show that if it is optimal to send the arriving type-3 packet to Q_2 in state x (i.e., $V^\beta(A_{23}x) \leq \min\{V^\beta(A_{13}x), V^\beta(A_3x)\}$, then the same decision remains optimal when there is an additional packet in the queues Q_1 and Q_3 (i.e., in states $A_{11}x, A_{12}x, A_{13}x$ and A_3x).

Furthermore, using the properties (P4) and (P5), we can show that if it is optimal to send the arriving type-3 packet to Q_3 when the system is in state x (i.e., $V^\beta(A_3x) \leq$

$\min\{V^\beta(A_{13}x), V^\beta(A_{23}x)\}$, then the same decision remains optimal when there is an additional packet in the queues Q_1 and Q_2 (i.e., in states $A_{11}x, A_{12}x, A_{13}x, A_{22}x$ and $A_{23}x$).

8.5.3 Service Priority

The considered queuing system receives packets of different classes, and the queues Q_1 and Q_2 contain different types of packets. Since the service in these queues is provided with respect to the priority of packets, we characterize in this section the service priority in the system.

In Q_1, type-1 packets have priority over type-2 and type-3 packets, and type-2 packets are prioritized over type-3 packets. When a prioritized packet enters Q_1 or Q_2 and finds a non-prioritized packet in service, it takes its place in service and puts it on hold. We characterize the service priority in Q_1 by the following proposition.

Proposition 8.3 (Service priority in Q_1)**.** *$\forall x \in E$, $\forall N \geq 0$, we have:*

$$V_N^\beta(D_{1i}x) \leq V_N^\beta(D_{1j}x) \tag{8.35}$$

for every $i > j$, with $i, j = 1, 2, 3$.

Proof. It is clear that V_0^β and V_1^β satisfy the inequality (8.35) above, since for every $x \in E$ we have $V_0^\beta(x) = 0$ and $V_1^\beta(x) = c(x,a)$.

Suppose that V_N^β satisfies the inequality (8.35) and let's show that V_{N+1}^β also satisfies this inequality.

We have $V_{N+1}^\beta = TV_N^\beta$. Hence, from the definition (8.32) of the operator T, it is sufficient to show that both functions:

$$\Delta_1(x) = \min\{V_N^\beta(A_{12}x), V_N^\beta(A_{22}x)\}$$

and

$$\Delta_2(x) = \min\{V_N^\beta(A_{13}x), V_N^\beta(A_{23}x), V_N^\beta(A_3x)\}$$

satisfy the inequality (8.35).

The additional elements in the definition (8.32) are linear functions, and we know that the inequality is maintained under linear combinations.

$\Delta_1(x)$ satisfies (8.35) is equivalent to:

$$\Delta_1(D_{1i}x) \leq \Delta_1(D_{1j}x) \tag{8.36}$$

$$\iff \min\{V_N^\beta(A_{12}D_{1i}x), V_N^\beta(A_{22}D_{1i}x)\} \leq \min\{V_N^\beta(A_{12}D_{1j}x), V_N^\beta(A_{22}D_{1j}x)\} \tag{8.37}$$

To prove the inequality (8.37), we deploy the following remark:

Remark 8.2. *We have the inequality* (8.37) *in the form:*

$$\min\{r_1, r_2\} \leq \min\{p_1, p_2\}. \tag{8.38}$$

If we show that $r_i \leq \min\{p_1, p_2\}$ *for a single* i $(i = 1,2)$*, then we have the inequality* (8.38)*, because* $\min\{r_1, r_2\} \leq r_i$.

For each case p_i, we choose r_i, and we will have the inequality $p_i - q_i \geq r_i - s_i$, which is satisfied by hypotheses for all $i = 1,2,3$.

Therefore, $\Delta_1(x)$ satisfies (8.35).

$\Delta_2(x)$ satisfies (8.35) is equivalent to:

$$\Delta_2(D_{1i}x) \leq \Delta_2(D_{1j}x) \tag{8.39}$$

$$\Longleftrightarrow \min\{V_N^\beta(A_{13}D_{1i}x), V_N^\beta(A_{23}D_{1i}x), V_N^\beta(A_3D_{1i}x)\} \leq \min\{V_N^\beta(A_{13}D_{1j}x), V_N^\beta(A_{23}D_{1j}x), V_N^\beta(A_3D_{1j}x)\} \tag{8.40}$$

We have the following remark:

Remark 8.3. *We have the inequality* (8.40) *in the form:*

$$\min\{r_1, r_2, r_3\} \leq \min\{p_1, p_2, p_3\}. \tag{8.41}$$

If we show that $r_i \leq \min\{p_1, p_2, p_3\}$ *for a single* i $(i = 1,2,3)$*, then we have the inequality* (8.40)*, because* $\min\{r_1, r_2, r_3\} \leq r_i$

For each p_i, we choose r_i, and we will have the inequality $p_i - q_i \geq r_i - s_i$, which is satisfied by hypotheses for all $i = 1,2,3$.

Therefore, $\Delta_2(x)$ satisfies (8.35).

Hence V_{N+1} satisfies the inequality (8.35) as a linear combination of the functions which satisfy this inequality.

Consequently, we have proved Proposition 8.3 by recurrence.

□

In Q_2, type-2 packets are prioritized over type-3 packets. We characterize this service priority in Q_2 by the following proposition.

Proposition 8.4 (Service priority in Q_2)**.** *$\forall x \in E$, $\forall N \geq 0$, we have:*

$$V_N^{\beta}(D_{22}x) \leq V_N^{\beta}(D_{23}x) \tag{8.42}$$

Proof. The proof of this proposition is derived by following the same procedure as in the previous proof of Proposition 8.3. □

8.5.4 Mean Sojourn Time

In order to prove the effectiveness of our proposed routing control of type-2 and type-3 packets, we evaluate the MST for these two types of packets and compare it to the case of no routing control. When no routing control is deployed and each traffic is forwarded to its designated queue, the system will consist of three independent M/M/1 queues. Each queue $Q_i, i = 1,2,3$ receives type-i packets following a Poisson process with the rate λ_i, and serves them following an exponential service time with the rate μ. In this case, for each $i = 1,2,3$, the MST of type-i packets is given by the following Little's result [111]:

$$T_i^* = \frac{S_i^*}{\lambda_i}, \tag{8.43}$$

where S_i^* is the average number of type-i packets in Q_i, given by

$$S_i^* = \frac{\lambda_i}{\mu_i - \lambda_i}. \tag{8.44}$$

In the case of optimal routing control, the MST of type-1 packets is not affected by the routing of type-2 and type-3 packets. That is because they have priority in Q_1 and are always put in the head of the queue. Therefore, we evaluate the MST for only type-2 and type-3 packets.

Type-2 packets are forwarded either to Q_1 or Q_2. Let α be the probability that an arriving type-2 packet is forwarded to Q_1. Thus, we can define the MST of type-2 packets as

$$T_2 = \alpha T_2^1 + (1-\alpha) T_2^2 \tag{8.45}$$

where T_2^i, $i = 1,2$ is the MST of type-2 packets in Q_i.

In Q_1, type-2 packets can be served after all the type-1 packets are forwarded. Hence, we have

$$T_2^1 = \frac{S_1^1 + S_2^1}{\alpha \lambda_2} \tag{8.46}$$

with S_i^1, $i = 1,2$ is the average number of type-i packets in Q_1.

We have

$$S_1^1 = \frac{\lambda_1}{\mu_1 - \lambda_1} \tag{8.47}$$

and

$$S_2^1 = \frac{\alpha\lambda_2}{\mu_1 - \alpha\lambda_2} \tag{8.48}$$

In Q_2, type-2 packets are prioritized and not affected by the presence of type-3 packets. Hence,

$$T_2^2 = \frac{S_2^2}{(1-\alpha)\lambda_2} \tag{8.49}$$

with S_2^2 being the average number of type-2 packets in Q_2 given by

$$S_2^2 = \frac{(1-\alpha)\lambda_2}{\mu_2 - (1-\alpha)\lambda_2} \tag{8.50}$$

Consequently, by substituting in (8.46), we obtain T_2 expressed as follows:

$$T_2 = \frac{(\alpha-1)\lambda_2(2\mu_1(\lambda_1+\alpha\lambda_2) - \mu_1^2 - 3\alpha\lambda_1\lambda_2) + \mu_2(\lambda_1\mu_1 + \alpha\lambda_2(\mu_1 - 2\lambda_1))}{\lambda_2(\mu_1-\lambda_1)(\mu_1-\alpha\lambda_2)(\mu_2+\alpha\lambda_2-\lambda_2)} \tag{8.51}$$

Type-3 packets are forwarded either to Q_1, Q_2 or Q_3. Let $\eta_i, i = 1,2,3$ be the probability that an arriving type-3 packet is forwarded to Q_i. Thus, we can define the MST of type-3 packets as

$$T_3 = \eta_1 T_3^1 + \eta_2 T_3^2 + \eta_3 T_3^3 \tag{8.52}$$

where T_3^i, $i = 1,2,3$ is the MST of type-3 packets in Q_i.

In Q_1, type-3 packets can be served only after type-1 and type-2 packets. Hence, we have

$$T_3^1 = \frac{S_1^1 + S_2^1 + S_3^1}{\eta_1\lambda_3} \tag{8.53}$$

with S_i^1, $i = 1,2,3$ being the average number of type-i packets in Q_1. S_1^1 and S_2^1 are respectively given by (8.47) and (8.48). S_3^1 is defined by

$$S_3^1 = \frac{\eta_1\lambda_3}{\mu_1 - \eta_1\lambda_3} \tag{8.54}$$

In Q_2, type-3 packets are served after type-2 packets. Hence,

$$T_3^2 = \frac{S_2^2 + S_3^2}{\eta_2\lambda_3} \tag{8.55}$$

with S_i^2, $i = 2,3$ being the average number of type-i packets in Q_2. S_2^2 is previously defined in (8.50) and S_3^2 is given by

$$S_3^2 = \frac{\eta_2\lambda_3}{\mu_1 - \eta_2\lambda_3} \tag{8.56}$$

Q_3 serves only type-3 packets, and hence their MST can be expressed as follows:

$$T_3^3 = \frac{S_3^3}{\eta_3 \lambda_3}, \tag{8.57}$$

where S_3^3 is the average number of type-3 packets in Q_3, given by

$$S_3^3 = \frac{\eta_3 \lambda_3}{\mu_3 - \eta_3 \lambda_3} \tag{8.58}$$

Finally, by substituting in (8.52), we obtain the expression of T_3 as follows:

$$T_3 = \frac{\lambda_1}{\lambda_3(\mu_1 - \lambda_1)} + \frac{\alpha\lambda_2}{\lambda_3(\mu_1 - \alpha\lambda_2)} + \frac{\eta_1}{\mu_1 - \eta_1\lambda_3} + \frac{(1-\alpha)\lambda_2}{\lambda_3(\mu_2 - (1-\alpha)\lambda_2)} + \frac{\eta_2}{\mu_1 - \eta_2\lambda_3} + \frac{\eta_3}{\mu_3 - \eta_3\lambda_3} \tag{8.59}$$

8.6 Numerical Evaluation and Discussion

This section analyzes and evaluates the optimal policy and the proposed properties. To obtain the stationary policy and associated optimal discounted cost, we use the VI algorithm to solve the dynamic programming equation (8.30), as depicted in Algorithm 8.1. The algorithm considers a finite set of the system states, the admissible actions, the rates of arrival and departure, the routing costs, the discount factor, and the stopping accuracy. The discounted cost is initiated with the value zero for all system states. Then at each iteration, it is updated using the dynamic programming equation. The optimal actions are saved as per the current policy. Therefore, we evaluate the stopping test using the updated differences in the discounted cost. Once the stopping test is verified, we stop the iterations process and save the current results as the optimal stationary policy and the optimal discounted cost.

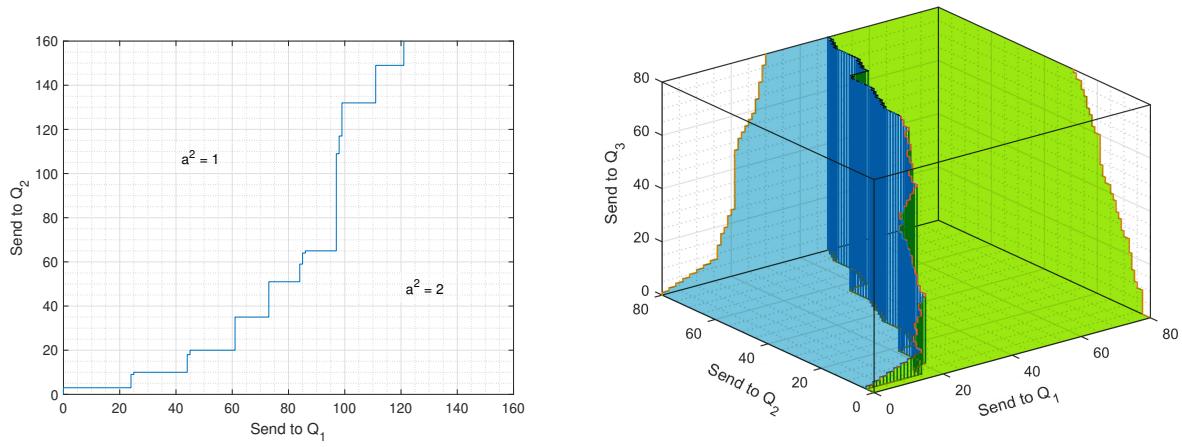
In the following results, we consider the following system setup: $I = \{0,1,2,3\}^6$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $\mu_1 = 8$, $\mu_2 = 6$, $\mu_3 = 4$, $\beta = 0.99$, $\varepsilon = 10^{-3}$. Figure 8.2 illustrates the optimal policy for the system when configured with costs: $c_1^1 = 5$, $c_1^2 = 4$, $c_2^1 = 5$, $c_2^2 = 4$, and $c_3 = 3$. Such a policy is composed of two optimal control decisions for type-2 and type-3 packets. Figure 8.2a depicts the optimal decisions to forward arriving packets of type-2 to either Q_1 or Q_2. We characterized by $a^2 = 1$ and $a^2 = 2$ the sending region of type-2 packets to Q_1 and Q_2, respectively. We observe that the switching curve between these two regions is monotonic, which validates the characterization of the structural property (P2) introduced in Section 8.5.2.

Furthermore, we present in Figure 8.2b the optimal control decisions for the arriving type-3 packets. The volume in green characterizes the sending region of type-3 packets to Q_1, the volume in blue characterizes the sending region of type-3 packets to Q_2, and the remaining volume characterizes the sending region of type-3 packets to Q_3. We observe that there is a

Algorithm 8.1 Value Iteration for routing control of multiclass traffic

Input: $I \in E$: Finite set of system states.
D: Set of admissible actions.
$\lambda_1, \lambda_2, \lambda_3$: Arrival rates.
μ_1, μ_2, μ_3: Service rates.
$c_1^1, c_1^2, c_2^1, c_2^2, c_3$: Routing costs.
β: Discount factor.
ε: Stopping accuracy.
Output: Optimal DSP π_n and optimal discounted costs V_n.

1: $V_0(i) \leftarrow 0$ for every $i \in I$
2: $n \leftarrow 1$
3: **for** every $i \in I$ **do**
4: $\quad V_n^\beta(i) \leftarrow \min_{a \in A(i)} \left\{ c(i,a) + \sum_{j \in I} \Pr\{j \mid i,a\} V_{n-1}(i) \right\}$
5: **end for**
6: $\pi_n \leftarrow \{a^*(i)\}_{\{i \in I\}}$, the resulting optimal DSP.
7: $m_n \leftarrow \min_{i \in I} \left\{ V_n^\beta(i) - V_{n-1}(i) \right\}$
8: $M_n \leftarrow \max_{i \in I} \left\{ V_n^\beta(i) - V_{n-1}(i) \right\}$
9: **if** $M_n - m_n \leq \varepsilon m_n$ **then**
10: $\quad$ Stop with policy π_n.
11: **else**
12: $\quad n \leftarrow n+1$ and go to Step 3.
13: **end if**
14: **return** π_n and V_n.



(a) Optimal control decisions for type-2 packets (b) Optimal control decisions for type-3 packets

Figure 8.2: Illustration of the optimal policy for $c_1^1 = 5$, $c_1^2 = 4$, $c_2^1 = 5$, $c_2^2 = 4$, and $c_3 = 3$.

switching surface that separates those regions. Additionally, the projection of the surface that separates each two sending regions to Q_i and Q_j $(i, j = 1,2,3,\ i \neq j)$ constructs a monotonic curve. Hence, validating our claim for the characterization of the structural properties (P3), (P4) and (P5) in Section 8.5.2.

Hereafter, we investigate the impact of the routing costs of arriving packets. Figure 8.3 shows the impact of the routing costs of type-2 packets. We evaluate the routing percentage

Conclusion and Perspectives

Conclusion

Throughout this thesis, we established modeling, evaluation, and optimization of IEEE 802.11ah-based networks. We mainly focus on developing new and accurate mathematical models for the periodic channel reservation introduced in IEEE 802.11ah through its new channel access mechanism, RAW. We first presented an overview of the IEEE 802.11ah standard and existing related research in the literature. Additionally, we presented some stochastic tools deployed in through our developed models. Since the legacy Markov chain models are not applicable for such scenarios of time-limited contention, we developed new adaptive analytical models with different stochastic approaches based on Renewal theory, such as arrival processes, counting processes, and renewal processes. Therefore, we develop several mathematical models for different considered network conditions. As IoT networks are typically composed of battery-powered devices, and reliable connectivity for a large number of devices requires good throughput, we mainly consider throughput and energy efficiency metrics for the performance evaluation of our proposals.

We first considered saturated traffic and ideal channel conditions to explore the maximum achievable performance of time-limited channel access under the RAW mechanism. Then, we examined the case of non-ideal channel conditions with transmission errors. We modeled the contention within the RAW slot period by representing transmissions as arrival to the channel timeline and developing a counting process to track the occurred transmissions continuously in terms of contention time. The proposed approach allows us to identify the different events that occurred during the RAW slot period. We then evaluate the performance of a RAW slot and a RAW consisting of several RAW slots.

Based on the closed-form results provided by this model, we proposed a RAW grouping scheme and an efficient RAW configuration algorithm maximizing the achievable throughput of the RAW grouping scheme. Our proposal proves effective in comparison to the legacy DCF protocol.

We additionally modeled a more realistic scenario for IoT networks deployed in dense urban environments by considering Rayleigh-fading channel conditions with the presence of

the capture effect. We considered a single-hop scenario of stations randomly distributed around the AP and the power attenuation of transmitted packets, and we modeled RAW-based channel access under capture awareness. Therefore, we evaluate the performance of both a single RAW slot and a RAW consisting of several RAW slots. Considering the power attenuation under the Rayleigh channel, we examined the impact of several parameters, including the distance between stations and AP, the capture threshold, assigned stations, RAW/RAW slot duration, and the number of groups within a RAW. Since the time-constrained contention within the RAW periods results in fluctuating behaviors of its performance, we developed a resource-efficient scheme to reconfigure RAW by eliminating the extra time wasted during each RAW slot period. The new configuration results in higher RAW throughput and energy efficiency while reducing channel utilization.

Moreover, we developed a mathematical model for unsaturated IEEE 802.11ah-based networks operating under a Rayleigh-fading channel with capture enabled. Considering the spatial distribution of stations, we proposed a LACA algorithm for RAW slot periods, ensuring the delivery of all packets designated stations carry. This scheme has proven to be effective in maximizing both PDR and CU, which is practical in enhancing the AoI at the AP for time-sensitive applications.

Finally, we developed an MDP model for multiclass traffic routing control within a relay in an IEEE 802.11ah network. We considered a multichannel relay with forwarding service priority in terms of the transmission data rate. Then, we derived a deterministic stationary optimal policy for the proposed routing control mechanism, which remarkably reduces the latency of packets within the relay.

To validate the accuracy of the analytical frameworks, we developed a discrete-event simulator on MATLAB software. The simulator mimics channel access using the RAW mechanism, integrating the network conditions of the scenario considered. Hereafter, all developed mathematical models were meticulously validated through extensive campaigns of simulations. Since our models are time-constrained, we further validated their accuracy via the convergence to the stationary cases given by Markov chain based approaches. Additionally, we carried out for each model several numerical results along with discussions and analysis of parameters that influence overall system performance, including network setup and channel conditions. Although we focus on the IEEE 802.11ah standard, the developed models apply to all alternative standards that integrate periodic channel reservations (e.g., .11aa, .11ad, and 11.ax). These models can be deployed under the considered scenarios and can also be implemented in global and practical scheduling schemes in heterogeneous dense networks.

Future work

Although there are several efficient communication technologies today, the fast growth of smart devices and IoT applications results in several limitations for deploying those technologies.

Thus, several amendments for different applications are in the development process. However, most new standards explain how their features should work without providing explicit deployment schemes. Since medium access protocols and other network setup parameters are crucial to overall network performance, it is essential to determine the best configuration and tune up the network parameters for every use case. However, such tasks are not obvious, as it is challenging to understand the behavior of a wireless network and assess its performance in large-scale scenarios. Deploying field trials for testing purposes is very costly, and running simulations is time-consuming due to the high complexity of systems. Thus, a practical and adequate alternative is to evaluate the network performance analytically using adaptive mathematical approaches. Analytical evaluations of network performance are not only limited to theoretical studies but are also significantly necessary for practical purposes such as optimal configuration, real-time resource management, and development. Stochastic approaches such as stochastic processes, Markov chains, Markov Decision Processes, and Game theory are suitable for many wireless network behaviors.

Several challenges and problems of IoT solutions need further attention for further enhancements. Since devices compete for channel access during RAW periods, some stations may not be able to transmit all their data. Therefore, we suggest extending the analytical models presented in this thesis to cover the dynamic conditions of the network in terms of buffered data in the stations. In addition, capture-based scheduling schemes for channel access will effectively enhance network performance if stations are grouped to maximize the capture of colliding packets.

The study of the distribution of stations around an urban environment can be modeled using the Stochastic Geometry (SG) approach. SG is a suitable approach to modeling the spatial distribution of large-scale networks. It represents a network by a stationary random model in the entire Euclidean plane or space and analyzes it probabilistically. The locations of stations are seen as the realizations of some point processes, and the probabilistic analysis provides a way of estimating spatial averages such as connectivity, stability, and capacity, etc.

The latency is an essential metric for many applications, and it is commonly presented by packet delay. However, time-sensitive applications such as telesurgery, autonomous driving, and UAV networks are prioritizing the metric of AoI. The AoI measures the freshness of information at the receiver, which is the time since the last received update was generated. AoI has recently gained more attention from many researchers and is commonly modeled using the queuing theory approaches. Therefore, the models proposed throughout this thesis can be a basis for developing adaptive models to optimize AoI for IoT communication technologies such as IEEE 802.11ah.

Conclusion et Perspectives

Conclusion

Tout au long de cette thèse, nous avons établi la modélisation, l'évaluation et l'optimisation des réseaux basés sur IEEE 802.11ah. Nous nous concentrons principalement sur le développement de nouveaux modèles mathématiques précis pour la réservation périodique des canaux introduite dans la norme IEEE 802.11ah à travers son nouveau mécanisme d'accès aux canaux, RAW. Nous avons d'abord présenté un aperçu de la norme IEEE 802.11ah et des recherches connexes existantes dans la littérature. En outre, nous avons présenté certains outils stochastiques déployés dans le cadre de nos modèles développés. Puisque les modèles de chaîne de Markov existants ne sont pas applicables à de tels scénarios de contention limitée dans le temps, nous avons développé de nouveaux modèles analytiques adaptatifs avec différentes approches stochastiques basées sur la théorie du renouvellement, telles que les processus d'arrivée, les processus de comptage et les processus de renouvellement. Par conséquent, nous développons plusieurs modèles mathématiques pour différentes conditions de réseau considérées. Comme les réseaux IoT sont généralement composés de dispositifs alimentés par batterie, et qu'une connectivité fiable pour un grand nombre de dispositifs nécessite un bon débit, nous considérons principalement les mesures de débit et d'efficacité énergétique pour l'évaluation des performances de nos propositions.

Nous avons d'abord considéré un trafic saturé et des conditions de canal idéales pour explorer la performance maximale réalisable de l'accès au canal limité dans le temps sous le mécanisme RAW. Ensuite, nous avons examiné le cas de conditions de canal non idéales avec des erreurs de transmission. Nous avons modélisé la contention dans la période du slot RAW en représentant les transmissions comme des arrivées sur la ligne temporelle du canal et en développant un processus de comptage pour suivre en continu les transmissions survenues en termes de temps de contention. L'approche proposée nous permet d'identifier les différents événements qui se sont produits pendant la période du créneau RAW. Nous évaluons ensuite les performances d'un slot RAW et d'un RAW composé de plusieurs slots RAW.

Sur la base des résultats à forme fermée fournis par ce modèle, nous avons proposé un schéma de regroupement RAW et un algorithme de configuration RAW efficace maximisant

le débit réalisable du schéma de regroupement RAW. Notre proposition s'avère efficace par rapport au protocole DCF.

Nous avons en outre modélisé un scénario plus réaliste pour les réseaux IoT déployés dans des environnements urbains denses en considérant les conditions de canal à évanouissement de Rayleigh avec la présence de l'effet de capture. Nous avons considéré un scénario à un seul saut de stations distribuées aléatoirement autour de l'AP et l'atténuation de la puissance des paquets transmis, et nous avons modélisé l'accès au canal basé sur RAW sous la présence de l'effet de capture. Par conséquent, nous évaluons les performances d'un seul créneau RAW et d'un RAW composé de plusieurs créneaux RAW. Compte tenu de l'atténuation de la puissance dans le canal de Rayleigh, nous avons examiné l'impact de plusieurs paramètres, notamment la distance entre les stations et l'AP, le seuil de capture, les stations assignées, la durée du slot RAW/RAW et le nombre de groupes dans un RAW. Puisque la contention limitée dans le temps au sein des périodes RAW entraîne des comportements fluctuants de sa performance, nous avons développé un schéma efficace en termes de ressources pour reconfigurer RAW en éliminant le temps supplémentaire perdu au cours de chaque période de slot RAW. La nouvelle configuration permet d'obtenir un débit RAW et une efficacité énergétique plus élevés tout en réduisant l'utilisation des canaux.

De plus, nous développons un modèle mathématique pour les réseaux non saturés basés sur IEEE 802.11ah fonctionnant sous un canal à évanouissement de Rayleigh avec capture activée. En tenant compte de la distribution spatiale des stations, nous avons proposé un algorithme d'allocation des canaux en fonction de la charge (LACA) pour les périodes de slot RAW, assurant la livraison de tous les paquets que les stations désignées transportent. Ce schéma s'est montré efficace pour maximiser à la fois le rapport de livraison des paquets (PDR) et l'utilisation du canal (CU), ce qui est pratique pour améliorer l'âge de l'information (AoI) au AP pour les applications sensibles au temps.

Finalement, nous avons développé un modèle MDP pour le contrôle du routage du trafic multiclasse au sein d'un relais dans un réseau IEEE 802.11ah. Nous avons considéré un relais multicanal avec une priorité de service d'acheminement en termes de débit de transmission. Ensuite, nous avons dérivé une politique optimale stationnaire déterministe pour le mécanisme de contrôle de routage proposé, qui réduit remarquablement la latence des paquets dans le relais.

Pour valider l'exactitude des modèles analytiques, nous avons développé un simulateur à événements discrets sur le logiciel MATLAB. Le simulateur imite l'accès aux canaux à l'aide du mécanisme RAW, en intégrant les conditions de réseau du scénario considéré. Par la suite, tous les modèles mathématiques développés ont été méticuleusement validés par de vastes campagnes de simulations. Puisque nos modèles sont contraints par le temps, nous avons validé davantage leur précision via la convergence vers les cas stationnaires donnés par les approches basées sur les chaînes de Markov. En outre, nous avons réalisé pour chaque modèle plusieurs résultats numériques ainsi que des discussions et des analyses des paramètres qui

influencent les performances globales du système, notamment la configuration du réseau et les conditions du canal. Bien que nous nous concentrions sur la norme IEEE 802.11ah, les modèles développés s'appliquent à toutes les normes alternatives qui intègrent des réservations périodiques de canaux (par exemple, .11aa, .11ad et 11.ax). Ces modèles peuvent être déployés dans les scénarios considérés et peuvent également être mis en œuvre dans des schémas d'ordonnancement globaux et pratiques dans les réseaux denses hétérogènes.

Travaux futurs

Bien qu'il existe aujourd'hui plusieurs technologies de communication efficaces, la croissance rapide des appareils intelligents et des applications IoT entraîne plusieurs limitations pour le déploiement de ces technologies. Ainsi, plusieurs amendements pour différentes applications sont en cours de développement. Cependant, la plupart des nouvelles normes expliquent comment leurs fonctionnalités doivent fonctionner sans fournir de schémas de déploiement explicites. Étant donné que les protocoles d'accès au support et les autres paramètres de configuration du réseau sont cruciaux pour les performances globales du réseau, il est essentiel de déterminer la meilleure configuration et de régler les paramètres du réseau pour chaque cas d'utilisation. Cependant, ces tâches ne sont pas évidentes, car il est difficile de comprendre le comportement d'un réseau sans fil et d'évaluer ses performances dans des scénarios à grande échelle. Le déploiement d'essais sur le terrain à des fins de test est très coûteux, et l'exécution de simulations prend beaucoup de temps en raison de la grande complexité des systèmes. Ainsi, une alternative pratique et adéquate consiste à évaluer les performances du réseau de manière analytique en utilisant des approches mathématiques adaptatives. Les évaluations analytiques de la performance du réseau ne se limitent pas seulement aux études théoriques, mais sont également très nécessaires à des fins pratiques telles que la configuration optimale, la gestion des ressources en temps réel et le développement. Les approches stochastiques telles que les processus stochastiques, les chaînes de Markov, les processus de décision de Markov et la théorie des jeux conviennent à de nombreux comportements de réseaux sans fil.

Plusieurs défis et problèmes des solutions IoT nécessitent une attention particulière pour être améliorés. Étant donné que les dispositifs se disputent l'accès aux canaux pendant les périodes RAW, certaines stations peuvent ne pas être en mesure de transmettre toutes leurs données. Par conséquent, nous suggérons d'étendre les modèles analytiques présentés dans cette thèse pour couvrir les conditions dynamiques du réseau en termes de données tamponnées dans les stations. En outre, les schémas d'ordonnancement basés sur la capture pour l'accès aux canaux amélioreront efficacement les performances du réseau si les stations sont regroupées pour maximiser la capture des paquets entrant en collision.

L'étude de la distribution des stations dans un environnement urbain peut être modélisée en utilisant l'approche de la géométrie stochastique (SG). La SG est une approche adaptée à

la modélisation de la distribution spatiale des réseaux à grande échelle. Elle représente un réseau par un modèle aléatoire stationnaire dans l'ensemble du plan ou de l'espace euclidien et l'analyse de manière probabiliste. Les emplacements des stations sont considérés comme les réalisations de certains processus ponctuels, et l'analyse probabiliste fournit un moyen d'estimer les moyennes spatiales telles que la connectivité, la stabilité et la capacité, etc.

La latence est une métrique essentielle pour de nombreuses applications, et elle est généralement présentée par le retard des paquets. Cependant, les applications sensibles au temps, comme la téléchirurgie, la conduite autonome et les réseaux de véhicules aériens sans pilote (UAV), donnent la priorité à la métrique de l'AoI. L'AoI mesure la fraîcheur des informations au niveau du récepteur, c'est-à-dire le temps écoulé depuis la génération de la dernière mise à jour reçue. L'AoI a récemment fait l'objet d'une attention accrue de la part de nombreux chercheurs et est généralement modélisé à l'aide des approches de la théorie des files d'attente. Par conséquent, les modèles proposés tout au long de cette thèse peuvent servir de base au développement de modèles adaptatifs pour optimiser l'AoI pour les technologies de communication IoT telles que IEEE 802.11ah.

Conclusión y Perspectivas

Conclusión

A lo largo de esta tesis se han desarrollado modelos estocásticos con el finde de evaluar las prestaciones y la optimización del funcionamiento de redes IEEE 802.11ah. La tesis se ha centrado principalmente en el desarrollo de modelos matemáticos nuevos y precisos para la reserva periódica de canales introducida en IEEE 802.11ah mediante su nuevo mecanismo de acceso a canales, RAW. En primer lugar, se ha presentado una visión general del estándar IEEE 802.11ah y de las investigaciones relacionadas existentes en la literatura. Además, se presentan algunas herramientas estocásticas utilizadas en el desarrollo de nuestros modelos. Dado que los modelos de cadena de Markov heredados no son aplicables a estos escenarios de contención limitada en el tiempo, se han creado nuevos modelos analíticos adaptativos con diferentes enfoques estocásticos basados en la teoría de la renovación, como los procesos de llegada, los procesos de recuento y los procesos de renovación. Por lo tanto, se han desarrollado varios modelos matemáticos para diferentes condiciones de red consideradas. Como las redes de IoT suelen estar compuestas por dispositivos alimentados por baterías, y la conectividad fiable para un gran número de dispositivos requiere un buen rendimiento, se ha considerado principalmente las métricas de rendimiento y eficiencia energética para la evaluación del rendimiento de nuestras propuestas.

En primer lugar, se ha considerado el tráfico saturado y las condiciones ideales del canal para explorar el rendimiento máximo alcanzable del acceso al canal limitado en el tiempo bajo el uso del mecanismo RAW. A continuación, se ha examinado el caso de condiciones de canal no ideales con errores de transmisión. Se ha proseguido con el modelado de la contención dentro del periodo de la ranura RAW representando las transmisiones como una llegada a la línea de tiempo del canal y desarrollando un proceso de recuento para rastrear las transmisiones ocurridas de forma continua en términos de tiempo de contención. El enfoque propuesto nos ha de permitir identificar los diferentes eventos ocurridos durante el periodo de ranura RAW. A continuación, se ha evaluado el rendimiento de una ranura RAW y de una RAW formada por varias ranuras RAW.

Basándonos en los resultados de forma cerrada proporcionados por este modelo, se he

propuesto un esquema de agrupación de RAW y un algoritmo de configuración de RAW eficiente que maximiza el rendimiento alcanzable del esquema de agrupación de RAW. Nuestra propuesta resulta eficaz en comparación con el protocolo DCF heredado.

Además, se ha modelado un escenario más realista para las redes IoT desplegadas en entornos urbanos densos considerando las condiciones de canal con desvanecimiento del tipo Rayleigh con la presencia del efecto de captura. Se ha considerado un escenario de un solo salto de estaciones distribuidas aleatoriamente alrededor del AP y la atenuación de la potencia de los paquetes transmitidos, y modelamos el acceso al canal basado en RAW bajo la conciencia de captura. Por lo tanto, se ha evaluado el rendimiento tanto de una única ranura RAW como de una RAW formada por varias ranuras RAW. Teniendo en cuenta la atenuación de la potencia bajo el canal Rayleigh, examinamos el impacto de varios parámetros, incluyendo la distancia entre las estaciones y el AP, el umbral de captura, las estaciones asignadas, la duración de la ranura RAW y el número de grupos dentro de un RAW. Dado que la contención por tiempo dentro de los periodos RAW da lugar a comportamientos fluctuantes de su rendimiento, se ha diseñado un esquema de eficiencia de recursos para reconfigurar el RAW eliminando el tiempo extra perdido durante cada periodo de ranura RAW. La nueva configuración da como resultado un mayor rendimiento de RAW y una mayor eficiencia energética, al tiempo que reduce la utilización del canal.

Además, se desarrolla un modelo matemático para redes no saturadas basadas en IEEE 802.11ah que operan bajo un canal con desvanecimiento de Rayleigh con captura habilitada. Teniendo en cuenta la distribución espacial de las estaciones, se propone un algoritmo de asignación de canales en función de la carga (LACA) para los periodos de ranura RAW, que garantiza la entrega de todos los paquetes que las estaciones designadas llevan. Este esquema ha demostrado ser eficaz para maximizar tanto la relación de entrega de paquetes (PDR) como el uso del canal (CU), lo que resulta práctico para mejorar la edad de la información (AoI) en el AP para las aplicaciones sensibles al tiempo.

Finalmente, desarrollamos un modelo MDP para el control de enrutamiento de tráfico multiclase dentro de un relé en una red IEEE 802.11ah. Consideramos un relé multicanal con prioridad de servicio de reenvío en función de la velocidad de transmisión de datos. A continuación, derivamos una política óptima estacionaria determinista para el mecanismo de control de enrutamiento propuesto, que reduce notablemente la latencia de los paquetes dentro del relé.

Para validar la precisión de los marcos analíticos, se ha desarrollado un simulador de eventos discretos en el software MATLAB. El simulador emula el acceso al canal mediante el mecanismo RAW, integrando las condiciones de red del escenario considerado. A continuación, todos los modelos matemáticos desarrollados se validaron meticulosamente mediante extensas campañas de simulaciones. Dado que nuestros modelos están limitados por el tiempo, se validan además su exactitud mediante la convergencia a los casos estacionarios dada por los enfoques basados en la cadena de Markov. Además, se han realizado para cada modelo

varios resultados numéricos junto con discusiones y análisis de los parámetros que influyen en el rendimiento general del sistema, incluyendo la configuración de la red y las condiciones del canal. Aunque el estudio se centra en el estándar IEEE 802.11ah, los modelos desarrollados pueden ser utilizados en le estudio de otros estándares que integren reservas de canal periódicas, tales como los estándares, .11aa, .11ad y 11.ax. Estos modelos pueden desplegarse en los escenarios considerados y también pueden implementarse en esquemas de programación globales y desrarollos de redes densas heterogéneas.

Trabajo futuro

Aunque hoy en día existen varias tecnologías de comunicación eficientes, el rápido crecimiento de los dispositivos inteligentes y de las aplicaciones IoT ha dado lugar a varias retos para el despliegue de dichas tecnologías. Al día de hoy, se vienen desarrollando varias enmiendas a los estándares con el fin de dar respuesta a las exigencias de nuevas aplicaciones. Las nuevas normas vienen incorporando nuevos mecanismos, pero sin precisar los detalles de su implementación. Dado que los protocolos de acceso al medio y otros parámetros de configuración de la red son cruciales para el rendimiento general de la misma, es esencial determinar la mejor configuración y afinar los parámetros de la red para cada caso de uso. Sin embargo, estas tareas no son obvias, ya que es un reto comprender el comportamiento de una red inalámbrica y evaluar su rendimiento en escenarios a gran escala. El despliegue de pruebas sobre el terreno para comprobarlo es muy costoso, y la realización de simulaciones requiere mucho tiempo debido a la gran complejidad de los sistemas. Por ello, una alternativa práctica y adecuada es evaluar el rendimiento de la red de forma analítica mediante enfoques matemáticos adaptativos. Las evaluaciones analíticas del rendimiento de la red no se limitan a los estudios teóricos, sino que también son significativamente necesarias para fines prácticos como la configuración óptima, la gestión de recursos en tiempo real y el desarrollo. Los enfoques probabilísticos, como los procesos estocásticos, las cadenas de Markov, los procesos de decisión de Markov y la teoría de juegos, son adecuados para muchos comportamientos de las redes inalámbricas.

Varios de los retos en el desarrollo delos sistemas IoT requieren una mayor atención para habilitar su despliegue. Dado que los dispositivos compiten por el acceso al canal durante los periodos RAW, es posible que algunas estaciones no puedan transmitir todos sus datos. Por lo tanto, una de las tareas inmediatas a abordar deberá enfocarse en el desarrollo de los modelos analíticos presentados en esta tesis para cubrir las condiciones dinámicas de la red en términos de datos almacenados en buffer en las estaciones. Además, los esquemas de programación basados en la captura para el acceso al canal mejorarán eficazmente el rendimiento de la red si las estaciones se agrupan para maximizar la captura de paquetes en colisión.

El estudio de la distribución de las estaciones en un entorno urbano puede modelarse mediante el enfoque de la geometría estocástica (SG). La SG es un enfoque adecuado para mode-

lar la distribución espacial de las redes a gran escala. Representa una red mediante un modelo aleatorio estacionario en todo el plano o espacio euclidiano y la analiza de forma probabilística. Las ubicaciones de las estaciones se ven como las realizaciones de unos procesos puntuales, y el análisis probabilístico proporciona una forma de estimar los promedios espaciales como la conectividad, la estabilidad y la capacidad, etc.

La latencia es una métrica esencial para muchas aplicaciones, y se presenta habitualmente mediante el retardo de los paquetes. Sin embargo, las aplicaciones sensibles al tiempo, como la telecirugía, la conducción autónoma y las redes de vehículos aéreos no tripulados (UAV), están dando prioridad a la métrica de la AoI. La AoI mide la frescura de la información en el receptor, que es el tiempo transcurrido desde su generación hasta la última actualización recibida. El AoI ha ganado recientemente más atención por parte de muchos investigadores y se suele modelar utilizando los enfoques de la teoría de colas. Por lo tanto, los modelos propuestos a lo largo de esta tesis pueden servir de base para desarrollar modelos adaptativos que optimicen el AoI para las tecnologías de comunicación del IoT, como el IEEE 802.11ah.

List of Publications

Research Papers

- **Hamid Taramit**, José Jaime Camacho-Escoto, Javier Gomez, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Accurate Analytical Model and Evaluation of Wi-Fi Halow Based IoT Networks under a Rayleigh-Fading Channel with Capture." Mathematics, 10(6): 952, 2022. https://doi.org/10.3390/math10060952

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "A renewal theory based performance and configuration framework of the IEEE 802.11 ah RAW mechanism." Digital Communications and Networks, 9(1):236-251, 2023. https://doi.org/10.1016/j.dcan.2022.02.009

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Energy Efficiency Framework for Time-limited Contention in the IEEE 802.11 ah Standard." In 2021 IEEE Globecom Workshops (GC Wkshps), pp. 1-6. IEEE, 2021. https://doi.org/10.1109/GCWkshps52748.2021.9681998

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Adaptive Channel Allocation for Wi-Fi HaLow Networks." In 2022 IEEE Conference on Standards for Communications and Networking (CSCN), pp. 211. IEEE, 2022. https://doi.org/10.1109/CSCN57023.2022.10051003

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Resource and Energy-Efficient Configuration of IEEE 802.11ah Networks under Rayleigh Channels." In 4th IEEE Middle East and North Africa COMMunications Conference (MENACOMM), pp. 177-184. IEEE, 2022. https://doi.org/10.1109/MENACOMM57252.2022.9998101

- **Hamid Taramit**, José Jaime Camacho-Escoto, Javier Gomez, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Load-Aware Channel Allocation for IEEE 802.11ah-Based Networks." IEEE Access, 2023. https://doi.org/10.1109/ACCESS.2023.3251896

Oral Communications

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Resource and Energy-Efficient Configuration of IEEE 802.11ah Networks under Rayleigh Channels". In 2022 4th IEEE Middle East and North Africa COMMunications Conference (MENACOMM), December 6-8, 2022. Amman, Jordan.
- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Adaptive Channel Allocation for Wi-Fi HaLow Networks". In 2022 IEEE Conference on Standards for Communications and Networking (CSCN), November 28-30, 2022. Thessaloniki, Greece.
- **Hamid Taramit**, José Jaime Camacho-Escoto, Javier Gomez, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Performance Evaluation of Wi-Fi HaLow Networks under Rayleigh Channel with Capture". In COST Action CA20120 - INTERACT 2nd Technical Meeting, June 13-15, 2022. Lyon, France.
- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Energy efficiency framework for time-limited contention in the IEEE 802.11ah standard". In COST Action CA20120 - INTERACT 1st Technical Meeting, February 8-11, 2022. Bologna, Italy.
- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Energy Efficiency Framework for Time-limited Contention in the IEEE 802.11 ah Standard". In 2021 IEEE Global Communications Conference (GLOBECOM 2021), December 7-11, 2021. Madrid, Spain.
- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Performance Evaluation and Optimization of RAW Mechanism in IEEE 802.11ah Technology". In 7ème Édition Journée Doctorant, May 2, 2019, FSTS, Settat, Morocco.

Posters

- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Perfromance Analysis of RAW Mechanism in IoT-Enabling IEEE 802.11ah Technology". In 6ème Édition Journée Doctorant, April 5, 2018, FSTS, Settat, Morocco.
- **Hamid Taramit**, Luis Orozco-Barbosa, and Abdelkrim Haqiq. "Performance Evaluation and Tuning of a Grouping Scheme for Dense IEEE 802.11ah Networks". In IX Jornadas Doctorales de la Universidad de Castilla-La Mancha, November 12, 2019. Ciudad Real, Spain.

Bibliography

[1] Wafa'a Kassab and Khalid A. Darabkh. A–z survey of internet of things: Architectures, protocols, applications, recent advances, future directions and recommendations. *Journal of Network and Computer Applications*, 163:102663, 2020.

[2] Pawan Kumar Verma, Rajesh Verma, Arun Prakash, Ashish Agrawal, Kshirasagar Naik, Rajeev Tripathi, Maazen Alsabaan, Tarek Khalifa, Tamer Abdelkader, and Abdulhakim Abogharaf. Machine-to-machine (m2m) communications: A survey. *Journal of Network and Computer Applications*, 66:83–105, 2016.

[3] Le Tian, Serena Santi, Amina Seferagić, Julong Lan, and Jeroen Famaey. Wi-fi halow for the internet of things: An up-to-date survey on IEEE 802.11ah research. *Journal of Network and Computer Applications*, 182:103036, 2021.

[4] IEEE standard for information technology–telecommunications and information exchange between systems - local and metropolitan area networks–specific requirements - part 11: Wireless lan medium access control (MAC) and physical layer (PHY) specifications amendment 2: Sub 1 GHz license exempt operation. *IEEE Std 802.11ah-2016 (Amendment to IEEE Std 802.11-2016, as amended by IEEE Std 802.11ai-2016)*, pages 1–594, May 2017.

[5] Giuseppe Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3):535–547, 2000.

[6] Andrea Goldsmith. *Wireless communications*. Cambridge university press, 2005.

[7] Carla-Fabiana Chiasserini, Marco Gribaudo, and Daniele Manini. *Analytical Modeling of Wireless Communication Systems*. John Wiley & Sons, 2016.

[8] Xianghao Yu, Chang Li, Jun Zhang, and Khaled B Letaief. *Stochastic Geometry Analysis of Multi-Antenna Wireless Networks*, volume 178. Springer, 2019.

[9] Robert G Gallager. *Stochastic processes: theory for applications*. Cambridge University Press, 2013.

[10] Abraham Wald. Sequential analysis. *john wiley & sons, New York, NY*, 1947.

[11] Vikram Krishnamurthy. *Partially observed Markov decision processes*. Cambridge university press, 2016.

[12] Md Hussain, Zaved Iqubal Ahmed, Nityananda Sarma, Dilip Kr Saikia, et al. An efficient tdma mac protocol for multi-hop wifi-based long distance networks. *Wireless Personal Communications*, 86(4):1971–1994, 2016.

[13] Matteo Ridolfi, Samuel Van de Velde, Heidi Steendam, and Eli De Poorter. Wifi ad-hoc mesh network and mac protocol solution for uwb indoor localization systems. In *2016 Symposium on Communications and Vehicular Technologies (SCVT)*, pages 1–6, 2016.

[14] Murad Murad and Ahmed M. Eltawil. A simple full-duplex mac protocol exploiting asymmetric traffic loads in wifi systems. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2017.

[15] Shiying Han, Ying-Chang Liang, Qian Chen, and Boon-Hee Soong. Licensed-assisted access for lte in unlicensed spectrum: A mac protocol design. *IEEE Journal on Selected Areas in Communications*, 34(10):2550–2561, 2016.

[16] Gang Wang and Yanyuan Qin. Mac protocols for wireless mesh networks with multi-beam antennas: A survey. In *Future of Information and Communication Conference*, pages 117–142. Springer, 2019.

[17] Ieee standard for local and metropolitan area networks–part 15.4: Low-rate wireless personal area networks (lr-wpans). *IEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006)*, pages 1–314, 2011.

[18] Carles Gomez, Joaquim Oller, and Josep Paradells. Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors*, 12(9):11734–11753, 2012.

[19] Usman Raza, Parag Kulkarni, and Mahesh Sooriyabandara. Low power wide area networks: An overview. *IEEE Communications Surveys & Tutorials*, 19(2):855–873, 2017.

[20] Victor Baños Gonzalez, M. Shahwaiz Afaqui, Elena López Aguilera, and Eduard García Villegas. IEEE 802.11ah: A technology to face the IoT challenge. *Sensors*, 16(11), 2016.

[21] Luis Felipe Del Carpio, Piergiuseppe Di Marco, Per Skillermark, Roman Chirikov, and Karin Lagergren. Comparison of 802.11 ah, ble and 802.15. 4 for a home automation use case. *International Journal of Wireless Information Networks*, 24(3):243–253, 2017.

[22] Nurzaman Ahmed, Hafizur Rahman, and Md I Hussain. A comparison of 802.11ah and 802.15.4 for iot. *ICT Express*, 2(3):100 – 102, 2016. Special Issue on ICT Convergence in the Internet of Things (IoT).

[23] C Kuhlins, B Rathonyi, A Zaidi, and M Hogan. Cellular networks for massive iot. *Ericsson White Paper*, 1217, 2020.

[24] Qi Pan, Xiangming Wen, Zhaoming Lu, Wenpeng Jing, and Linpei Li. Cluster-based group paging for massive machine type communications under 5g networks. *IEEE Access*, 6:64891–64904, 2018.

[25] Evgeny Khorov, Andrey Lyakhov, Alexander Krotov, and Andrey Guschin. A survey on IEEE 802.11ah: An enabling networking technology for smart cities. *Computer Communications*, 58:53 – 69, 2015. Special Issue on Networking and Communications for Smart Cities.

[26] Behnam Badihi Olyaei, Juho Pirskanen, Orod Raeesi, Ali Hazmi, and Mikko Valkama. Performance comparison between slotted ieee 802.15.4 and ieee 802.1 lah in iot based applications. In *2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 332–337, 2013.

[27] Weiping Sun, Munhwan Choi, and Sunghyun Choi. Ieee 802.11 ah: A long range 802.11 wlan at sub 1 ghz. *Journal of ICT Standardization*, 1(1):83–108, 2013.

[28] Dmitry Bankov, Evgeny Khorov, and Andrey Lyakhov. The study of the centralized control method to hasten link set-up in ieee 802.11ah networks. In *Proceedings of European Wireless 2015; 21th European Wireless Conference*, pages 1–6, 2015.

[29] Dmitry Bankov, Evgeny Khorov, and Andrey Lyakhov. The study of the distributed control method to hasten link set-up in ieee 802.11ah networks. In *2016 XV International Symposium Problems of Redundancy in Information and Control Systems (REDUNDANCY)*, pages 13–17, 2016.

[30] Ieee standard for information technology–telecommunications and information exchange between systems–local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 10: Mesh networking. *IEEE Std 802.11s-2011 (Amendment to IEEE Std 802.11-2007 as amended by IEEE 802.11k-2008, IEEE 802.11r-2008, IEEE 802.11y-2008, IEEE 802.11w-2009, IEEE 802.11n-2009, IEEE 802.11p-2010, IEEE 802.11z-2010, IEEE 802.11v-2011, and IEEE 802.11u-2011)*, pages 1–372, 2011.

[31] Y. Zhou, H. Wang, S. Zheng, and Z. Z. Lei. Advances in IEEE 802.11ah standardization for machine-type communications in sub-1GHz WLAN. In *2013 IEEE International Conference on Communications Workshops (ICC)*, pages 1269–1273, June 2013.

[32] M. Park. Ieee 802.11ah: sub-1-ghz license-exempt operation for the internet of things. *IEEE Communications Magazine*, 53(9):145–151, Sep. 2015.

[33] B. Domazetović, E. Kočan, and A. Mihovska. Performance evaluation of IEEE 802.11ah systems. In *2016 24th Telecommunications Forum (TELFOR)*, pages 1–4, Nov 2016.

[34] Yue Zhao, Osman N. C. Yilmaz, and Anna Larmo. Optimizing M2M Energy Efficiency in IEEE 802.11ah. In *2015 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2015.

[35] L. Tian, J. Famaey, and S. Latré. Evaluation of the IEEE 802.11ah restricted access window mechanism for dense iot networks. In *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–9, June 2016.

[36] O. Raeesi, J. Pirskanen, A. Hazmi, J. Talvitie, and M. Valkama. Performance enhancement and evaluation of IEEE 802.11ah multi-access point network using restricted access window mechanism. In *2014 IEEE International Conference on Distributed Computing in Sensor Systems*, pages 287–293, May 2014.

[37] O. Raeesi, J. Pirskanen, A. Hazmi, T. Levanen, and M. Valkama. Performance evaluation of IEEE 802.11ah and its restricted access window mechanism. In *2014 IEEE International Conference on Communications Workshops (ICC)*, pages 460–466, June 2014.

[38] Amina Šljivo, Dwight Kerkhove, Le Tian, Jeroen Famaey, Adrian Munteanu, Ingrid Moerman, Jeroen Hoebeke, and Eli De Poorter. Performance Evaluation of IEEE 802.11ah Networks With High-Throughput Bidirectional Traffic. *Sensors*, 18(2), 2018.

[39] Lei Zheng, Lin Cai, Jianping Pan, and Minming Ni. Performance analysis of grouping strategy for dense ieee 802.11 networks. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 219–224, 2013.

[40] L. Zheng, M. Ni, L. Cai, J. Pan, C. Ghosh, and K. Doppler. Performance analysis of group-synchronized DCF for dense IEEE 802.11 networks. *IEEE Transactions on Wireless Communications*, 13(11):6180–6192, Nov 2014.

[41] T. C. Chang, C. H. Lin, K. C. J. Lin, and W. T. Chen. Load-balanced sensor grouping for IEEE 802.11ah networks. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Dec 2015.

[42] M. Qutab-ud-din, A. Hazmi, B. Badihi, A. Larmo, J. Torsner, and M. Valkama. Performance analysis of IoT-enabling IEEE 802.11ah technology and its RAW mechanism with non-cross slot boundary holding schemes. In *2015 IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–6, June 2015.

[43] N. Nawaz, M. Hafeez, S. A. R. Zaidi, D. C. McLernon, and M. Ghogho. Throughput enhancement of restricted access window for uniform grouping scheme in IEEE 802.11ah. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–7, May 2017.

[44] M Zulfiker Ali, Jelena Mišić, and Vojislav B Mišić. Efficiency of restricted access window scheme of IEEE 802.11 ah under non-ideal channel condition. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 251–256. IEEE, 2018.

[45] M Zulfiker Ali, Jelena Mišić, and Vojislav B Mišić. Performance evaluation of heterogeneous IoT nodes with differentiated QoS in IEEE 802.11 ah RAW mechanism. *IEEE Transactions on Vehicular Technology*, 68(4):3905–3918, 2019.

[46] Evgeny Khorov, Andrey Lyakhov, and Ruslan Yusupov. Two-slot based model of the ieee 802.11ah restricted access window with enabled transmissions crossing slot boundaries. In *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, pages 1–9, 2018.

[47] Evgeny Khorov, Alexander Krotov, Andrey Lyakhov, Ruslan Yusupov, Massimo Condoluci, Mischa Dohler, and Ian Akyildiz. Enabling the internet of things with wi-fi halow—performance evaluation of the restricted access window. *IEEE Access*, 7:127402–127415, 2019.

[48] Evgeny Khorov, Andrey Lyakhov, Ivan Nasedkin, Ruslan Yusupov, Jeroen Famaey, and Ian F. Akyildiz. Fast and reliable alert delivery in mission-critical wi-fi halow sensor networks. *IEEE Access*, 8:14302–14313, 2020.

[49] Elizaveta Zazhigina, Ruslan Yusupov, Evgeny Khorov, and Andrey Lyakhov. Analytical study of periodic restricted access window mechanism for short slots. *Electronics*, 10(5), 2021.

[50] Yanru Wang, Yun Li, Kok Keong Chai, Yue Chen, and John Schormans. Energy-aware adaptive restricted access window for IEEE 802.11ah based networks. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1211–1215, Aug 2015.

[51] Y. Wang, Y. Li, K. K. Chai, Y. Chen, and J. Schormans. Energy-aware adaptive restricted access window for IEEE 802.11ah based smart grid networks. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 581–586, Nov 2015.

[52] Y. Wang, K. K. Chai, Y. Chen, J. Schormans, and J. Loo. Energy-aware restricted access window control with retransmission scheme for IEEE 802.11ah (wi-fi halow) based networks. In *2017 13th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*, pages 69–76, Feb 2017.

[53] E. Khorov, A. Krotov, and A. Lyakhov. Modelling machine type communication in IEEE 802.11ah networks. In *2015 IEEE International Conference on Communication Workshop (ICCW)*, pages 1149–1154, June 2015.

[54] Serena Santi, Le Tian, Evgeny Khorov, and Jeroen Famaey. Accurate energy modeling and characterization of IEEE 802.11ah RAW and TWT. *Sensors*, 19(11), 2019.

[55] Tung-Chun Chang, Chi-Han Lin, Kate Ching-Ju Lin, and Wen-Tsuen Chen. Traffic-Aware Sensor Grouping for IEEE 802.11ah Networks: Regression Based Analysis and Design. *IEEE Transactions on Mobile Computing*, 18(3):674–687, 2019.

[56] Le Tian, Michael Mehari, Serena Santi, Steven Latré, Eli De Poorter, and Jeroen Famaey. Ieee 802.11ah restricted access window surrogate model for real-time station grouping. In *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, pages 14–22, 2018.

[57] Le Tian, Michael T Mehari, Serena Santi, Steven Latré, Eli De Poorter, and Jeroen Famaey. Multi-objective surrogate modeling for real-time energy-efficient station grouping in IEEE 802.11 ah. *Pervasive and Mobile Computing*, 57:33–48, 2019.

[58] Le Tian, Elena Lopez-Aguilera, Eduard Garcia-Villegas, Michael Tetemke Mehari, Eli De Poorter, Steven Latré, and Jeroen Famaey. Optimization-Oriented RAW Modeling of IEEE 802.11ah Heterogeneous Networks. *IEEE Internet of Things Journal*, 6(6):10597–10609, 2019.

[59] Dirk Gorissen, Ivo Couckuyt, Piet Demeester, Tom Dhaene, and Karel Crombecq. A surrogate modeling and adaptive sampling toolbox for computer based design. *Journal of machine learning research.-Cambridge, Mass.*, 11:2051–2055, 2010.

[60] Aleksandr Ometov, Nader Daneshfar, Ali Hazmi, Sergey Andreev, Luis Felipe Del Carpio, Parth Amin, Johan Torsner, Yevgeni Koucheryavy, and Mikko Valkama. System-level analysis of ieee 802.11 ah technology for unsaturated mtc traffic. *International Journal of Sensor Networks*, 26(4):269–282, 2018.

[61] M. Zulfiker Ali, Jelena Misic, and Vojislav B. Misic. Differentiated qos to heterogeneous iot nodes in ieee 802.11ah raw mechanism. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, 2018.

[62] Chien-Erh Weng and Hsing-Chung Chen. The performance evaluation of IEEE 802.11 DCF using Markov chain model for wireless LANs. *Computer Standards & Interfaces*, 44:144–149, 2016.

[63] Hamid Taramit, Luis Orozco Barbosa, and Abdelkrim Haqiq. Energy efficiency framework for time-limited contention in the IEEE 802.11ah standard. In *2021 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2021.

[64] Taramit, Hamid and Orozco-Barbosa, Luis and Haqiq, Abdelkrim. A renewal theory based performance and configuration framework of the IEEE 802.11ah RAW mechanism. *Digital Communications and Networks*, 2022.

[65] Hamid Taramit, José Jaime Camacho-Escoto, Javier Gomez, Luis Orozco-Barbosa, and Abdelkrim Haqiq. Accurate Analytical Model and Evaluation of Wi-Fi Halow Based IoT Networks under a Rayleigh-Fading Channel with Capture. *Mathematics*, 10(6), 2022.

[66] Evgeny Khorov, Andrey Lyakhov, Alexander Ivanov, and Ian F Akyildiz. Modeling of real-time multimedia streaming in wi-fi networks with periodic reservations. *IEEE Access*, 8:55633–55653, 2020.

[67] Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pages 1–2793, 2012.

[68] Evgeny Khorov, Artem Krasilov, Alexander Krotov, and Andrey Lyakhov. Will mcca revive wireless multihop networks? *Computer Communications*, 104:159–174, 2017.

[69] Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 2: Mac enhancements for robust audio video streaming. *IEEE Std 802.11aa-2012 (Amendment to IEEE Std 802.11-2012 as amended by IEEE Std 802.11ae-2012)*, pages 1–162, 2012.

[70] Pei Zhou, Kaijun Cheng, Xiao Han, Xuming Fang, Yuguang Fang, Rong He, Yan Long, and Yanping Liu. Ieee 802.11ay-based mmwave wlans: Design challenges and solutions. *IEEE Communications Surveys & Tutorials*, 20(3):1654–1681, 2018.

[71] Iso/iec/ieee international standard - information technology–telecommunications and information exchange between systems - local and metropolitan area networks–specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 2: Sub 1 ghz license exempt operation. *ISO/IEC/IEEE 8802-11:2018/Amd.2:2019(E)*, pages 1–596, 2019.

[72] Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications amendment 1: Enhancements for high-efficiency wlan. *IEEE Std 802.11ax-2021 (Amendment to IEEE Std 802.11-2020)*, pages 1–767, 2021.

[73] Evgeny Khorov, Anton Kiryanov, Andrey Lyakhov, and Giuseppe Bianchi. A tutorial on ieee 802.11ax high efficiency wlans. *IEEE Communications Surveys & Tutorials*, 21(1):197–216, 2019.

[74] Caihong Kai, Jiaojiao Zhang, Xiangru Zhang, and Wei Huang. Energy-efficient sensor grouping for IEEE 802.11 ah networks with max-min fairness guarantees. *IEEE Access*, 7:102284–102294, 2019.

[75] Xiaomin Zhang. A new method for analyzing nonsaturated IEEE 802.11 DCF networks. *IEEE Wireless Communications Letters*, 2(2):243–246, 2013.

[76] Sami Khairy, Mengqi Han, Lin X Cai, Yu Cheng, and Zhu Han. A renewal theory based analytical model for multi-channel random access in IEEE 802.11 ac/ax. *IEEE Transactions on Mobile Computing*, 18(5):1000–1013, 2019.

[77] Xinhua Ling, Yu Cheng, Jon W Mark, and Xuemin Shen. A renewal theory based analytical model for the contention access period of IEEE 802.15. 4 MAC. *IEEE Transactions on Wireless Communications*, 7(6):2340–2349, 2008.

[78] Richard L Burden, Douglas J Faires, and Annette M Burden. *Numerical Analysis*. Cengage Learning, 10th edition, 2015.

[79] International Data Corporation. IoT Growth Demands Rethink of Long-Term Storage Strategies, says IDC. https://www.idc.com/getdoc.jsp?containerId=prAP46737220, 2020 (accessed 20 September 2020).

[80] Sunitha Safavat, Naveen Naik Sapavath, and Danda B. Rawat. Recent advances in mobile edge computing and content caching. *Digital Communications and Networks*, 6(2):189 – 194, 2020.

[81] Yuan Ai, Mugen Peng, and Kecheng Zhang. Edge computing technologies for internet of things: a primer. *Digital Communications and Networks*, 4(2):77 – 86, 2018.

[82] Andrea Biral, Marco Centenaro, Andrea Zanella, Lorenzo Vangelista, and Michele Zorzi. The challenges of m2m massive access in wireless cellular networks. *Digital Communications and Networks*, 1(1):1 – 19, 2015.

[83] David Malone, Peter Clifford, and Douglas J Leith. On buffer sizing for voice in 802.11 wlans. *IEEE Communications Letters*, 10(10):701–703, 2006.

[84] Ahmed Boujnoui, Luis Orozco Barbosa, and Abdelkrim Haqiq. Performance evaluation and tuning of an IEEE 802.11 audio video multicast collision prevention mechanism. *Wireless Networks*, 2020.

[85] P Chatzimisios, AC Boucouvalas, and V Vitsas. Influence of channel BER on IEEE 802.11 DCF. *Electronics letters*, 39(23):1687–1689, 2003.

[86] Kaveh Pahlavan. *Understanding Communications Networks for Emerging Cybernetics Applications*. Stylus Publishing, LLC, 2021.

[87] Fred Daneshgaran, Massimiliano Laddomada, Fabio Mesiti, and Marina Mondin. Unsaturated throughput analysis of IEEE 802.11 in presence of non ideal transmission channel and capture effects. *IEEE Transactions on Wireless Communications*, 7(4):1276–1286, 2008.

[88] Fred Daneshgaran, Massimiliano Laddomada, Fabio Mesiti, Marina Mondin, and Massimiliano Zanolo. Saturation throughput analysis of IEEE 802.11 in the presence of non ideal transmission channel and capture effects. *IEEE Transactions on Communications*, 56(7):1178–1188, 2008.

[89] Gordon J. Sutton, Ren Ping Liu, and Iain B. Collings. Modelling IEEE 802.11 dcf heterogeneous networks with rayleigh fading and capture. *IEEE Transactions on Communications*, 61(8):3336–3348, 2013.

[90] Zebo Yang, Ali Ghubaish, Devrim Unal, and Raj Jain. Factors affecting the performance of sub-1 ghz iot wireless networks. *Wireless Communications and Mobile Computing*, 2021, 2021.

[91] Davide Cocco and Massimiliano Giona. Generalized counting processes in a stochastic environment. *Mathematics*, 9(20), 2021.

[92] Run-Fa Liao, Hong Wen, Jinsong Wu, Huanhuan Song, Fei Pan, and Lian Dong. The rayleigh fading channel prediction via deep learning. *Wireless Communications and Mobile Computing*, 2018, 2018.

[93] Le Tian, Elena Lopez-Aguilera, Eduard Garcia-Villegas, Michael Tetemke Mehari, Eli De Poorter, Steven Latré, and Jeroen Famaey. Optimization-oriented RAW modeling of IEEE 802.11ah heterogeneous networks. *IEEE Internet of Things Journal*, 6(6):10597–10609, 2019.

[94] Celestine Iwendi, Praveen Kumar Reddy Maddikunta, Thippa Reddy Gadekallu, Kuruva Lakshmanna, Ali Kashif Bashir, and Md. Jalil Piran. A metaheuristic optimization approach for energy efficiency in the iot networks. *Software: Practice and Experience*, n/a(n/a).

[95] Nurzaman Ahmed, Debashis De, Ferdous Ahmed Barbhuiya, and Md. Iftekhar Hussain. MAC Protocols for IEEE 802.11ah-Based Internet of Things: A Survey. *IEEE Internet of Things Journal*, 9(2):916–938, 2022.

[96] Xiaolin Jiang, Hossein Shokri-Ghadikolaei, Gabor Fodor, Eytan Modiano, Zhibo Pang, Michele Zorzi, and Carlo Fischione. Low-latency networking: Where latency lurks and how to tame it. *Proceedings of the IEEE*, 107(2):280–306, 2019.

[97] Rajat Talak and Eytan H. Modiano. Age-delay tradeoffs in queueing systems. *IEEE Transactions on Information Theory*, 67(3):1743–1758, 2021.

[98] Roy D. Yates, Yin Sun, D. Richard Brown, Sanjit K. Kaul, Eytan Modiano, and Sennur Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, 2021.

[99] Huimin Hu, Ke Xiong, Gang Qu, Qiang Ni, Pingyi Fan, and Khaled Ben Letaief. Aoi-minimal trajectory planning and data collection in uav-assisted wireless powered iot networks. *IEEE Internet of Things Journal*, 8(2):1211–1223, 2021.

[100] Xingran Chen, Konstantinos Gatsis, Hamed Hassani, and Shirin Saeedi Bidokhti. Age of information in random access channels. *IEEE Transactions on Information Theory*, 2022.

[101] R Gnanaselvam and M.S. Vasanthi. Signal coverage analysis with link adaptation in narrowband-iot. *Sustainable Computing: Informatics and Systems*, 33:100644, 2022.

[102] Ola Salman, Imad Elhajj, Ali Chehab, and Ayman Kayssi. Iot survey: An sdn and fog computing perspective. *Computer Networks*, 143:221–246, 2018.

[103] Henk C Tijms. *A first course in stochastic models*. John Wiley and sons, 2003.

[104] Steven A Lippman. On dynamic programming with unbounded rewards. *Management Science*, 21(11):1225–1233, 1975.

[105] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[106] Prakash Narayan and Ioannis Lambadaris. Jointly optimal admission and routing controls at a network node. *Communications in Statistics. Stochastic Models*, 10(1):223–252, 1994.

[107] Bruce Hajek. Optimal control of two interacting service stations. *IEEE Transactions on Automatic Control*, 29(6):491–499, 1984.

[108] Panqanamala Ramana Kumar and Pravin Varaiya. *Stochastic systems: Estimation, identification, and adaptive control*. SIAM, 2015.

[109] Zvi Rosberg, P Varaiya, and J Walrand. Optimal control of service in tandem queues. *IEEE Transactions on Automatic Control*, 27(3):600–610, 1982.

[110] Jean Walrand. *An introduction to queueing networks*. Prentice Hall, 1988.

[111] AO Allen. Probability, statistics and queueing theory with data science applications, 2014.

APPENDIX A

Proof of Proposition 8.2: Structural Properties

For convenience, we replace the notation V_N^{β} by V_N. Additionally, for every $x \in E$, we note

$$\begin{cases} \Delta_{11}V_N(x) = V_N(D_{11}x) \\ \Delta_{12}V_N(x) = V_N(D_{12}x)\mathbb{1}_{\{x^{11}=0\}} \\ \Delta_{13}V_N(x) = V_N(D_{13}x)\mathbb{1}_{\{x^{11}=0,x^{12}=0\}} \\ \Delta_{22}V_N(x) = V_N(D_{22}x) \\ \Delta_{23}V_N(x) = V_N(D_{23}x)\mathbb{1}_{\{x^{22}=0\}} \\ \Delta_{3}V_N(x) = V_N(D_{3}x) \\ \Gamma_{11}V_N(x) = V_N(A_{11}x) \\ \Gamma_{2}V_N(x) = \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \Gamma_{3}V_N(x) = \min\{V_N(A_{13}x), V_N(A_{23}x), V_N(A_{3}x)\} \end{cases}$$

We shall prove the structural properties by recurrence.

We have

$$\begin{cases} V_{N+1}(x) = c(x,a) + \frac{1}{\nu(\delta+\nu)}\left[\mu_1\left(\Delta_{11}V_N(x) + \Delta_{12}V_N(x) + \Delta_{13}V_N(x)\right)\right. \\ \qquad\qquad +\mu_2\left(\Delta_{22}V_N(x) + \Delta_{23}V_N(x)\right) + \mu_3\Delta_3V_N(x) \\ \qquad\qquad \left.+\lambda_1\Gamma_{11}V_N(x) + \lambda_2\Gamma_2V_N(x) + \lambda_3\Gamma_3V_N(x)\right], \ \forall N \geq 0 \\ \textit{with } V_0(.) = 0. \end{cases} \tag{A.1}$$

Remark A.1.

1) Since H contains the function that is identically zero, then $V_0 \in H$.

2) It is clear that the function:

$$V_1 = TV_0 : x \longmapsto c(x,a)$$

is an element of H.

3) Any product of a function that belongs to H by a positive number is in H, and any sum of functions that belong to H is in H.

We suppose that $V_N \in H$, and show that $V_{N+1} \in H$ for every $N \in \mathbb{N}$. That is, we need to show that V_{N+1} satisfies the properties (P1), (P2), (P3), (P4) and (P5).

According to the definition (A.1) of V_{N+1}, and Remark A.1, it is sufficient to show that the functions $\Delta_{11}V_N(x)$, $\Delta_{12}V_N(x)$, $\Delta_{13}V_N(x)$, $\Delta_{22}V_N(x)$, $\Delta_{23}V_N(x)$, $\Delta_3V_N(x)$, $\Gamma_{11}V_N(x)$, $\Gamma_2V_N(x)$ and $\Gamma_3V_N(x)$ verify these properties.

Property (P1)

$\Delta_{11}V_N$ satisfies the property (P1) is equivalent to:

For every $x \in E$ such that $x^{11} \geq 1$

$$\begin{cases} \Delta_{11}V_N(A_{11}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{12}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{13}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{22}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{23}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_3x) \geq \Delta_{11}V_N(x) \end{cases} \iff \begin{cases} V_N(A_{11}D_{11}x) \geq V_N(D_{11}x) \\ V_N(A_{12}D_{11}x) \geq V_N(D_{11}x) \\ V_N(A_{13}D_{11}x) \geq V_N(D_{11}x) \\ V_N(A_{22}D_{11}x) \geq V_N(D_{11}x) \\ V_N(A_{23}D_{11}x) \geq V_N(D_{11}x) \\ V_N(A_3D_{11}x) \geq V_N(D_{11}x) \end{cases}$$

which is true because V_N verifies the property (P1).

Similarly, we show that $\Delta_{12}V_N$ and $\Delta_{22}V_N$ satisfy the property (P1).

$\Delta_{13}V_N$ satisfies the property (P1) is equivalent to:

For every $x \in E$

$$\begin{cases} \Delta_{11}V_N(A_{11}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{12}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{13}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{22}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_{23}x) \geq \Delta_{11}V_N(x) \\ \Delta_{11}V_N(A_3x) \geq \Delta_{11}V_N(x) \end{cases} \tag{A.2}$$

If $x^{13} \geq 1$, we have

$$(A.2) \iff \begin{cases} V_N(A_{11}D_{13}x) \geq V_N(D_{13}x) \\ V_N(A_{12}D_{13}x) \geq V_N(D_{13}x) \\ V_N(x) \geq V_N(D_{13}x) \\ V_N(A_{22}D_{13}x) \geq V_N(D_{13}x) \\ V_N(A_{23}D_{13}x) \geq V_N(D_{13}x) \\ V_N(A_3D_{13}x) \geq V_N(D_{13}x) \end{cases}$$

which is true, because V_N verifies the property (P1).

If $x^{13} = 0$,

$$(A.2) \Longleftrightarrow \begin{cases} V_N(A_{11}x) & \geq & V_N(x) \\ V_N(A_{12}x) & \geq & V_N(x) \\ V_N(A_{13}x) & \geq & V_N(x) \\ V_N(A_{22}x) & \geq & V_N(x) \\ V_N(A_{23}x) & \geq & V_N(x) \\ V_N(A_3x) & \geq & V_N(x) \end{cases}$$

which is true, because V_N verifies the property (P1).

Similarly, we show that $\Delta_{23}V_N$ and $\Delta_3 V_N$ satisfy the property (P1).

$\Gamma_{11}V_N$ satisfies the property (P1) is equivalent to:

For every $x \in E$

$$\begin{cases} \Gamma_{11}V_N(A_{11}x) & \geq & \Gamma_{11}V_N(x) \\ \Gamma_{11}V_N(A_{12}x) & \geq & \Gamma_{11}V_N(x) \\ \Gamma_{11}V_N(A_{13}x) & \geq & \Gamma_{11}V_N(x) \\ \Gamma_{11}V_N(A_{22}x) & \geq & \Gamma_{11}V_N(x) \\ \Gamma_{11}V_N(A_{23}x) & \geq & \Gamma_{11}V_N(x) \\ \Gamma_{11}V_N(A_3x) & \geq & \Gamma_{11}V_N(x) \end{cases} \Longleftrightarrow \begin{cases} V_N(A_{11}A_{11}x) & \geq & V_N(A_{11}x) \\ V_N(A_{12}A_{11}x) & \geq & V_N(A_{11}x) \\ V_N(A_{13}A_{11}x) & \geq & V_N(A_{11}x) \\ V_N(A_{22}A_{11}x) & \geq & V_N(A_{11}x) \\ V_N(A_{23}A_{11}x) & \geq & V_N(A_{11}x) \\ V_N(A_3A_{11}x) & \geq & V_N(A_{11}x) \end{cases}$$

which is true because V_N verifies the property (P1).

To show that $\Gamma_2 V_N(x)$ satisfies the property (P1), We shall demonstrate inequalities in the form:

$$\min\{p_1, p_2\} \geq \min\{r_1, r_2\} \tag{A.3}$$

and use the following remark.

Remark A.2. *If we show that*

$$\min\{p_1, p_2\} \geq r_i$$

for a single i $(i = 1, 2)$, then we have the inequality (A.3), because $r_i \geq \min\{r_1, r_2\}$

$\Gamma_2 V_N$ satisfies the property (P1) is equivalent to:

For every $x \in E$,

$$\begin{cases} \Gamma_2 V_N(A_{11}x) & \geq & \Gamma_2 V_N(x) \\ \Gamma_2 V_N(A_{12}x) & \geq & \Gamma_2 V_N(x) \\ \Gamma_2 V_N(A_{13}x) & \geq & \Gamma_2 V_N(x) \\ \Gamma_2 V_N(A_{22}x) & \geq & \Gamma_2 V_N(x) \\ \Gamma_2 V_N(A_{23}x) & \geq & \Gamma_2 V_N(x) \\ \Gamma_2 V_N(A_3x) & \geq & \Gamma_2 V_N(x) \end{cases}$$

i.e., for every $x \in E$,

$$\begin{cases} \min\{V_N(A_{11}A_{12}x), V_N(A_{11}A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \min\{V_N(A_{12}A_{12}x), V_N(A_{12}A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \min\{V_N(A_{13}A_{12}x), V_N(A_{13}A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \min\{V_N(A_{22}A_{12}x), V_N(A_{22}A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \min\{V_N(A_{23}A_{12}x), V_N(A_{23}A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \\ \min\{V_N(A_3A_{12}x), V_N(A_3A_{22}x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x)\} \end{cases} \tag{A.4}$$

For each case of $p_i = \min\{p_1, p_2\}$, $i = 1,2$ in (A.3), we apply Remark A.2 with r_i for all inequalities of (A.4). We obtain

$$(A.4) \Longleftrightarrow \begin{cases} V_N(A_{11}A_{12}x) & \geq & V_N(A_{12}x) \\ V_N(A_{12}A_{12}x) & \geq & V_N(A_{12}x) \\ V_N(A_{13}A_{12}x) & \geq & V_N(A_{12}x) \\ V_N(A_{22}A_{12}x) & \geq & V_N(A_{12}x) \\ V_N(A_{23}A_{12}x) & \geq & V_N(A_{12}x) \\ V_N(A_3A_{12}x) & \geq & V_N(A_{12}x) \end{cases} \text{ or } (A.4) \Longleftrightarrow \begin{cases} V_N(A_{11}A_{22}x) & \geq & V_N(A_{22}x) \\ V_N(A_{12}A_{22}x) & \geq & V_N(A_{22}x) \\ V_N(A_{13}A_{22}x) & \geq & V_N(A_{22}x) \\ V_N(A_{22}A_{22}x) & \geq & V_N(A_{22}x) \\ V_N(A_{23}A_{22}x) & \geq & V_N(A_{22}x) \\ V_N(A_3A_{22}x) & \geq & V_N(A_{22}x \end{cases}$$

which is true because V_N verifies the property (P1).

To show that $\Gamma_3 V_N(x)$ satisfies the property (P1), We shall demonstrate inequalities in the form:

$$\min\{p_1, p_2, p_3\} \geq \min\{r_1, r_2, r_3\} \tag{A.5}$$

using the following remark.

Remark A.3. *If we show that*

$$\min\{p_1, p_2, p_3\} \geq r_i$$

for a single i ($i = 1,2,3$), then we have the inequality (A.5), because $r_i \geq \min\{r_1, r_2, r_3\}$

$\Gamma_3 V_N$ satisfies the property (P1) is equivalent to:

For every $x \in E$

$$\begin{cases} \Gamma_3 V_N(A_{11}x) & \geq & \Gamma_3 V_N(x) \\ \Gamma_3 V_N(A_{12}x) & \geq & \Gamma_3 V_N(x) \\ \Gamma_3 V_N(A_{13}x) & \geq & \Gamma_3 V_N(x) \\ \Gamma_3 V_N(A_{22}x) & \geq & \Gamma_3 V_N(x) \\ \Gamma_3 V_N(A_{23}x) & \geq & \Gamma_3 V_N(x) \\ \Gamma_3 V_N(A_3x) & \geq & \Gamma_3 V_N(x) \end{cases}$$

That is, for every $x \in E$

$$\begin{cases} \min\{V_N(A_{11}A_{13}x), V_N(A_{11}A_{23}x), V_N(A_{11}A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x), V_N(A_3x)\} \\ \min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x), V_N(A_3x)\} \\ \min\{V_N(A_{13}A_{13}x), V_N(A_{13}A_{23}x), V_N(A_{13}A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x), V_N(A_3x)\} \\ \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x), V_N(A_3x)\} \\ \min\{V_N(A_{23}A_{13}x), V_N(A_{23}A_{23}x), V_N(A_{23}A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x, V_N(A_3x))\} \\ \min\{V_N(A_3A_{13}x), V_N(A_3A_{23}x), V_N(A_3A_3x)\} & \geq & \min\{V_N(A_{12}x), V_N(A_{22}x), V_N(A_3x)\} \end{cases} \tag{A.6}$$

For each case of $p_i = \min\{p_1, p_2, p_3\}$, $i = 1,2,3$ in (A.5), we apply Remark A.3 with r_i for all inequalities of (A.6). We obtain

$$\begin{cases} V_N(A_{11}A_{13}x) & \geq & V_N(A_{13}x) \\ V_N(A_{12}A_{13}x) & \geq & V_N(A_{13}x) \\ V_N(A_{13}A_{13}x) & \geq & V_N(A_{13}x) \\ V_N(A_{22}A_{13}x) & \geq & V_N(A_{13}x) \\ V_N(A_{23}A_{13}x) & \geq & V_N(A_{13}x) \\ V_N(A_3A_{13}x) & \geq & V_N(A_{13}x) \end{cases} \text{ or } \begin{cases} V_N(A_{11}A_{23}x) & \geq & V_N(A_{23}x) \\ V_N(A_{12}A_{23}x) & \geq & V_N(A_{23}x) \\ V_N(A_{13}A_{23}x) & \geq & V_N(A_{23}x) \\ V_N(A_{22}A_{23}x) & \geq & V_N(A_{23}x) \\ V_N(A_{23}A_{23}x) & \geq & V_N(A_{23}x) \\ V_N(A_3A_{23}x) & \geq & V_N(A_{23}x) \end{cases} \text{ or } \begin{cases} V_N(A_{11}A_3x) & \geq & V_N(A_3x) \\ V_N(A_{12}A_3x) & \geq & V_N(A_3x) \\ V_N(A_{13}A_3x) & \geq & V_N(A_3x) \\ V_N(A_{22}A_3x) & \geq & V_N(A_3x) \\ V_N(A_{23}A_3x) & \geq & V_N(A_3x) \\ V_N(A_3A_3x) & \geq & V_N(A_3x) \end{cases}$$

which is true because V_N verifies the property (P1).

Consequently, V_{N+1} verifies the property (P1).

Property (P2)

$\Delta_{11}V_N$ satisfies the property (P2) is equivalent to:
For every $x \in E$ such that $x^{11} \geq 1$

$$\begin{cases} \Delta_{11}V_N(A_{12}A_{11}x) - \Delta_{11}V_N(A_{22}A_{11}x) & \geq & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \\ \Delta_{11}V_N(A_{12}^2x) - \Delta_{11}V_N(A_{22}A_{12}x) & \geq & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \\ \Delta_{11}V_N(A_{12}A_{13}x) - \Delta_{11}V_N(A_{22}A_{13}x) & \geq & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \\ \Delta_{11}V_N(A_{12}A_{22}x) - \Delta_{11}V_N(A_{22}^2x) & \leq & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \\ \Delta_{11}V_N(A_{12}A_{23}x) - \Delta_{11}V_N(A_{22}A_{23}x) & \leq & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \\ \Delta_{11}V_N(A_{12}A_3x) - \Delta_{11}V_N(A_{22}A_3x) & = & \Delta_{11}V_N(A_{12}x) - \Delta_{11}V_N(A_{22}x) \end{cases}$$

$$\iff \begin{cases} V_N(A_{12}x) - V_N(A_{22}x) & \geq & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \\ V_N(A_{12}^2D_{11}x) - V_N(A_{22}A_{12}D_{11}x) & \geq & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \\ V_N(A_{12}A_{13}D_{11}x) - V_N(A_{22}A_{13}D_{11}x) & \geq & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \\ V_N(A_{12}A_{22}D_{11}x) - V_N(A_{22}^2D_{11}x) & \leq & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \\ V_N(A_{12}A_{23}D_{11}x) - V_N(A_{22}A_{23}D_{11}x) & \leq & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \\ V_N(A_{12}A_3D_{11}x) - V_N(A_{22}A_3D_{11}x) & = & V_N(A_{12}D_{11}x) - V_N(A_{22}D_{11}x) \end{cases}$$

which is true because V_N verifies the property (P2).

Similarly, we show that $\Delta_{12}V_N$ and $\Delta_{22}V_N$ satisfy the property (P2).

$\Delta_{13}V_N$ satisfies the property (P2) is equivalent to:

For every $x \in E$,

$$\begin{cases} \Delta_{13}V_N(A_{12}A_{11}x) - \Delta_{13}V_N(A_{22}A_{11}x) & \geq & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \\ \Delta_{13}V_N(A_{12}^2x) - \Delta_{13}V_N(A_{22}A_{12}x) & \geq & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \\ \Delta_{13}V_N(A_{12}A_{13}x) - \Delta_{13}V_N(A_{22}A_{13}x) & \geq & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \\ \Delta_{13}V_N(A_{12}A_{22}x) - \Delta_{13}V_N(A_{22}^2x) & \leq & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \\ \Delta_{13}V_N(A_{12}A_{23}x) - \Delta_{13}V_N(A_{22}A_{23}x) & \leq & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \\ \Delta_{13}V_N(A_{12}A_3x) - \Delta_{13}V_N(A_{22}A_3x) & = & \Delta_{13}V_N(A_{12}x) - \Delta_{13}V_N(A_{22}x) \end{cases}$$

$$\iff \begin{cases} V_N(A_{12}A_{11}D_{13}x) - V_N(A_{22}A_{11}D_{13}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}^2D_{13}x) - V_N(A_{22}A_{12}D_{13}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_{13}D_{13}x) - V_N(A_{22}A_{13}D_{13}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_{22}D_{13}x) - V_N(A_{22}^2D_{13}x) & \leq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_{23}D_{13}x) - V_N(A_{22}A_{23}D_{13}x) & \leq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_3D_{13}x) - V_N(A_{22}A_3D_{13}x) & = & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \end{cases} \tag{A.7}$$

If $x^{13} \geq 1$,

$$(A.7) \iff \begin{cases} V_N(A_{12}A_{11}D_{13}x) - V_N(A_{22}A_{11}D_{13}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}^2D_{13}x) - V_N(A_{22}A_{12}D_{13}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}x) - V_N(A_{22}x) & \geq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_{22}D_{13}x) - V_N(A_{22}^2D_{13}x) & \leq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_{23}D_{13}x) - V_N(A_{22}A_{23}D_{13}x) & \leq & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \\ V_N(A_{12}A_3D_{13}x) - V_N(A_{22}A_3D_{13}x) & = & V_N(A_{12}D_{13}x) - V_N(A_{22}D_{13}x) \end{cases}$$

which is true because V_N verifies the property (P2).

If $x^{13} = 0$,

$$(A.7) \iff \begin{cases} V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) & \geq & V_N(A_{12}x) - V_N(A_{22}x) \\ V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) & \geq & V_N(A_{12}x) - V_N(A_{22}x) \\ V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) & \geq & V_N(A_{12}x) - V_N(A_{22}x) \\ V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) & \leq & V_N(A_{12}x) - V_N(A_{22}x) \\ V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) & \leq & V_N(A_{12}x) - V_N(A_{22}x) \\ V_N(A_{12}A_3x) - V_N(A_{22}A_3x) & = & V_N(A_{12}x) - V_N(A_{22}x) \end{cases}$$

which is true because V_N verifies the property (P2).

Similarly, we show that $\Delta_{23}V_N$ and $\Delta_3 V_N$ satisfy the property (P2).

$\Gamma_{11}V_N$ satisfies the property (P2) is equivalent to:

For every $x \in E$

$$\begin{cases} \Gamma_{11}V_N(A_{12}A_{11}x) - \Gamma_{11}V_N(A_{22}A_{11}x) & \geq & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \\ \Gamma_{11}V_N(A_{12}^2x) - \Gamma_{11}V_N(A_{22}A_{12}x) & \geq & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \\ \Gamma_{11}V_N(A_{12}A_{13}x) - \Gamma_{11}V_N(A_{22}A_{13}x) & \geq & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \\ \Gamma_{11}V_N(A_{12}A_{22}x) - \Gamma_{11}V_N(A_{22}^2x) & \leq & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \\ \Gamma_{11}V_N(A_{12}A_{23}x) - \Gamma_{11}V_N(A_{22}A_{23}x) & \leq & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \\ \Gamma_{11}V_N(A_{12}A_3x) - \Gamma_{11}V_N(A_{22}A_3x) & = & \Gamma_{11}V_N(A_{12}x) - \Gamma_{11}V_N(A_{22}x) \end{cases}$$

$$\iff \begin{cases} V_N(A_{12}A_{11}A_{11}x) - V_N(A_{22}A_{11}A_{11}x) & \geq & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ V_N(A_{12}^2A_{11}x) - V_N(A_{22}A_{12}A_{11}x) & \geq & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ V_N(A_{12}A_{13}A_{11}x) - V_N(A_{22}A_{13}A_{11}x) & \geq & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ V_N(A_{12}A_{22}A_{11}x) - V_N(A_{22}^2A_{11}x) & \leq & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ V_N(A_{12}A_{23}A_{11}x) - V_N(A_{22}A_{23}A_{11}x) & \leq & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ V_N(A_{12}A_3A_{11}x) - V_N(A_{22}A_3A_{11}x) & = & V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \end{cases}$$

which is true because V_N verifies the property (P2).

To show that $\Gamma_2 V_N$ satisfies the property (P2), We shall demonstrate inequalities in the form:

$$\min\{p_1,p_2\} - \min\{q_1,q_2\} \geq \min\{r_1,r_2\} - \min\{s_1,s_2\} \tag{A.8}$$

using the following remark.

Remark A.4.

1. *If we show that*
$$\min\{p_1,p_2\} - q_i \geq r_j - \min\{s_1,s_2\}$$
for unique i and j, then we have the inequality (A.8), because
$$q_i \geq \min\{q_1,q_2\} \text{ and } r_j \geq \min\{r_1,r_2\}$$
2. *Pour each case of p_i and s_i, we choose q_i and r_i, because the inequality $p_i - q_i \geq r_i - s_i$ is verified by hypotheses for every $i,j \in \{1,2\}$.*

Hereafter, $\Gamma_2 V_N$ satisfies the property (P2) is equivalent to:

For every $x \in E$,

$$\begin{cases} \Gamma_2 V_N(A_{12}A_{11}x) - \Gamma_2 V_N(A_{22}A_{11}x) & \geq & \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) \\ \Gamma_2 V_N(A_{12}A_{12}x) - \Gamma_2 V_N(A_{22}A_{12}x) & \geq & \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) \\ \Gamma_2 V_N(A_{12}A_{13}x) - \Gamma_2 V_N(A_{22}A_{13}x) & \geq & \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) \\ \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) & \geq & \Gamma_2 V_N(A_{12}A_{22}x) - \Gamma_2 V_N(A_{22}A_{22}x) \\ \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) & \geq & \Gamma_2 V_N(A_{12}A_{23}x) - \Gamma_2 V_N(A_{22}A_{23}x) \\ \Gamma_2 V_N(A_{12}A_{23}x) - \Gamma_2 V_N(A_{22}A_{23}x) & = & \Gamma_2 V_N(A_{12}x) - \Gamma_2 V_N(A_{22}x) \end{cases}$$

$$\iff \begin{cases} \min\{V_N(A_{12}A_{11}A_{12}x), V_N(A_{12}A_{11}A_{22}x)\} - \min\{V_N(A_{22}A_{11}A_{12}x), V_N(A_{22}A_{11}A_{22}x)\} \\ \geq \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\}, \\ \min\{V_N(A_{12}^3x), V_N(A_{12}^2A_{22}x)\} - \min\{V_N(A_{22}A_{12}^2x), V_N(A_{22}A_{12}A_{22}x)\} \\ \geq \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\}, \\ \min\{V_N(A_{12}A_{13}A_{12}x), V_N(A_{12}A_{13}A_{22}x)\} - \min\{V_N(A_{22}A_{13}A_{12}x), V_N(A_{22}A_{13}A_{22}x)\} \\ \geq \min\{V_N(A_{12}A_{12}x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}A_{22}x)\}, \\ \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\}, \\ \geq \min\{V_N(A_{12}A_{22}A_{12}x), V_N(A_{12}A_{22}^2x)\} - \min\{V_N(A_{22}^2A_{12}x), V_N(A_{22}^3x)\} \\ \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\} \\ \geq \min\{V_N(A_{12}A_{23}A_{12}x), V_N(A_{12}A_{23}A_{22}x)\} - \min\{V_N(A_{22}A_{23}A_{12}x), V_N(A_{22}A_{23}A_{22}x)\}, \\ \min\{V_N(A_{12}A_3A_{12}x), V_N(A_{12}A_3A_{22}x)\} - \min\{V_N(A_{22}A_3A_{12}x), V_N(A_{22}A_3A_{22}x)\} \\ \geq \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\}, \\ \min\{V_N(A_{12}^2x), V_N(A_{12}A_{22}x)\} - \min\{V_N(A_{22}A_{12}x), V_N(A_{22}^2x)\} \\ \geq \min\{V_N(A_{12}A_3A_{12}x), V_N(A_{12}A_3A_{22}x)\} - \min\{V_N(A_{22}A_3A_{12}x), V_N(A_{22}A_3A_{22}x)\}. \end{cases}$$

For the case p_1, s_2, we choose q_1, r_2 for all inequalities. i.e., we must show that

$$V_N(A_{12}A_{11}A_{12}x) - V_N(A_{22}A_{11}A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x), \tag{A.9}$$

$$V_N(A_{12}^3x) - V_N(A_{22}A_{12}^2x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x), \tag{A.10}$$

$$V_N(A_{12}A_{13}A_{12}x) - V_N(A_{22}A_{13}A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x), \tag{A.11}$$

$$V_N(A_{12}A_{22}^2x) - V_N(A_{22}^3x) \leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x), \tag{A.12}$$

$$V_N(A_{12}A_{23}A_{22}x) - V_N(A_{22}A_{23}A_{22}x) \leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x), \tag{A.13}$$

$$V_N(A_{12}A_3A_{12}x) - V_N(A_{22}A_3A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x), \tag{A.14}$$

and

$$V_N(A_{12}A_3A_{22}x) - V_N(A_{22}A_3A_{22}x) \leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x). \tag{A.15}$$

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{12}, x^{11} and decreasing in x^{22}, then

$$\begin{aligned} V_N(A_{12}A_{11}A_{12}x) - V_N(A_{22}A_{11}A_{12}x) &\geq V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \end{aligned}$$

  Hence we have the inequality (A.9).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{12} and decreasing in x^{22}, then

$$\begin{aligned} V_N(A_{12}^3x) - V_N(A_{22}A_{12}A_{12}x) &\geq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \end{aligned}$$

  Hence we have the inequality (A.10).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{12}, x^{13} and decreasing in x^{22}, then

$$\begin{aligned} V_N(A_{12}A_{13}A_{12}x) - V_N(A_{22}A_{13}A_{12}x) &\geq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \end{aligned}$$

  Hence we have the inequality (A.11).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{22} and increasing in x^{12}, then

$$\begin{aligned} V_N(A_{12}A_{22}^2x) - V_N(A_{22}^3x) &\leq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \end{aligned}$$

  Hence we have the inequality (A.12).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{22}, x^{23} and increasing in x^{12}, then

$$\begin{aligned} V_N(A_{12}A_{23}A_{22}x) - V_N(A_{22}A_{23}A_{22}x) &\leq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \end{aligned}$$

  Hence we have the inequality (A.13).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{12}, constant in x^3 and decreasing in x^{22}, then

$$\begin{aligned} V_N(A_{12}A_3A_{12}x) - V_N(A_{22}A_3A_{12}x) &\geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \\ &= V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \end{aligned}$$

  Hence we have the inequality (A.14).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{22}, constant in x^3 and increasing in x^{12}, then

$$\begin{aligned} V_N(A_{12}A_3A_{22}x) - V_N(A_{22}A_3A_{22}x) &\leq V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \\ &= V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \end{aligned}$$

  Hence we have the inequality (A.15).

For the case of p_2, s_1, we chose q_1, r_2 for all inequalities. i.e., we must show that

$$V_N(A_{12}A_{11}A_{22}x) - V_N(A_{22}A_{11}A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.16}$$

$$V_N(A_{12}^2A_{22}x) - V_N(A_{22}A_{12}^2x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.17}$$

$$V_N(A_{12}A_{13}A_{22}x) - V_N(A_{22}A_{13}A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.18}$$

$$V_N(A_{12}^3x) - V_N(A_{22}A_{22}A_{12}x) \leq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.19}$$

$$V_N(A_{12}A_{23}A_{22}x) - V_N(A_{22}A_{23}A_{12}x) \leq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.20}$$

$$V_N(A_{12}A_3A_{22}x) - V_N(A_{22}A_3A_{12}x) \geq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x), \tag{A.21}$$

and

$$V_N(A_{12}A_3A_{22}x) - V_N(A_{22}A_3A_{12}x) \leq V_N(A_{12}A_{22}x) - V_N(A_{22}A_{12}x). \tag{A.22}$$

which is obvious.

Finally, we have $\Gamma_2 V_N$ satisfies the property (P2).

To show that $\Gamma_3 V_N$ satisfies the property (P2), we shall demonstrate inequalities in the form:

$$\min\{p_1,p_2,p_3\} - \min\{q_1,q_2,q_3\} \geq \min\{r_1,r_2,r_3\} - \min\{s_1,s_2,s_3\} \tag{A.23}$$

using the following remark.

Remark A.5.

1. *If we show that*
$$\min\{p_1,p_2,p_3\} - q_i \geq r_j - \min\{s_1,s_2,s_3\}$$
for a single i and a single j, then we have the inequality (A.23), because
$$q_i \geq \min\{q_1,q_2,q_3\} \text{ and } r_j \geq \min\{r_1,r_2,r_3\}$$
2. *For each case of p_i, s_i we choose q_i and r_i, because the inequality $p_i - q_i \geq r_i - s_i$ is verified by by assumptions for every $i, j \in \{1,2,3\}$.*

Hereafter, $\Gamma_3 V_N(x)$ satisfies the property (P2) is equivalent to:

For every $x \in E$,

$$\begin{cases} \Gamma_3V_N(A_{12}A_{11}x) - \Gamma_3V_N(A_{22}A_{11}x) & \geq & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \\ \Gamma_3V_N(A_{12}^2x) - \Gamma_3V_N(A_{22}A_{12}x) & \geq & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \\ \Gamma_3V_N(A_{12}A_{13}x) - \Gamma_3V_N(A_{22}A_{13}x) & \geq & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \\ \Gamma_3V_N(A_{12}A_{22}x) - \Gamma_3V_N(A_{22}^2x) & \leq & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \\ \Gamma_3V_N(A_{12}A_{23}x) - \Gamma_3V_N(A_{22}A_{23}x) & \leq & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \\ \Gamma_3V_N(A_{12}A_3x) - \Gamma_3V_N(A_{22}A_3x) & = & \Gamma_3V_N(A_{12}x) - \Gamma_3V_N(A_{22}x) \end{cases}$$

i.e., for every $x \in E$,

$$\begin{cases}
\min\{V_N(A_{12}A_{11}A_{13}x), V_N(A_{12}A_{11}A_{23}x), V_N(A_{12}A_{11}A_{3}x)\} - \min\{V_N(A_{22}A_{11}A_{13}x), V_N(A_{22}A_{11}A_{23}x), V_N(A_{22}A_{11}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\}, \\
\min\{V_N(A_{12}^2A_{13}x), V_N(A_{12}^2A_{23}x), V_N(A_{12}^2A_{3}x)\} - \min\{V_N(A_{22}A_{12}A_{13}x), V_N(A_{22}A_{12}A_{23}x), V_N(A_{22}A_{12}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\}, \\
\min\{V_N(A_{12}A_{13}^2x), V_N(A_{12}A_{13}A_{23}x), V_N(A_{12}A_{13}A_{3}x)\} - \min\{V_N(A_{22}A_{13}^2x), V_N(A_{22}A_{13}A_{23}x), V_N(A_{22}A_{13}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\}, \\
\min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{22}A_{13}x), V_N(A_{12}A_{22}A_{23}x), V_N(A_{12}A_{22}A_{3}x)\} - \min\{V_N(A_{22}^2A_{13}x), V_N(A_{22}^2A_{23}x), V_N(A_{22}^2A_{3}x)\}, \\
\min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{23}A_{13}x), V_N(A_{12}A_{23}^2x), V_N(A_{12}A_{23}A_{3}x)\} - \min\{V_N(A_{22}A_{23}A_{13}x), V_N(A_{22}A_{23}A_{23}x), V_N(A_{22}A_{23}A_{3}x)\}, \\
\min\{V_N(A_{12}A_{3}A_{13}x), V_N(A_{12}A_{3}A_{23}x), V_N(A_{12}A_{3}^2x)\} - \min\{V_N(A_{22}A_{3}A_{13}x), V_N(A_{22}A_{3}A_{23}x), V_N(A_{22}A_{3}^2x)\} \\
\geq \min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\}, \\
\min\{V_N(A_{12}A_{13}x), V_N(A_{12}A_{23}x), V_N(A_{12}A_{3}x)\} - \min\{V_N(A_{22}A_{13}x), V_N(A_{22}A_{23}x), V_N(A_{22}A_{3}x)\} \\
\geq \min\{V_N(A_{12}A_{3}A_{13}x), V_N(A_{12}A_{3}A_{23}x), V_N(A_{12}A_{3}^2x)\} - \min\{V_N(A_{22}A_{3}A_{13}x), V_N(A_{22}A_{3}A_{23}x), V_N(A_{22}A_{3}^2x)\}.
\end{cases}$$

For the case of p_1, s_2, we choose q_1, r_2 for all inequalities. i.e., we must show that

$$V_N(A_{12}A_{11}A_{13}x) - V_N(A_{22}A_{11}A_{13}x) \geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x), \tag{A.24}$$

$$V_N(A_{12}^2A_{13}x) - V_N(A_{22}A_{12}A_{13}x) \geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x), \tag{A.25}$$

$$V_N(A_{12}A_{13}^2x) - V_N(A_{22}A_{13}^2x) \geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x), \tag{A.26}$$

$$V_N(A_{12}A_{22}A_{23}x) - V_N(A_{22}^2A_{23}x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x), \tag{A.27}$$

$$V_N(A_{12}A_{23}^2x) - V_N(A_{22}A_{23}^2x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x), \tag{A.28}$$

$$V_N(A_{12}A_{3}A_{13}x) - V_N(A_{22}A_{3}A_{13}x) \geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x), \tag{A.29}$$

and

$$V_N(A_{12}A_{3}A_{23}x) - V_N(A_{22}A_{3}A_{23}x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x). \tag{A.30}$$

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, x^{11} and decreasing in x^{23}, then

$$\begin{aligned}
V_N(A_{12}A_{11}A_{13}x) - V_N(A_{22}A_{11}A_{13}x) &\geq V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\
&\geq V_N(A_{12}x) - V_N(A_{22}x) \\
&\geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x)
\end{aligned}$$

  Hence we have the inequality (A.24).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, x^{12} and decreasing in x^{23}, then

$$\begin{aligned}
V_N(A_{12}^2A_{13}x) - V_N(A_{22}A_{12}A_{13}x) &\geq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \\
&\geq V_N(A_{12}x) - V_N(A_{22}x) \\
&\geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x)
\end{aligned}$$

Hence we have the inequality (A.25).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13} and decreasing in x^{23}, then

$$\begin{aligned} V_N(A_{12}A_{13}^2x) - V_N(A_{22}A_{13}^2x) &\geq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) \end{aligned}$$

Hence we have the inequality (A.26).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{23}, x^{22} and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_{22}A_{23}x) - V_N(A_{22}^2A_{23}x) &\leq V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.27).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{23} and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_{23}^2x) - V_N(A_{22}A_{23}^2x) &\leq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.28).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, constant in x^3 and decreasing in x^{23}, then

$$\begin{aligned} V_N(A_{12}A_3A_{13}x) - V_N(A_{22}A_3A_{13}x) &\geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \\ &= V_N(A_{12}x) - V_N(A_{22}x) \\ &\geq V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) \end{aligned}$$

Hence we have the inequality (A.29).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is decreasing in x^{23}, constant in x^3 and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_3A_{23}x) - V_N(A_{22}A_3A_{23}x) &\leq V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \\ &= V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.30).

For the case of p_1, s_3, we choose q_1, r_3 for all inequalities. i.e., we must show that

$$V_N(A_{12}A_{11}A_{13}x) - V_N(A_{22}A_{11}A_{13}x) \geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x), \tag{A.31}$$

$$V_N(A_{12}^2A_{13}x) - V_N(A_{22}A_{12}A_{13}x) \geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x), \tag{A.32}$$

$$V_N(A_{12}A_{13}^2x) - V_N(A_{22}A_{13}^2x) \geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x), \tag{A.33}$$

$$V_N(A_{12}A_{22}A_3x) - V_N(A_{22}^2A_3x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x), \tag{A.34}$$

$$V_N(A_{12}A_{23}A_3x) - V_N(A_{22}A_{23}A_3x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x), \tag{A.35}$$

$$V_N(A_{12}A_3A_{13}x) - V_N(A_{22}A_3A_{13}x) \geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x), \tag{A.36}$$

and

$$V_N(A_{12}A_3^2x) - V_N(A_{22}A_3^2x) \leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x). \tag{A.37}$$

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, x^{11} and constant in x^3, then

$$\begin{aligned} V_N(A_{12}A_{11}A_{13}x) - V_N(A_{22}A_{11}A_{13}x) &\geq V_N(A_{12}A_{11}x) - V_N(A_{22}A_{11}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &= V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \end{aligned}$$

  Hence we have the inequality (A.31).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, x^{12} and constant in x^3, then

$$\begin{aligned} V_N(A_{12}^2A_{13}x) - V_N(A_{22}A_{12}A_{13}x) &\geq V_N(A_{12}^2x) - V_N(A_{22}A_{12}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &= V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \end{aligned}$$

  Hence we have the inequality (A.32).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13} and constant in x^3, then

$$\begin{aligned} V_N(A_{12}A_{13}^2x) - V_N(A_{22}A_{13}^2x) &\geq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \\ &\geq V_N(A_{12}x) - V_N(A_{22}x) \\ &= V_N(A_{12}A_3x) - V_N(A_{22}A_3x) \end{aligned}$$

  Hence we have the inequality (A.33).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is constant in x^3, décroissante en x^{22} and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_{22}A_3x) - V_N(A_{22}^2A_3x) &= V_N(A_{12}A_{22}x) - V_N(A_{22}^2x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.34).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is constant in x^3, décroissante en x^{23} and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_{23}A_3x) - V_N(A_{22}A_{23}A_3x) &= V_N(A_{12}A_{23}x) - V_N(A_{22}A_{23}x) \\ &\leq V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.35).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is increasing in x^{13}, then

$$V_N(A_{12}A_3A_{13}x) - V_N(A_{22}A_3A_{13}x) \geq V_N(A_{12}A_3x) - V_N(A_{22}A_3x)$$

Hence we have the inequality (A.36).

- Since $V_N(A_{12}x) - V_N(A_{22}x)$ is constant in x^3, and increasing in x^{13}, then

$$\begin{aligned} V_N(A_{12}A_3^2x) - V_N(A_{22}A_3^2x) &= V_N(A_{12}x) - V_N(A_{22}x) \\ &\leq V_N(A_{12}A_{13}x) - V_N(A_{22}A_{13}x) \end{aligned}$$

Hence we have the inequality (A.37).

Consequently, V_{N+1} verifies the property (P2).

Properties (P3), (P4), and (P5)

Similarly to the proof of the property (P2), one can prove the properties (P3), (P4), and (P5) using Remark A.4 and Remark A.5.