UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Energy efficient offloading techniques for heterogeneous networks

Vasileios Miliotis

Ph.D. Thesis

Energy Efficient Offloading Techniques for Heterogeneous Networks

Author: Vasileios Miliotis

Advisors: Luis Alonso, Ph.D.
Associate Professor
Universitat Politècnica de Catalunya (UPC)

Christos Verikoukis, Ph.D.
Senior Researcher
Telecommunications Technological Center
of Catalonia (CTTC)

Department of Signal Theory and Communications
Universitat Politècnica de Catalunya

Barcelona, December 2015

Abstract

Mobile data offloading has been proposed as a solution for the network congestion problem that is continuously aggravating due to the increase in mobile data demand. The concept of offloading refers to the exploitation of network heterogeneity with the objective to mitigate the load of the cellular network infrastructure. In this thesis a multicast protocol for short range networks that exploits the characteristics of physical layer network coding is presented. In the proposed protocol, named CooPNC, a novel cooperative approach is provided that allows collision resolutions with the use of an indirect inter-network cooperation scheme. Through this scheme, a reliable multicast protocol for partially overlapping short range networks with low control overhead is provided. It is shown that with CooPNC, higher throughput and energy efficiency are achieved, while it presents lower delay compared to state-of-the-art multicast protocols. A detailed description of the proposed protocol is provided, with a simple scenario of overlapping networks and also for a generalised scalable scenario. Through mathematical analysis and simulations it is proved that CooPNC presents significant performance gains compared to other state-of-the-art multicast protocols for short range networks. In order to reveal the performance bounds of Physical Layer Network Coding, the so-called Cross Network is investigated under diverse Network Coding (NC) techniques. The impact of Medium Access Control (MAC) layer fairness on the throughput performance of the network is provided, for the cases of pure relaying, digital NC with and without overhearing and physical layer NC with and without overhearing. A comparison among these techniques is presented and the throughput bounds, caused by MAC layer limitations, are discussed. Furthermore, it is shown that significant coding gains are achieved with digital and physical layer NC and the energy efficiency performance of each NC case is presented, when applied on the Cross Network.

In the second part of this thesis, the uplink offloading using IP Flow Mobility (IFOM) is also investigated. IFOM allows a LTE mobile User Equipment (UE) to maintain two concurrent data streams, one through LTE and the other through WiFi access technology, that presents uplink limitations due to the inherent fairness design of IEEE 802.11 DCF. To overcome these limitations, a weighted proportionally fair bandwidth allocation algorithm is proposed, regarding the data volume that is being offloaded through WiFi, in conjunction with a pricing-based rate allocation algorithm for the rest of the data volume needs of the UEs that are transmitted through the LTE uplink. With the proposed approach, the energy efficiency of the

UEs is improved, and the offloaded data volume is increased under the concurrent use of access technologies that IFOM allows. In the weighted proportionally fair WiFi bandwidth allocation, both the different upload data needs of the UEs, along with their LTE spectrum efficiency are considered, and an access mechanism is proposed that improves the use of WiFi access in uplink offloading. In the LTE part, a two-stage pricing-based rate allocation is proposed, under both linear and exponential pricing approaches, with the objective to satisfy all offloading UEs regarding their LTE uplink access. The proposed algorithms are theoretically analysed and their performance is evaluated through simulations. Through the evaluation of energy efficiency, offloading capabilities and throughput performance, it is proved that the proposed uplink access scheme for UEs that operate with IFOM for uplink offloading presents higher performance compared to other state-of-the-art access schemes. The existence of a malicious UE is also considered that aims to exploit the WiFi bandwidth against its peers in order to upload less data through the energy demanding LTE uplink and a reputation based method is proposed to combat its selfish operation. This approach is theoretically analysed and its performance is evaluated, regarding the malicious and the truthful UEs in terms of energy efficiency. It is shown that while the malicious UE presents better energy efficiency before being detected, its performance is significantly degraded with the proposed reaction method.

Resumen

La derivación del tráfico de datos móviles (en inglés data offloading) ha sido propuesta como una solución al problema de la congestión de la red, un problema que empeora continuamente debido al incremento de la demanda de datos móviles. El concepto de offloading se entiende como la explotación de la heterogeneidad de la red con el objetivo de mitigar la carga de la infraestructura de las redes celulares. En esta tesis se presenta un protocolo multicast para redes de corto alcance (short range networks) que explota las características de la codificación de red en la capa física (physical layer network coding). En el protocolo propuesto, llamado CooPNC, se implementa una solución cooperativa que permite la resolución de colisiones mediante la utilización de un esquema indirecto de cooperación entre redes. Gracias a este esquema, se consigue un protocolo multicast fiable i con poco overhead de control para redes de corto alcance parcialmente solapadas. Se demuestra que el protocolo CooPNC consigue una mayor tasa de transmisión neta (throughput) y una mejor eficiencia energética, a la vez que el retardo se mantiene por debajo del obtenido con los protocolos multicast del estado del arte. La tesis ofrece una descripción detallada del protocolo propuesto, tanto para un escenario simple de redes solapadas como también para un escenario general escalable. Se demuestra mediante análisis matemático y simulaciones que CooPNC ofrece mejoras significativas en comparación con los protocolos multicast para redes de corto alcance del estado del arte. Con el objetivo de encontrar los límites de la codificación de red en la capa física (physical layer network coding), se estudia el llamado Cross Network bajo distintas técnicas de Network Coding (NC). Se proporciona el impacto de la equidad (fairness) de la capa de control de acceso al medio (Medium Access Control, MAC), para los casos de repetidor puro (pure relaying), NC digital con y sin escucha del medio, y NC en la capa física con y sin escucha del medio. La tesis presenta una comparación de todas estas técnicas y analiza los límites del throughput, causados por las limitaciones de la capa MAC. Además, se muestra que con NC digital y NC en la capa física se consiguen ganancias de codificación significativas. Para cada uno de estos casos, se presenta la eficiencia energética cuando se utiliza Cross Network.

En la segunda parte de la tesis se investiga el offloading en el enlace ascendente mediante IP Flow Mobility (IFOM). El IFOM permite a los usuarios móviles de LTE mantener dos flujos de datos concurrentes, uno a través de LTE y el otro a través de la tecnología de acceso WiFi, que presenta limitaciones en el enlace ascendente debido a la equidad (fairness) inherente del diseño de IEEE 802.11 DCF. Para

superar estas limitaciones, se propone un algoritmo proporcional ponderado de asignación de banda para el volumen de datos derivado a través de WiFi, junto con un algoritmo de asignación de tasa de transmisión basado en pricing para el volumen de datos del enlace ascendente de LTE. Con la solución propuesta, se mejora la eficiencia energética de los usuarios móviles, y se incrementa el volumen de datos que se pueden derivar gracias a la utilización concurrente de tecnologías de acceso que permite IFOM. En el algoritmo proporcional ponderado de asignación de banda de WiFi, se toman en consideración tanto las distintas necesidades de los usuarios en el enlace ascendente como su eficiencia espectral en LTE, y se propone un mecanismo de acceso que mejora el uso de WiFi para el tráfico derivado en el enlace ascendente. En cuanto a la parte de LTE, se propone un algoritmo en dos etapas de asignación de tasa de transmisión basada en pricing (con propuestas de pricing exponencial y lineal) con el objetivo de satisfacer el enlace ascendente de los usuarios en LTE. Los algoritmos propuestos son analizados teóricamente y su funcionamiento es evaluado mediante simulaciones. Mediante la evaluación de la eficiencia energética, la capacidad de offloading y el rendimiento del throughput se demuestra que el esquema de acceso propuesto para el enlace ascendente para usuarios que utilizan IFOM presenta mayor rendimiento en comparación con otros esquemas del estado del arte. También se contempla la existencia de usuarios maliciosos, que pretenden utilizar el ancho de banda WiFi contra sus iguales para transmitir menos datos a través del enlace ascendente de LTE (menos eficiente energéticamente). Para ello se propone un método basado en la reputación que combate el funcionamiento egoísta (selfish). Esta propuesta se analiza teóricamente y su rendimiento se evalúa, en términos energéticos, para los usuarios maliciosos y honestos (truthful). Se demuestra que los usuarios maliciosos presentan una eficiencia energética mejor antes de ser detectados, pero su rendimiento se degrada significativamente con el método reactivo propuesto.

Στους γονείς μου Κώστα και Ευνίκη
και στην αδερφή μου Χριστίνα.

Acknowledgements

It has been a long and rocky path towards this PhD. It was a life changing experience in many aspects and it would never be possible for me without the contribution of the people that I am grateful to and I will never forget.

First of all, I would like to thank my supervisors Professor Christos Verikoukis and Professor Luis Alonso for their detailed guidance throughout our weekly GREENET meetings. Their uncompromising devotion to detail and their persistence for high quality work has been a great influence for my future. The interaction with them has deeply affected my perception of research.

I would also like to express my gratefulness to the Professors Leandros Tassiuilas, Iordanis Koutsopoulos and Thanasis Korakis, in the order I met them, that introduced me in the world of research and inspired my passion for research and technology, each one with his own unique way.

I could not forget to thank Melani Gurdiel García that efficiently managed every non-technical detail during my PhD and gladly provided all her help during my stay in Barcelona. I will never forget our short coffee breaks in UPC.

During a PhD you also live your life, although you do not usually have much time. For the bad times that get even harder when you are abroad, I would like to thank my friends Dimitris Chandolias and Stathis Tsolakidis for their full support. I would also like to thank once more my supervisor Professor Christos Verikoukis for his sincere understanding and his discreet interest during these hard times.

I am deeply thankful to the Greek State Scholarship Foundation (IKY), as this thesis was funded by IKY, following a procedure of individualised assessment of the academic year 2011-2012, through the resources of the educational program "Education and Lifelong Learning" of the European Social Fund (ESF) and the NSRF 2007-2013.

Last but not least, I would like to thank my parents Evniki and Kostas, and my sister Christina for their full support and understanding. Words are not enough to express how grateful I am to my family!

Contents

List of Tables

List of Figures

Acronyms

3GPP	3rd Generation Partnership Project
ACK	Acknowledgement
AF-PNC	Amplify and Forward Physical layer Network Coding
AP	Access Point
ARQ	Automatic Repeat reQuest
BMMM	Batch Mode Multicast MAC Protocol
BMW	Broadcast Medium Window
BS	Base Station
CFC	Call For Cooperation
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear To Send
DBP	Delayed Feedback-based Protocol
DCF	Distributed Coordination Function
DF-PNC 1/ 2	Decode and Forward Physical layer Network Coding type 1/ 2
DPMM	Double Piggyback Mode Multicast
EDGE	Enhanced Data rates for GSM Evolution
ELBP	Enhanced LBP
eNodeB	Evolved Node B
HSPA	High Speed Packet Access
IEEE	Institute of Electrical and Electronics Engineers
IFOM	IP Flow Mobility
IP	Internet Protocol
LAN	Local Area Network
LBP	Leader Based Protocol
LIPA	Local IP Access

LTE	Long Term Evolution
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
MCTS	Multicast CTS
MIMO	Multiple Input and Multiple Output
MIMO_NC	MIMO_Network Coding
MRTS	Multicast Request to Send
NAK	Negative Acknowledgment
NAV	Network Allocation Vector
NC	Network Coding
PNC	Physical layer Network Coding
QoE	Quality of Experience
QoS	Quality of Service
RAMP	Reliable Access Multicast Protocol
R-DSTC	Randomized Distributed Space Time Codes
RLNC	Random Linear Network Coding
SEQ	Sequence
SETL	Smart Exponential-Threshold-Linear
SIPTO	Selected IP Traffic Offload
SNR	Signal to Noise Ratio
STA	Station
UDP	User Datagram Protocol
UE	User Equipment
UFM	Unicast-Friendly Multicast
WLAN	Wireless LAN

Chapter 1

Introduction

1.1 Motivation

The explosion of the data demand that we are already witnessing is the main reason that drives cellular network operators into the upgrade of cellular access to 4G systems like LTE-A aiming to be able to serve the requested traffic by their users. Due to the dramatic increase in population density in both residential and business areas, even with the upgrades of the cellular infrastructure the pace of the increase of data traffic demand requires further improvements besides the infrastructure upgrades, that will take advantage of the diversity of access through the exploitation of heterogeneous networks. Aiming to confront this "Tidal Effect", as described in [1], the research community has proposed offloading techniques, that will help to mitigate the overload of the cellular network spectrum.

Mobile data offloading pertains to traffic, created or requested from dual-mode devices and support cellular and WiFi connectivity, when it is routed over WiFi and small-networks. Mobile offloading occurs the user device leven when switched from a cellular connection to WiFi/ small-cell access or when both technologies are being used concurrently. According to cisco's global mobile data traffic forecast, as seen in Fig. 1.1 [2], as a percentage of total mobile data traffic from all mobile-connected devices, mobile offload increases from 45 percent (1.2 exabytes/month) in 2014 to 54 percent (28.9 exabytes/month) by 2019. Without offload, Global mobile data traffic would grow at a compound annual growth rate (CAGR) of 62 percent instead of 57 percent. Offload volume is determined by smartphone penetration, dual-mode share of handsets, percentage of home-based mobile Internet use, and percentage of dual-mode smartphone owners with WiFi fixed Internet access at home.

The objective of this thesis is dual. First, to present medium access control algorithms for heterogeneous networks that will provide effective ways to avoid congestion conditions at the cellular infrastructure and second to provide to the end users energy efficient access without compromising in terms of throughput and de-

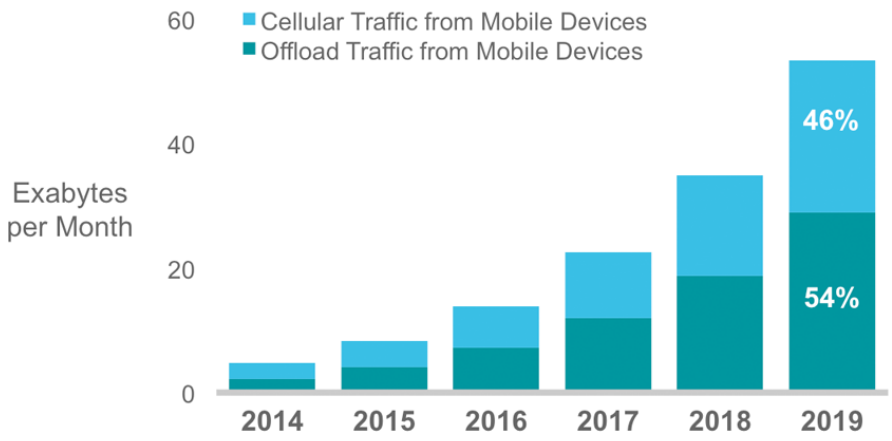


Figure 1.1: Offloading mobile data traffic forecast [2].

lay. Towards this objective, the considered models in this thesis are composed by concurrent LTE/ WiFi connectivity of the mobile users and the proposed offloading schemes are based on connection sharing for cooperative downlink and on proportional fair sharing access on the uplink combined with pricing strategies. In addition to the classic network performance metrics of throughput and delay, the energy efficiency of the network operation is presented, using the network bit delivery energy efficiency metric. This metric focuses on the amount of energy spent per delivered bit and is measured in bits per Joule.

This thesis provides a contribution to the field of MAC layer protocol design for energy efficient offloading for heterogeneous networks, by proposing and evaluating mechanisms that enhance different aspects of the network performance. The main motivation for this work has stemmed from the following two factors:

- *Cooperative communications* can achieve spatial diversity gains by requesting from neighboring stations to retransmit the overheard information to the final destination. By combining cooperation with Physical Layer Network Coding (PNC), further gains in terms of throughput, delay and energy efficiency are achieved. The density of short range networks in urban environments creates more room for improvement in the context of partially overlapping wireless networks that can be used to further enhance cellular communications by mitigating user created congestion and at the same time augmenting the network lifetime by reserving energy at the user-end.
- The support of multiple access technologies in contemporary personal wireless devices in conjunction with the evolution of modern cellular technologies that include heterogeneity, create challenges on fair and effective access schemes. Already mature technologies like WiFi need to be adapted to serve better new offloading techniques. Aspects like data volume needs and challenge conditions have to be considered for fairness in the increasing complexity of access. In addition, economic tools like pricing can also give insights and solutions towards more efficient and fair utilization of the network resources.

The main contributions of this work and the structure of the thesis will be discussed in detail in the following section.

1.2 Structure of the Thesis and Contributions

In this thesis we focus on two different methods of exploiting short range networks for the leverage of mobile data offloading. The first method refers to the use of cooperative networking for partially overlapping short range networks where users that are interested in the same information form a short range network. Instead of maintaining multiple sessions with the cellular network, one user is receiving the requested information and concurrently disseminates the data to its neighboring stations through multicast transmission. In the overlapping areas, collisions are solved by means of Physical Layer Network Coding. The second method is focused on the uplink offloading with the concurrent use of short range and cellular communication. While most of the users would prefer to offload all their mobile data through a more energy efficient and less expensive short range network, a fair approach is proposed, where users' data volume needs and channel conditions are the weighting factors for the fair access. In addition, pricing schemes are proposed to distributively control the uplink access of the cellular network.

The remaining part of the thesis consists of three chapters. Chapter 2 is focused in offloading with cooperative Physical Layer Network Coding. In Section 2.2 the state-of-the-art regarding reliable multicast MAC protocols is presented. In Section 2.3 a literature overview and categorization of cooperative network coding techniques is presented. The basic model and characteristics of CooPNC protocol are analysed in Section 2.4, where the proposed cooperative mechanism for collisions resolution is also described for the simple scenario of two partially overlapping networks. In the same Section, a scalable scenario of N partially overlapping networks is presented. In this scenario one central and $N-1$ peripheral networks are considered. The mathematical analysis of the performance of the central network in this scalable scenario, operating under the CooPNC protocol and a problem decomposition approach to explore the performance of the peripheral networks, is also included. Subsections 2.4.4 to 2.4.6 contain the evaluation of CooPNC. The part of the thesis that refers to network coding is finalised with Chapter 3 with an investigation of the upper performance bounds of diverse digital and physical layer network coding techniques in the context of enforced fairness, regarding throughput and energy efficiency.

Chapter 4 is focused on uplink offloading through concurrent WiFi and LTE transmissions. In Section 4.2 the state-of-the-art on offloading algorithms using WiFi networks is presented and in Section 4.3 the system model of our proposal. An analytical presentation of the proposed weighted proportionally fair WiFi access algorithm is in Section 4.4 and in Section 4.5 the analysis of the two proposed LTE pricing schemes. Section 4.6 includes the evaluation of the proposed uplink offloading scheme with IFOM. In Section 4.7 we a system model, where a UE presents malicious operation is described and in Section 4.8, the impact of the malicious operation on

the energy efficiency of the uplink offloading scheme is investigated. The Chapter is finalised in Section 4.9 where the performance evaluation in terms of throughput and energy efficiency under selfish misbehavior is presented.

1.3 Research Contributions

The novel proposals discussed in this thesis have been published in several research contributions. The work presented in Chapters 2 and 3, has been published in two journals and four international conferences, cited next:

[J1] V. Miliotis, L. Alonso, and C. Verikoukis, "Combining Cooperation and Physical Layer Network Coding to Achieve Reliable Multicast," *Recent Advances in Communications and Networking Technology*, vol. 2, pp. 41–49, Aug. 2013.

[J2] V. Miliotis, L. Alonso, and C. Verikoukis, "CooPNC: A Cooperative Multicast Protocol Exploiting Physical Layer Network Coding," *Ad Hoc Networks*, vol. 14, pp. 35–50, Mar. 2014.

[C1] V. Miliotis, L. Alonso, and C. Verikoukis,, "Multicast Performance Bounds Exploiting Cooperative Physical Layer Network Coding," in *IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 135–139, Sep. 2012, Barcelona, Spain.

[C2] V. Miliotis, L. Alonso, C. Skianis, and C. Verikoukis, "The Impact of Cooperative Physical Layer Network Coding on Multicast Short Range Networks," in *Proc. of IEEE International Conference on Communications (ICC 2013)*, pp. 3547–3551, June 2013, Budapest, Hungary.

[C3] V. Miliotis, L. Alonso, and C. Verikoukis, "Cooperative Multicast Exploiting Physical Layer Network Coding: A Performance Analysis," in *Proc. of IEEE International Conference on Communications 2013: IEEE ICC'13 - 3rd IEEE International Workshop on Smart Communication Protocols and Algorithms (SCPA 2013)*, pp. 1010–1014, June 2013, Budapest, Hungary.

[C4] V. Miliotis, L. Alonso, and C. Verikoukis, "Digital and Physical Layer Network Coding Performance in the Context of Enforced Fairness," in *Proc. of 21st International Conference on Telecommunications (ICT)*, pp. 72–76, May 2014, Lisbon, Portugal.

The proposed uplink offloading schemes, presented in Chapter 4 of this thesis, have been submitted in two journal papers under review and published in three international conferences:

[J3] V. Miliotis, L. Alonso, and C. Verikoukis, "Weighted Proportional Fairness and Pricing Based Resource Allocation for Uplink Offloading Using IP Flow Mobility," *Under Review in Elsevier Ad Hoc Networks*

[**J4**] V. Miliotis, L. Alonso, and C. Verikoukis, "Resource Allocation Techniques for Heterogeneous Networks Under User Misbehavior," *Under Review in IEEE Communication Letters*

[**C5**] V. Miliotis, L. Alonso, and C. Verikoukis, "Energy Efficient Proportionally Fair Uplink Offloading for IP Flow Mobility," in *IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 6–10, Dec. 2014, Athens, Greece.

[**C6**] V. Miliotis, L. Alonso, and C. Verikoukis, "Offloading With IFOM: The Uplink Case," in *Proc. of IEEE GLOBECOM'14*, pp. 2661–2666, Dec. 2014, Austin, Texas.

[**C7**] V. Miliotis, L. Alonso, and C. Verikoukis, "Combating Selfish Misbehavior with Reputation Based Uplink Offloading for IP Flow Mobility," to appear in *IEEE International Conference on Communications (ICC 2015)*, June 2015, London, UK.

Chapter 2

Offloading with Cooperative Physical Layer Network Coding

2.1 Introduction

We have been recently witnessing a rapid transition on the way internet is being used. New web-based applications appear frequently, providing social networking services, multimedia streaming, content sharing, remote storage space and remote processing of several instances of the everyday computer-aided activity. The need for being connected has started to be considered as a basic need. Social networks like Facebook and Twitter lead people to transform their social life from a conventional to an on-line type, sharing large amounts of data concerning their lives. The tremendous progress of connectivity speed along with the high storage spaces in the continuously deploying data centers opened the way for streaming services and cloud computing. Activities like watching videos, listening to the music, storing everyday work files and even processing files are no longer connected to local stored data in a user's personal computer but to transparent hardware accessed by web applications.

In parallel to the transition of the use habits of internet users, the wide spread of smart-phones surges the cellular networks' subscribers to access mobile internet services, providing them the opportunity to transfer large amount of data through wireless networks. This fact places a substantial pressure on operator's network capacity. As operators will not likely be able to follow the current pace of mobile data demand, they respond by rolling out WLANs to public areas to offload data traffic. WLANs are an appropriate solution, as they are easy to deploy and cost less than upgrading the existing cellular infrastructure gear. Despite the fact that the heterogeneity in the design and deployment of the networks tries to give a solution

to mitigate the pressure of mobile data demand, there is not yet any cooperative approach in connectivity, adopted to alleviate the pressure on the capacity of the operator's networks. Most devices act selfishly and frequently, separate requests from a 3G or LTE network are sent for the same streaming services from users close to each other that are interested in the same information. This fact not only creates redundant traffic to the cellular infrastructure, but it is also energy inefficient, while short range communication networks can be used for the same functionality. Instead of separate requests for streaming services, one connection could be established with the concurrent dissemination of the information through short range communication, exploiting a multicast scheme. At the same time, the intense augmentation of the density of short range networks in urban environments exacerbates the phenomenon of collisions, as most of the short range networks suffer from the fact that they partially overlap with other analogous neighbouring networks that operate in the same channel. This phenomenon is analyzed in [3] for realistic workload in unplanned multi-cell WLANs through testbed experimentation. In this Chapter, we focus on the performance of an offloading scheme for data dissemination and we provide analysis of a central network that suffers from collisions provoked by partially overlapping peripheral networks. We present our basic concepts as published in [4–7] and a generalised scalable scenario that is titled CooPNC (Cooperative multicast exploiting Physical Layer Network Coding) published in [8], and we compare its performance to other state of the art multicast protocols. The motivation behind our work is the offloading of the cellular networks, in cases where multiple requests for the same content originate from personal devices of a group of people that are close to each other. For example, when a group of friends is willing to watch a video (live streaming or stored to a server) at the same time. Instead of multiple connections to the cellular network, one connection with the concurrent dissemination of the information to the group would be adequate and as proven hereby more energy efficient. The main contributions of this Chapter are:

- The proposal of CooPNC protocol, a novel reliable multicast protocol that exploits the features of PNC to resolve collisions of partially overlapping multicast networks.
- The presentation of a scalable scenario that operates under CooPNC and consists of N neighbouring partially overlapping networks. Specifically one central and $N-1$ peripheral networks
- The analysis of the performance of CooPNC for the central and one peripheral network, in terms of throughput, delay and energy efficiency metrics, which are evaluated through extensive simulations and through comparison to state of the art multicast protocols.

The Chapter is organized as follows. In Section 2.2 we present a literature overview regarding reliable multicast MAC protocols and in Section 2.3 the related work and categorization of cooperative network coding techniques. In Section 2.4, we present the basic characteristics of CooPNC protocol and the cooperative mechanism for collisions resolution. We also provide the mathematical analysis of the performance

of the central network in a scalable scenario operating under the CooPNC protocol and a problem decomposition approach to explore the performance of the peripheral networks. We finalise this Chapter with the evaluation of our proposal in Sections 2.4.4 to 2.4.6.

2.2 MAC Protocols for Reliable Multicast Communications

Multicast communication in short range wireless networks represents the operation of sending data to a group of recipients, which are scattered in the range of the transmitter. The source address is a unicast address, whereas the destination address is a group address of some specific type. Unlike broadcasting, multicasting allows every member to choose whether to be part of the multicast group or not. Both broadcast and multicast communication protocols, as described in IEEE 802.11 [9], are not reliable. Nevertheless, research efforts have led to solutions that add reliability in such communication paradigms. In [10], a network coding-based reliable broadcast protocol for wireless mesh networks is proposed, that succeeds high packet delivery ratio with low delay. In [11] several routing metrics are studied for high-throughput in multicast communications. The authors also propose a low-overhead adaptive algorithm to incorporate link-quality metrics to a representative multicast routing protocol and they investigate the performance improvement achieved by using different link-quality-based routing metrics.

Multicasting is a way to reduce network load and end-to-end delay. It is therefore beneficial to the source of the transmission as well as to users interested in the same information. However, efficient multicast communication demands for special capabilities and specific algorithms at various layers of the protocol stack. The fact that, in prevalent short range communication protocols like IEEE 802.11, MAC layer retransmissions are not present in multicasting, has led the research community towards the proposal and design of several multicast MAC protocols. These protocols try to add reliability to this type of communications. Recently, several MAC protocols have been proposed to enhance the reliability and the efficiency of the IEEE 802.11 multicast protocol. These protocols can be categorized into two main categories: *Non-cooperative* and *Cooperative* MAC layer multicast protocols for short range communication networks.

2.2.1 Non-Cooperative Multicast

The non-cooperative MAC layer multicast protocols that have been proposed to enhance the reliability and the efficiency of the standard's multicast scheme can be classified into two main subcategories. The first is based on negative feedback [12–14] and the other is based on positive feedback [15–18]. In [12], the Leader Based Protocol (LBP) attempts to extend the IEEE 802.11 multicast protocol with a handshaking mechanism and a recovery mechanism. The protocol assumes that a

receiver is selected as the leader. Only the leader transmits a multicast CTS (MCTS) frame in reply to the sender. If the data are correctly received, the leader sends an ACK, otherwise sends a NACK. If any other receiver detects a transmission error, a NACK is also sent. This NACK frame will collide with the ACK, if any, sent by the leader. This leads to the AP not hearing any ACK, and thus retransmitting the lost frame. Applying LBP to IEEE 802.11 suffers from some problems. If an error occurs at any non-leader receiver, it will send a NACK, regardless of whether this erroneous frame has been received successfully before or not, which results in potential redundant retransmissions. In [14], the Enhanced LBP (ELBP) is proposed to solve this problem. In ELBP the MCTS transmitted from the leader is followed by a control frame named SEQ (sequence) from the AP, which informs the multicast group for the sequence number of the frame that will be transmitted. A different approach named 802.11MX, also based on negative feedback, is proposed in [13]. An extension to the IEEE 802.11 MAC protocol is proposed, aiming to provide reliability to multicast data communications. The extension is NACK based and uses tones, instead of conventional packets, to signal a NACK.

One of the most referenced positive feedback based protocol is proposed in [15]. The Broadcast Medium Window (BMW) is introduced to provide a reliable broadcast MAC. The basic idea of BMW is to treat each broadcast request as multiple unicast requests. BMW protocol is reliable but not very efficient. In order to improve the efficiency, Batch Mode Multicast MAC Protocol (BMMM) is proposed in [16]. To achieve that, in BMMM the sender uses RTS frames to sequentially instruct each intended receiver to transmit a CTS. In [17], the Double Piggyback Mode Multicast (DPMM) protocol is proposed to address the extra overhead problem. The protocol piggybacks the ACK in the CTS frame and also piggybacks the priority information in multicast data aggregated frames. Another reliable multicast MAC protocol is proposed in [18]. This protocol aims at providing high packet delivery ratio as well as low control overhead. The Reliable Access Multicast Protocol (RAMP), is fully compliant with the IEEE 802.11 protocol and is suitable for both infrastructure and wireless multihop networks. The key points of the RAMP scheme are that sender and receivers carry out an efficient handshaking procedure to ensure reliability and that nodes receiving or overhearing a frame on the channel can properly update their NAV so as to increase the channel utilization without increasing the collision probability on the wireless medium. In Fig. 2.1, we present the multicast transmission procedure of three key protocols. Namely, the ELBP, the BMMM and the RAMP protocols are presented. In ELBP the AP unicasts to a selected leader a RTS packet (u-RTS) and the leader responds with a MCTS packet. Following, a sequence packet (SEQ) is multicasted by the AP for the prevention of duplicates and then the multicast data packet is transmitted. In BMMM, the reliability is obtained by the exchange of unicast RTS - CTS pairs with all n subscribed multicast receivers. In RAMP, a MRTS that contains information on the response sequence of the multicast receivers is followed by unicast CTS by all subscribed receivers. After this handshake procedure the multicast data packet is transmitted.

The majority of the protocols and algorithms that have been proposed in non-cooperative MAC layer multicast mainly focus on the reliability of multicast transmission networks in terms of packet delivery ratio. Towards this aim in the BMMM,

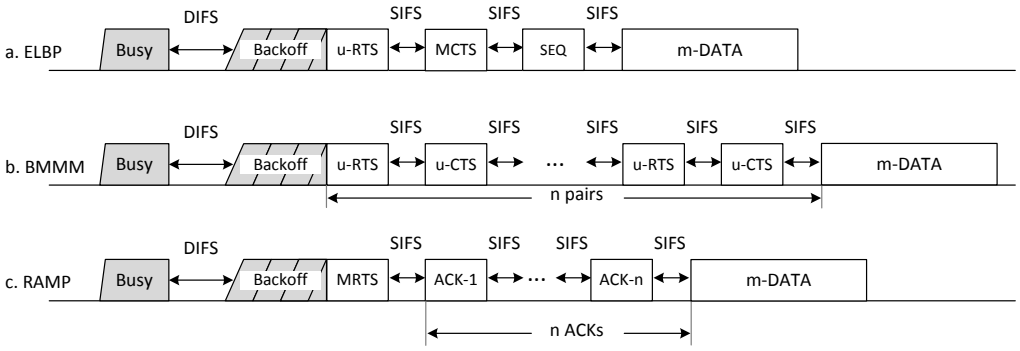


Figure 2.1: Multicast Procedures of (a) ELBP, (b) BMMM and (c) RAMP.

BMW and RAMP protocols, extra control packets are introduced to assure that a multicast frame is successfully received by all subscribed users in the multicast service. RAMP protocol presents the higher reliability among the non-cooperative multicast protocols. Nonetheless, the control overhead of the existing state of the art algorithms remains high, as all proposals present a high control to payload ratio. Our target in CooPNC is to preserve high reliability with high energy efficiency and low control information exchange. The low control overhead is achieved by only including in the protocol a Call for Cooperation (CFC) packet, as proposed in [19], to request a retransmission of the lost data due to a collision. The CFC packets are control packets of the same length and structure as the RTS packets with the difference being that the empty field for address 4 is exploited, as done in [20], to distinguish the packet from a normal RTS. The additional functionality of the CFC packets is that a retransmission is asked by a relay while concurrently the base station is informed for a collision.

2.2.2 Cooperative Multicast

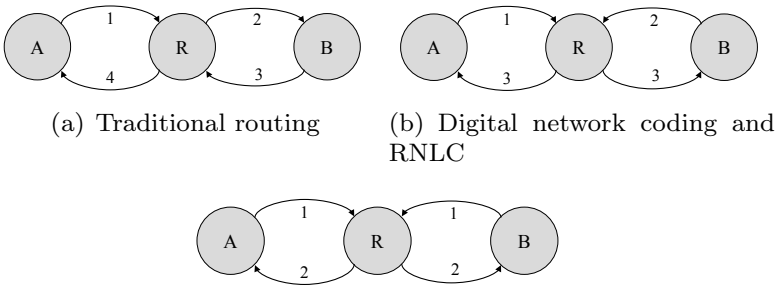
In order to increase the reliability, to allow for higher data rate transmission from the AP and to increase the coverage range, cooperative communications can prove useful by allowing nodes to receive or recover lost data from surrounding nodes. An AP may transmit multicast data packets to the relay nodes and the relay nodes forward them to the surrounding nodes, so that the AP may transmit at a higher data rate. Moreover the relay nodes may forward data to nodes that did not experience a successful reception. In the survey in [21], a broad spectrum of research effort is presented on the analysis of multicasting over wireless access networks from the late nineties onwards. Though, little attention is given to cooperative multicast communications. In [22] a joint relay node and channel selection algorithm is proposed to ensure that, within an interference range, each relay node uses a unique channel. In [23], the authors proposed that the source and the relays are allocated different subcarriers in order to transmit data at the same time. In certain circumstances, like video transmission over multicast in wireless infrastructure networks, errors are frequently location dependent. Hence, each user in the multicast group will most

likely lose different packets. Therefore, a simple Automatic Repeat reQuest (ARQ) based scheme is not appropriate for video multicast over wireless channels, since it will induce a large volume of retransmissions. In conventional multicast design, the receivers with a good channel quality unnecessarily suffer and experience a lower quality of service than the one they would have if the system were targeted at good receivers. In [24] a cooperative multicast mechanism is proposed, where the receivers are divided into two groups according to the average channel quality. Receivers located close to the sender, experience better conditions and receive data in higher rate. Thereafter they act as relays and transmit the received data to the rest of the intended receivers that present lower channel quality. This approach transforms the problem into a two layer multihop scheme and with this strategy substantial gain in signal quality is achieved. In [25], the cooperation at the medium access control (MAC) layer is explored and a new protocol called CoopMAC is proposed. CoopMAC is based upon the existing IEEE 802.11 DCF mode and is verified that it can achieve substantial throughput and delay performance improvements, without incurring significant complexity overhead in the system design. In conventional multicast wireless communications the rate of the transmission is bottlenecked by the data rate of the weakest client, degrading system performance. It is noteworthy that, neither for non-cooperative, nor for cooperative MAC layer multicast protocols proposed in the literature, energy efficiency is scarcely considered in the design and evaluation of the proposed approaches.

2.3 Cooperative Network Coding Techniques

Multicast, as a bandwidth efficient mechanism to provide wireless services for a group of terminals/ users, has been continuously investigated and combined with new ideas and techniques in networking. Augmenting the concept of cooperation in multicast communications, Network Coding (NC) has given new potentials. In the literature, research on Network Coding based cooperative multicast presents an augmenting pace. With the seminal work of R. Ahlswede et al [26], a simple but important observation was made, that in communication networks nodes can be allowed to not only forward but also process the incoming independent information flows. At the network layer for example, intermediate nodes can perform binary addition of independent bit-streams, whereas at the physical layer, intermediate nodes can receive superimposed incoming signals. In other words, data streams that are independently produced and consumed do not necessarily need to be kept separate when they are transported through the network. There are ways to combine and later extract independent information [27]. The research effort on exploitation of Network Coding concept in cooperative communications has led in the following categories:

1. Digital Network Coding and Random Linear Network Coding (RNLC)
2. Hybrid Channel/ Network Coding
3. Physical Layer Network Coding (PNC)



(a) Traditional routing

(b) Digital network coding and RNLC

(c) Analog network coding and PNC

Figure 2.2: Proposed protocols for the two-way relay topology.

4. MIMO Network Coding

Certainly, the use of NC techniques in cooperative communications reduces the number of needed steps for data delivery from source to destination through relays. In Fig. 2.2(a)-2.2(c), different protocols for the two-way relay topology are depicted and as shown, the needed slots for a successful frame transmission from each node can be reduced from four to two.

2.3.1 Digital Network Coding and RNLC

The intense research activity on the field of Network Coding firstly led in the concept of Digital Network Coding (DNC) and Random Linear Network Coding (RNLC). Digital NC and RNLC have broadly found field for applications in cooperative networks. In [28], the properties of the wheel network and diverse cooperative schemes of relaying are presented. The three schemes of cooperation are: (i) pure relaying (ii) pure network coding and (iii) network coding in teams. The energy efficiency of the implementation of digital network coding in relaying networks is also analysed. An application layer NC scheme for photo exchange is proposed in [29]. This work evaluates the wireless capacity used, the time needed for exchange of data and the energy consumption. The benefits of letting more groups work cooperatively for data dissemination, exploiting NC in short range communications, aiming to mitigate the traffic of cellular network access, are explored in [30], while in [31] and [32], the use of a short range scheme as an error recovery secondary network is proposed, in order for the nodes to be able to exchange missing data from a broadcast transmission of a primary network (e.g. LTE network). Consider a system that operates as information relay, such as a router, a node in an WLAN network, or a node in a peer to peer distribution network. Traditionally, when forwarding an information packet destined to some other node, the relay simply repeats it. With RNLC, a node is allowed to combine a number of packets it has received or created into one or several outgoing packets. This combination is linear. The reason for choosing a linear framework is that the algorithms for coding and decoding are well understood.

Linear combination is not concatenation. If we linearly combine packets of length L, the resulting encoded packet also has size L. An encoded packet generally carries information about several original packets, but in contrast to concatenation, just by itself it does not allow to recover any part of the original packets. One can think of linear network coding as a form of information spreading. This is also a reason why RNLC is considered a method that induces security in the communication.

2.3.2 Hybrid Channel/ Network Coding

The combination of channel coding and network coding in a cooperative network has led into this hybrid category of network coding. Commonly, the network code is designed independently from the channel code and the decoding of the channel code is carried out separately from the decoding of the network code [33]. Further improvements are expected when a joint design of the network and the channel code is applied and joint decoding is performed. The studies on hybrid channel/ network coding have come to a few important inferences. Firstly, the performance gap between separate and joint channel/ network coding is about 2-3 dB according to [34–36]. Also, the usage of incorrectly decoded frames to generate additional redundancy leads to bandwidth expansion [37]. Finally, the achievable diversity order increases as the code rate is decreased [36, 38].

2.3.3 Physical Layer Network Coding

Analog NC or Physical layer NC (PNC) [39] has attracted much research attention. In multiple access phase of PNC, two nodes are allowed to transmit simultaneously as depicted in Fig. 2.2(c). The channel additivity that happens during collisions naturally computes a linear combination of two frames and this outcome may then be broadcasted to the two nodes by a relay. The clear advantage is that just 2 slots are required to deliver the frames, rather than 3 in DNC or 4 in traditional routing. According to [33], in ANC three main approaches have emerged:

1. **Amplify and Forward PNC (AF-PNC)**, also known as Analog Network Coding (ANC) [40–43]. According to this idea, the relay deals only with the analog signals that have collided during the frames reception. The resulting signal is amplified and broadcasted and the two colliding nodes decode the intended frame after subtracting the frame that they sent. The advantage of this scheme is its simplicity, but the relay also amplifies the noise with which the final destinations must cope.

2. **Decode and forward PNC type 1 (DF-PNC 1)**, also known as compute and forward [44, 45]. This strategy suggests decoding the sum of two colliding frames at the relay, but neither the one nor the other frame individually. This strategy is more complex and extremely sensitive to timing and carrier synchronization. In other words, the colliding frames must be symbol and phase synchronous when received by the relay, which is hard to achieve.

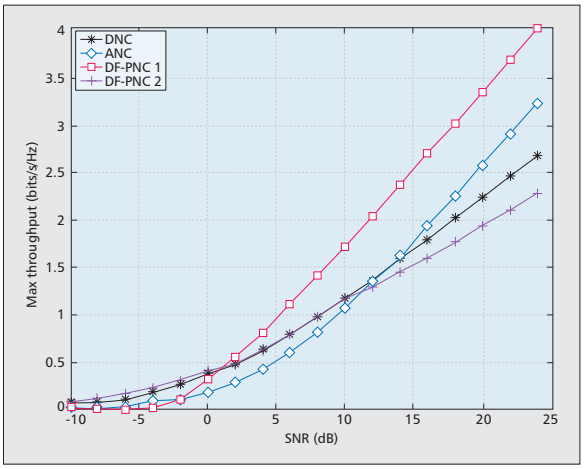


Figure 2.3: Comparison of DNC and the three main approaches of PNC [33].

3. **Decode and forward PNC type 2 (DF-PNC 2)**. In this approach the relay decodes both the superimposed signals by means of multiuser detection [46]. After that, it transmits a linear combination of the digital packets. This approach is not affected by noise propagation but requires decoding more information than DF-PNC 1 and this fact can heavily affect the system performance.

In Fig. 2.3 [33], a comparison between the main four protocols of NC is presented. Namely DNC is compared to ANC, DF-PNC 1 and DF-PNC 2. As it can be noticed, DF-PNC 1 is the overall winner, because it can successfully suppress the noise at the relay, hence it is not affected by error propagation, unlike ANC. On the other hand, the necessity for DF-PNC 2 to decode both the superimposed signals induces a big performance loss also with respect to DNC.

Up to our knowledge, little research has been done on applying PNC features in a suitable MAC protocol for multicast traffic. The seminal work of S. Zhang et al [39] led the research community to focus its attention to practical implementations of this topic. According to [47], "PNC was originally proposed as a way to exploit the network coding operation that occurs naturally in superimposed electromagnetic (EM) waves". The first implementation of PNC for a two-way relay network was presented in [48] and [49]. In these works the relay has to deal with symbol and carrier-phase asynchronies of the concurrent signals received from the two end nodes. A simple implementation case of PNC called Analog Network Coding (ANC) was presented in [40]. In ANC the assumption of perfect collision synchronization is relaxed. The decoding is performed by combining a two-packet collision with a priori information of one of the two packets, gained by previous clean reception. One proposal that builds on the idea of ANC is ZigZag decoding [50]. ZigZag is a new form of interference cancellation that exploits asynchrony across successive collisions. Specifically, 802.11 retransmissions, in the case of hidden terminals, cause successive collisions. These collisions have different interference-free stretches at

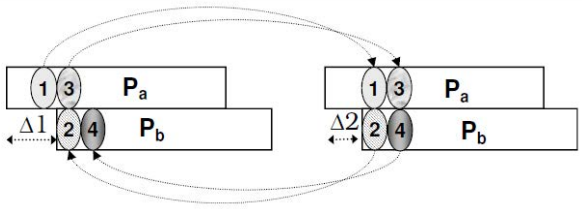


Figure 2.4: ZigZag decoding with two successive collisions of the same packets [50].

their start, as shown in Fig. 2.4 [50], which ZigZag exploits to bootstrap its decoding. ZigZag makes no changes to the 802.11 MAC and introduces no overhead when there are no collisions. In the cases that senders frames collide, ZigZag attains the same throughput as if the colliding packets were a priori scheduled in separate time slots. The idea is based on a hidden terminal scenario, where two users are out of the transmission range of each other and at the same time they are associated with the same AP, hence their uplink transmissions present high probability of collisions. The retransmission of the collided packets will provoke a new collision with different interference-free stretches at their start. While PNC and ANC collision resolution only applies for two collided packets, ZigZag is able to decode more than two collided packets under the assumption that it receives an equal number of differently overlapped collisions. In [51], a cooperative protocol allows two interfering senders to retransmit selected lost packets forcing a collision and using Analog Network Coding (ANC) to resolve the collision when having a priori information for one of the two colliding packets. This Cooperative Retransmission Through Collision (CRTC) protocol, which we use for comparison with our proposed scheme, presents higher throughput compared to traditional ARQ retransmission schemes. While ZigZag, from which we were partly inspired, is proposed to resolve collisions during the uplink transmissions of stations (STAs) to their base station (BS), we follow a different approach in downlink, where the collisions are provoked from multicast transmissions of different BSs in their overlapping areas. The cooperative approach we propose aims at the leverage of the collision resolution using ANC.

2.3.4 MIMO Network Coding

MIMO is well known to provide diversity and is robust to errors and noise. Given this property, MIMO is able to retrieve information even if some of the antennas are subject to strong fading. Such features are especially desirable in NC, as the loss of a coded packet may delay the whole decoding process. This area is still rather unexplored. In [38], the usage of coherent MIMO signal processing for NC detection is investigated. The underlying principle is to use the channel state estimates and NC coefficients to perform joint demodulation and NC decoding based on the received analog signals. A straightforward application of this principle is Phoenix [52], where MIMO-NC is proposed. This work has shown how to get the diversity gain of cooperation with the throughput efficiency of NC by means of signal processing techniques borrowed from MIMO. Through MIMO-NC it was shown that

separation of network coding and channel coding implies non-negligible performance degradations. Combining these two techniques entails more complexity in the system, and the question of what is the trade-off between performance, redundancy, and complexity has only been partly investigated and is still an open territory.

2.4 CooPNC: A Cooperative Multicast Protocol Exploiting Physical Layer Network Coding

2.4.1 CooPNC Overview

For the sake of simplicity in the description of the basic operation of CooPNC, we firstly consider only two partially overlapping short range communication networks. Each network consists of a multi-radio Base Station (BS) that receives data through a cellular network connection (e.g. 3G, LTE) and disseminates the information through a short range multicast scheme to its associated users, which are the stations (STAs) of each network. As presented in Fig. 2.5, STA1 and STA2 are associated with BS1 and BS2 respectively, and lie in an overlapping section of the transmission ranges of the two BSs. Each network also includes other STAs, like R1 and R2 that are subscribed to the multicast address and can also act as relays to carry out the resolution of collisions. The relay selection problem is out of the scope of this work. Every STA is associated with one BS but at the same time it is able to receive frames from other adjacent networks. This functionality can be implemented using for example the promiscuous mode in IEEE 802.11. As in IEEE 802.11 [9] that supports multicast transmissions by simply transmitting without any feedback, we assume that there is also no use of acknowledgements in this multicast scenario. In this situation there is a high probability of collisions within neighbouring short range communication networks. The collisions occur in the overlapping areas like area C in Fig. 2.5. This fact leads to a lower QoS in terms of throughput and delay for wireless applications that are based on the UDP transport layer protocol. It is worth mentioning that STAs located in area C suffer from the exact same collisions, as BS1 is hidden to BS2 and vice versa. The relays R1 and R2, that are situated outside the collision area, may facilitate the reliable reception of collided packets P1 and P2 originated from BS1 and BS2 respectively, by retransmissions at a higher rate than the multicast transmissions. In a conventional cooperative scheme, STA1 and STA2 would wait for an exponentially increasing back-off time and then send a control packet, namely a CFC packet, to request a retransmission of the lost data, while the BSs receiving a CFC packet would infer that a collision has occurred and will defer from transmitting new packets, invoking the execution of their own back-off algorithms. This conventional cooperative approach would induce high delay in the successful retransmission of the lost packets and at the same time the idle time of the network would be noticeably extended due to the back-off algorithms running in STAs and BSs. Both effects would affect negatively the overall throughput and energy consumption of the neighbouring networks.

Our proposal aims to mitigate both transmission delay and energy consumption

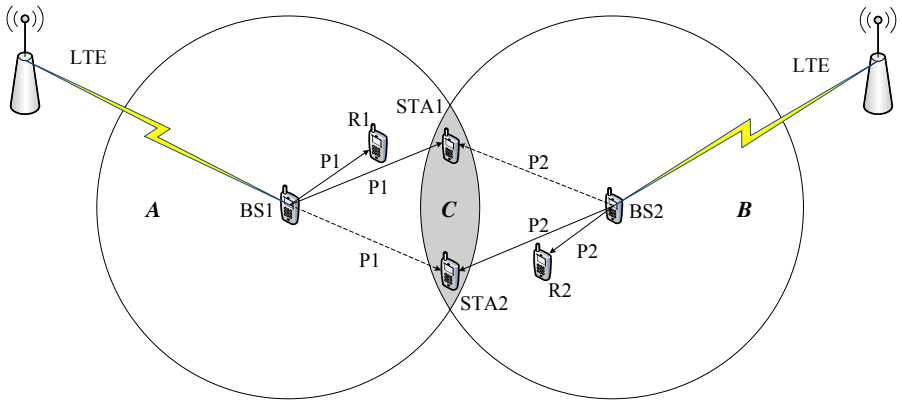


Figure 2.5: Multicast Transmission Scenario.

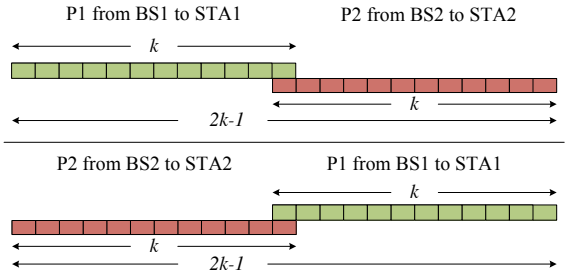


Figure 2.6: Distinguishing Symmetrical Collision Cases.

of the network through indirect inter-group cooperation, by exploiting ANC. In the scenario presented in Fig. 2.5, let's assume that there is a collision of Packet 1 (P1) and Packet 2 (P2) originated from BS1 and BS2, respectively. In order to solve this collision, we propose to use Zag or ANC, depending on the conditions the collision has occurred, and we assume that CFC packets are always sent by the STA that had first started receiving its packet before the collision occurred. This is why we consider the symmetrical collisions depicted in Fig. 2.6 as different collision cases, as in the first case a CFC packet will be transmitted by STA1, while in the second case by STA2. In case both packets arrive almost at the same time instant, resulting in a situation where none of the STAs is able to identify its own address in the Destination Address (DA) field, we introduce the use of a register, stored in the wireless interface card's driver of each STA, that is OFF for the last STA that sent a CFC during the previous collision, and ON for the STA that had benefit in resolving the collision by a relay of the neighbouring network. In these cases, the CFC packet will be only sent by the STA having its register set to ON and the values of the registers will change. With this mechanism we ensure that only one CFC will be sent by colliding STAs after each collision. In Fig. 2.7 STA1 has sent a CFC and a cooperation with R1 is in progress. When R1 transmits P1 to STA1, it is overheard by STA2, which now can exploit the ANC feature and export its own packet (P2) from the last collision. In the case that STA2 resolves the collision by a

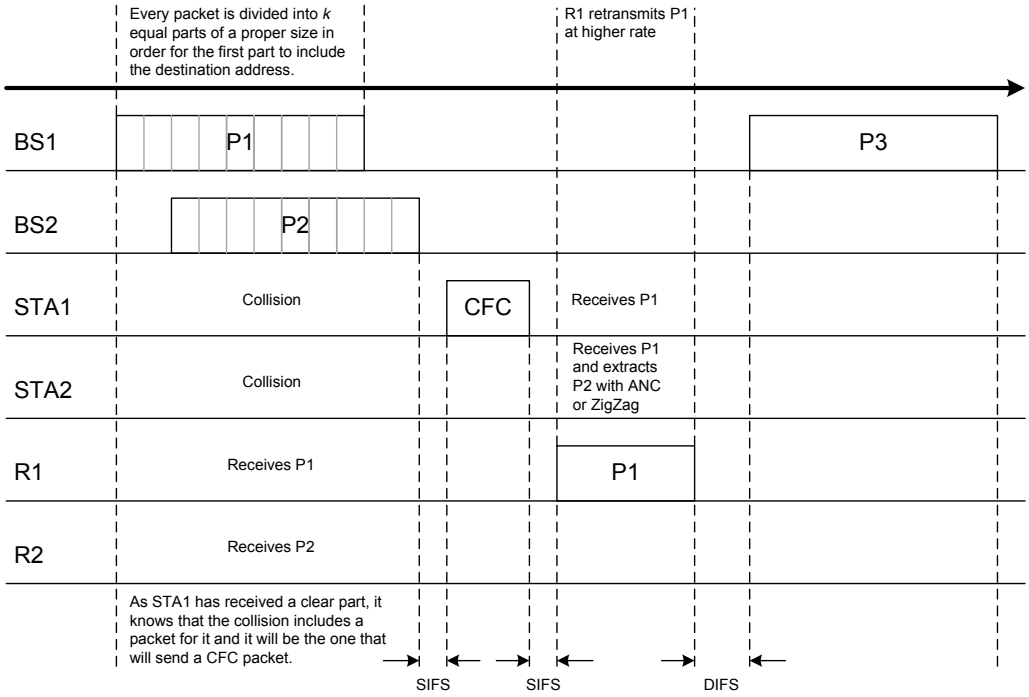


Figure 2.7: One Relay Collision Resolution.

retransmission from R1, it sets its CFC register set to ON, while STA1 sets it OFF. The side effect of this cooperation is that the collision is not only solved by STA1, but at the same time by STA2. This creates an indirect inter-group cooperation between the STAs of BS1 and BS2 networks.

In our analysis we consider an ideal error free wireless channel where the only source of errors are the occurring collisions. With this approach we are able to investigate the upper bounds of the performance improvements that CooPNC can provide. We assume that the time is slotted and in every slot three possible cases can occur: a successful transmission, no transmission or a collision. We define τ_i as the duration of an idle time slot and τ_s the duration of a successful transmission time slot. During τ_s one packet can be transmitted. We assume that $\tau_s = \tau_i$ and for the rest of the Chapter we will refer to both durations as τ_s. The proximity of the relays to the STAs allows them for retransmissions at higher rate from the multicast transmissions. Consequently, we define $\tau_r < \tau_s$ the duration of the retransmission of a packet from a relay to a STA and $\bar{\tau_c}$ the mean duration of one collision. In order to analytically present $\bar{\tau_c}$, we have first to consider the different possible cases of collisions. To that end, we divide each frame into k equal parts and calculate all the overlapping cases between two colliding frames, under the assumption that all data frames are of the same length. The k parts are of suitable length, in order for the first part to include the destination address of the frame. Sliding the frames from the first to the last possible collision, the number of different collision cases is equal to $2k - 1$. Quantizing the possible collision cases, we can assume without loss of generality that any of these collisions happens with equal probability. The

average length of a collision, excluding the case where the collision duration is approximately equal to the packet length (almost complete time overlapping of packets), is expressed as $\bar{l_c}$, where:

$$\bar{l_c} = \frac{2\sum_{m=1}^{k-1}(k+m)}{2k-2} = \frac{3k}{2} \tag{2.1}$$

We can distinguish two different cases in resolving a collision, depending on the actual length of a collision denoted by l_c:

(i) $l_c = k$, which happens with probability $\frac{1}{2k-1}$

(ii) $l_c \neq k$, which happens with probability $\frac{2k-2}{2k-1}$

In case (i) the duration of the collision of high synchronization is equal to τ_s, which is continued by a SIFS, a CFC packet, another SIFS, a retransmission of the collided packet by the corresponding relay and finally a DIFS. The CFC packet is sent by the STA that has its register set to ON. The total collision time of this case τ_{c_1}, excluding the retransmission time is presented in (2.2).

$$\tau_{c_1} = \tau_s + \tau_{CFC} + 2\tau_{SIFS} + \tau_{DIFS} \tag{2.2}$$

In case (ii) the mean duration of the collision of imperfect synchronization is equal to $(\tau_s/k)\bar{l_c}$, which is also continued by a SIFS, a CFC packet, again a SIFS, a retransmission of the collided packet by the corresponding relay, and a DIFS. The total collision time of this case τ_{c_2}, excluding the retransmission time is presented in (2.3).

$$\tau_{c_2} = \frac{\tau_s}{k}\bar{l_c} + \tau_{CFC} + 2\tau_{SIFS} + \tau_{DIFS} \tag{2.3}$$

Thus, the average duration of a collision $\bar{\tau_c}$ taking into account the probabilities according to which different cases of collisions occur, the retransmission time, and equations (2.2) and (2.3), can be expressed as:

$$\bar{\tau_c} = \frac{1}{2k-1}(\tau_{c_1} + \tau_r) + \frac{2k-2}{2k-1}(\tau_{c_2} + \tau_r) \tag{2.4}$$

It is clear that a portion of the duration of a collision includes the retransmission time by the relay, and this time has to be considered in the calculation of the throughput of the network. Rewriting (2.4) in (2.5), we separate $\bar{\tau_c}$ into the useful retransmission time τ_r that follows after each collision and the average duration of the collisions.

$$\bar{\tau_c} = \tau_r + \frac{1}{2k-1}\tau_{c_1} + \frac{2k-2}{2k-1}\tau_{c_2} \tag{2.5}$$

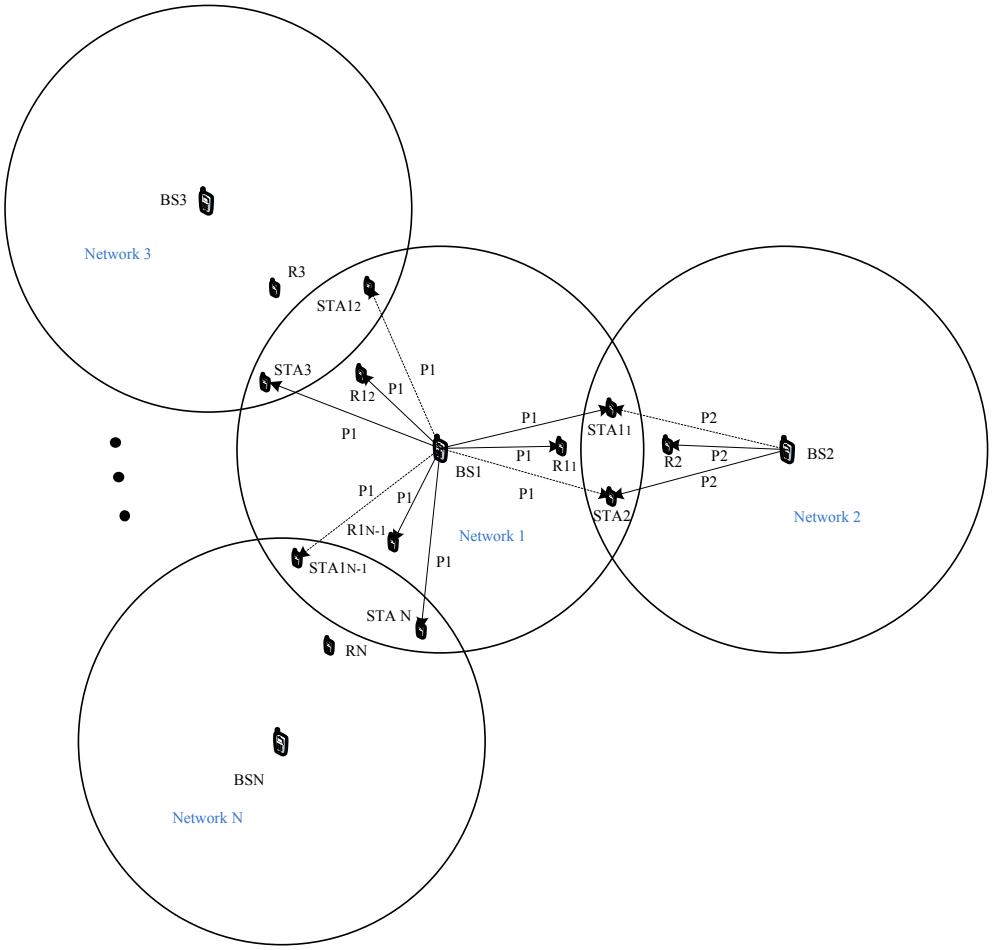


Figure 2.8: Scalable Multicast Scenario.

2.4.2 Performance of the CooPNC Scalable Scenario

Hereunder we present a scalable scenario of multicast communication short range networks operating with CooPNC and we investigate the performance in terms of throughput, delay and energy efficiency. We consider N partially overlapping short range networks. Each network consists of a multi-radio Base Station (BS), that is also connected to a cellular network (e.g. 3G, LTE) by which it receives data, that are disseminated through the short range network to its associated STAs through multicast transmission. As presented in Fig. 2.8, $\text{STA1}_1, ..., \text{STA1}_{N-1}$ are associated to BS1, while STA2, ...,STAN are associated to BS2, ...,BSN respectively. STA2, ...,STAN lie in sections of their corresponding BSs' ranges, that are overlapping with the range of BS1. Networks $2, ..., N$ are also multicast networks, with other associated receivers that are situated outside the overlapping areas. For the sake of simplicity we omit these receivers in the illustration of Fig. 2.8. Each network also includes other stations subscribed to the multicast address. These extra stations, like R1_1,...,R1_{N-1} for Network 1, may also act as relays to leverage the re-

solve of the collisions that occur in the $N-1$ overlapping areas. These areas present overlap only between the central and each surrounding network. We also assume that the relay transmissions for a collision that happens in a specific overlapping area do not interfere with neighbour overlapping areas and that all networks use the same channel to transmit. In this section we focus on the central network and we explore the lower bounds of its performance regarding throughput, delay and energy efficiency. Towards this aim, we assume that the collisions that occur are of worst case. Namely, the colliding packets are overlapping as less as possible resulting to the maximum duration of a collision. We also assume that these collision cases happen with equal probability and consequently the CFC transmissions are shared between STAs situated in the same collision areas. In this way, an indirect inter-group cooperation is achieved between the central and the peripheral overlapping networks.

Throughput analysis of the central network

In each BS, traffic arrives from the cellular network and information is disseminated to the associated STAs of each BS through multicast transmissions in its short range network. These multicast transmissions are modelled as statistically independent Bernoulli processes. The probability that a packet is transmitted from BSi is equal to p_i, for $i \in [1, 2, ..., N]$. We focus on the analysis of the throughput of BS1's network in relation to the activity of the surrounding networks. As we want to investigate how BS1's throughput performance is affected by the neighbouring networks BS2,...,BSN, we start by expressing the average idle period of BS1's network. Frame transmissions from BS1 follow a Bernoulli process with probability of transmission equal to p_1 and the number of consecutive idle slots until a frame transmission occurs is equal to the geometric random variable X, with probability mass function (PMF) equal to $p_X(n) = (1-p_1)^{n-1}p_1$. Thus, the expected idle period of BS1's network is defined as $E[T_i]$ and is expressed in (2.6).

$$E[T_i] = \tau_s \sum_{n=1}^{\infty} np_X(n) = \tau_s \sum_{n=1}^{\infty} n(1-p_1)^{n-1}p_1 = \frac{\tau_s}{p_1} \tag{2.6}$$

Following the same rationale as above, we express in (2.7) the average transmission period $\bar{E}[T_s]$ of BS1, in the case that its network was isolated of any other transmission.

$$\bar{E}[T_s] = \tau_s \sum_{n=1}^{\infty} np_1^{n-1}(1-p_1) = \frac{\tau_s}{1-p_1} \tag{2.7}$$

As the existence of the surrounding networks affects the success of BS1's transmissions, the quantity in (2.8) is divided into two parts: the expected duration of successful transmission periods $E[T_s]$, and the expected duration of collision periods $E[T_c]$. Applying the total expectation theorem for independent random variables,

we can express the expected duration of successful transmission periods, $E[T_s]$ as:

$$E[T_s] = \tau_s \prod_{k=2}^{N}(1-p_k)\sum_{n=1}^{\infty} np_1^{n-1}(1-p_1) = \tau_s \frac{\prod_{k=2}^{N}(1-p_k)}{1-p_1} \tag{2.8}$$

Collisions occur when BS1 transmits and at least one of the BSk, $k \in [2, ..., N]$ transmits. This happens with probability equal to $1-\prod_{k=2}^{N}(1-p_k)$. Thus, the expected duration of collision periods is equal to:

$$\begin{aligned} E[T_c] &= \tau_c(1-\prod_{k=2}^{N}(1-p_k))\sum_{n=1}^{\infty} np_1^{n-1}(1-p_1) \Rightarrow \\ E[T_c] &= \tau_c \frac{1-\prod_{k=2}^{N}(1-p_k)}{1-p_1} \end{aligned} \tag{2.9}$$

,where τ_c is the duration of the worst case collision. To express the duration of the worst case collision, we divide each frame into k equal parts and calculate the maximum length of a collision. The k parts are of suitable length, in order for the first part to include the destination address of the frame. These worst case collisions are of length equal to $l_c = (2k-1)\tau_s/k$. Every collision is continued by a SIFS, a CFC packet, another SIFS, a higher rate retransmission of the collided packet by the corresponding relay and finally a DIFS. The total collision time τ_c is presented in (2.10).

$$\tau_c = (2k-1)\tau_s/k + \tau_r + \tau_{CFC} + 2\tau_{SIFS} + \tau_{DIFS} \tag{2.10}$$

Distinguishing that every collision occurrence consists of the useful retransmission time and the non-beneficial time (the specific collision, the CFC, and the interframe spaces), we can express the expected duration of collision periods of (2.9) as:

$$E[T_c] = E[T_{c_u}] + E[T_{c_l}] \tag{2.11}$$

where $E[T_{c_u}]$ consists the expected duration of the retransmission portion of the expected duration of the collision periods and $E[T_{c_l}]$ is the expected duration of the non-beneficial parts of collision periods. Based on (2.9) we can express these expected durations in (2.12) and (2.13).

$$E[T_{c_u}] = \tau_r \frac{1-\prod_{k=2}^{N}(1-p_k)}{1-p_1} \tag{2.12}$$

$$E[T_{c_l}] = (\tau_c - \tau_r)\frac{1-\prod_{k=2}^{N}(1-p_k)}{1-p_1} \tag{2.13}$$

Using the multicast transmission rate R_m and the retransmission rate of the relays R_r, the throughput U of BS1 can be written as the ratio of successful transmission periods, including the retransmissions by the relays, weighted by the corresponding rates, to the total time of the network's operation, as shown in (2.14).

$$U = \frac{R_m E[T_s] + R_r E[T_{c_u}]}{E[T_s] + E[T_i] + E[T_c]} \tag{2.14}$$

Delay and energy efficiency analysis

The delay of the multicast transmission of BS1 is induced due to the non-beneficial part of collisions. We define as $E[T_D]$ the average per packet delay. We calculate the delay by dividing the non-beneficial expected duration of collision periods by the total number of successfully transmitted packets including the retransmissions. $E[T_D]$ can be expressed as:

$$E[T_D] = \frac{E[T_{c_l}]}{\frac{1}{\tau_s} E[T_s] + \frac{1}{\tau_r} E[T_{c_u}]} \tag{2.15}$$

Using equations (2.8), (2.12) and (2.13) we can rewrite (2.15) as:

$$E[T_D] = (\tau_c - \tau_r)[1 - \prod_{k=2}^{N}(1 - p_k)] \tag{2.16}$$

During idle periods of BS1's network, the wireless interfaces of BS1, its $N-1$ relays and its $N-1$ associated STAs are idle. The expected energy consumption of idle periods can be expressed as:

$$\mathcal{E}_I = (2N-1) P_I E[T_i] = (2N-1) P_I \frac{\tau_s}{p_1} \tag{2.17}$$

During successful transmission periods, BS1 transmits and its STAs and relays receive. Thus, the expected energy consumption for successful transmission periods can be expressed as:

$$\begin{aligned} \mathcal{E}_S = &\quad [P_T + 2(N-1)P_R] E[T_s] \Rightarrow \\ \mathcal{E}_S = &\quad [P_T + 2(N-1)P_R] \tau_s \frac{\prod_{k=2}^{N}(1-p_k)}{1-p_1} \end{aligned} \tag{2.18}$$

To analytically represent the energy consumption of BS1's network during collision periods, we need to calculate all the possible combinations of colliding networks. For $N-1$ surrounding networks, all the possible combinations c are equal to:

$$c = \sum_{l=1}^{N-1} \binom{N-1}{l} = 2^{N-1} - 1 \tag{2.19}$$

We denote as $\mathcal{E}_1$ the consumed energy by a STA and its assigned relay of Network 1 that are involved in a collision, $\mathcal{E}_2$ the consumed energy of a non-involved station and its assigned relay again of Network 1, and $\mathcal{E}_{BS}$ the consumed energy of BS1, all referring to one collision period. We construct a matrix A of binary elements, representing all the discrete combinations of transmissions. A consists of $2^{N-1}-1$ rows and every row represents a binary number in $[1, ..., 2^{N-1}-1]$. Thus, matrix A will have $N-1$ columns, each one corresponding to one of the surrounding networks and every line corresponding to a discrete combination of transmissions. Let $(\alpha_{i,1}, \alpha_{i,2}, ..., \alpha_{i,N-1})$ be equal to the corresponding elements of row i of the matrix A. Let $\bar{\alpha}_{i,1}$ be the one's complement of the binary number $\alpha_{i,1}$ of matrix A. When the corresponding element to a surrounding network, $\alpha_{i,j}$, is equal to 1, its BS transmits. Otherwise it is idle. Taking (2.9) into account, we can now express the energy consumed during collision periods as:

$$\begin{aligned}\mathcal{E}_c &= \sum_{i=1}^{2^{N-1}-1} [(\prod_{j=1}^{N-1} p_{j+1}^{\alpha_{i,j}}(1-p_{j+1})^{\bar{\alpha}_{i,j}}) \cdot \\ &\cdot(\mathcal{E}_1 \sum_{k=1}^{N-1} \alpha_{i,1} + \mathcal{E}_2 \sum_{k=1}^{N-1} \bar{\alpha}_{i,1} + \mathcal{E}_{BS})] \frac{1}{1-p_1} \end{aligned} \tag{2.20}$$

,where $\alpha_{i,j} \in A$. The mean energy consumptions $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_{BS}$ can be expressed as:

$$\begin{aligned}\mathcal{E}_1 &= \tau_s(2P_R) + 2\tau_{SIFS}(2P_I) + \frac{\tau_{CFC}}{2}(P_T + P_R) + \frac{\tau_{CFC}}{2}(2P_I) \\ &+\tau_{DIFS}(2P_I) + \frac{\tau_r}{2}(P_T + P_R) + \frac{\tau_r}{2}(P_R + P_I) \end{aligned} \tag{2.21}$$

$$\mathcal{E}_2 = \tau_s 2P_R + (\tau_c - \tau_s)2P_I \tag{2.22}$$

$$\mathcal{E}_{BS} = \tau_s P_T + \tau_{CFC} P_R + (\tau_c - \tau_s - \tau_{CFC})P_I \tag{2.23}$$

2.4.3 Peripheral Networks Performance

Aiming to explore how the performance of a peripheral network is affected under different traffic conditions, we consider the same scenario of Section 2.4.2 and we focus on the performance of the peripheral Network N. We decompose this scenario as depicted in Fig. 2.9 and we try to find the traffic p_e of an equivalent virtual network, composed by the combination of the rest $N-1$ networks, with which Network N interacts. The interplay between the peripheral Network N and the equivalent network is done through the central Network 1. Thus, we start by expressing the expected durations for idle, collision and successful transmission periods of Network 1, when isolated from Network N, as $E[T_i]$, $E[T_c]$ and $E[T_s]$ respectively.

$$E[T_i] = \tau_s / p_1 \tag{2.24}$$

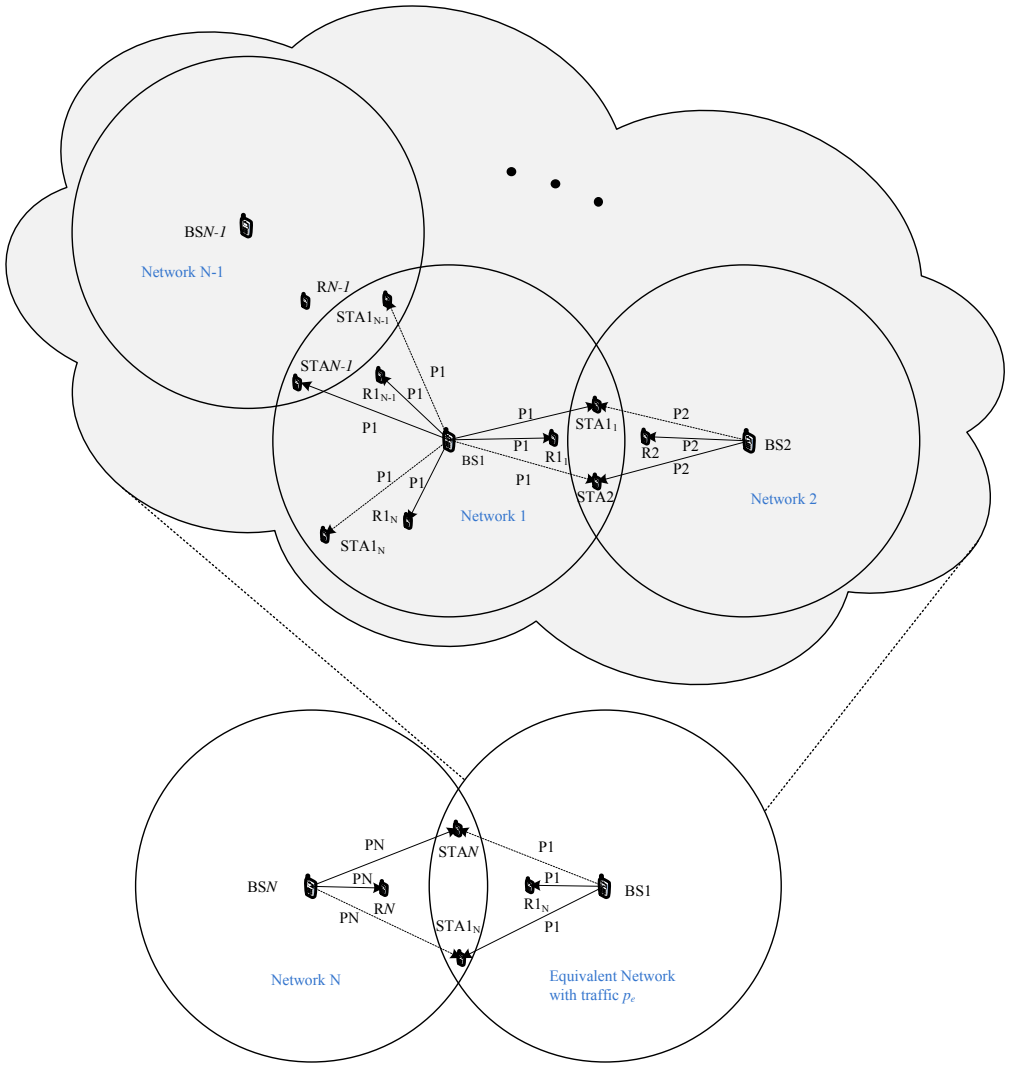


Figure 2.9: Decomposed Approach with Equivalent Network.

$$E[T_c] = \bar{\tau_c}\left(1 - \prod_{k=2}^{N-1}(1-p_k)\right)\frac{1}{1-p_1} \tag{2.25}$$

$$E[T_s] = \tau_s \prod_{k=2}^{N-1}(1-p_k)\frac{1}{1-p_1}. \tag{2.26}$$

The equivalent network will present activity during $E[T_s]$ and it will be idle during $E[T_i]$. From the perspective of Network N, the expected collision duration of (2.25) can be divided into two parts: the occupied expected duration $E[T_{c_o}]$ that includes the transmissions of Network 1 during collisions with any of the rest peripheral networks $(2, ..., N-1)$, and the free expected duration $E[T_{c_f}]$ that includes the rest of the expected duration of collision periods, during which Network N would be able to transmit, without confronting collisions. Hence, the joint virtual network would present an equivalent expected duration of idle period equal to $E[T_i] + E[T_{c_f}]$ and

an equivalent expected duration of transmission period equal to $E[T_s] + E[T_{c_o}]$.

Following the rationale of (2.6) we can calculate the equivalent probability p_e that describes the traffic of the virtual network, and provide an analytical expression through the solution of the following discrete integral equation:

$$\begin{aligned} E[T_i] + E[T_{c_f}] &= \tau_s \sum_{n=1}^{\infty} n(1-p_e)^{n-1} p_e = \tau_s / p_e \Rightarrow \\ p_e &= \tau_s / (E[T_i] + E[T_{c_f}]) \end{aligned} \tag{2.27}$$

The free period of the expected collision duration is:

$$E[T_{c_f}] = (\bar{\tau_c} - \tau_s) \left(1 - \prod_{k=2}^{N-1} (1-p_k)\right) \frac{1}{1-p_1} \tag{2.28}$$

Combining (2.24), (2.27) and (2.28) we can express p_e as:

$$p_e = \frac{\tau_s p_1 (1-p_1)}{\tau_s (1-p_1) + (\bar{\tau_c} - \tau_s) p_1 \left(1 - \prod_{k=2}^{N-1} (1-p_k)\right)}$$

Hereby, according to equations (2.6), (2.8) and (2.9), we can express the expected duration of idle, successful transmission and collision periods of Network N in relation to p_e and though, in relation to the activity of the equivalent network comprised by Networks 1 to $N-1$. We express these durations as:

$$E_N[T_i] = \tau_s / p_N \tag{2.29}$$

$$E_N[T_s] = \tau_s (1-p_e)/(1-p_N) \tag{2.30}$$

$$E_N[T_c] = \bar{\tau_c} p_e / (1-p_N) \tag{2.31}$$

The expected duration of collision periods of Network N includes the retransmissions of lost data, that we consider as the useful expected duration of collision periods $E_N[T_{c_u}]$. We express this expected duration as:

$$E_N[T_{c_u}] = \tau_r p_e / (1-p_N). \tag{2.32}$$

Throughput and delay of the peripheral networks

The throughput U_N of Network N can be written as the ratio of successful transmission periods, including the retransmissions by the relays, weighted by the corresponding rates, to the total expected duration of the network's operation:

$$U_N = \frac{R_m E_N[T_s] + R_r E_N[T_{c_u}]}{E_N[T_s] + E_N[T_i] + E_N[T_c]} \tag{2.33}$$

The delay of the multicast transmissions of BSN is induced due to the collisions, excluding the retransmissions that are considered beneficial transmissions. We define as $E_N[T_D]$ the average delay per packet for Network N. We calculate this delay by dividing the non-beneficial expected duration of collision periods by the total number of successfully transmitted packets including the retransmissions. $E_N[T_D]$ can be expressed as:

$$E_N[T_D] = \frac{E_N[T_c] - E_N[T_{c_u}]}{\frac{1}{\tau_s}E_N[T_s] + \frac{1}{\tau_r}E_N[T_{c_u}]} \tag{2.34}$$

Using equations (2.30)-(2.32) we can rewrite (2.34) as:

$$E_N[T_D] = (\bar{\tau_c} - \tau_r)p_e \tag{2.35}$$

Energy efficiency of the peripheral networks

Considering that the wireless interfaces of BSN, RN and STAN are idle during idle periods of the Network N, we can express its expected energy consumption for idle periods $\mathcal{E}_I$ as:

$$\mathcal{E}_I = 3P_I E_N[T_i] = 3P_I\tau_s/p_N \tag{2.36}$$

During successful transmission periods, BSN transmits and RN and STAN receive. Thus, the expected energy consumption for successful transmission periods can be expressed as:

$$\mathcal{E}_S = (P_T + 2P_R)E_N[T_s] = (P_T + 2P_R)\tau_s\frac{1-p_e}{1-p_N} \tag{2.37}$$

To calculate the expected energy consumption during collision periods, we take into consideration that all cases of collisions occur with equal probability, hence while half of the CFCs are transmitted by STAN, also half of the total retransmissions will be done by RN, while for the other half retransmissions, STAN will take benefit from the neighbouring network. Distinguishing perfect and imperfect synchronized collisions, we define as $\bar{\mathcal{E}}_{C_1}$ and $\bar{\mathcal{E}}_{C_2}$ the average energy consumption in Network N for each collision case respectively. Based on (2.2) and (2.3), we can express the mean energy consumption for each collision case, including retransmissions. Consequently, in the case of high synchronization we have:

$$\begin{aligned}\bar{\mathcal{E}}_{C_1} &= \tau_s(P_T + 2P_R) + 2\tau_{SIFS}(3P_I) + \frac{\tau_{CFC}}{2}(P_T + 2P_R) + \frac{\tau_{CFC}}{2}(3P_I) \\ &\quad +\tau_{DIFS}(3P_I) + \frac{\tau_r}{2}(P_I + P_T + P_R) + \frac{\tau_r}{2}(P_R + 2P_I)\end{aligned} \tag{2.38}$$

Table 2.1: Simulation Parameters

Parameter	Value
Packet length	1534 bytes
Data/ Control Transmission Rate	24/ 6 Mbps
Retransmission Rate	54 Mbps
Transmission/ Reception/ Idle Power	1900/ 1340/ 1340 mW
SIFS/ DIFS	10/ 50 μsec
k (parts of a frame)	100

and in the case of imperfect synchronization we have:

$$\begin{aligned}\bar{\mathcal{E}}_{C_2} &= \tau_s(P_T+2P_R)+(\frac{\tau_s}{k}\bar{l_c}-\tau_s)(P_R+2P_I)+2\tau_{SIFS}(3P_I)\\ &+\frac{\tau_{CFC}}{2}(P_T+2P_R)+\frac{\tau_{CFC}}{2}(3P_I)+\tau_{DIFS}(3P_I)\\ &+\frac{\tau_r}{2}(P_I+P_T+P_R)+\frac{\tau_r}{2}(P_R+2P_I) \qquad (2.39)\end{aligned}$$

In both expressions (2.38) and (2.39) we calculated the consumed energy taking into account the discrete parts of a collision. In our calculations we include the fact that half of the retransmissions are done by RN. The expected energy consumption of collision periods, based on the frequency that collision cases occur can be expressed as:

$$\begin{aligned}\mathcal{E}_C &= (\frac{1}{2k-1}\bar{\mathcal{E}}_{C_1}+\frac{2k-2}{2k-1}\bar{\mathcal{E}}_{C_2})\frac{E_N[T_c]}{\bar{\tau_c}} \Rightarrow\\ \mathcal{E}_C &= (\frac{1}{2k-1}\bar{\mathcal{E}}_{C_1}+\frac{2k-2}{2k-1}\bar{\mathcal{E}}_{C_2})\frac{p_e}{1-p_N} \qquad (2.40)\end{aligned}$$

2.4.4 CooPNC Performance Evaluation

To evaluate and validate the performance of CooPNC protocol, we performed simulations using MATLAB^{TM}, for the scenario of one central and three peripheral networks. The simulations were performed for diverse traffic combinations assuming constant packet length. We simulated repeatedly the transmission of a 50 MBytes file in each network with 24Mbps multicast rate, 6 Mbps rate for the control packets transmission and 54 Mbps for the relays' retransmissions. For the energy efficiency evaluation we used the average values of transmission, reception and idle power levels measured in [53]. Namely, $P_T = 1900mW$ and $P_R = P_I = 1340mW$. The simulation parameters are presented in Table 2.1. We also compared CooPNC with two state of the art multicast protocols, the non-cooperative protocol RAMP [18] and the cooperative protocol CRTC [51] which uses analog network coding.

We evaluated RAMP in our scenario by considering its extra control frames.

A Multicast Request to Send (MRTS) frame is sent by each BS to its multicast group, including coordination information for the sequence of the responses of the receivers. According to that coordination, each intended receiver sends its Multicast Clear to Send (MCTS). After the transmission of multicast data frame, the receivers send their Multicast Acknowledgement (MACK) frames with the same order. A retransmission is done only if one or more MACKs are not received by the BS.

Aiming to evaluate CRTC and compare it to CooPNC, we extend its operation according to our scenario considering three collision areas instead of one. CRTC forces the retransmissions to be held by the BSs at the same rate as the multicast transmissions by pairs of neighbouring BSs, aiming to generate a collision only in the cases that the two STAs of an overlapping area have received the other of the intended packets. As both protocols assume that retransmissions are held by the BSs, the relays in the evaluation of RAMP and CRTC are considered as normal receivers, situated outside the collision areas.

Following, we provide the results of the mathematical analysis along with the simulation outcomes for the central network in Section 2.4.5 and for the peripheral networks in Section 2.4.6, both performed for $N = 4$ as presented in Fig. 2.10.

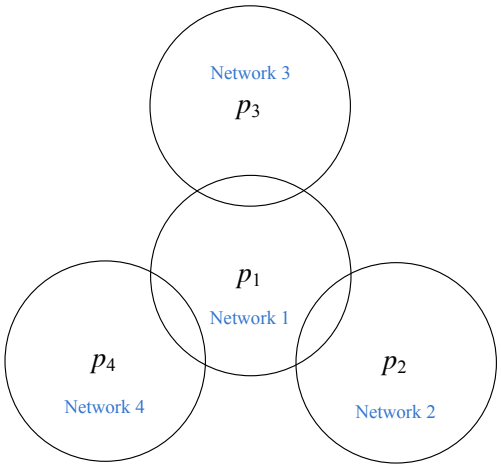


Figure 2.10: Evaluation scenario of one central and 3 overlapping networks ($N = 4$).

2.4.5 Performance Evaluation of the Central Network

Hereunder, the graphical representations of the throughput, the delay and the energy efficiency are provided for the central network. The results refer to high traffic conditions at the central network ($p_1 = 0.95$). We keep the traffic of Network 4 constant at medium level ($p_4 = 0.5$) and variate the traffic of Networks 2 and 3, ($p_2, p_3 \in (0, 1)$) aiming to show their impact on the performance of the central network. In Figs. 2.11, 2.12 and 2.13 the throughput, the delay and the energy efficiency of the central network is presented in comparison to the performance of RAMP and CRTC protocols. We notice that for $(p_1, p_2, p_3, p_4) = (0.95, 0.95, 0.5, 0.5)$

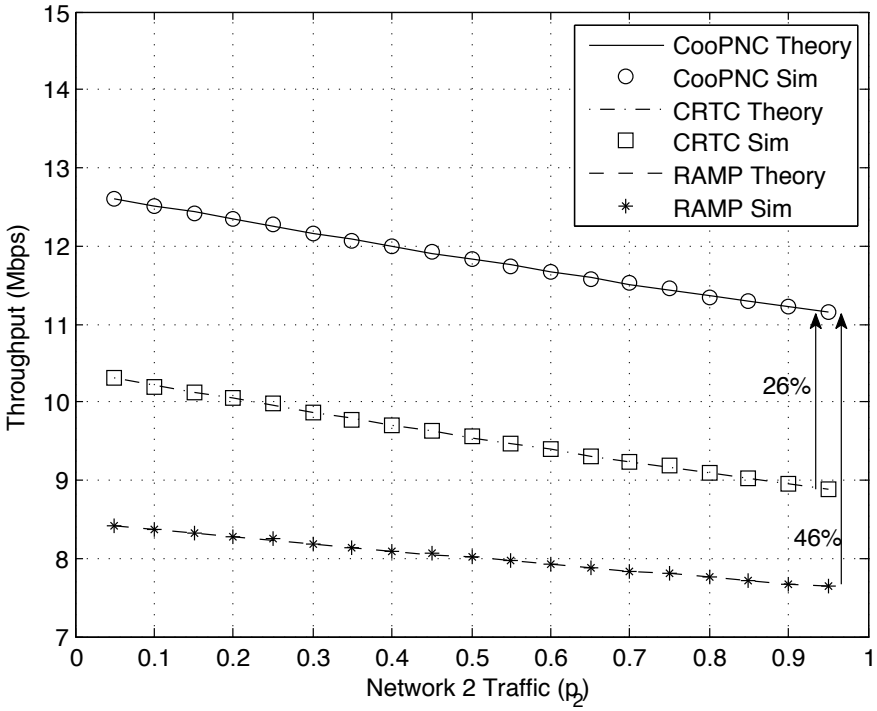


Figure 2.11: Throughput of the central network for $N = 4$, with $(p_1, p_3, p_4) = (0.95, 0.5, 0.5)$ and $p_2 \in (0, 1)$.

CooPNC presents 46% better throughput than RAMP and 26% better throughput than CRTC. For the same conditions CooPNC presents 37% less delay than RAMP and 25% less delay than CRTC. The exploitation of ANC along with the proposed indirect cooperative scheme in CooPNC leads to 23% improvement in energy efficiency compared to RAMP and 14% improvement compared to CRTC. The higher energy efficiency, the lower delay and the higher throughput of our protocol in comparison with the state of the art is a result of the combination of ANC with the proposed cooperative scheme, as collisions are resolved by just one retransmission and the retransmissions are shared between the relays of the neighbouring networks. Otherwise, in every collision that a pair of STAs of an overlapping area receives, one retransmission from each one of the corresponding relays should follow to resolve the collision in both STAs.

In Table 2.2 we present the performance gain of the central network for various traffic conditions of the 4 interacting networks. The first case refers to low traffic at the central network and medium traffic at the peripheral networks, the second to medium traffic at all 4 networks, the third to high traffic at the central and one peripheral network and medium traffic at the other two peripheral networks and the last case refers to high traffic at the central and two peripheral networks and medium traffic at one peripheral network. In all the evaluated protocols, the central network faces the same plurality of collisions, leading to different delay levels but constant delay gain, as it is only affected by the needed time to resolve a collision.

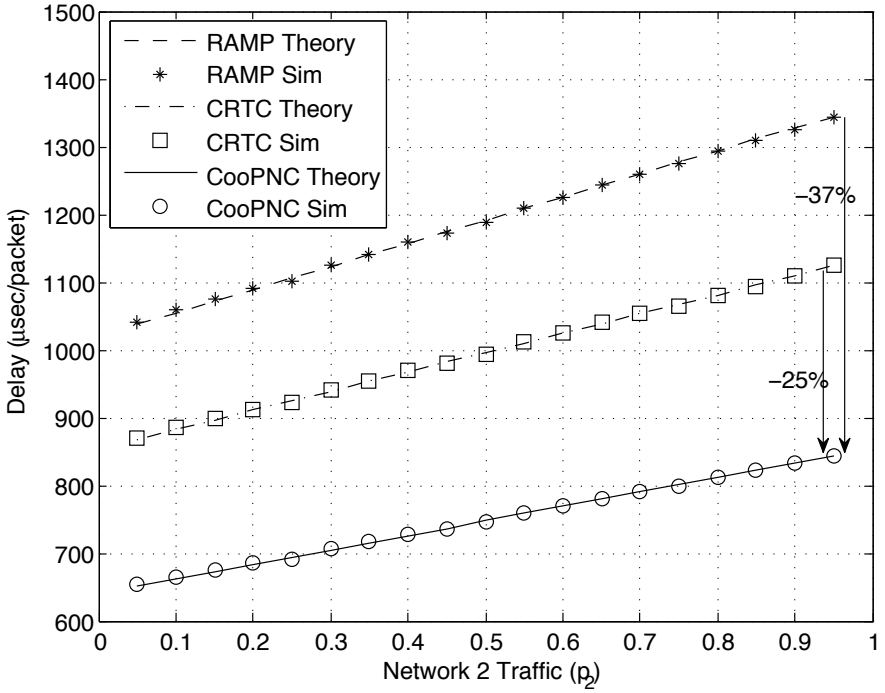


Figure 2.12: Delay of the central network for $N = 4$, with $(p_1, p_3, p_4) = (0.95, 0.5, 0.5)$ and $p_2 \in (0, 1)$.

It is noteworthy to mention that for high traffic at the central network and for high traffic at more than one peripheral networks, as shown in the last combination, the energy efficiency is degraded as more peripheral STAs of Network 1 suffer from collisions. The throughput gain, though, remains stable as it is already degraded if only one of the peripheral networks presents high traffic. When more than one peripheral networks present high traffic, the throughput is not significantly affected as it is already close to its lower bound but the energy efficiency is decreasing, as more STAs are involved in collisions.

2.4.6 Performance Evaluation of the Peripheral Networks

We evaluate the impact of CooPNC on the performance of the peripheral networks for the same scenario of one central and two peripheral networks ($N = 4$). We keep the traffic of Network 4 constant at high level ($p_4 = 0.95$) and of Network 3 at medium level ($p_3 = 0.5$) and we perform simulations for all traffic combinations of Networks 1 and 2.

In Fig. 2.14, we notice that Network 4 suffers from low throughput under the conditions of low traffic level in Network 2 ($p_2 = 0.05$) and for high traffic level in the central network ($p_1 = 0.75$), because it partially overlaps with the central network. Under these circumstances, where frequent collisions occur with Network 1, the positive impact of CooPNC is revealed, as the collisions are resolved with

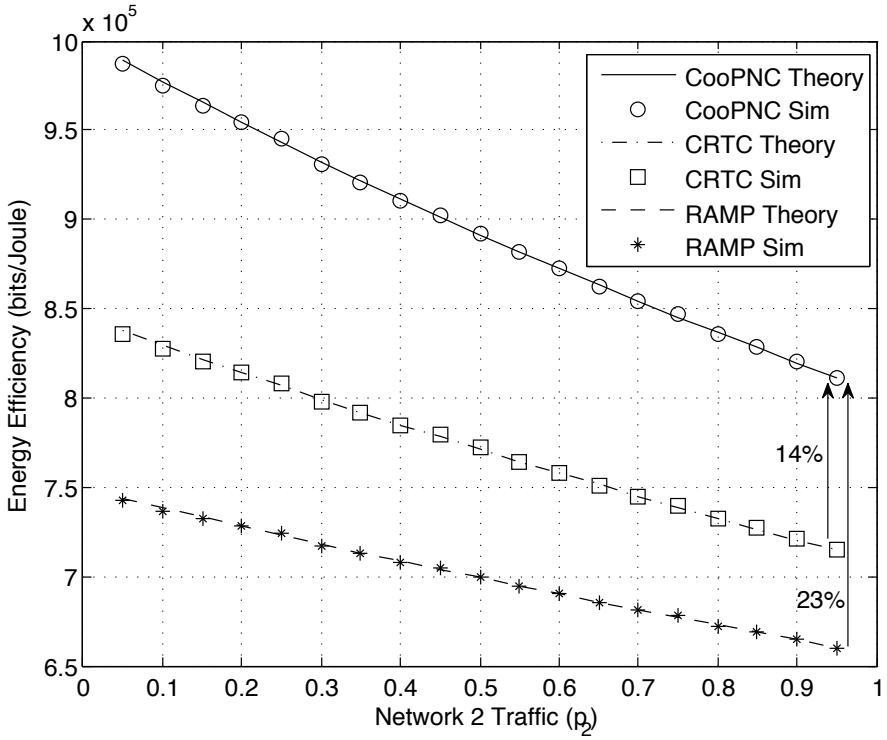


Figure 2.13: Energy efficiency of the central network for $N = 4$, with $(p_1, p_3, p_4) = (0.95, 0.5, 0.5)$ and $p_2 \in (0, 1)$.

our proposed cooperative scheme. Indeed, for $(p_1, p_2, p_3, p_4) = (0.75, 0.05, 0.5, 0.95)$, Network 4 presents the highest gain in energy efficiency compared to CRTC (21%), as shown in Fig. 2.15, where the relative gain in energy efficiency of CooPNC is presented when compared to CRTC. As the traffic of the central network p_1 increases, the plurality of collisions with Network 4 also increases. At the same time, the plurality of collisions augments between the central network and Network 3, creating opportunities for the peripheral Network 4 to transmit. This is why the throughput of Network 4 augments for $p_1 > 0.75$, as shown in Fig. 2.14. Aiming to illustrate the impact of the different traffic levels of one peripheral network to the performance of another, we provide the throughput, the delay and the energy efficiency of Network 4 in Figs. 2.16, 2.17 and 2.18 respectively. We keep the traffic levels of Network 1, Network 3 and Network 4 constant $(p_1, p_3, p_4) = (0.75, 0.5, 0.95)$ and we present our results for all traffic levels of Network 2 ($p_2 \in (0, 1)$). We notice that for $(p_1, p_2, p_3, p_4) = (0.75, 0.05, 0.5, 0.95)$ CooPNC presents 60% and 20% better throughput compared to RAMP and CRTC respectively. It is also 75% and 21% more energy efficient than RAMP and CRTC respectively, while it presents 47% and 25% less delay. As the peripheral networks present symmetric characteristics, the performance evaluation of Network 4 is alike for the other peripheral Networks of the evaluated scenario. In Table 2.3 we present the performance gain in throughput, delay and energy efficiency of Network 4, for low ($p_3 = 0.05$), medium ($p_3 = 0.5$) and high traffic ($p_3 = 0.95$) in Network 3 and for $(p_1, p_2, p_4) = (0.75, 0.05, 0.95)$. In this

Table 2.2: Performance Gain of Network 1 for Diverse Traffic Levels of the Peripheral Networks

$(p_1, p_2, p_3, p_4) = (0.05, 0.5, 0.5, 0.5)$	Throughput	Delay	Energy Eff.
CooPNC vs RAMP	4.6%	-37%	3.5%
CooPNC vs CRTC	2.3%	-25%	2%
$(p_1, p_2, p_3, p_4) = (0.5, 0.5, 0.5, 0.5)$	Throughput	Delay	Energy Eff.
CooPNC vs RAMP	32.5%	-37%	20.5%
CooPNC vs CRTC	16.3%	-25%	11.7%
$(p_1, p_2, p_3, p_4) = (0.95, 0.95, 0.5, 0.5)$	Throughput	Delay	Energy Eff.
CooPNC vs RAMP	46%	-37%	23%
CooPNC vs CRTC	26%	-25%	14%
$(p_1, p_2, p_3, p_4) = (0.95, 0.95, 0.95, 0.5)$	Throughput	Delay	Energy Eff.
CooPNC vs RAMP	45%	-37%	18%
CooPNC vs CRTC	25%	-25%	9%

table the interplay between the traffic of Network 3 and the performance of Network 4 is revealed. As traffic increases in Network 3, the plurality of its collisions with Network 1 also increases, creating more opportunities for Network 4 to transmit. This is the reason why the performance gain in throughput and energy efficiency of Network 4 increases with the traffic increase of another peripheral network.

Table 2.3: Impact of Network 3 traffic on the Performance Gain of Network 4

$p_3 = 0.05$	Throughput	Delay	Energy Efficiency
CooPNC vs RAMP	60%	-47%	75%
CooPNC vs CRTC	20%	-25%	21%
$p_3 = 0.5$	Throughput	Delay	Energy Efficiency
CooPNC vs RAMP	54%	-55%	64%
CooPNC vs CRTC	11%	-25%	12%
$p_3 = 0.95$	Throughput	Delay	Energy Efficiency
CooPNC vs RAMP	47.5%	-58%	54%
CooPNC vs CRTC	7.8%	-25%	8%

For our results we have assumed perfect channel conditions and comparable reception power levels at the STAs that suffer from the same collisions. According to the experimental results of ZigZag presented in [50], over a threshold of 11dB in

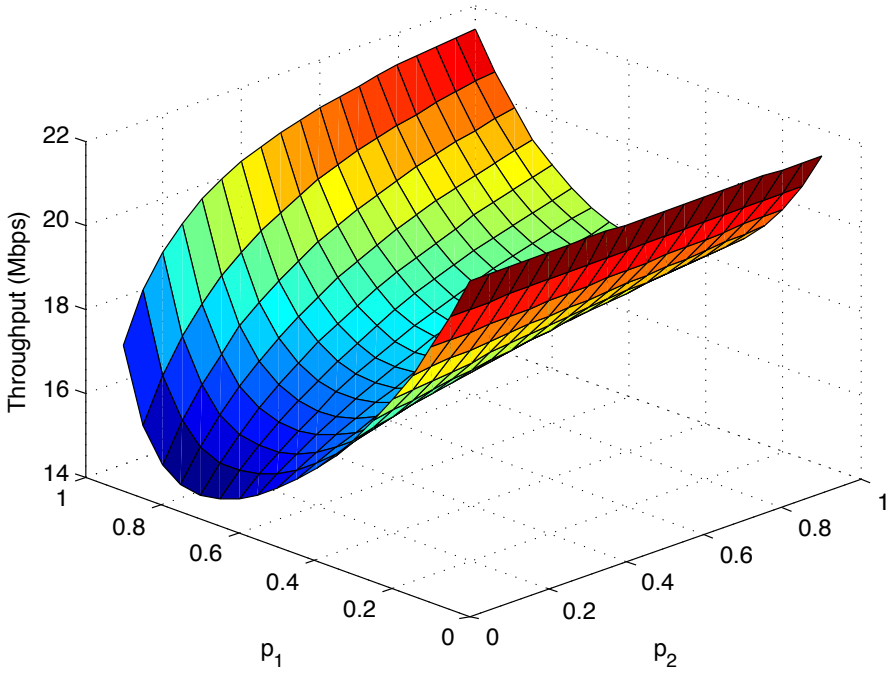


Figure 2.14: Throughput of Network 4 for $N = 4$, with $(p_3, p_4) = (0.5, 0.95)$ and $(p_1, p_2) \in (0, 1)$.

the difference of the signal reception power at two STAs belonging to the overlapping area of two neighbouring networks, due to the capture effect the performance of the network with the lower reception power significantly degrades. The capture effect occurs when a stronger packet, a packet with higher signal strength, comes before a weaker packet, a packet with lower signal strength, because the radio locks onto the stronger packet and the weaker signal may not cause substantial interference. Based on these experimental results, in Fig. 2.19 we present the degradation of the throughput performance of the STA of Network 4 of the evaluated scenario, for which we assume that it suffers from weaker signal from its BS. In Fig. 2.20 we show how the delay in Network 4 is affected. We use $SINR_1$ to represent the difference between the received power level and the interference in the overlapping area where the two STAs, associated with two neigbouring BSs are situated. For the ANC feature to operate efficiently we should have $SINR_1 \leq -11dB$.

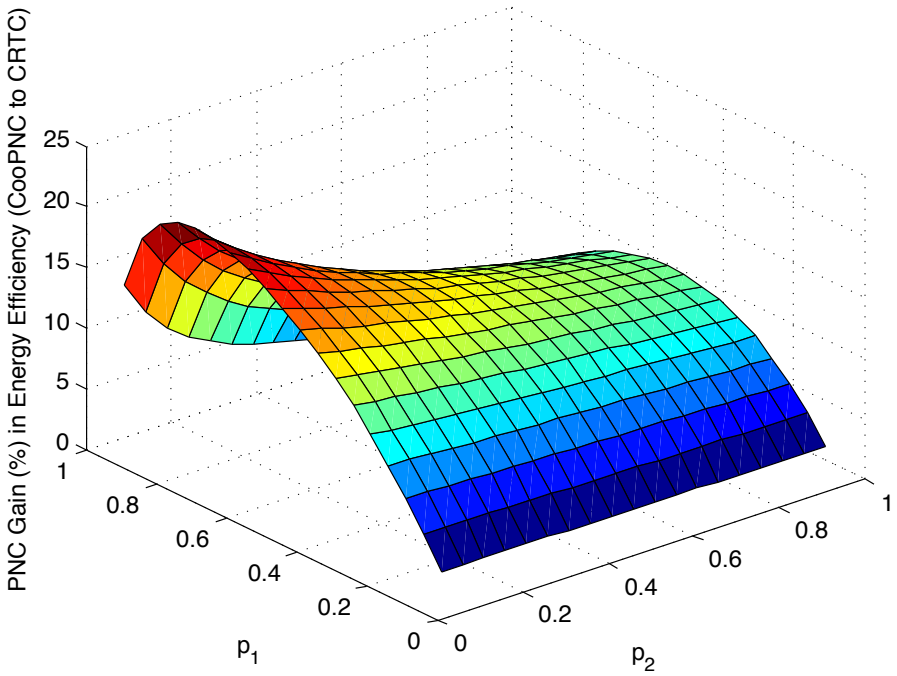


Figure 2.15: CooPNC gain(%) in energy efficiency compared to CRTC.

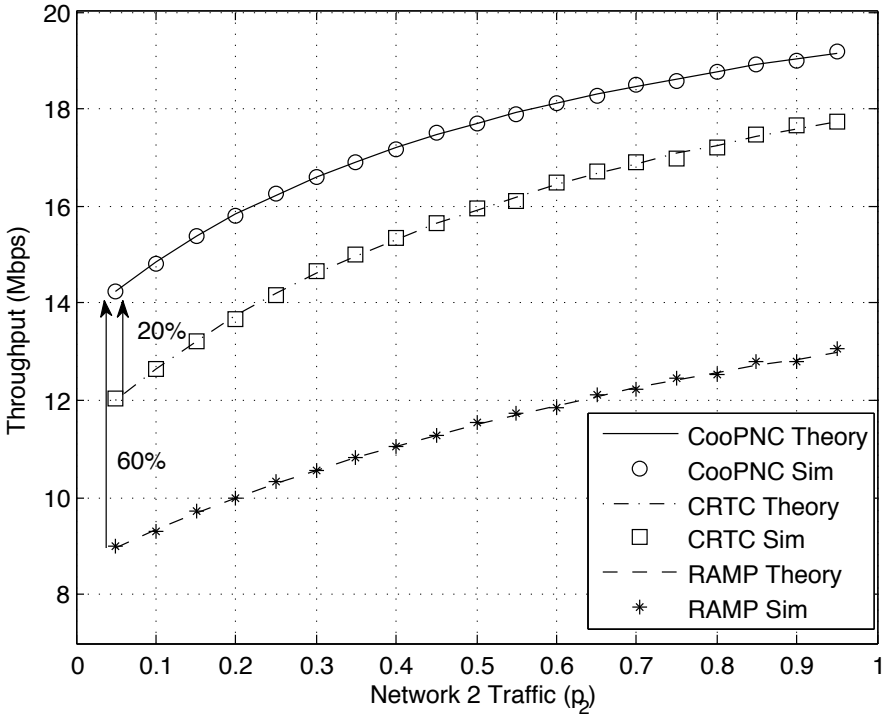


Figure 2.16: Throughput of Network 4 for $(p_1, p_3, p_4) = (0.75, 0.5, 0.95)$ and $p_2 \in (0, 1)$.

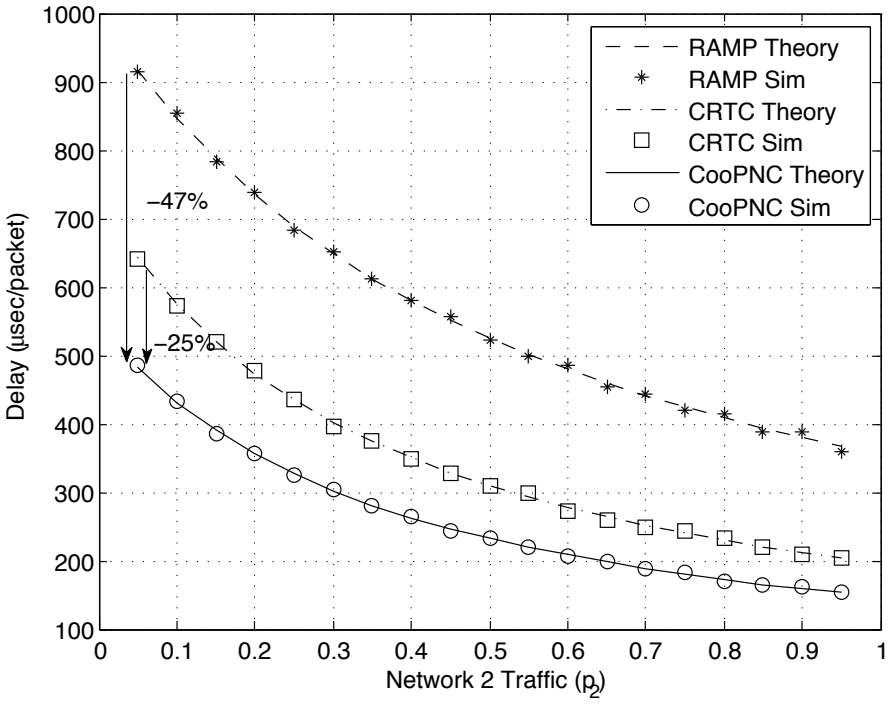


Figure 2.17: Delay of Network 4 for $(p_1, p_3, p_4) = (0.75, 0.5, 0.95)$ and $p_2 \in (0, 1)$.

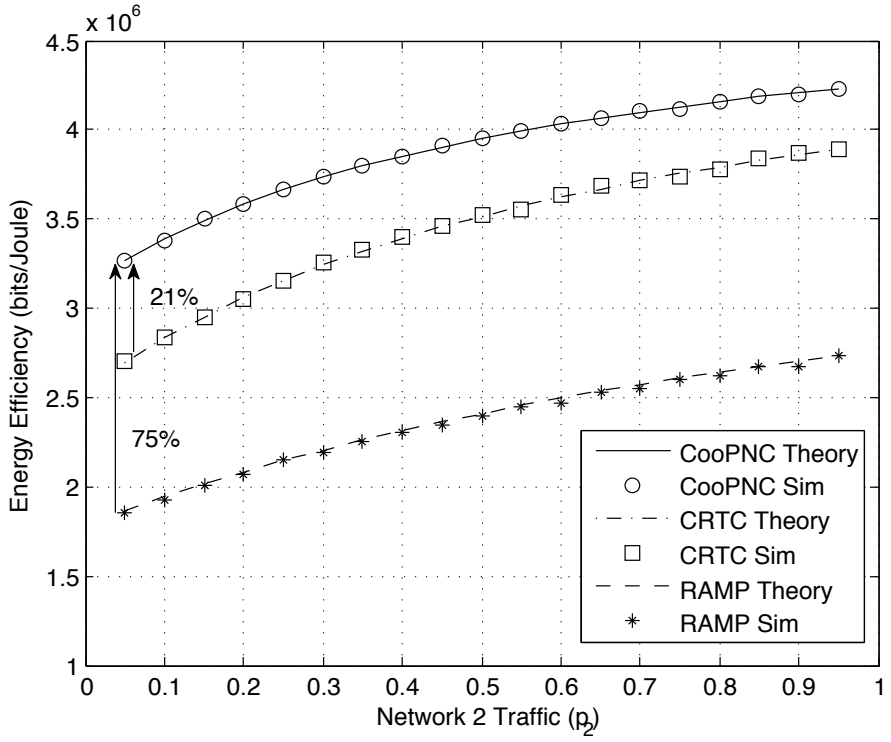


Figure 2.18: Energy efficiency of Network 4 for $(p_1, p_3, p_4) = (0.75, 0.5, 0.95)$ and $p_2 \in (0, 1)$.

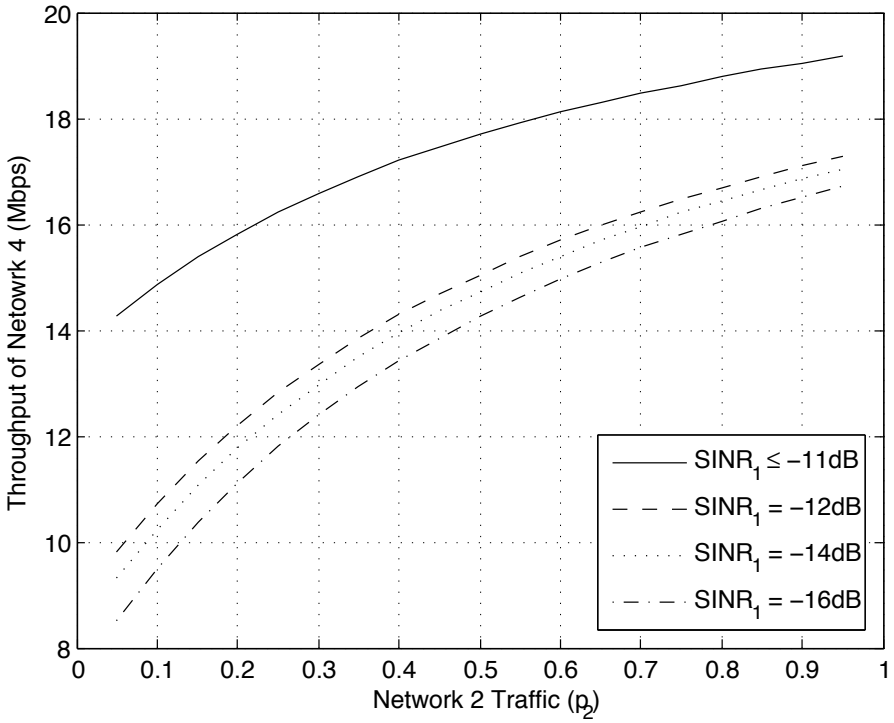


Figure 2.19: Impact of the capture effect on the throughput of Network 4.

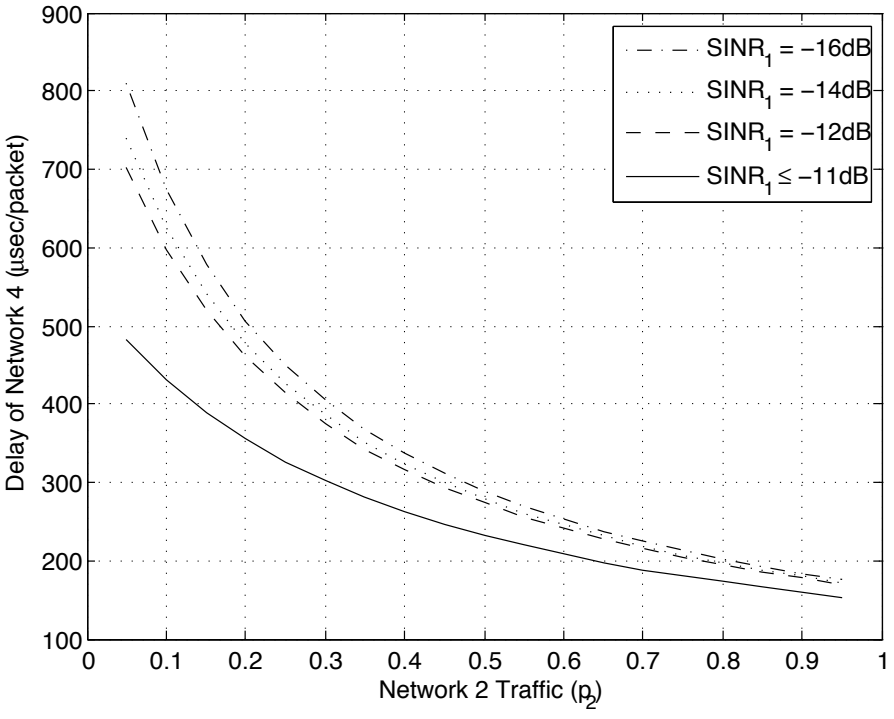


Figure 2.20: Impact of the capture effect on the delay of Network 4.

Chapter 3

Digital and Physical Layer Network Coding Performance in the Context of Enforced Fairness

In the previous Chapter, we focused on the combination of a cooperative collision resolution technique that exploits PNC to resolve collisions and provide a reliable multicast communication for partially overlapping short range networks. In this Chapter we investigate the upper performance bounds of PNC in a broadly used scenario for benchmarking the performance of diverse Network Coding techniques. Following, we present both the digital NC and the ANC/ PNC throughput performance upper bounds, for networks operating under protocols that provide fairness in medium access and we provide an energy efficiency assessment of each NC technique. We focus on the Cross Network topology operating under several NC techniques and we investigate the impact of the MAC layer fairness on the network's throughput. The main contribution of this Chapter is the investigation of the upper performance bounds of PNC for the Cross Network scenario in comparison with other NC techniques.

The Chapter is organized as follows. In Section 3.1 we present a literature overview regarding network coding throughput performance in Section 3.2 the system under investigation and the performance analysis of different network coding techniques. We finalise this Chapter with Section 3.3, which contains the comparison of the different network coding techniques in terms of throughput and energy efficiency.

3.1 Background

In [54], Katti et al. presented COPE, which is an architecture for wireless mesh networks that largely increases network throughput. In COPE, opportunistic listening is introduced along with opportunistic coding from a relay node. The throughput gain is studied for several broadly adopted topologies. The first was the two-way relay topology, with two flows of information intersecting at the relay. In addition, the 'X' topology with four nodes and two intersecting flows at the relay, as well as the Cross topology with four nodes and four intersecting flows at the relay were also presented. In these topologies, the relay where the flows intersect is able to opportunistically code packets and broadcast them to the nodes of the network.

As noted in [55], the performance of the two-way relay, the 'X' and the Cross topology is strongly affected by the limitation of fair access to the wireless channel imposed by existing protocols like IEEE 802.11 [9]. In this work, the authors introduce an extra degree of freedom to the Cross topology by adding unicast independent flows from the central relay to the nodes and they propose the use of a virtual queue to give priority to coded packets aiming to improve the network throughput. While the research interest for applying digital NC in diverse topologies has been wide, little research effort has been devoted on practical solutions for embracing analog and physical layer network coding (ANC and PNC). While digital NC takes place at the network layer, ANC and PNC take place at the physical layer and packets are encoded by means of channel additivity.

3.2 System Under Investigation

We consider the Cross scenario as depicted in Fig. 3.1, where four nodes want to bidirectionally communicate in pairs through a fifth node that acts as a relay. In this scenario node A and C as well as B and D want to exchange packets through the relay R. We assume without loss of generality that the total bandwidth is 1 and that five nodes in the network share the wireless channel. The relay is only forwarding packets and does not generate any traffic. The bandwidth allocated to nodes A, B, C, D and R is denoted as BW_a, BW_b, BW_c, BW_d and BW_r, respectively. For the sake of simplicity we assume that nodes A, B, C and D contribute equally to the offered load of the network and the wireless channel is lossless. We study the throughput performance of the Cross scenario for the pure relaying case (without NC) and for several NC cases including both digital and physical layer NC.

Pure Relaying (without NC): In the case of pure relaying, every node transmits its packet to the relay and the relay forwards the packet to the intended receiver. If every node sends one packet to the relay, then four slots would be needed for the relay to receive the four packets. Following, the relay would need four more slots to deliver the packets to the corresponding nodes. Through this simple case it is shown that in order for the relay to be able to serve the offered load by the nodes, the bandwidth allocated to it should be equal to the sum of the bandwidth allocated to the transmitting nodes. The system throughput is always equal to BW_r. Due to

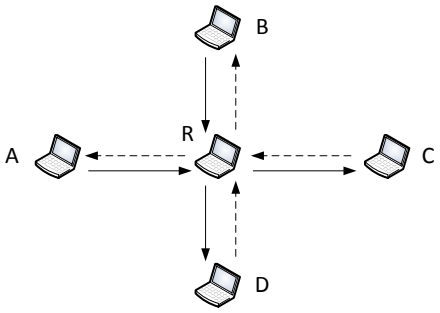


Figure 3.1: The Cross Scenario.

the MAC fairness of the IEEE 802.11 we can distinguish three different phases related to the offered load l by the nodes, $l = \sum_n BW_n$, for $n = \{a, b, c, d\}$. In the first phase, for $0 < l \leq 1/2$, the throughput increases linearly, as the sum of the offered load by the nodes and the served load by the relay is less than the channels capacity, and reaches its peak when:

$$BW_a = BW_b = BW_c = BW_d = 1/8 \text{ and } BW_r = 1/2. \tag{3.1}$$

At the peak point, the total throughput of the Cross is equal to $BW_r = 1/2$. In the second phase, as long as the total offered load l by nodes A, B, C and D is $1/2 < l \leq 4/5$, the total throughput linearly decreases until the channel is equally shared between the four nodes and the relay. In this range of offered load the served load by the relay, namely the system throughput, is equal to $1 - l$ and packets are being backlogged at the nodes. The last phase of saturation occurs for load $l > 4/5$. At this phase, due to the MAC fairness, the throughput is stabilised at $1/5$, while the bandwidth is equally shared between the four nodes and the relay. At this phase the assigned bandwidth at each node is:

$$BW_a = BW_b = BW_c = BW_d = BW_r = 1/5. \tag{3.2}$$

The maximum achievable throughput $C_1(l)$, depending on the offered load by nodes A, B, C and D, can be expressed as:

$$C_1(l) = \begin{cases} l, & \text{if } 0 < l \leq 1/2 \\ 1 - l, & \text{if } 1/2 < l \leq 4/5 \\ 1/5, & \text{if } l > 4/5 \end{cases} \tag{3.3}$$

Digital NC Without Overhearing (NC W/O OH): When the relay is able to code received packets in pairs, for every packet it transmits, two packets are delivered. We assume symmetric load at the nodes A, B, C and D and that the relay, after receiving a packet from A and C, broadcasts back a coded packet. The same procedure is followed for nodes B and D. Thus, six slots are needed for four packets to be delivered to the corresponding nodes. In the case of NC without

overhearing, the throughput linearly increases and reaches its peak when:

$$BW_a = BW_b = BW_c = BW_d = 1/6 \text{ and } BW_r = 1/3. \tag{3.4}$$

While every transmission from the relay delivers two packets, the peak throughput is 2/3. When the offered load is further increased, namely for $2/3 < l \leq 4/5$, the throughput is decreasing linearly and is equal to $2(1-l)$, while packets start being backlogged at the nodes. The phase of saturation for total load $l > 4/5$ presents a stable throughput of 2/5 and the assigned bandwidth at each node is:

$$BW_a = BW_b = BW_c = BW_d = BW_r = 1/5. \tag{3.5}$$

In the case of Digital NC Without Overhearing, the maximum achievable throughput $C_2(l)$ can be expressed as:

$$C_2(l) = \begin{cases} l, & \text{if } 0 < l \leq 2/3 \\ 2(1-l), & \text{if } 2/3 < l \leq 4/5 \\ 2/5, & \text{if } l > 4/5 \end{cases} \tag{3.6}$$

Digital NC With Overhearing (NC W OH): In this case we assume that every node is able to overhear the transmissions of all the other nodes except its opposite. For example, when node A is transmitting, node B and D can overhear the transmission but not the intended receiver, which in this case is node C. As the nodes intend to communicate in opposite pairs, as soon as the relay receives one packet from every node, it produces a coded packet that it broadcasts to all nodes. Thus, one transmission by the relay is enough for four packets to be delivered to their intended nodes. Five slots are adequate for four packets to be delivered to the corresponding nodes. Here, the throughput linearly increases and reaches its peak when:

$$BW_a = BW_b = BW_c = BW_d = BW_r = 1/5. \tag{3.7}$$

While every transmission from the relay delivers four packets, the peak throughput is equal to 4/5. At this point, the medium is also equally shared between the four nodes and the relay, leading to the fact that for total offered load $l > 4/5$, the throughput stays stable at 4/5. The maximum achievable throughput of this case $C_3(l)$ can be written as:

$$C_3(l) = \begin{cases} l, & \text{if } 0 < l \leq 4/5 \\ 4/5, & \text{if } l > 4/5 \end{cases} \tag{3.8}$$

PNC in Pairs Without Overhearing (PNC in Pairs W/O OH): In this case, we assume that the nodes that want to communicate, namely the pairs (A, C) and (B, D), are able of transmitting their packets concurrently. The relay is also able to resolve the provoked collisions and transmit back a coded version of the collided packets of each pair. With this procedure two collisions from the node pairs and two coded packets from the relay would be adequate for four packets to be

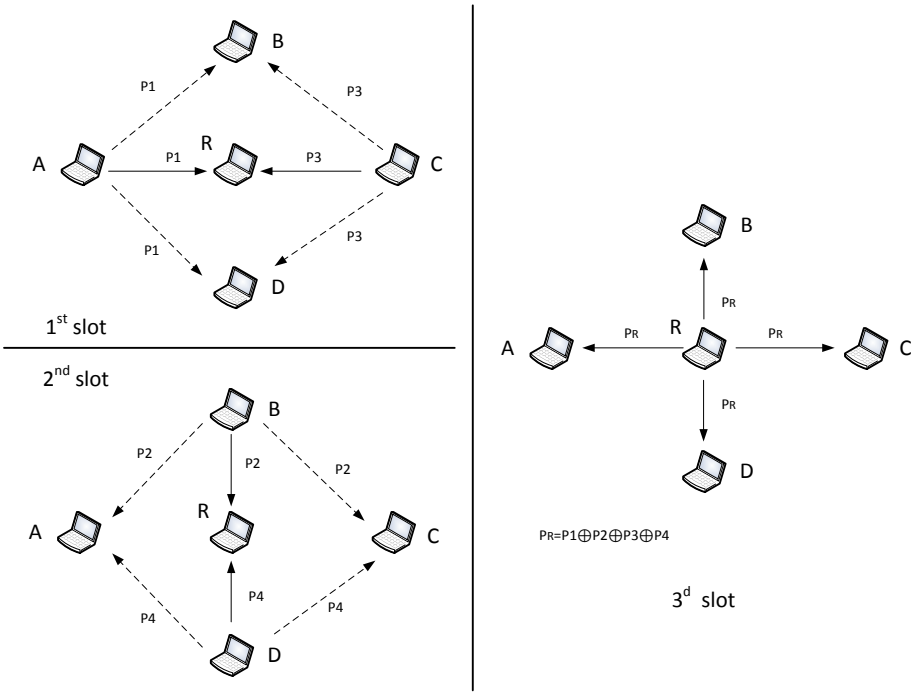


Figure 3.2: PNC in pairs with overhearing.

delivered. This means that the relay would need half of the bandwidth to deliver the packets to the corresponding nodes. We assume that the bandwidth is shared to three entities of the network, the pair (A, C) with bandwidth $BW_{a,c}$, the pair (B, D) with bandwidth $BW_{b,d}$ and the relay. For the case that the offered load by the pairs is symmetric and the pair transmissions synchronised, the throughput would increase linearly until it reaches its peak when:

$$BW_{a,c} = BW_{b,d} = 1/4 \text{ and } BW_r = 1/2. \tag{3.9}$$

At this point the throughput is equal to 1, as for every transmission by the relay two packets are delivered. For every transmitting pair two packets are also delivered and the total offered load l is equal to 1. When $1 < l \leq 4/3$, the throughput decreases and is equal to $2 - l$. For $l > 4/3$ the throughput stays stable at $2/3$. We express the maximum achievable throughput $C_4(l)$ for this case as:

$$C_4(l) = \begin{cases} l, & \text{if } \ 0 < l \leq 1 \\ 2 - l, & \text{if } \ 1 < l \leq 4/3 \\ 2/3, & \text{if } \ l > 4/3 \end{cases} \tag{3.10}$$

PNC in Pairs With Overhearing (PNC in Pairs W OH): When overhearing is possible, every collision provoked by the transmitting pair is also received by the non transmitting pair. Consequently, when every pair has transmitted, the relay has received four packets by the nodes and every node has received all pack-

ets except the one that was destined to it. Following, the relay broadcasts a coded version of all the four packets, by combining the two received collisions, and each node is able to receive its corresponding packet. This procedure is shown in Fig. 3.2. In the first slot, two packets from node A and C are delivered to the relay, in the second slot two packets from node B and D are delivered to the relay, and in the third slot one coded packet that includes the four previously received packets is broadcasted by the relay to all the nodes. In this case, the channel is equally shared among the pairs (A, C), (B, D) and the relay, and the throughput is linearly increased until the channel is equally shared between the three entities of the network where it presents its peak, namely when:

$$BW_{a,c} = BW_{b,d} = BW_r = 1/3. \tag{3.11}$$

At this point, as for every transmission by the relay four packets are delivered, the throughput is equal to 4/3 and remains stable as the offered load by the pairs is increasing. The maximum achievable throughput $C_5(l)$ in this case equals:

$$C_5(l) = \begin{cases} l, & \text{if } 0 < l \leq 4/3 \\ 4/3, & \text{if } l > 4/3 \end{cases} \tag{3.12}$$

The use of these NC techniques in the Cross Network presents significant gains in terms of energy efficiency compared to pure relaying. In order to assess this gain, we distinguish three different states for every node, namely idle, transmitting and receiving states with their corresponding durations for the case that every node wants to send to its opposite a file of the same specific size. We denote as α_i^A, α_t^A and α_r^A, the idle, transmission and reception time respectively, that is needed by node A to transmit a certain amount of traffic to node C and to receive an equal amount of traffic from node C. For instance, node A is in idle state for a duration equal to α_i^A, in transmitting state for α_t^A and in receiving state for α_r^A until its file is delivered to the opposite node C and a file of the same size by node C is received through the relay. We assume that all the transmissions for this interchange of equal files is done under the same rate. The energy consumed by node A is equal to $E_A = \alpha_i^A P_I + \alpha_t^A P_T + \alpha_r^A P_R$, where P_I, P_T and P_R are the power levels of the wireless interface cards of the nodes when idle, transmitting and receiving respectively. The time variables α_i^A, α_t^A and α_r^A depend on the NC case that is used. Following the same rationale, we express the energy consumed by the relay as $E_R = \alpha_i^R P_I + \alpha_t^R P_T + \alpha_r^R P_R$, where α_i^R, α_t^R and α_r^R express the time that the relay is in idle, transmitting and receiving state respectively until the four nodes of the Cross Network are served.

3.3 Evaluation of Different Network Coding Techniques

We evaluate the performance of each presented case of Section II and we provide in Fig. 3.3 the representation of the analysis of the throughput performance of the Cross Network for all possible offered loads as described in equations (3.3), (3.6), (3.8), (3.10) and (3.12). The presented throughput is considered for symmetric load, for lossless channel and for channel capacity equal to 1 Mbps. It is noticeable that for the cases of pure relaying, NC without overhearing and PNC in pairs without overhearing, as the offered load is increased, the throughput presents a peak and decreases to a level where it stays stable. On the other hand, NC with overhearing and PNC with overhearing present a peak and their throughput keeps stable at the higher performance as the offered load is increasing. This happens because the point of maximum throughput coincides with the point where the medium is equally shared between the entities of the network. In the case of NC with overhearing we have five entities in the network, namely, the four nodes and the relay, and the Cross presents maximum throughput when each entity shares one fifth of the bandwidth. In the PNC with overhearing case the network is comprised by two pairs of nodes and the relay, and the maximum throughput is achieved when these three entities share one third of the bandwidth.

Aiming to asses the energy efficiency of each presented case of Section II, we assume that every node of the Cross Network has a 10 MB file to send to the opposite node and that the transmission rate is 54 Mbps. The transmitted frames are of 1534 bytes long with a payload of 1500 bytes. We calculate the energy efficiency of one of the four symmetrical nodes and of the relay in bits per Joule. For our evaluation of the energy efficiency of the nodes of the Cross Network, we assumed that the average values of idle, transmission and reception power levels are as measured in [53], namely, $P_I = P_R = 1340$ mW and $P_T = 1900$ mW. We take into consideration the fact that the relay has to serve four flows of 10 MB each.

Regarding the saturation throughput, we notice in Fig. 3.3 that NC without overhearing presents two times the performance of pure relaying. This happens because with pure relaying in saturation conditions, due to enforced fairness, the relay occupies the channel for 1/5 of the time, leading to system throughput equal to 1/5 of the capacity, while with NC without overhearing, the relay occupies the channel for 1/5 of the time but it leads to system throughput equal to 2/5 of the capacity as two packets are delivered for every transmission. NC with overhearing presents four times higher performance than that of pure relaying, as in saturation conditions while it occupies the channel for 1/5 of the time it delivers 4 packets per transmission leading to a system throughput equal to 4/5 of the capacity. In addition, we notice that the saturation throughput of PNC in pairs with overhearing is two times better than that of PNC in pairs without overhearing. In PNC in pairs with overhearing the channel is shared between three entities, the relay and the two node pairs. When the relay occupies the channel for 1/3 of the time, it achieves a throughput equal to 2/3 of the capacity as two packets are delivered for every transmission. On the other hand, PNC in pairs with overhearing is delivering four

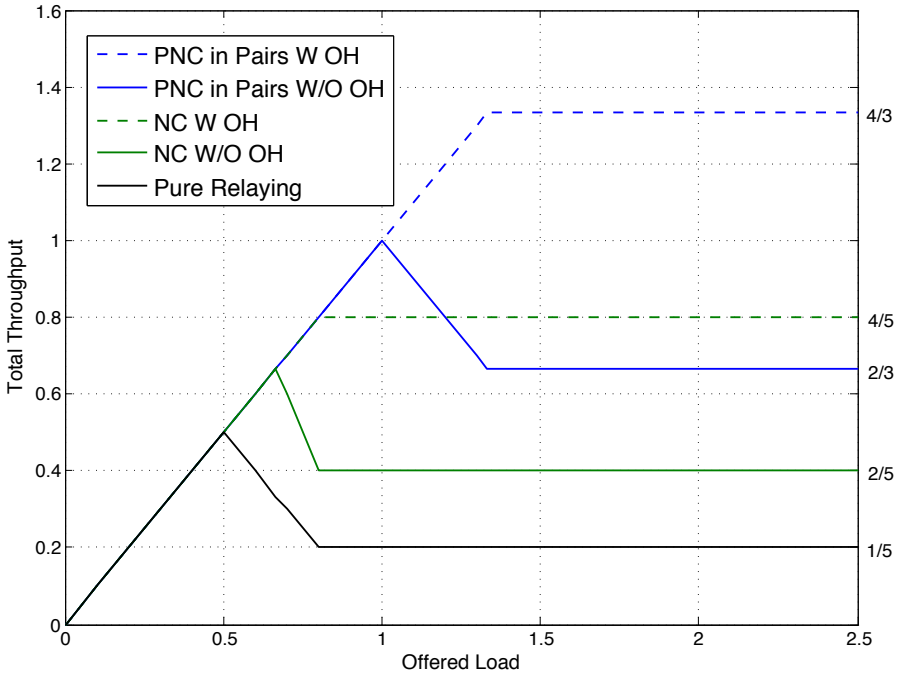


Figure 3.3: Throughput comparison of the different NC cases.

packets per transmission while occupying the channel for 1/3 of the time, achieving a throughput equal to 4/3 of the capacity, two times higher than PNC in pairs without overhearing.

We compare the performance of the NC techniques with the performance of pure relaying in terms of coding gain. We denote as coding gain the number of packets the relay would send with pure relaying divided by the number of packets sent with a NC technique for the transmission of the same amount of data. In Fig. 3.4, the coding gain of digital NC without overhearing, digital NC with overhearing, PNC in pairs without overhearing and PNC in pairs with overhearing is presented. In saturation conditions NC without overhearing present coding gain equal to two, while NC with overhearing has a coding gain equal to four. For the same conditions PNC without overhearing presents coding gain equal to 3.3 and PNC with overhearing a coding gain equal to 6.6.

The performance of pure relaying, NC without overhearing and PNC in pairs without overhearing, presented in Fig. 3.3, presents a peak and then decreases until the medium is equally shared to the network entities of each case. An optimized approach for these cases would use scheduling or flow control to maintain the network offered load at the peak performance point. The throughput performance of these specific NC cases along with the non decreasing cases is presented in Fig. 3.5.

For the case of PNC in pairs with and without overhearing, the throughput in Fig. 3.3 presents their performance under ideal lossless channels and scheduling of the transmitting entities. The assumption of ideal channels will be relaxed in future

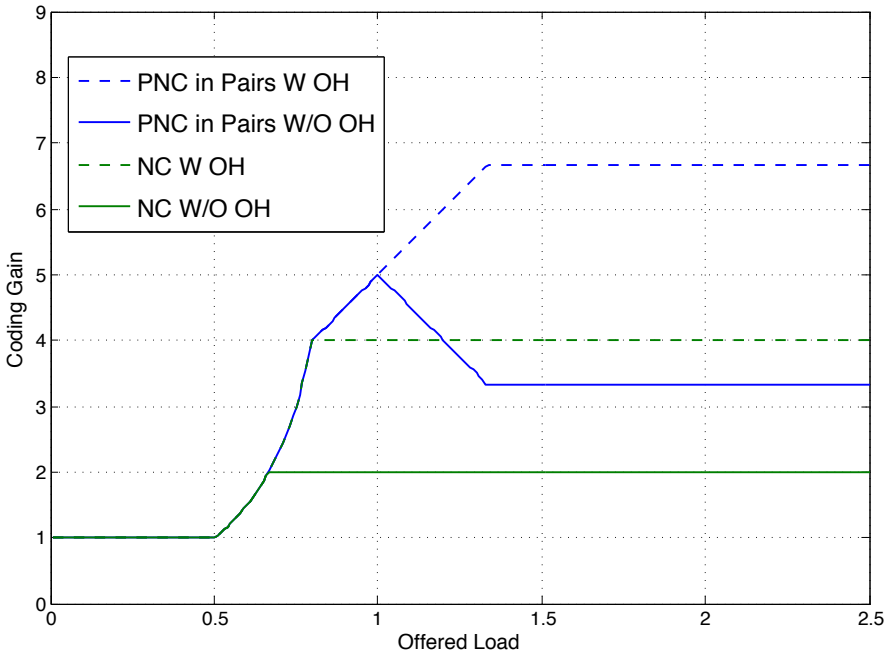


Figure 3.4: Coding gain of the different NC cases.

work. In a realistic approach, the scheduling should be done by the relay node, that would be aware of coding opportunities.

In Fig. 3.6 and 3.7, we present the results of the energy efficiency analysis of node A and of the relay for the pure relaying and every NC case for the load that the Cross Network presents maximum throughput. To that end, we present the energy efficiency of the pure relaying for $l = 1/2$, of the Digital NC without overhearing for $l = 2/3$, of the Digital NC with overhearing for $l = 4/5$, of the PNC without overhearing for $l = 1$ and of the PNC with overhearing for $l = 4/3$. Through this comparison we notice in Fig. 3.6 that with PNC with overhearing node A presents 31.4%, 63.2%, 94.6% and 158.2% better energy efficiency from PNC without overhearing, NC with overhearing, NC without overhearing and pure relaying respectively. In Fig. 3.7, we notice that with PNC with overhearing the relay presents 28.7%, 47.2%, 85.9% and 162.9% better energy efficiency from PNC without overhearing, NC with overhearing, NC without overhearing and pure relaying, respectively.

The throughput decrease that is observed after the peak in the cases of pure relaying, NC without overhearing and PNC in pairs without overhearing is due to the fact that there is room for the nodes to increase the offered load until the wireless medium is shared equally among the network entities. When the total offered load reaches the point where the medium is equally shared, the throughput of each case stays stable regardless of any further increase in the offered load. Under these conditions packets are backlogged and dropped when the queues exceed the buffers capacities. New MAC protocols could by applied for these schemes to maintain the throughput to its peak by applying flow control or by provisioning for explicit

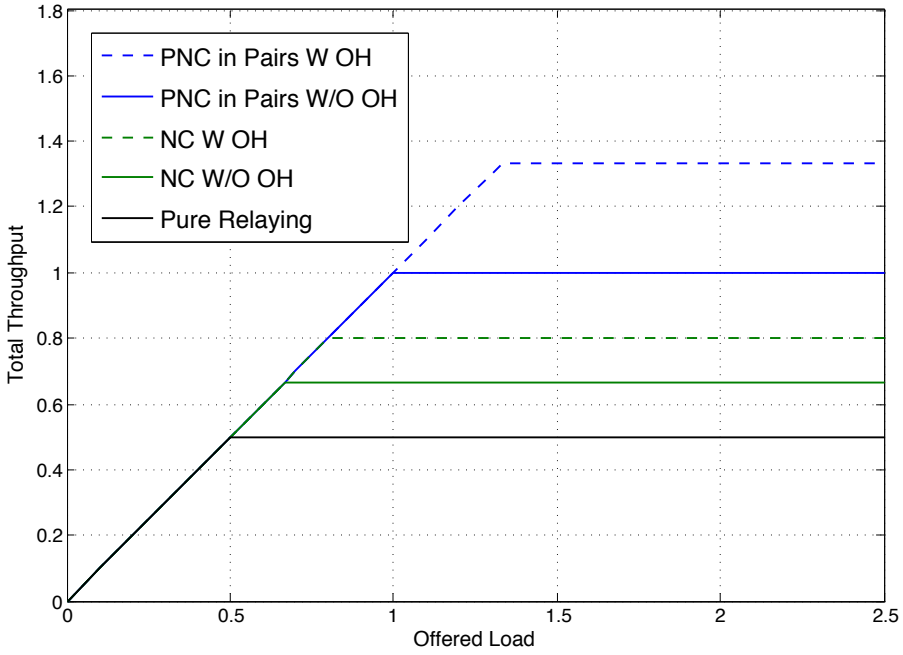


Figure 3.5: Performance of the different NC cases for optimized throughput.

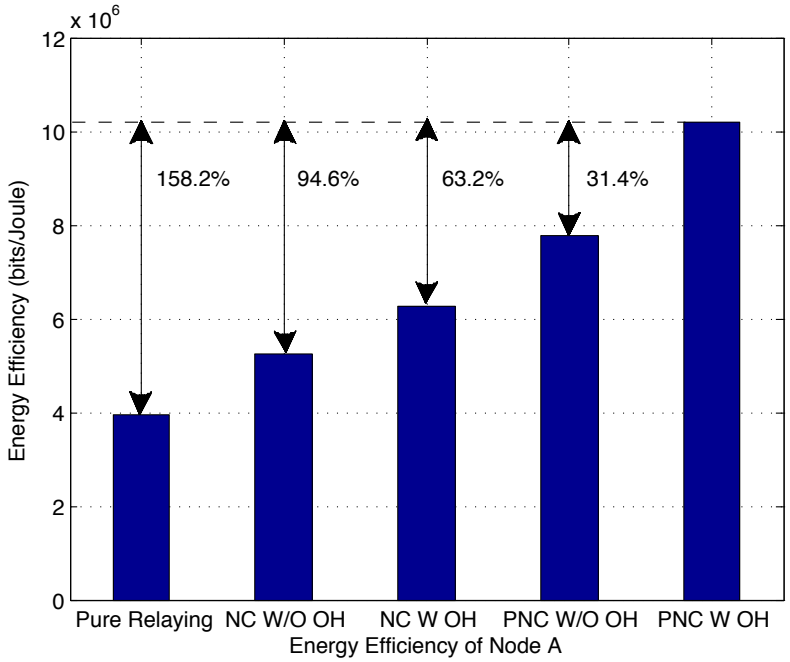


Figure 3.6: Energy efficiency of node A for optimized throughput for different NC cases.

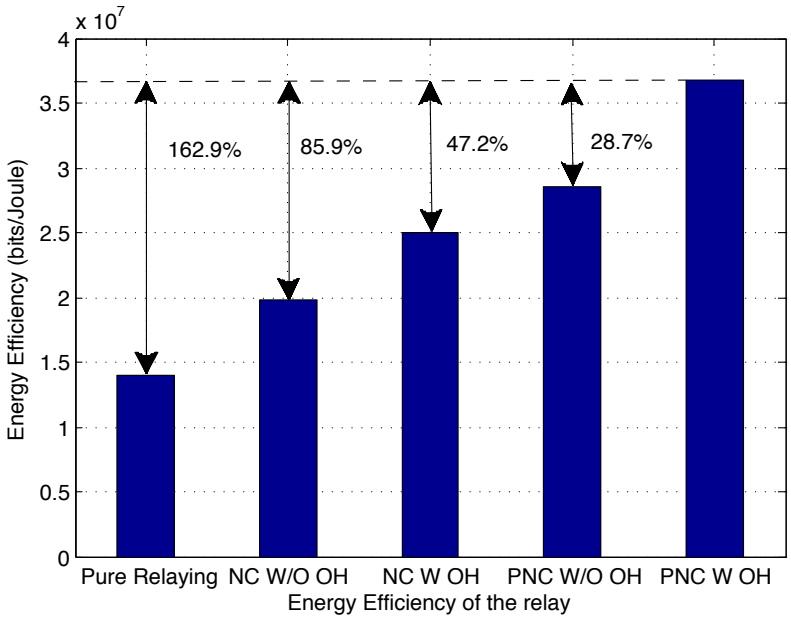


Figure 3.7: Energy efficiency of the relay for optimized throughput for different NC cases.

scheduling in the medium access.

In the case of NC with overhearing the throughput reaches its peak and stays stable as the offered load is increased. This happens because the conditions for the maximum throughput coincide with the equal sharing of the channel among the entities of the network. In the case of PNC in pairs with and without overhearing, a new MAC protocol that includes scheduling in the medium access and synchronization of the transmitting nodes should be designed in order for the maximum available throughput to be achievable.

Chapter 4

Uplink Offloading with IP Flow Mobility

4.1 Introduction

The continuous increase of cellular data demand that is already witnessed, is the main driving force for cellular network operators towards the capital investments on upgrades of their cellular network infrastructures into 4G systems, as LTE. With the upgrade of their networks, cellular providers aim to be able to serve the requested traffic by their customers. Despite the upgrade of the cellular infrastructures, the pace of the increase of the data traffic demand [2] puts pressure on the cellular network providers, as traffic congestion is not avoided. These facts have led the research community to propose offloading techniques that will leverage the mitigation of the overload of the cellular network spectrum and the network's traffic congestion.

According to the work of Paul *et al.* [56] on the dynamics of cellular data networks, downloads dominate uploads with more than 75% of the traffic coming from download traffic. On the other hand, smartphone applications slowly change the users attitude, transforming them into content creators. Facebook, Twitter, Youtube and Instagram are some of the main applications that let users upload their content (videos, photos, audio, text and combinations of them) at the time of creation. This change of use habits is highly demanding in terms of energy consumption, as in LTE, uploading is nearly eight times more energy consuming compared to downloading according to the extensive measurements of [57]. In the same work it is experimentally measured that LTE consumes two times the energy of WiFi for uploading small files of size equal to 10 kB and 2.53 times the energy of WiFi for larger files of size equal to 10 MB. Considering the solution of offloading the uplink traffic of users that are in the range of WiFi Access Points (APs), the battery life of mobile users will be extended and at the same time the uplink load of an eNodeB will be mitigated. According to Cisco's mobile data traffic forecast [2], mobile

offload increases from 45% (1.2 exabytes/month) that was in 2014 to 54% (28.9 exabytes/month) by 2019. As operators will not likely be able to keep pace with the current pace of mobile data demand, they respond by rolling out WiFi APs to public areas to offload data traffic. WiFi is an appropriate solution, as WiFi APs are easier to deploy and they cost less than upgrading existing cellular infrastructure gear.

With the release-10 of 3GPP, a UE in LTE networks is able to concurrently maintain connections with the cellular network and a WiFi AP, in order to offload part of its traffic through WiFi access and upload the rest through LTE. The scheme that allows this connectivity is named IP Flow Mobility (IFOM) [58]. The other two offloading techniques are Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO). IP Flow Mobility is currently being standardized by 3GPP [59]. This technology allows an operator or a UE to shift an IP flow to a different radio access technology, without disrupting any ongoing communication. Consider a UE connected to a cellular base station having multiple simultaneous flows. For example, it maintains a voice call and a file upload, and it is moving into the range of a WiFi AP. The UE may shift the file upload on the WiFi network and when it moves out of the AP coverage it will make a seamless shift of the flow back to the cellular network. Another example is the division of a UE's data flow into two sub-flows and the service of each sub-flow by different radio access technologies, as proposed in [60].

A question that arises from the IFOM uplink offloading scheme is how the UEs will offload part of their data through WiFi with fairness, where their different upload data needs and their LTE connection with the eNodeB will be considered, and how the rest of the data will be uploaded through LTE. Although the access method in 802.11, DCF (Distributed Coordinated Function), uses the CSMA/CA protocol in conjunction with a binary exponential backoff algorithm to share radio resources in a fair way, it treats all users equally. This access scheme creates unfairness considering the different data needs of each UE and the different channel conditions of their connection with the eNodeB. In cases where different queue lengths are considered [61] or in multi-rate conditions [62], fair resource allocation is achieved by weighted proportional fairness. While the downlink of a WiFi AP can be adaptive, based on priority queuing of data, the uplink does not present the same flexibility. In uplink, all transmitting users are treated equally, following the binary exponential backoff algorithm of 802.11 DCF. Based on this fact we focus on providing an effective access scheme for uplink offloading through WiFi that will treat all UEs on a weighted proportionally fair way, which includes the UEs uplink data needs in conjunction with their LTE channel conditions. The main objective of this approach is to achieve energy efficiency and throughput improvement in the uplink offloading with IFOM.

The main challenge is to provide an efficient uplink offloading algorithm that takes into consideration the different uplink data volume needs of UEs that are associated with the same WiFi AP and eNodeB, and present different channel conditions regarding their LTE uplink. The main questions that are tackled throughout this chapter are the following: (i) How the different data needs of UEs under the

coverage of the same eNodeB and WiFi AP should be divided into two sub-flows per UE that will be concurrently routed through the available access technologies? (ii) How can we improve the WiFi uplink access to maximize the uplink offloading of the data volume needs of the UEs? and (iii) How can we provide an efficient resource allocation for the LTE uplink of the data volume needs of the UEs that are not offloaded through WiFi? In this Chapter we discuss on the limitations of IEEE 802.11 DCF uplink access and we propose an offloading method for IFOM that combines weighted proportional fairness in the WiFi access and price-based resource allocation in the LTE upload. UEs that have larger upload data needs or experience worse LTE connection are favoured in the WiFi offloading part. This is achieved by choosing appropriate weights for the proportional fairness. The LTE uplink rate allocation we propose is a two-stage pricing algorithm. In the first stage, the LTE operator decides the price p per unit of a UE's LTE uplink rate. In the second stage, the UEs decide the rate for which they intend to pay, based on the price and the spectrum efficiency that they experience. Data pricing has been recently adopted as a promising economics tool that provides effective solutions for resource allocation aiming to mitigate network congestion [63]. We follow two different pricing schemes. A linear pricing scheme, that was used in [64] and [65] and an exponential pricing scheme, that was used in [66]. The main contributions of this part are the following:

- To the best of our knowledge this is the first work that considers uplink offloading methods for WiFi and LTE networks that operate under the IFOM offloading technique.
- We propose a weighted Proportionally Fair Bandwidth (PFB) allocation algorithm for the WiFi, aiming to improve the uplink offloading in terms of offloaded data volume, energy efficiency and fairness. We include in the fairness criteria the different data needs of the UEs and their LTE uplink spectrum efficiency.
- For the rest of each UE's data we propose a price-based rate allocation for the LTE uplink, and we follow a linear and an exponential pricing scheme. Our major focus is to investigate the effect of different pricing schemes on the energy efficiency and throughput performance of UEs under IFOM uplink offloading.

We compare the PFB algorithm with 802.11 DCF and with a state of the art uplink access scheme [67] in terms of UEs' energy efficiency for both linear and exponential pricing of the LTE rate allocation. We investigate the conditions under which exponential pricing performs better than linear pricing and we reveal the effect of the UEs' data needs and spectrum efficiency on their energy efficiency and throughput performance. In addition, we evaluate the offloading capabilities of PFB and we show that a greater data volume is offloaded using our proposed algorithm.

The Chapter is organized as follows. In Section 4.2 we present the state-of-the-art on offloading algorithms using WiFi networks and in Section 4.3 we present the system model of our approach. In Section 4.4 we present analytically our weighted

proportionally fair WiFi access algorithm and in Section 4.5 the analysis of the two proposed LTE pricing schemes. In Section 4.6, the evaluation of our uplink offloading scheme with IFOM is presented. In Section 4.7 we present a system model, where a UE presents malicious operation and in Section 4.8 we present its impact on the energy efficiency of the truthful UEs and of its own. We finalize this Chapter in Section 4.9 where we present the performance evaluation in terms of throughput and energy efficiency under selfish misbehavior.

4.2 Offloading Techniques for Heterogeneous Networks

Hereunder we present the the state-of-the-art regarding already proposed offloading algorithms through WiFi networks, along with recently published works related to misbehavior of users in offloading scenarios.

The relatively low deployment costs of WiFi APs has led the providers and the research community to investigate offloading techniques for the cellular networks through WiFi. In [68], the authors indicate that WiFi already offloads in US about 65% of the total mobile data traffic and saves 55% of battery power. In the same work, the offloading capabilities of WiFi are investigated under trace-driven simulations, based on mobility habits of mobile users and useful insights are provided on temporal offloading. In [69], offloading through opportunistic communications is explored, where a user offloads to another peer user, which in its turn maintains a short range connection (e.g. WiFi or Bluetooth) or a cellular connection (e.g. EDGE or HSPA). The problem of device-to-device communications over locally formed *ad hoc* networks is also addressed in [70], in the context of the download process. According to their model, the base station transmits different parts of the content to selected mobile devices. Following, the mobile devices multicast the received data to each other. The combination of device-to-device communication with delay-tolerant traffic was proposed in RoCNet [71], where a user terminal under the coverage of a high traffic loaded BS, forwards its traffic through a WLAN or a Bluetooth connection to another user terminal, which will be physically moved under the coverage of another cellular BS with low traffic load to offload its peers uplink data. In [72], the authors have proposed Wiffler, which is an application that is used to predict WiFi connectivity aiming to leverage the exploitation of offloading opportunities. Through the conducted measurements in city-wide testbeds they found that cellular and WiFi availability are negatively correlated, a fact that expands the benefit of offloading through WiFi in terms of network coverage. The authors in [73] study the economics of mobile data offloading through third-party WiFi or femtocell APs and they propose a market-based offloading scenario, aiming to investigate the market outcome with game theory. In [74], the authors propose a framework named iDeal, that allows providers to use resources from third-party resource owners, by leasing capacity in cases of congestion through reverse auctions. Several third-party resource owners are considered, that compete to lease their resources to the cellular provider, leading to significant savings for the provider's side. An optimal delayed

WiFi offloading algorithm is proposed in [75]. The authors consider the case of file downloading by mobile users that move under the BreadCrumbs mobility model, proposed in [76], and they provide an optimal algorithm that minimizes the mobile user's communication cost. In [77], methods for session continuity are proposed during non-seamless WiFi offloading in LTE networks. The performance of these methods is analysed in terms of throughput and energy consumption. In [78], the authors analyse the behaviour of the network decision-making and reconfiguration process in terms of handled handover requests, aiming to achieve load balancing by guiding the relocation of mobile terminals to achieve offloading. An enhanced framework for Wi-Fi-based offloading is proposed in [79], where the author investigates how to raise QoE by evaluating the deadline assurance in offloading, while saving a significant amount of 3G resources. In [80], the authors consider a model where a subset of users under the coverage of a base station are at the same time under WiFi coverage, and they aim to maximize the per-user-throughput by selecting which WiFi connected users will offload, taking into consideration the induced collisions for non-saturated traffic. The energy efficiency under this offloading scheme is left open. The recent published works related to offloading are mainly focused on the downlink traffic offloading and do not consider the increasing tendency of uploading user created content [81]. In our work we raise awareness of the uplink traffic offloading and its impact on the energy efficiency of the modern mobile communication devices. Contrary to [80] we consider that all UEs under the concurrent coverage of an eNodeB and a WiFi AP are given the opportunity to offload following a weighted fair allocation.

In our work we present a resource allocation approach for uplink offloading with IP Flow Mobility that is based on weighted proportional fairness for the WiFi access and on linear and exponential pricing for the LTE access. The WiFi access is affected by the data volume needs declared by the UEs and their spectrum efficiency regarding their LTE channel conditions. Thus, a UE offloads part of its data through WiFi and the rest are routed through its LTE connection. Taking into consideration the shared nature of WiFi resources, the access algorithms for uplink offloading need to be fair. Namely, the UEs must be truthful when declaring their uplink data needs. In [82] the authors analyse the types of misbehavior in wireless networks and identify untrusted partners as a usual class of vulnerabilities. These types of attacks can be identified by means of reputation based detection as described in [83]. In [84] and [85] selfish detection mechanisms are proposed for WLAN and WiFi tethering respectively.

Table 4.1: List of notations and their physical meanings.

Symbols	PhysicalMeanings
UE_i	UEs under the coverage of the same eNodeB and WiFi AP, $i = (1, ..., N)$
K_i	Uplink data volume needs of UE_i
ΔT	Duration of an uplink offloading period
θ_i	Normalized spectrum efficiency of UE_i, $\theta_i \in [0, 1]$
P_i^{LTE}	Power level of UE_i's LTE interface during uplink transmission
R_i^{LTE}	LTE uplink rate of UE_i (in Mbps)
R_{max}^{LTE}	Maximum value of LTE uplink rate (in Mbps)
α_u	Uplink transmission power per Mbps
β	Base power of the LTE card
$S(N)$	Per UE WiFi uplink throughput (in Mbps), for N contending UEs
$EE(N)$	Per UE WiFi energy efficiency (in bits/Joule), for N contending UEs
R^{WiFi}	WiFi data transmission rate
w_i	Offloading factor of UE_i, $w_i \in [0, 1]$
EC_{Tx}^{WiFi}	Average per UE energy consumption during WiFi transmission phase (Joule)
EC_{sleep}^{WiFi}	Average per UE energy consumption during WiFi sleep phase (Joule)
EC^{LTE}	Average per UE energy consumption during LTE transmission (Joule)
E_{eff}^{PFB}	Average per UE energy efficiency of IFOM offloading under PFB (bits/Joule)
Off_{PFB}	Offloading index of the PFB algorith
R_i^{LTE}	LTE uplink transmission rate of UE_i
p	Price per unit of LTE uplink transmit rate for linear pricing
p_e	Price per unit of LTE uplink transmit rate for exponential pricing
U_i^{lin}	Payoff function of UE_i, for linear pricing
U_i^{exp}	Payoff function of UE_i, for exponential pricing

4.3 System Model for Uplink Offloading

The main notations used in this Chapter are summarized and explained in Table 4.1 for the ease of reading. We consider a LTE macro-cell and we focus on its coverage area that is also partially covered by several WiFi APs that belong to the same LTE provider, as shown in Fig. 4.1 and published in our work in [86]. All UEs are equipped with a WiFi network interface card in addition to their LTE connectivity. We assume that N LTE UEs are concurrently under the coverage of the macro-cell and one of the deployed APs, and they need to upload a file (e.g. a photo or a video) through a mobile application. The used applications are assumed to be able to divide an IP flow into two sub-flows and to define the size of each one. The UEs are able to use concurrently the two access technologies with IFOM and direct one sub-flow to LTE and the other to WiFi. The data volume UE_i needs to upload is equal to K_i bits, where $i = (1, ..., N)$. The data needs K_i are assumed a priori known to the WiFi AP. The AP has a high bandwidth backbone (e.g. fiber connection). Thus, the bottleneck of this route lies in the wireless uplink access of the WiFi connection.

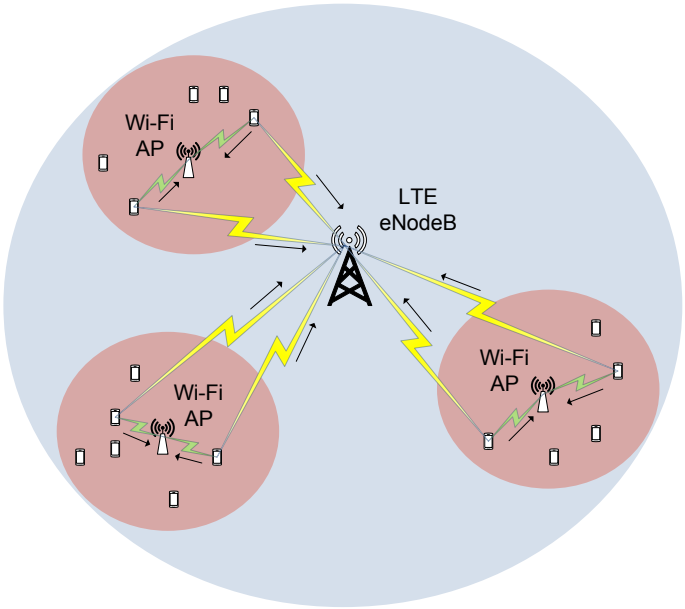


Figure 4.1: Uplink offloading scenario with IP Flow Mobility (IFOM).

The described scheme is applied to each one of the WiFi APs and we investigate the uplink data offloading for a time horizon equal to ΔT. Each UE_i offloads part of each data needs K_i. The rest is uploaded through its LTE connection. We assume that the channel characteristics between each UE_i and the LTE macro-cell are described by a normalized spectrum efficiency $\theta_i \in [0, 1]$, such that for a bandwidth allocation that gives to UE_i the ability to upload with an uplink rate equal to R_i^{LTE} under ideal channel conditions, the actual achieved uplink rate is equal to $\theta_i R_i^{LTE}$. As we focus on the access layer of the heterogeneous network, we assume that θ_i abstracts the physical layer characteristics including the frequency selectivity that the UEs may experience due to transmitting in different frequencies even with the same bandwidth. With this abstraction we provide a plug-in parameter to our access layer study, available to be used over a physical layer analysis. In Fig. 4.2 we provide a schematic representation of our proposed model, where we present which decisions are centralised at the operators side, and which are distributed at the UEs side. The operator decides how the weighted fair bandwidth allocation will be held at the WiFi offloading. The WiFi AP uses an exclusive access scheme to implement the bandwidth allocation decided by the operator. Regarding the LTE rate allocation, the operator decides centrally the price p per unit of LTE uplink rate and following, the UEs distributively decide the LTE rate to request, based on the already decided price p and their spectrum efficiency θ_i. After these decisions, each UE_i transmits its data by concurrently using both access technologies.

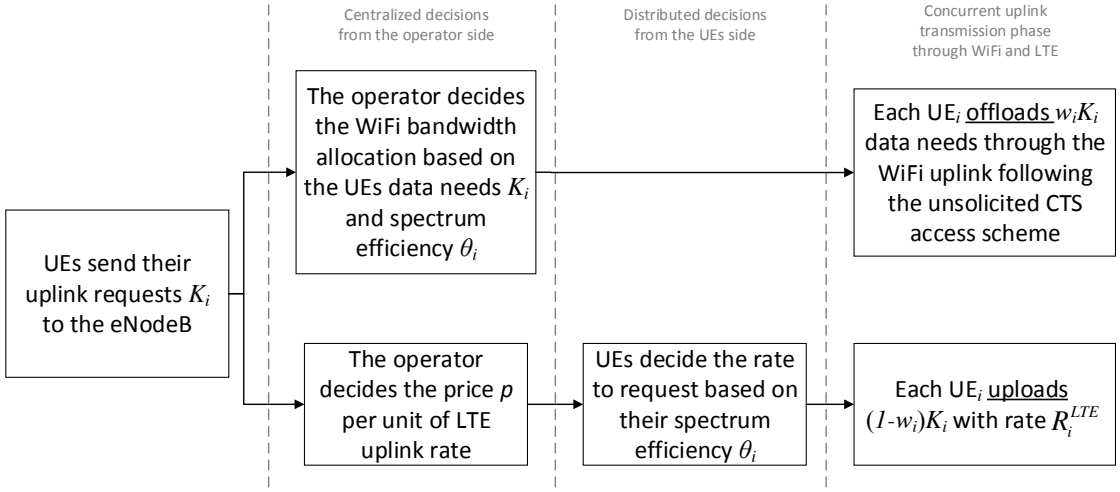


Figure 4.2: Schematic representation of the system model.

4.3.1 LTE Uplink Power Model

Regarding the LTE uplink power level of the UE, we adopt the energy model proposed by Huang *et al.* in [57]. According to this model the power level of the UE$_i$'s LTE interface during uplink transmission is expressed as

$$P_i^{LTE} = \alpha_u R_i^{LTE} + \beta \text{ [mW]} \tag{4.1}$$

where α_u is the uplink transmission power per Mbps, R_i^{LTE} is the LTE uplink rate (in Mbps) and β is the base power of the LTE card.

4.3.2 IEEE 802.11 DCF Energy Consumption in the Uplink

The uplink access mechanism of IEEE 802.11 DCF [9] is based on contention among users that are willing to transmit data to the AP and try to avoid collisions following the standard's binary exponential back-off algorithm. Following Bianchi's analysis [87] for saturated traffic conditions we notice that the throughput of a user that tries to upload data through WiFi is significantly affected by the number of users that are under the coverage of the same AP. The per user uplink throughput $S(N)$ (in Mbps), where N is the number of contending users, is expressed as

$$S(N) = \frac{P_s(N)P_{tr}(N)E[P]}{N\left[(1-P_{tr}(N))\sigma + P_{tr}(N)P_s(N)T_s + P_{tr}(N)(1-P_s(N)T_c)\right]} \tag{4.2}$$

$E[P]$, T_s, T_c and σ correspond to the average payload of a packet, the duration of a successful transmission, the duration of a collision and the time slot's duration respectively. $P_{tr}(N)$ is the probability that there is at least one transmission in a considered time slot and $P_s(N)$ is the probability that an occurring transmission

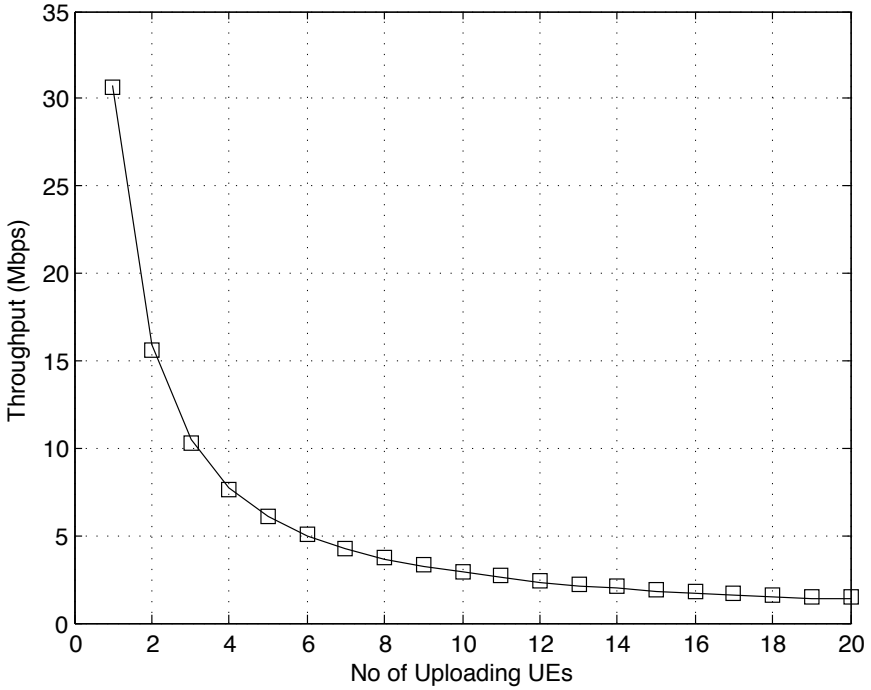


Figure 4.3: Degradation of per user throughput in IEEE 802.11 DCF.

is successful. In Fig. 4.3 the degradation of a user's throughput is presented for the cases of one to 20 users associated with the same AP under saturated traffic conditions and transmission rate equal to $R^{ul}_{WiFi} = 54$ Mbps. A user's energy efficiency $EE(N)$ (in bits/Joule), as a function of the number of contending users N is expressed as

$$EE(N) = \frac{P_s(N)P_{tr}(N)E[P]}{N\left[(1-P_{tr}(N))E_i + P_{tr}(N)P_s(N)E_s + P_{tr}(N)(1-P_s(N)E_c)\right]} \quad (4.3)$$

where E_i, E_s and E_c correspond to the energy consumption of a user during an idle, a successful transmission and a collision period. The duration of a successful transmission is equal to $T_s = T_H + T_P + T_{SIFS} + T_{ACK} + T_{DIFS}$. The duration of a collision period is equal to $T_c = T_H + T_P + T_{DIFS}$, and the duration of an idle period is equal to a time slot σ. Where T_H is the transmission duration of the PHY and MAC headers and T_P the transmission duration of a packet's payload for transmission rate equal to $R^{WiFi} = 54$ Mbps. Taking these duration expressions into consideration we analytically express the energy consumption values of (4.3) in (4.4).

$$\begin{aligned} E_s &= P_{Tx}(T_H + T_P) + P_{idle}(T_{SIFS} + T_{DIFS}) + P_{Rx}T_{ACK} \\ E_c &= P_{Tx}(T_H + T_P) + P_{idle}T_{DIFS} \\ E_i &= \sigma P_{idle} \end{aligned} \quad (4.4)$$

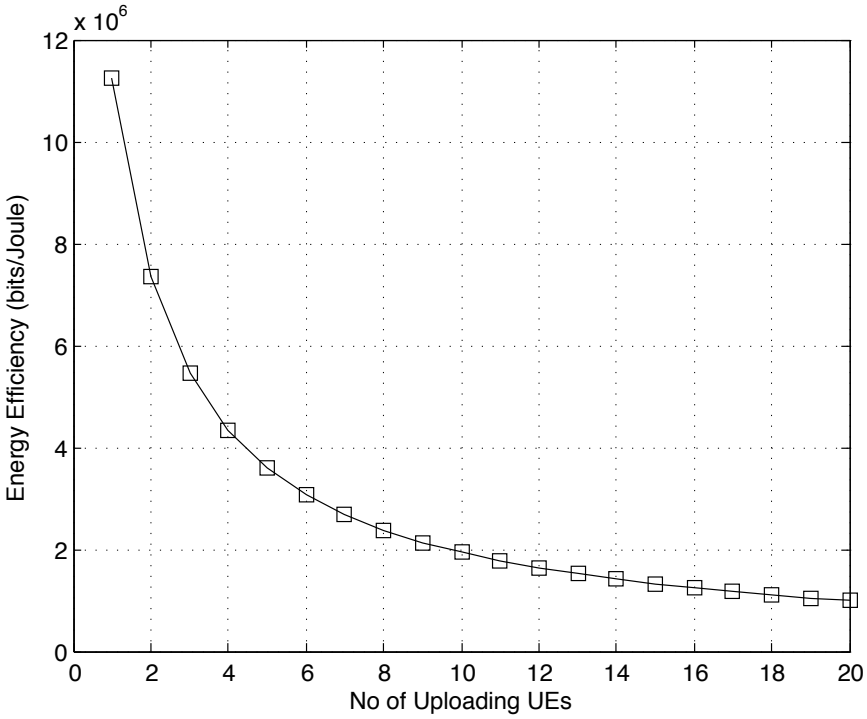


Figure 4.4: Degradation of per user energy efficiency in IEEE 802.11 DCF.

where P_{idle}, P_{Tx} and P_{Rx} are the power levels of the user's 802.11 network interface card. In Fig. 4.4 the degradation of a user's energy efficiency is presented for different number of uploading users.

4.3.3 Uplink Offloading Energy Consumption

Every UE under the concurrent coverage of the two access technologies will have the opportunity to offload $w_i K_i$ bits through the WiFi AP, where $w_i \in [0, 1]$ for $i = (1, ..., N)$. The remainder data volume $(1 - w_i)K_i$ is transmitted through the LTE connection of each UE. Every UE_i with data needs equal to K_i that offloads its uplink according to w_i will present energy consumption $EC_i(N)$, which is expressed as

$$EC_i(N) = (1 - w_i)K_i \frac{P_i^{LTE}}{\theta_i R_i^{LTE}} + w_i K_i \frac{1}{EE(N)} \text{ [Joule]} \tag{4.5}$$

For the WiFi uplink offloading we provide a weighted proportionally fair allocation algorithm over the data needs and the LTE channel conditions of the UEs. For the LTE uplink of the $(1 - w_i)K_i$ data volume of each UE_i we provide a two stage pricing algorithm for the LTE uplink rate allocation. Based on these two parts of the IFOM uplink offloading we are able to calculate each UE_i's energy consumption according to (4.5). Considering that without offloading a UE_i would upload with throughput equal to $\theta_i R_i^{LTE}$, we express the equivalent throughput of UE_i with offloading as $\theta_i R_i^{LTE}/(1 - w_i)$, assuming that the LTE uploading continues after

the WiFi offloading.

4.4 Weighted Proportionally Fair WiFi Access

The UEs offload part of their data needs through the WiFi, according to the Proportionally Fair Bandwidth (PFB) allocation algorithm that we propose [88]. Each UE_i is allocated bandwidth equal to r_i, $i = (1, ..., N)$, such as $\sum_{i=1}^{N} r_i \leq R_i^{WiFi}$. The allocation is proportionally fair over the ratio $\rho_i = K_i/\theta_i$. According to the definition of proportional fairness by Kelly *et al.* [89], a vector of rate allocation $\mathbf{r} = (r_1, ..., r_N)$ is proportionally fair if it is feasible, that is $\mathbf{r} \geq 0$ and $\sum_{i=1}^{N} r_i \leq R_i^{WiFi}$ and if for any other feasible vector $\mathbf{r}^*$, regarding the proportional fairness over the ratio ρ_i of each UE_i, the aggregate of proportional changes is zero or negative and is expressed as

$$\sum_{i=1}^{N} \rho_i \frac{r_i^* - r_i}{r_i} \leq 0 \tag{4.6}$$

which can be rewritten as

$$\sum_{i=1}^{N} \rho_i (log(r_i))' dr_i \leq 0 \tag{4.7}$$

It follows from (4.7) that the proportionally fair allocation solution represents a maximum of the utility function $U_i(\mathbf{r}) = \sum_{i=1}^{N} \rho_i log(r_i)$. Consequently, in order to find the proportionally fair solution we have to solve the maximization problem described as follows

$$\begin{aligned} \max_{\mathbf{r}} \quad & \sum_{i=1}^{N} \rho_i log(r_i) \\ \text{subject to} \quad & \sum_{i=1}^{N} r_i \leq R^{WiFi} \\ \text{and} \quad & r_i \geq 0, \forall i = 1, ...N \end{aligned} \tag{4.8}$$

The problem has a unique solution since the objective function is strictly concave and the constraint set is convex. To solve this problem, we relax the constraints and define the Lagrangian [90], changing $r_i \geq 0$ to $-r_i \leq 0$

$$L(\mathbf{r}, \mu) = \sum_{i=1}^{N} \rho_i (log(r_i)) - \mu_0 \left(\sum_{i=1}^{N} r_i - R^{WiFi} \right) + \sum_{i=1}^{N} \mu_i r_i \tag{4.9}$$

where $\mu_0 \geq 0$ and $\mu_i \geq 0$, $i = 1, ..., N$. Following we take the Karush-Kuhn-Tucker (KKT) optimality conditions. Starting with the stationarity condition we have

$$\nabla_{r_i} L(\mathbf{r}, \mu) = \frac{\rho_i}{r_i} - \mu_0 + \mu_i = 0 \tag{4.10}$$

since $\rho_i > 0$, then $\mu_0 > \mu_i$, which also means $\mu_0 > 0$. From the complementary slackness conditions we have

$$\mu_0 \left(R^{WiFi} - \sum_{i=1}^{N} r_i \right) = 0 \tag{4.11}$$

$$\mu_i r_i = 0 \tag{4.12}$$

$$\mu_0 \geq 0 \text{ and } \mu_i \geq 0,\ i = 1, ..., N \tag{4.13}$$

and since $\mu_0 > 0$, we know that

$$\sum_{i=1}^{N} r_i = R^{WiFi} \tag{4.14}$$

which means that r_i, $i = 1, ..., N$ cannot be zero. Thus, forcing $\mu_i = 0$, $\forall i = 1, ...N$ we have from (4.10)

$$r_i = \frac{\rho_i}{\mu_0} \tag{4.15}$$

Combining (4.14) and (4.15) we have the optimal solution which represents the weighted proportionally fair solution

$$r_i = \frac{\rho_i}{\sum_{i=1}^{N} \rho_i} R^{WiFi} \tag{4.16}$$

Following, we provide an access method that allocates exclusive access to each UE following the PFB algorithm.

4.4.1 Implementation Consideration

In the PFB algorithm we aim to allocate exclusive access periods to each UE_i equal to t_i, for $i = (1, ..., N)$. In these periods the UEs will be able to transmit through the WiFi AP with throughput $R^{WiFi} = S(1)$. Namely, the transmitting UE will face no contention with other peers. We transform the proportionally fair bandwidth allocation into proportionally fair airtime allocation by having $r_i \Delta T = t_i S(1)$. Now, the weighted proportionally fair airtime allocation is equal to

$$t_i = \frac{\rho_i}{\sum_{i=1}^{N} \rho_i} \Delta T \tag{4.17}$$

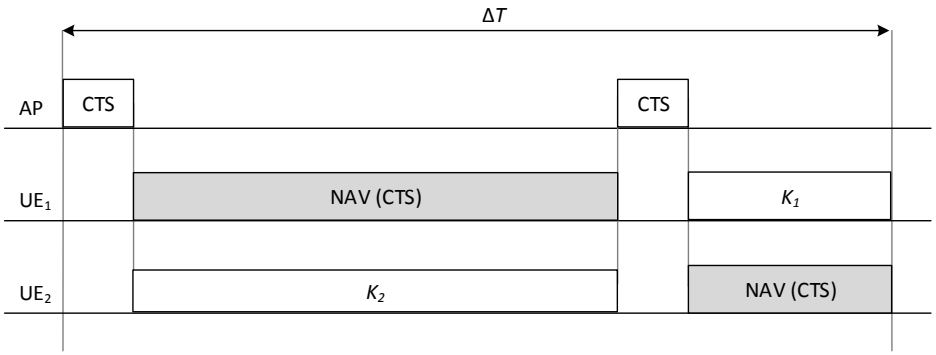


Figure 4.5: An example of the PFB algorithm for two UEs.

Regarding the implementation of the PFB algorithm we aim to give exclusive access to the WiFi AP to each UE_i for a period equal to t_i. To achieve that, we adopt the idea of unsolicited Clear To Send (CTS) frames initiated by the AP that was proposed in [91]. With a CTS frame the AP protects a specific UE to upload its data through WiFi, while all other UEs put their 802.11 network interface cards into sleep mode for a duration equal to the NAV information of the CTS. A timeline example for the WiFi access of the PFB algorithm for two UEs is presented in Fig. 4.5. We notice that due to non optimally scheduled user access, UE_2 is obliged to wait for a long period in comparison to its own access time. Even though during this waiting period UE_2's WiFi card is in sleep mode, it consumes energy. We can further improve our algorithm by applying the optimal scheduling for one machine and non-preemptive jobs which is a shortest-job-first fashion approach.

4.4.2 Energy Efficiency of PFB

The average per UE energy consumption of the WiFi network interface card, during the uploading phase, is expressed as

$$EC_{Tx}^{WiFi} = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{\rho_i}{\sum_{i=1}^{N}\rho_i}\Delta T\frac{S(1)}{EE(1)}\right) \text{ [Joule]} \tag{4.18}$$

After scheduling the exclusive time periods t_i in augmenting order of duration, the average per UE energy consumption of the WiFi network interface card while in sleep mode with power level P_{sleep}^{WiFi}, is expressed as

$$EC_{sleep}^{WiFi} = \frac{1}{N}\sum_{i=1}^{N-1}(N-i)t_i P_{sleep}^{WiFi} \text{ [Joule]} \tag{4.19}$$

The average per UE energy consumption of the LTE network interface card, due to the concurrent transmission through WiFi, is equal to

$$EC^{LTE} = \frac{1}{N}\sum_{i=1}^{N}\left((K_i - t_i S(1))\frac{P_i^{LTE}}{\theta_i R_i^{LTE}}\right) \text{ [Joule]} \tag{4.20}$$

Combining (4.18)–(4.20) the average per UE energy efficiency of IFOM offloading under the PFB algorithm is expressed in (4.21).

$$E_{eff}^{PFB} = \frac{\sum_{i=1}^{N} K_i}{N(EC_{Tx}^{WiFi} + EC_{sleep}^{WiFi} + EC^{LTE})} \text{ [bits/Joule]} \tag{4.21}$$

4.4.3 Offloading Index

Aiming to reveal the performance improvement of the PFB algorithm in terms of data volume offloading, we define the WiFi offloading index Off_{PFB}. The Off_{PFB} is expressed in (4.22) and is equal to the ratio of the total offloaded data volume through the WiFi following the PFB algorithm to the data volume that would be uploaded by the standard 802.11 DCF if only one user was accessing the AP to offload. The WiFi offloading index of the PFB algorithm, Off_{PFB} is equal to

$$Off_{PFB} = \frac{\sum_{i=1}^{N} t_i S(1)}{S(1)\Delta T} \tag{4.22}$$

It follows from (4.22) that $Off_{PFB} = 1$, which means that the PFB algorithm fully exploits the offloading capabilities of the WiFi AP, as every UE is allocated exclusive offloading access to the AP.

4.5 LTE Pricing Scheme

The LTE uplink power of each UE_i, following the power model of (4.1), is a function of its LTE uplink transmission rate, R_i^{LTE}. Hereunder, we propose a two-stage LTE pricing scheme, where the LTE operator decides the price p per unit of uplink transmit rate in the first step and in the second step the UEs decide the rate for which they intend to pay as a function of the price and the spectrum efficiency they experience. We approach the pricing problem using backward induction, examining first the UEs demands (Stage II) and then the operator's decision on the price (Stage I). We propose two pricing models, one linear and one exponential.

4.5.1 LTE Uplink Rate With Linear Pricing

Stage II: The payoff function of the UE_i, for acquiring R_i^{LTE} quantity of uplink rate with a price p per unit of rate, following the linear pricing model, is expressed as

$$U_i^{lin}(R_i^{LTE}) = ln(1 + \theta_i R_i^{LTE}) - pR_i^{LTE} \tag{4.23}$$

This payoff function of a UE_i, with normalized spectrum efficiency θ_i, is equal to the logarithmic utility function, that expresses the diminishing return of getting additional resources, minus the linear price that the UE_i has to pay for acquiring R_i^{LTE} quantity of rate. We notice that $U_i^{lin}(R_i^{LTE})$ is a concave function, since $U_i^{lin}(R_i^{LTE})'' = -\left(\theta_i/(1 + \theta_i R_i^{LTE})\right)^2 < 0$. Thus, it has only one maximum, and therefore the local maximum is also the global maximum. Differentiating (4.23) we have

$$\frac{\partial U_i^{lin}}{\partial R_i^{LTE}} = \frac{\theta_i}{1 + \theta_i R_i^{LTE}} - p = 0 \tag{4.24}$$

The optimal value of rate that maximizes UE_i's payoff is

$$R_i^{*LTE} = \begin{cases} \frac{1}{p} - \frac{1}{\theta_i}, & \text{if } p \leq \theta_i \\ 0, & \text{otherwise} \end{cases} \tag{4.25}$$

Stage I: We take into consideration that the N UEs that are under the coverage of the same WiFi AP are close enough to present similar channel statistics regarding their LTE connection. Thus, we assume that their spectrum efficiency is such that $\max(\theta_i) - \min(\theta_i) < \varepsilon$, where $\varepsilon > 0$. Under this assumption, the operator's choice of price p is such, that the UE with the $\max(\theta_i)$ is allocated the maximum value of the LTE uplink rate $R_{\max}^{LTE}$, aiming to provide the best available service to UEs with better channel conditions compared to the rest of the UEs situated under the same WiFi AP coverage. We also assume that the eNodeB has adequate available resources to satisfy the requests of all UEs. The price is formed according to (4.26).

$$p = \frac{\max(\theta_i)}{1 + \max(\theta_i) R_{\max}^{LTE}} \tag{4.26}$$

The provider aims to give to every UE_i the opportunity to transmit through the LTE. This means that even for the UE with the $\min(\theta_i)$, the quantity $1/p - 1/\min(\theta_i)$ is positive. Using (4.26) we find the range of values of ε under which this rate allocation is feasible. This range is expressed as

$$0 < \varepsilon \leq \max(\theta_i) \min(\theta_i) R_{\max}^{LTE} \tag{4.27}$$

The allocated rate to each UE_i following the linear pricing model is expressed as

$$R_i^{LTE} = \frac{1 + \max(\theta_i) R_{\max}^{LTE}}{\max(\theta_i)} - \frac{1}{\theta_i} \tag{4.28}$$

4.5.2 LTE Uplink Rate With Exponential Pricing

For the rate allocation with the exponential pricing model, where the price per unit of uplink transmit rate is denoted by p_e, we follow the same steps as described in the linear pricing approach.

Stage II: The payoff function of UE_i, for acquiring R_i^{LTE} quantity of uplink rate when applying the exponential pricing model is expressed as

$$U_i^{exp}(R_i^{LTE}) = ln(1 + \theta_i R_i^{LTE}) - p_e(e^{R_i^{LTE}} - 1) \tag{4.29}$$

This payoff function of a UE_i under exponential pricing, with normalized spectrum efficiency θ_i, is equal to the logarithmic utility function, that expresses the diminishing return of getting additional resources, minus the exponential price that the UE_i has to pay for acquiring R_i^{LTE} quantity of rate. We notice that $U_i^{exp}(R_i^{LTE})$ is a concave function, since $U_i^{exp}(R_i^{LTE})'' = -\left(\theta_i/(1 + \theta_i R_i^{LTE})\right)^2 - p_e e^{R_i^{LTE}} < 0$. Thus, it has only one maximum, and therefore the local maximum is also the global maximum. Differentiating (4.29) we have

$$\frac{\partial U_i^{exp}}{\partial R_i^{LTE}} = \frac{\theta_i}{1 + \theta_i R_i^{LTE}} - p_e e^{R_i^{LTE}} = 0 \tag{4.30}$$

We need to solve this non-linear equation with respect to R_i^{LTE}. (4.30) can be rewritten as

$$ln\left(\frac{1}{p_e}\right) + \frac{1}{\theta_i} = \left(R_i^{LTE} + \frac{1}{\theta_i}\right) + ln\left(R_i^{LTE} + \frac{1}{\theta_i}\right) \tag{4.31}$$

For $x = R_i^{LTE} + \frac{1}{\theta_i}$ and $y = ln\left(\frac{1}{p_e}\right) + \frac{1}{\theta_i}$, (4.31) can be written as

$$y = x + lnx \tag{4.32}$$

which after some straight forward mathematical manipulations can be written as

$$xe^x = e^y \tag{4.33}$$

Taking the value of the Lambert W function [92] of each part of (4.33) and using the Lambert W function identity $W(xe^x) = x$ we have $x = W(e^y)$. Replacing x and y we have

$$R_i^{LTE} = W\left(\frac{e^{\frac{1}{\theta_i}}}{p_e}\right) - \frac{1}{\theta_i} \tag{4.34}$$

Stage I: The price p_e that the provider decides in the exponential pricing model is such, that the UE with the $\max(\theta_i)$ is allocated the maximum value of the LTE uplink rate $R_{\max}^{LTE}$. The price is formed according to (4.35).

$$p_e = \frac{\max(\theta_i)}{(1 + \max(\theta_i) R_{\max}^{LTE})\, e^{R_{\max}^{LTE}}} \tag{4.35}$$

As the provider aims to give to all N UEs under the coverage of the AP the opportunity to upload part of their data needs through the eNodeB, the rate that will be allocated to the user with the $\min(\theta_i)$ should be positive. This means that the range of the spectrum efficiency of the N UEs is such, that

$$W\left(\frac{e^{\frac{1}{\min(\theta_i)}}}{p_e}\right) - \frac{1}{\min(\theta_i)} > 0 \tag{4.36}$$

The allocated rate to each UE_i following the exponential pricing model is expressed as

$$R_i^{LTE} = W\left(\frac{(1+\max(\theta_i)R_{\max}^{LTE})e^{R_{\max}^{LTE}+\frac{1}{\theta_i}}}{\max(\theta_i)}\right) - \frac{1}{\theta_i} \tag{4.37}$$

4.6 Evaluation of Uplink Offloading with IP Flow Mobility

We evaluate our offloading schemes by running extensive simulations using MATLAB™. We run the PFB algorithm for a diverse number of UEs under the concurrent coverage of an eNodeB and a WiFi AP, namely for four to 20 UEs. During the examined offloading periods, each UE_i under the concurrent coverage of the same eNodeB and WiFi AP has K_i data volume needs. Hence, every UE has frames to transmit in its buffer and the traffic is considered under saturation conditions. By following this assumption, we provide results for the worst case traffic scenario. We compare the performance of PFB in terms of energy efficiency and offloading capabilities with the standard 802.11 DCF and with an access mechanism titled Smart Exponential-Threshold-Linear (SETL) that was proposed in [67] and presents better uplink throughput compared to IEEE 802.11. In the backoff algorithm proposed in SETL the Contention Window (CW) of a 802.11 user is increasing exponentially up to a threshold that is equal to $CW_{th} = (CW_{max}/2 + CW_{min})$. After this threshold, it is increasing linearly up to CW_{max} according to $CW_{th} + kCW_{min}$, where k is a positive integer. Regarding the LTE part of IFOM, we conduct the simulations applying the presented pricing models, i.e. the linear and the exponential pricing approaches. In more detail, the WiFi AP is aware of the upload data needs of each UE_i. Based on these data needs and the LTE spectrum efficiency of each UE_i, it allocates access time for offloading. The remaining data of each UE_i are uploaded in parallel through its LTE connection with an uplink rate equal to $\theta_i R_i^{LTE}$, where R_i^{LTE} is defined by the applied pricing scheme.

The simulations are repetitively conducted for an offloading time period equal to $\Delta T = 5$ sec. The data volume needs of the UEs are assumed to follow a uniform distribution of the file sizes between $5 - 15$ MB. These data needs represent the volume of a photo to a small video, created by contemporary smartphones. The simulations are performed for different value ranges of the spectrum efficiency θ_i, aiming to explore the performance of our offloading scheme for offloading regions situated in different distances from the eNodeB. The uplink power level of UE_i's

LTE interface card, P_i^{LTE}, is assumed to follow (4.1). The 802.11 network interface card power levels P_{Tx}^{WiFi}, P_{Rx}^{WiFi}, P_{idle}^{WiFi} and P_{sleep}^{WiFi} are assumed to follow the measurements provided in [53]. The numerical values of the simulation parameters are presented in Table 4.2.

4.6.1 Energy Efficiency with Linear Pricing

In Fig. 4.6(a) we present the energy efficiency of PFB, 802.11 and SETL, for different number of UEs ranging from four to 20, with $\theta_i \in [0.8, 1]$. We notice that as the number of UEs increases the average energy efficiency of a UE decreases because the WiFi bandwidth is shared between more UEs and consequently they have to upload a larger part of their data needs through the energy demanding LTE uplink. The LTE uplink rate allocation follows the linear pricing model. In Fig. 4.6(b) the energy efficiency gain (%) is presented, comparing PFB to the 802.11 standard and the SETL algorithm. For $\theta_i \in [0.6, 0.8]$ we present the energy efficiency of PFB, 802.11 and SETL in Fig. 4.7(a), while in Fig. 4.7(b) the energy efficiency gain (%) is presented, comparing PFB to the 802.11 standard and the SETL algorithm.

4.6.2 Energy Efficiency with Exponential Pricing

Applying the exponential pricing for the LTE rate allocation we achieve a slight improvement in the energy efficiency of the IFOM offloading scheme. While UEs with lower θ_i are facilitated to acquire more resources, UEs with higher θ_i are pushed to acquire less than they would do by following the linear pricing. In Fig. 4.8(a) we present the energy efficiency results following the exponential pricing for

Table 4.2: Simulation Parameters

Parameter	Value
LTE max uplink rate $R_{\max}^{LTE}$	5 Mbps
LTE uplink power per Mbps α_u	438.39 mW/Mbps
LTE base power β	1288.04 mW
LTE uplink power	$\alpha_u R_i^{LTE} + \beta$ mW
WiFi packet payload	1500 bytes
WiFi Data/ Ctrl. transmission rate	54/ 6 Mbps
WiFi Tx/ Rx/ Idle/ Sleep Power	1900/ 1340/ 1340/ 75 mW
SIFS/ DIFS	10/ 50 μsec
Offloading period ΔT	5 sec
Number of UEs	8-20
Uplink data volume per UE$_i$, (K_i)	5-15 MB

$\theta_i \in [0.8, 1]$, while in Fig. 4.8(b) the energy efficiency gain is presented for the same range of θ_i. In Fig. 4.9(a) and 4.9(b) we present results for $\theta_i \in [0.6, 0.8]$.

Comparing the energy efficiency of uplink offloading between linear and exponential pricing schemes, we notice that for $\theta_i \in [0.8, 1]$ and $\theta_i \in [0.6, 0.8]$ the performance is similar. Though, for larger range of θ_i, e.g. for $\theta_i \in [0.2, 1]$, and for large number of UEs under the coverage of the same AP, namely from 10 to 20, the exponential pricing scheme presents 20% better energy efficiency compared to the linear pricing scheme. Thus, for greater diversity in the LTE channel quality of the UEs that share the resources of the same AP, the exponential pricing scheme performs better because it gives to UEs that experience worse LTE channel quality the opportunity to purchase more LTE resources, while it pushes UEs with better LTE channel conditions to purchase less resources compared to the linear pricing scheme. Simulation results regarding this comparison are presented in Fig. 4.10.

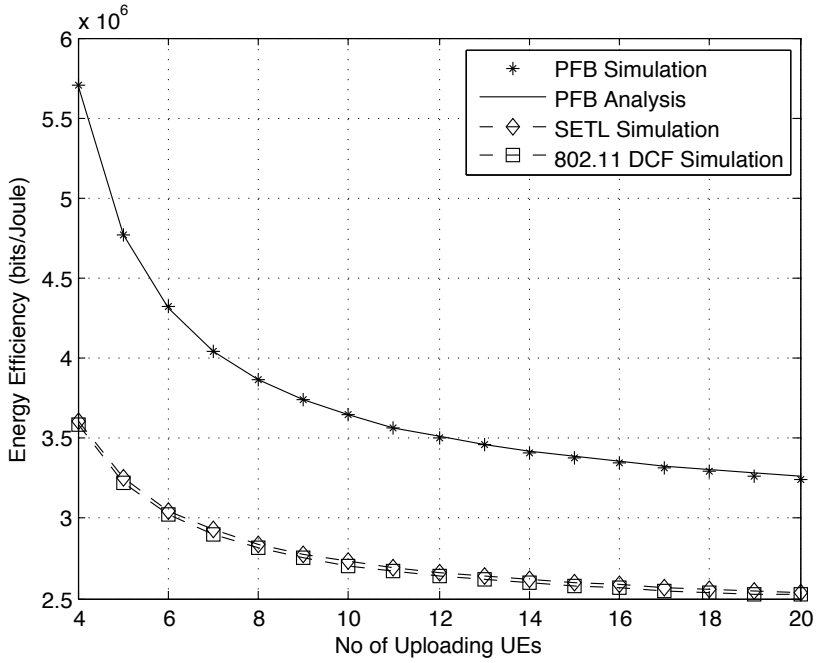
4.6.3 Evaluation of Offloading Capabilities

A comparison of the offloading capability of each WiFi uplink approach is presented in Fig. 4.11. As expected from (4.22), PFB presents an offloading index equal to one. This means that with PFB we achieve the maximum exploitation of the AP's capability for offloading. With SETL we achieve an offloading index near 0.94 for high contention conditions (20 offloading UEs) and with 802.11 the offloading index has a value near 0.86 under the same high contention conditions. While PFB provides exclusive WiFi access to the UEs, SETL and 802.11 operate under the CSMA/CA protocol leading to frequent collisions, especially when the number of uploading UEs is increasing.

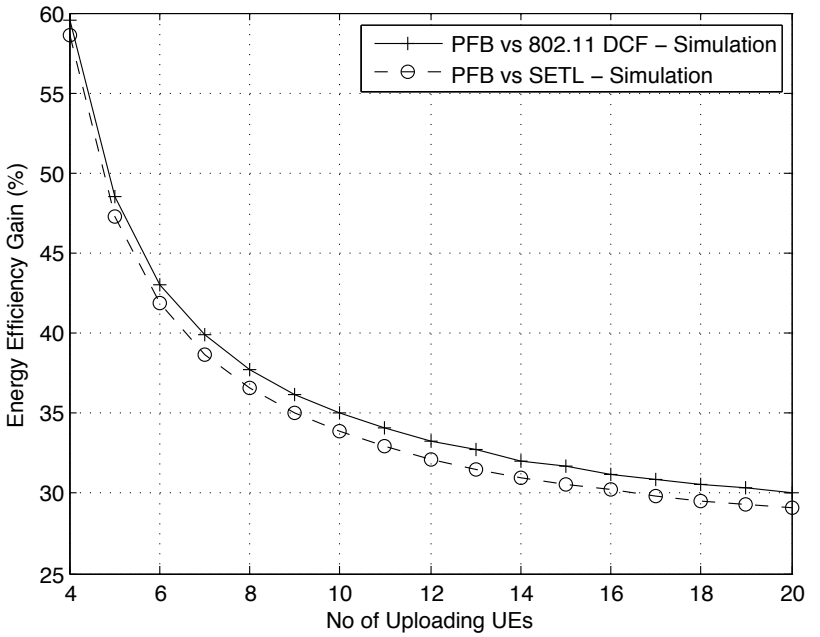
4.6.4 Trade-off Between Spectrum Efficiency and Data Volume Needs

Aiming to reveal the trade-off between the spectrum efficiency and the data volume needs of the UEs, we examine our proposed uplink offloading scheme for $N = 20$ UEs. We let UE_i for $i \in (1, ..., 19)$ with uniformly distributed $\theta_i \in [0.8, 1]$ and uniformly distributed data needs $K_i \in [5, 15]$, while we set specific values of UE_N's spectrum efficiency and data needs. In Fig. 4.13 we notice that the higher the θ_N the less favoured UE_N is to offload, which is expected as the proportionally fair allocation is weighted by $\rho_i = K_i/\theta_i$. Comparing different data needs K_N for the same θ_N we see that the offloading percentage lowers, while the absolute value of the actual offloaded volume raises, which is also expected as K_i is in the nominator of the weights ρ_i. We also compare the aggregate equivalent throughput of UE_N for the concurrent transmission through LTE and WiFi with the case that there was no opportunity to offload and all data were transmitted through LTE, following the linear pricing scheme. Throughput results for UE_N's data needs $K_N = (5, 10, 15)$ MB, are presented in Fig. 4.12(a), 4.12(b) and 4.12(c) respectively, for $\theta_i \in [0.8, 1]$, with $i = (1, ..., N)$. For these data volume needs and spectrum efficiency

range, we achieve a throughput improvement of 7% to 15.5%, due to the concurrent transmission through LTE and WiFi. Comparing the three figures, we notice that despite the fact that greater θ_N gives less access to offload, the aggregate throughput of UE_N raises due to its improved LTE uplink channel conditions.

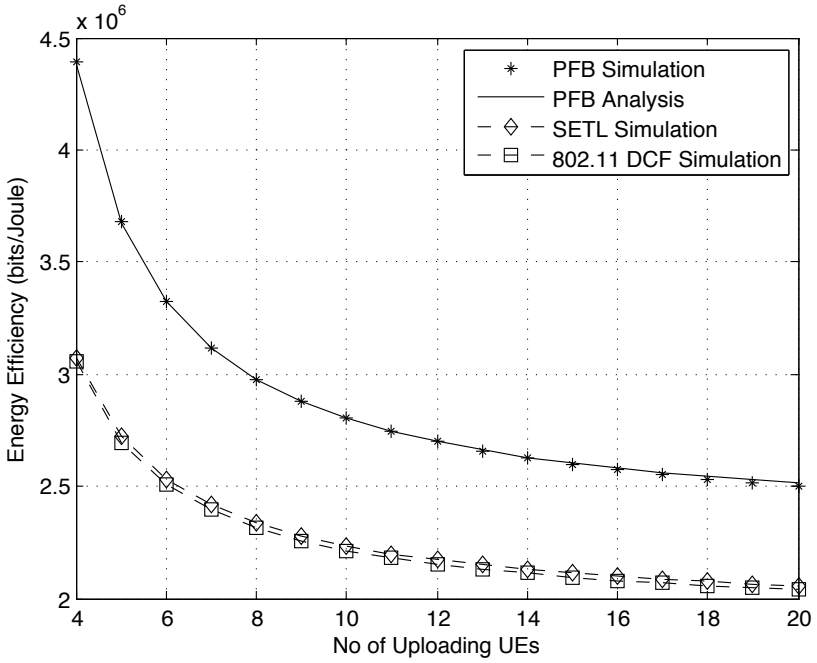


(a) Energy Efficiency.

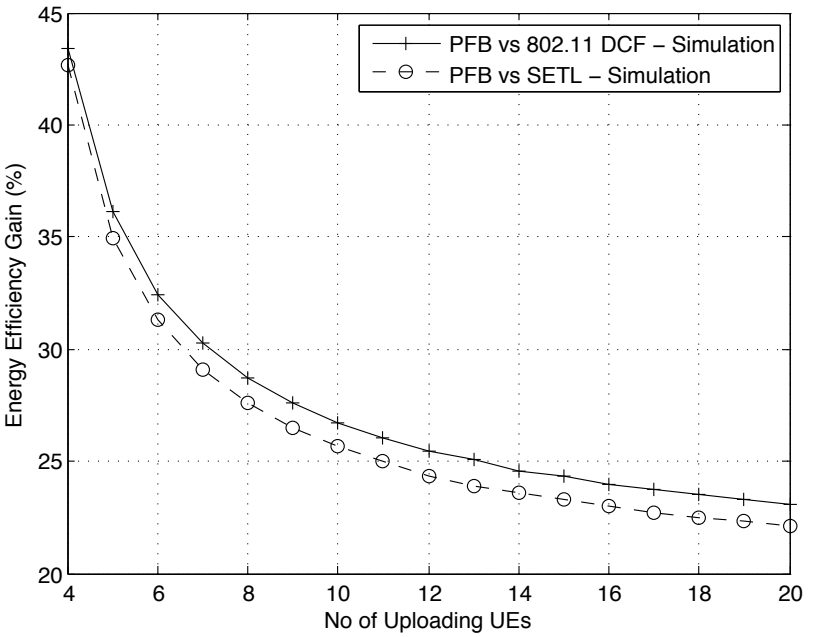


(b) Energy Efficiency Gain.

Figure 4.6: Energy efficiency and energy efficiency gain for $\theta_i \in [0.8, 1]$ with linear pricing.

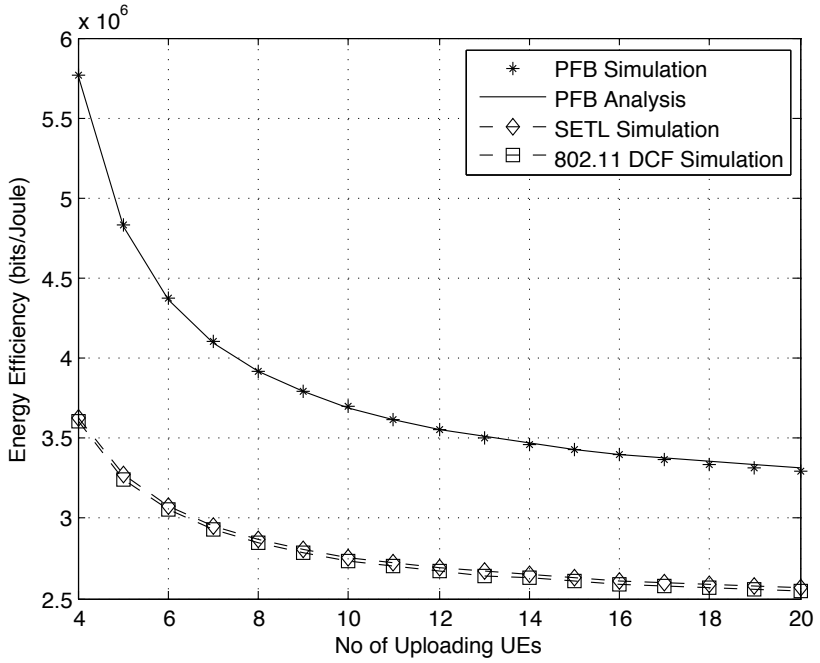


(a) Energy Efficiency.

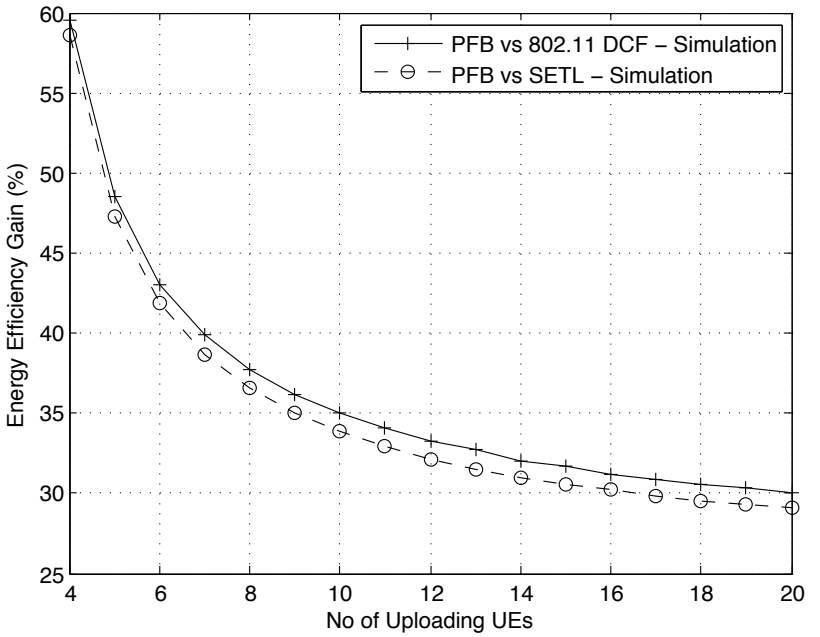


(b) Energy Efficiency Gain.

Figure 4.7: Energy efficiency and energy efficiency gain for $\theta_i \in [0.6, 0.8]$ with linear pricing.

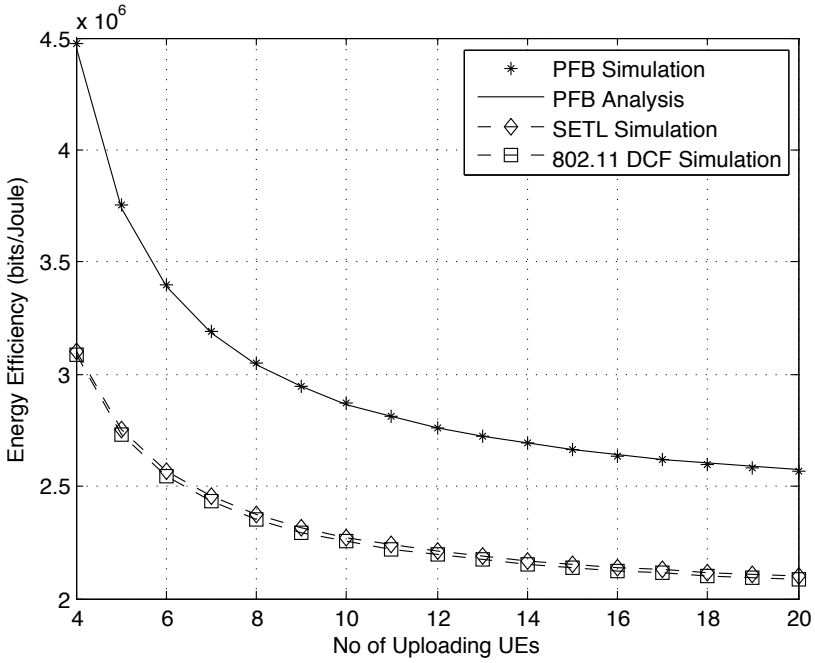


(a) Energy Efficiency.

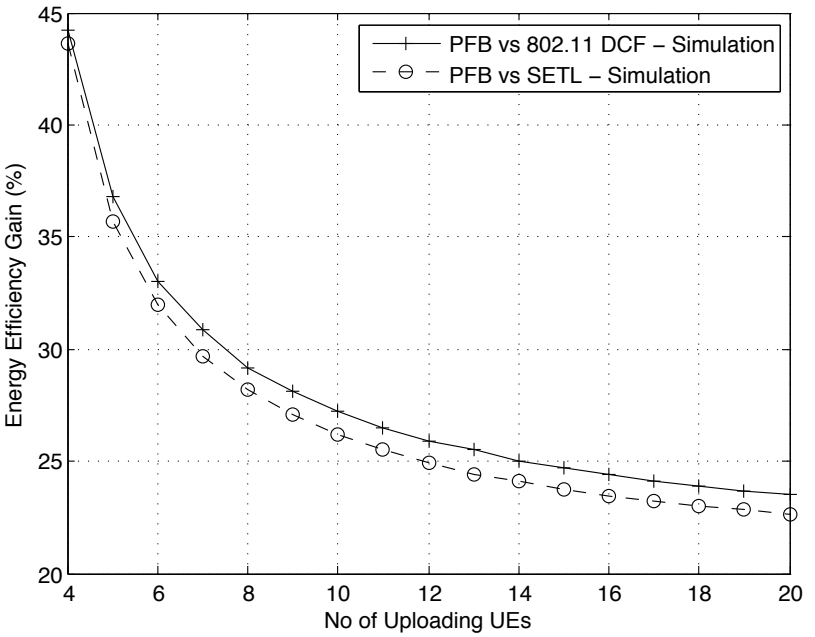


(b) Energy Efficiency Gain.

Figure 4.8: Energy efficiency and energy efficiency gain for $\theta_i \in [0.8, 1]$ with exponential pricing.



(a) Energy Efficiency.



(b) Energy Efficiency Gain.

Figure 4.9: Energy efficiency and energy efficiency gain for $\theta_i \in [0.6, 0.8]$ with exponential pricing.

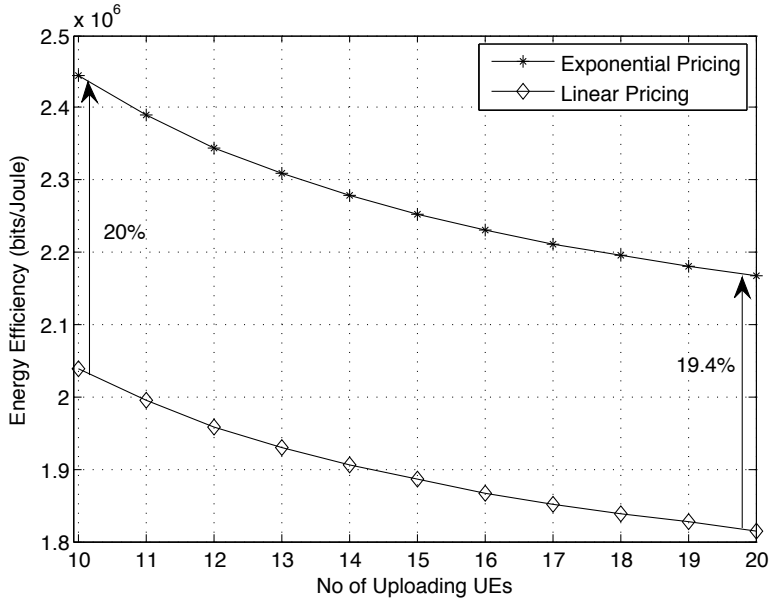


Figure 4.10: Energy efficiency comparison between linear and exponential pricing scheme.

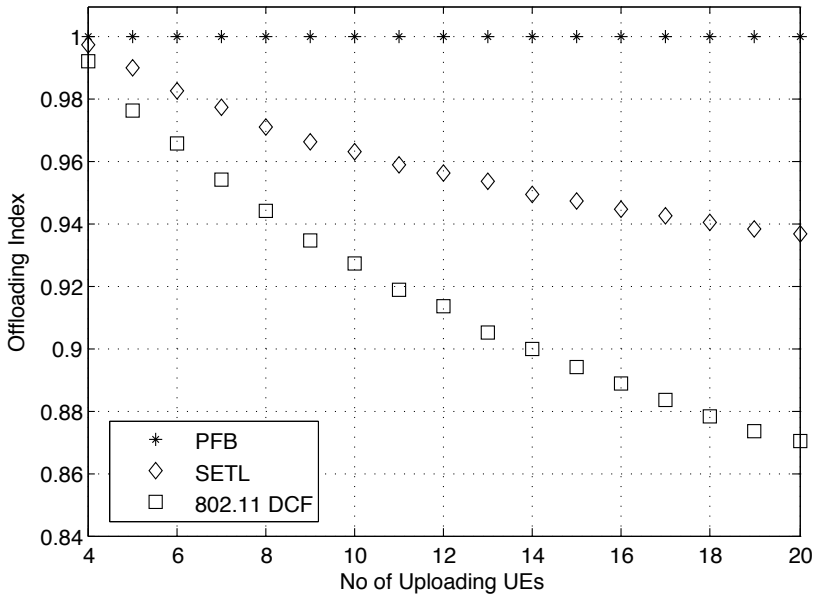
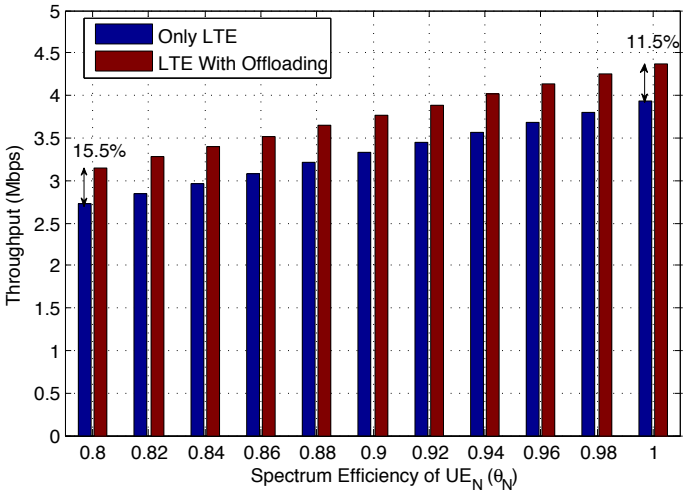
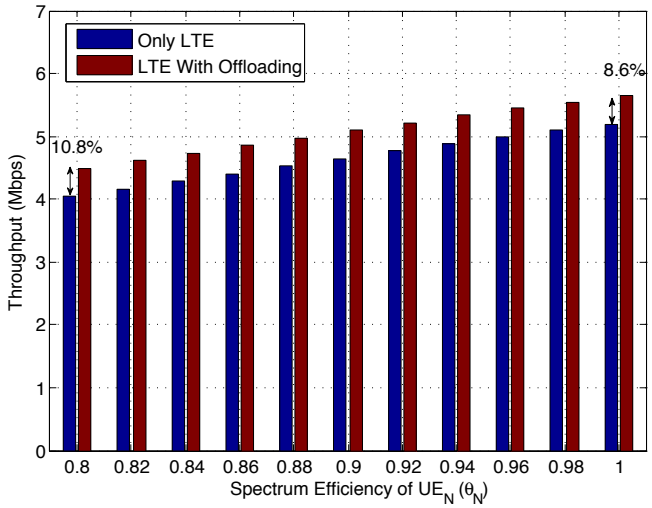


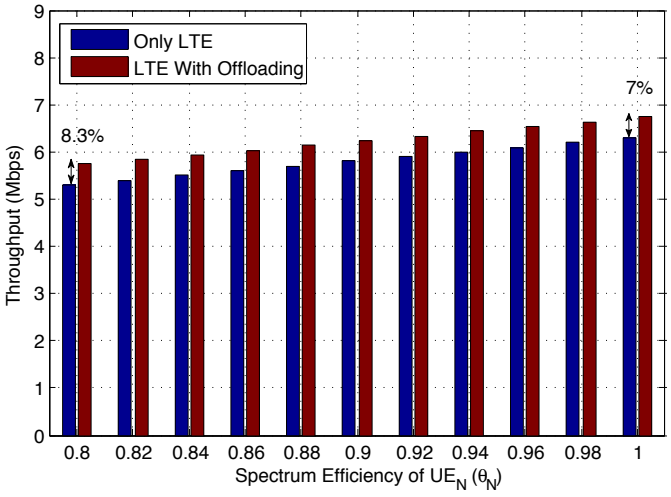
Figure 4.11: Offloading Index for different number of UEs.



(a) $K_N = 5$ MB and $\theta_N \in [0.8, 1]$.



(b) $K_N = 10$ MB and $\theta_N \in [0.8, 1]$.



(c) $K_N = 15$ MB and $\theta_N \in [0.8, 1]$.

Figure 4.12: Throughput comparison for $K_N = (5, 10, 15)$ MB and $\theta_i \in [0.8, 1]$ with linear pricing.

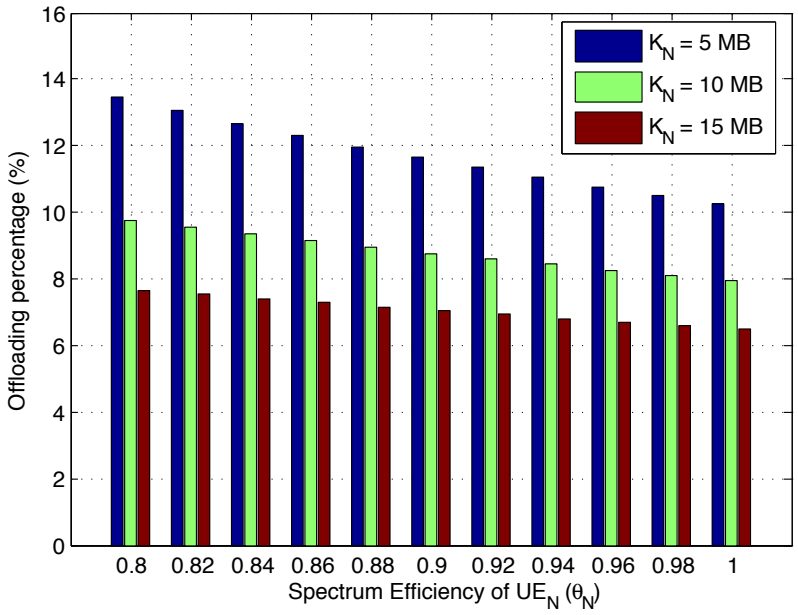


Figure 4.13: Impact of the trade-off between spectrum efficiency and data volume needs.

4.7 Combating Selfish Misbehavior with Reputation Based Offloading

In this Section, as shown in Fig. 4.14, we consider the existence of one untruthful UE that declares upload needs equal to K_{max}, that are more than its real needs, aiming to gain more WiFi bandwidth to offload, while the rest of the UEs are truthful. This part of the Chapter is published in [93]. The WiFi bandwidth is allocated based on weighted proportional fairness, and the upload needs of the users are part of the weighting factor. Hence, the more the declared upload data needs, the more WiFi bandwidth is allocated to the user. After the end of each offloading period, the eNodeB is able to identify a malicious operation and inform the WiFi AP for future allocation. We define a reputation vector $\mathbf{v}_j$, with $\mathrm{v}_j(i) \in (0, 1]$ to represent the truthfulness of the UEs during the j^{th} offloading period of duration equal to ΔT. At the start, every UE_i is considered truthful and its reputation value is equal to $\mathrm{v}_1(i) = 1$. After an offloading period j, the reputation vector is updates as follows

$$\mathrm{v}_{j+1}(i) = \begin{cases} 1, & \text{if } \mathrm{UE}_i \text{ truthful in } j^{th} \text{ period} \\ K_i/K_{max}, & \text{if } \mathrm{UE}_i \text{ untruthful in } j^{th} \text{ period} \end{cases} \tag{4.38}$$

A UE_i that is untruthful during the offloading period j is punished according to (4.38) for the following offloading period. After this punishment period the reputation of the untruthful UE_i is reset to $\mathrm{v}_{j+2}(i) = 1$. Truthful UEs maintain their reputation value equal to one. Each UE_i offloads part of its data needs K_i and the rest is uploaded through its LTE connection. We assume that the channel characteristics between each UE_i and the LTE eNodeB are described by a normalized spectrum efficiency $\theta_i \in [0, 1]$, such that for a bandwidth allocation that gives to UE_i the ability to upload with an uplink rate equal to R_i^{LTE}, the actual achieved uplink rate is equal to $\theta_i R_i^{LTE}$.

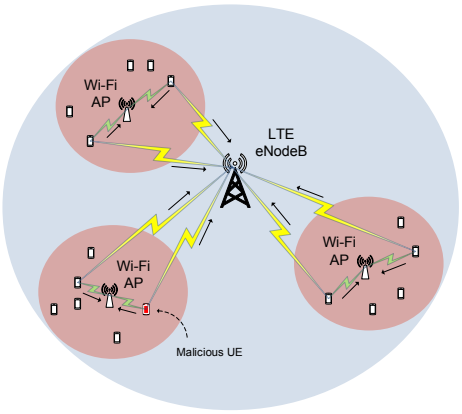


Figure 4.14: Uplink offloading scenario with IFOM and a malicious UE.

4.8 Energy Efficiency of UEs

Following we provide analytical expressions for the average energy consumption of the offloading UEs for both truthful and malicious UEs.

4.8.1 Energy Efficiency of the Truthful UEs

The average per truthful UE energy consumption of the WiFi network interface card, during the uploading phase, is expressed as

$$\overline{EC}_{Tx}^{WiFi} = \frac{1}{N-1}\left(\sum_{i=1}^{N-1}\frac{K_i/\theta_i}{\sum_{i=1}^{N}\rho_i}\Delta T\frac{S(1)}{EE(1)}\right) \text{ [Joule]} \tag{4.39}$$

After scheduling the exclusive time periods t_i in augmenting order of duration, the average per UE energy consumption of the WiFi network interface card while in sleep mode with power level P_{sleep}^{WiFi}, is expressed as

$$\overline{EC}_{sleep}^{WiFi} = \frac{1}{N-1}\sum_{i=1}^{N-2}(N-i-1)t_i P_{sleep}^{WiFi} \text{ [Joule]} \tag{4.40}$$

The average per UE energy consumption of the LTE network interface card is equal to

$$\overline{EC}^{LTE} = \frac{1}{N-1}\sum_{i=1}^{N-1}\left((K_i - t_i S(1))\frac{P_i^{LTE}}{\theta_i R_i^{LTE}}\right) \text{ [Joule]} \tag{4.41}$$

Combining (4.39)-(4.41) the average per UE energy efficiency of IFOM offloading under the PFB algorithm is expressed in (4.42).

$$\overline{E}_{eff}^{PFB} = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}K_i}{\overline{EC}_{Tx}^{WiFi} + \overline{EC}_{sleep}^{WiFi} + \overline{EC}^{LTE}} \text{ [bits/Joule]} \tag{4.42}$$

4.8.2 Energy Efficiency of the Malicious UE

The malicious UE's (UE_N) energy consumption of its WiFi network interface card, during the uploading phase, is expressed as

$$EC_{Tx}^{WiFi} = \frac{\mathrm{v}_j(N)K_N/\theta_N}{\sum_{i=1}^{N}\rho_i}\Delta T\frac{S(1)}{EE(1)} \text{ [Joule]} \tag{4.43}$$

While in sleep mode its energy consumption is expressed as

$$EC_{sleep}^{WiFi} = \left(\Delta T - \sum_{i=1}^{N-1} t_i \right) P_{sleep}^{WiFi} \text{ [Joule]} \tag{4.44}$$

The malicious UE's energy consumption of the LTE network interface card is equal to

$$EC^{LTE} = (K_N - t_N S(1)) \frac{P_N^{LTE}}{\theta_N R_N^{LTE}} \text{ [Joule]} \tag{4.45}$$

Combining (4.43)-(4.45) the malicious UE's energy efficiency of IFOM uplink offloading under the PFB algorithm is equal to

$$E_{eff}^{PFB} = \frac{K_N}{EC_{Tx}^{WiFi} + EC_{sleep}^{WiFi} + EC^{LTE}} \text{ [bits/Joule]} \tag{4.46}$$

4.9 Performance Evaluation Under Selfish Misbehavior

We evaluate the system under consideration by running extensive simulations using MATLAB^{TM}. We present the performance of a malicious UE in comparison to truthful UEs which are situated under the concurrent coverage of the same eNodeB and WiFi AP for diverse number of offloading UEs, namely for eight to 20 UEs. We compare the performance of a malicious UE to truthful UEs in terms of energy efficiency before and after the update of the reputation vector. The simulations are repetitively conducted for an offloading time period $\Delta T = 5$ sec. The data volume needs of the UEs are assumed to follow a uniform distribution of file sizes between $5 - 15$ MB. These data needs represent the volume of a photo to a small video, created by contemporary smartphones. The UE that operates in malicious mode declares its data volume equal to $K_{max} = 15$ MB. The uplink power level of UE$_i$'s LTE interface card, P_i^{LTE}, is assumed to follow (4.1) as a function of its LTE uplink rate, which is defined by the linear pricing algorithm presented in Section IV.

4.9.1 Performance Evaluation Under Linear Pricing

We perform the simulations for two different ranges of θ_i. Specifically for uniformly distributed $\theta_i \in [0.8, 1]$ and $\theta_i \in [0.6, 0.8]$. The IEEE 802.11 network interface card power levels P_{Tx}^{WiFi}, P_{Rx}^{WiFi}, P_{idle}^{WiFi} and P_{sleep}^{WiFi} are assumed to follow the measurements provided in [53]. The numerical values of the simulation parameters are presented in Table 4.2.

In Fig. 4.15 we present the energy efficiency results for spectrum efficiency $\theta_i \in [0.8, 1]$ and we can see that analysis and simulations perfectly fit. It is notable that the malicious UE performs better compared to the average energy consumption of truthful UEs ranging from 18.2% for eight offloading UEs to 5.7% for 20 UEs. This

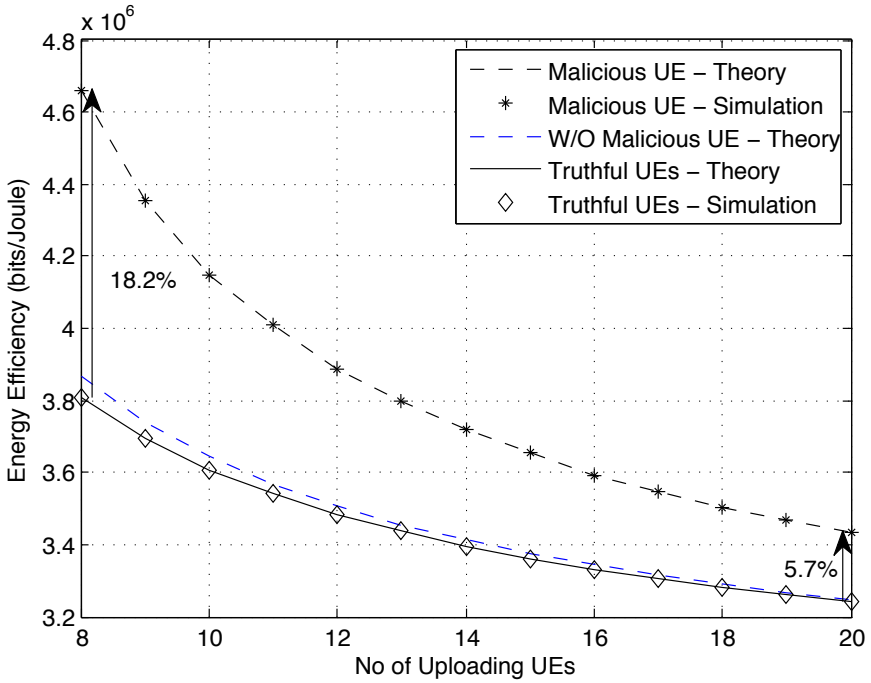


Figure 4.15: Energy efficiency during malicious operation ($\theta_i \in [0.8, 1]$).

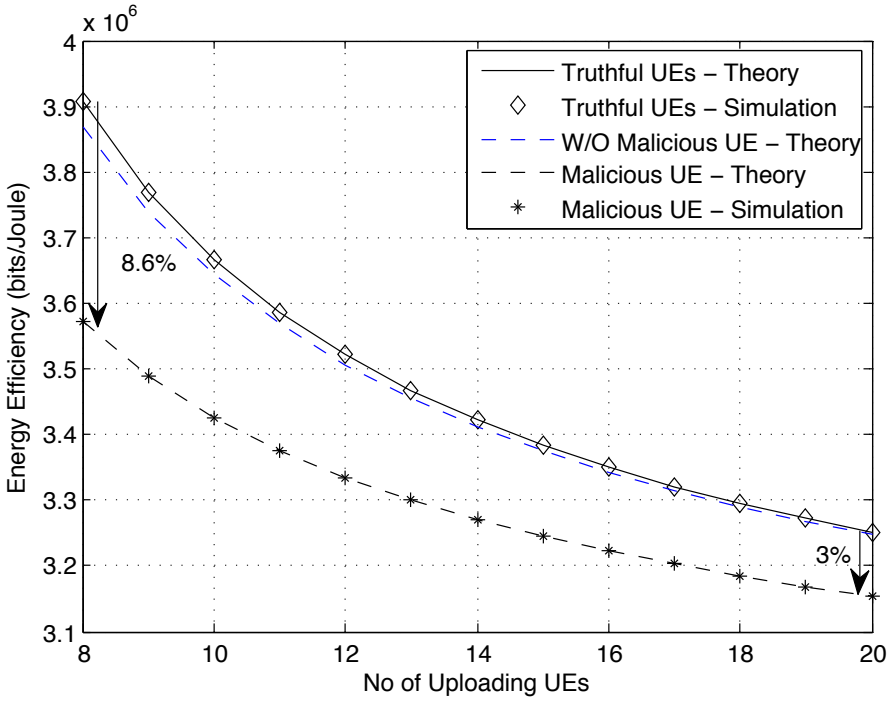


Figure 4.16: Energy efficiency after reputation vector update ($\theta_i \in [0.8, 1]$).

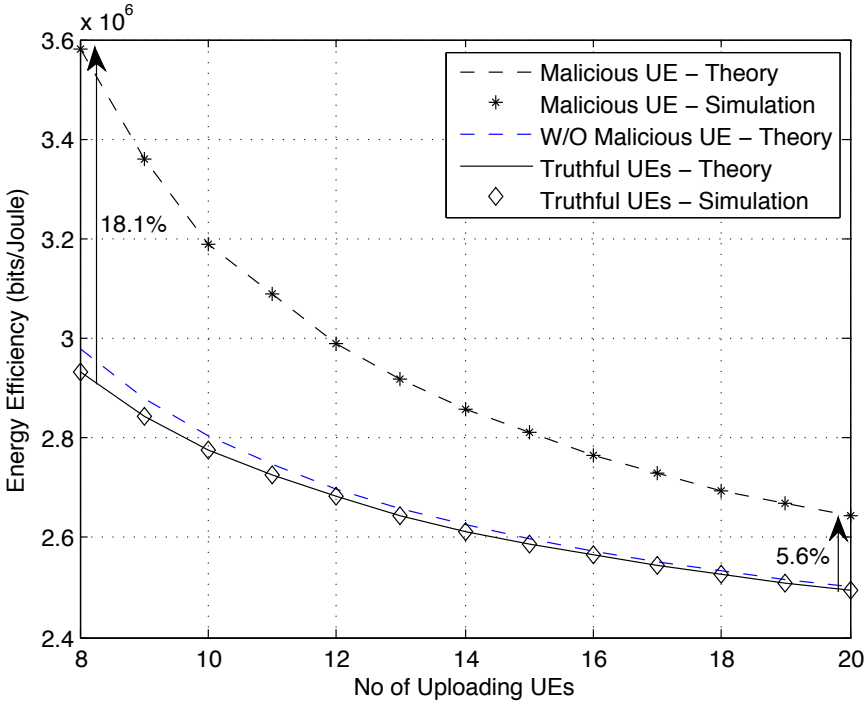


Figure 4.17: Energy efficiency during malicious operation ($\theta_i \in [0.6, 0.8]$).

happens because the malicious UE, by declaring more uplink data volume needs than its real, is allocated more WiFi bandwidth and uploads less data through its LTE connection, which is more energy consuming. The truthful UEs present a slight reduction in their energy efficiency compared to the case where no malicious UE exists. In Fig. 4.16, we can see that after the update of the reputation vector the energy efficiency of the malicious UE is deteriorated as the punishment rule applies. This deterioration varies from 8.6% for eight UEs to 3% for 20 UEs compared to a truthful UE. We also notice that the truthful UEs present a slight improvement compared to the case of absence of a malicious UE. This happens because the WiFi bandwidth that is not allocated to the malicious UE is proportionally allocated to the truthful UEs, helping them upload less data through their LTE connections. In Fig. 4.17, we present the energy efficiency of the malicious UE compared to the truthful UEs for $\theta_i \in [0.6, 0.8]$ and in Fig. 4.18 the energy efficiency results after the update of the reputation vector. We notice analogous performance gain and loss with the case of $\theta_i \in [0.8, 1]$ but for lower values of the energy efficiency.

4.9.2 Performance Evaluation Under Linear Pricing

We perform the simulations for $\theta_i \in [0.8, 1]$. The IEEE 802.11 network interface card power levels P_{Tx}^{WiFi}, P_{Rx}^{WiFi}, P_{idle}^{WiFi} and P_{sleep}^{WiFi} are assumed to follow the measurements provided in [53]. The numerical values of the simulation parameters are presented in Table 4.2.

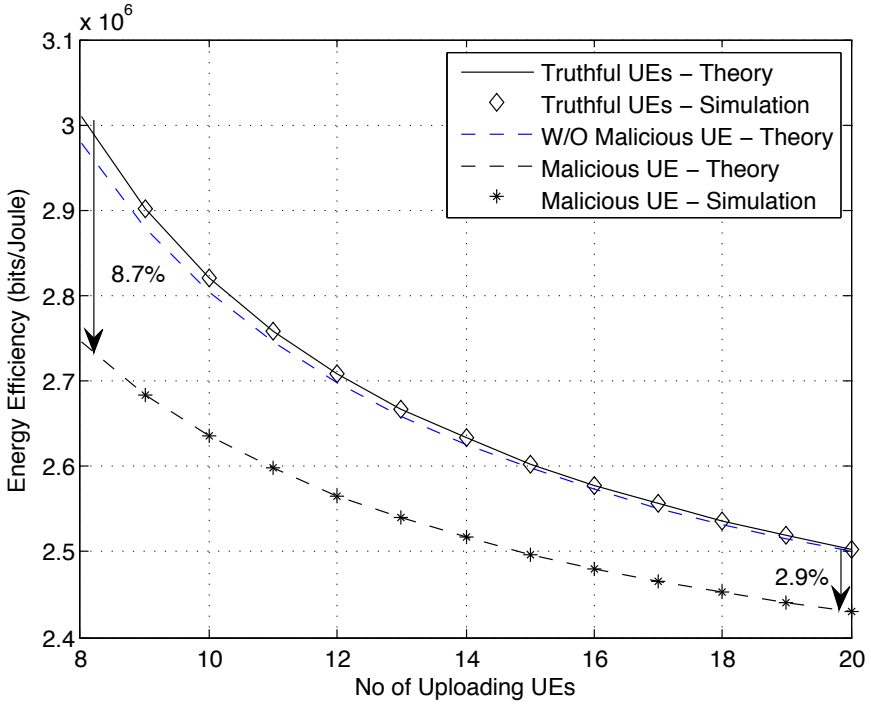
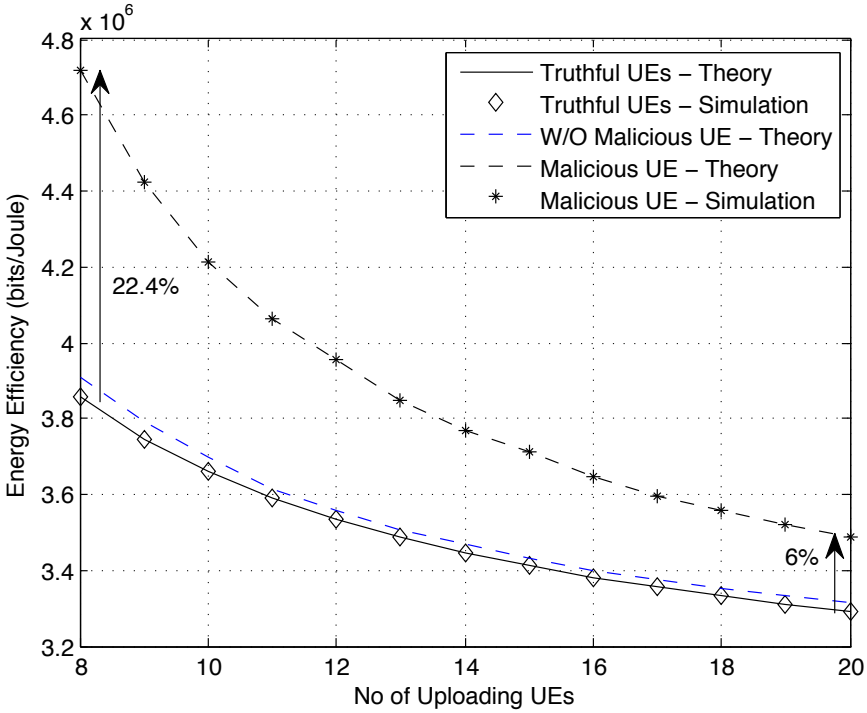
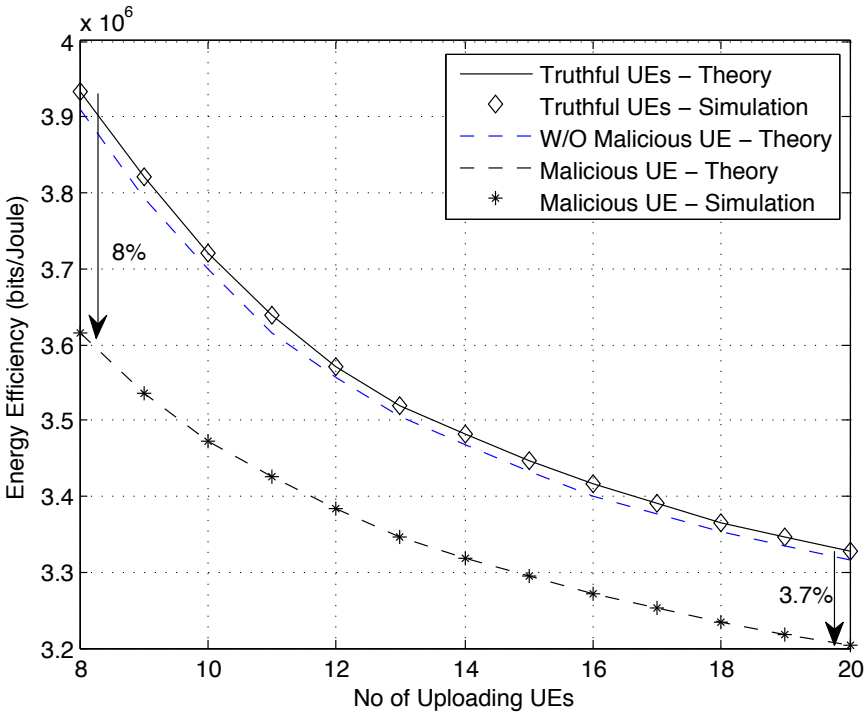


Figure 4.18: Energy efficiency after reputation vector update ($\theta_i \in [0.6, 0.8]$).

In Fig. 4.19(a) we present the energy efficiency results for spectrum efficiency $\theta_i \in [0.8, 1]$ and we can see that analysis and simulations perfectly fit. It is notable that the malicious UE performs better compared to the average energy consumption of truthful UEs ranging from 22.4% for eight offloading UEs to 6% for 20 UEs. This happens because the malicious UE, by declaring more uplink data volume needs than its real, is allocated more WiFi bandwidth and uploads less data through its LTE connection, which is more energy consuming. The truthful UEs present a slight reduction in their energy efficiency compared to the case where no malicious UE exists. In Fig. 4.19(b), we can see that after the update of the reputation vector the energy efficiency of the malicious UE is deteriorated as the punishment rule applies. This deterioration varies from 8% for eight UEs to 3.7% for 20 UEs compared to a truthful UE. We also notice that the truthful UEs present a slight improvement compared to the case of absence of a malicious UE. This happens because the WiFi bandwidth that is not allocated to the malicious UE is proportionally allocated to the truthful UEs, helping them upload less data through their LTE connections. In Fig. 4.20(a), we present a throughput comparison between a truthful and a malicious UE assuming they experience the same channel conditions $\theta_i \in [0.8, 0.1]$ with a step equal to 0.02. This comparison refers to an offloading period with malicious operation. We notice that a malicious UE presents from 5% to 6.2% better throughput. In Fig. 4.20(b) the throughput comparison is presented for the offloading period after the update of the reputation vector, where the malicious UE is punished. In this case we notice that the malicious UE presents from 3.2% to 4% less throughput compared to the truthful UE.

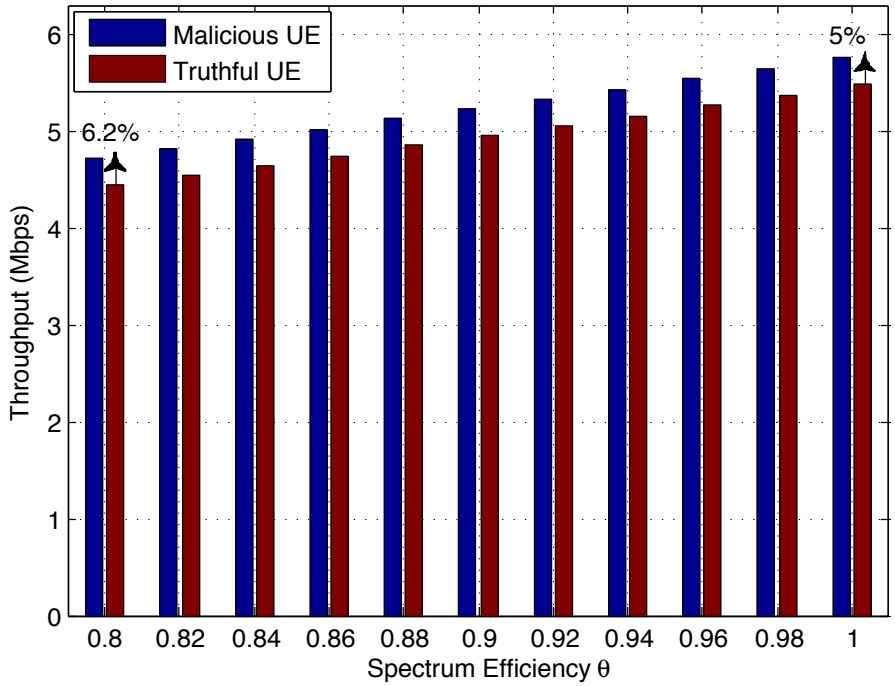


(a) During malicious operation.

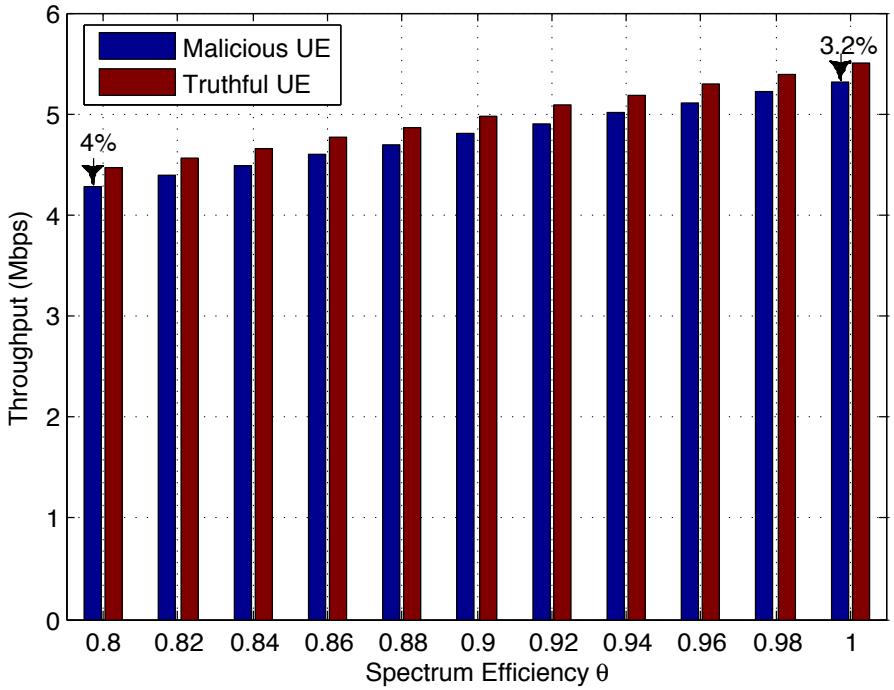


(b) After reputation vector update.

Figure 4.19: Energy efficiency for $\theta_i \in [0.8, 1]$.



(a) During malicious operation.



(b) After reputation vector update.

Figure 4.20: Throughput comparison between a malicious and a truthful UE.

Chapter 5

Conclusions

The continuously increasing mobile data traffic along with the data demanding mobile applications of contemporary mobile devices is creating congestion conditions on the cellular network infrastructures. Offloading techniques that exploit heterogeneity of access technologies are able to mitigate these conditions by providing more energy efficient access. In this thesis, efficient schemes and algorithms are proposed to combine short range networks with cellular access. Regarding downlink, a cooperative physical layer network coding scheme is proposed for data dissemination, while for uplink offloading a weighted proportional fair access in conjunction with pricing schemes for cellular access are investigated.

A new multicast protocol for short range networks called CooPNC is proposed, in order to mitigate downlink load of cellular operators under the conditions of common interest in the same time of neighboring users. In this protocol, an indirect cooperative scheme at the medium access control (MAC) layer is proposed. As verified by the mathematical analysis and the extensive simulations, CooPNC can achieve substantial throughput, delay and energy efficiency performance improvements compared to state of the art multicast protocols. Moreover, the improvement of the multicast performance with CooPNC is achieved without incurring extra control overhead. This result is due to the features of ANC and the ability of CooPNC to share the retransmissions of collided packets between the neighbouring networks. This characteristic allows STAs situated in the same overlapping area to experience benefit from the retransmissions of the relays of both neighbouring networks, while STAs acting as relays to share the needed retransmissions for the multicast communication to be reliable.

The performance improvement of data dissemination with CooPNC was achieved by cooperative Physical Layer Network Coding. The upper bounds of this technique were revealed in detail for a broadly used scenario in NC techniques, the Cross Network scenario. The impact of MAC fairness on the throughput performance of the Cross Network was presented and analysed, when applying different NC techniques. In addition to the pure relaying the performance of both digital NC and

PNC with and without overhearing was investigated in each case. A throughput and energy efficiency comparison of these techniques was provided and the need for MAC protocol improvements that will exploit the improved throughput offered by NC techniques was discussed.

Considering uplink offloading, the proposed access schemes were focused on the IFOM technique for uplink offloading that allows the concurrent transmission through WiFi and LTE. UEs are scheduled to upload through the WiFi following a weighted proportionally fair allocation algorithm, where both the data load and the LTE spectrum efficiency is considered in the weighted fairness. Regarding the WiFi access scheme we presented our implementation considerations and proposed a scheme that provides for maximum exploitation of the WiFi resources. For the LTE uplink rate allocation two pricing algorithms were proposed, one linear and one exponential. These algorithms are consisted of two stages. In the first stage the operator chooses the price and in the second stage UEs decide the quantity of resources they intend to acquire based on their payoff functions. An energy efficiency evaluation of this offloading approach was presented, for both the linear and the exponential pricing models, comparing at the same time the performance of the proportionally fair access with the standard 802.11 and a state-of-the-art access algorithm. The benefit of exponential pricing in terms of energy efficiency for high diversity of the LTE channel conditions of the UEs was also revealed and the results for the interplay of spectrum efficiency and uplink data needs were presented. In addition the offloading index and the UEs aggregate throughput was investigated.

The proposed uplink offloading approach with IFOM, was also investigated for the case of existence of a malicious UE among other truthful UEs under the coverage of the same eNodeB and WiFi AP. A weighted proportionally fair algorithm was proposed for the WiFi access and two pricing based algorithm, one linear and one exponential, were examined for the LTE access. The energy efficiency performance of the malicious UE compared to the truthful UEs was investigated, and a reputation based reaction method was proposed to combat the malicious operation. Through analysis and simulation it is shown that the performance of a malicious UE significantly decreases after applying the proposed reaction method.

Concluding, this thesis has advanced the state of the art first by presenting CooPNC, a new efficient MAC layer protocol for data dissemination combining cooperative techniques with Physical Layer Network Coding and, second, by introducing an efficient hybrid access scheme for uplink offloading in heterogeneous condition consisting of cellular and WiFi networks. The two parts of the thesis have provided valuable results that improve MAC layer access in the aspects of throughput, delay, and energy efficiency. Even though they have been treated independently throughout this dissertation, it is possible to envision a system where both parts are combined. This joint scenario could consist of the use of CooPNC in the described downlink cases for data dissemination, combined with the proposed offloading scheme in the uplink.

Bibliography

[1] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (c-ran): A primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, 2015.

[2] Cisco Systems Inc, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper." [Online].Available:http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html[Accessed:12Mar.2015].

[3] M. A. Ergin, K. Ramachandran, and M. Gruteser, "An experimental study of inter-cell interference effects on system performance in unplanned wireless lan deployments," *Computer Networks*, vol. 52, no. 14, pp. 2728 – 2744, 2008.

[4] V. Miliotis, L. Alonso, and C. Verikoukis, "Multicast performance bounds exploiting cooperative physical layer network coding," in *Proc. IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 135 –139, 2012.

[5] V. Miliotis, L. Alonso, and C. Verikoukis, "Combining cooperation and physical layer network coding to achieve reliable multicast," *Recent Advances in Communications and Networking Technology (Formerly Recent Patents on Telecommunication)*, vol. 2, no. 1, pp. 41–49, 2013.

[6] V. Miliotis, L. Alonso, C. Skianis, and C. Verikoukis, "The impact of cooperative physical layer network coding on multicast short range networks," in *IEEE International Conference on Communications (ICC)*, pp. 3547–3551, IEEE, 2013.

[7] V. Miliotis, L. Alonso, and C. Verikoukis, "Cooperative multicast exploiting physical layer network coding: A performance analysis," in *IEEE International Conference on Communications Workshops (ICC)*, pp. 1010–1014, IEEE, 2013.

[8] V. Miliotis, L. Alonso, and C. Verikoukis, "Coopnc: A cooperative multicast protocol exploiting physical layer network coding," *Ad Hoc Networks*, vol. 14, pp. 35–50, 2014.

[9] "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 802.11-2007 (Revision of IEEE Std 802.11-2007)*, 2012.

[10] Z. Yang, M. Li, and W. Lou, "R-code: Network coding-based reliable broadcast in wireless mesh networks," *Ad Hoc Networks*, vol. 9, no. 5, pp. 788–798, 2011.

[11] S. Roy, D. Koutsonikolas, S. Das, and Y. C. Hu, "High-throughput multicast routing metrics in wireless mesh networks," *Ad Hoc Networks*, vol. 6, no. 6, pp. 878–899, 2008.

[12] J. Kuri and S. K. Kasera, "Reliable Multicast in Multi-Access Wireless LANs," *Wireless Networks*, vol. 7, no. 4, pp. 359–369, 2001.

[13] S. Gupta, V. Shankar, and S. Lalwani, "Reliable multicast MAC protocol for wireless LANs," in *Proc. IEEE International Conference on Communications, ICC '03*, vol. 1, pp. 93–97, 2003.

[14] C.-W. Bao and W. Liao, "Performance analysis of reliable MAC-layer multicast for IEEE 802.11 Wireless LANs," in *Proc. IEEE International Conference on Communications, ICC '05*, vol. 2, pp. 1378–1382, 2005.

[15] K. Tang and M. Gerla, "MAC reliable broadcast in ad hoc networks," in *Proc. IEEE Military Communications Conference, MILCOM '01. Communications for Network-Centric Operations: Creating the Information Force.*, vol. 2, pp. 1008–1013, 2001.

[16] M.-T. Sun, L. Huang, A. Arora, and T.-H. Lai, "Reliable MAC layer multicast in IEEE 802.11 wireless networks," in *Proc. International Conference on Parallel Processing.*, pp. 527–536, 2002.

[17] M.-T. Sun, L. Huang, S. Wang, A. Arora, and T.-H. Lai, "Reliable MAC layer multicast in IEEE 802.11 wireless networks," *Wireless Communications and Mobile Computing*, vol. 3, no. 4, pp. 439–453, 2003.

[18] C. Campolo, A. Molinaro, C. Casetti, and C.-F. Chiasserini, "An 802.11-Based MAC Protocol for Reliable Multicast in Multihop Networks," in *IEEE 69th Vehicular Technology Conference. VTC Spring '09.*, pp. 1–5, 2009.

[19] J. Alonso-Zarate, L. Alonso, and C. Verikoukis, "Performance analysis of a persistent relay carrier sensing multiple access protocol," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5827 –5831, 2009.

[20] J. Alonso-Zarate, L. Alonso, and C. Verikoukis, "Performance analysis of a persistent relay carrier sensing multiple access protocol," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5827 –5831, 2009.

[21] J. Vella and S. Zammit, "A Survey of Multicasting over Wireless Access Networks," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–36, 2012.

[22] K. J. Yoon and Y. Y. Kim, "Relayed Multicast Scheme in IEEE 802.11 Wireless LAN Systems," in *Proc. 25th IEEE International Conference on Computer Communications. INFOCOM 2006*, pp. 1–2.

[23] B. Rong and A. Hafid, "Cooperative Multicast for Mobile IPTV Over Wireless Mesh Networks: The Relay-Selection Study," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2207–2218, 2010.

[24] O. Alay, T. Korakis, Y. Wang, E. Erkip, and S. Panwar, "Layered wireless video multicast using omni-directional relays," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP 2008*, pp. 2149–2152.

[25] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "Coopmac: A cooperative mac for wireless lans," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 340–354, 2007.

[26] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[27] C. Fragouli and E. Soljanin, *Network Coding Fundamentals*. NOW Publishers, 2007.

[28] Heide, J. and Pedersen, M.V. and Fitzek, FH and Larsen, T., "Cautious view on network coding-from theory to practice," *Journal of Communications and Networks (JCN)*, vol. 10, no. 4, pp. 403–411, 2008.

[29] M. Pedersen, J. Heide, F. Fitzek, and T. Larsen, "PictureViewer - A mobile application using network coding," in *Proc. European Wireless Conference. EW 2009*, pp. 151–156.

[30] L. Militano, F. Fitzek, A. Iera, and A. Molinaro, "Group Interactions in Wireless Cooperative Networks," in *Proc. IEEE 73rd Vehicular Technology Conference (VTC Spring), 2011*, pp. 1–5.

[31] P. Vingelmann, M. Pedersen, F. Fitzek, and J. Heide, "On-the-Fly Packet Error Recovery in a Cooperative Cluster of Mobile Devices," in *Proc. IEEE Global Telecommunications Conference, GLOBECOM 2011*, pp. 1–6.

[32] A. Rezaee, L. Zeger, and M. Médard, "Multi packet reception and network coding," in *MILCOM 2010*, pp. 1393–1398, IEEE, 2010.

[33] F. Rossetto and M. Zorzi, "Mixing network coding and cooperation for reliable wireless communications," *IEEE Wireless Communications*, vol. 18, no. 1, pp. 15–21, 2011.

[34] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON)*, vol. 3, pp. 817–822, IEEE, 2006.

[35] X. Bao and J. Li, "Adaptive network coded cooperation (ancc) for wireless relay networks: matching code-on-graph with network-on-graph," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 574–583, 2008.

[36] L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello, "A network coding approach to cooperative diversity," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3714–3722, 2007.

[37] S. Yang and R. Koetter, "Network coding over a noisy relay: a belief propagation approach," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 801–804, IEEE, 2007.

[38] E. Fasolo, F. Rossetto, and M. Zorzi, "Network coding meets mimo," in *Fourth Workshop on Network Coding, Theory and Applications (NetCod)*, pp. 1–6, IEEE, 2008.

[39] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical-layer network coding," in *Proc. of the 12th annual international conference on Mobile computing and networking*, pp. 358–365, ACM MobiCom 2006.

[40] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *ACM SIGCOMM Computer Communication Review*, vol. 37, pp. 397–408, ACM, 2007.

[41] S. Katti, I. Marie, A. Goldsmith, D. Katabi, and M. Médard, "Joint relaying and network coding in wireless networks," in *IEEE International Symposium on Information Theory(ISIT)*, pp. 1101–1105, IEEE, 2007.

[42] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *IEEE 63rd Vehicular Technology Conference, VTC 2006-Spring*, vol. 2, pp. 588–593, IEEE, 2006.

[43] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, "On the study of network coding with diversity," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1247–1259, 2009.

[44] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3498–3516, 2007.

[45] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *IEEE International Conference on Communications, ICC'06*, vol. 9, pp. 3885–3890, IEEE, 2006.

[46] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, 2007.

[47] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," *Physical Communication*, vol. 6, pp. 4–42, 2013.

[48] L. Lu, T. Wang, S. C. Liew, and S. Zhang, "Implementation of physical-layer network coding," in *Proc. IEEE International Conference on Communications, ICC '12*, pp. 4734–4740, 2012.

[49] L. Lu, T. Wang, S. C. Liew, and S. Zhang, "Implementation of physical-layer network coding," *Physical Communication*, vol. 6, pp. 74–87, 2013.

[50] S. Gollakota and D. Katabi, "Zigzag decoding: combating hidden terminals in wireless networks," in *Proc. of the ACM SIGCOMM 2008 conference on Data communication*, (New York, NY, USA), pp. 159–170, ACM, 2008.

[51] J. Qureshi, J. Cai, and C. H. Foh, "Cooperative Retransmissions through Collisions," in *Proc. IEEE International Conference on Communications, ICC '11*, pp. 1–5, 2011.

[52] E. Fasolo, A. Munari, F. Rossetto, and M. Zorzi, "Phoenix: a hybrid cooperative-network coding protocol for fast failure recovery in ad hoc networks," in *5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON'08*, pp. 404–412, IEEE, 2008.

[53] J. Ebert, S. Aier, G. Kofahl, A. Becker, B. Burns, and A. Wolisz, "Measurement and simulation of the energy consumption of an wlan interface," *Technical University of Berlin, Telecommunication Networks Group, Tech. Rep. TKN-02-010*, 2002.

[54] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," in *Proc. ACM SIGCOMM '06*, vol. 36, pp. 243–254.

[55] F. Zhao and M. Médard, "On analyzing and improving cope performance," in *Proc. IEEE Information Theory and Applications Workshop (ITA)*, pp. 1–6, 2010.

[56] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding Traffic Dynamics in Cellular Data Networks," in *Proc. of IEEE INFOCOM 2011*, Apr. 2011.

[57] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "A Close Examination of performance and Power Characteristics of 4G LTE Networks," in *Proc. of the 10th ACM International Conference on Mobile Systems, Applications, and Services*, June 2012.

[58] C. Sankaran, "Data Offloading Techniques in 3GPP Rel-10 Networks: A Tutorial," *IEEE Communications Magazine*, vol. 50, pp. 46–53, June 2012.

[59] ETSI, "3GPP TS 23.261: IP flow mobility and seamless Wireless Local Area Network (WLAN) offload; Stage 2 (v11.0.0)." Sep. 2012.

[60] H. Son, S. Lee, S.-C. Kim, and Y.-S. Shin, "Soft Load Balancing Over Heterogeneous Wireless Networks," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 2632–2638, July 2008.

[61] A. Eryilmaz and R. Srikant, "Fair Resource Allocation in Wireless Networks using Queue-length-based Scheduling and Congestion Control," in *Proc. of IEEE INFOCOM 2011*, IEEE, Mar. 2005.

[62] A. Banchs, P. Serrano, and H. Oliver, "Proportional Fair Throughput Allocation in Multirate IEEE 802.11e Wireless LANs," *Wireless Networks*, vol. 13, no. 5, pp. 649–662, 2007.

[63] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "Smart Data Pricing (SDP): Economic Solutions to Network Congestion," *SIGCOMM eBook on Recent Advances in Networking*, 2013.

[64] C. U. Saraydar, N. B. Mandayam, and D. Goodman, "Efficient Power Control via Pricing in Wireless Data Networks," *IEEE Trans. on Communications*, vol. 50, no. 2, pp. 291–303, 2002.

[65] P. Liu, P. Zhang, S. Jordan, and M. L. Honig, "Single-Cell Forward Link Power Allocation Using Pricing in Wireless Networks," *IEEE Trans. on Wireless Communications*, vol. 3, no. 2, pp. 533–543, 2004.

[66] E.-E. Tsiropoulou, G. K. Katsinis, and S. Papavassiliou, "Distributed Uplink Power Control in Multiservice Wireless Networks via a Game Theoretic Approach with Convex Pricing," *IEEE Trans. on Parallel and Distributed Systems*, vol. 23, no. 1, pp. 61–68, 2012.

[67] C. H. Ke, C. C. Wei, K. W. Lin, and J. W. Ding, "A Smart Exponential-Threshold-Linear Backoff Mechanism for ieee 802.11 WLANs," *International Journal of Communication Systems*, vol. 24, pp. 1033–1048, Aug. 2011.

[68] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?," *IEEE/ACM Trans. on Networking*, vol. 21, pp. 536–550, Apr. 2013.

[69] B. Han, P. Hui, V. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular Traffic Offloading Through Opportunistic Communications: A Case Study," in *Proc. of the 5th ACM Workshop on Challenged Networks*, Sep. 2010.

[70] L. Al-Kanj, H. Poor, and Z. Dawy, "Optimal Cellular Offloading via Device-to-Device Communication Networks With Fairness Constraints," *IEEE Trans. on Wireless Communications*, vol. 13, pp. 4628–4643, Aug. 2014.

[71] H. Izumikawa and J. Katto, "Rocnet: Spatial mobile data offload with user-behavior prediction through delay tolerant networks," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pp. 2196–2201, IEEE, 2013.

[72] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting Mobile 3G Using WiFi," in *Proc. of the 8th ACM International Conference on Mobile Systems, Applications, and Services*, June 2010.

[73] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "An Iterative Double Auction for Mobile Data Offloading," in *Proc of 11th International Symposium on Modeling Optimization in Mobile, Ad Hoc Wireless Networks (WiOpt '13)*, May 2013.

[74] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. W. Cho, "iDEAL: Incentivized Dynamic Cellular Offloading via Auctions," *IEEE/ACM Trans. on Networking*, vol. 22, pp. 1271–1284, Aug. 2014.

[75] M. H. Cheung and J. Huang, "Optimal Delayed Wi-Fi Offloading," in *Proc. of 11th IEEE International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt '13)*, May 2013.

[76] A. J. Nicholson and B. D. Noble, "Breadcrumbs: Forecasting Mobile Connectivity," in *Proc. of the 14th ACM International Conference on Mobile Computing and Networking*, Sep. 2008.

[77] W. Yoon and B. Jang, "Enhanced Non-Seamless Offload for LTE and WLAN Networks," *IEEE Communications Letters*, vol. 17, pp. 1960–1963, Oct. 2013.

[78] E. Patouni, N. Alonistioti, and L. Merakos, "Modeling and Performance Evaluation of Reconfiguration Decision Making in Heterogeneous Radio Network Environments," *IEEE Trans. on Vehicular Technology,*, vol. 59, no. 4, pp. 1887–1900, 2010.

[79] S.-I. Sou, "Mobile Data Offloading With Policy and Charging Control in 3GPP Core Network," *IEEE Trans. on Vehicular Technology,*, vol. 62, pp. 3481–3486, Sept 2013.

[80] B. H. Jung, N.-O. Song, and D. K. Sung, "A Network-Assisted User-Centric WiFi-Offloading Model for Maximizing Per-User Throughput in a Heterogeneous Network," *IEEE Trans. on Vehicular Technology,*, vol. 63, pp. 1940–1945, May 2014.

[81] J. Parikh, "Technical talk in Open Compute Summit," July 2013.

[82] L. Guang, C. Assi, and A. Benslimane, "Mac Layer Misbehavior in Wireless Networks: Challenges and Solutions," *IEEE Wireless Commun. Mag.*, vol. 15, pp. 6–14, Aug. 2008.

[83] S. Buchegger, J. Mundinger, and J.-Y. Le Boudec, "Reputation Systems for Self-Organized Networks," *IEEE Technology and Society Magazine*, vol. 27, no. 1, pp. 41–47, 2008.

[84] P. Serrano, A. Banchs, V. Targon, and J. Kukielka, "Detecting Selfish Configurations in 802.11 WLANs," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 142–144, 2010.

[85] J. Choi, A. W. Min, and K. G. Shin, "On Selfish Configuration in Wi-Fi Tethering," *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 841–843, 2013.

[86] V. Miliotis, L. Alonso, and C. Verikoukis, "Offloading with ifom: The uplink case," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 2661–2666, IEEE, 2014.

[87] G. Bianchi, "Performance Analysis of the ieee 802.11 Distributed Coordination Function," *IEEE JSAC*, vol. 18, pp. 535–547, Mar. 2000.

[88] V. Miliotis, L. Alonso, and C. Verikoukis, "Energy efficient proportionally fair uplink offloading for ip flow mobility," in *IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 6–10, IEEE, 2014.

[89] F. P. Kelly, A. K. Maulloo, and D. K. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research society*, pp. 237–252, 1998.

[90] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009.

[91] Y. Li, D. Papagiannaki, and A. Sheth, "Uplink Traffic Control in Home 802.11 Wireless Networks," in *Proc. of the 2nd ACM SIGCOMM Workshop on Home Networks*, Aug. 2011.

[92] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W Function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[93] V. Miliotis, L. Alonso, and C. Verikoukis, "Combating selfish misbehavior with reputation based uplink offloading for ip flow mobility," in *IEEE International Conference on Communications (ICC)*, IEEE, 2015.