Energy-efficient enhancements for IEEE 802.11 WLANs

On the way to enable Cellular/Wi-Fi networks interworking

FERNANDO VALDENEBRO GONZÁLEZ

KTH Information and
Communication Technology

Degree project in
Communication Systems
Second level, 30.0 HEC
Stockholm, Sweden

Energy-efficient enhancements for IEEE 802.11 WLANs

On the way to enable Cellular/Wi-Fi networks interworking

Fernando Valdenebro González

2014-02-28

Examiner and academic adviser: Professor Gerald Q. Maguire Jr.

Industrial adviser: Henrik Lundqvist, Huawei Technologies Co., Ltd.

School of Information and Communication Technology (ICT)
KTH Royal Institute of Technology
Stockholm, Sweden

Abstract

Globally, the number of mobile broadband subscriptions is growing and the amount of mobile data traffic is expected to continue to grow rapidly. In the next five years the number of smartphone subscriptions is expected to more than double, while the amount of mobile traffic per active subscription per month of these subscribers is expected to nearly quadruple.

As a consequence, mobile network operators (MNOs) aim to increase radio network capacity and coverage through heterogeneous deployments. In such heterogeneous networks, wireless local area networks (WLANs) are integrated with wireless wide area networks (WWANs), and there exist a tight interaction between them.

The almost-ubiquitous support for IEEE802.11 WLANs (usually referred to as Wi-Fi®) makes this radio access technology a potential integrated component of near-future mobile broadband. With Wi-Fi completely integrated into mobile access, MNOs would optimize user experience and use of resources by controlling device's choice of connectivity. In addition to guaranteeing the best user experience, optimal use of access networks should care about energy-efficiency in order to extend device's battery life.

However, the performance of Wi-Fi is far from meet neither energy-efficiency nor quality of service (QoS) user's requirements. This radio access technology employs an energy-consuming medium access control (MAC) protocol that wastes both bandwidth and device's energy resources. Therefore, enhanced MAC protocols, cleverly combined with standardized power saving mechanisms such as automatic power save delivery (APSD), would improve both energy-efficiency and QoS in order to enhance WLANs performance and meet user's expectations.

In addition, current WLAN discovery mechanisms neither meet requirements of the integrated scenario. Handover operations must be improved in terms of energy efficiency and latency. Consequently, enhanced handover schemes should reduce overall device's energy consumption during the process, and enable seamless handover between Wi-Fi APs and between cellular/Wi-Fi networks.

During this thesis project, the main challenges of Wi-Fi towards its integration into mobile access broadband have been analyzed. Consequently, a solution has been designed in order to address the identified challenges, which have been introduced in the previous paragraphs. The solution consists of enhancements for IEEE 802.11 WLANs based on current standards that achieve energy-efficiency and QoS, and facilitate Wi-Fi/cellular networks interworking. Finally, a custom-designed simulator has been used to evaluate the proposed solution.

Keywords: energy-efficiency, WLAN, Wi-Fi, heterogeneous networks, seamless handover, network discovery

Sammanfattning

Globalt sett är antalet mobila bredbandsabonnemang ökar och mängden av mobil datatrafik förväntas fortsätta att växa snabbt. Under de kommande fem åren kommer antalet smartphone-abonnemang väntas mer än fördubblas, medan mängden av mobiltrafiken per aktiv prenumeration per månad för dessa abonnenter väntas nästan fyrdubbla. Som en följd av mobiloperatörer som mål att öka sin radio nätkapacitet och täckning genom heterogena distributioner. I sådana heterogena nätverk, är trådlösa lokala nätverk (WLAN) integrerad med trådlösa WAN-nätverk (WWAN), och det finns en tät interaktion mellan dem.

För att möta denna efterfrågan ämnar operatörer av mobila nätverk att öka kapacitet och täckning genom att bygga ut heterogena nätverk. I sådana heterogena nätverk integreras trådlösa lokala nätverk (WLAN) med nätverk med större yttäckning (cellulära nät) med täta interaktioner mellan de olika näten.

Det mycket utbredda stödet för IEEE 802.11-standarden (ofta kallad för Wi-Fi®) för WLAN gör denna radioaccessteknik till en potentiell integrerad komponent för mobilt bredband i den nära framtiden. Med Wi-Fi som en integrerad i det mobila accessnätet kan mobilnätsoperatörer optimera användarupplevelsen och resursanvädningen genom att styra de mobila enheternas val av uppkoppling. Förutom att garantera den bästa användarupplevelsen så bör valet av accessnät ta hänsyn till energieffektiviteten för att förlänga batteridrifttiden för den mobila enheten.

Wi-Fi är dock långt ifrån att uppfylla användarnas krav på energieffektivitet och tjänstekvalitet, eftersom denna radioaccessteknik använder ett mediumaccessprotokoll (MAC) som varken använder bandbredd eller batterienergi effektivt. Därför kan förbättrade MAC-protokoll kombinerade med standardiserade energibesparingslösningar såsom automatic power save delivery (APSD) ge bättre energieffektivitet och tjänstekvalitet, och därmed förbättra WLANs möjligheter att möta användarnas förväntningar.

Dessutom har nuvarande nätverksidentifieringsmekanismer i WLAN svårt att uppfylla kraven i ett scenario med integrerade nätverk, eftersom den nuvarande sökmetoden är långsam och använder mycket energi. En förbättrad lösning bör minska energikonsumtionen under hela processen, och möjliggöra avbrottsfri övergång mellan Wi-Fi accessnoder och mellan cellulära och Wi-Fi-nätverk.

Under detta examensarbete har de största utmaningarna för Wi-Fi under integrationen med mobil bredbandsaccess analyserats. En lösning har utvecklats för att lösa de identifierade problemen som beskrivits ovan. Lösningen består av förbättringar av IEEE 802.11 accessnät, som bygger vidare på existerande standardens energieffektivitets- och tjänstekvalitetslösningar och underlättar samverkan mellan Wi-Fi och cellulära nätverk. Slutligen har en egenutvecklad simulator använts för att utvärdera den föreslagna lösningen.

Nyckelord: energieffektivitet, WLAN, Wi-Fi, avbrottsfri handover, nätverksidentifiering, heterogena nätverk

Acknowledgements

First of all I would like to express my gratitude to Huawei Sweden R&D Center for giving me the opportunity to work in such a leader company in ICT sector. It has been a privilege to work here and I have got a great personal and working experience. I would like to thank my colleagues in Huawei and especially to my supervisor Henrik for his help and precious advises. Many thanks as well to Kari for his support.

I wish to express a special thanks to my KTH supervisor, professor Gerald Maguire Jr., who helped me to find this opportunity and has also been there every time I needed any advice and guidance.

Some special thanks as well to my home university, the technical university of Madrid (UPM), for the academic training during the first four years of my degree studies, and for giving me the opportunity to finish my degree at KTH University. Many thanks to KTH as well for this year and a half.

Finally, I would like to deeply thank my family, girlfriend, and friends for their constant affection and encouragements. They have been really important for me and without their support it would have been impossible.

Table of contents

Abstract i
Sammanfattning iii
Acknowledgements v
Table of contents vii
List of Figures xi
List of Tables xv
List of Acronyms and Abbreviations xvii
1 Introduction 1
1.1 General introduction to the area 1
1.2 Problem definition 2
1.3 Goals 3
1.4 Assumptions and limitations 3
1.5 Research Methodology 3
1.6 Structure of the thesis 4
2 Background 5
2.1 LTE 5
2.2 Wi-Fi 6
2.2.1 IEEE 802.11 Architecture 6
2.2.2 IEEE 802.11 Physical layer 7
2.2.3 IEEE 802.11 Medium Access Control (MAC) 7
2.2.4 Client Access Process: IEEE 802.11 scanning 8
2.2.4.1 Passive scanning 9
2.2.4.2 Active scanning 9
2.2.5 Infrastructure timing synchronization 10
2.2.6 IEEE 802.11 Roaming Task Groups 11
2.2.6.1 IEEE 802.11k-2008: Radio Resource Management of Wireless LAN 11
2.2.6.2 IEEE 802.11r-2008: Fast BSS Transition 11
2.3 Horizontal and Vertical Handovers 12
2.4 Enhanced scanning mechanisms 12
2.4.1 Solutions based on scanning times optimization 13
2.4.2 Selective scanning solutions 13
2.4.3 Reduction of the overhead of transmitting management messages in active scanning schemes 14
2.5 The IEEE 802.21 standard 14
2.5.1 Overview of the standard 14
2.5.2 IEEE 802.21-based solutions 17
2.6 Access Network Discovery and Selection Function 19
2.7 Wi-Fi Alliance® Hotspot 2.0 20
2.8 IEEE 802.11 Power Saving Mechanisms 21
2.8.1 IEEE 802.11 MAC Power Save Mode 21
2.8.2 Automatic Power Save Delivery (APSD) 23
2.8.2.1 Unscheduled APSD (U-APSD) 24

2.8.2.2 Scheduled APSD (S-APSD) 25
2.9 VHO decision: Energy-efficiency QoS tradeoff for RAT selection based upon traffic patterns 27
3 Method 29
3.1 Proposed solution to enhance WLANs' performance 30
3.1.1 Enhanced MAC protocol 31
3.1.2 S-APSD in depth 33
3.1.3 Seamless handover mechanism: "*Seamless-S-APSD*" 35
3.1.4 Discovery information mechanism 35
3.1.5 Neighbor report mechanism 37
3.1.6 Background selective active scanning 37
3.1.7 Overall operation example 41
3.2 *Next step* towards cellular/WLANs interworking 44
4 Analysis 47
4.1 Description of the simulator 47
4.1.1 Simple mobile wireless network simulator 47
4.1.2 Topology 48
4.1.3 Beacon generation 48
4.1.4 S-APSD traffic generation 48
4.1.5 Active scanning procedure 50
4.1.6 U-APSD traffic generation 52
4.1.7 Summary of input parameters 53
4.1.8 Simulator's outputs 54
4.2 Simulation results 57
4.2.1 Results concerning the *four-periods* version 57
4.2.1.1 Scanning operation's performance 57
4.2.1.1.1 Influence of the number of APs to be scanned 57
4.2.1.1.2 Influence of the number of MNs attached to the target APs 61
4.2.1.1.3 Influence of S-APSD traffic intensity at scanning node 63
4.2.1.2 S-APSD operations' performance 65
4.2.1.2.1 Influence of S-APSD data packet's size over the bandwidth efficiency 65
4.2.1.2.2 Throughput and energy consumption per bit achieved by S-APSD flows 66
4.2.1.3 U-APSD operations' performance 67
4.2.2 Comparison between both versions 69
4.2.3 Verification of the results and comparison with other solutions 71
4.2.3.1 Validity and evaluation of scanning results 71
4.2.3.2 Validation and evaluation of the results regarding S-APSD operations 72
5 Conclusions and Future work 77
5.1 Conclusions 77
5.2 Future work 78
5.3 Reflections 79

References 81
Appendix A Simulation output 85
Appendix B Cumulative distribution functions (CDFs) 90

List of Figures

Figure 2-1: IEEE 802.11 architecture ... 7
Figure 2-2: IEEE 802.11 passive scan scheme ... 9
Figure 2-3: IEEE 802.11 active scan scheme ... 10
Figure 2-4: Example of IEEE 802.21-based network model (adapted from figure 3 of [23]) ... 15
Figure 2-5: MIHF relationships (adapted from figure 5 of [23]) ... 16
Figure 2-6: IEEE 802.21-based network-initiated handover operation ... 17
Figure 2-7: ANDSF-based network architecture ... 19
Figure 2-8: Power saving mode operation ... 22
Figure 2-9: Unscheduled APSD operations (adapted from figure 1 of [35]) ... 24
Figure 2-10: Scheduled-APSD operation (adapted from figure 1 of [35]) ... 26
Figure 3-1: Example of network architecture for a single ESS scenario ... 30
Figure 3-2: During the CFP, all STAs must access the channel using S-APSD ... 31
Figure 3-3: Every active SP must be reallocated in order to use the gap left by the deleted Traffic Stream ... 32
Figure 3-4: Allocation of S-APSD flows using four CFPs-CPs every beacon interval ... 32
Figure 3-5: Frame format of TSPEC elements (adapted from figure 46ti of [34]) ... 34
Figure 3-6: Frame format of Schedule elements (adapted from figure 46tt of [34]) ... 34
Figure 3-7: Neighboring APs: An AP is considered a neighbor when it is within the coverage of any associated station, i.e. those inside the big (gray shaded) circle in the figure ... 37
Figure 3-8: Active scanning operation ... 39
Figure 3-9: Overall operation scenario ... 41
Figure 3-10: Overall operation example ... 43
Figure 3-11: IEEE 802.21-based heterogeneous network ... 44
Figure 4-1: Illustration of the grouping algorithm's output shown in Table 4-2 ... 49
Figure 4-2: Operation within S-APSD SPs ... 50
Figure 4-3: Contention and IFS regarding a single unicast scanning attempt ... 52
Figure 4-4: Contention and IFS associated to U-APSD traffic ... 53
Figure 4-5: Example of a MN's power states graphical output from the simulator ... 55
Figure 4-6: Power states graphic with horizontal zoom applied ... 55
Figure 4-7: Influence of the number of APs to be scanned over the scanning latency (The average scanning latency in ms is shown above each set of three columns, i.e. for a given number of APs to be scanned.) ... 58
Figure 4-8: Influence of the number of APs to be scanned over the energy consumption (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of APs to be scanned.) ... 58

Figure 4-9: Handover operations' timing concerning one single unicast active scanning, plus the handover execution ... 59
Figure 4-10: Influence of the number of APs to be scanned over the scanning latency, when half of the APs are out of the scanning node's coverage (The average scanning latency in ms is shown above each set of three columns, i.e. for a given number of APs to be scanned.) ... 60
Figure 4-11: Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's range. (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of APs to be scanned.) ... 60
Figure 4-12: Influence of the number of MNs associated with the target AP on the scanning MN's energy consumption (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given total number of MNs per AP.) ... 61
Figure 4-13: Influence of the number of MNs associated with the target AP on the listening times associated to scanning operations (The average listening time in ms is shown above each set of three columns, i.e. for a given number of MNs per AP.) ... 62
Figure 4-14: Influence of the number of MNs associated with the target AP on the scanning MN's energy consumption when frames associated with scanning operations use the same CW as other frames (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of MNs per AP.) ... 62
Figure 4-15: Influence of the number of MNs associated with the target AP on the listening times associated to scanning operations when frames associated to scanning operations use the same CW as other frames. (The average listening time in ms is shown above each set of three columns, i.e. for a given number of MNs per AP.) ... 63
Figure 4-16: Probe success rate over different S-APSD traffic conditions (The average probe success rate is shown above each set of three columns, i.e. for a given number of flows.) ... 64
Figure 4-17: Energy consumption of the scanning node as a function of the number of active flows (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of flows.) ... 64
Figure 4-18: Flow's bandwidth efficiency concept ... 65
Figure 4-19: Flows' bandwidth efficiency as a function of data packets' size (The average CFP bandwidth efficiency is shown above each set of three columns, i.e. for a given data packet size.) ... 66
Figure 4-20: Throughput achieved by a flow as a function of its TXOP ... 67
Figure 4-21: Mobile node's energy consumption (in nJ) per bit while it has one S-APSD active flow as a function of the flow's TXOP ... 67
Figure 4-22: Efficiency within CPs as a function of the number of MNs associated with an AP ... 68
Figure 4-23: MN's energy consumption per bit (in nJ) as a function of the number of MNs associated with an AP ... 68

Figure 4-24: Time that a MN spends in listening state while using U-APSD, as a function of the number of MNs associated with an AP 69
Figure 4-25: Efficiency within CPs as a function of unscheduled SPs' length.................. 69
Figure 4-26: Scanning latency comparison between both versions of the proposed solution (The average scanning latency for each version is shown above each set of three columns, i.e. for a given number of APs to be scanned. The larger value in all cases is for the one-period version.)........... 70
Figure 4-27: Comparison between MNs' energy consumption per bit as a function of the number of mobile nodes of both versions 70
Figure 4-28: Scanning phase latency in ms for various scanning schemes (Our proposed solution's performance is shown in the right most pair of columns.) 71
Figure 4-29: MNs' energy consumption in mJ during the scanning phase for various scanning schemes (Our proposed solution's performance is shown in the right most pair of columns.).. 72
Figure 4-30 Average energy consumption (in J) for a single STA as a function of the class of traffic carried for two different S-APSD schemes 73
Figure 4-31: Average power consumption (in Watts) for a single STA as a function of the class of traffic carried for two different S-APSD schemes 74
Figure 4-32: Energy consumption per bit (in nJ) for a single STA using U-APSD and S-APSD respectively 75

Appendix Figure B-1: CDF correspondent to Figure 4 7 (Influence of the number of APs to be scanned over the scanning latency) 90
Appendix Figure B-2: CDF correspondent to Figure 4 11 (Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's coverage) 90
Appendix Figure B-3: CDF correspondent to Figure 4 10 (Influence of the number of APs to be scanned over the scanning latency, when half of the APs are out of the scanning node's coverage) 91
Appendix Figure B-4: CDF correspondent to Figure 4 11 (Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's coverage) 91
Appendix Figure B-5: CDF correspondent to Figure 4 12 (Influence of the number of MNs attached to the target APs over the scanning MN's energy consumption) 92
Appendix Figure B-6: CDF correspondent to Figure 4 13 (Influence of the number of MNs attached to the target APs over the listening times associated to scanning operations).. 92
Appendix Figure B-7: CDF correspondent to Figure 4 14 (Influence of the number of MNs attached to the target APs over the scanning MN's energy consumption when frames associated to scanning operations use the same CW as the other frames).. 93
Appendix Figure B-8: CDF correspondent to Figure 4 15 (Influence of the number of MNs attached to the target APs over the listening times associated to

scanning operations when frames associated to scanning operations use the same CW as the other frames) 93
Appendix Figure B-9: CDF correspondent to Figure 4 16 (Probe success rate over different S-APSD traffic conditions)... 94
Appendix Figure B-10: CDF correspondent to Figure 4 17 (Energy consumption of the scanning node as a function of the number of active flows) 94
Appendix Figure B-11: CDF correspondent to Figure 4 19 (Flows' bandwidth efficiency as a function of data packets' size).. 95
Appendix Figure B-12: CDF correspondent to Figure 4 21 (Mobile node's energy consumption per bit while it has one S-APSD active flow as a function of flow's TXOP) 95
Appendix Figure B-13: CDF correspondent to Figure 4 22.. 96
Appendix Figure B-14: CDF correspondent to Figure 4 23 (MN's energy consumption per bit as a function of the number of MNs associated to an AP).. 96
Appendix Figure B-15: CDF correspondent to Figure 4 24 (Time that a MN spends in listening state while using U-APSD, as a function of the number of MNs associated to an AP) 97
Appendix Figure B-16: CDF correspondent to Figure 4 25 (CPs' efficiency as a function of unscheduled SPs' length)... 97

List of Tables

Table 2-1: IEEE 802.11 physical layer specifications 7
Table 3-1: Example of data that could be collected by the proposed discovery information mechanism 36
Table 3-2: The discovery information that MN1 has at time t1 42
Table 4-1: Random input that was given to the grouping allocation algorithm 49
Table 4-2: Grouping algorithm's output 49
Table 4-3: Simulation input parameters 53
Table 4-4: Devices' power consumption in the different power states 56
Table 4-5: Traffic characteristics 73

List of Acronyms and Abbreviations

3G	third generation
3GPP	third generation partnership project
4G	fourth generation
AC	access category
ACPI	Advanced Configuration and Power Interface
ADDTS	*add traffic stream*
AM	active mode
ANDSF	access network discovery and selection function
AP	access point
APSD	automatic power save delivery
BS	base station
BSS	basic service set
BSSID	BSS identifier
CFP	Contention free period
CID	Channel information based discovery
CLID	Channel and location information based discovery
CP	Contention period
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
DCF	distributed coordination function
DIFS	DCF interframe space
DS	distribution system
EDCA	enhanced distributed channel access
EOSP	end of service period
ESS	extended service set
FA	foreign agent
GAS	Generic Advertisement Service
GPS	global positioning system
GSM	Global System for Mobile Communications
H-ANDSF	Home ANDSF
HCCA	HCF controlled channel access
HCF	hybrid coordination function
HHO	Horizontal handover
HO	handover
HS2.0	Hotspot 2.0
ID	Identifier
IEEE	Institute of Electrical and Electronics Engineers
IFS	Inter frame space
IP	Internet protocol
ISMP	Inter system mobility policy
ISRP	Inter system routing policy

LID	Location information based discovery
LTE	Long term evolution
MAC	media access and control
MIH	media independent handover
MIHF	MIH functions
MIIS	MIH information server
MN	mobile node
MNO	mobile network operator
PCF	point coordination function
PLMN	public land mobile network
PoA	point of attachment
PSM	power saving mode
PS-poll	power-save poll
QoS	quality of service
RAN	radio access network
RAT	radio access technology
S-APSD	scheduled APSD
SE	schedule element
SI	service interval
SLA	service-level agreement
SP	service period
SSID	service set identifier
SST	service start time
STA	station
TIM	traffic information map
TS	Traffic Stream
TSF	the timing synchronization function
TSPEC	traffic specification
TXOP	transmission opportunity
U-APSD	unscheduled APSD
UE	User Equipment
V-ANDSF	visited ANDSF
VHO	Vertical Handover
Wi-Fi™	Wireless Fidelity
WLAN	Wireless Local Area Networks

1 Introduction

This chapter provides an introduction to the problem area and states the specific problem that this thesis project addresses. The chapter also describes the project's goals and gives an overview of the structure of the thesis.

1.1 General introduction to the area

Total mobile subscriptions are at around 6.6 billion nowadays, and they are expected to reach around 9.3 billion by 2019 [1]. Moreover, the majority of mobile broadband devices are smartphones. The total smartphone subscriptions are around 1.9 billion these days, and are expected to grow to 5.6 billion by 2019. Regarding mobile technologies, despite Global System for Mobile communications (GSM) and Enhanced Data Rates for GSM Evolution (EDGE) representing the largest share of mobile subscriptions today, Long Term Evolution (LTE), with 150 million subscriptions, is growing rapidly and is expected to reach 2.6 billion subscriptions by 2019. Concerning mobile traffic, while voice traffic (excluding Voice over IP) remained stable throughout the last years, mobile data traffic per active smartphone subscription per month is expected to grow from 600MB to 2.2GB between 2014 and 2019.

The next generation of mobile communication networks is based upon heterogeneous wireless access network environments. This is the mobile broadband strategy of most mobile network operators (MNOs) in order to achieve additional radio network capacity and coverage. For instance, the development of fourth generation (4G) mobile broadband systems based on the Third Generation Partnership Project (3GPP) LTE includes enhanced support for heterogeneous deployments where wireless local area networks (WLANs) are tightly integrated with wireless wide area networks (WWANs) [2].

Therefore, WLANs, particularly IEEE 802.11 WLANs (commonly referred to as Wi-Fi®), are expected to play an important role in future wireless communication networks. The main factor that gives Wi-Fi technology such an important role is the almost ubiquitous support for this this technology in modern devices. The Wi-Fi Alliance® expects the number of IEEE 802.11-enabled devices to surpass 2 billion by 2015 [3], and 97.5 percent of smartphones currently support Wi-Fi [4]. In addition, other key factors are Wi-Fi's ability to offload traffic and the evolution of smart cells supporting both cellular and Wi-Fi access network technologies.

On the other hand, Wi-Fi must be fully integrated into mobile access networks to allow MNOs to control the Wi-Fi/cellular handoff process in order to guarantee their users the best user experience and to optimize the use of resources. This concept is referred to as *carrier Wi-Fi* in [5]. The decision about the choice of connectivity would be made according to much more complex criteria than the currently used algorithm, which is basically to use Wi-Fi if it is available. Moreover, the end goal of mobile operators is to achieve a user experience as convenient, responsive, and seamless as possible. Therefore, one of the main challenges is to support seamless handovers between Wi-Fi and cellular networks and between Wi-Fi access points (APs), to have seamless vertical and horizontal handovers.

There are some integration tools that aim to facilitate the integration of Wi-Fi, such as Hotspot 2.0 (HS2.0), the IEEE 802.21 standard, and the access network discovery and selection function (ANDSF). The Hotspot 2.0 program [6] focuses on enabling mobile devices to automatically discover and associate with WLAN APs. This program is one example of MNOs and service providers attempting to increase the densification of their networks. The IEEE 802.21 standard aims to provide a framework that allows interworking within several

IEEE 802 systems and between non-IEEE 802 and IEEE 802 systems. The standard defines media independent handover (MIH) functions that enable the optimization of handover in heterogeneous networks. A similar framework has been defined by 3GPP with ANDSF, which provides network discovery and assistance to user devices.

Optimal selection of an access network would reduce devices' energy consumption, as was analyzed by Jose María in his thesis *Energy-Efficient Vertical Handovers* [7]. Nevertheless, current IEEE 802.11 WLANs do **not** meet either the energy-efficiency or quality-of-service (QoS) demands of mobile users. The main challenges that must be addressed to improve the performance of Wi-Fi are:

- Enhance the medium access control (MAC) protocol in order to reduce stations' energy consumption and increase bandwidth efficiency,
- Support QoS,
- Improve WLAN discovery operations to reduce the power consumption of the scanning process and to allow seamless handovers between APs, and
- Facilitate Wi-Fi/cellular networks interworking.

These challenges are further explained in the following section.

1.2 Problem definition

The use of Wi-Fi technology by smartphones and other mobile devices is expected to increase dramatically due to the fact that this technology will be an integrated component of future mobile broadband access networks. Therefore, IEEE 802.11 WLANs' operation must be optimized in order to meet a number of service requirements. The main challenges that must be addressed are described in the next paragraphs.

WLANs employ a contention-based MAC protocol to manage user access to the shared medium. IEEE 802.11 uses a MAC protocol called carrier sense multiple access with collision avoidance (CSMA/CA). CSMA/CA is considered an energy consuming MAC protocol that limits WLANs' performance [8]. Stations using CSMA/CA access mechanisms, such as the distributed coordination function (DCF), point coordination function (PCF), or the latest enhanced distributed channel access (EDCA), and hybrid coordination function (HCF) controlled access (HCCA) mechanisms, must listen to the WLAN channel in order to contend for channel access. This fact, in addition to other MAC overheads such as inter-frame spaces (IFSs) and acknowledgments (ACKs), wastes both bandwidth and devices' energy resources.

Unfortunately, the majority of available solutions that aim to reduce energy consumption sacrifice the user's QoS. Such solutions seek to minimize the period that a WLAN interface stays awake. The IEEE 802.11 standard [9] first introduced a power saving mechanism called power saving mode (PSM). Afterwards, another mechanism called automatic power save delivery (APSD) was proposed in order to achieve energy-efficiency while taking care of QoS.

Consequently, enhanced MAC solutions based on coordinated channel access, cleverly combined with standardized power saving mechanisms (such as APSD) would reduce devices' energy-consumption *while* maintaining users' QoS due to efficient use of bandwidth resources.

In addition, WLANs discovery mechanisms must be improved in order to meet the requirements of the integrated scenario. Enhanced discovery mechanisms should enable seamless handover between APs, preserve QoS after handovers are performed, and reduce overall energy consumption of stations during this process. The handover process can be divided into three phases: network discovery and selection, re-authentication, and re-association. The first phase, also known as scanning phase, accounts for almost 90% of the handover's total duration [10], and also consumes a considerable amount of energy. Therefore, enhanced discovery mechanisms should focus on improving the scanning phase.

Finally, WLANs APs must facilitate Wi-Fi integration by collecting relevant information about available resources (e.g. other nearby APs) and network conditions, and deliver this information to the operator. Thus, with the operator in control of Wi-Fi/cellular switching process the connectivity decision can be based on more complex criteria.

1.3 Goals

The start point of this master thesis project was a research study about the state of the art of current mobile communications. This study was done in order to identify the main challenges of Wi-Fi technology with respect to its integration into a mobile broadband access network. Those challenges were presented in the previous two sections and will be further addressed later in this thesis.

The primary goal was to design a solution to address the identified challenges. This solution needed to incorporate enhancements based on those current standards that achieve energy-efficiency and QoS, while facilitating Wi-Fi/cellular networks interworking.

Finally, the performance of the proposed solution was to be evaluated. This evaluation has been done via a custom-designed simulation framework.

1.4 Assumptions and limitations

A decision algorithm that controls stations' behavior concerning whether to switch or not to the Wi-Fi access network or to switch back to cellular access network to optimize both user experience and resource utilization is outside the scope of this thesis. We will simply assume that there exists a suitable access network selection mechanism, which makes decisions about when to perform a handover and which access network to use at any moment.

Moreover, both security issues and operator policies are not addressed. Both of these areas require their own careful treatment. As a result this thesis focuses only on enhanced mechanisms to achieve energy-efficiency and QoS.

1.5 Research Methodology

The research question was: *what are the main challenges of Wi-Fi technology with respect to its integration into mobile broadband access network?*

This project has adopted *evaluation research* method because the main goal was to design and evaluate a solution that addresses those challenges. The main phases that this project went through are detailed below.

i. The initial research phase was based on *document analysis* and *existing data* collection methods. It consisted of an extensive *literature review* in order to identify the challenges that we were looking for. Those identified challenges constitute the hypothesis of the research.

ii. Afterwards, it was the design phase that consisted of the design of a complete solution that addresses the identified challenges.

iii. Finally, the analysis phase was based on a *simulation-based analysis* method in order to get quantitative results regarding the performance of the proposed solution. The results derived from the simulations were compared with results of other solutions in order to evaluate the proposed solution. As a result, the initial hypothesis (i.e. the identified challenges) was verified: *by addressing the identified challenges, Wi-Fi performance will be improved in terms of energy-efficiency, QoS support, and optimal use of resources in order to facilitate its integration into mobile access*.

Other methodologies that were considered are in-depth interviews, in order to identify users' expectations, mathematical-based analysis of the proposed solution. The former was rejected because I considered that this information could be easily found in the research community, and the later was rejected because we considered that a simulation-based analysis would provide more accurate results.

1.6 Structure of the thesis

Following this initial chapter the thesis continues with an extensive literature study of relevant background material (in Chapter 2) that the reader will find useful to understand the rest of this thesis. The literature study provides the basis for the proposed solution's design, such as the IEEE 802.11 Standard, enhanced scanning solutions, and integration tools (e.g. the IEEE 802.21 Standard, ANDSF, and HS2.0). Chapter 3 describes the solution that has been designed to address the challenges stated in section 1.2, and a brief description of a proposed IEEE 802.21-based heterogeneous network architecture where IEEE 802.11 WLANs use the proposed solution. Chapter 4 describes the custom-designed simulator used to evaluate the proposed solution and analyzes the results of the simulations that were conducted. In addition, the performance of the proposed solution is compared with other solutions found in the research literature. Chapter 4 concludes the thesis and suggests some future work. There are also two appendices which contain an example of the simulator's output and the cumulative distribution function of the results presented in section 4.2.

2 Background

Current wireless devices usually have multiple radio interfaces in order to support multiple radio access technologies (RATs). These radio interfaces allow the device to utilize different radio access networks (RANs). In the following subsections, the most relevant RANs for this thesis project will be introduced. These RANs are 3GPP's Long Term Evolution (LTE) and IEEE 802.11 WLANs (hereafter referred to by its trade name Wi-Fi®). We will place special emphasis on the IEEE 802.11 standard [9].

2.1 LTE

The fourth generation mobile broadband systems that this thesis project will consider are based on 3GPP long term evolution (LTE) RAT. LTE is the latest standard in 3GPP's evolution from the earlier Global System for Mobile Communications (GSM) and the Universal Mobile Telecommunications System (UMTS) standards.

This section gives a high-level overview of LTE, focusing on its evolution via different releases and the likely next major steps. For further technical details, the reader is referred to [11].

LTE is a highly flexible platform which is continuously evolving to address new requirements and potential future scenarios. Current commercial LTE deployments are based on 3GPP Releases 8 and 9. These are the first 3GPP releases to include LTE. The main features of current deployments are:

- A physical layer using orthogonal frequency-division multiplexing (OFDM) with a cyclic prefix for the downlink multiple access scheme; while a single-carrier frequency-division multiple access (SC-FDMA) scheme is used for the uplink scheme, in order to enable power-efficient transmission by mobile devices;
- High peak data rates of up to 300 Mb/s on the downlink and 75 Mb/s on the uplink;
- Support for multiple-antennas; and
- Flexible bandwidth - allowing operation with up to 20 MHz bandwidths, in order to allow operators to operate with many different sized spectrum allocations.

The next step in the evolution of LTE is included in 3GPP Releases 10 and 11; and is referred to as LTE-Advanced. LTE-Advanced enhances the system allowing it to meet the International Mobile Telecommunications-Advanced (IMT-A) requirements for a 4G system [12]. For that purpose, LTE-Advanced introduces some additional techniques in order to improve spectrum flexibility:

- Carrier aggregation;
- Enhanced multi-antenna transmissions, based on an extended and more flexible reference signal structure;
- Basic functionality for coordinated multipoint (CoMP) transmission and reception; and
- Enhanced support for heterogeneous networks, which are based on the presence of complementary low-power nodes, physically closer to the terminals, distributed within the coverage area of an existing macrocellular network. These are referred to as local area wireless access. The types of low-power nodes include microcells, picocells, femtocells, and relays.

The future Release 12, whose main ideas are introduced in [2], is supposed to include further support for enhanced heterogeneous networks, by tightening interaction between the wide area and the local access networks. This interaction between wide area and local area access networks may be based on the soft cell concept, which is a combination of dual connectivity with ultra-lean transmissions:

Dual connectivity	Dual connectivity refers to the fact that a mobile terminal is connected to two network nodes, for instance, a cellular base station (BS) and a Wi-Fi AP. Dual connectivity enables throughput aggregation. Dual connectivity also provides other benefits, such as uplink-downlink separation and increased robustness to mobility. Uplink-downlink separation uses the node with strongest received signal for downlink transmissions, while using the node with the lowest path lost for the uplink. An example of mobility robustness would be a scenario where the wide area access is used as an anchor connection for signaling and low data rate transmissions, while the local access network is used as an *assisting connection* to carry large amounts of high data rate user traffic.
Ultra-lean transmissions	Ultra-lean transmissions impose the minimum possible amount of overhead on the *assisting connection*. For instance, the assisting connection transmission would carry only data frames, as the overhead is reduced by using the anchor connection to carry management frames such as acknowledgements.

Finally, it is worth noting that IEEE 802.11 WLANs can provide the local wireless access network. These types of networks are already used by many operators for offloading purposes. 3GPP Release 12 will include signaling mechanisms to support network control of the RAT used by the terminal. With respect to the previously mentioned scenario for mobility robustness, the WLAN interface could be used as an assisting connection, while LTE provides an anchor connection.

2.2 Wi-Fi

The Wi-Fi Alliance® [13] defines Wi-Fi as "wireless local area network (WLAN) products that are based on the Institute of Electrical and Electronics Engineers (IEEE) 802.11 standards". The following subsections describe the essential aspects of IEEE 802.11 infrastructure networks in relation to the scope of this master’s thesis project. All of this material is derived from the IEEE 802.11-2012 standard [9].

2.2.1 IEEE 802.11 Architecture

The basic building block of IEEE 802.11 WLANs is the basic service set (BSS), which consists of some number of stations (STAs) executing the same MAC protocol and contending for access to the same shared wireless medium. The most basic type of BSS is the Independent BSS (IBSS) which may consist of only two STAs. However, the most common type of BSS is the infrastructure BSS, where client stations do not communicate directly but rather communicate through an AP. A BSS may be isolated or may be connected to a backbone distribution system (DS) through the AP.

Infrastructure BSSs may also be part of an extended network composed of multiple BSSs. This network is referred to as extended service set (ESS) consists of two or more BSSs interconnected by a DS, as shown in Figure 2-1. Each BSS is identified by its BSS identifier (BSSID), which corresponds to the MAC address of the AP. The ESS is identified by its service set identifier (SSID), which is advertised by all the APs within the ESS in beacon frames.

Note that IBSS and mesh BSS (MBBS) will not be addressed since they are not relevant for this thesis project.

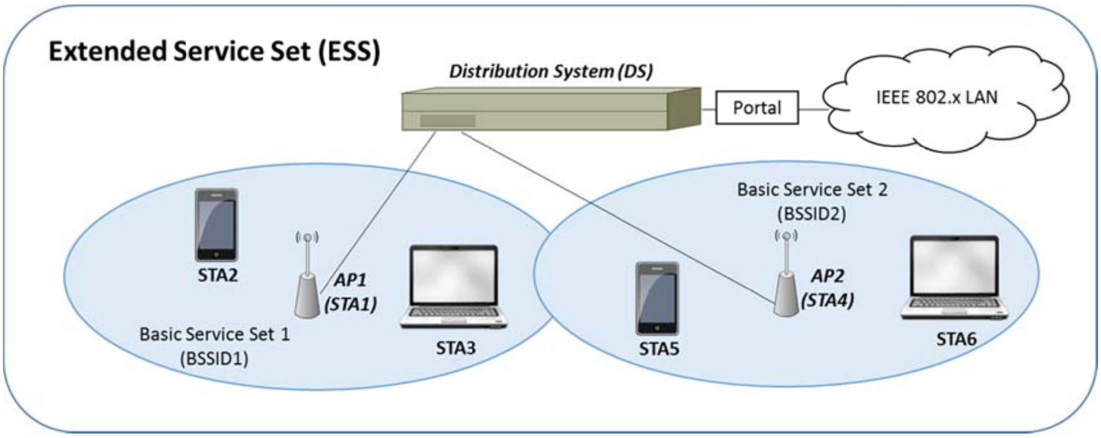


Figure 2-1: IEEE 802.11 architecture

2.2.2 IEEE 802.11 Physical layer

The different IEEE 802.11 family standards introduced modifications regarding operating bands, channel assignments, spread-spectrum technology, etc. The main characteristics of the physical layer of those standards are shown in Table 2-1. For further information about the IEEE 802.11 physical layer, the reader is referred to Chapter 4 of [14].

Table 2-1: IEEE 802.11 physical layer specifications

Standard	Date	Frequency band (GHz)	Number of channels (non-overlapping)	Theoretical throughput (Mbps)	Available channel bandwidth (MHz)	Spread-spectrum technology
802.11a	1999	5.8	12 (12)	54	20	OFDM
802.11b	1999	2.4	14* (3)	11	5	CCK
802.11g	2003	2.4	14 (3)	54	5	CCK, OFDM
802.11n	2009	2.4 and 5	24 (3)	600	20 or 40	OFDM, MIMO

2.2.3 IEEE 802.11 Medium Access Control (MAC)

The main drawback of the IEEE 802.11 physical and link layers, as for any other wireless network, is the use of a shared medium which leads to greater impairments than in the case of cable or fiber. Due to the fact that the medium is shared, stations must contend for the channel access before initiating a transmission. Additionally, noise, interference, and other propagation effects mean unreliability and can result in the loss of a large number of frames.

The fundamental MAC technique is the distributed coordination function (DCF), defined by the standard in section 9 [9, Sec. 9]. It uses a contention-based media access-mechanism that combines carrier sense multiple access with collision avoidance (CSMA/CA) mechanism [15], with a binary exponential backoff mechanism. The basic operation is as follows: before transmission the STA must sense the channel to be idle for a short period of time, called distributed inter-frame space (DIFS). Otherwise, if the channel is sensed busy

* Not all of these channels are available to be used in a given regulatory domain.

during that period of time, the STA must persist to sense the channel until it is measured idle for a DIFS. At this point, the STA must defer its access to the channel for an extra random period called a backoff interval in order to increase robustness against collisions.

For highly congested IEEE 802.11 radio environments, the standard supports a Request to Send/Clear to Send (RTS/CTS) mechanism, which can be used by stations in order to gain access to the medium before transmitting. This mechanism consist of the exchange of short control frames between the transmitting and receiving STAs, prior to transmission of a longer frame, and after determining that the medium is idle and after any backoffs and deferrals.

The IEEE 802.11 MAC may also include an optional contention-free access mechanism called the Point Coordination Function (PCF). This is a centralized method in which a point coordinator (PC) placed at the AP of the BSS determines which STA currently has the right to transmit. When this optional mechanism is used, the PCF and the DCF access methods alternate, with a contention free period (CFP) followed by a contention period (CP).

Moreover, the carrier sensing may be performed through physical and virtual mechanisms. Through the former, STAs sense the channel physically. In contrast, in the virtual carrier sense mechanism STAs use the duration field of MAC headers to define the period of time that the medium is to be reserved for the current transmission. This duration field is also referred to as the network allocation vector (NAV), and should be used during CPs when all the STAs must listen to the channel before initiating a transmission.

The IEEE 802.11e standard introduced a hybrid coordination function (HCF), which combines DCF and PCF functions in order to support QoS guarantees. The standard defines two methods of channel access: HCF controlled channel access (HCCA) and enhanced distributed channel access (EDCA). The basic unit of allocation of the right to transmit is called a transmission opportunity (TXOP). These two methods are further described as:

HCCA Allows STAs to reserve TXOPs for several traffic streams, characterized by a traffic specification (TSPEC) element. HCCA can be used during both CFP and CP.

EDCA EDCA is used during the CP, and therefore, the TXOP may be obtained through channel contention. Moreover, QoS is achieved using differentiated user priorities varying the amount of time a STA senses the channel to be idle before transmission, size of backoff contention windows, and the maximum duration that a STA can use the channel. Differentiation is achieved by using four access categories for data frames: *AC_BE*, *AC_BK*, *AC_VO*, and *AC_VI*.

2.2.4 Client Access Process: IEEE 802.11 scanning

The client access process involves three steps: scanning, authentication, and association. These three steps are the three phases of the handover procedure, i.e., the process of roaming from one BSS to another. This thesis project will focus on the first phase, the scanning process.

A STA becomes aware of the existence of a BSS through channel scanning. The IEEE 802.11-2012 standard defines two modes of operation: passive and active scanning modes. Prior to the description of each of these modes, it is important to describe the beacon frames generated by infrastructure networks. The AP periodically broadcasts beacon frames for synchronization purposes. These beacon frames include information about the AP, specifically the maximum transmit power and which channels will be used. The inter-beacon period is fixed, but the transmission of a beacon frame may be delayed because of the CSMA/CA mechanism. However, subsequent beacons shall be broadcasted according to the original schedule.

Active scanning is not allowed in some frequency bands and regulatory domains. In these cases the MAC of a STA should check the regulatory domain information *before* performing an active scan.

The SSID parameter is present in both scanning modes. It indicates the SSID for which the device should scan. It is possible to use one or more SSID list elements or a wildcard SSID (the latter is used when the desired SSID is unknown). If passive scanning is used, a STA iteratively scans the channels for a beacon frame, while in active scanning the STA searches for a SSID to which it wishes to associate by sending Probe request frames for this SSID.

2.2.4.1 Passive scanning

The STA passively listens to each channel. Moreover, the STA might not scan all the channels, but could decide upon a list of channels to scan. However, it is necessary to wait on each channel for a period of time longer than the inter-beacon period of the APs, because the STA can only receive a beacon frame if it is listening on the channel when the beacon frame is sent. The passive scanning procedure is shown in Figure 2-2.

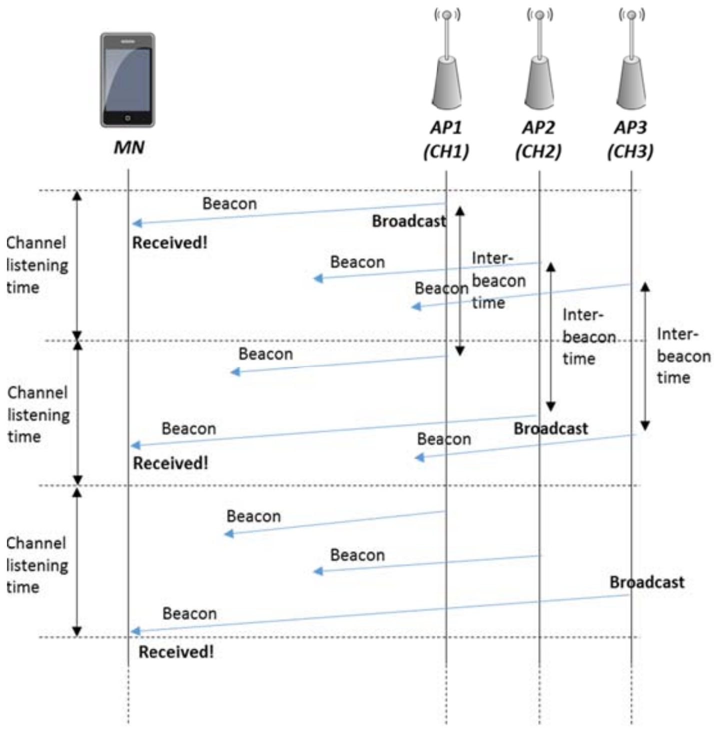


Figure 2-2: IEEE 802.11 passive scan scheme

2.2.4.2 Active scanning

In the active scanning mode, the STA actively sends Probe request frames on each channel specifying the desired SSID that the STA is looking for. First, the STA tunes its radio interface to the appropriate channel, listens in order to avoid a collision, and sends a Probe request frame. Then, the STA listens to that channel for a defined period of time in order to receive a Probe response or a beacon frame from the AP. Probe request frames may include requests for other elements, such as the received power indicator (RCPI) value.

The APs send a Probe response directly to the STA that generated the request when the AP receives a Probe request in which the SSID field contains the wildcard SSID, the SSID of the AP, or the AP's SSID is included in the SSID List element.

The active scanning procedure is shown Figure 2-3. As can be observed, the waiting time on each channel is irregular and controlled by two timers (in contrast with the fixed time period in the passive scan procedure). The channels are scanned one by one. The STA sends a Probe request and waits MinChannelTime for a probe response; if at least one packet is received, then the sensing interval is extended to MaxChannelTime in order to obtain more responses; in contrast, if during MinChannelTime there is no activity in the channel, then the channel is declared inactive and the STA continues by scanning the next channel.

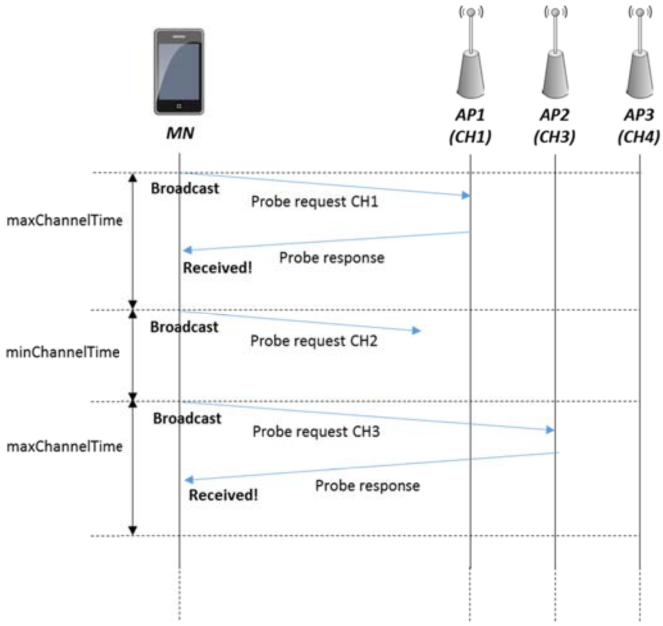


Figure 2-3: IEEE 802.11 active scan scheme

The main drawback of active scanning is the amount of overhead generated by probe request/response transmissions. Some enhancements, which aim to mitigate this problem, will be further explained in section 2.4.

After the scanning procedure has completed, the STA has learnt and recorded the channels being used and the SSIDs advertised by the available APs. In addition, it has gathered capability information from the capability fields and the information elements in the beacon/probe response frames. Now the STA can make a decision about which BSS to join. While typically it might decide to associate with the AP providing the strongest received signal, enhanced IEEE 802.11 STAs may apply other criteria such as security, network load, backhaul capacity, or even consider the network operator's policies.

Once the STA decides upon the target AP, it initiates the authentication and association phases, which are not addressed in this thesis project. For further details about these phases and alternatives to the traditional binary authentication see [16].

2.2.5 Infrastructure timing synchronization

In addition to data distribution, the AP is the central coordinator for the timing synchronization in an infrastructure BSS. It is the timing master for the timing synchronization function (TSF), and AP's timer value is included in the beacon's *timestamp* field. Stations associating with this AP must accept this timing value in order to remain synchronized with the AP - even if some beacon frames are missed. Moreover, timing values are also included in Probe Response frames to assist active scanning STAs.

2.2.6 IEEE 802.11 Roaming Task Groups

The IEEE 802.11 standard has evolved due to the activities of several IEEE 802.11 task groups. Several improvements have been introduced to the standard in order to enhance the performance of the handover process. The most relevant enhancements were codified by the IEEE 802.11k and IEEE 802.11r task groups. Details of these enhancements are described in the following paragraphs for each of the standards developed by these task groups.

2.2.6.1 IEEE 802.11k-2008: Radio Resource Management of Wireless LAN

The IEEE 802.11k task group developed standards for measuring and reporting statistics about signal strength, load levels, channel usage, etc. These radio resource measurements aim to provide a service to enable STAs to understand the radio environment in which they exist. An STA may choose to make measurements locally or request a measurement from another STA. This information can be very useful in reducing the energy consumption of an STA looking for available APs. Moreover, these measurements provide valuable information to upper layers. For example, applications, such as VoIP, require a certain minimum performance in terms of radio environment measurements in order to achieve the application's desired QoS.

Radio resource measurements, such as Channel Load request/report and the Neighbor request/report, may be used to collect transition information. This information can drastically speed up handovers between BSSs *within* and *without* the same ESS.

The most interesting measurements relevant to faster and more efficient scanning are:

Beacon	The beacon request/report pair enables a STA to request from another STA a list of APs whose beacons it can receive on a specified channel or channels.
Measurement Pilot	The measurement pilot frame provides a subset of the information provided in a beacon frame. It is smaller than a beacon frame and its main purpose is to assist a STA with scanning.
Channel load	The channel load request/report pair returns the channel utilization measurement as observed by the measuring STA.
Noise histogram	The noise histogram request/report pair returns a power histogram measurement of non-IEEE 802.11 noise power.
STA statistics	The STA statistics request/report pair returns groups of values for STA counters and for BSS Average Access Delay.
Location	The location request/report pair returns a requested location in terms of latitude, longitude, and altitude.
Neighbor report	The neighbor report request is sent to an AP which returns a neighbor report containing information about known neighboring APs. This measurement enables a STA to collect information about the neighbors of the current AP to which it is associated and these neighbors can be used as potential roaming candidates in order to perform selective scanning.

2.2.6.2 IEEE 802.11r-2008: Fast BSS Transition

The IEEE 802.11r standard attempts to combine the roaming-work of other IEEE task groups, specifically: 802.11i (security), 802.11e (QoS provisioning), and 802.11k (measurement and reporting).

IEEE 802.11r attempts to define the procedures and protocols required to enable a faster, secure, and QoS-sensitive BSS transition. This standard does **not** consider issues such as detecting the need to roam or the discovery and selection of a target AP. However, these issues are very relevant in context of this thesis project; hence these issues will be addressed in the next chapter.

The main goal of the IEEE 802.11r standard is to reduce the security overhead during BSS transitions. For that purpose, this standard allows a dialog between a STA and the target AP (i.e. potential handover candidate) *prior* to association, i.e., while the STA is associated with another AP. Two mechanisms are defined to enable such a dialog: *over-the-air* (OTA), where the STA and the target AP communicate directly, and *over-the-DS* (ODS), where the STA and target AP communicate through the DS.

2.3 Horizontal and Vertical Handovers

The process in which a mobile device switches from one point of attachment (PoA) to another PoA is known as a handover. The reasons why a handover might be performed can be diverse, for example, a STA moves from one coverage area to another or the currently offered QoS is insufficient for an application's requirements.

In this thesis project we are primarily concerned with two different types of handovers: a Horizontal Handover (HHO) occurs when the STA switches between different network APs using the same RAT; while a Vertical Handover (VHO) involves two different RATs, for example, a handover from LTE to IEEE 802.11.

The first step in the handover process is network discovery, which implies that the mobile device makes measurements in order to discover available access networks. However, this discovery process can cause unnecessary power consumption. This scanning process was described for Wi-Fi networks in section 2.2.4. One of the goals of this thesis project is to suggest improvements regarding the WLAN discovery process in order to achieve energy-efficiency.

In addition to HHO and VHO there are two different types of handovers with regard to the hardness/softness of the transition: a hard handover implies an abrupt break in the path that the user signaling and data takes and typically introduces longer communication latencies; while a soft handover utilizes both the current and future AP to receive user traffic, enabling a seamless transition.

2.4 Enhanced scanning mechanisms

This section reviews the state of the art concerning enhanced scanning mechanisms. Several different solutions will be explained, starting from timing based solutions and moving to solutions with greater interaction between the cellular and Wi-Fi networks.

Rebai and Hanafi [10] have stated that the scanning phase is the most costly of the three handover phases in terms of duration and power consumption. Checking for available APs consumes a considerable amount of energy, and this energy cost is directly related to the duration of this process. In addition, handover latency plays a key role in multimedia applications' QoS requirements.

Enhanced scanning solutions can be classified into three groups: the first group of algorithms aim to optimize scanning operations by reducing channel scanning times; meanwhile, the second group of scanning solutions are referred to as selective scanning schemes and they aim to reduce the number channels to be scanned based on information about the access points in the vicinity; finally, a third class of methods attempts to reduce the measurements made by the device by using information from other sources. The following subsections describe some of the scanning approaches found in the research community. Each of them has been grouped into one of the three classes defined here.

2.4.1 Solutions based on scanning times optimization

This group of algorithms comprise those solutions that strive to reduce the time spent scanning each channel. The most interesting solution is SyncScan algorithm, proposed by Ishwar Ramani, and Stefan Savage in [17]. It is a passive scanning approach based on the synchronization of beacons between APs working on different channels. Non-AP STAs passively monitors other channels in the background for the presence of nearby APs. The absence delay of the non-AP STA with its current AP (i.e. the period of time during which the non-AP STA is not listening the channel of its associated AP) is minimized by using short listening periods while scanning other channels. The use of such short listening periods is possible due to the fact that non-AP STAs synchronize themselves with the timing of beacon broadcasts on each channel. Therefore, a non-AP STA passively scans other channels by switching channels exactly when a beacon is about to arrive. However, when a non-AP STA examines other channels it cannot communicate with its associated AP, so that it may miss packets during scanning periods.

2.4.2 Selective scanning solutions

The second group of algorithms improves scanning performance by using information about potential handover candidates, i.e. APs located close to the MN. Then, once a non-AP STA is aware of its adjacent APs, it can use selective scanning to avoid scanning all channels. However, different solutions propose different ways to collect such information. Sometimes, information about nearby APs is provided by servers, either through the currently attached AP or using another radio access network. Meanwhile, other proposals are based on the idea of cooperation between APs, where APs exchange information between themselves.

In [18], Huang et al. propose a selective scan technique that combines an enhanced neighbor graph scheme and an enhanced inter access point protocol (IAPP) scheme. The neighbor graph (NG) is an undirected graph with each edge representing a mobility path between vertices, which represent the APs. Moreover, IAPP allows an AP to communicate with other APs within a distribution system in order to exchange information about their associated non-AP STAs. This information exchange enables fast authentication and association between non-AP STAs and APs. The NG is provided by a NG Server located in the IEEE 802.11 network. Furthermore, the NG information can be set manually or can be collected from non-AP STAs' history (of associations).

A similar solution is proposed by Tuysuz et al. in [19]. In this paper they aim to reduce the energy consumption of mobile stations by enhancing handover operations. Selective channel scanning is performed in the background making use of an IEEE 802.21 information server (IS). The IS provides as much information as possible to the non-AP STAs; for example, the available APs and their coordinates. More information about the IEEE 802.21 standard is given in section 2.5. Based on this information, STAs first actively scan for the AP at the closest distance. Furthermore, during the handover process, non-AP STAs associate with an AP that is expected to allow them to consume the least energy among the entire set of candidate APs.

Finally, the Athanasiou, Korakis, and Tassiulas [20] introduce the idea of cooperation between APs. APs exchange neighbor statistics between them, and this information is provided to non-AP STAs in management frames (such as the beacons). The cooperation between neighbors takes advantage of the IEEE 802.11k framework, so that these neighbors interchange beacon and neighbor report frames. The IEEE 802.11k framework was described in section 2.2.6.1. The information provided allows non-AP STAs to avoid sequential channel scanning and AP probing.

2.4.3 Reduction of the overhead of transmitting management messages in active scanning schemes

The first solution in this class that was evaluated was proposed by Chang, et al. in [21]. They proposed two enhancements to the active scanning mechanism based on reducing the *unnecessary* transmission of management messages. Both enhancements were based on the draft IEEE 802.11ai standard, which defines a fast initial link setup (FILS) in order to reduce the time that STAs spend in the initial phase. This idea is further explained by Ong in [22]. The proposed improvements aim to decrease the number of probe request/response exchange frames. The first scheme establishes two rules for an AP when receiving a probe request:

- APs shall **not** respond - if there is already a probe response in the transmission buffer.
- APs shall **not** respond - if the AP is included in the request's filter list. This filter list contains the APs from whom the STA has already received a probe response.

The second scheme introduces an interesting improvement: APs **shall** include information of all neighbor APs in their probe response(s). Thus, if an AP receives a probe *response* from another AP, it will flush all of its own probe response messages from its own buffer as these only contain information that is already in the received probe response.

Simulation studies by Chang, et al. demonstrate that the proposed algorithms are generally more efficient than existing schemes. The main conclusion is that the proposed interaction and cooperation between neighboring APs could lead to an enhanced AP detection process when using an active scanning scheme. However, the proposed algorithms imply non-trivial modifications to the existing control messages exchange which may be difficult to implement in practice.

The following section introduces the IEEE 802.21 standard, which defines MIH mechanisms to optimize and facilitate handovers in *heterogeneous* networks.

2.5 The IEEE 802.21 standard

This section consist of two subsections: the former introduces the IEEE 802.21 standard; while the later section presents some interesting solutions proposed in the research community that make use of this standard to improve the network discovery process.

2.5.1 Overview of the standard

The IEEE 802.21 standard [23] provides a framework that allows a MN to achieve seamless connectivity while switching between different RATs. It makes this possible by making cooperative use of the information available at the MN (which is well placed to detect available networks) and at the network entities (which are suitable for storing and sharing useful information).

Although cellular and Wi-Fi networks already have mechanisms for learning a list of neighboring BSs or APs (respectively), the IEEE 802.21 standard provides a unified mechanism that can be used by higher layer entities to provide handover-related information in a heterogeneous network environment at a given geographical location. Figure 2-4 illustrates a potential IEEE 802.21-based network model that enables Wi-Fi/cellular interworking.

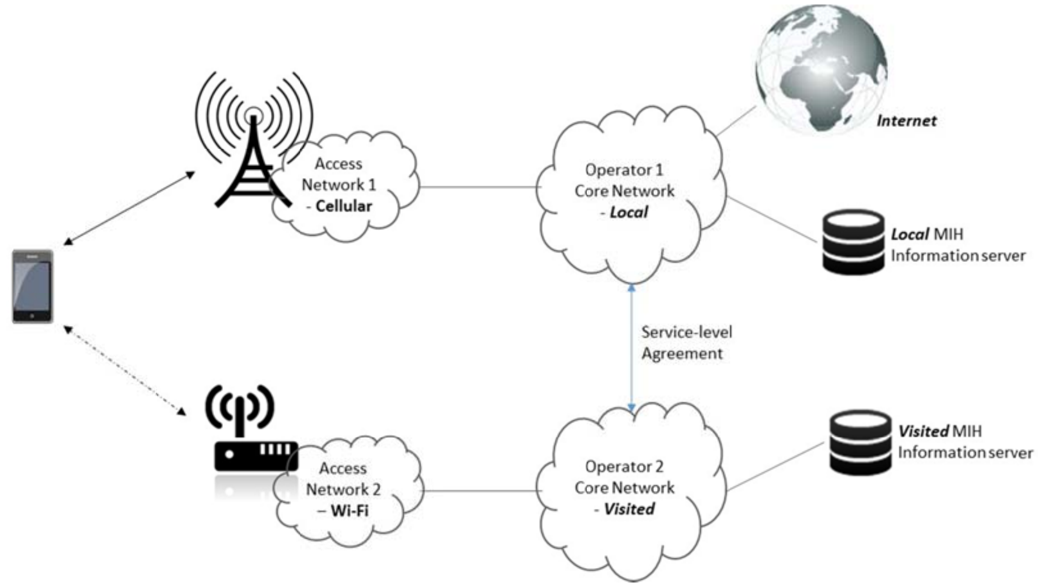


Figure 2-4: Example of IEEE 802.21-based network model (adapted from figure 3 of [23])

The model consists of two access networks that belong to two different core operators' networks. Moreover, the provisioning provider is the cellular network. At any given time, the MN may choose to get service from the subscriber's *home* (cellular) network or from a *visited* (Wi-Fi) network. Both *home* and *visited* core networks have access to the *home* and *visited* MIH Information server (respectively). These servers contain handover-related information, and can be accessed from other networks based on subscriptions or roaming agreements. Therefore, the MN is able to get MIH information from any access network of its service provider or from visited networks that have service-level agreements (SLAs) with the MN's service provider.

The IEEE 802.21 communication model is mainly based on two components: MIH functions (MIHFs) and MIH users. Those components are:

MIH functions (MIHF)	Logical entities, which are located within the protocol stack of either network entities or MNs, and provide MIH services to MIH users
MIH users	Entities located in both MNs and networks that get MIH services and generate MIH commands. Examples of MIH users are network layer mobility modules, such as Mobile Internet Protocol (MIP) and Radio Resource Management modules

Figure 2-5 shows a summary of the relationships between MIHFs and MIH users of network nodes found in the network model. As is illustrated in the figure, the communication model allows MIH-related information to be interchanged between two different MIHFs in any network entity.

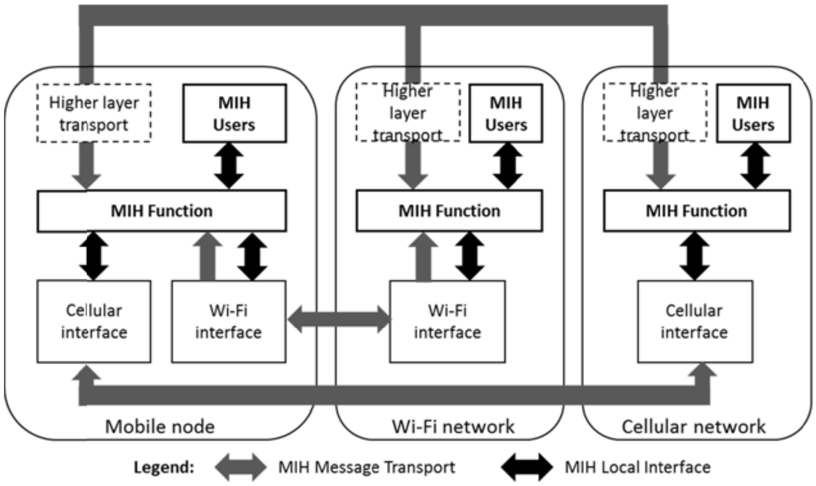


Figure 2-5: MIHF relationships (adapted from figure 5 of [23])

On the other hand, the standard defines three different services that are provided by MIHFs. These services facilitate handovers between diverse PoAs in heterogeneous networks. An overview of MIH services follows:

Media independent event service The media independent event service detects changes in link-layer properties and generates events (i.e., triggers) to notify others of the changes of MIH users from local to remote interfaces or *vice versa*. These events are used to detect the need for handovers.

Media independent command service This service provides a set of commands that enables MIH users to manage and control link behavior. For instance, these commands allow both MNs and network entities to initiate handovers, exchange information about available networks, and negotiate the best available network under different conditions (notice that these commands could be generated by a local or remote MIH entity).

Media independent information service (MIIS) This service provides information about different available networks and their services within a geographical area. The information typically consists of static link-layer parameters such as channel information, MAC addresses, or security information about a PoA. Moreover, the information may be stored in *MISS servers* within different networks. Details of this information are outside the scope of the standard.

Concerning use cases of this standard, the standard allows a mobile device to save energy while looking for available wireless networks, by avoiding the need to power up multiple interfaces to get that information. For example, when connected to a cellular network, the MN can access information about IEEE 802 networks within a given region. This capability allows the MN to use its current access network to inquire about other potentially available access networks within a geographic region. This example, referred to as *mobile-initiated handover*, is further described in Appendix C of the standard. Furthermore, this model allows a MNO to control MN’s choice of connectivity, in order for it to use the best access network at any moment. This operation is referred to as *network-initiated handover*, and is explained in Appendix C-2 of the standard. A simplified version of that explanation is shown in Figure 2-6.

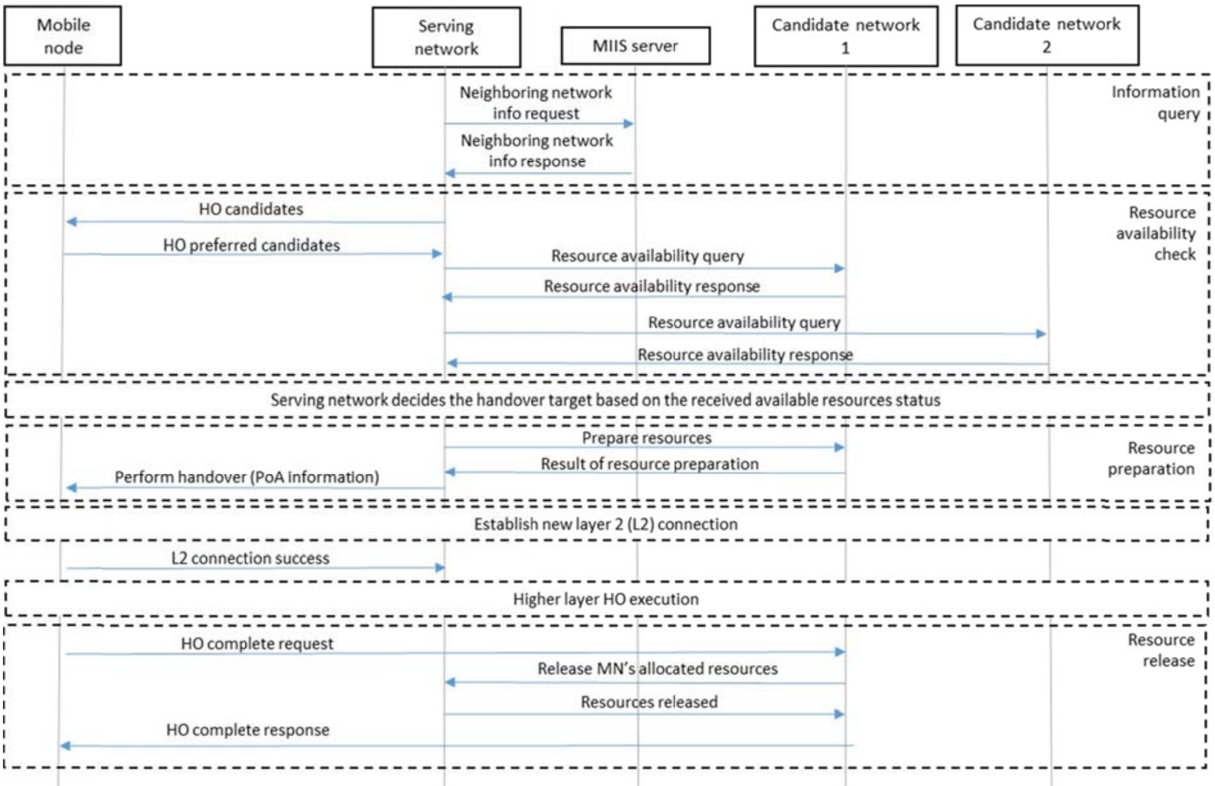


Figure 2-6: IEEE 802.21-based network-initiated handover operation

A MIH user entity in the cellular network commands the MN through the MIHF of the cellular network to perform a handover to a specified access network. Information about available PoAs could have previously been obtained by the serving MIHF by querying the MIIS server of another network. Moreover, the serving MIHF could request candidate networks to provide information about their available resources in order to decide upon handover targets.

Finally, the standard provides a mechanism to use when a MN does not know its current location. In this case, an observed link-layer address or cell ID is sent to the information server and the information server uses this as a hint to estimate the client's current location.

Some interesting solutions based on the IEEE 802.21 standard have been proposed in the research community and will be described below.

2.5.2 IEEE 802.21-based solutions

In [24], Lim, et al. propose three enhanced WLAN discovery schemes based on the IEEE 802.21 standard. They assume that a MIH-capable MN obtains neighbor information by requesting this information from a MIIS server. Two problems can occur: (i) a MIIS cannot manage information for all APs since many are individually managed and (ii) the MN is not always able to know its current location without a significant energy cost. As a result they propose three different schemes according to the available information:

Channel information based discovery (CID) The MN performs a selective scan using a list of channels used by nearby APs as announced by the MIIS server. If no AP is detected after a certain number of attempts, then the MN performs a full scan and updates the channel list when a new AP is found.

Location information based discovery (LID)	Assuming that the MN knows its current location, it can construct a location map of nearby APs based upon the information it received from the MIIS server. The MN will set its channel switching interval to a shorter period if it is within the coverage of an AP. LID includes a learning mechanism so that the MN registers (in a location map located in the server) unmanaged APs that it discovered while scanning.
Channel and location information based discovery (CLID)	The location map includes the channel number of each AP. This enables a multimode terminal* to perform selective scanning when it is located within the coverage of an AP.

The simulations by Lim, et al. suggests that both CLID and LID achieve similar and the best performance, while CID is the worst among these proposed schemes - although is much better than conventional discovery scheme in terms of energy consumption and duration. However, these proposals do not make the most of the powerful services already offered by this standard.

In contrast, the solutions proposed by Khan and Andresen in [25] take greater advantage of IEEE 802.21 services. Their solutions are based on the use of an MIIS server to reduce the scanning latency. When a MN detects that a handover is imminent, it sends a query, which includes the MN's GPS coordinates, to the MIIS server in order to get channel configuration information for APs that in the future will be nearby thus avoiding the need to perform a full scan. The configuration information includes at least one BSSID and a currently operating channel number for a future AP. If this information is available, then the MN either scans only one channel (referred to in their paper as *Intelligent-ScanOne)* or directly associates on the indicated channel number and BSSID, thus avoiding scanning altogether. Their simulation results show a 93% and 99% scanning improvement respectively over classic IEEE 802.11 scanning, in terms of reduced delay.

While Khan and Andresen show that the use of the IEEE 802.21 standard leads to huge improvements over classic IEEE 802.11 scanning, they did not test if a large reduction in energy consumption can be achieved. In order to evaluate this we must consider the energy consumption incurred to obtain GPS coordinates and the possibility that the MIIS server returns no information or incorrect information. In the scope of this thesis project we are interested in evaluating the performance of MIIS-based solutions with respect to performing vertical handovers between cellular and IEEE 802.11 networks.

Regarding the power consumption of mobile devices within IEEE 802.21 capable infrastructure networks, it is difficult to find relevant papers. In [26], Wassie, et al. evaluate the energy consumption of the main steps of VHO operations relevant to the IEEE 802.21 framework. Their methodology is based on the Advanced Configuration and Power Interface (ACPI) measurement subsystem. The main conclusion drawn from their results is that the exchange of handover signaling messages consumes much less energy than the network scanning and association processes. Furthermore, both mobile and network-initiated handover schemes consume almost the same energy on average.

The previous proposals do not make use of one of the most interesting features of IEEE 802.21, which is the interaction between different access networks in order to improve the handover process. This interaction is used by Bae, Chung, and So in [27] to propose a vertical handover triggering mechanism, named data rate based vertical handover triggering

---

* A multimode terminal is a terminal capable of communicating via different RATs.

mechanism. Moreover, they adopt a time-to-trigger mechanism standardized in LTE to minimize unnecessary handovers. This time-to-trigger mechanism is based on the use of a timer to carefully execute handovers, while avoiding unnecessary handover executions. The MIH command and event services are used to request and obtain the available data rate of the candidate WLAN AP or cellular BS, even when the MN is currently attached to a cellular BS or WLAN AP. Their simulations show that their proposed mechanism achieves better capacity for not only the LTE cellular network but also better overall network capacity. Nevertheless, their work does not address the issue of energy efficiency, since the interface of the not currently attached network must perform some measurements (such as scanning) and the device must perform some computation.

A similar framework for specifying and delivering access network policies to a mobile device is defined by 3GPP in the Access Network Discovery and Selection Function (ANDSF). This framework will be described in the next section.

2.6 Access Network Discovery and Selection Function

3GPP's technical specification 3GPP TS 23.402 (ETSI 123 402) [20] defines an access network discovery and selection function (ANDSF). ANDSF provides network discovery and assistance data to mobile devices. Management objects that can be used by ANDSF and MNs are defined in 3GPP TS 24.321 (ETSI TS 124 321) [29].

Figure 2-7 illustrates a simplified network architecture based on ANDSF. An ANDSF-based network architecture must contain at least one ANDSF element located in the home public land mobile network (PLMN). This ANDSF element is called the home ANDSF (H-ANDSF), while an ANDSF element located in the visited PLMN is called visited ANDSF (V-ANDSF). Moreover, a MN (called in 3GPP's terminology - User Equipment (UE)) that is roaming can access both the H-ANDSF and V-ANDSF.

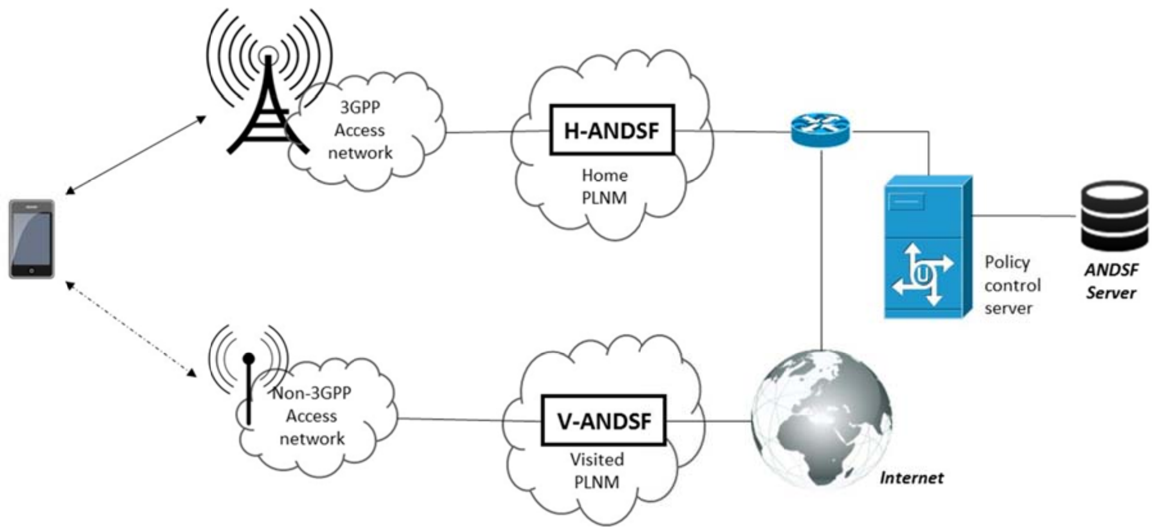


Figure 2-7: ANDSF-based network architecture

The ANDSF management object represents the ANDSF information, which is classified into three groups:

Inter system mobility policy (ISMP) A set of operator-defined rules and preferences that affect inter system mobility decisions (i.e. which RAN to use depending on conditions, such as location area or time) taken by the UE. ISMP is used when the UE can only route IP traffic over a single radio interface at a given time.

Inter system routing policy (ISRP)	ISRP is used when the UE can route IP traffic simultaneously over multiple radio interfaces, and indicates how to distribute traffic over the available access networks depending on whether this IP traffic matches specific criteria.
Access network discovery information	The ANDSF server may send access network discovery information to the UE upon request. This information consists of a list of access networks that are available in the vicinity of the UE, including the access technology type (e.g., IEEE 802.11 WLAN), and the RAN identifier (e.g. the BSSID of the WLAN AP).

The typical usage scenario in an ANDSF-based architecture is:

- The UE sends its current location to the ANDSF (H-ANDSF or V-ANDSF). The UE location information can be based on geographical coordinates; a cellular cell identifier (ID) or area ID; or a WLAN location (SSID, BSSID, and/or HSSID).
- The ANDSF sends to the UE the discovery information ("in this network access area, SSID x should be available"), an ISMP ("in this validity area, Wi-Fi SSID x is preferable"), or an ISRP ("in this validity area, Wi-Fi SSID x is preferable for these types of IP flows").

It is important to note (once again) that discovery and selection information regarding non-3GPP access networks (e.g. WLAN) can be provided to the UE even when the WLAN interface is not switched on – assuming that the cellular interface is switched on.

Consequently, the ANDSF mechanism provides operators with improved control over the decision-making process of MNs. However, the final decision is still made by the MN, and is dependent on vendor specific implementations. Furthermore, it does not provide real-time coordination with other nodes in the area.

In this thesis project we assume that the IEEE 802.21-based architecture is sufficiently complete and is better suited to support energy-efficient solutions than architectures based on ANDSF. The main reason for this is that the IEEE 802.21 standard enables more flexible interaction between the MN and other network entities, thus energy-efficient solutions can be achieved based on this standard or some small improvements may be needed to this standard.

Another integration tool developed to ease the integration of Wi-Fi is Hotspot 2.0 (HS2.0), developed by the Wi-Fi Alliance®. Further information about HS2.0 can be found in the following section.

2.7 Wi-Fi Alliance® Hotspot 2.0

The Wi-Fi Alliance® is working on a new program to enable seamless Wi-Fi access in public hotspots. The name of this program is Wi-Fi CERTIFIED Passpoint™ [6]. Service providers would like to increase the densification of their networks, and Wi-Fi seems to be an appropriate option due to its wide bandwidth and the fact that it is generally available in all locations where people congregate. MNOs would also like to offload traffic from their congested networks, as Stetson states in [30]: "Wi-Fi has emerged as a cost-effective solution for operators who face serious capacity problems on their mobile networks". The main goal of this program is to allow users to connect to visited networks as easily as they do to their home network. This process is called "Wi-Fi roaming". The roaming partners can include multiple-system operators (MSOs), MNOs, wireline operators, public venues, and a wide variety of enterprises (hotels, hospitals, airports, etc.). This work is known as Hotspot 2.0 (HS2.0) and will be further explained below. In order to simplify the explanation and to focus on those areas most relevant to the scope of this thesis, security issues will *not* be addressed further.

HS2.0 focuses on enabling a MN to automatically discover APs that have a roaming agreement with the user's home network and to establish secure network access. HS2.0 is built on the recently ratified IEEE 802.11u specification [31], which introduces a new entity called a subscription service provider (SSP) in order to manage users' subscriptions and access credentials. When a HS2.0 capable MN comes within range of a HS2.0 capable AP, IEEE 802.11u enables the unauthenticated MN to initiate a dialog with this AP to determine the AP's capabilities. This dialog is supported by the Generic Advertisement Service (GAS) and the access network query protocol (ANQP). The most relevant capabilities of an AP include:

- The domain name of the network operator. If the AP is not connected to the user's home network, then roaming is required.
- When roaming is required, the list of roaming partners that are supported by the AP is sent to the MN*.

Once the mobile device learns the list of roaming partners and the identity of the AP's operator, then the decision of whether to join the AP or not is based on built-in home network policies (i.e. whether there exist suitable agreements between home and visited networks).

The impact of HS2.0 could be enormous as it will change the nature of SSIDs, breaking the relationship between SSID and access credentials. What really matters is whether the visited AP has a roaming agreement with the user's home network provider. Although the main goal of HS2.0 is not energy efficiency, the use of IEEE 802.11u reduces the overhead of beacons and other management frames, thus potentially reducing battery power consumption and decreasing the duration of the network discovery process.

Many MNOs are working on deploying HS2.0. However, Wi-Fi CERTIFIED Passpoint™ - based solutions differ from the objective of this thesis project, as HS2.0 does **not** fully exploit the cooperation of cellular and Wi-Fi networks to reduce energy consumption.

A lot of work is being done in order to decide upon the best interworking architecture that integrates the cellular network and IEEE 802.11 WLANs. However, improvements regarding power saving mechanisms in IEEE 802.11 WLANs should be done in any case since a lot of energy is consumed to maintain connectivity. The existing solutions that aim to mitigate this problem will be illustrated in the following sections.

2.8 IEEE 802.11 Power Saving Mechanisms

IEEE 802.11 defines several different power saving mechanisms. This section will review these mechanisms.

2.8.1 IEEE 802.11 MAC Power Save Mode

This subsection reviews the power save mode in IEEE 802.11 infrastructure networks as defined in the IEEE 802.11-2012 standard [4].

A wireless network interface card (WNIC)† may be in one of two power states: awake and doze. During the awake state, the WNIC is fully powered on and is able to transmit and receive frames. In contrast, during the doze state, the WNIC consumes very little power, but is neither able to transmit nor receive frames. Moreover, two power management modes of

---

* This would seem to lead to scalability problems due to the potentially very large number of operators.

† We will refer to the interface as a network interface card as this is common usage, although generally the network interface need not be on a card that can be inserted or removed from the device.

operation are defined: active mode (AM) and power saving mode (PSM). In AM the WNIC is always in the awake state, while in the PSM it can change its state between the awake and the doze state depending upon the traffic pattern.

In an infrastructure BSS, the power management mechanism is centralized in the AP. The AP maintains a power-management status for each currently associated STA. Thus STAs who change their WNIC's power management mode should notify their associated AP. The AP is also the coordinator for the time synchronization function (TSF); hence the AP includes its TSF timer value in beacon frames which enables all the STAs within the BSS to be synchronized. In PSM, a WNIC periodically wakes up every listening interval and listens for beacon frames which contain a **traffic information map** (TIM). The TIM identifies the STAs for which traffic is pending (i.e., incoming frames for this STA have been buffered in the AP while the STA's WNIC was dozing). Another kind of TIM, referred to as a **delivery indication message (DTIM)**, is sent periodically (not at every beacon) within a beacon frame to advertise multicast/broadcast transmissions frames which have been buffered by the AP. Once the STA receives a new TIM (or DTIM):

- If the TIM (DTIM) indicates the presence of a buffered frame for the STA, then the STA's WNIC remains in the awake state, contends for the channel, and sends a power-save poll (PS-poll) frame to request the transmission of the buffered packets. A single buffered frame is sent by the AP after a PS-poll is received. Moreover, the AP indicates whether or not there are more data frames outstanding, using the "More Data" bit in the data frames. The STA's WNIC stays in the awake state until it has retrieved all pending data (i.e. receives a frame with more data bit set to 0).
- If the TIM (DTIM) indicates that the AP has no buffered frames for the STA, then the STA's WNIC immediately enters the doze state.

Figure 2-8 shows an example of the operation of PSM in an IEEE 802.11 infrastructure network.

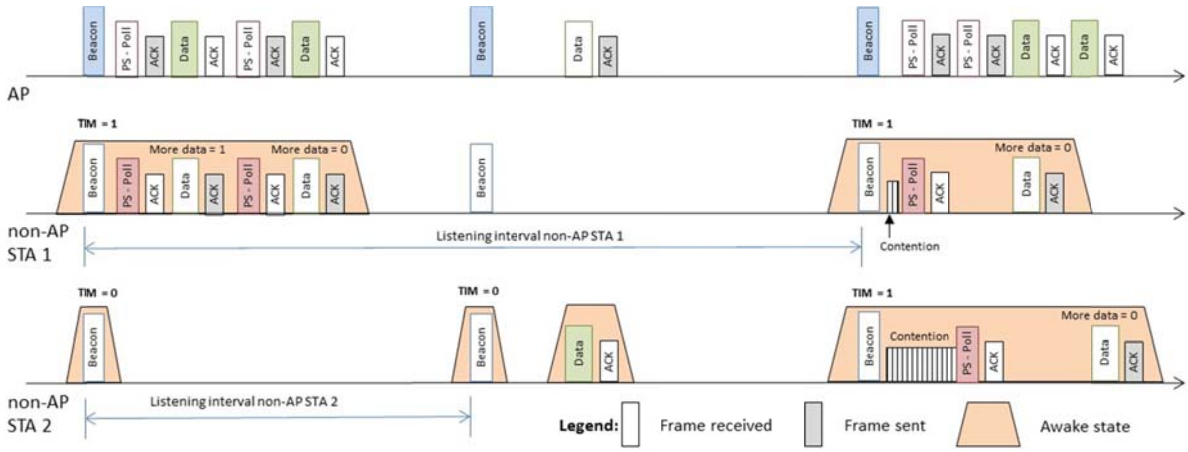


Figure 2-8: Power saving mode operation

According to Ding, et al. in [32] there are two major PSM implementations: Static PSM and Dynamic PSM. Static PSM corresponds to the classic operation of PSM, which was explained above. If the received TIM indicates pending frames buffered at the AP, and then the STA switches its WNIC to the awake state, sends PS-poll requests, and the WNIC remains in the awake state receiving buffered frames until the AP indicates it has no additional buffered frames for this STA. At this point the STA immediately switches its WNIC to the doze state. Static PSM achieves good energy savings, but introduces considerable delay since every packet needs to be polled and because packets that arrive shortly after the start of a doze period must wait to be delivered during the next beacon interval, because the WNIC switched to the doze state immediately after receiving the last frame. In contrast, Dynamic PSM avoids the use of PS-poll requests and prevents the STA's WNIC from going to sleep in the middle of a sequence of frames by keeping the WNIC in the awake state for a pre-defined duration (called the PSM timeout) following transmitting or receiving a frame. However, dynamic PSM can lead to unnecessarily long periods in the awake state, thus reducing the energy saving. The main challenge is that the PSM timeout should be set large enough to accommodate most of the higher layer protocol's round trip times (RTTs). In [32], Ding, et al. suggest the use of a local proxy server at the AP to improve the energy savings for web surfing flows. The proxy server provides a constant short delay in responses to the client, hence allowing the PSM timeout to be shorter. Moreover, it allows the client to enter PSM while the proxy is getting/sending data from/to the remote server. Their experiments show that this solution reduces the energy-consumption by 44-67% compared to various Dynamic PSM configurations. Unfortunately, this proxy server would violate all of the user's personal integrity and privacy.

In [33], He et al. claim that the use of PSM can extend the operating lifetime of a mobile device by up to 250% with a moderate traffic load at the AP. However, the power saved would be considerably reduced if there is even a limited amount of additional traffic (to/from other MNs) via this AP. The reason is that the delivery of data frames for a STA can be interrupted by other competing traffic being set to or from the AP. This other traffic forces the STA's WNIC to remain in the awake state for a longer period and thereby increasing the STA's energy consumption. Both analytic and simulation results show that the expected decrease in energy efficiency ratio is roughly a function of the number of STAs associated with the AP. For example, if there are k other STAs generating traffic (with a data rate over 1000 kbps) in the same area, then the time spent in the data retrieving process will be around (k + 1) times higher than in the case without competing traffic. This follows from the fact that each of the k+1 STAs is competing for the same resource: channel capacity, hence if each of them is generating a comparable number of frames then they effectively divide the channel rate by k+1.

On the other hand, the downlink delays due to using PSM may be unacceptable for certain QoS-sensitive applications, such as real-time applications (e.g. VoIP). In order to provide QoS support and achieve power saving, IEEE 802.11e [34] defines a mechanism called automatic power save delivery (APSD). This mechanism is described in the following section.

2.8.2 Automatic Power Save Delivery (APSD)

IEEE 802.11e defines the automatic power save delivery (APSD) mechanism as an enhancement of PSM. APSD's main function is to manage the delivery of data frames to STAs in PSM, while avoiding the necessity of PS-Poll frames. APSD defines two delivery mechanisms: a distributed power saving scheme, called unscheduled APSD, and a centralized one, called scheduled APSD. These mechanisms are in line with IEEE 802.11e's QoS mechanisms, which have been already described in section 2.2.3. Summarizing, the standard defines two channel access mechanisms: enhanced distributed channel access (EDCA), which is a contention-based channel access method that provides support for traffic prioritization; and the centralized hybrid coordination function (HCF) controlled channel access (HCCA), which is a contention-free channel access method that provides support for parameterized QoS. APSD also defines service period (SP) as the period of time when the device is awake in order to receive buffered frames from the AP. An overview of both mechanisms is provided below. Further details can be found in section 3.1.2.

2.8.2.1 Unscheduled APSD (U-APSD)

The novel idea is to use uplink frames (sent from the STA to the AP) as indications (triggers) of STAs being awake, triggering the initiation of unscheduled SPs. So that, the AP can deliver data frames destined to the STA that were buffered when the STA was asleep. Therefore, an unscheduled SP is initiated by an STA sending a trigger frame or QoS data frame to the AP. In contrast, the end of a SP is indicated by the AP by setting the end of service period (EOSP) flag found in data frames. Figure 2-9 illustrates U-APSD operations.

U-APSD uses EDCA access method (remember that EDCA defines four access categories (ACs) for data frames: *AC_BE, AC_BK, AC_VO,* and *AC_VI*). Moreover, U-PASD introduces two different labels for data frames of each AC of each station: *delivery-enabled* and *trigger-enabled* data frames. Moreover, legacy IEEE 802.11 PSM is used for frames that are neither *delivery-enabled* nor *trigger-enabled*. When an AP receives a frame of a *trigger-enabled* AC, an SP is started if one is not in progress. After that, the AP delivers the buffered frames corresponding to *delivery-enabled* ACs using EDCA. The number of data frames of *delivery-enabled* ACs delivered by the AP in a single unscheduled SP is limited by the *Max_SP_length* parameter.

In order to guarantee backward compatibility, the TIM indicates the buffer status only of the *non-delivery-enabled* ACs, unless all ACs of a station are *delivery-enabled*. Only in this later case, TIM indicates the buffer status of *delivery-enabled* ACs. Notice that buffered *delivery-enabled* frames may not be advertised by the AP because they are delivered during unscheduled SPs triggered by uplink frames. Furthermore, in the case of an STA that does not have data frames to transmit via the uplink, QoS null frames, which are substitutes of *PS-poll* frames in PSM, are sent in order to trigger unscheduled SPs.

Furthermore, the configuration of the different ACs as *trigger-enabled* or *delivery-enabled* can be done during the association period, or using special frames for that purpose. These frames are referred to as *add traffic stream* (ADDTS), and they contain *traffic specification* (TSPECs) elements that are used to set the different ACs.

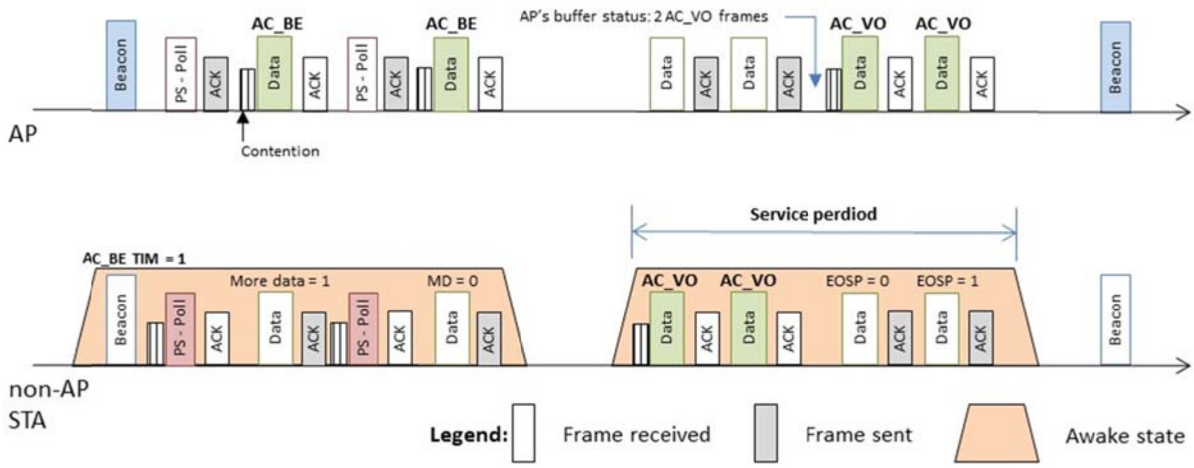


Figure 2-9: Unscheduled APSD operations (adapted from figure 1 of [35])

Figure 2-9 was an example of U-APSD operations. The STA has the best effort access class (*BE_AC*) configured as neither *delivery-* nor *triggered-enabled*. In addition, the voice access class (*VO_AC*) is both *trigger-enabled* and *delivery-enabled*. Initially, the STA's WNIC wakes up to listen for the beacon frame, which indicates that there are buffered frames belonging to the *BE_AC* access class. Then, those frames are retrieved using legacy PSM operation because this AC is neither *delivery-enabled* nor *triggered-enabled*. On the other hand, the STA'WNIC wakes up later in order to send data frames of *AC_VO* access class. Since that type of AC is *trigger-enabled*, and then an unscheduled SP has started so the STA's WNIC must remain in the awake state listening for data frames from the AP after it has sent the frames via the uplink. Finally, the AP sends some data frames of *AC_VO* access class (because it is *delivery-enabled as well*), and indicates the end of the SP setting the EOSP flag to 1 in the last data frame.

2.8.2.2 Scheduled APSD (S-APSD)

The AP schedules when different stations using S-APSD should wake up to receive buffered frames or send data frames in the uplink. A STA can request the AP use a specific delivery mechanism (EDCA for access categories or HCCA for traffic streams). STAs have active flows for which scheduled SPs are reserved. Every active flow is characterized by: *service start time* (SST), *service interval* (SI), and TXOP associated with the flow. The SST is the initial time of the first SP, the SI is the fixed interval time between two successive SPs, and the TXOP is the maximum duration of the SPs.

In order to set up a new flow during the initial phase, which is called TSPEC negotiation, the STA sends an ADDTS request that describes the new traffic flow. Then, the admission control function at the AP decides whether the new flow can be accepted or not. If the AP can satisfy the requested service, it will indicate this to the STA in the *schedule element* (SE) of the response, which includes SST, SI, and TXOP parameters associated with the flow. Moreover, the scheduling function at the AP is in charge of set up these parameters based on resource allocation algorithms. Further information about these different resource allocation approaches can be found in [36].

Regarding the transmission procedure, flows can be set in Downlink or Uplink directions. It is also possible to set up a bidirectional flow, which is equivalent to a downlink and an uplink flow. APs can directly send data frames belonging to a downlink active flow at every SP associated with the flow. Moreover, the STA's WNIC will automatically wake up at the scheduled time of each SP of each downlink active flow, and will remain awake until it receives a frame with the EOSP flag set to 1. On the other hand, the STA's WNIC must wake up at scheduled times of each SP of its uplink active flows, and send data frames to the AP. In that case, the STA is the one that indicates the end of scheduled SPs to the AP.

An example of S-APSD operation is shown in Figure 2-10. This STA has, an active bidirectional flow (referred to as TS1 in the figure) configured to use S-APSD with HCCA. It could be, for instance, a VoIP call. In addition, the STA has *AC_VI* access class configured to use S-APSD with EDCA in the downlink. It could be, for instance, streaming video traffic. Firstly, the STA's WNIC wakes up at the scheduled instant of a downlink SP of the TS1. Later, it wakes up at scheduled instant of an SP of *AC_VI* in order to receive video frames. Then, there is a scheduled uplink SP of the TS1, in which the STA sends voice frames to the AP. After that, the STA's WNIC continues waking up at scheduled instants, which are indicated by the service interval value of each flow.

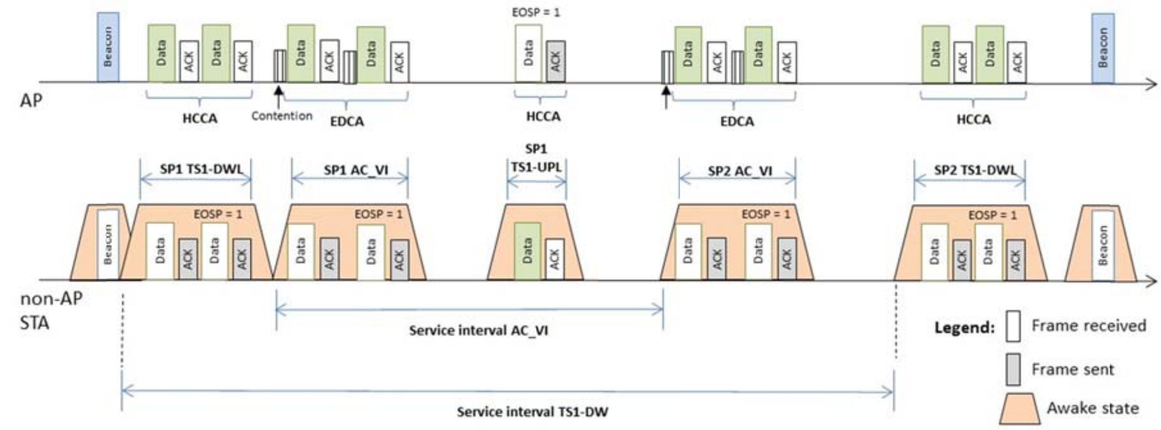


Figure 2-10: Scheduled-APSD operation (adapted from figure 1 of [35])

In the following paragraphs, some interesting prior research results regarding the APSD mechanism will be reviewed.

In [35], Perez-Costa and Camps-Mur evaluate the performance of the different possible combinations of IEEE 802.11e QoS and power saving mechanisms. Their analysis also includes the baseline IEEE 802.11 access mechanism, called the distributed coordination function (DCF), and power saving mode (PSM). The five combinations are: HCCA + S-APSD, EDCA + S-APSD, EDCA + U-APSD, EDCA + PSM, and DCF + PSM. The proposed scenario is based on basic clusters composed of four stations, running VoIP, video streaming, web browsing, and FTP download applications respectively. The number of clusters increases from three to 63 clusters, and one-third of the stations of each application type operate at 54 Mb/s, one-third at 24 Mb/s, and one-third at 12 Mb/s. The results for real-time and non-real-time applications are summarized below:

Real-time applications:

- PSM based configurations achieve the worst performance, since PSM requires more signaling and also increases the collision probability due to the fact that all STAs attempt to request buffered frames immediately after the beacon from the AP is received.
- S-APSD-based configurations outperform all other approaches, thanks to the fact that no signaling is required to directly start the downlink transmission.
- Regarding throughput, EDCA + S-APSD is the best performing combination for a VoIP application, because EDCA gives higher priority to voice flows, while HCCA + S-APSD gives greater priority to video streaming, due to the fact that HCCA does not take care about a real-time flow's duration.
- Both S-APSD and U-APSD successfully achieve delays within acceptable values.
- Finally, with respect to energy consumption, the same results as for delay are reached. The obvious reason is that the longer the STA's WNIC remains awake *waiting* to transmit or receive, the more energy is consumed.

Non-real-time applications:

- FTP applications reach on average a TCP connection throughput of around 25 Mb/s, while web applications achieve a maximum of 600 kb/s. The reason for this is that the STAs running web applications WNICs go to sleep *before* the next window of TCP packets arrives, due to the large number of small sized web objects. One solution to this problem is the use of a local proxy server at the AP. Perez-Costa and Camps-Mur propose to use the optional pipelining mechanism of HTTP 1.1.
- EDCA + S-APSD configuration performs best.

The main conclusion is that S-APSD configurations outperform the other configurations in almost all cases due to the lower required signaling load for downlink transmissions due to the S-APSD mechanism. Furthermore, both S-APSD and U-APSD mechanisms lead to acceptable delays.

Several improvements to the network discovery process and power saving mechanisms for IEEE 802.11 WLANs have been studied. The remaining challenge is: Which RAT should a device associate with at each moment? Regarding the handover operation, this process is referred to as the handover decision mechanism, whose main function is to choose the best available network based upon the information gathered during the discovery phase.

2.9 VHO decision: Energy-efficiency QoS tradeoff for RAT selection based upon traffic patterns

The majority of WLAN-capable devices always connect via WLAN if WLAN connectivity is available, and then all user data traffic is sent over the WLAN interface. This behavior could lead to lower performance in those cases when the cellular interface could provide better performance than the WLAN interface. For this reason, enhanced access network selection mechanisms could lead to improvements in performance from both the user's QoS experience and from a network perspective. The following paragraphs describe some previous work regarding the performance of WLAN and cellular networks from a power consumption perspective.

Jose Maria Rodriguez [7] showed that Wi-Fi is more energy efficient than LTE with the current implementations of both interfaces. He analyzed the energy consumption of both technologies under different traffic assumptions. The conclusions of his analysis were:

VoIP and video calls	Using Wi-Fi is much more efficient, *even if* the LTE interface must be left on during the call.
Downloads	If the total amount of the data is larger than 10 Mbytes, then the LTE interface is less energy consuming when conditions are good. However, if the Wi-Fi signal is good, then it is not efficient to perform a handover from LTE to WLAN *unless* LTE network load is very low. For small downloads the potential energy savings of using LTE decreases significantly.
Uploads	VHOs from Wi-Fi to LTE will only save some energy if the LTE uplink is in a "lower power mode" *and* Wi-Fi conditions are poor.
Web browsing sessions	VHOs from Wi-Fi to LTE are *not* efficient due to the very low energy consumption of these sessions.

In [37], Kellokoski, Koskinen, and Hamalainen draw similar conclusions for UMTS to what Rodriguez concluded for LTE. Their experiment showed WLAN performance was superior to UMTS. WLAN technology achieves greater energy efficiency because it offers higher data rates. Since both UMTS and WLAN interfaces have similar power consumption as a function of time, slower connections lead to longer transmissions, and consequently higher energy consumption. The results reported by Petander in [38] show that the energy consumption for data transfer over UMTS can be up to three orders of magnitude higher than over WLAN for bulk transfers. Even though a vertical handover consumes a significant amount of power, loading a large web page would generate enough traffic to justify a handover from UMTS to WLAN.

Despite the fact that WLAN usually achieves better energy efficiency than cellular networks, the network load and the signal quality have a significant impact on power consumption. In [39], Trestian, et al. analyzed the impact of networking load and signal quality in the context of video delivery. In the radically different cases they studied, they compared the ideal scenario (where the MN is close to the AP without background traffic) with the extreme one (where the MN is far from the AP with background traffic), the average energy consumption considering a high video quality level is up to 30% higher in the second scenario. However, as shown by Trestian, et al. in [40], the average power consumption considering the same video quality level is even higher when using the cellular network than using the WLAN network over poor conditions.

The attributes considered while making VHO decisions are crucial for the performance of the VHO process. For this reason power-consumption and QoS parameters should play an important role in such a decision. In [41], Petander proposed and analyzed an energy *and* QoS aware handover decision mechanism. Simulation results show better performance regarding power consumption, availability, and QoS requirements than other algorithms.

3 Method

Wi-Fi technology is expected to be fully integrated into mobile access and core networks. Within such a multi-RAT environment, so called heterogeneous networks, resources should be utilized in an optimal way. Therefore, with Wi-Fi fully integrated into mobile access, MNOs would be able to control devices' behavior concerning connectivity decisions in order to guarantee the best experience for users and optimizing use of resources at the same time. The main challenges that operators are focusing on to achieve this objective are detailed in [42], and in summary are:

- Support for seamless handover between cellular and Wi-Fi networks,
- Offload data traffic from cellular networks to WLANs,
- Support QoS,
- Maintaining data throughput, and
- Support seamless handover between different Wi-Fi access points.

On the other hand, due to the fact that the use of Wi-Fi technology is expected to grow considerably, a lot of work must be done in order to improve its performance as well. The main challenges that should be addressed to improve the performance of such a technology are:

- Reduce STAs' WNIC energy consumption by using and improving power saving mechanisms such as power saving mode (PSM) and automatic power save delivery (APSD)
- Enhance QoS support, by using the mechanisms proposed in the IEEE 802.11e standard [34]
- Improve WLAN discovery schemes reducing both energy consumption of the scanning phase and overall handover latency
- Support seamless handover between APs
- Improve MAC protocol performance. The CSMA/CA MAC protocol used in WLANs is an energy-consuming protocol, as was mentioned in section 2.2.3. Enhanced MAC protocols would reduce STAs' WNIC energy consumption and improve network performance, so that APs can deliver higher throughput rates with the same equipment.
- When there are too many STAs associated with an AP, the network performance degrades. Therefore, the AP must ensure that the service for STAs that are already associated with it will not be dramatically degraded by the addition of another STA.
- Assist cellular/WLANs interworking schemes to facilitate Wi-Fi integration into mobile access. In order to do that, APs shall collect real-time information about all the available resources nearby (e.g. neighbor APs, network conditions, etc.). This information can subsequently be used to optimize network selection decisions in order to improve both the use of the available resources and the user's experience.

In this thesis project, a solution has been designed that addresses the above challenges so as to improve WLANs' performance and facilitate their integration into mobile access networks. The proposed solution addresses STAs' energy consumption while supporting QoS. In addition, the proposed solution includes an energy-efficient scanning scheme that allows seamless handovers between APs.

The proposed solution is described in section 3.1, while section 3.2 explains how this solution could facilitate Wi-Fi/cellular networks interworking. Additionally, a simulator has been implemented in order to detect faults in the different operations and to evaluate the solution's performance. The description of this simulator and the results of simulations are given in the next chapter.

3.1 Proposed solution to enhance WLANs' performance

The proposed solution aims to enhance WLANs' performance in terms of energy-efficiency, optimal use of resources, and users' experience. In order to simplify the discussion, a single ESS scenario has been considered, although the solution has been designed to be scalable to both multiple-ESS and heterogeneous networks. Figure 3-1 shows a possible network architecture for such a single ESS scenario. This single ESS is referred to as ESS_1 and it uses a single SSID *x*. This ESS has several APs, each of which has its own BSSID. The APs are connected via a DS, typically simply an Ethernet switch, to an Internet gateway. We will assume that this Internet gateway also acts as a DHCP server to provide each MN with an IP address for use within ESS_1. Based upon this description we can see that this scenario represents a single IP domain with the APs acting as layer 2 bridges and the Internet gateway acting as the default router for this IP domain.

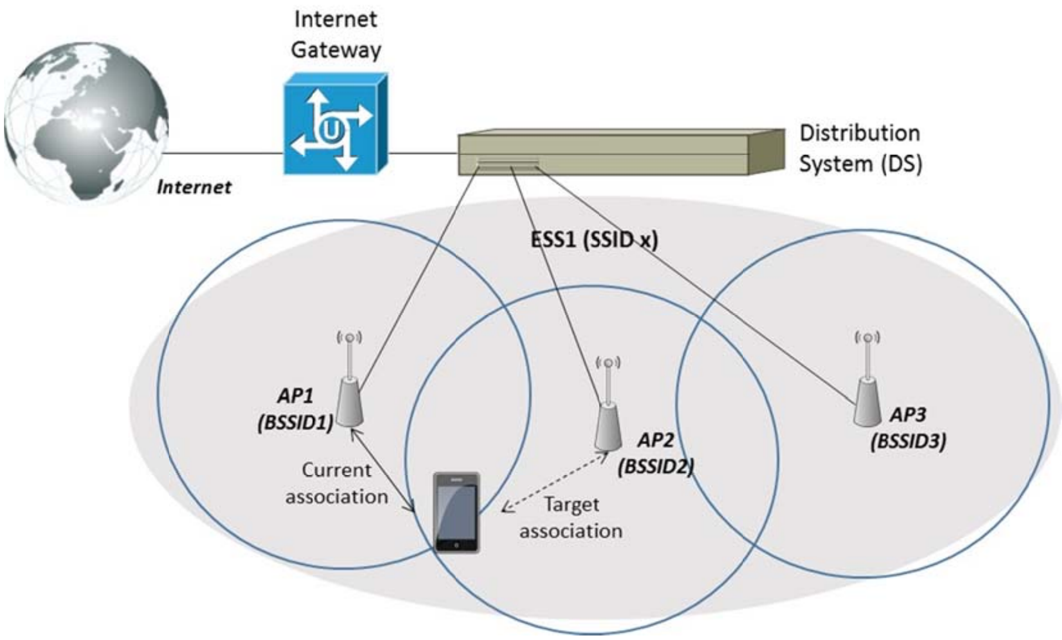


Figure 3-1: Example of network architecture for a single ESS scenario

The proposed solution aims to reduce STAs' energy consumption by combining an enhanced energy-efficient MAC protocol and APSD mechanisms introduced by the IEEE 802.11e standard [34]. The enhanced MAC protocol that has been designed improves channel access by using both contention free periods (CFPs) and contention periods (CPs). This allows optimizing use of resources by reducing contention and inter-frame spaces (IFSs). In addition, energy-efficiency is achieved by using APSD mechanisms that minimize the time that the STA's WNIC operates in power consuming states (i.e. transmitting and receiving) while supporting QoS. Sections 3.1.1 and 3.1.2 describe the proposed MAC protocol and further information about APSD mechanisms.

In addition, the proposed solution allows seamless handover between APs. This feature is achieved due to an innovative idea referred to as "*seamless-S-APSD*", which enables continuity of S-APSD active flows after a handover to another AP is performed. This idea is presented in section 3.1.3. Moreover, the solution includes a new scanning scheme to make "*seamless-S-APSD*" possible. In this scanning scheme, MNs actively scans in the background for a selected list of APs located in the vicinity of the MN. This is referred to as *background selective active scanning*, and is explained in section 3.1.6.

Finally, a *discovery information mechanism* and a *neighbor report mechanism* have been designed to collect relevant handover-related information in order to improve scanning operations. These mechanisms are introduced in sections 3.1.4, and 0 respectively. Finally, section 3.1.7 contains an example of the overall proposed solution's operation.

3.1.1 Enhanced MAC protocol

An introduction to IEEE 802.11 MAC protocol was given in section 2.2.3. As stated previously the IEEE 802.11 MAC protocol is an energy-consuming protocol. Enhanced MAC protocols aim to reduce power consumption and increase bandwidth efficiency by reducing IFSs & contentions and minimizing the number of acknowledgments. A survey of energy efficient MAC protocols is presented in [8]. This section proposes a simplified enhanced MAC protocol. The protocol utilizes both unscheduled and scheduled APSD at the same time, thus minimizing energy consumption - while allowing QoS for data transfers. The reader is referred to section 2.8.2 for an introduction to the APSD mechanism.

The beacon interval is divided in two periods: a contention free period (CFP) located at the beginning of the beacon interval, and a contention period (CP) that starts right after the end of the CFP. These CFP and CP concepts are already defined in the IEEE 802.11 standard to support both DCF and PCF access methods. During the CFP, all STAs must access the channel using S-APSD, i.e. all the SPs of S-APSD active flows are allocated within the CFP. This mechanism allows non-AP STAs to access the WLAN channel using prescheduled TDMA-like access. This idea is illustrated in Figure 3-2**.** Allocating all the SPs one after the other is known as a *grouping approach*. This forces the AP to use a fixed service interval common to all active flows. The reader is referred to Mur, et al. [36] to gain further knowledge about this and other resource allocation approaches.

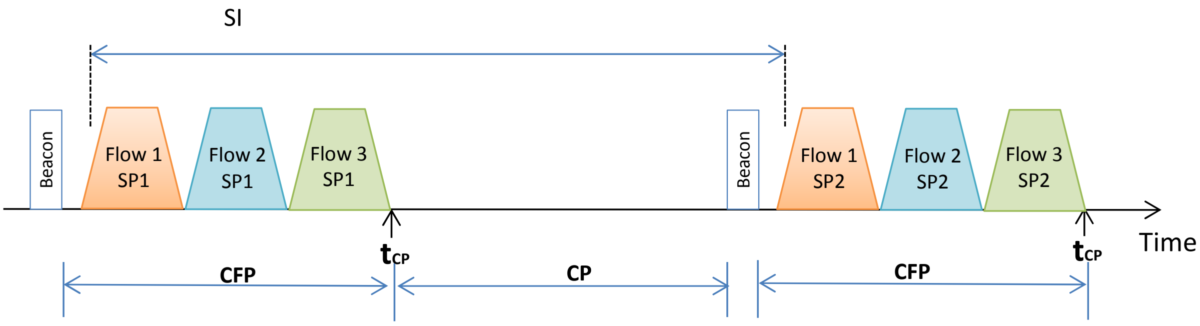


Figure 3-2: During the CFP, all STAs must access the channel using S-APSD

During the CFP, SPs are reserved for their corresponding STAs. These STAs' WMICs will wake up at the scheduled instants when their SPs start. This approach aims to reduce mobile devices' energy consumption, while assuring QoS. Since the station that is associated with the currently active SP is the only station that has the right to access the channel, STAs do *not* have to contend for channel access within the SPs. Therefore, bandwidth efficiency during the CFP is increased by avoiding channel contention and minimizing IFS. The end of the CFP, which is labeled t_{cp} in Figure 3-2, is known by the AP, and it announces t_{cp}'s value in beacon frames. In order to avoid wasting bandwidth, all the SPs must be reallocated after a flow leaves the system. The reallocation process is done flow by flow, i.e. every active SP must be reallocated in order to use the gap left by the deleted flow. The reallocation process results in a new SST for each flow, but each flow retains its SI. Notice that the AP shall advertise the STAs' new schedule by sending a *Schedule Element* (SE) with the updated SST parameter. When all the active flows have been reallocated, the new value of t_{cp} will be updated in the beacon frames. This idea is illustrated in Figure 3-3.

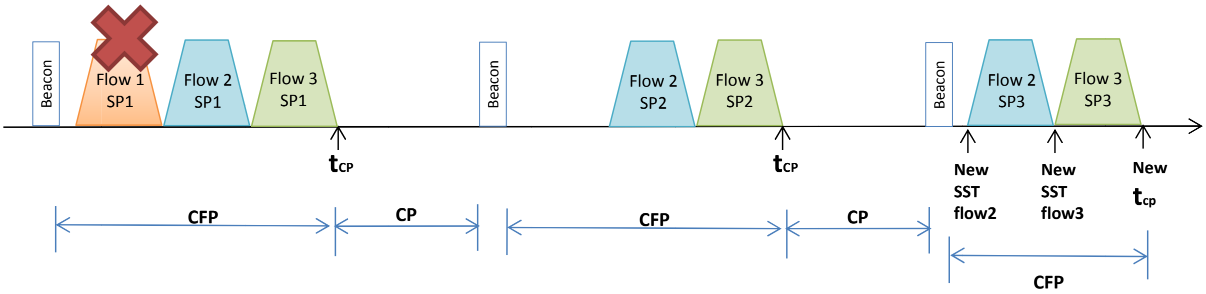


Figure 3-3: Every active SP must be reallocated in order to use the gap left by the deleted Traffic Stream

However, this approach implies the use of S-APSD SIs equal to the duration of the beacon interval, which is typically 100ms. This SI does not meet the QoS requirements of some real time applications, such as VoIP. The requirements necessary to support multimedia services are specified by ITU-T in [43]. The recommended one-way delay for voice services is below 50 milliseconds. In order to solve this problem, the beacon interval can be divided into four periods, and each period must contain a CFP and a CP. This approach is illustrated in Figure 3-4. Therefore, there would be three different values of t_{cp} for every beacon interval, and all of them must be announced via the beacon frames. Moreover, different SI values could be used for a defined flow, as multiples of $\frac{beacon\_interval}{4}$. If we consider a beacon interval equal to its typical value, i.e. 100ms, then typical SI's values might be, for instance, 25, 50, 75, and 100ms.

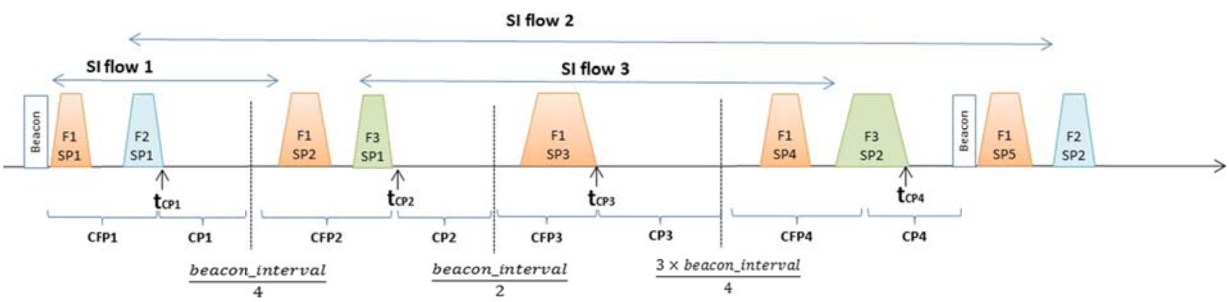


Figure 3-4: Allocation of S-APSD flows using four CFPs-CPs every beacon interval

On other hand, during the CPs STAs must use the U-APSD mechanism or DCF for channel access. Note that non-QoS STAs can also access the channel using normal DCF operations during the CP. Both mechanisms imply a STA must contend for the channel before sending a packet. U-APSD was described in section 2.8.2. U-APSD operations are complex, thus in this thesis a simplified version of the U-APSD mechanism is used. Since S-APSD is used during the CFP to ensure QoS for those flows that have special requirements, the goal when using U-APSD during the CP is to reduce energy-consumption while continuing to deliver short transmissions.

Concerning U-APSD operation, the TIM field, which is present in all beacon frames, indicates whether there are pending frames buffered at the AP to be delivered to non-AP STAs. The data frames in the uplink (non-AP STA → AP) direction are sent whenever non-AP STAs need to transmit. These frames trigger the initiation of unscheduled SPs during which the AP delivers buffered data frames to the STAs. Moreover, when a non-AP STA realizes that there are buffered frames for it at the AP, it sends a QoS null fame (i.e. a QoS frame that carries no Data field) to trigger a new unscheduled SP. Therefore, all non-AP STAs that receive a positive TIM should wake up at the beginning of one of the AP's CPs, and contend for channel access in order to send the uplink frame that triggers the unscheduled SP. The fact that STAs are aware of the beginning of the CP allows a station to save energy by avoiding channel contention during the CFP since it will be impossible to access the channel during the CFP because it is *never* idle. This approach involves all issues of an active STA using the DCF access, including the use of virtual carrier sense mechanisms, since all the STAs' WNICs that contend for the channel access are awake. Nevertheless, within a U-APSD SP, the STA does not have to contend for the channel since the channel is reserved for this STA. The fact that there is no contention within an unscheduled SP avoids other stations attempting to access the channel and interrupting the unscheduled SP.

With regard to broadcast/multicast traffic, the DTIM present in the beacon indicates the presence of buffered broadcast/multicast frames. Therefore, STAs' WNICs receiving a positive DTIM should wake up at the beginning of the CP to retrieve these buffered frames.

The following subsection further describes the S-APSD mechanism, which was introduced in section 2.8.2. The reader is referred to [34] for further information.

3.1.2 S-APSD in depth

The AP schedules the instants when its attached QoS-capable STAs' WNICs should be awake in order to retrieve buffered data frames or send frames in the uplink. A scheduled SP starts at fixed intervals of time as specified by the SI parameter. The scheduling process is flow-based, i.e. each active flow, also called a traffic stream, has its own assigned SI, which indicates the instants when the service periods associated with that traffic stream (TS) start.

In order to set up a new flow, a non-AP STA sends an ADDTS request frame to the AP, in which it includes the TSPEC element for this flow. A TSPEC element describes the characteristics and the QoS expectations for a traffic flow. The non-AP STA sets only some known parameters of the TSPEC element in the ADDTS request. These parameters typically are the variables of a token bucket model that describes the TS (i.e., mean data rate, peak data rate, and burst size). The AP may change the value of those parameters that have been left unspecified by the STA.

	Element ID	Length	TS Info	Nominal MSDU size	Maximum MSDU size	Minimum Service Interval	Maximum Service Interval	Inactivity Interval	Suspension Interval
Octets:	1	1	3	2	2	4	4	4	4

	Service Start Time	Minimum Data Rate	Mean Data Rate	Peak Data Rate	Burst Size	Delay Bound	Minimum PHY Rate	Surplus Bandwidth Allowance	Minimum Time
Octets:	4	4	4	4	4	4	4	2	2

Figure 3-5: Frame format of TSPEC elements (adapted from figure 46ti of [34])

The AP uses a resource allocation algorithm to check whether the new flow can be served, and computes the scheduling parameters. Readers are referred to [36], where different resource allocation algorithms are presented. Next, the AP decides whether to admit the TSPEC as specified, refuse it, or not admit it but rather propose an alternative TSPEC. After this, the AP replies with an ADDTS response frame to the non-AP STA containing a TSPEC and a SE. The latter defines the scheduled intervals reserved for this flow. The structure of such a schedule element is defined in Figure 3-6. The SST indicates the instant when the first SP of that flow starts, while the SI indicates the period of time between two successive scheduled SPs associated with that flow.

	Element ID	Length	Schedule Info	Service Start Time	Service Interval	Specification Interval
Octets:	1	1	2	4	4	2

Figure 3-6: Frame format of Schedule elements (adapted from figure 46tt of [34])

Finally, it is worth noting that to send a frame during the CP is more energy consuming than during the CFP, due to the fact that a non-AP STA's WNIC must be awake for longer periods in order to contend for channel access within the CP. Therefore, how non-AP STAs distribute different applications' data traffic using the different available access methods is crucial for the overall performance. The S-APSD TSs must be set appropriately in order to deliver frames associated with applications with special QoS requirement or for those applications that require exchanging large amounts of data packets. Examples of such flows are VoIP calls, streaming audio and video sessions, and large downloads. For instance, prior to initiating a download a non-AP STA negotiates the scheduling intervals associated with that flow, and then after a successful negotiation the AP reserves those intervals for that flow. Conversely, frames associated with applications that do not have special QoS requirements nor require exchanging large numbers of data packets, such as email, web browsing, messaging, and control management frames must be delivered using U-APSD or DFS during the CP. Therefore, when a data frame belonging to an active traffic stream arrives at the AP, it is first buffered, and subsequently delivered during the CFP of the correspondent SP associated with that TS, while buffered data frames that are not associated with any active flow are retrieved during unscheduled service periods or using normal DCF operations (in the case of non-QoS STAs) during the CP.

Section 3.1.3 describes the innovative idea called "seamless-S-APSD" that allows seamless handovers between APs.

3.1.3 Seamless handover mechanism: "*Seamless-S-APSD*"

IEEE 802.11 standard states that APs shall maintain power management status information for each currently associated station that indicates in which power mode each station is currently operating. In order to support S-APSD, APs shall also maintain schedule elements and TSPECs of active flows belonging to each currently associated station.

The main target of the proposed solution is to enable continuity of S-APSD active flows of a STA when that STA performs a handover to another AP. This feature reduces both delay and energy-consumption that the ADDTS re-negotiation procedure for each active flow would generate. This operation has been referred to as "seamless-S-APSD". The main challenges regarding this mechanism are presented below.

In order to enable the decision algorithm to make decisions based on QoS expectations that target APs offer, the TSPECs of the current S-APSD active flows belonging to the non-AP STA must be sent to the target APs while scanning. Then, the target APs should reply with the TSPEC that it can offer for each requested flow.

Once a handover candidate has been chosen, the non-AP STA again sends the agreed upon TSPECs of currently active flows to the target AP. In order to maintain these flows upon association with the new AP, the target AP replies with a TSPEC and SE for each accepted flow. The target AP uses a resource reallocation algorithm to reallocate each flow, as done in the normal S-APSD operation after receiving an ADDTS request.

On the other hand, the serving and candidate APs could use an Inter-Access Point Protocol defined in IEEE Std 802.11F^{TM}-2003[44] to check resource availability at candidate networks. Then, TSPECs concerning active flows associated to the STA can be interchanged between the APs across the distribution system. This fact enables the non-AP STA to avoid using wireless link to carry information regarding to its S-APSD active flows that is also known by the serving AP.

In order to reduce delays experienced by the active flows, the target AP starts to receive all of the data frames destined to the STA after it receives a handover request[*]. After the handover is performed, data frames belonging to accepted flows are delivered using S-APSD at recently scheduled SPs during the new AP's CFP. In contrast, frames that belong to refused flows and other frames are delivered using U-APSD.

The normal ADDTS Request-Response procedure is used to set up new flows once the STA is attached to the new AP. Note that this negotiation shall be done during the CP.

The proposed solution utilizes a set of new mechanisms in order facilitate seamless handovers. These mechanisms, further described in the following paragraphs, enable a STA to have information about nearby APs and perform scanning in the background.

3.1.4 Discovery information mechanism

The main objective of the discovery information mechanism is to allow non-AP STAs to have information about available nearby APs in order to simplify the scanning process. This mechanism enables APs to collect discovery information and deliver this information to its associated STAs.

In order to advertise its own information to the other APs, each AP broadcasts a new frame, called an *advertisement frame*, throughout the distribution system. This frame allows

[*] The details of directing copies of the data frame via the DS to the new AP are outside the scope of this thesis.

the other APs within the same DS to maintain up-to-date information about neighboring APs. Advertisement frames are broadcasted by APs after changes, but can be also broadcasted periodically to detect disconnected APs. The main information parameters shared by the APs via those advertisement frames are described below. Table 3-1 shows an example of the information that would be collected by the AP1 in Figure 3-1.

Table 3-1: Example of data that could be collected by the proposed discovery information mechanism

BSSID	CH	CP offset	Neighboring
BSSID2	6	CP_off_2	1 (YES)
BSSID3	11	CP_off_3	0 (NO)

BSSID MAC address of the AP's WLAN interface

Channel Operating channel of the AP

CPs offset Time instant when the CPs of the AP starts. This value is used by STAs associated with other APs while performing active scanning, in order to identify the beginning of the CPs of a target AP. As will be explained later, this offset is crucial for the performance of the active scanning procedure. Probe requests must be sent at the beginning of the target APs' CPs in order to minimize contention times, and therefore, scanning times. Further information regarding the active scanning mechanism is given later in this chapter. In order to distribute this parameter, the APs can include time stamps in the advertisement frames that allow other APs to obtain the aforementioned t_{cp}'s values. Note that a similar procedure is used to synchronize stations within an infrastructure BSS, as APs include time stamps in their beacon and probe response frames. These values are updated by the AP whenever they change, i.e. every time a scheduled flow is removed from the AP's CFP. It is worth noting that if a non-AP STA that is performing an scanning has an incorrect t_{cp} value(s) of target APs, then the probe requests sent to those APs may experience longer delays - due to longer waiting times caused by channel contention.

Neighboring An AP is considered a neighbor if it can be "sensed" by any associated STA. The neighbor report mechanism that allows an AP to identify its neighbor APs is explained later.

Finally, the rules that drive the delivery of discovery information by APs to its associated STAs are as follows:

- The whole information table about neighboring APs is delivered directly to a STA upon association. This information could be included in the association response frame.
- When an AP detects a change in any information field of the discovery information table, it should broadcast an update to its associated stations. Therefore, a new control frame is defined for this purpose. Note that broadcast traffic will be delivered using the procedure described earlier to indicate the presence of buffered broadcast frames in the DTIM of the beacon.
- Non-AP STAs keep their discovery information up-to-date, and use this information to enhance the scanning process (as will be described later).

3.1.5 Neighbor report mechanism

The purpose of this neighbor report mechanism is to allow an AP to identify which APs belonging to the same distribution system are within range of its associated STAs. These APs are potential handover candidates (i.e., target APs) for its associated stations. Figure 3-7 illustrates this idea. However, it might be possible that some of the APs present in the discovery information table are *not* within the coverage of a specific STA. If the AP has information about the approximate physical location of its associated nodes and the other APs, then this information could also be delivered to the associated STAs. Given this additional information a non-AP STA could start scanning for APs at the closest distance, thus avoiding scanning for APs that are farther away.

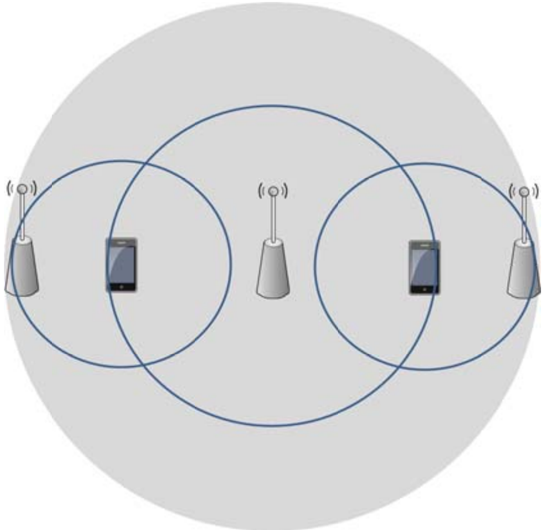


Figure 3-7: Neighboring APs: An AP is considered a neighbor when it is within the coverage of any associated station, i.e. those inside the big (gray shaded) circle in the figure

The neighbor report concept is already defined in the IEEE 802.11 standard. In the proposed solution, non-AP stations actively participate in this mechanism by sending neighbor report information about nearby APs to the associated AP, *after* a full scan has been performed. Note that a full scan should be performed by a non-AP STA after an *unsuccessful* active scanning.

3.1.6 Background selective active scanning

Background selective active scanning is the most important mechanism of the proposed solution. This allows a STA to greatly reduce its power consumption during the discovery phase of the handover procedure. Moreover, it enables “seamless-S-APSD”. The main characteristics of this scanning mechanism are:

Background Scanning is performed during periods when the station is idle, but after the decision algorithm detects that a handover is imminent. Therefore, the decision algorithm determines, based on criteria that are not covered here, when the mobile node shall start the *network discovery phase*. It is worth noting that the pitfall of possible miss-synchronization with the AP, which is common in other background scanning procedures, is not present in this mechanism. This pitfall is explained in section 5.2.7 of [45]. The reason that the solution proposed here avoids this problem is that the AP does *not* send frames to the STA during idle periods when the scanning is performed, because the AP is no allowed by the APSD operation to transmit to the STA during this time period – hence the STA cannot miss any frames sent to it by its current AP nor does it loose synchronization with its current AP. Furthermore, the scanning process does not influence active flows' perceived QoS since this scanning is performed by the STA only during idle periods.

Selective STAs only perform unicast single scanning for those APs present in the discovery information table, which is provided and updated by the current AP (as explained earlier). Note that some of those APs might not be in range of the STA (for details see the previous subsection).

Active Each time this mechanism is triggered a single active scan is performed. This scan consists of sending a probe request directly to a potential target AP, followed by waiting for a defined timeout period for a response.

This single active scanning operation is shown in Figure 3-8. The non-AP STA includes a TSPEC of each of its S-APSD active flows in the *probe request* that is sent to the target AP, and then it waits a period of time (probe timeout) for a response. It is worth remembering that each flow's TSPEC defines its QoS expectations. The probe timeout is variable, and depends on the number of active flows that the target AP has to process. The target AP receiving such a *probe request* uses its own resource allocation algorithm to compute a SE and other TSPEC parameters for each requested flow. Then, the AP sends a *probe response* including the TSPECs that it can currently offer for each of the requested flows. Notice that the *probe request/response* frames' formats need to be modified to incorporate these TSPEC elements.

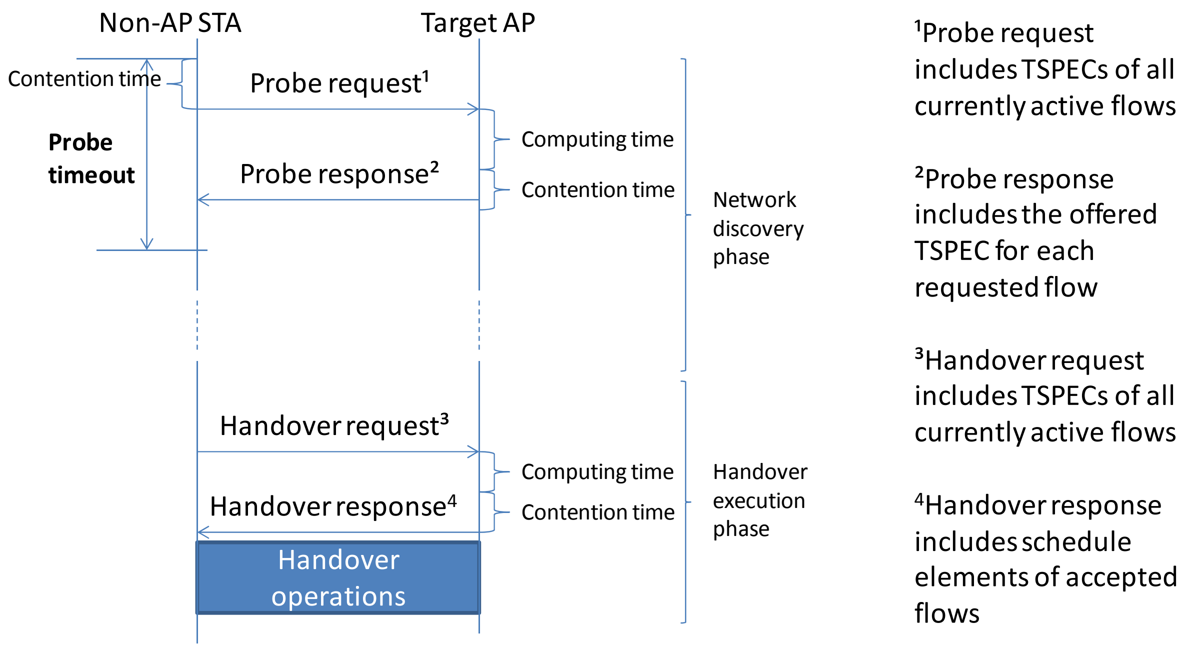


Figure 3-8: Active scanning operation

The probe timeout value can be determined by the *MinChannelTime* and *MaxChannelTime* values. The most appropriated values for those parameters concerning the proposed solution will be explained below, but the reader is referred to section 3 of [10] and [46] for further information.

$$MinChannelTime \leq probe\ timeout \leq MaxChannelTime$$

$$MinChannelTime \geq time\ to\ transmit\ a\ probe\ request + time\ to\ transmit\ a\ probe\ response$$

$$time\ to\ transmit\ a\ probe\ request \geq T_{switch} + DIFS + backoff\ time + transmission\ time$$

$$time\ to\ transmit\ a\ probe\ response \geq T_{switch} + DIFS + backoff\ time + transmission\ time$$

$$MaxChannelTime \leq \max(MinChannelTime) + probe\ computing\ time$$

These parameters are:

T_{switch} The switching time incurred while the STA is altering one channel to another. It is identified as negligible in [10], and varies between 40 and 150 µs.

DIFS The minimum waiting time for a frame to access the channel. Its value is specified in table 22-25 of [9] as 34µs.

Backoff time The backoff mechanism aims to randomize the period of time that a STA must wait prior to initiating a transmission while contending for the channel. This mechanism will be further explained in section 4.1.5. Moreover, it depends on the value of the contention window used. Probe and handover request/responses use a shorter contention window in order to minimize this backoff time. Its value is equal to: $CW \times slot\_time$. The CW's minimum and maximum values are 2 and 15 respectively. In [9], $slot\_time$ is specified as 9 µs. However, the optimal backoff time is difficult to determine because it depends highly on the number of STAs contending for the channel. Simulations concerning the results presented in section 0 suggest that this value has an upper bound of 2ms.

Transmission times Transmission times depend on probe request/responses size and data rates, but they can be considered negligible since their values will be shorter than 1 µs.

Probe computing time The probe computing time is the time that the AP spends to compute the probe request. Probe requests include TSPECs for active flows, and the AP must compute those TSPECs in order to include information about the service that it can give to the requesting STA. There is no information in the literature about this computing time, but for the purposes of this thesis 0.5 ms will be considered as an upper bound time to compute the relevant information about each flow.

Based on the explanation above, and in line with typical values used in [46], and [21], the probe timeout value that a mobile node must use is:

$$probe\ timeout = 5\ ms + number\ of\ active\ flows\ \times 0.5\ ms$$

The non-AP STA should use the timing information present in its discovery information table to send the probe request **at the beginning** of the CPs. Note that the STA needs to contend for channel access before sending this probe request. In order to improve the performance of the scanning procedure, probe requests must use a shorter backoff contention window (CW) than the contention window used by the stations currently associated with the target AP, thus giving priority to these probe requests. Sending probe requests at the beginning of the target AP's CPs using shorter CWs may lead to shorter scanning delays, but this will be evaluated later in section 4.2.1.1.

After a probe request has been sent to each AP present in the discovery information table, the non-AP STA has generated a list of handover candidates, based upon the probe responses that have been received. However, the decision algorithm might determine that a handover must be performed even if not all of the APs in the discovery information table have been scanned. At this point, the decision algorithm will decide which of these AP should be chosen as the target AP based on the TSPECs received in their probe responses, and other selection criteria. After this decision, the *handover execution phase* starts, and the non-AP STA sends a *handover request* to the target AP, including the previously negotiated TSPEC for each currently active flow. Note that such a control frame is *not* defined in the standard, but could be easily defined as an extension of the *association request* control frame. After receiving this *handover request* the AP again computes the TSPECs and generates a SE for each accepted flow. These SEs are included in the *handover response*. Moreover, the AP reserves these scheduled periods, and starts receiving and buffering all of the data frames destined to the newly associated station. Starting to receive and buffering all of the data frames for this STA may be possible since both APs share the same distribution system. While this cannot be considered a soft handover, as the non-AP STA does not receive data frames from both APs at the same time, it does reduce the delay experienced by S-APSD active flows. After all of the handover operations have been performed, the non-AP STA's WNIC shall wake up at the new SIs to retrieve/send data frames associated with the accepted flows from the new AP.

3.1.7 Overall operation example

This section contains an example of the proposed solution's operation. This scenario consists of five access points and two non-AP stations, called MN1 and MN2. MN1 is initially associated with AP1 and has three S-APSD active flows reserved: flow 1 (*F1*) is set in downlink direction with SI equal to 50 ms; flow 2 (*F3*) is set in uplink direction with SI equal to 50 ms; and flow 3 (*F3*) is set in downlink direction with SI equal to 100 ms. For instance, *F1* and *F2* could be associated to a VoIP call, while *F3* could be associated to a download. Moreover, MN2 is initially associated with AP2, and also has some S-APSD active flows. Figure 3-9 illustrates this scenario and Figure 3-10 shows the packet exchange. Note that sent and received frames are drawn filled and unfilled respectively. This example considers only **two** CFP and one CP per beacon interval, in order to simplify the explanation.

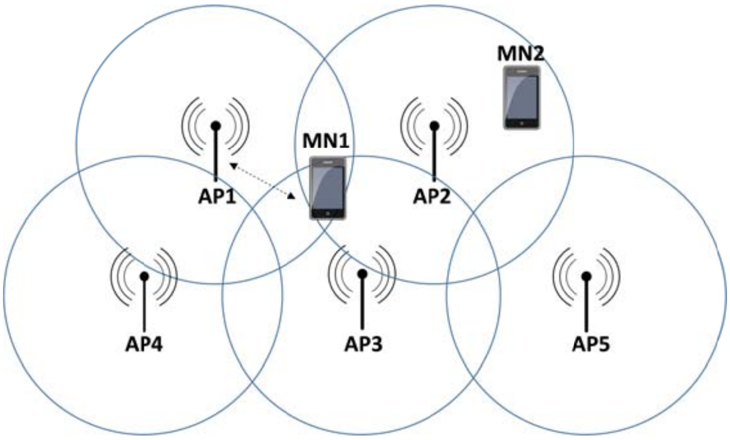


Figure 3-9: Overall operation scenario

Initially, we assume that some frames that are not associated with any S-APSD active flow (e.g. frames belonging to email client application) destined to MN1 and MN2 have been received at AP1 and AP2 respectively. Therefore, these frames have been buffered at both APs. This fact is then advertised by AP1 in the TIM element of its beacon frame, and these buffered frames are retrieved by MN1 using U-APSD during the first AP1's CP1. On the other hand, MN2 retrieves it buffered frame during the first AP2's CP2.

We assumed that MN1 has three S-APSD active flows: F1, F2, and F3. As a consequence, the corresponding SPs are allocated during the CFPs of AP1. Moreover, additional S-APSD traffic that is not shown in the figure causes each AP's beacon interval to be divided into two CFPs and two CPs as well. The instants when each CP of every AP's starts are tagged in the figure as $t_{cpY\_ap\_X}$.

At instant **t1**, the decision algorithm of MN1 decides that it must start the *network discovery phase* looking for potential handover candidates. The discovery information that MN1 has at this moment is shown in Table 3-2.

Table 3-2: The discovery information that MN1 has at time t1

BSSID	CH	CP offset
BSSID2	6	CP_off_12
BSSID3	11	CP_off_13
BSSID4	7	CP_off_14

Note: AP5 is **not** present because it is **not** announced by AP1 since it is **not** considered a neighbor (i.e. no neighbor report about that AP has been received)

Next, MN1 sends a probe request to each target AP at the beginning of one of its CPs. Note that there is no response from AP4 due to the fact that it cannot receive the probe request sent by the STA – as it out of the communication range of MN1. MN1 listens to each channel until the probe timer expires of the probe response is received. The situation represented in Figure 3-10 is quite ideal because MN1 is idle during each target AP's t_{cp}, so MN1 can send these probe requests. This situation does not always occur, because a STA might not be able to send a probe request at these instants because it is already busy. However, if the beacon interval is divided into four periods as described earlier, then it is much more likely that STAs can send a probe request at these instants since there are four different CPs' initial instants in every beacon interval. Another solution would be to announce in the advertisement frames not only the beginning of CPs, but also its duration. Therefore, STAs could send probe requests at any instant within the target APs' CP. The performance of these different solutions will be evaluated later (see section 4.2.2).

Notice that at the initial instant of the first AP2's CP2 (i.e. $\mathbf{t}_{CP2\_AP2}$) both MN1 and MN2 attempt to send a probe request and a trigger data frame respectively to AP2. Then, contention occurs and MN1 gets firstly access to the CH6 due to the fact that the CW used for that transmission is shorter than the one used by MN2 to send the data frame.

Finally, at instant **t2**, the decision algorithm commands the MN1 to perform a handover to AP2. The decision about *which* target AP to associate with is out of the scope of this proposed solution. Moreover, such a decision should be based on multiple criteria, but basically the decision is derived from the QoS expectations received in the probe responses. In response to the handover request sent by the MN1, AP3 sends a handover response which includes the new SEs for the accepted flows as previously negotiated. These SEs specify the SST associated with each S-APSD active flow, along with the SIs. After MN1 has successfully associated with AP3, it wakes up at the new SST associated to F1 in order to retrieve frames associated to F1.

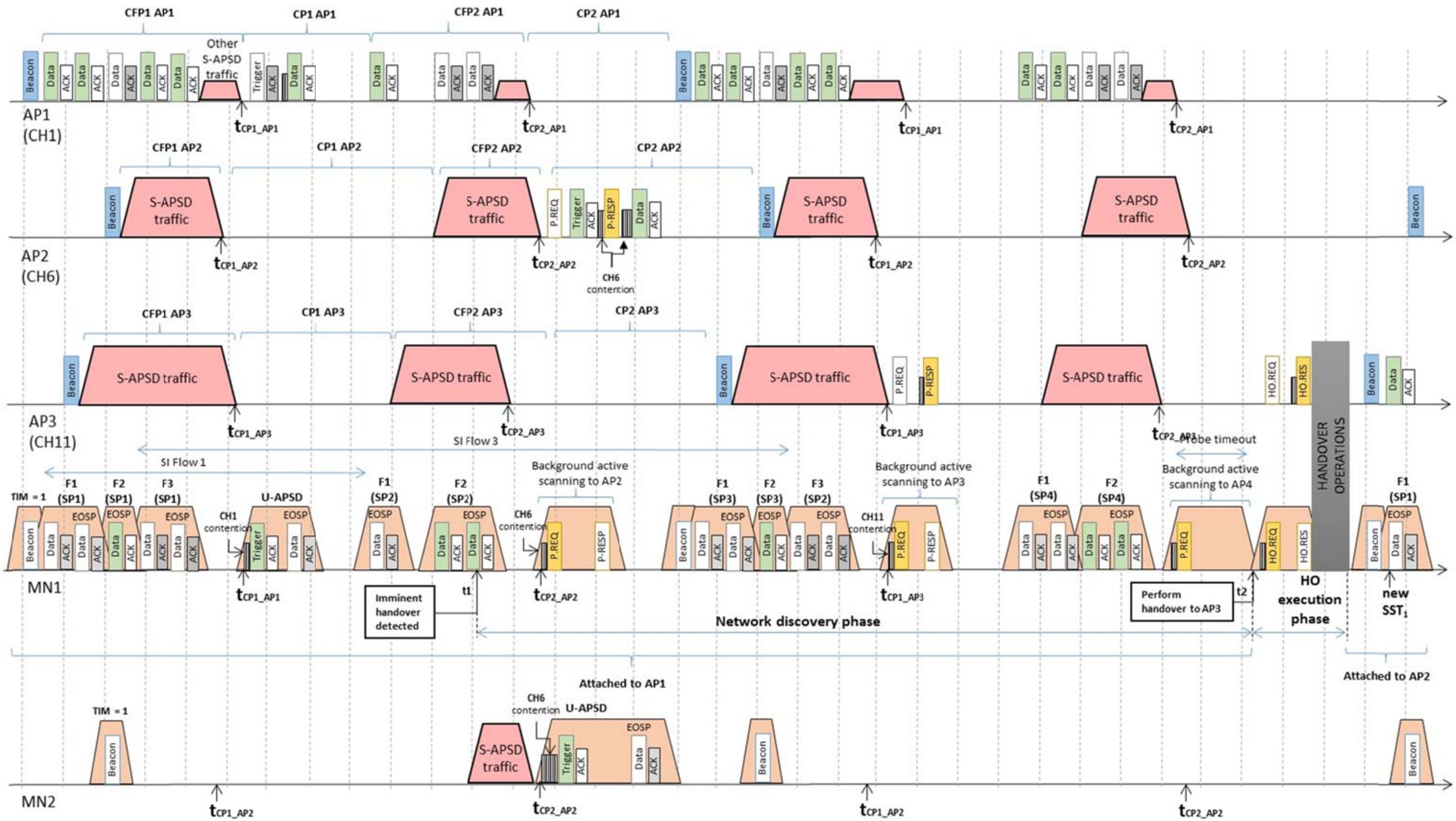


Figure 3-10: Overall operation example

3.2 *Next step* towards cellular / WLANs interworking

This section describes an integrated scenario for Wi-Fi/cellular networks interworking based on the IEEE 802.21 standard [23]. Figure 3-11 illustrates the IEEE 802.21-based network architecture that provides MIH services within a heterogeneous network.

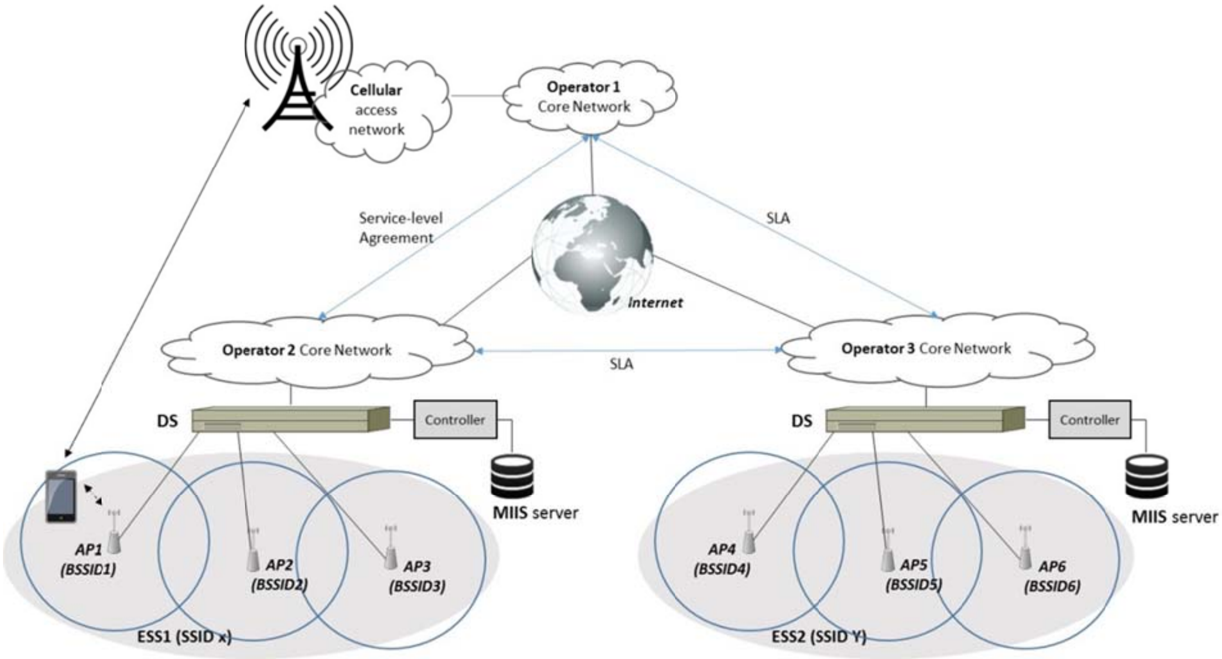


Figure 3-11: IEEE 802.21-based heterogeneous network

The heterogeneous network consists of the following components:

Single-ESSs	Each ESS's operation is based on the proposed solution. The MIISs, which are present in both networks, contain handover-related information.
Cellular network	The cellular network can access MIH information found in other networks' MIISs according to service-level agreements between network operators.

Handover (HO) related information within a single-ESS is collected by the controller, which is connected to the DS, and stored in the MIIS server. The controller collects HO-related information found in the *advertisement frames* that are broadcasted by the APs through the DS. HO-related information may include available APs in this network, QoS capabilities, operator policies, and security related information.

The proposed heterogeneous network architecture provides IEEE 802.21 functionalities that include:

- Remote access from an external network to the MIIS server of another network in order to get HO-related information.
- Dialogue between the current serving network and other candidate networks' PoAs, so as to interchange MIH messages such as resource availability queries, HO preparation requests, and HO initiation commands.

The most interesting feature provided by this heterogeneous architecture is that it allows MNOs to control, monitor, and predict devices' choice of connectivity. This mode of operation is known as *network-initiated handover*, and was illustrated in Figure 2-6 on page 17. The main phases of this operation are summarized below.

i. The serving network gets information about available resources by querying the MIIS server of another network. HO-related information found in the MIIS may contain information about PoAs in the vicinity, operator policies associated with those PoAs, and security issues.

ii. The MN receives the HO-related information from the serving PoA, and the MN can include a list of its preferred PoAs in its response.

iii. The serving network dialogues with one or more candidate PoAs to check the availability of resources. After the decision about the target PoA is made, the serving network requests the target PoA to prepare resources at the target network *prior* to handover initiation.

iv. Finally, the serving network commands the MN to perform a handover to the specified network. After that, the new serving network commands the previous serving network to release resources allocated to the MN.

Regarding the proposed solution, resource availability can be checked in terms of S-APSD active flows. Therefore, prior to the initiation of a handover to a Wi-Fi AP, the serving network must check if the candidate network is able to serve the currently active flows associated with this MN (e.g. active VoIP calls, video streaming sessions, or downloads). Furthermore, during the handover preparation phase, the serving and target network must negotiate S-APSD active flows parameters (i.e. TSPECs and SEs), and the target network must reserve resources associated with active flows *prior* to the handover initiation. Consequently, seamless handover can be achieved based on the continuity of S-APSD active flows as proposed above.

In addition, the IEEE 802.21-based architecture is high flexible and provide many more features. For instance, it allows a MN to perform a handover based on HO-related information found in MIIS servers. This mode of operation is referred to as *mobile-initiated handover*, and consists of:

i. The MN queries HO-related information from a MIIS server of another network.

ii. Through the serving network, the MN dialogues with other candidate networks' PoAs to get information about resource availability.

iii. Once the target network is chosen, the MN commands its serving network to prepare a handover to a specified network. This operation includes resource preparation at target network.

iv. Finally, the MN initiates a handover to the target network, and then commands the previous serving network to release allocated resources associated to the MN.

In summary, the main challenges for cellular and Wi-Fi network integration that this heterogeneous network architecture addresses are:

- MNOs would be able to get real-time information about available resources in the vicinity of the MNs.
- The serving network would prepare vertical handovers in advance, through a dialogue with candidate networks' PoAs in order to enable seamless handovers.
- MNOs would control devices' choice of connectivity. Moreover, decision-making processes could be based on a complex set of criteria such as real-time information about the service-levels offered by candidate networks.

- In the case of *network-initiated handover*, the involvement of mobile devices in handover operations would be minimized, thus achieving large energy-savings.

4 Analysis

This section consists of a simulation based analysis of the proposed solution's performance. This simulator has been developed in MathWorks' Matlab®. The first section of this chapter describes the simulator that has been designed for this thesis project, while the second section contains some of the main results obtained by using this simulator.

4.1 Description of the simulator

The main goal of the simulator is to evaluate the performance of the proposed solution, as well as to detect faults in its design. As my starting point, I used the simple mobile wireless network present in [47]. Due to the fact that I have proposed a new version of the MAC protocol and a set of innovative mechanisms to improve both the scanning process' power consumption and performance, I have made a lot of improvements in this simulator in order to emulate the proposed MAC protocol's operations.

The following subsections give a simple description of the simulator that was used as a starting point, as well as all the improvements that have been made in order to evaluate the proposed solution's innovative operations.

4.1.1 Simple mobile wireless network simulator

The simple mobile wireless network simulator, available as a SourceForge project [47], is an event-based simulator. It is in the public domain and its original author is anonymous. It allows simulation of a simple packet exchange from the application layer of a sending mobile station to the same entity at the receiving station. It provides three different radio propagation models: free space, two-ray, and lognormal shadowing. Moreover, SNR-based packet capture, broadcast, and dynamic transmission rate and power are provided as physical layer mechanisms. The MAC layer includes some of the mechanisms described in the IEEE 802.11 standard, such as CSMA/CA, virtual carrier sense, and RTS/CTS. Additionally, the simulator includes some additional simple mechanisms at higher layers that I have not used since this thesis project focused on MAC layer operations.

The improvements made in order to adapt this simulator to the proposed solution are:

Network topology The network architecture of the simulator has been improved to simulate infrastructure networks, i.e. to differentiate between APs and non-AP STAs.

Physical layer A simple improvement has been made to allow transmissions in different channels. However, this addition does not consider interference between different channels.

MAC layer Most of the improvements have been made in the MAC layer, in order to emulate the proposed MAC protocol described in section 3.1.1. These improvements include: CFP-CP differentiation within a beacon interval, beacons, S-APSD and U-APSD traffic generation, and active scanning procedures. Furthermore, two different versions of the MAC layer have been implemented: in the first version, referred to as *one-period* version, every beacon interval is divided into one CFP and one CP; in the second version, referred to as the *four-periods* version, every beacon interval is divided into four CFPs and four CPs. This idea was illustrated in Figure 3-4 on page 32. The following descriptions of the different operations focus on the *four-periods* version. Section 4.2.1 contains simulation results of the *four-periods* version and section 4.2.2 consists of a quantitative comparison between both versions.

Outputs	The simulator gives several outputs after every execution, such as energy consumption of each node. Those outputs are further explained in section 4.1.8.
Multiple executions	The simulator allows the user to run several executions sequentially. Some parameters can be modified from one execution to another, in order to evaluate the proposed solution's performance under different conditions. Moreover, the simulators' output of every execution includes computations of average, minimum, and maximum values, as well as the cumulative distribution function of every simulation's output.

These and other improvements are further described in detail in the following subsections.

4.1.2 Topology

In order to emulate infrastructure networks, STA are differentiated between APs and non-AP STAs. For each AP it is possible to modify the number of associated non-AP STAs (also called MNs), as well as the physical location of every station. Moreover, the operational channel of each AP can also be modified.

4.1.3 Beacon generation

Beacon frames are broadcasted by APs every beacon interval. The offset for the first beacon of each AP (i.e. time between the initial instant and the first beacon of the AP) is calculated as a random value between 0 and 100ms. Moreover, the beacon frame's size has been set to 80 octets.

4.1.4 S-APSD traffic generation

The simulator allows the user to modify the number of currently active flows associated with each MN, as well as their TXOPs. In the case of the *four-periods* version, each flow has an associated SI that can be a multiple of $\frac{beacon\_interval}{4}$. However, this simulator only contemplates three possible SI values: 25, 50, and 100 ms, since the beacon interval is set to 100 ms.

The S-APSD SPs are allocated following the grouping approach described earlier in section 3.1.1. A smart grouping approach has been implemented, which allocates every active flow's SP in the different CFPs, respecting each flow's SI and minimizing each CFP's duration. In order to show the operation of that algorithm, an example is provided in the following. Table 4-1 contains the information about active flows given to the grouping allocation algorithm. It contains the mobile node that the flow belongs to, the associated AP, and flows' SIs and TXOPs. The number of flows of each MN and their associated SIs were generated randomly, while the TXOP parameter was fixed to 2 ms for all the flows in order to simplify this explanation. Table 4-2 contains the generated allocation that contains the different SST values associated with each flow. Moreover, it is illustrated as well in Figure 4-1. These SSTs correspond to the time distance between the instant when the beacon frame is sent and beginning of the flow's first SP. Notice that these SPs are all allocated within the AP's CFPs, whose starting times have an offset of 0, 25, 50, and 75 ms respectively from the AP's beacon frame.

Table 4-1: Random input that was given to the grouping allocation algorithm

Flow	MN	AP	SI (s)	TXOP (s)
1	2	1	0.025	0.002
2	2	1	0.1	0.002
3	3	1	0.025	0.002
4	3	1	0.05	0.002
5	3	1	0.1	0.002
6	4	1	0.025	0.002
7	4	1	0.05	0.002
8	4	1	0.1	0.002

Table 4-2: Grouping algorithm's output

Flow	MN	AP	SST (s)	CFP	SI (s)	TXOP (s)
1	2	1	2.00E-05	1	0.025	0.002
2	2	1	0.00202	1	0.1	0.002
3	3	1	0.00402	1	0.025	0.002
4	3	1	0.02702	2	0.05	0.002
5	3	1	0.05202	3	0.1	0.002
6	4	1	0.00602	1	0.025	0.002
7	4	1	0.00802	1	0.05	0.002
8	4	1	0.03302	2	0.1	0.002

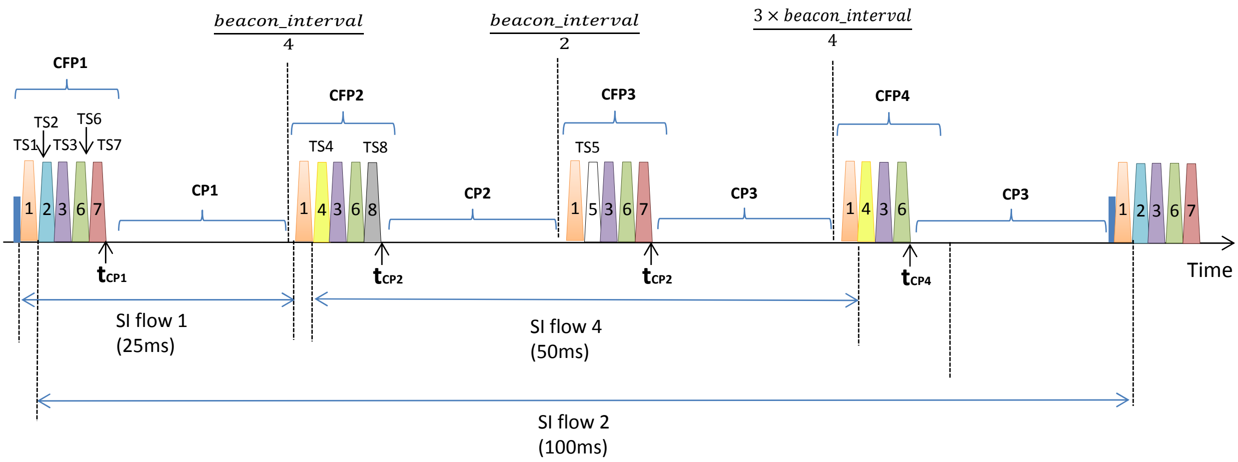


Figure 4-1: Illustration of the grouping algorithm's output shown in Table 4-2

As was illustrated in Figure 4-1, CFPs start at fixed instants within a beacon interval. Furthermore, the first CFP within a beacon interval starts just after the transmission of the beacon frame. The end of the last S-APSD service period of every CFP determines the beginning of the next CP. Those transition instants between CFP and CP are denoted as t_{cp1}, t_{cp2}, t_{cp3}, and t_{cp4} respectively. Furthermore, their values, i.e. the CFPs' duration, cannot be longer that 25 ms. On the other hand, the sum of all CFP's durations cannot be longer than a defined maximum CFP duration, which has been set to 70 ms, in order to allow other operations within the CPs.

The operation within scheduled SPs is shown in Figure 4-2. Channel contention is not needed since these SPs are scheduled and reserved for the associated MN. Moreover, the interframe spacing between data packets and acknowledgments is equal to the short interframe space (SIFS). SIFS is the smallest gap that can be used between transmissions of different stations. During this gap other stations will **not** try to access the channel. Each flow has a defined TXOP, which determines the maximum duration of its SP. In order to simplify the simulator, at each SP either the AP or a MN is randomly selected to be the sender of data frames. Each of these data frames should be acknowledged by the receiver. Furthermore, the selected sending station sends a new data frame after receiving the acknowledgment for the previous data frame until it notices that the next data-ACK transmission would exceed the defined TXOP.

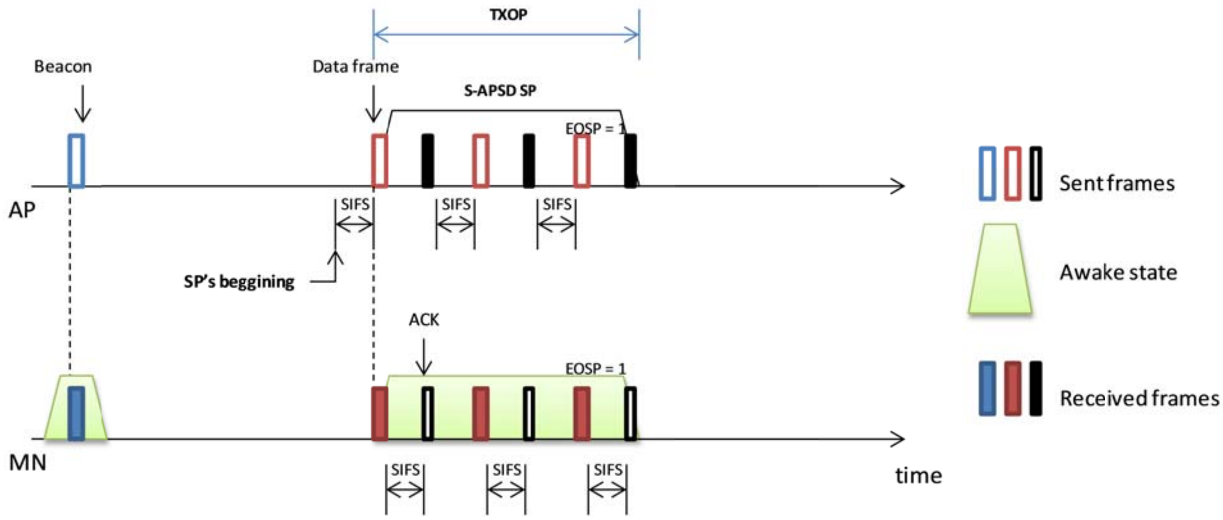


Figure 4-2: Operation within S-APSD SPs

The influence of different data packets' size on the performance during the CFP will be analyzed in section 4.2.1.2.1.

4.1.5 Active scanning procedure

The simulator allows the user to select those instants when handovers are imminent for every MN, i.e. when a MN should start the *network discovery phase*, and therefore, send probe requests to the target APs. During this phase, a *probe request* must be sent to every neighboring AP found in its network discovery information whenever it is possible. In order to send such a *probe request*, the MN starts contending for the target AP's channel at any of its t_{cp} instants. Note that it is possible that a MN is unable to send probe requests to an AP because the MN is within a scheduled SP at the beginning of all CPs of the target AP. The average probability for a MN to be able to send these probe requests will be analyzed in section 4.2.1.1.3. Moreover, it is possible that some of these APs are not within the MN´s range. There is a probe timeout that starts when the MN starts to contend for the channel. If there is no answer prior to this timeout, then the scanning attempt finishes unsuccessfully. The MN continues to perform other operations, including other scanning attempts to the rest of APs at their t_{cp} instants.

Figure 4-3 shows the IFS and contention associated with this procedure. Note that this figure shows only one scanning attempt, but the procedure might consist of one scanning attempt for each neighboring AP, plus a *handover request/response* interchange with the selected AP sent during the *handover realization phase*. The MN that wishes to send a *probe request* starts contending for the channel at a certain target AP's t_{cp} instant. After sensing the channel is idle for a DCF interframe space (DIFS) time, the MN starts a backoff counter. The backoff time associated with the backoff mechanism is calculated as:

$$backoff\ time = random(0, CW) \times slot\_time$$

The random value used is an integer between zero and the current value of the contention window (CW). The CW's value is initially equal to *CWmin*. Furthermore, the value doubles each time the backoff counter expires and the channel is busy. The maximum value is *CWmax*. The *CWmin* and *CWmax* values associated with *probe* and *handover requests/response* are 2 and 15 respectively. The CW's size for these *probe* and *handover requests/responses* should be smaller than the CW used for data frames in order to give priority to these probe requests. Moreover, the backoff counter decrements only when the STA senses that the channel is idle.

A *handover request* is sent to one of the APs whose *probe response* has been received just after the last scanning attempt has finished. Moreover, the time period between the last scanning attempt and the *handover request* can be set to a value different from zero. Concerning proposed solution's actual operation, the *handover request* is sent as soon as the decision algorithm determines that the MN shall start the *handover initiation phase*. This means that the *network discovery phase* is finished, but not all of the neighboring APs need to have been already scanned.

Regarding the *probe* and *handover requests* and *probe responses*' size, all of these need to include a TSPEC element for each active flow associated with the MN. However, *handover responses* also include the SE for each active flow. The *handover request/responses*' size has been set as a typical *association request/responses*' size plus the TSPEC's and SE's sizes. The resulting size is calculated:

$$probe\_request\_size = basic\_probe\_request\_size\ +\ n\_active\_flows \times TSPEC\_size$$

$$probe\_response\_size = basic\_probe\_response\_size\ +\ n\_active\_flows \times TSPEC\_size$$

$$handover\_request\_size = association\_request\_size +\ n\_active\_flows \times TSPEC\_size$$

$$handover\_response\_size = association\_response\_size +\ n\_active\_flows \times SE\_size$$

$$Where, basic\_probe\_request\_size = 56\ octets, basic\_probe\_response\_size = 85\ octets,$$

$$association\_request\_size =\ 60\ octets, and\ association\_response\_size = 58\ octets.$$

$$TSPEC\_size =\ 57\ octets, \qquad SE\_size\ = 14\ octets.$$

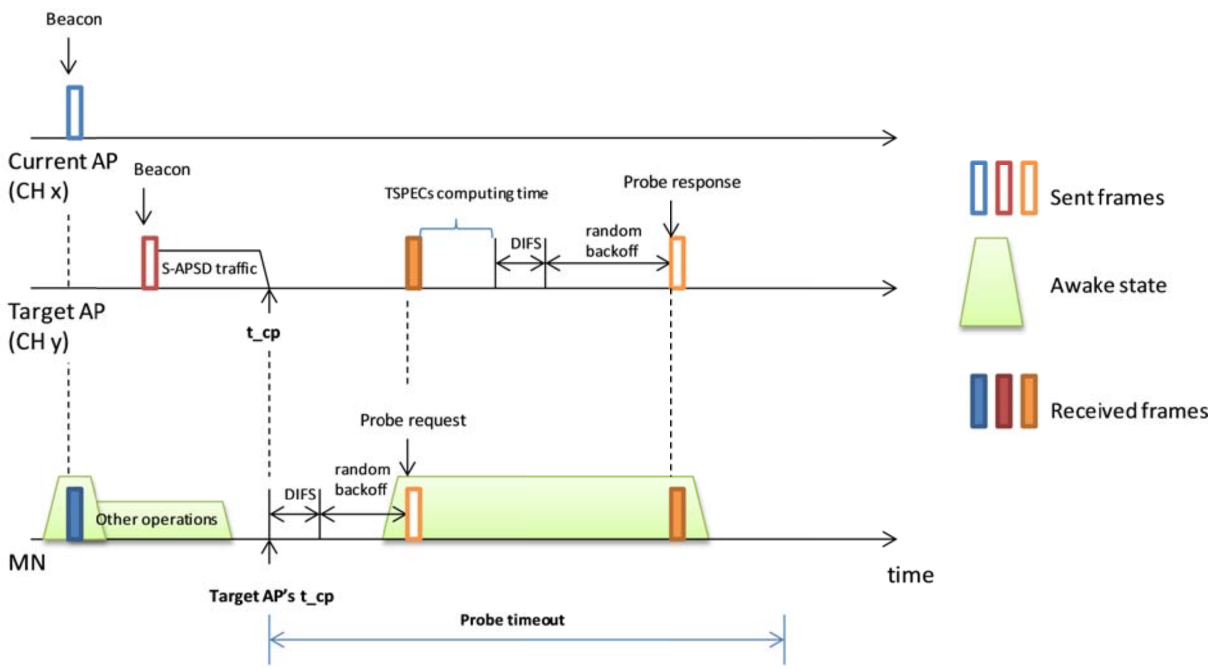


Figure 4-3: Contention and IFS regarding a single unicast scanning attempt

4.1.6 U-APSD traffic generation

The simulator generates a random TIM element at every beacon, which contains a list of those MNs that have pending frames at the AP. These MNs's WNICs wake up at one of the next four t_{cp} instants of its associated AP to retrieve these buffered frames. In this simulation one of the four possible t_{cp} instants in the next beacon interval is randomly chosen by each MN. The amount of buffered frames that each MN can retrieve during an unscheduled SP is fixed, as will be explained in the next paragraph. Following receipt of the TIM element the MNs starts contending for the channel at the chosen t_{cp} instant in order to send a trigger frame to its associated AP so as to trigger a new unscheduled SP. This contention consists of a DIFS period plus a random backoff time. The backoff time is calculated as explained earlier, but using a longer CW. The values used for *CWmin* and *CWmax* for data frames are 15 and 1023 respectively. Note that these and other values are specified in the IEEE 802.11 standard in the MIMO PHY characteristics subsection included in section 20 of the standard.

Figure 4-4 shows the contention and IFS associated with U-APSD traffic. There is an initial CP prior to the initiation of the unscheduled SP. Afterwards there is no contention within the SP, as SIFS is applied between data frames and acknowledgements to avoid other STAs attempting to access the channel. The SP's duration is determined by the *max_SP_length* parameter, which can be set. This parameter determines the maximum number of data frames that can be sent within an unscheduled SP. Moreover, the SP's end is announced by the AP when it sets the EOSP field of the last data frame to 1. Note that this operation realizes a simplified version of U-APSD traffic generation since the AP is always the sender of data frames within an unscheduled SP and it *always* sends as many frames as specified in max_SP_length parameter.

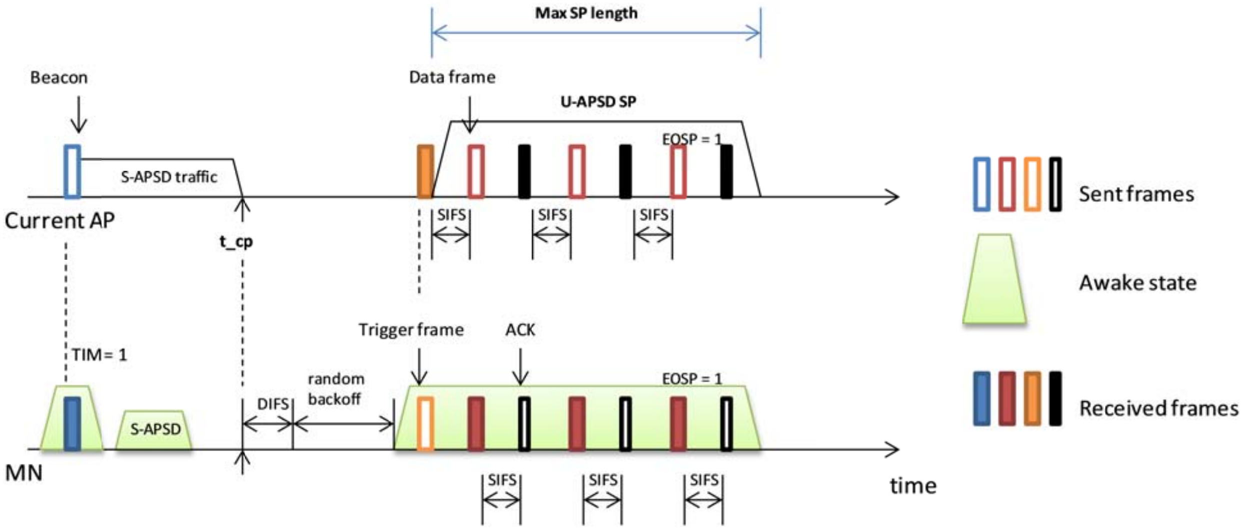


Figure 4-4: Contention and IFS associated to U-APSD traffic

The simulator allows the user to run several executions sequentially. It is possible to modify the value of some parameters in order to simulate different situations. Furthermore, the result of every execution is saved by the simulator. The following subsection consists of a summary of the different input parameters that can be modified. Later, section 4.1.8 explains the outputs or results given by the simulator after every individual execution, and after several sequential executions.

4.1.7 Summary of input parameters

This subsection presents the set of parameters that can be modified in the simulator, along with their typical values. These parameters and their values are shown in Table 4-3.

Table 4-3: Simulation input parameters

	Parameter	**Comments and typical values**
Topology	Number of APs	-
	Operational channel used by each AP	Non overlapping and no interference between channels is considered
	Number of MNs associated with each AP	-
	Physical location of every node	-
S-APSD	Number of active flows associated with each MN	-
	TXOP or SP’s duration of each flow	1 – 10 ms
	Service interval of each flow	25, 50, or 100 ms
	Data frames’ size	1500 octets
	ACK frames’ size	14 octets
	Maximum cumulative CFP duration	70 ms
Beacons	Beacon interval	100 ms

	Parameter	Comments and typical values	
	Beacon offset associated with each AP	Random value between 0 and 99 ms from the initial instant	
	Beacon frames' size	80 octets	
U-APSD	TIM element at each beacon instant	Random. The probability of whether a MN has buffered frames at the AP at every beacon instant is uniformly distributed therefore.	
	Data frames' size	150 octets	
	ACK frames' size	14 octets	
	Unscheduled SP length (max_SP_length)	4 – 8 frames	
Active scanning	Probe timeout	5 ms + n_active_flows x 0.5 ms	
	TSPECs computing time	0.5 ms per each active flow	
	Probe requests' size	56 + n_active_flows x 57 octets	
	Probe responses' size	85 + n_active_flows x 57 octets	
	Handover requests' size	60 + n_active_flows x 57 octets	
	Handover responses' size	58+ n_active_flows x 14 octets	
IFS **(See table 20-25 of the standard IEEE802.11-2012)**	SIFS	16 µs	
	Slot time	9 µs	
	DIFS	SIFS + 2 x slot_time	
	Clear channel assessment (CCA) time	4 µs	
	Short CW_min (Out of the standard)	2	(Applied to probe requests/responses and handover requests/responses)
	Short CW_max (Out of the standard)	15	
	Long CW_min	15	
	Short CW_max	1023	
	Default rate	48 Mbps	

4.1.8 Simulator's outputs

First, the simulator shows the data packets that are sent and received by AP and non-AP STAs. In Appendix A, an example of this output is shown. This output was generated by a simple simulation of four APs (nodes 1, 2, 3, and 4) with one mobile node associated with each AP (nodes 5, 6, 7, and 8). Moreover, AP1, AP2, AP3, and their associated nodes are all within the range of each other, meanwhile AP4 and its associated mobile node are out of range of the other nodes. Every MN has three active flows, with service intervals of 25 ms, 50 ms, and 100 ms respectively. Those flows have a quite short TXOP that allows interchanging only one data frame. Furthermore, mobile node 5 starts its *network discovery phase* at the instant 0.1 seconds.

The simulator also generates graphical output that presents the power states during the simulation for each MN. The graphical output generated for node 5 is shown in Figure 4-5. Figure 4-6 has a horizontal zoom applied in order to show the power states concerning scanning operations that start just following 0.1 seconds.

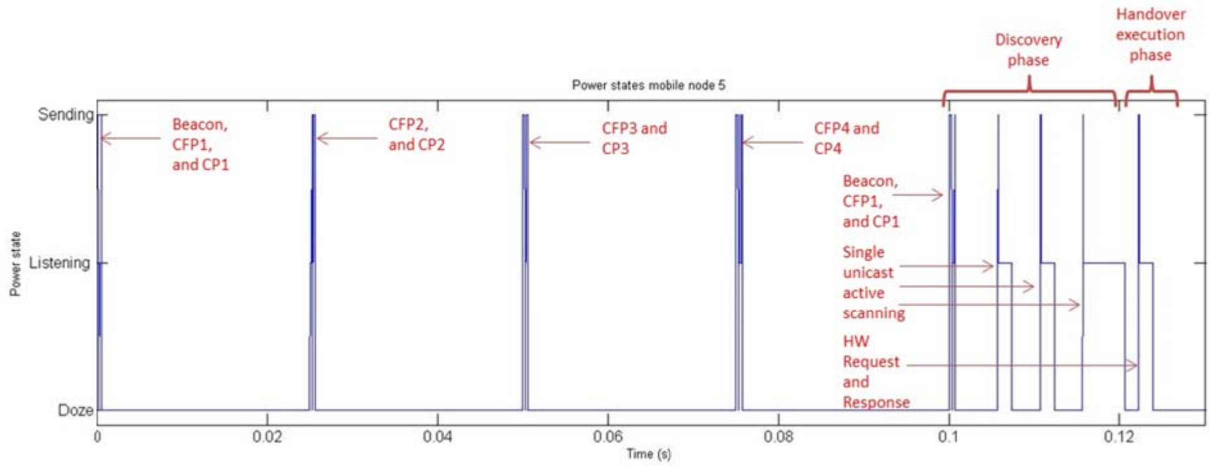


Figure 4-5: Example of a MN's power states graphical output from the simulator

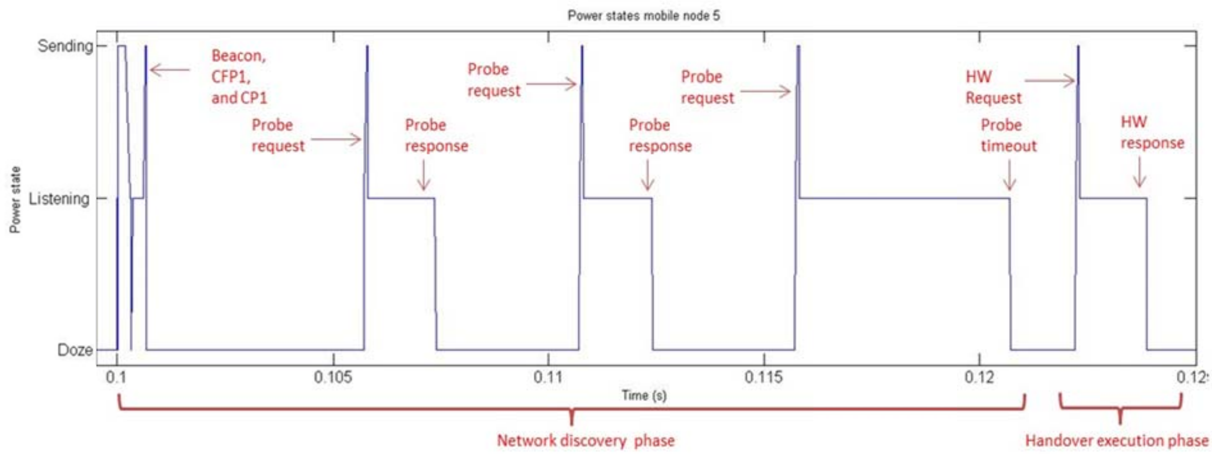


Figure 4-6: Power states graphic with horizontal zoom applied

Moving to quantitative results, the simulator provides the amount of time that each MN has been in doze, listening, and sending states. Moreover, it gives *doze*, *listening*, and *sending times* regarding scanning, S-APSD, and U-APSD operations. This information is used to calculate the *energy consumption* of each node, as well as *energy consumption per bit*.

In order to calculate the energy consumption, the values present by Tsao and Huang in Table 3 of [8] have been used. Thus, the average power consumption in different states that the simulator uses to obtain nodes' power consumption is shown in Table 4-4.

Table 4-4: Devices' power consumption in the different power states

Doze state	RX state	TX state
10 mW	424 mW	484 mW

Concerning the scanning process, the simulator calculate a *scanning latency* for each MN, which is the amount of time from the beginning of the *network discovery phase* (i.e. the instant when the MN starts sending probe requests) to the end of the *handover execution phase* (i.e. when the handover response is received from the selected AP). This scanning latency is expected to be higher when the number of APs to be scanned increases. In addition, the energy consumption of the scanning process is computed by considering scanning listening and sending times together with the power consumption values in Table 4-4. Listening times for the scanning process will increase as the number of MNs associated with the target AP increases. This leads to more MNs contending for the channel at target APs' t_{cp} instants, and therefore, to longer CPs. Finally, due to the fact that a MN could be busy at every t_{cp} instant of a target AP, so that it could not scan that AP, the simulator computes a *probe success rate*. This rate is the number of probe response received divided by the number of APs that are within range of the MN (i.e. APs that are able to receive probe requests sent by the MN). Therefore, this probe success rate is expected to be lower when the MN that is scanning has several active S-APSD flows with longer TXOPs (i.e. a greater intensity of S-APSD traffic).

Regarding S-APSD, the simulator computes the *bandwidth efficiency during the CFPs* as the average value of the bandwidth efficiency of each S-APSD flow, which is determined as follows:

$$BW\ Efficiency\ per\ flow = \frac{\frac{S-APSD\ data\ bit\ sent}{Channel\ busy\ time}}{data\ rate} = \frac{\frac{S-APSD\ data\ bit\ sent}{\left(\frac{TXOP}{SI}\right)*total\ time}}{data\ rate}$$
$$= \frac{S-APSD\ data\ bit\ transmited}{Amount\ of\ bit\ that\ could\ be\ transmited\ during\ the\ time\ that\ the\ channel\ is\ busy}$$

Finally, in relation to U-APSD the simulator provides the *efficiency during the CPs*, which is calculated as follows:

$$CP\ Efficiency\ per\ node = \frac{\frac{U-APSD\ data\ bit}{time\ awake\ during\ the\ CPs}}{data\ rate}$$
$$= \frac{U-APSD\ data\ bit\ transmited}{Amout\ of\ bit\ that\ could\ be\ transmited\ during\ the\ time\ that\ the\ MN\ was\ awake}$$

Notice that the time that a mobile node spends awake during the CPs consist of listening and sending times regarding U-APSD operations, which includes contention, IFSs, and acknowledgments Furthermore, such listening times will be longer when the number of MNs associated with the same AP increases. This assumes that the number of MNs contending for the channel at every t_{cp} increases, which causes longer CPs. Therefore, the efficiency during the CPs is expected to be lower when the number of MNs attached to the AP increases.

The reader will find more information about the simulator's outputs in the following section.

4.2 Simulation results

This section contains the simulation results and conclusions. Several different simulation configurations have been used to evaluate the performance of the proposed solution under different circumstances. Since scanning, S-APSD, and U-APSD operations' performance depends on different parameters, they have been evaluated separately through different simulations. Moreover, **each simulation configuration was run fifty times** in order to obtain average values and cumulative distribution functions of the simulator's outputs. It is worth noting that for a single execution, to simulate one second of real-time takes roughly 30 seconds in normal traffic conditions, but up to 5 minutes under high traffic conditions. Cumulative distribution functions of the results presented in this section can be found in Appendix B.

The results regarding the *four-periods* version (i.e., the beacon interval is divided into four CFPs-CPs) are shown in the first subsection. The second subsection compares the performance of both versions (i.e. *four-periods* versus *one-period*). Finally, the last subsection compares the performance of the *four-periods* version with other solutions found in the literature.

4.2.1 Results concerning the *four-periods* version

As it was mentioned earlier, scanning, S-APSD, and U-APSD operations' performance has been evaluated separately through different simulations. In the following subsections, such simulations and their associated results are explained in detail.

4.2.1.1 Scanning operation's performance

A set of simulations were used to evaluate the scanning operation's performance under different circumstances. The first simulation configuration evaluates the influence of the number of APs that the MN must actively scan. Then, the second one aims to estimate the influence of the number of MNs associated with the same AP. Finally, scanning's performance is evaluated when the MN that is scanning has different S-APSD traffic intensities.

4.2.1.1.1 Influence of the number of APs to be scanned

According to the proposed solution, the decision algorithm indicates to the MN the instant when it shall initiate the *network discovery phase*. After that, the MN starts performing single unicast active scanning attempts in the background. Then, the decision algorithm can indicate to the MN that the handover must be performed based on the information currently collected, even if all the APs found in the network discovery information list have been not yet been scanned. Therefore, the most suitable APs would be placed at the beginning of the discovery information list, so that those APs would be the first to be scanned. As was explained in section 3.1.4, it is also possible that some of the APs found in the discovery information list are not within the coverage of the MN.

The number of APs to be scanned evidently influences the total scanning latency and energy consumption. In order to evaluate that influence two different simulation configurations have been used. In the first one, all the APs to be scanned are within the range of the MN that is scanning. Meanwhile, in the second some of the APs are *not* within the scanning MN's range, which leads to longer listening periods waiting for a probe response in those cases when the AP did not receive the probe request.

In the first simulation configuration, in which all the APs are within range of the scanning MN, the number of APs to be scanned varies from 1 to 12. S-APSD traffic conditions are randomized for every execution following these settings: the number of active flows associated with each MN can be from 0 to 6, their SIs can be 25, 50, or 100 ms, and their TXOPs can range from 0.5 to 1 ms. Concerning U-APSD traffic, the TIM is estimated randomly every beacon interval, as was explained in section 4.1.6, and the unscheduled SP length is equal to 5 frames.

Figure 4-7 and Figure 4-8 show the average, minimum, and maximum values of the scanning latency and MN's energy consumption after fifty executions of this simulation configuration. In addition, the cumulative distribution function can be found in Appendix B.

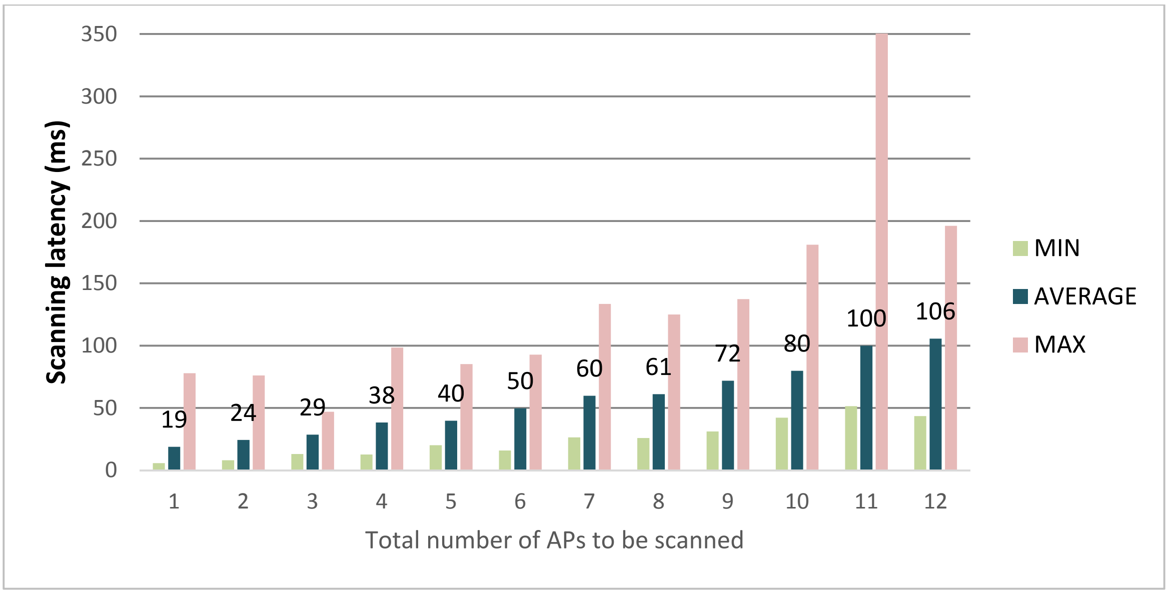


Figure 4-7: Influence of the number of APs to be scanned over the scanning latency (The average scanning latency in ms is shown above each set of three columns, i.e. for a given number of APs to be scanned.)

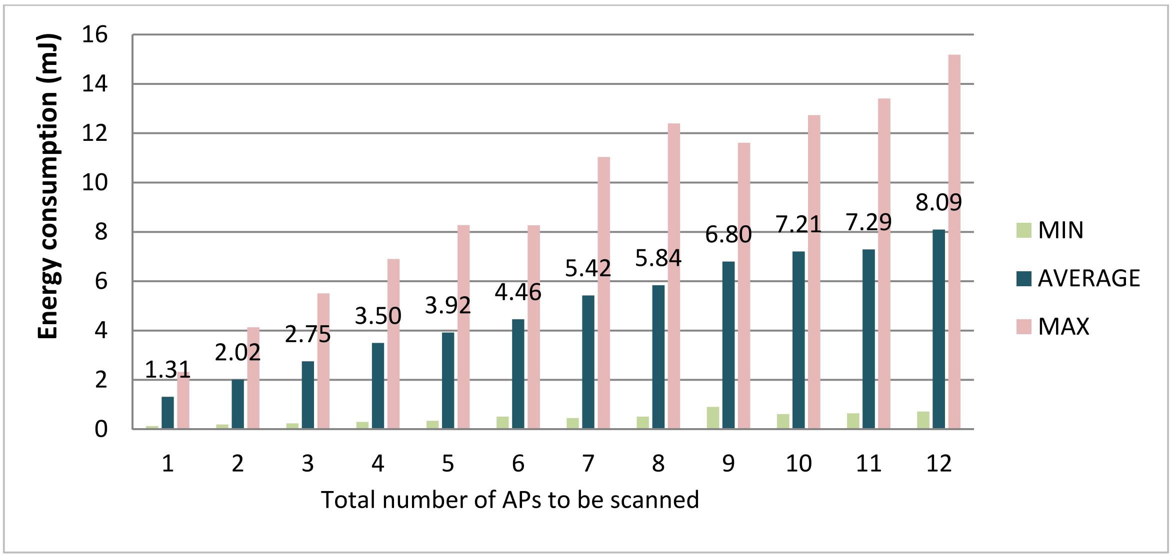
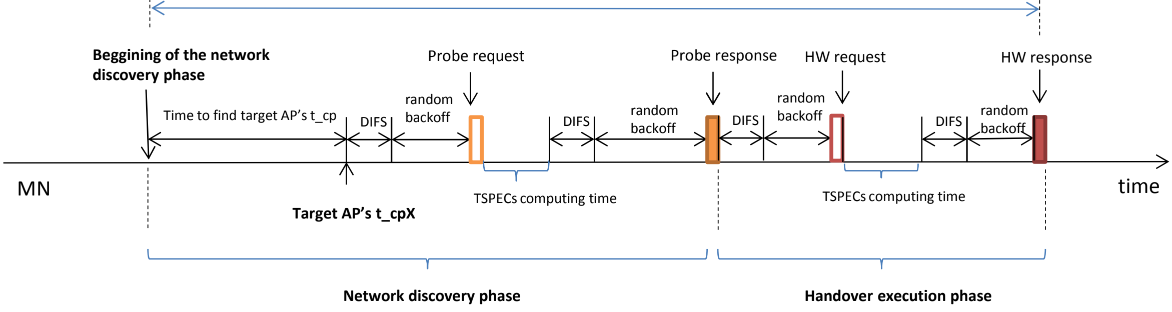


Figure 4-8: Influence of the number of APs to be scanned over the energy consumption (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of APs to be scanned.)

The average value of the scanning latency varies from 19 to 106 ms when the number of APs to be scanned is 1 and 12 respectively. In order to understand those values, Figure 4-9 illustrates the timing associated with a scanning operation that consists of one single unicast active scanning during the network discovery phase, plus the handover execution. Note that " ≅ " denotes average value.

$$Handover\ latency \cong 12.5 + 4 = 16.5ms$$



$$Time\ to\ find\ the\ target\ AP's\ t_{cp} \in (0{,}25ms) \cong 12.5ms$$

$$4 \times DIFS + 4 \times backoff + 2 \times TSPECs\ computing \cong 4ms$$

$$DIFS = 34us$$

$$\text{Backoff} = random(0, CW) \times slot\_time = random(0{,}2) \times 9us \cong 9us$$

$$TSPECs\ computing = number\ of\ active\ flows \times 0.5ms \cong 3 \times 0.5ms = 1.5ms$$

Figure 4-9: Handover operations' timing concerning one single unicast active scanning, plus the handover execution

In the second simulation configuration, the number of APs that are within and outside the MN's range varies from 1 to 6, so that the number of APs that must be scanned varies from 2 to 12 with half of the APs within and outside the MN's range. The S-APSD and U-APSD conditions are the same as in the previous simulation configuration. Scanning latency and energy consumption results are shown in Figure 4-10 and Figure 4-11, and the cumulative distribution functions can be found in Appendix B.

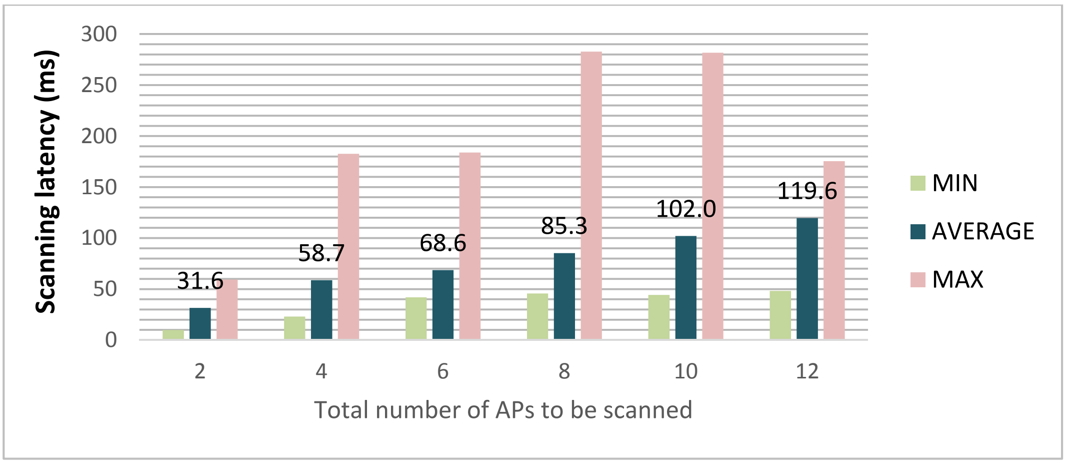


Figure 4-10: Influence of the number of APs to be scanned over the scanning latency, when half of the APs are out of the scanning node's coverage (The average scanning latency in ms is shown above each set of three columns, i.e. for a given number of APs to be scanned.)

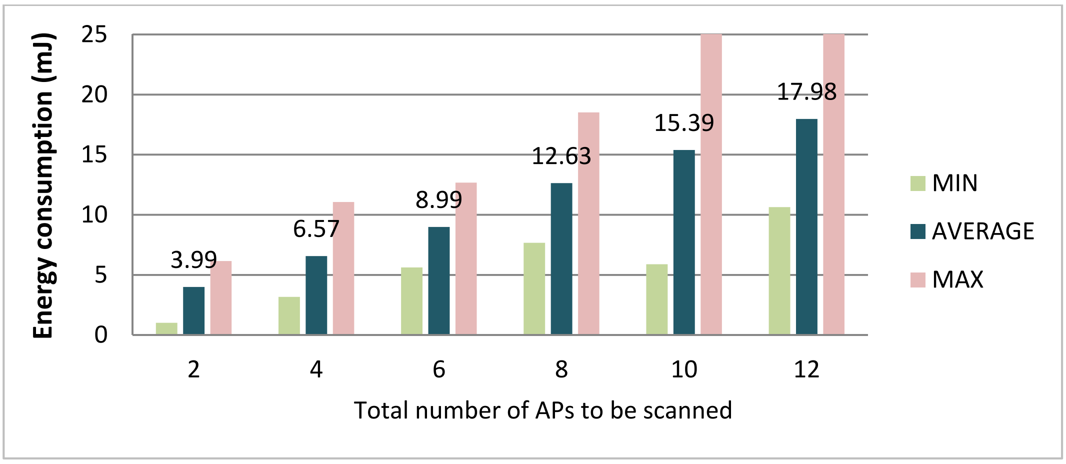


Figure 4-11: Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's range. (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of APs to be scanned.)

Comparing the results obtained from both simulation configurations, we conclude that scanning's performance is penalized considerably when the MN has to scan APs that are out of its range, even when short probe timeout values are used. Therefore, MNs should start scanning for the closest APs, while avoiding scanning for those that are far away. In order to do so, the currently associated AP should have information about the approximate physical location* of its associated nodes and the physical location of other APs, and then the discovery information given to a certain MN could be both ordered and filtered.

* Note that here were are assuming that it is physical distance that has the largest effect upon the signal strength on the up and down links of the MN. This may not be true, especially indoors; however, it is left for future work to consider more detailed propagation modeling and use of knowledge about the radio environment of the MNs and APs.

4.2.1.1.2 Influence of the number of MNs attached to the target APs

The number of MNs associated with the target AP is proportional to the number of MNs contending for channel access at the target AP's t_{cp} instants. Even though probe responses use a shorter contention window than the other frames, when the number of MNs contending for the channel increases, the probe requests are likely to experience longer backoffs. The following equations show the minimum and maximum possible backoff durations regarding probe/handover requests/responses and other frames respectively.

$$backoff = rand(0, CW) \times slot_{time} = rand(0, CW) \times 9us$$

$$Probe/Handover\ frames:$$

$$Initial\ CW: CWmin = 2 \rightarrow backoff \in (0, 18us)$$

$$Maximum\ CW: CWmax = 15 \rightarrow backoff \in (0, 135us)$$

$$Other\ frames:$$

$$Initial\ CW: CWmin = 5 \rightarrow backoff \in (0, 18us)$$

$$Maximum\ CW: CWmax = 1023 \rightarrow backoff \in (0, 9.207ms)$$

This simulation configuration consists of 4 APs, i.e. the MN shall scan 3 APs, and all of them are within the MN's coverage. Moreover, S-APSD and U-APSD traffic are randomized in the same way as in the previous simulation configurations.

Figure 4-12 and Figure 4-13 show the energy consumption and listening times associated with the node that performs scanning operations as the number of MNs per target AP increases. Figure 4-14 and Figure 4-15 show those results but in this case the frames associated with the scanning operations use the same CW values as other frames. In the second case, we can observe that listening times are longer when the number of mobile nodes associated with the target APs increases due to longer backoff periods. Therefore, the MN's energy consumption associated with scanning operations is higher. However, in the first case energy consumption is less affected by the number of MNs. Consequently, shorter CWs should be used for scanning frames in order to improve the scanning operations' performance.

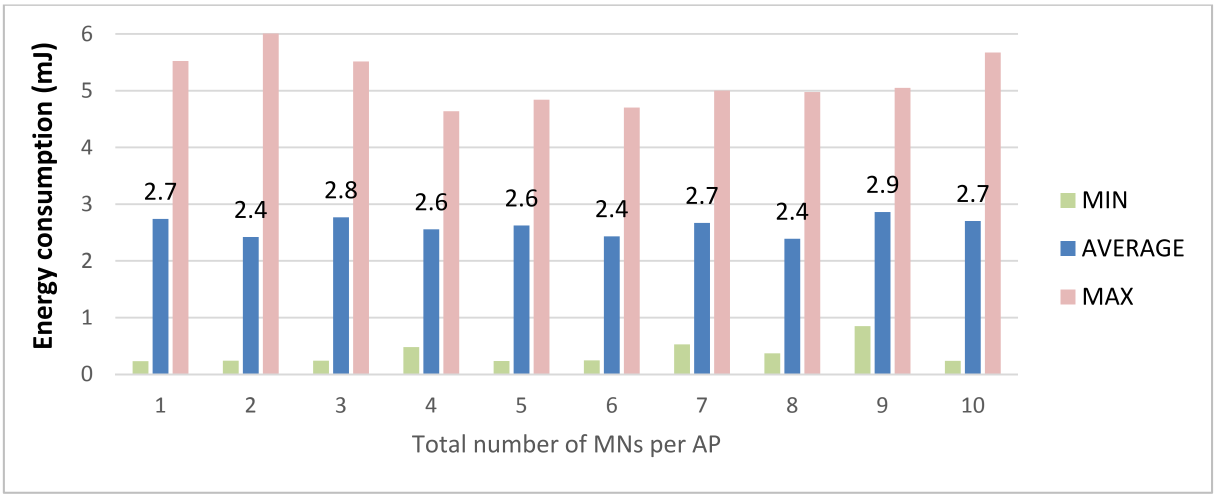


Figure 4-12: Influence of the number of MNs associated with the target AP on the scanning MN's energy consumption (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given total number of MNs per AP.)

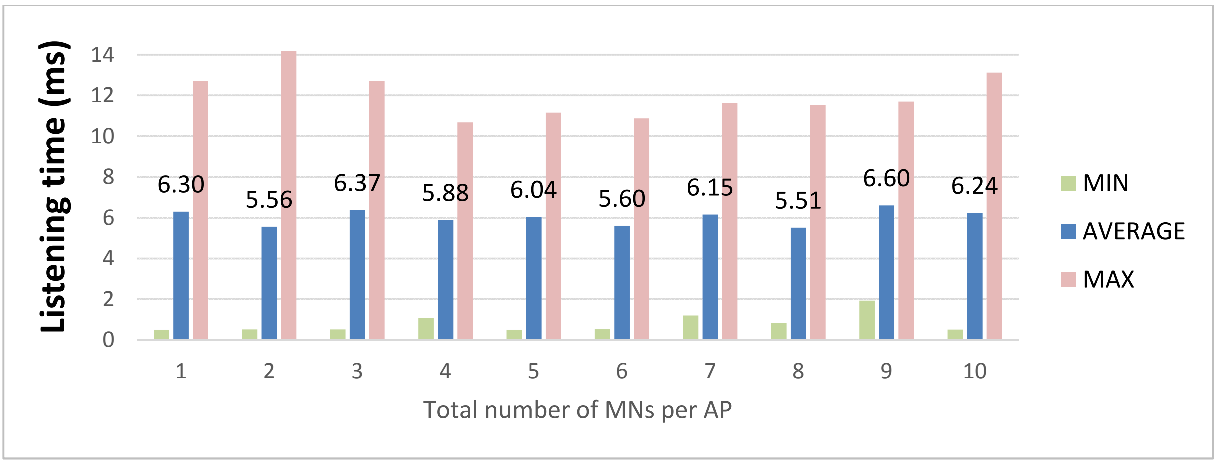


Figure 4-13: **Influence of the number of MNs associated with the target AP on the listening times associated to scanning operations (The average listening time in ms is shown above each set of three columns, i.e. for a given number of MNs per AP.)**

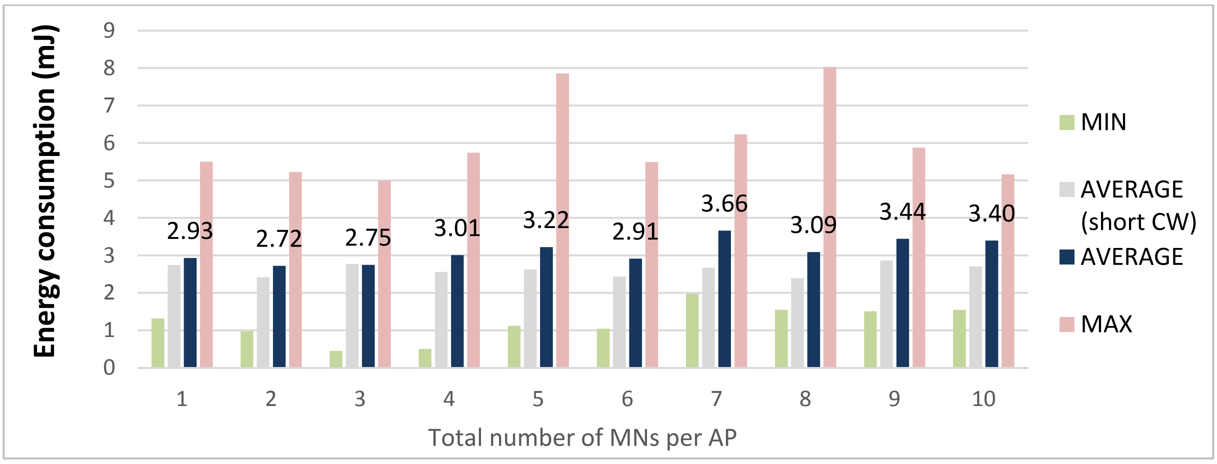


Figure 4-14: **Influence of the number of MNs associated with the target AP on the scanning MN's energy consumption when frames associated with scanning operations use the same CW as other frames (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of MNs per AP.)**

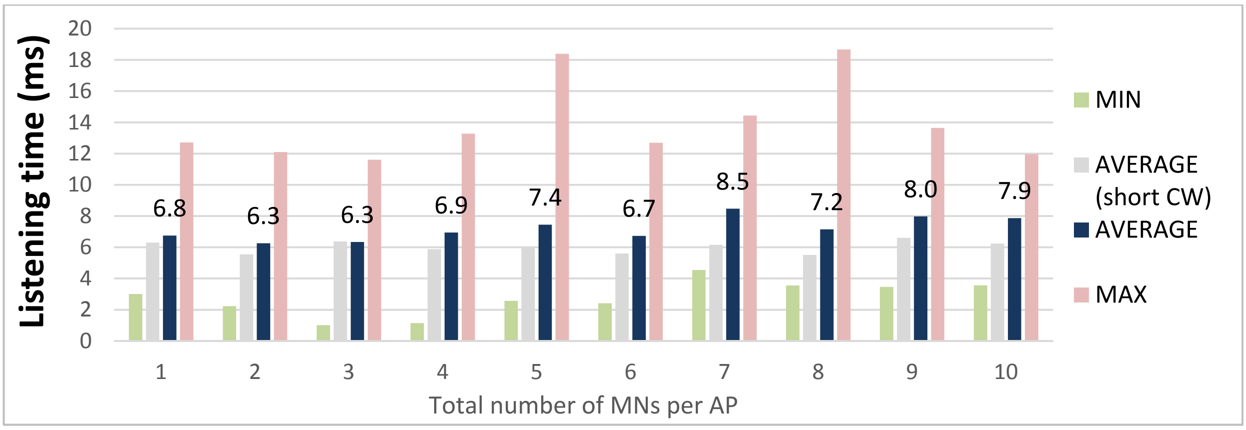


Figure 4-15: Influence of the number of MNs associated with the target AP on the listening times associated to scanning operations when frames associated to scanning operations use the same CW as other frames. (The average listening time in ms is shown above each set of three columns, i.e. for a given number of MNs per AP.)

4.2.1.1.3 Influence of S-APSD traffic intensity at scanning node

The S-APSD traffic intensity associated with the scanning node, i.e. the number of active flows and their TXOPs, affects the scanning performance as the MN that performs the scanning is busy for longer periods of time. The probability that a MN is able to successfully send a probe request to a target AP at one of the target AP's t_{cp} instants (referred to as *probe success rate*) decreases with increasing traffic intensity.

The fact that the background scanning does not influence the S-APSD active flows' achievements is one of the strengths of this proposed solution. However, the fact that a MN shall send probe request at the target AP's t_{cp} instants reduces contention times, and as a consequence also reduces energy consumption. However, when the *probe success rate* becomes extremely low then we should consider allowing MNs to send a probe request at any time during the target AP's CPs. To do so the MNs need to be aware of each target AP's CP durations. This information could be provided by the AP in the discovery information list.

Figure 4-16 shows the evolution of the *probe success rate* for a set of simulation configurations in which the S-APSD traffic intensity associated with the scanning node has been modified. The basic simulation configuration consists of four APs, one MN associated with each AP, and the same S-APSD and U-APSD conditions as the previous simulation configuration. In addition, S-APSD traffic conditions associated with the scanning node are as follows: number of active flows varies from 1 to 8, each flow has TXOP of 1 ms, and SI is 25 ms.

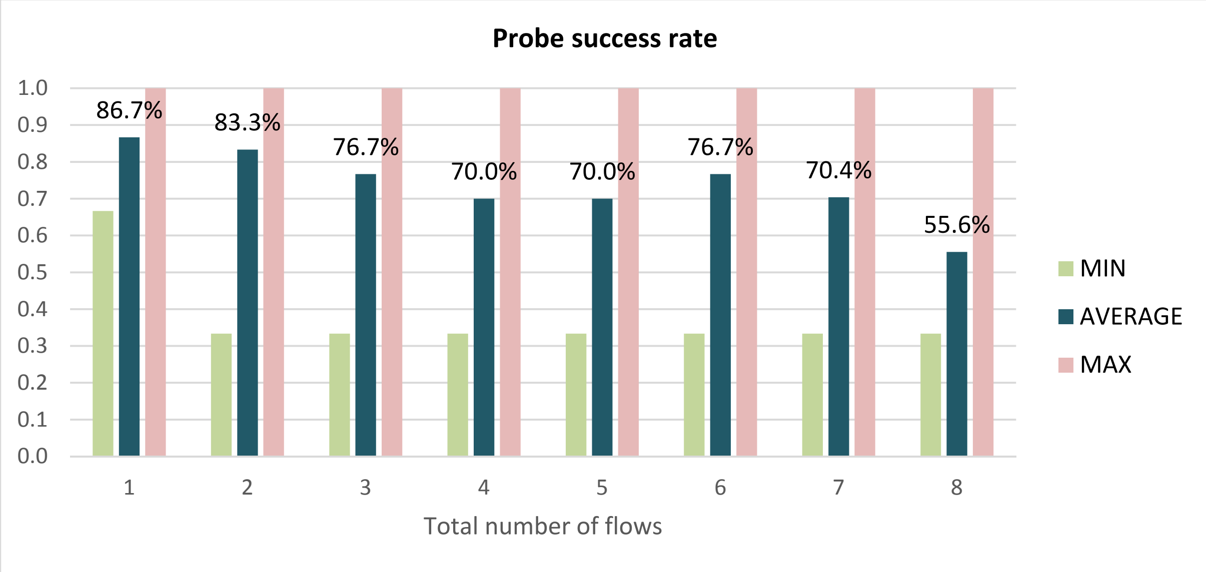


Figure 4-16: **Probe success rate over different S-APSD traffic conditions (The average probe success rate is shown above each set of three columns, i.e. for a given number of flows.)**

The *probe success rate* becomes dramatically low when the scanning node has high S-APSD traffic intensity. Notice that the total throughput achieved by 8 flows using TXOP and SI equal to 1 ms and 25 ms respectively is around 10 Mbps. The following equation illustrates this:

$$Throughput\ per\ flow = efficiency \times \left(\frac{TXOP}{SI}\right) \times data\ rate \cong 0.7 \times \left(\frac{1}{25}\right) \times 48Mbps$$
$$= 1.344Mbps$$

$$8\ flows \times 1.344\ Mbps = 10.752\ Mbps$$

On the other hand, the total number of active flows associated with the scanning node also influences scanning performance. Probe/Handover request/responses carry information about each active flow, so their size depends on the number of active flows. In addition, each TSPEC must be processed by the target AP that receives the probe response, so it takes more time for that AP to send back a probe response when the number of active flows increases. Figure 4-17 shows the scanning node's energy consumption regarding scanning operations in this simulation configuration.

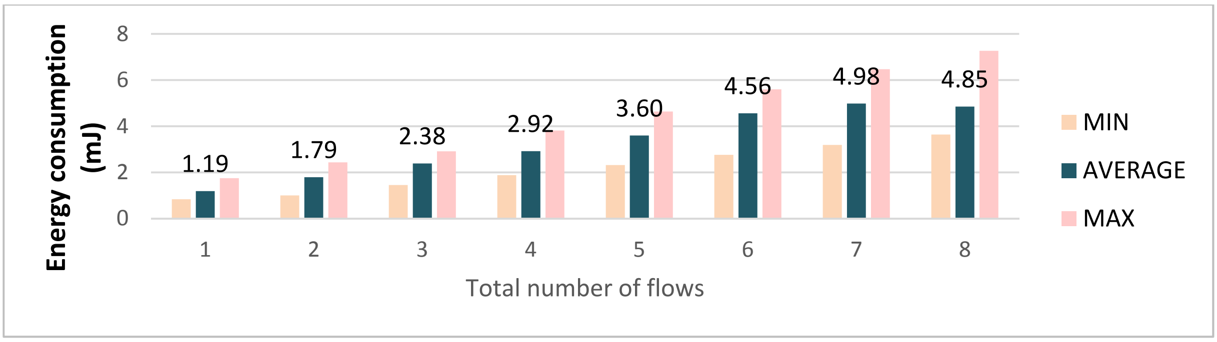


Figure 4-17: **Energy consumption of the scanning node as a function of the number of active flows (The average energy consumption in mJ is shown above each set of three columns, i.e. for a given number of flows.)**

4.2.1.2 S-APSD operations' performance

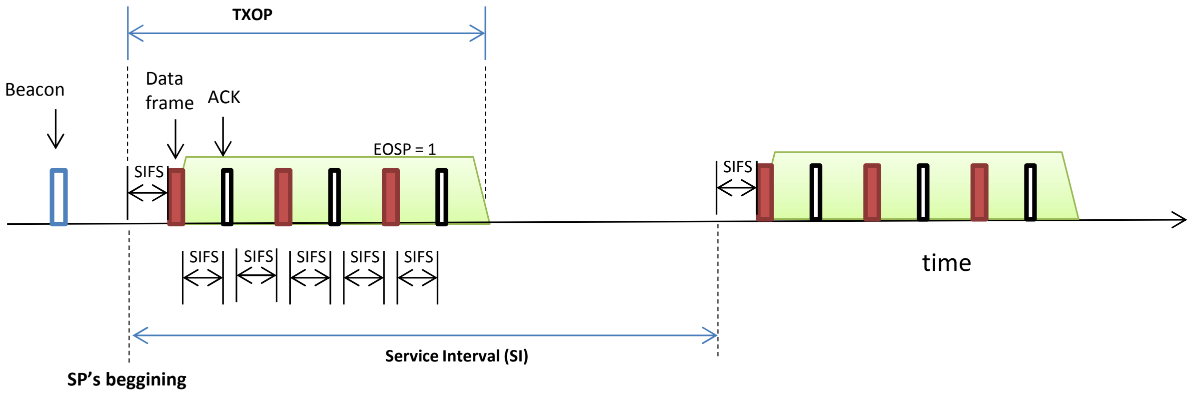
There are three main strengths of the propose solution's S-APSD operations:

- Neither background nor U-APSD/DCF traffic during the CPs influence the QoS of S-APSD flows.
- Channel contention is unneeded and the shortest IFS can be used within scheduled SPs enabling efficient use of bandwidth resources.
- S-APSD active flows allow transmissions with low energy consumption since MNs are only awake at scheduled instants associated with their active flows.

In order to evaluate S-APSD operations' performance, two different simulation configurations have been used. The first simulation configuration evaluates the *bandwidth efficiency* achieved by each node using S-APSD. The reader is referred to section 4.1.8 for more information about how the simulator estimates this parameter. The the second simulation configuration evaluates throughputs and energy consumption per bit achieved by S-APSD flows. Results from both simulation configurations are presented in the following subsections.

4.2.1.2.1 Influence of S-APSD data packet's size over the bandwidth efficiency

Although a single *flow's bandwidth efficiency* was defined in section 4.1.8, it deserves further explanation. Figure 4-18 illustrates this concept.



$$flow's\ BW\ efficiency = \frac{S - APSD\ data\ bits}{\left(\frac{TXOP}{SI}\right) * data\ rate}$$

Figure 4-18: Flow's bandwidth efficiency concept

As a result increasing the data packets' size, adapting flows' TXOP to the data packet's size, and other mechanisms (such as block acknowledgment as defined in the standard [9]) lead to higher bandwidth efficiency, thus reducing the overhead due to IFSs and ACKs. In order to evaluate the bandwidth efficiency achieved by S-APSD flows in the propose solution a simulation configuration has been constructed where the data packet's size varies from 250 to 2000 octets. The TXOP associated with every SP of each flow is a random value uniformly distributed between 2 and 3.5 ms. Figure 4-19 shows the results of this simulation configuration.

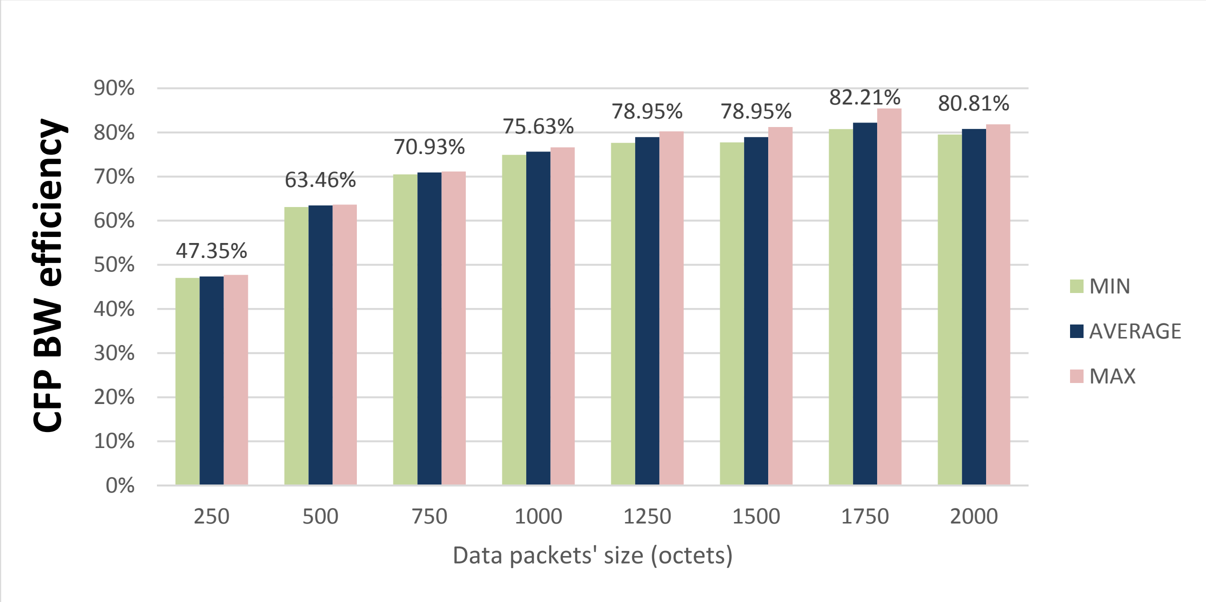
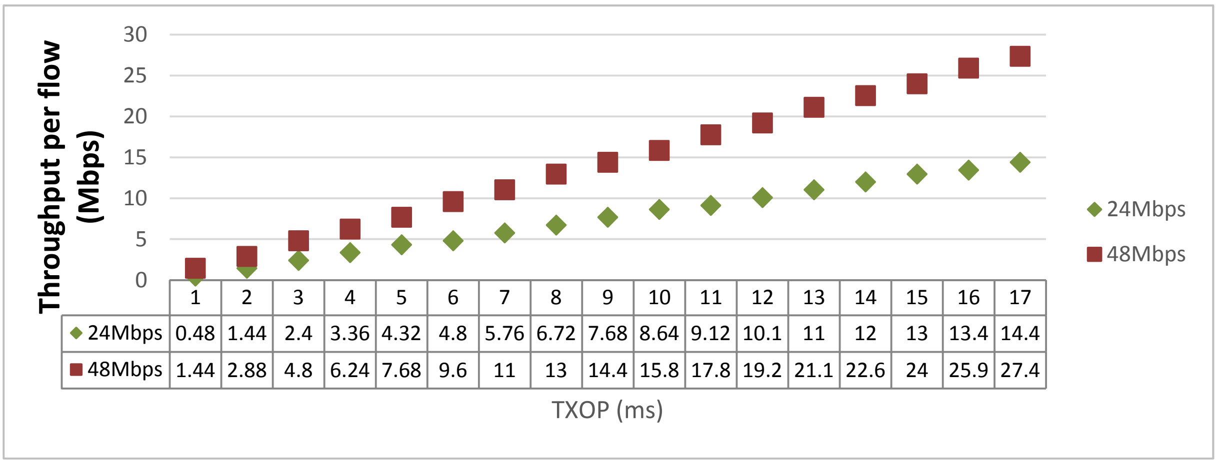


Figure 4-19: Flows' bandwidth efficiency as a function of data packets' size (The average CFP bandwidth efficiency is shown above each set of three columns, i.e. for a given data packet size.)

4.2.1.2.2 Throughput and energy consumption per bit achieved by S-APSD flows

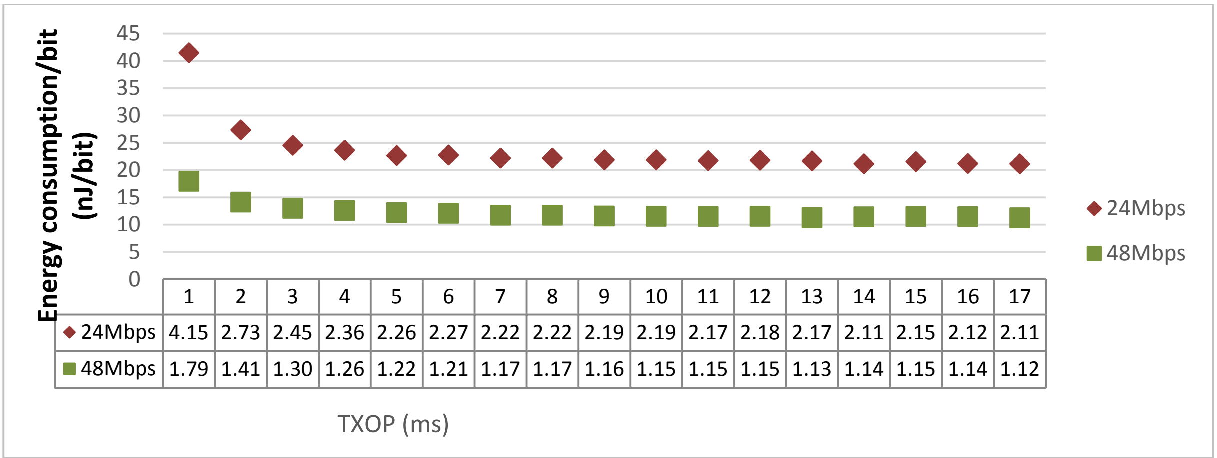
This simulation configuration evaluates throughput and energy consumption per bit achieved by S-APSD flows. The S-APSD traffic parameters used are: each flow's TXOP switches from 1 to 17.5 ms; every flow has SI equal to 25 ms; and the data packet's size was 1500 octets. Since the cumulative maximum duration of the CFPs must be lower or equal to 0.7 times the inter beacon time, then each CFP duration cannot be longer than 0.7 times 25 ms, i.e. 17.5 ms. Therefore, by using one flow with SI and TXOP equal to 25 ms, and 17.5 ms (respectively), we are using the full resources that the solution allows for S-APSD traffic.

Figure 4-20 and Figure 4-21 show the throughput and MN's energy consumption per bit when the default data rate is equal to 48 Mbps and 24 Mbps (respectively). Note that MN's energy consumption includes energy consumption during the times periods when the MN is in doze state.



TXOP (ms)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
24Mbps	0.48	1.44	2.4	3.36	4.32	4.8	5.76	6.72	7.68	8.64	9.12	10.1	11	12	13	13.4	14.4
48Mbps	1.44	2.88	4.8	6.24	7.68	9.6	11	13	14.4	15.8	17.8	19.2	21.1	22.6	24	25.9	27.4

Figure 4-20: Throughput achieved by a flow as a function of its TXOP



TXOP (ms)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
24Mbps	4.15	2.73	2.45	2.36	2.26	2.27	2.22	2.22	2.19	2.19	2.17	2.18	2.17	2.11	2.15	2.12	2.11
48Mbps	1.79	1.41	1.30	1.26	1.22	1.21	1.17	1.17	1.16	1.15	1.15	1.15	1.13	1.14	1.15	1.14	1.12

Figure 4-21: Mobile node's energy consumption (in nJ) per bit while it has one S-APSD active flow as a function of the flow's TXOP

4.2.1.3 U-APSD operations' performance

APs announce through the beacon frames the instants when their CPs begin. Thus, a MN that intends to send or receive a frame must wake up at any t_{cp} instant of its associated AP's CP and contend for the channel.

Regarding U-APSD operations, APs announce through the TIM, which is included in the beacon frames, those MNs that have buffered frames at the AP. Afterward, those MNs must choose one of the next CPs of the AP to retrieve those frames. In order to do so, the MN must contend for the channel at the selected t_{cp} instant, and send a U-APSD trigger frame to initiate an unscheduled SP.

The number of MNs associated with an AP is correlated with the number of nodes contending for the channel at a certain t_{cp} instant. Furthermore, when the number of MNs increases the MNs suffer longer contention times while attempting to send a frame, and therefore, the MNs' energy consumption rises.

A simulation configuration was created in order to evaluate the influence of the number of MNs attached to an AP on the energy consumption of a MN while using U-APSD. The simulation configuration consists of one AP that has different numbers of MNs associated with it. The number of MNs varies from 1 to 20. Moreover, there is no S-APSD traffic associated to any MN. Concerning U-APSD traffic conditions: the TIM is randomly calculated every beacon interval and the unscheduled SPs' length is 5 frames (i.e. MNs always retrieve five data packets within an unscheduled SP). Figure 4-22 and Figure 4-23 show average, minimum, and maximum values of the efficiency and energy consumption per bit as a function of the MNs associated with an AP. This efficiency was defined in section 4.1.8. Note that both parameters are correlated with the time that the MNs have spent in the listening state as shown in Figure 4-24, where this amount of time increases due to increasing contention times.

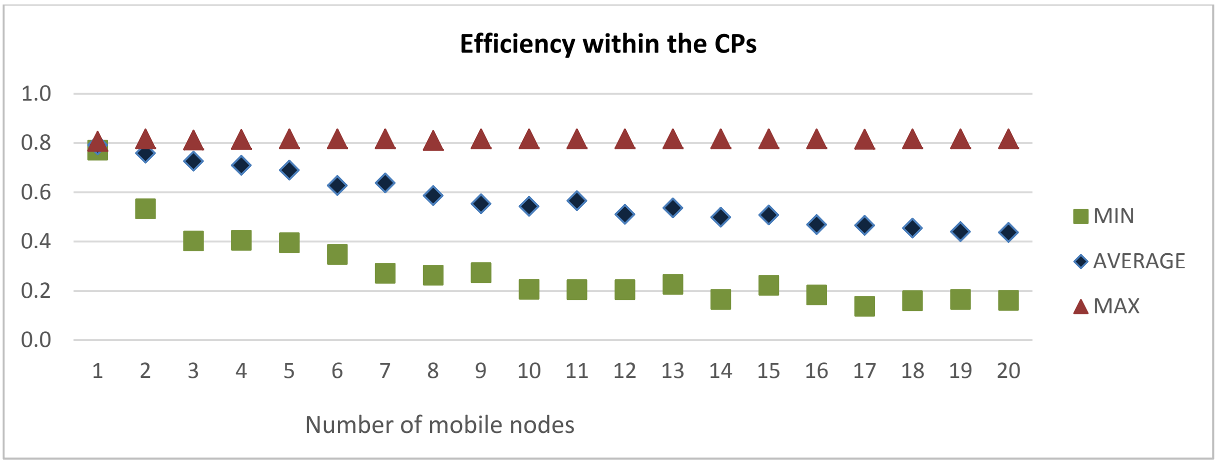


Figure 4-22: Efficiency within CPs as a function of the number of MNs associated with an AP

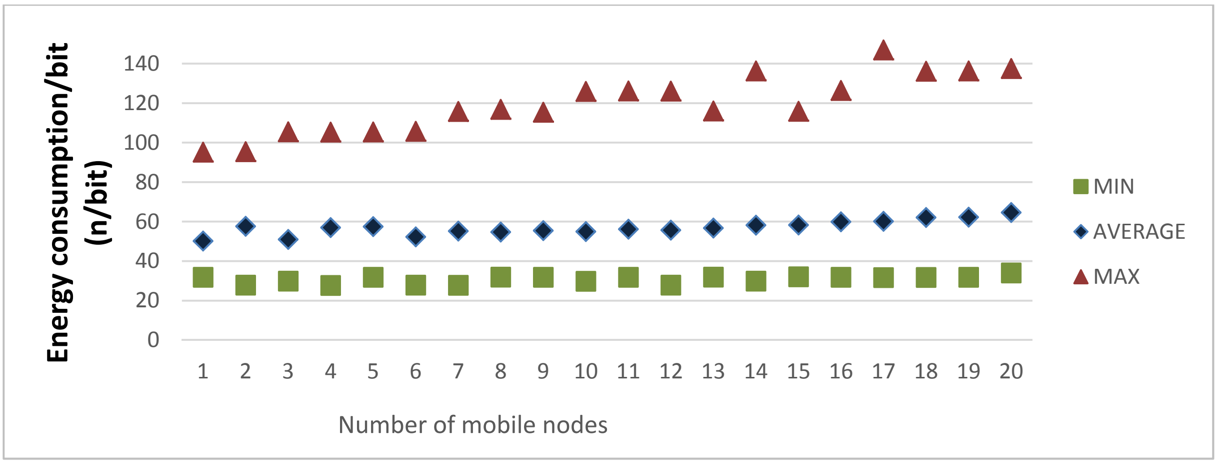


Figure 4-23: MN's energy consumption per bit (in nJ) as a function of the number of MNs associated with an AP

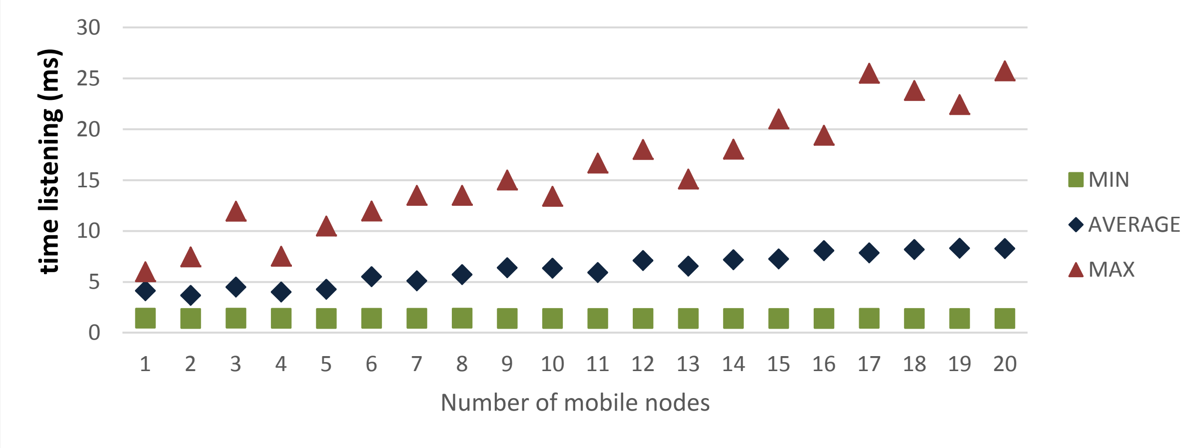


Figure 4-24: Time that a MN spends in listening state while using U-APSD, as a function of the number of MNs associated with an AP

Finally, the maximum number of frames that a MN is allowed to retrieve during an unscheduled SP, i.e. the SP's length, has an effect on U-APSD performance as well. Through a simulation configuration where the SP's length has been varied from 1 to 10, we can conclude that the optimal value for SP's length is equal to 8. Longer values of this parameter reduce the overhead created by contention times for access to a single channel, but also increase the overall duration of the contention times. Figure 4-25 shows the efficiency as a function of the unscheduled SP's length value. The cumulative distribution function of the efficiency can be found Appendix B.

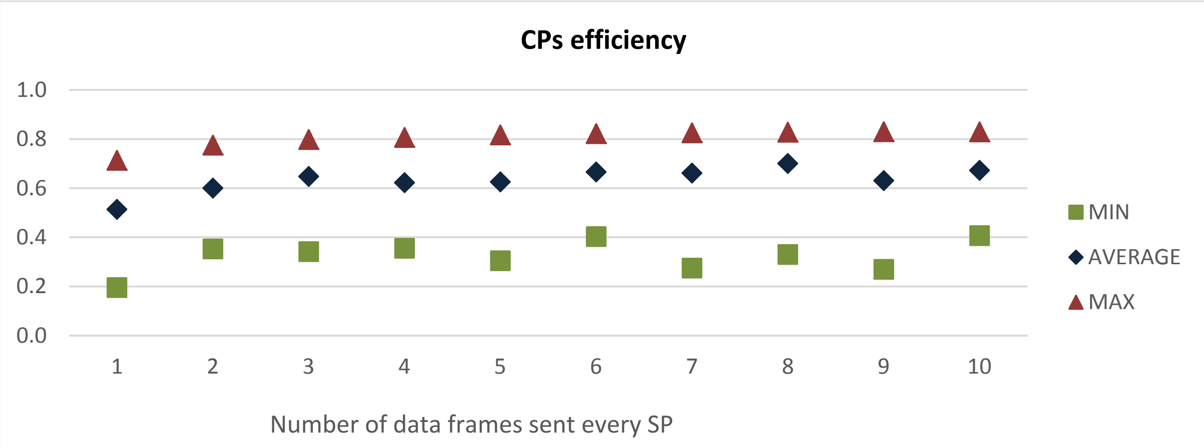


Figure 4-25: Efficiency within CPs as a function of unscheduled SPs' length

The *four-periods* version, which divides the beacon into four CFPs-CPs intervals, outperforms the *one-period* version. The following section compares both versions and shows the main differences between them.

4.2.2 Comparison between both versions

This section highlights the main differences between both versions' performance. Concerning scanning operations, the *four-periods* version achieves shorter scanning latencies. In addition, the number of MNs associated with an AP has a worse influence on the performance of the *one-period* version. Moreover, S-APSD flows do not fulfill some QoS requirements in the *one-period* version because we have used a fixed SI of 100 ms.

In order to compare the scanning latencies achieved by both versions, the simulation configuration described in section 4.2.1.1.1, in which the number of APs to be scanned varies from 1 to 12, has been repeated using the *one-period* version. Figure 4-26 compares the scanning latencies achieved by both versions. The main difference is that the time to find a t_{cp} instant of each target AP lies within the interval zero to 99 ms in the *one-period* version, while in the *four-periods* version it lies between zero and 25 ms, as was illustrated in Figure 4-9.

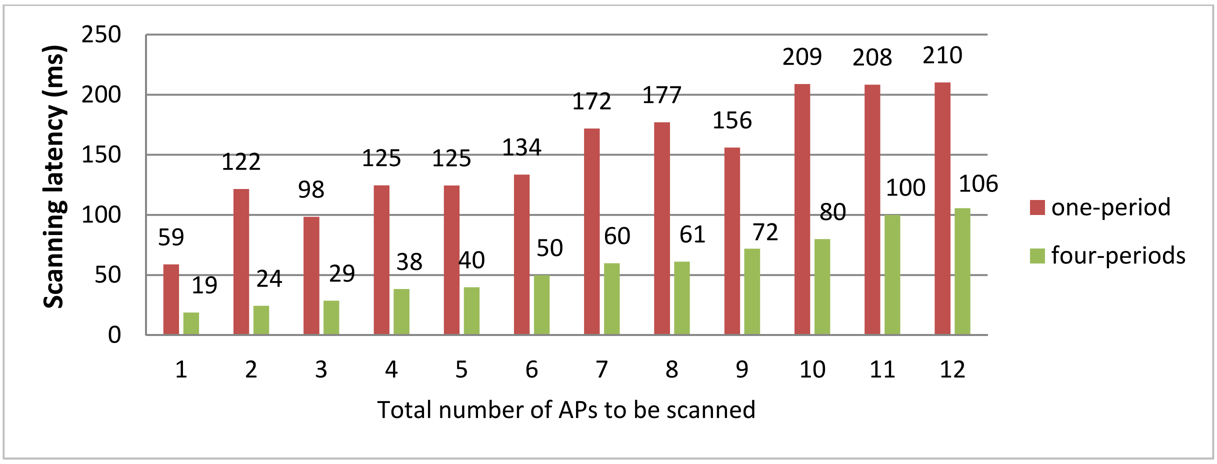


Figure 4-26: Scanning latency comparison between both versions of the proposed solution (The average scanning latency for each version is shown above each set of three columns, i.e. for a given number of APs to be scanned. The larger value in all cases is for the one-period version.)

The influence of the number of MNs associated with an AP on the U-APSD performance of both versions is illustrated in Figure 4-27. In order to obtain these values, the simulation configuration described in section 4.2.1.3 has been applied to the *one-period* version. We can observe that the *one-period* version is affected worse by the number of MNs associated with an AP. This is because the number of MNs contending for the channel rises as well, therefore the contention times are longer. In contrast, in the *four-periods* version the AP has four CPs every beacon interval, which reduces the number of MNs contending for the channel at every t_{cp} instant.

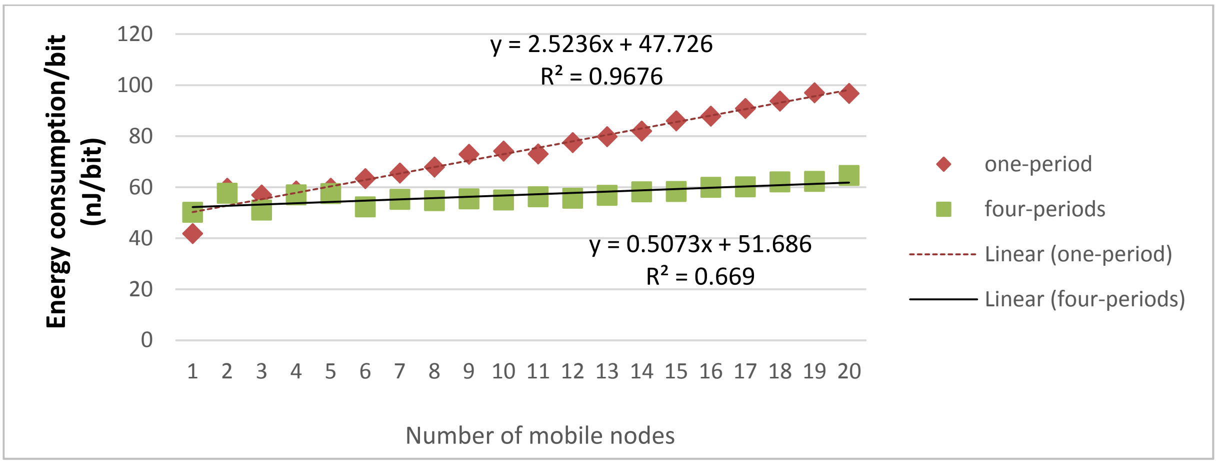


Figure 4-27: Comparison between MNs' energy consumption per bit as a function of the number of mobile nodes of both versions

Finally, S-APSD flows in the *one-period* version, whose SI is fixed and equal to 100 ms, do not fulfill the packetization delay requirement of some multimedia applications, such as VoIP's 40ms [43].

The following section compares the results described above with typical values found in the research literature.

4.2.3 Verification of the results and comparison with other solutions

This section consist of some comparisons in order to verify the validity of the previous simulations' results, as well as to quantify improvements of this solution's performance in comparison with that of other interesting solutions found in the research literature. In the same way as it was done presenting the results in the previous subsections, scanning and S-APSD are evaluated separately. The following subsections compare and contrast scanning and S-APSD simulations' results.

4.2.3.1 Validity and evaluation of scanning results

This section compares scanning latency as well as the MN's energy consumption derived in the previously presented simulations with values found in the research literature.

The latency of the proposed solution's scanning phase is the total time from the beginning of the network discovery phase until the end of the handover execution phase. This will be compared with the total time spent for channel scanning in some other schemes, such as passive scanning, active scanning, selective scanning, and unicast scanning. We will compare with the values presented by Tuysuz et al. in [19] and the values presented by Rebai and Hanafi in [10]. Figure 4-28 shows a comparison of different schemes in terms of the scanning phase's latency when there are 11 and 3 APs respectively in the IEEE 802.11 network.

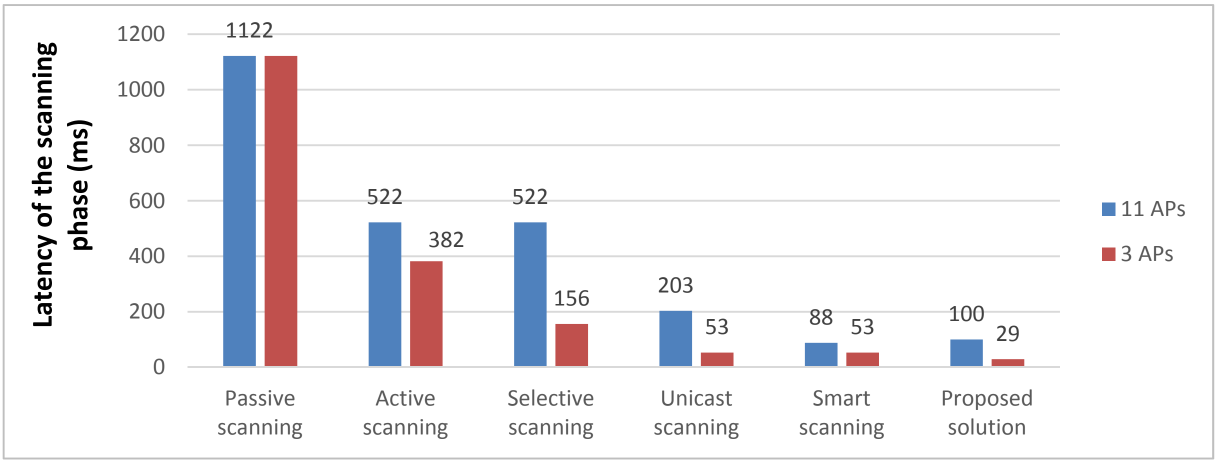


Figure 4-28: Scanning phase latency in ms for various scanning schemes (Our proposed solution's performance is shown in the right most pair of columns.)

The overall MN's energy consumption during the scanning phase is presented in Figure 4-29. In order to obtain the values shown for the previous scanning schemes, it has been assumed that the MN is in the listening state during the total scanning time. This assumption is based on the fact that the sending times of the previously reported scanning schemes are negligible in comparison with their listening times.

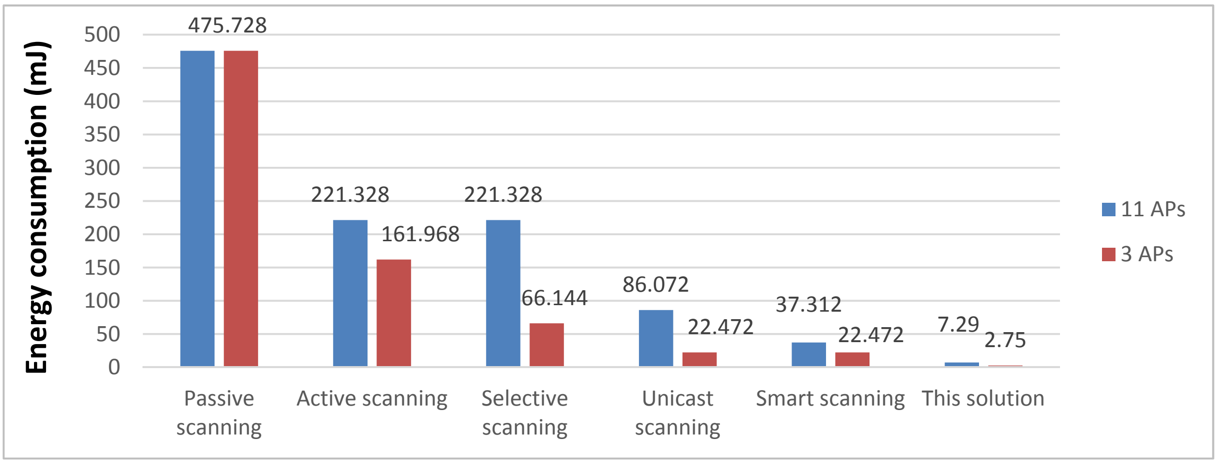


Figure 4-29: MNs' energy consumption in mJ during the scanning phase for various scanning schemes (Our proposed solution's performance is shown in the right most pair of columns.)

On the other hand, in other schemes such as SincScan [17] non-AP STAs may miss packets that were sent while the MN was scanning other channels. Moreover, scanning latencies need to be shorter than the maximum required for some multimedia applications (i.e. 50 ms). As the proposed scanning scheme performs scanning in the background there is no possibility of a packet being missed since the APSD operation does not allow the AP to send packets to the MN when it is either *dozing* or *busy* performing other operations. In addition, even when the scanning latency is longer than the QoS requirement (this may happen when the number of APs to be scanned increases), the requirement is not violated for currently associated MNs' active flows because the scanning operations do not interrupt active flows.

4.2.3.2 Validation and evaluation of the results regarding S-APSD operations

The best feature of this solution regarding S-APSD performance is that active flows are unaffected by operations such as scanning, handovers, or U-APSD operations. Apart from this, there are no major difference between this solution's S-APSD operations and that of other schemes. Actually, the main difference between the different S-APSD schemes is the scheduling algorithm used to allocate flows' SPs. The scheduling algorithm used in the simulations reported earlier in this chapter is a simple grouping approach, which was described in sections 3.1.1 and 4.1.4. In the following discussion this algorithm will be referred to as "this algorithm". In this section, S-APSD performance is compared with that of a low-complexity class-based scheduling algorithm for S-APSD (CLCS-APSD) presented by Lee and Hsieh in [48].

In order to compare both algorithms, the same simulation configuration shown in the performance evaluation section of Lee and Hsieh' algorithm has been made using the simulator described in this chapter. The scenario that has been considered is composed of one AP that broadcasts beacons every 100 ms. Regarding S-APSD traffic, there are three classes of traffic streams associated with real-time gaming, real-time voice, and real-time video applications (respectively). Moreover, the number of non-AP STAs associated with the AP is fixed to 30 and the number of active traffic streams (i.e. flows) is also 30 (one TS per mobile node and 10 TSs for each class). The traffic characteristics used by both algorithms are listed in Table 4-5.

Table 4-5: Traffic characteristics

		CLCS-APSD algorithm		This algorithm	
Class	Application	Data throughput	Service interval	Data throughput	Service interval
1	Real-time gamming	20 Kbps	100 ms	20 kbps	100 ms
2	Real-time Voice	26 Kbps	40 ms	30 kbps	25 ms
3	Real-time video	58 Kbps	60 ms	70 kbps	50 ms

A simulation was performed to model 600 seconds of real time. In addition, the MNs' power consumption in different states is as follows: The sending and receiving states take 1.4 W; while the doze state takes 0.045 W. Finally, the PHY data rate is fixed to 24 Mbps. The average energy consumption of the different non-AP STAs running the different applications for both algorithms is shown in Figure 4-30. Notice that the figure also includes the estimated energy consumption in case of stations working in active mode, i.e. their WNICs are never in doze state.

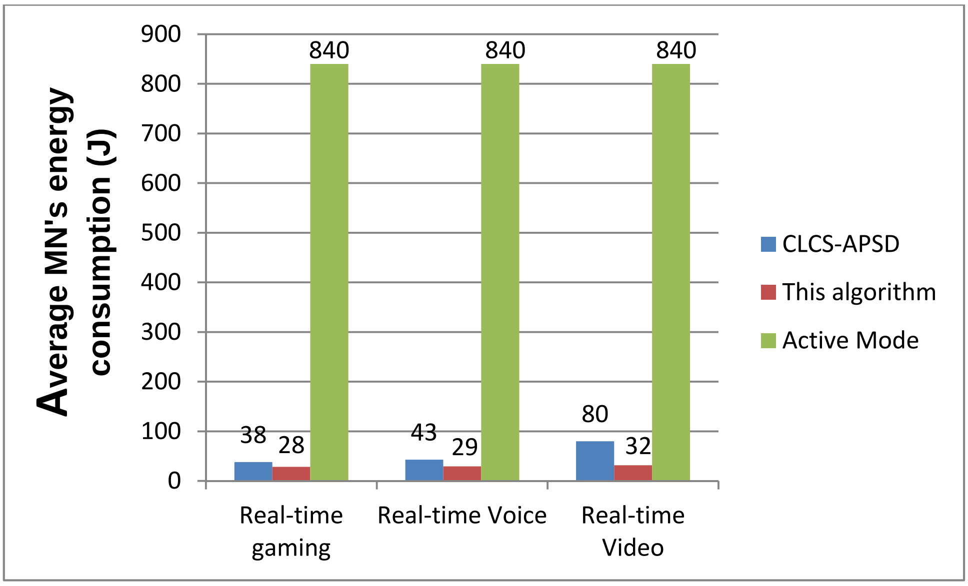


Figure 4-30 Average energy consumption (in J) for a single STA as a function of the class of traffic carried for two different S-APSD schemes

To explain the results shown in Figure 4-30, we can see that the energy consumption increases due to longer time duration for delivering data associated with these multimedia applications. The energy consumption associated with the proposed solution is lightly lower, but it should be due to the fact that the simulator does not consider the energy consumption consumed by switchovers between different power states. Nevertheless, we observe that stations' energy consumption concerning both S-APSD schemes is one order of magnitude lower that in the case of active mode. In addition, Figure 4-31 shows the average power consumption per MN, which has been calculated dividing the average energy consumption by the total time (i.e. 600 s).

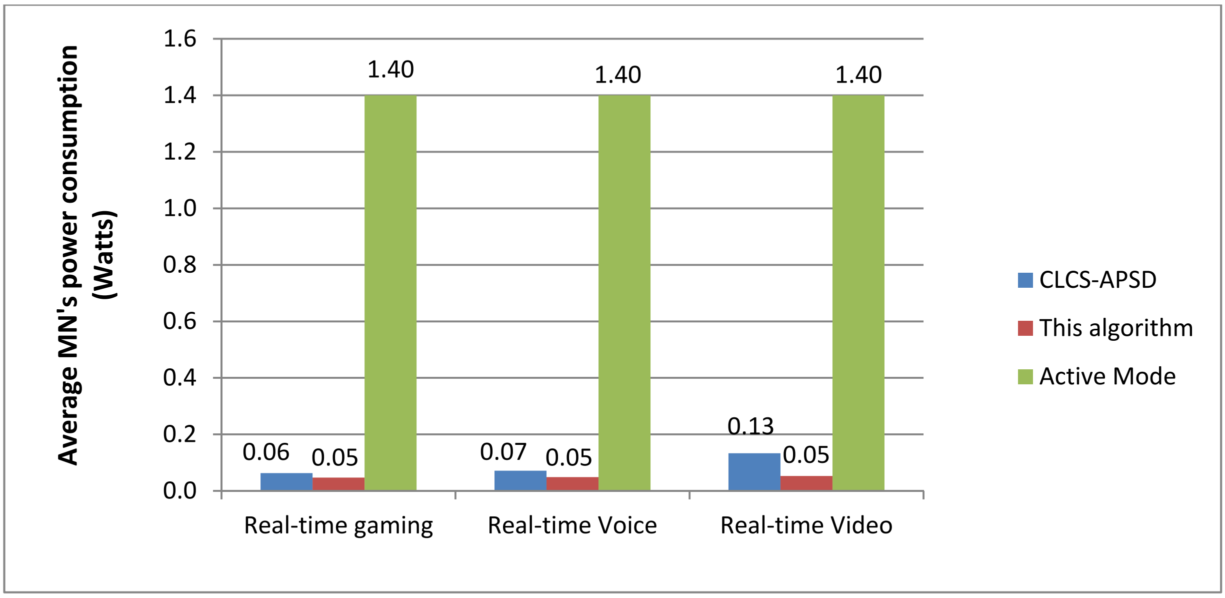


Figure 4-31: Average power consumption (in Watts) for a single STA as a function of the class of traffic carried for two different S-APSD schemes

I could not find suitable values in the literature to compare with the simulations' U-APSD results, due to the fact that U-APSD operations have been simplified in the simulator. However, if we compare results regarding MN's energy consumption per bit for U-APSD with that for S-APSD, we observe that a MN's energy consumption per bit is lightly higher for U-APSD traffic as would be expected since the MN has to spend longer listening while contenting for the channel. Figure 4-32 shows minimum, average, and maximum values for the energy consumption per bit of a single STA using S-APSD, and U-APSD, among all the previous simulation configurations. Those simulation configurations can be found in sections 4.2.1.2 and 4.2.1.3 respectively. Note that the results regarding U-APSD do not consider those for SP lengths shorter than 5 frames, these values are slightly smaller than the energy consumption per bit for S-APSD different due to the fact that MNs consume some extra energy contending for the channel when using U-APSD. Consequently, U-APSD results can be considered reliable.

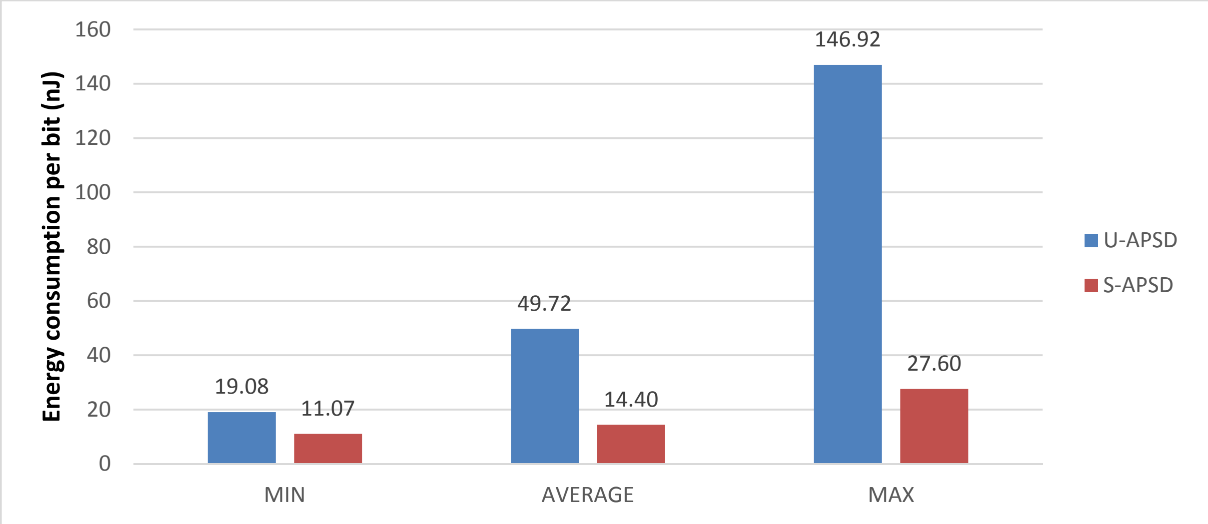


Figure 4-32: Energy consumption per bit (in nJ) for a single STA using U-APSD and S-APSD respectively

5 Conclusions and Future work

This chapter concludes the thesis with a summary of conclusions, suggestions for future work, and some reflections on the social, economic, and ethical aspects of this thesis project.

5.1 Conclusions

The evolution of mobile access towards heterogeneous deployments is a common mobile-broadband strategy of most MNOs. In fact, Wi-Fi technology is expected to be a key component in realizing the promise of future of heterogeneous networks. With Wi-Fi fully integrated in mobile access, MNOs will be able to control devices' choice of connectivity; hence they will be able to increase their network capacity and coverage, and guarantee the best possible user experience and use of resources.

However, current IEEE 802.11 WLANs' performance is far from meet the requirements of this integrated scenario. Fortunately, these requirements can be meet with some enhancements. IEEE 802.11-2012 already provides a number of interesting mechanisms to enhance WLANs' performance, but they are not typically implemented in current deployments. Regarding IEEE 802.11 MAC mechanisms, current solutions typically employ the distributed access function (DCF) mechanism. Nevertheless, this mechanism, which is based on the CSMA/CA access method, generates high overhead due to channel contention, long IFSs, and acknowledgements that waste both bandwidth and devices' energy resources. In contrast, other standardized access mechanisms such as enhanced distributed channel access (EDCA) and hybrid coordination function (HCF) channel access (HCCA) achieve better performance, but they are not usually included in current deployments. In addition, the standard defines APSD mechanisms that can achieve major energy savings while supporting QoS. WLAN discovery mechanisms require further enhancements to meet the latency requirements necessary to enable seamless handover between APs. There exist several vendor-specific solutions and other proposals that aim to enhance WLAN discovery operations. Most of these solutions and proposals are based on selective scanning schemes, and background scanning in order to reduce the latency of the scanning phase and devices' energy consumption.

The proposed solution includes an enhanced MAC protocol that cleverly uses contention free period (CFP) and contention period (CP) concepts, in combination with APSD mechanisms found in the standard in order to enhance IEEE 802.11 WLANs' performance in terms of energy-efficiency, QoS, and use of resources. In addition, the proposed solution includes a background selective active scanning mechanism whose operation is enabled by the proposed MAC protocol. The main features and performance advantages of the proposed solution are summarized below.

WLAN Discovery operations	• The *background selective active scanning* mechanism achieves large **energy savings** and **reduces** the **scanning latency** by using information about other APs in the vicinity. • Scanning operations are performed in the **background,** but they do **not** influence current active flows' performance. Furthermore, **miss-synchronization with the current AP is avoided** by using a coordinated access scheme based on APSD. • The scanning scheme allows STAs to check resource availability at handover candidates *prior* to perform a handover. Moreover, the target AP reserves requested resources for the STA that is performing a handover. Consequently, the proposed solution enables **seamless handovers between APs** by allowing continuity of S-APSD active flows.
Enhanced MAC Protocol	• The proposed MAC protocol is based on a **coordinated channel access scheme** that combines contention-free periods and contention periods in order to **reduce** the main MAC overheads (i.e. **channel contention** and **inter-frame spaces**). • The combination of a coordinated channel access scheme with APSD power saving mechanisms enables **low-power transmissions, high bandwidth efficiency**, and **QoS support**.

On the other hand, there exist several integration tools to enable Wi-Fi/cellular networks interworking, such as the IEEE 802.21 standard, and access network selection and discovery function (ANDSF). The IEEE 802.21 standard provides a suitable framework to integrate Wi-Fi into heterogeneous networks. Therefore, the architecture of heterogeneous networks should be based on this standard. This thesis project proposed a simplified Wi-Fi/cellular integration model that uses a IEEE 802.21-based network architecture. In addition, the proposed solution for IEEE 802.11 WLANs solves some of the main challenges of Wi-Fi integration. The proposed solution allows APs to collect information about other APs in the vicinity and provides this information to the MNs. Furthermore, the proposed solution enables QoS support and seamless handovers between different APs.

5.2 Future work

A IEEE 802.21-based solution that aims to enable Wi-Fi/cellular interworking has only been introduced in this thesis. A detailed design of an IEEE 802.21-based solution should be addressed in future work. This design should include a detailed definition of the network architecture with all the main IEEE 802.21 elements, such as MIH functions and MIH users, suitably placed within the different network entities. Moreover, different algorithms based on MIH message interchange should be designed to allow different interworking operations, such as *network-initiated handovers* or remote access to MIH information servers from external networks.

Additionally, the simulator should be extended to evaluate the performance of a IEEE 802.21-based solution. Concerning the simulator, it might be interesting to implement channel layer operations as the current simulator considers only operations at link and physical layer. The application layer of the simulator could be extended to include support for different applications in order to generate traffic based on application-based patterns (such as VoIP traffic or video streaming). In addition, the simulator needs to include support for non-QoS STAs accessing the channel using legacy DCF/PSM operations. This later addition is important to understand the interoperation with existing equipment.

Finally, issues concerning network security and operator policies have not been addressed in this thesis project. Therefore, future works could focus on how security should be added to the proposed solution and how operator policies will affect the proposed interworking solution.

5.3 Reflections

New business models based on the integration of Wi-Fi into mobile access can be included in operator offerings, thus MNOs offer fully integrated Wi-Fi could provide optimal user experience. Potentially this would enable them to achieve greater market success than MNOs that do not offer heterogeneous services. However, there is also a question of whether the cellular operators *should* be allowed to make such aggressive use (and control) of publically accessible bands that today are utilized by individuals and others for their Wi-Fi networks. From a societal point of view there needs to be discussion of whether these operators should be allowed to take over a public good or whether they should be forced to license additional spectrum that they would implement the modified MAC protocol in, combined with agreements with the WLAN interface manufacturers to extend their devices to operate in these licensed frequencies.

The proposed solution achieves high energy savings that extend the operating time of a mobile devices' battery. Reducing energy consumption also has a positive effect on sustainability as it reduces the amount of electricity needed to charge all of these battery powered devices. Additionally, being able to perform handovers from cellular access networks to WLANs can potentially reduce the electrical energy needed by the cellular access network, further improving the sustainability of these integrated heterogeneous networks.

As noted earlier this thesis project did not look at the security aspects of the proposed solution, hence the problems associated with tracking of mobile devices has been ignored in this thesis project. It is suggest that future research concerning the addition of security to this proposed solution also look at how to minimize the loss of privacy and personal integrity due to the mechanism suggest in this thesis concerning tracking devices and retaining information about the physical location (and locality of all MNs in an MNO's network). In addition to the issues of privacy and integrity of the MNs location, there are also questions about who should have access to the information concerning a MN and its flows that have been ignored in this thesis. These issues must also be addressed in future research.

References

[1] “Ericsson Mobility Report, November 2013,” 15-Nov-2012. [Online]. Available: http://www.ericsson.com/mobility-report. [Accessed: 15-Feb-2014].

[2] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, “LTE release 12 and beyond [Accepted From Open Call],” *IEEE Communications Magazine*, vol. 51, no. 7, pp. 154–160, 2013.

[3] K. Stetson, “Wi-Fi expands as the center of leading-edge technologies in 2011,” AUSTIN, TX, 06-Jan-2011.

[4] “Wi-Fi in heterogeneous networks - an integrated approach to delivering the best user experience,” 25-Jun-2013. [Online]. Available: http://www.ericsson.com/news/130625-wifi-heterogeneous-networks_244129228_c?tagsFilter=Wifi. [Accessed: 06-Feb-2014].

[5] R. Guerra, “Carrier Wi-Fi: the next generation.” 20-Dec-2013.

[6] Wi-Fi® Alliance, “Wi-Fi CERTIFIED Passpoint™: A new program from the Wi-Fi Alliance® to enable seamless Wi-Fi® access in hotspots (2012).” Jun-2012.

[7] R. Castillo and J. María, “Energy-Efficient Vertical Handovers,” Technology, KTH, School of Information and Communication Technology (ICT), Communication Systems, CoS, 2013.

[8] S.-L. Tsao and C.-H. Huang, “A survey of energy efficient MAC protocols for IEEE 802.11 WLAN,” *Computer Communications*, vol. 34, no. 1, pp. 54–67, Jan. 2011.

[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 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pp. 1–2793, 2012.

[10] A. R. Rebai and S. Hanafi, “An Adaptive Multimedia-Oriented Handoff Scheme for IEEE 802.11 WLANs,” *IJWMN*, vol. 3, no. 1, Feb. 2011.

[11] E. Dahlman, S. Parkvall, and J. Sköld, *4G: LTE/LTE-Advanced for mobile Broadband.* Academic Press Inc, 2013.

[12] “Radiocommunication Sector (ITU-R) - ITU global standard for international mobile telecommunications ´IMT-Advanced´.” [Online]. Available: http://www.itu.int/ITU-R/index.asp?category=information&rlink=imt-advanced&lang=en. [Accessed: 11-Oct-2013].

[13] Wi-Fi® Alliance, “Wi-Fi CERTIFIED™.” [Online]. Available: https://www.wi-fi.org/certification. [Accessed: 06-Sep-2013].

[14] J. L. Burbank, J. Andrusenko, J. S. Everett, and W. T. M. Kasch, *Wireless networking understanding internetworking challenges*. Hoboken, NJ; Piscataway, NJ: Wiley ; IEEE Prerss, 2013.

[15] A. F. Molisch, “Spread Spectrum Systems - Chapter18,” in *Wireless Communications*, Chichester: Wiley, 2011, pp. 387 – 416.

[16] J. Guo, “A New Authenticator,” Student thesis, KTH, School of Information and Communication Technology (ICT), 2010.

[17] I. Ramani and S. Savage, “SyncScan: practical fast handoff for 802.11 infrastructure networks,” in *Proceedings IEEE INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2005, vol. 1, pp. 675–684 vol. 1.

[18] P.-J. Huang, Y.-C. Tseng, and K.-C. Tsai, “A Fast Handoff Mechanism for IEEE 802.11 and IAPP Networks,” in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, 2006, vol. 2, pp. 966–970.

[19] M. F. Tuysuz and H. A. Mantar, “A novel energy-efficient QoS-aware handover scheme over IEEE 802.11 WLANs,” in *2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2013, pp. 1045–1049.

[20] G. Athanasiou, T. Korakis, and L. Tassiulas, “An 802.11k Compliant Framework for Cooperative Handoff in Wireless Networks,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, p. 350643, Sep. 2009.

[21] Z. Chang, O. Alanen, E. H. Ong, and J. Kneckt, “Enhanced channel scanning schemes for next generation WLAN system,” in *2012 1st IEEE International Conference on Communications in China (ICCC)*, 2012, pp. 676–681.

[22] E. H. Ong, “Performance analysis of fast initial link setup for IEEE 802.11ai WLANs,” in *2012 IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2012, pp. 1279–1284.

[23] “IEEE Standard for Local and Metropolitan Area Networks- Part 21: Media Independent Handover,” *IEEE Std 802.21-2008*, pp. c1–301, 2009.

[24] W.-S. Lim, D.-W. Kim, Y.-J. Suh, and J.-J. Won, “Efficient WLAN Discovery Schemes Based on IEEE 802.21 MIH Services in Heterogeneous Wireless Networks,” in *IEEE Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008*, 2008, pp. 1–5.

[25] M. Q. Khan and S. H. Andresen, “Zero Scanning Time for 802.11 Networks by Using Media Independent Information Server (MIIS),” in *2012 IEEE 26th International Conference on Advanced Information Networking and Applications (AINA)*, 2012, pp. 467–473.

[26] D. A. Wassie, D. Loukatos, L. Sarakis, K. Kontovasilis, and C. Skianis, “On the energy requirements of vertical handover operations: Measurement-based results for the IEEE 802.21 framework,” in *2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2012, pp. 145–149.

[27] S. J. Bae, M. Y. Chung, and J. So, “Handover triggering mechanism based on IEEE 802.21 in heterogeneous networks with LTE and WLAN,” in *2011 International Conference on Information Networking (ICOIN)*, 2011, pp. 399–403.

[28] 3GPP Technical specification 23.402, “Architecture enhancements for non-3GPP accesses,” Sep-2011.

[29] 3GPP Technical specification 23.312, “Access Network Discovery and Selection Function (ANDSF) Management Object (MO) - Release 11,” 15-Mar-2013.

[30] K. Stetson, “Wi-Fi® takes center stage as capacity expansion solution for operator networks,” *Austin, Texas*, 21-May-2013.

[31] “IEEE Standard for Information Technology-Telecommunications and information exchange between systems-Local and Metropolitan networks-specific requirements-Part II: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 9: Interworking with External Networks.” 2011.

[32] N. Ding, A. Pathak, D. Koutsonikolas, C. Shepard, Y. C. Hu, and L. Zhong, “Realizing the full potential of PSM using proxying,” in *2012 Proceedings IEEE INFOCOM*, 2012, pp. 2821–2825.

[33] Y. He, R. Yuan, X. Ma, and J. Li, “Analysis of the impact of background traffic on the performance of 802.11 power saving mechanism,” *IEEE Communications Letters*, vol. 13, no. 3, pp. 164–166, 2009.

[34] “IEEE Standard for Information technology–Local and metropolitan area networks–Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements,” *IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003)*, pp. 1–212, 2005.

[35] X. Perez-Costa and D. Camps-Mur, “IEEE 802.11E QoS and power saving features overview and analysis of combined performance [Accepted from Open Call],” *IEEE Wireless Communications*, vol. 17, no. 4, pp. 88–96, 2010.

[36] D. C. Mur, X. Pérez-Costa, V. Marchenko, and S. S. Ribes, “On centralized schedulers for 802.11e WLANs distribution versus grouping of resources allocation,” *Wireless Communications and Mobile Computing*, vol. 12, no. 13, pp. 1175–1190, 2012.

[37] J. Kellokoski, J. Koskinen, and T. Hamalainen, “Power Consumption Analysis of the Always-Best-Connected User Equipment,” in *2012 5th International Conference on New Technologies, Mobility and Security (NTMS)*, 2012, pp. 1–5.

[38] H. Petander, “Energy-aware network selection using traffic estimation,” in *Proceedings of the 1st ACM workshop on Mobile internet through cellular networks*, New York, NY, USA, 2009, pp. 55–60.

[39] R. Trestian, A.-N. Moldovan, O. Ormond, and G. Muntean, “Energy consumption analysis of video streaming to Android mobile devices,” in *2012 IEEE Network Operations and Management Symposium (NOMS)*, 2012, pp. 444–452.

[40] R. Trestian, O. Ormond, and G.-M. Muntean, “Energy-Quality-Cost Tradeoff in a Multimedia-Based Heterogeneous Wireless Network Environment,” *IEEE Transactions on Broadcasting*, vol. 59, no. 2, pp. 340–357, 2013.

[41] V. Kantubukta, S. Maheshwari, S. Mahapatra, and C. S. Kumar, “Energy and quality of service aware FUZZY-technique for order preference by similarity to ideal solution based vertical handover decision algorithm for heterogeneous wireless networks,” *IET Networks*, vol. 2, no. 3, pp. 103–114, 2013.

[42] M. Paolini, “Carrier Wi-Fi® for mobile operators: The economics of Wi-Fi small cells (2013).” Wi - Fi Alliance®, 2013.

[43] ITU-T, “International telephone connections and circuits – General Recommendations on the transmission quality for an entire international telephone connection - G.114,” May 2003.

[44] “IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability Via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation,” *IEEE Std 802.11F-2003*, pp. 0_1–67, 2003.

[45] P. Goransson and R. Greenlaw, *Secure Roaming in 802.11 Networks*. Newnes, 2011.

[46] H. Jeon and B. Lee, “A low latency background scan in IEEE 802.11 networks using a triangular cooperation,” in *2013 International Conference on ICT Convergence (ICTC)*, 2013, pp. 841–842.

[47] “Wireless Network Simulator in Matlab.” [Online]. Available: http://wireless-matlab.sourceforge.net/. [Accessed: 11-Dec-2013].

[48] T.-H. Lee and J.-R. Hsieh, “Low-Complexity Class-Based Scheduling Algorithm for Scheduled Automatic Power-Save Delivery for Wireless LANs,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 3, pp. 571–580, 2013.

Appendix A Simulation output

```
===== Network size = 8  maxx = maxy = 516.3978 =====
time 0:             AP1 +---- BEACON (ch 1) ->
time 1.7792e-05:    AP1+-     -- BEACON (ch 1)----------> MN5
time 2e-05:         MN5 +------- SAPSD_DATA (flow 1, SP 1, ch 1) ->     AP1
time 0.00029613:    MN5+-     -- SAPSD_DATA (flow 1, SP 1, ch 1)------------> AP1
time 0.00031213:    MN5       <- SAPSD_ACK (flow 1, SP 1, ch 1)----------------+ AP1
time 0.00031492:    MN5 <---------- SAPSD_ACK (flow 1, SP 1, ch 1)--     -+ AP1
The end of SP 1 of flow 1 node 5 has been reached
time 0.00037:       AP1 +------- SAPSD_DATA (flow 2, SP 1, ch 1) ->     MN5
time 0.00064613:    AP1+-     -- SAPSD_DATA (flow 2, SP 1, ch 1)------------> MN5
time 0.00066213:    AP1       <- SAPSD_ACK (flow 2, SP 1, ch 1)----------------+ MN5
time 0.00066492:    AP1 <---------- SAPSD_ACK (flow 2, SP 1, ch 1)--     -+ MN5
The end of SP 1 of flow 2 node 5 has been reached
time 0.005:         AP2 +---- BEACON (ch 2) ->
time 0.0050178:     AP2+-     -- BEACON (ch 2)----------> MN6
time 0.00502:       MN6 +------- SAPSD_DATA (flow 4, SP 1, ch 2) ->     AP2
time 0.0052961:     MN6+-     -- SAPSD_DATA (flow 4, SP 1, ch 2)------------> AP2
time 0.0053121:     MN6       <- SAPSD_ACK (flow 4, SP 1, ch 2)----------------+ AP2
time 0.0053149:     MN6 <---------- SAPSD_ACK (flow 4, SP 1, ch 2)--     -+ AP2
The end of SP 1 of flow 4 node 6 has been reached
time 0.00537:       MN6 +------- SAPSD_DATA (flow 5, SP 1, ch 2) ->     AP2
time 0.0056461:     MN6+-     -- SAPSD_DATA (flow 5, SP 1, ch 2)------------> AP2
time 0.0056621:     MN6       <- SAPSD_ACK (flow 5, SP 1, ch 2)----------------+ AP2
time 0.0056649:     MN6 <---------- SAPSD_ACK (flow 5, SP 1, ch 2)--     -+ AP2
The end of SP 1 of flow 5 node 6 has been reached
time 0.01:          AP3 +---- BEACON (ch 3) ->
time 0.010018:      AP3+-     -- BEACON (ch 3)----------> MN7
time 0.01002:       AP3 +------- SAPSD_DATA (flow 7, SP 1, ch 3) ->     MN7
time 0.010296:      AP3+-     -- SAPSD_DATA (flow 7, SP 1, ch 3)------------> MN7
time 0.010312:      AP3       <- SAPSD_ACK (flow 7, SP 1, ch 3)----------------+ MN7
time 0.010315:      AP3 <---------- SAPSD_ACK (flow 7, SP 1, ch 3)--     -+ MN7
The end of SP 1 of flow 7 node 7 has been reached
time 0.01037:       MN7 +------- SAPSD_DATA (flow 8, SP 1, ch 3) ->     AP3
time 0.010646:      MN7+-     -- SAPSD_DATA (flow 8, SP 1, ch 3)------------> AP3
time 0.010662:      MN7       <- SAPSD_ACK (flow 8, SP 1, ch 3)----------------+ AP3
time 0.010665:      MN7 <---------- SAPSD_ACK (flow 8, SP 1, ch 3)--     -+ AP3
The end of SP 1 of flow 8 node 7 has been reached
time 0.015:         AP4 +---- BEACON (ch 4) ->
time 0.015018:      AP4+-     -- BEACON (ch 4)----------> MN8
time 0.01502:       MN8 +------- SAPSD_DATA (flow 10, SP 1, ch 4) ->     AP4
time 0.015296:      MN8+-     -- SAPSD_DATA (flow 10, SP 1, ch 4)------------> AP4
time 0.015312:      MN8       <- SAPSD_ACK (flow 10, SP 1, ch 4)----------------+ AP4
time 0.015315:      MN8 <---------- SAPSD_ACK (flow 10, SP 1, ch 4)--     -+ AP4
The end of SP 1 of flow 10 node 8 has been reached
time 0.01537:       MN8 +------- SAPSD_DATA (flow 11, SP 1, ch 4) ->     AP4
time 0.015646:      MN8+-     -- SAPSD_DATA (flow 11, SP 1, ch 4)------------> AP4
time 0.015662:      MN8       <- SAPSD_ACK (flow 11, SP 1, ch 4)----------------+ AP4
time 0.015665:      MN8 <---------- SAPSD_ACK (flow 11, SP 1, ch 4)--     -+ AP4
The end of SP 1 of flow 11 node 8 has been reached
time 0.02502:       AP1 +------- SAPSD_DATA (flow 1, SP 2, ch 1) ->     MN5
time 0.025296:      AP1+-     -- SAPSD_DATA (flow 1, SP 2, ch 1)------------> MN5
time 0.025312:      AP1       <- SAPSD_ACK (flow 1, SP 2, ch 1)----------------+ MN5
time 0.025315:      AP1 <---------- SAPSD_ACK (flow 1, SP 2, ch 1)--     -+ MN5
```

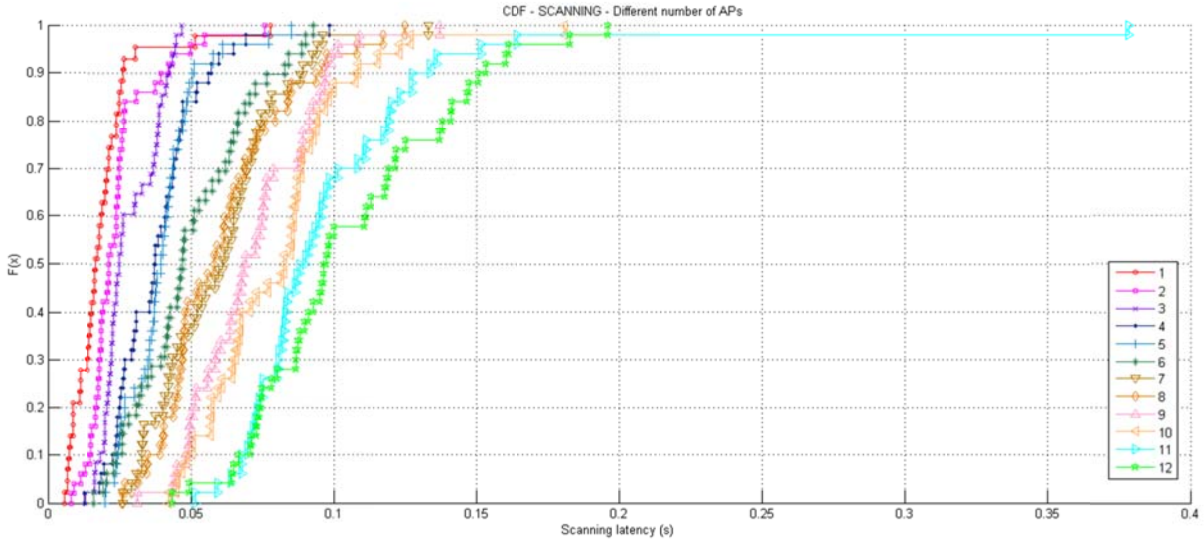
```
The end of SP 2 of flow 1 node 5 has been reached
time 0.02537:      MN5 +------- SAPSD_DATA (flow 3, SP 1, ch 1) ->      AP1
time 0.025646:     MN5+-    -- SAPSD_DATA (flow 3, SP 1, ch 1)------------> AP1
time 0.025662:     MN5        <- SAPSD_ACK (flow 3, SP 1, ch 1)----------------+ AP1
time 0.025665:     MN5 <---------- SAPSD_ACK (flow 3, SP 1, ch 1)--     -+ AP1
The end of SP 1 of flow 3 node 5 has been reached
time 0.03002:      AP2 +------- SAPSD_DATA (flow 4, SP 2, ch 2) ->      MN6
time 0.030296:     AP2+-    -- SAPSD_DATA (flow 4, SP 2, ch 2)------------> MN6
time 0.030312:     AP2        <- SAPSD_ACK (flow 4, SP 2, ch 2)----------------+ MN6
time 0.030315:     AP2 <---------- SAPSD_ACK (flow 4, SP 2, ch 2)--     -+ MN6
The end of SP 2 of flow 4 node 6 has been reached
time 0.03037:      MN6 +------- SAPSD_DATA (flow 6, SP 1, ch 2) ->      AP2
time 0.030646:     MN6+-    -- SAPSD_DATA (flow 6, SP 1, ch 2)------------> AP2
time 0.030662:     MN6        <- SAPSD_ACK (flow 6, SP 1, ch 2)----------------+ AP2
time 0.030665:     MN6 <---------- SAPSD_ACK (flow 6, SP 1, ch 2)--     -+ AP2
The end of SP 1 of flow 6 node 6 has been reached
time 0.03502:      AP3 +------- SAPSD_DATA (flow 7, SP 2, ch 3) ->      MN7
time 0.035296:      AP3+-    -- SAPSD_DATA (flow 7, SP 2, ch 3)------------> MN7
time 0.035312:     AP3        <- SAPSD_ACK (flow 7, SP 2, ch 3)----------------+ MN7
time 0.035315:     AP3 <---------- SAPSD_ACK (flow 7, SP 2, ch 3)--     -+ MN7
The end of SP 2 of flow 7 node 7 has been reached
time 0.03537:      MN7 +------- SAPSD_DATA (flow 9, SP 1, ch 3) ->      AP3
time 0.035646:     MN7+-    -- SAPSD_DATA (flow 9, SP 1, ch 3)------------> AP3
time 0.035662:     MN7        <- SAPSD_ACK (flow 9, SP 1, ch 3)----------------+ AP3
time 0.035665:     MN7 <---------- SAPSD_ACK (flow 9, SP 1, ch 3)--     -+ AP3
The end of SP 1 of flow 9 node 7 has been reached
time 0.03572:      MN7 +------- UAPSD_SP_TRIGGER (ch 3) ->      AP3
time 0.035802:     MN7+-    -- UAPSD_SP_TRIGGER (ch 3)------------> AP3
time 0.035802:     MN7     <- UAPSD_DATA (ch 3) -----------+ AP3
time 0.036079:     MN7<----------- UAPSD_DATA ch(3)--     -+ AP3
time 0.036095:     node 7+------------- UAPSD_ACK ch(3)-->     node 3
time 0.036097:     MN7+-       -- UAPSD_ACK (ch 3)---------------> AP3
The end of an U-APSD SP  of node 7 has been reached
time 0.04002:      AP4 +------- SAPSD_DATA (flow 10, SP 2, ch 4) ->      MN8
time 0.040296:     AP4+-    -- SAPSD_DATA (flow 10, SP 2, ch 4)------------> MN8
time 0.040312:     AP4        <- SAPSD_ACK (flow 10, SP 2, ch 4)----------------+ MN8
time 0.040315:     AP4 <---------- SAPSD_ACK (flow 10, SP 2, ch 4)--     -+ MN8
The end of SP 2 of flow 10 node 8 has been reached
time 0.04037:      MN8 +------- SAPSD_DATA (flow 12, SP 1, ch 4) ->      AP4
time 0.040646:     MN8+-    -- SAPSD_DATA (flow 12, SP 1, ch 4)------------> AP4
time 0.040662:     MN8        <- SAPSD_ACK (flow 12, SP 1, ch 4)----------------+ AP4
time 0.040665:     MN8 <---------- SAPSD_ACK (flow 12, SP 1, ch 4)--     -+ AP4
The end of SP 1 of flow 12 node 8 has been reached
time 0.05002:      MN5 +------- SAPSD_DATA (flow 1, SP 3, ch 1) ->      AP1
time 0.050296:     MN5+-    -- SAPSD_DATA (flow 1, SP 3, ch 1)------------> AP1
time 0.050312:     MN5        <- SAPSD_ACK (flow 1, SP 3, ch 1)----------------+ AP1
time 0.050315:     MN5 <---------- SAPSD_ACK (flow 1, SP 3, ch 1)--     -+ AP1
The end of SP 3 of flow 1 node 5 has been reached
time 0.05037:      MN5 +------- SAPSD_DATA (flow 2, SP 2, ch 1) ->      AP1
time 0.050646:     MN5+-    -- SAPSD_DATA (flow 2, SP 2, ch 1)------------> AP1
time 0.050662:     MN5        <- SAPSD_ACK (flow 2, SP 2, ch 1)----------------+ AP1
time 0.050665:     MN5 <---------- SAPSD_ACK (flow 2, SP 2, ch 1)--     -+ AP1
The end of SP 2 of flow 2 node 5 has been reached
time 0.05502:      AP2 +------- SAPSD_DATA (flow 4, SP 3, ch 2) ->      MN6
time 0.055296:     AP2+-    -- SAPSD_DATA (flow 4, SP 3, ch 2)------------> MN6
```

```
time 0.055312:    AP2      <- SAPSD_ACK (flow 4, SP 3, ch 2)----------------+ MN6
time 0.055315:    AP2 <---------- SAPSD_ACK (flow 4, SP 3, ch 2)--     -+ MN6
The end of SP 3 of flow 4 node 6 has been reached
time 0.05537:     MN6 +------- SAPSD_DATA (flow 5, SP 2, ch 2) ->      AP2
time 0.055646:    MN6+-    -- SAPSD_DATA (flow 5, SP 2, ch 2)------------> AP2
time 0.055662:    MN6      <- SAPSD_ACK (flow 5, SP 2, ch 2)----------------+ AP2
time 0.055665:    MN6 <---------- SAPSD_ACK (flow 5, SP 2, ch 2)--     -+ AP2
The end of SP 2 of flow 5 node 6 has been reached
time 0.06002:     AP3 +------- SAPSD_DATA (flow 7, SP 3, ch 3) ->      MN7
time 0.060296:    AP3+-    -- SAPSD_DATA (flow 7, SP 3, ch 3)------------> MN7
time 0.060312:    AP3      <- SAPSD_ACK (flow 7, SP 3, ch 3)----------------+ MN7
time 0.060315:    AP3 <---------- SAPSD_ACK (flow 7, SP 3, ch 3)--     -+ MN7
The end of SP 3 of flow 7 node 7 has been reached
time 0.06037:     MN7 +------- SAPSD_DATA (flow 8, SP 2, ch 3) ->      AP3
time 0.060646:    MN7+-    -- SAPSD_DATA (flow 8, SP 2, ch 3)------------> AP3
time 0.060662:    MN7      <- SAPSD_ACK (flow 8, SP 2, ch 3)----------------+ AP3
time 0.060665:    MN7 <---------- SAPSD_ACK (flow 8, SP 2, ch 3)--     -+ AP3
The end of SP 2 of flow 8 node 7 has been reached
time 0.06502:     MN8 +------- SAPSD_DATA (flow 10, SP 3, ch 4) ->      AP4
time 0.065296:    MN8+-    -- SAPSD_DATA (flow 10, SP 3, ch 4)------------> AP4
time 0.065312:    MN8      <- SAPSD_ACK (flow 10, SP 3, ch 4)----------------+ AP4
time 0.065315:    MN8 <---------- SAPSD_ACK (flow 10, SP 3, ch 4)--     -+ AP4
The end of SP 3 of flow 10 node 8 has been reached
time 0.06537:     AP4 +------- SAPSD_DATA (flow 11, SP 2, ch 4) ->      MN8
time 0.065646:    AP4+-    -- SAPSD_DATA (flow 11, SP 2, ch 4)------------> MN8
time 0.065662:    AP4      <- SAPSD_ACK (flow 11, SP 2, ch 4)----------------+ MN8
time 0.065665:    AP4 <---------- SAPSD_ACK (flow 11, SP 2, ch 4)--     -+ MN8
The end of SP 2 of flow 11 node 8 has been reached
time 0.07502:     MN5 +------- SAPSD_DATA (flow 1, SP 4, ch 1) ->      AP1
time 0.075296:    MN5+-    -- SAPSD_DATA (flow 1, SP 4, ch 1)------------> AP1
time 0.075312:    MN5      <- SAPSD_ACK (flow 1, SP 4, ch 1)----------------+ AP1
time 0.075315:    MN5 <---------- SAPSD_ACK (flow 1, SP 4, ch 1)--     -+ AP1
The end of SP 4 of flow 1 node 5 has been reached
time 0.07537:     MN5 +------- UAPSD_SP_TRIGGER (ch 1) ->      AP1
time 0.075443:    MN5+-    -- UAPSD_SP_TRIGGER (ch 1)------------> AP1
time 0.075443:    MN5     <- UAPSD_DATA (ch 1) -----------+ AP1
time 0.07572:     MN5<----------- UAPSD_DATA ch(1)--      -+ AP1
time 0.075736:    node 5+------------- UAPSD_ACK ch(1)-->     node 1
time 0.075738:    MN5+-       -- UAPSD_ACK (ch 1)---------------> AP1
The end of an U-APSD SP  of node 5 has been reached
time 0.08002:     MN6 +------- SAPSD_DATA (flow 4, SP 4, ch 2) ->      AP2
time 0.080296:    MN6+-    -- SAPSD_DATA (flow 4, SP 4, ch 2)------------> AP2
time 0.080312:    MN6      <- SAPSD_ACK (flow 4, SP 4, ch 2)----------------+ AP2
time 0.080315:    MN6 <---------- SAPSD_ACK (flow 4, SP 4, ch 2)--     -+ AP2
The end of SP 4 of flow 4 node 6 has been reached
time 0.08502:     MN7 +------- SAPSD_DATA (flow 7, SP 4, ch 3) ->      AP3
time 0.085296:    MN7+-    -- SAPSD_DATA (flow 7, SP 4, ch 3)------------> AP3
time 0.085312:    MN7      <- SAPSD_ACK (flow 7, SP 4, ch 3)----------------+ AP3
time 0.085315:    MN7 <---------- SAPSD_ACK (flow 7, SP 4, ch 3)--     -+ AP3
The end of SP 4 of flow 7 node 7 has been reached
time 0.09002:     MN8 +------- SAPSD_DATA (flow 10, SP 4, ch 4) ->      AP4
time 0.090296:    MN8+-    -- SAPSD_DATA (flow 10, SP 4, ch 4)------------> AP4
time 0.090312:    MN8      <- SAPSD_ACK (flow 10, SP 4, ch 4)----------------+ AP4
time 0.090315:    MN8 <---------- SAPSD_ACK (flow 10, SP 4, ch 4)--     -+ AP4
The end of SP 4 of flow 10 node 8 has been reached
```

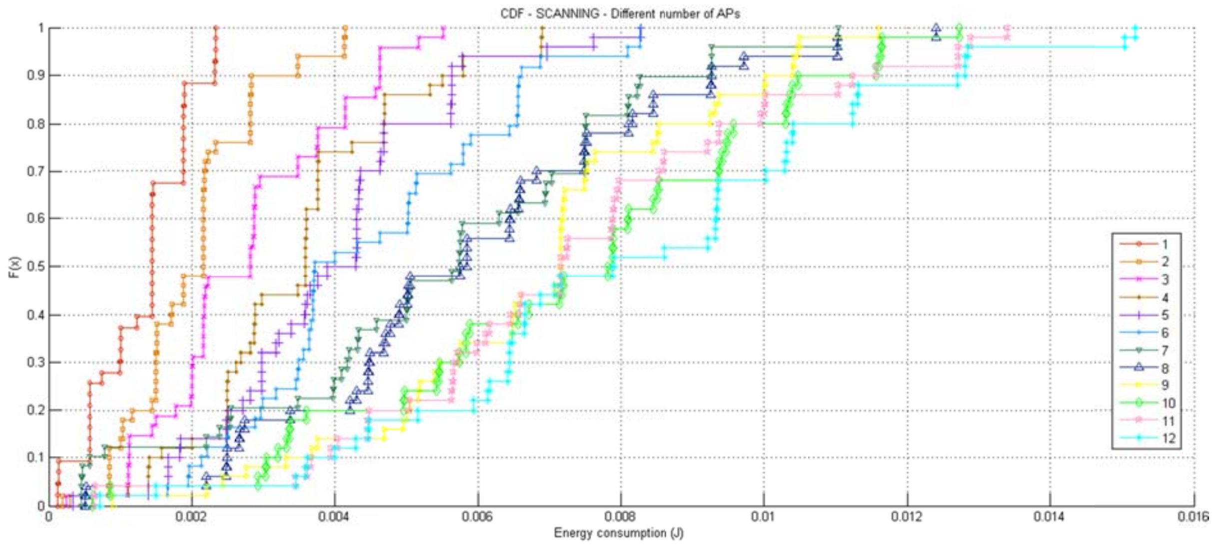
```
time 0.1:          AP1 +---- BEACON (ch 1) ->
time 0.10002:      AP1+-      -- BEACON (ch 1)----------> MN5
time 0.10002:      MN5 +------- SAPSD_DATA (flow 1, SP 5, ch 1) ->      AP1
time 0.1003:       MN5+-     -- SAPSD_DATA (flow 1, SP 5, ch 1)------------> AP1
time 0.10031:      MN5        <- SAPSD_ACK (flow 1, SP 5, ch 1)----------------+ AP1
time 0.10031:      MN5 <---------- SAPSD_ACK (flow 1, SP 5, ch 1)--     -+ AP1
The end of SP 5 of flow 1 node 5 has been reached
time 0.10037:      AP1 +------- SAPSD_DATA (flow 2, SP 3, ch 1) ->      MN5
time 0.10065:      AP1+-     -- SAPSD_DATA (flow 2, SP 3, ch 1)------------> MN5
time 0.10066:      AP1        <- SAPSD_ACK (flow 2, SP 3, ch 1)----------------+ MN5
time 0.10066:      AP1 <---------- SAPSD_ACK (flow 2, SP 3, ch 1)--     -+ MN5
The end of SP 3 of flow 2 node 5 has been reached
time 0.105:        AP2 +---- BEACON (ch 2) ->
time 0.10502:      AP2+-      -- BEACON (ch 2)----------> MN6
time 0.10502:      AP2 +------- SAPSD_DATA (flow 4, SP 5, ch 2) ->      MN6
time 0.1053:       AP2+-     -- SAPSD_DATA (flow 4, SP 5, ch 2)------------> MN6
time 0.10531:      AP2        <- SAPSD_ACK (flow 4, SP 5, ch 2)----------------+ MN6
time 0.10531:      AP2 <---------- SAPSD_ACK (flow 4, SP 5, ch 2)--     -+ MN6
The end of SP 5 of flow 4 node 6 has been reached
time 0.10537:      AP2 +------- SAPSD_DATA (flow 5, SP 3, ch 2) ->      MN6
time 0.10565:      AP2+-     -- SAPSD_DATA (flow 5, SP 3, ch 2)------------> MN6
time 0.10566:      AP2        <- SAPSD_ACK (flow 5, SP 3, ch 2)----------------+ MN6
time 0.10566:      AP2 <---------- SAPSD_ACK (flow 5, SP 3, ch 2)--     -+ MN6
The end of SP 3 of flow 5 node 6 has been reached
time 0.10572:      Node 5 +---- PROBE_REQUEST (ch 2) ->      AP2
time 0.10572:      MN6 +------- UAPSD_SP_TRIGGER (ch 2) ->      AP2
time 0.10581:      MN5+-     -- PROBE_REQUEST (ch 2)------------> AP2
time 0.10588:      MN6+-     -- UAPSD_SP_TRIGGER (ch 2)------------> AP2
time 0.10588:      MN6      <- UAPSD_DATA (ch 2) -----------+ AP2
time 0.10616:      MN6<----------- UAPSD_DATA ch(2)--      -+ AP2
time 0.10617:      node 6+------------- UAPSD_ACK ch(2)-->      node 2
time 0.10617:      MN6+-        -- UAPSD_ACK (ch 2)---------------> AP2
The end of an U-APSD SP  of node 6 has been reached
time 0.10731:      Node 5    <- PROBE_RESPONSE (ch 2)---------+ AP2
time 0.1074:       MN5<--------- PROBE_RESPONSE (ch 2)-      -+ AP2
time 0.11:         AP3 +---- BEACON (ch 3) ->
time 0.11002:      AP3+-      -- BEACON (ch 3)----------> MN7
time 0.11002:      MN7 +------- SAPSD_DATA (flow 7, SP 5, ch 3) ->      AP3
time 0.1103:       MN7+-     -- SAPSD_DATA (flow 7, SP 5, ch 3)------------> AP3
time 0.11031:      MN7        <- SAPSD_ACK (flow 7, SP 5, ch 3)----------------+ AP3
time 0.11031:      MN7 <---------- SAPSD_ACK (flow 7, SP 5, ch 3)--     -+ AP3
The end of SP 5 of flow 7 node 7 has been reached
time 0.11037:      MN7 +------- SAPSD_DATA (flow 8, SP 3, ch 3) ->      AP3
time 0.11065:      MN7+-     -- SAPSD_DATA (flow 8, SP 3, ch 3)------------> AP3
time 0.11066:      MN7        <- SAPSD_ACK (flow 8, SP 3, ch 3)----------------+ AP3
time 0.11066:       MN7 <---------- SAPSD_ACK (flow 8, SP 3, ch 3)--     -+ AP3
The end of SP 3 of flow 8 node 7 has been reached
time 0.11072:      Node 5 +---- PROBE_REQUEST (ch 3) ->      AP3
time 0.11081:      MN5+-     -- PROBE_REQUEST (ch 3)------------> AP3
time 0.11231:      Node 5    <- PROBE_RESPONSE (ch 3)---------+ AP3
time 0.11241:      MN5<--------- PROBE_RESPONSE (ch 3)-      -+ AP3
time 0.115:        AP4 +---- BEACON (ch 4) ->
time 0.11502:      AP4+-      -- BEACON (ch 4)----------> MN8
time 0.11502:      MN8 +------- SAPSD_DATA (flow 10, SP 5, ch 4) ->      AP4
time 0.1153:       MN8+-     -- SAPSD_DATA (flow 10, SP 5, ch 4)------------> AP4
```

```
time 0.11531:	MN8 	<- SAPSD_ACK (flow 10, SP 5, ch 4)----------------+ AP4
time 0.11531:	MN8 <---------- SAPSD_ACK (flow 10, SP 5, ch 4)--	-+ AP4
The end of SP 5 of flow 10 node 8 has been reached
time 0.11537:	MN8 +------- SAPSD_DATA (flow 11, SP 3, ch 4) ->	AP4
time 0.11565:	MN8+-	-- SAPSD_DATA (flow 11, SP 3, ch 4)------------> AP4
time 0.11566:	MN8 	<- SAPSD_ACK (flow 11, SP 3, ch 4)----------------+ AP4
time 0.11566:	MN8 <---------- SAPSD_ACK (flow 11, SP 3, ch 4)--	-+ AP4
The end of SP 3 of flow 11 node 8 has been reached
time 0.11572:	Node 5 +---- PROBE_REQUEST (ch 4) ->	AP4
Time 0.12072 : Probe timeout has expired, and probe response has not been received. Node 5 AP4
time 0.12222:	MN5 +---- HANDOVER_REQUEST (ch 3) ->	AP3
time 0.12231:	MN5+-	-- HANDOVER_REQUEST (ch 3)------------> AP3
time 0.12381:	MN5	<- HANDOVER_RESPONSE (ch 3)---------+ AP3
time 0.12388:	MN5<--------- HANDOVER_RESPONSE (ch 3)-	-+ AP3
time 0.13002:	AP2 +------- SAPSD_DATA (flow 4, SP 6, ch 2) ->	MN6
time 0.1303:	AP2+-	-- SAPSD_DATA (flow 4, SP 6, ch 2)------------> MN6
time 0.13031:	AP2	<- SAPSD_ACK (flow 4, SP 6, ch 2)----------------+ MN6
time 0.13031:	AP2 <---------- SAPSD_ACK (flow 4, SP 6, ch 2)--	-+ MN6
The end of SP 6 of flow 4 node 6 has been reached
time 0.13037:	MN6 +------- SAPSD_DATA (flow 6, SP 2, ch 2) ->	AP2
time 0.13065:	MN6+-	-- SAPSD_DATA (flow 6, SP 2, ch 2)------------> AP2
time 0.13066:	MN6	<- SAPSD_ACK (flow 6, SP 2, ch 2)----------------+ AP2
time 0.13066:	MN6 <---------- SAPSD_ACK (flow 6, SP 2, ch 2)--	-+ AP2
The end of SP 2 of flow 6 node 6 has been reached
time 0.13502:	AP3 +------- SAPSD_DATA (flow 7, SP 6, ch 3) ->	MN7
time 0.1353:	AP3+-	-- SAPSD_DATA (flow 7, SP 6, ch 3)------------> MN7
time 0.13531:	AP3	<- SAPSD_ACK (flow 7, SP 6, ch 3)----------------+ MN7
time 0.13531:	AP3 <---------- SAPSD_ACK (flow 7, SP 6, ch 3)--	-+ MN7
The end of SP 6 of flow 7 node 7 has been reached
time 0.13537:	AP3 +------- SAPSD_DATA (flow 9, SP 2, ch 3) ->	MN7
time 0.13565:	AP3+-	-- SAPSD_DATA (flow 9, SP 2, ch 3)------------> MN7
time 0.13566:	AP3	<- SAPSD_ACK (flow 9, SP 2, ch 3)----------------+ MN7
time 0.13566:	AP3 <---------- SAPSD_ACK (flow 9, SP 2, ch 3)--	-+ MN7
The end of SP 2 of flow 9 node 7 has been reached
time 0.14002:	MN8 +------- SAPSD_DATA (flow 10, SP 6, ch 4) ->	AP4
time 0.1403:	MN8+-	-- SAPSD_DATA (flow 10, SP 6, ch 4)------------> AP4
time 0.14031:	MN8	<- SAPSD_ACK (flow 10, SP 6, ch 4)----------------+ AP4
time 0.14031:	MN8 <---------- SAPSD_ACK (flow 10, SP 6, ch 4)--	-+ AP4
The end of SP 6 of flow 10 node 8 has been reached
time 0.14037:	AP4 +------- SAPSD_DATA (flow 12, SP 2, ch 4) ->	MN8
time 0.14065:	AP4+-	-- SAPSD_DATA (flow 12, SP 2, ch 4)------------> MN8
time 0.14066:	AP4	<- SAPSD_ACK (flow 12, SP 2, ch 4)----------------+ MN8
time 0.14066:	AP4 <---------- SAPSD_ACK (flow 12, SP 2, ch 4)--	-+ MN8
The end of SP 2 of flow 12 node 8 has been reached
time 0.15502:	MN6 +------- SAPSD_DATA (flow 4, SP 7, ch 2) ->	AP2
time 0.1553:	MN6+-	-- SAPSD_DATA (flow 4, SP 7, ch 2)------------> AP2
time 0.15531:	MN6	<- SAPSD_ACK (flow 4, SP 7, ch 2)----------------+ AP2
time 0.15531:	MN6 <---------- SAPSD_ACK (flow 4, SP 7, ch 2)--	-+ AP2
The end of SP 7 of flow 4 node 6 has been reached
```

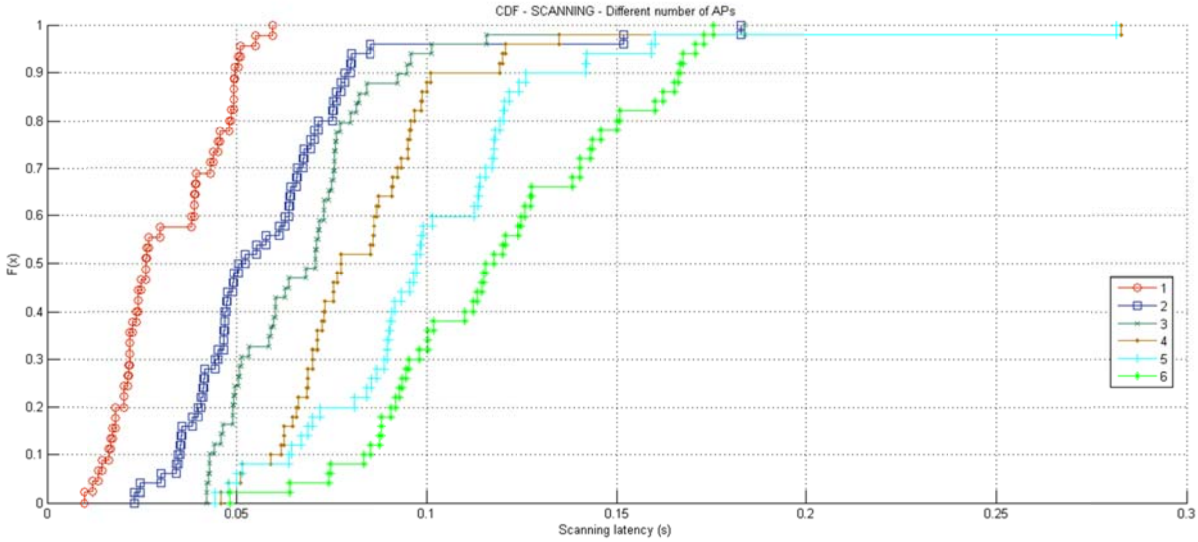
Appendix B Cumulative distribution functions (CDFs)



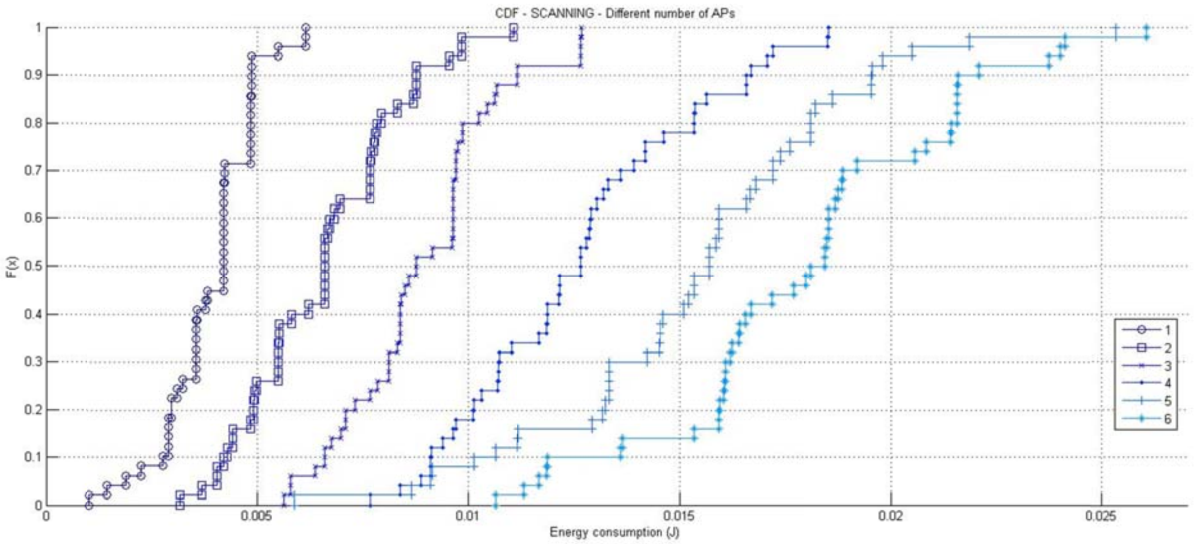
Appendix Figure B-1: CDF correspondent to Figure 4 7 (Influence of the number of APs to be scanned over the scanning latency)



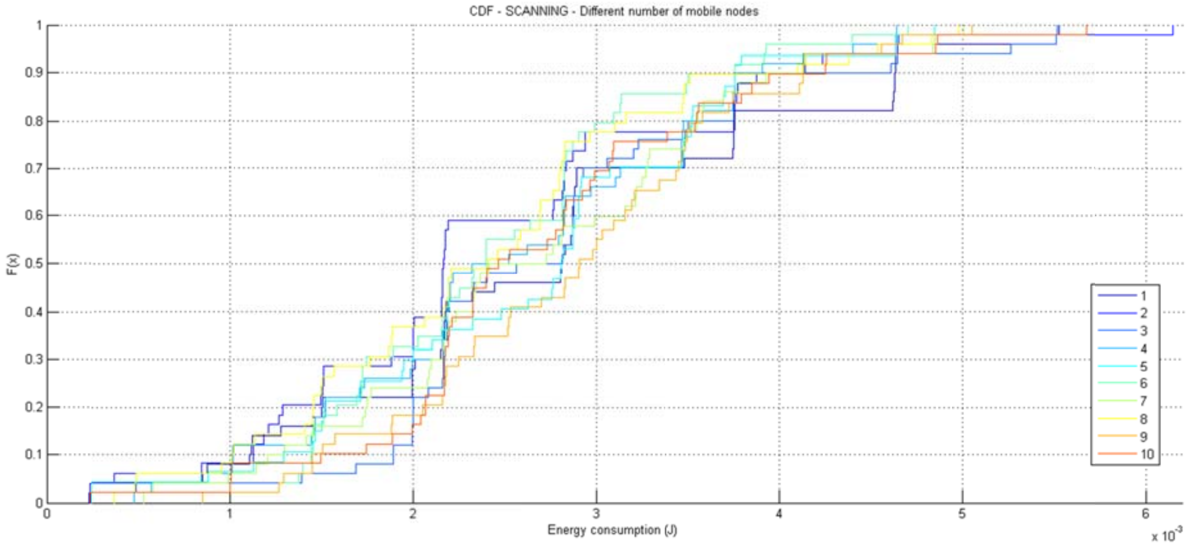
Appendix Figure B-2: CDF correspondent to Figure 4 11 (Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's coverage)



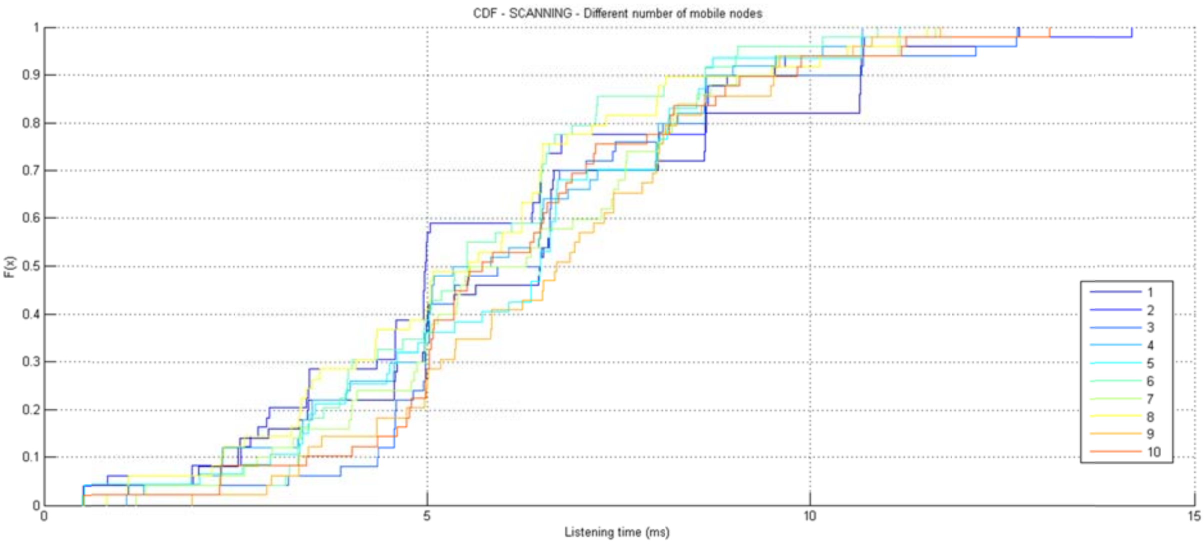
Appendix Figure B-3: CDF correspondent to Figure 4 10 (Influence of the number of APs to be scanned over the scanning latency, when half of the APs are out of the scanning node's coverage)



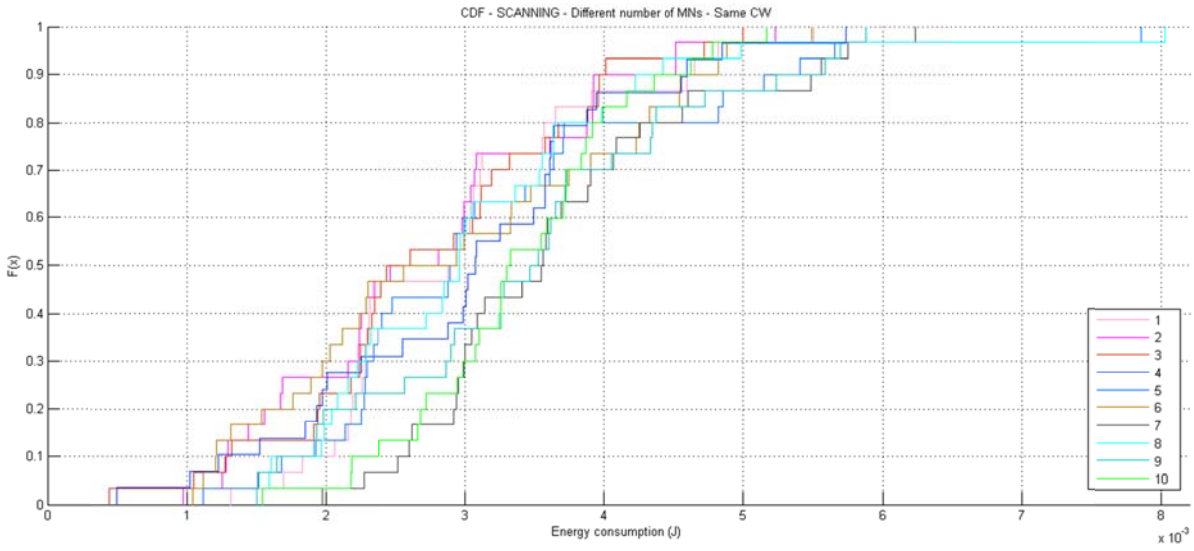
Appendix Figure B-4: CDF correspondent to Figure 4 11 (Influence of the number of APs to be scanned over the energy consumption, when half of the APs are out of the scanning node's coverage)



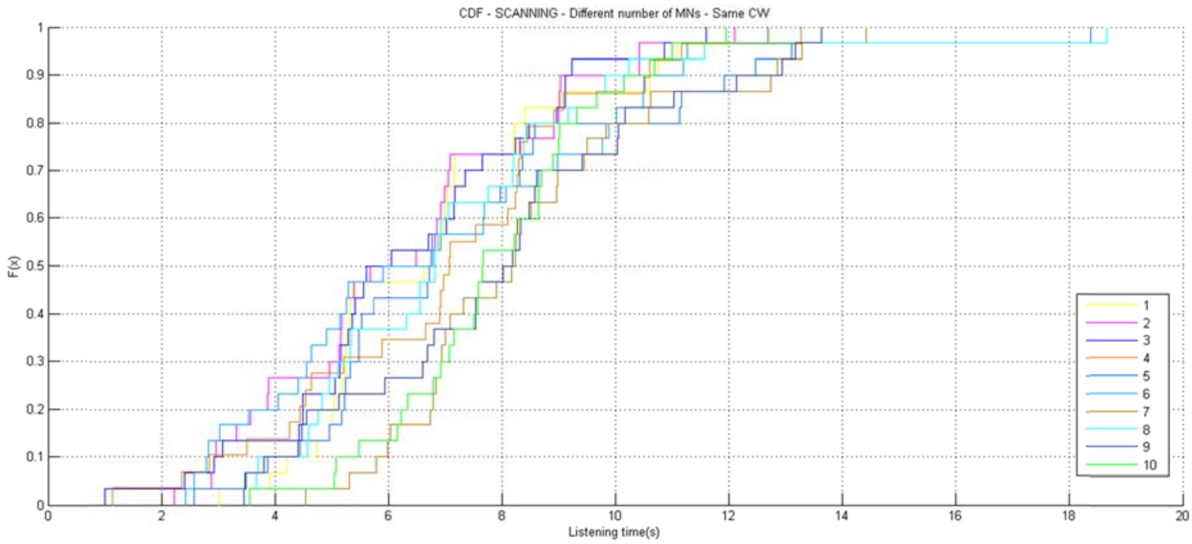
Appendix Figure B-5: CDF correspondent to Figure 4 12 (Influence of the number of MNs attached to the target APs over the scanning MN's energy consumption)



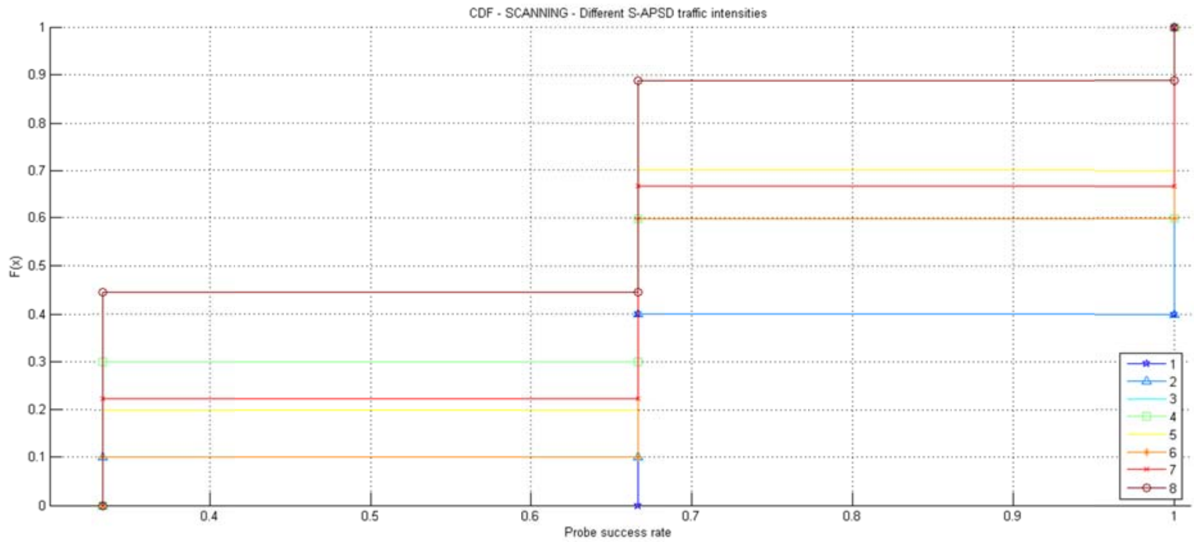
Appendix Figure B-6: CDF correspondent to Figure 4 13 (Influence of the number of MNs attached to the target APs over the listening times associated to scanning operations)



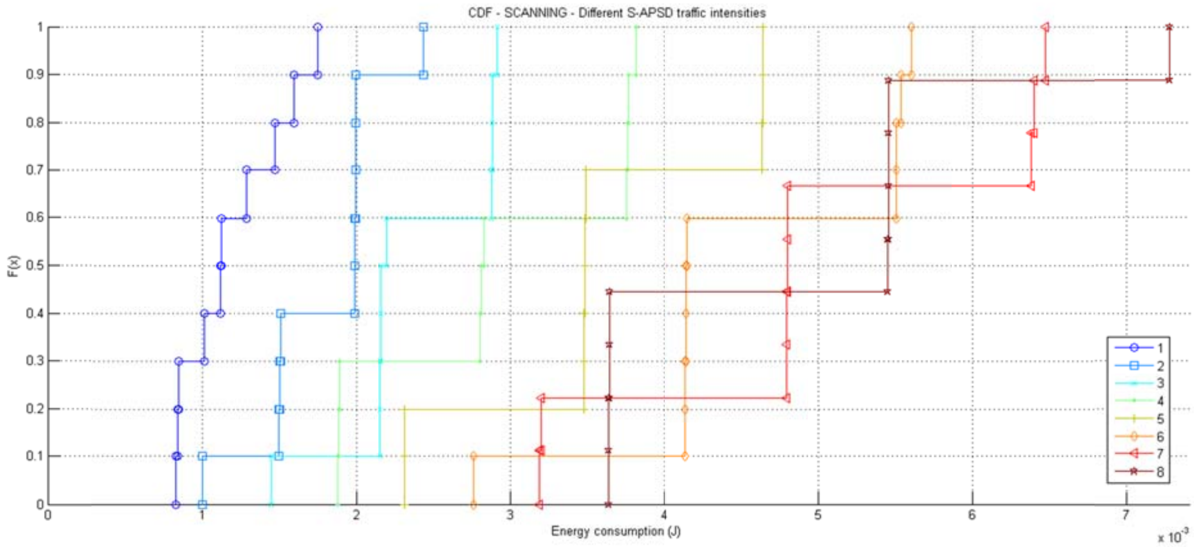
Appendix Figure B-7: CDF correspondent to Figure 4 14 (Influence of the number of MNs attached to the target APs over the scanning MN's energy consumption when frames associated to scanning operations use the same CW as the other frames)



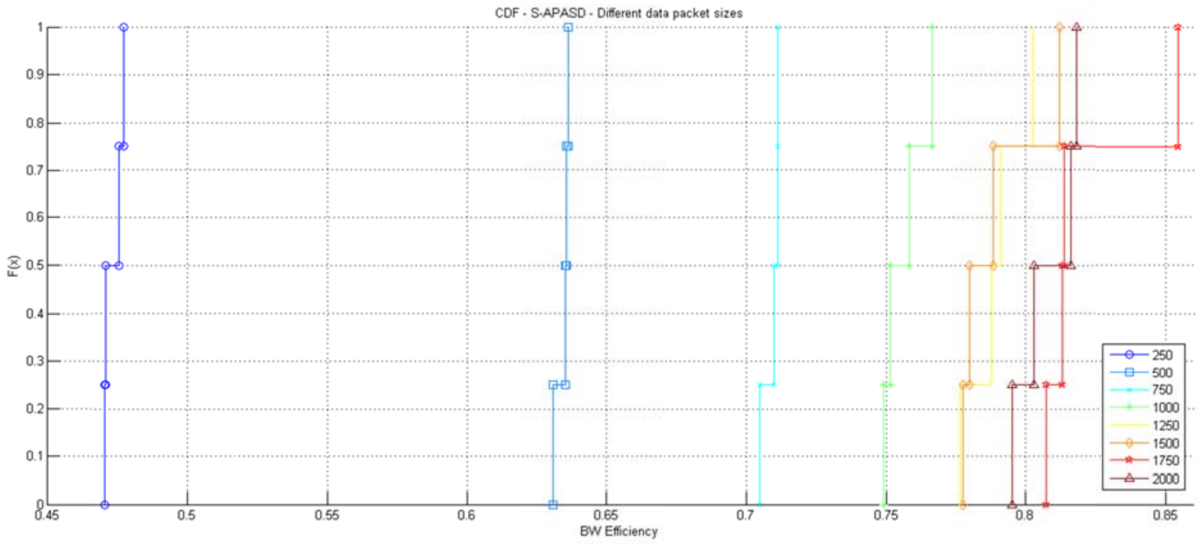
Appendix Figure B-8: CDF correspondent to Figure 4 15 (Influence of the number of MNs attached to the target APs over the listening times associated to scanning operations when frames associated to scanning operations use the same CW as the other frames)



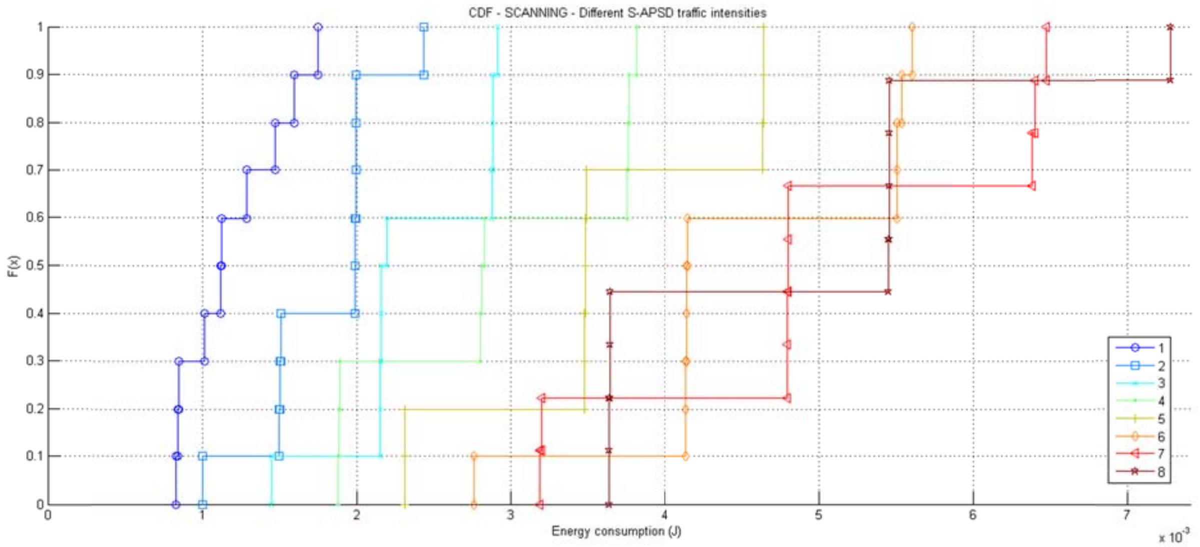
Appendix Figure B-9: CDF correspondent to Figure 4 16 (Probe success rate over different S-APSD traffic conditions)



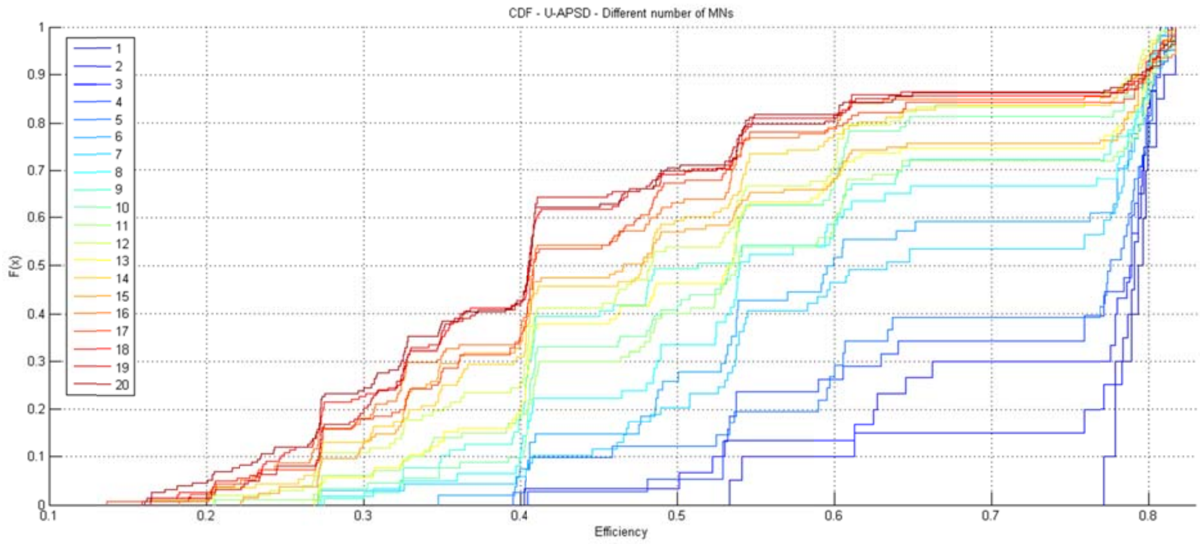
Appendix Figure B-10: CDF correspondent to Figure 4 17 (Energy consumption of the scanning node as a function of the number of active flows)



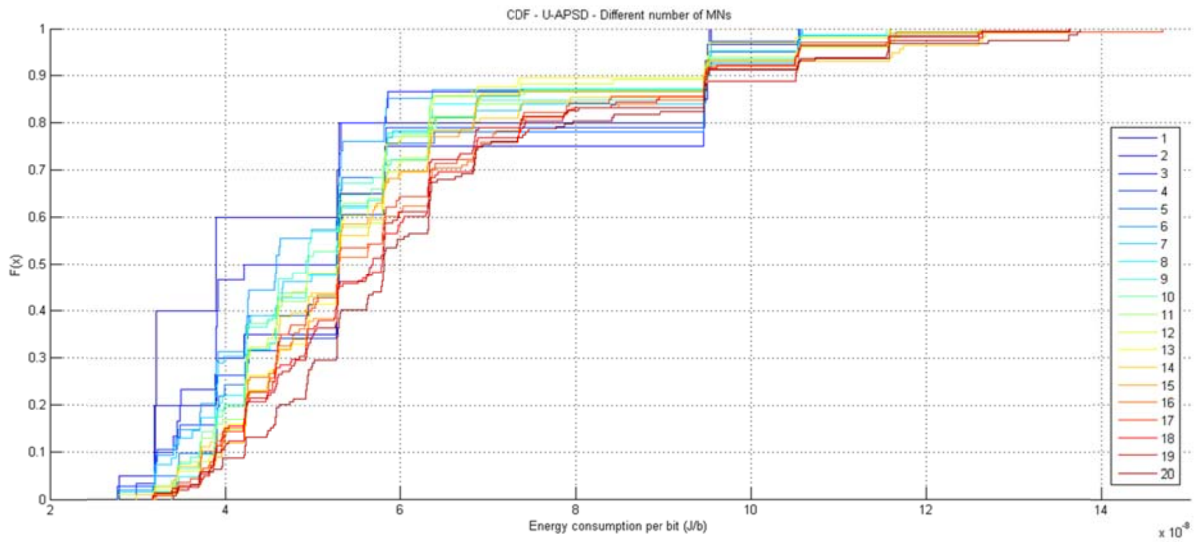
Appendix Figure B-11: CDF correspondent to Figure 4 19 (Flows' bandwidth efficiency as a function of data packets' size)



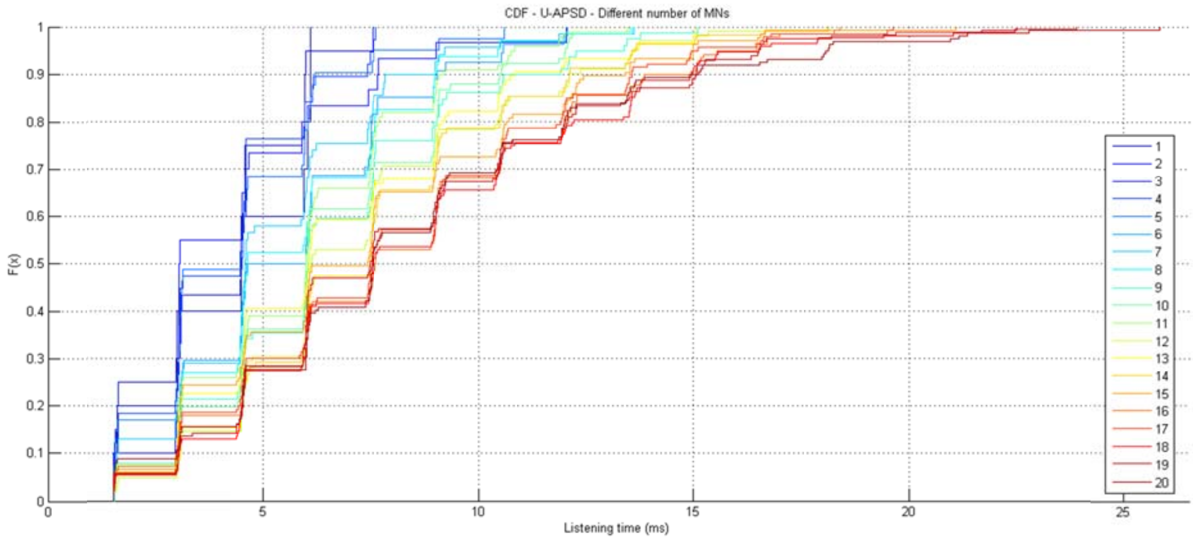
Appendix Figure B-12: CDF correspondent to Figure 4 21 (Mobile node's energy consumption per bit while it has one S-APSD active flow as a function of flow's TXOP)



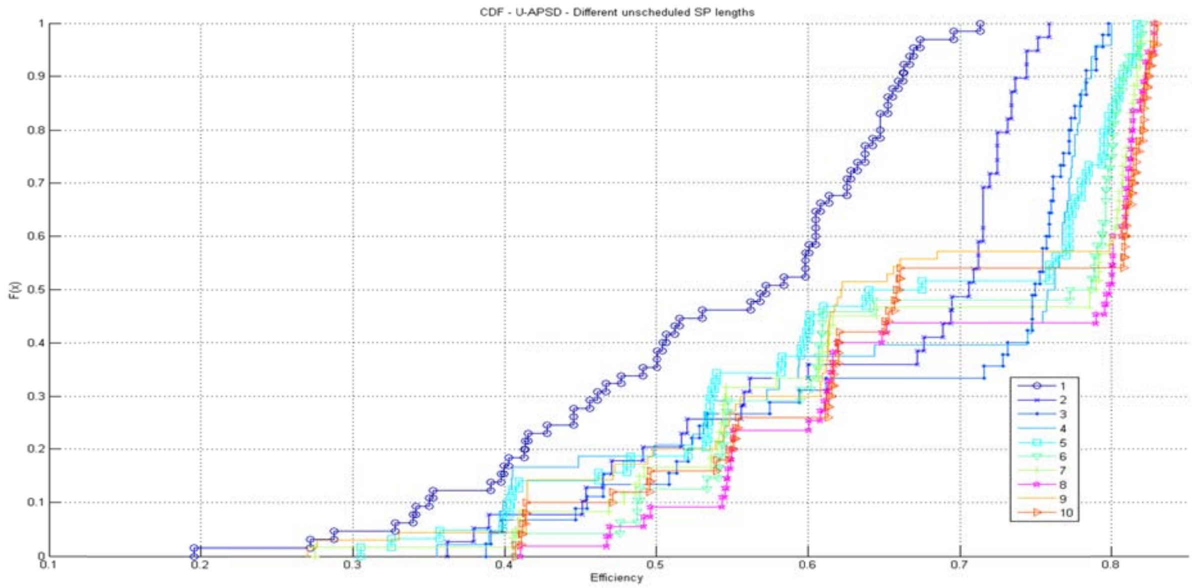
Appendix Figure B-13: CDF correspondent to Figure 4 22



Appendix Figure B-14: CDF correspondent to Figure 4 23 (MN's energy consumption per bit as a function of the number of MNs associated to an AP)



Appendix Figure B-15: CDF correspondent to Figure 4 24 (Time that a MN spends in listening state while using U-APSD, as a function of the number of MNs associated to an AP)



Appendix Figure B-16: CDF correspondent to Figure 4 25 (CPs' efficiency as a function of unscheduled SPs' length)

TRITA-ICT-EX-2014:16