Contributions à l'amélioration de l'utilisation des ressources dans les réseaux de paquets sans fil

Yacine Ghamri-Doudane

Mémoire présenté par

Mohamed Yacine GHAMRI DOUDANE

en vue de l'obtention du diplôme

d'Habilitation à Diriger des Recherches
de l'Université Paris-Est

CONTRIBUTIONS A L'AMELIORATION DE L'UTILISATION DES RESSOURCES DANS LES RESEAUX DE PAQUETS SANS FIL

Habilitation prévue le *13 décembre 2010*, devant le Jury :

Rapporteur :	RAOUF BOUTABA (PR., UNIVERSITY OF WATERLOO, CANADA)
Rapporteur :	OLIVIER FESTOR (DR INRIA, INRIA NANCY – GRAND-EST)
Rapporteur :	JOSE-MARCOS NOGUEIRA (PR., UNIVERSIDADE FEDERAL DE MINAS GERAIS, BRÉSIL)
Examinateur :	ANDRÉ-LUC BEYLOT (PR., INPT/ENSEEIHT)
Examinateur :	WALID DABBOUS (DR INRIA SOPHIA ANTIPOLIS)
Examinateur :	MICHEL DIAZ (DR CNRS, LAAS TOULOUSE)
Examinateur :	GILLES ROUSSEL (PR., UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE)

Abstract

Since two decades, we are observing a continuous evolution in wireless communications and networking technologies. This evolution creates nowadays an unprecedented demand for accessing communication services anywhere, anytime and using any device. This trend is also encouraging the rapid development of more and more novel communication services and applications. These latter require more capacity and more quality of service from the network, as well as more efficiency from the communication device. One of the main goals in current research is to design new communication solutions that are more robust and resource-efficient. In this context, our aim is to propose novel mechanisms and protocols that target the "*improvement of the resource usage in wireless packet networks*". Improving the resource usage can be realized at two complementary levels: the packet-level and the connection-level. In our research, we addressed and solved different facets of the resource usage issue at both levels. At the connection level, our main concern was to maintain the experienced quality of service while the users are willing to move transparently over the diverse available wireless packet networks; while at the packet-level the issue we addressed was to improve the quality of service experienced by the packet flows generated by novel data and multimedia communication services. This report presents the major research results we obtained in this field.

Keywords: Wireless Local Area Networks, Mobile Ad hoc Networks, Broadband Wireless Access Networks, Vehicular Networks, QoS, Energy-Efficiency, Performance Evaluation, Simulation, Analytical Models, and Experimentation.

Content

Part 1. Research Activity

Chapter 1. Introduction

1.1 CONTEXT OF THE RESEARCH

Over the last two decades, we saw a major evolution of the telecommunication sector regulation. This lead to the definition of new business models and the emergence of new players such as service providers, content providers, virtual operators, and so on. This period also seen the ascend of mobile wireless access thanks to the great success of the Global System for Mobile communication (GSM) and the huge number of Wireless Local Area Networks (WLANs) deployments. Consequently, the Internet has been witnessing a perpetual growth in data traffic and more and more users are getting mobile and nomadic. Internet access through smartphones and notebooks became though a reality. For instance, 22% of mobile phone subscribers in the US were accessing Internet in 2008 while being mobile and it is expected that this percentage will double by 2013. Thanks to the increasing rise in smartphone popularity, similar trends are also experienced in Europe and Asia. All these have fostered the development of new data-intensive and multimedia services that are coming out day after day.

The proliferation of new data-intensive and multimedia services has created a rising demand for robustness, capacity and throughput increase especially in the wireless access part. These multimedia services impose to the network a certain number of constraints in order to operate correctly. In addition to the bandwidth requirements, these constraints can be expressed in terms of delay, jitter, packet loss rates or even terminal's energy consumption. The development of mechanisms that allow using efficiently the wireless and terminal resources in such a realm is the mandatory step in order to be able to offer guarantees to the above mentioned constraints. Indeed, an efficient usage of the scarce wireless and terminal resources is a must in order for the user to get a quality of experience that is close to the one offered by wired networks.

In parallel to that, the evolution of the telecommunication and Internet sectors also fueled the emergence of a multitude of new wireless technologies and paradigms allowing extending the boundaries of telecommunication and Internet access. From this multitude of wireless packet networks, one can quote Wireless Local Area Networks (WLANs), Broadband Wireless Access Networks (BWANs), as well as Mobile Ad hoc Networks (MANETs) and their diverse use cases. One of the major challenges became though to offer means to improve the use of the wireless resources in order to handle efficiently Quality of Service (QoS)-demanding applications by this new set of wireless packet networks. The wireless resource usage issues raised by these networks are twofold:

1. *Packet-level issues*: or how to improve the Quality of Service experienced by the packet flows generated by data and multimedia services ;

2. *Connection-level issues*: or how to maintain the experienced Quality of Service while the users are mobile.

Thus, within this general context, our research work concentrated on the "*improvement of the resource usages in wireless packet networks*". Our target is to address both the packet-level issues in different wireless packet network contexts as well as the connection-level issues while the user requires roaming seamlessly across a heterogeneous wireless packet network environment.

1.2 IMPROVING THE RESOURCE USAGE IN WIRELESS PACKET NETWORKS

The emergence of new wireless technologies in the late 90s has given rise to a set of novel usage modes as well as new application opportunities. Indeed, these new wireless technologies can either be used to seamlessly connect a mobile user and its usual applications on the Internet, or be used to connect directly mobile users among themselves (i.e. *ad hoc*) offering a new range of applications suited to such environments. These new wireless technologies as well as their novel usage modes pose multitude challenges and still require a major research effort before being "*deployable*". It is in this context that our research activity had been conducted. More specifically, we are interested on the "*improvement of the resource usages in wireless packet networks*". Different facets of this problem are discussed. More specifically, these concerns:

1. Mechanisms for improving the resource usage at the connection-level. This concerns the seamless mobility management in heterogeneous wireless packet networks.

2. Mechanisms for improving the resource usage at the packet-level. We concentrated on the following case studies: *(i)* the improvement of the Transport Control Protocol (TCP) behavior in wireless packet networks such as WLANs and MANETs, *(ii)* the improvement of the resource management within wireless packet networks such as 802.11e WLANs and Relay-based BWANs, and *(iii)* the improvement of the unicast and dissemination traffic performances in vehicular packet networks.

In the following we briefly summarize each of these facets and case studies. A more detailed discussion on the results we obtained will then be given throughout this document.

1.2.1 Mobility over Heterogeneous Wireless Packet Networks

Advances in wireless communication systems and handheld devices are driving an evolution towards ubiquitous and seamless service delivery across multiple wireless access systems. Mobile users will be Always Best Connected (ABC) anywhere and at anytime to diverse access technologies. Future users are expected to use this diversity and take advantage from the interworking between wireless access systems in order to maximize their profitability and/or improve the perceived QoS. New network selection and mobility management mechanisms are therefore needed to handle the complexity of the seamless handover and to select the best available wireless access network that satisfies the user's QoS requirements at the lowest cost and energy use. Our contributions in this context consist on proposing and validating these different mechanisms that will achieve the above mentioned objective. These contributions are summarized in the following.

Utility-Based Access Network Selection: Network selection is one of the most important elements of the mobility management process when the user is willing to move across heterogeneous wireless packet networks. It is the key to being "Always Best Connected". This is imperative not only in providing end-users with the most suitable access network but also in providing operators with the highest spectrum utilization and revenue. To design such a mechanism, we first analyze the utility theory with the aim of defining an appropriate decision metric for access network selection. After reviewing existing utility models and highlighting their limitations, we propose new single-criterion and multi-criteria utility forms to best capture user satisfaction and sensitivity to varying access network characteristics. We thus demonstrate that a network selection based on these new utility forms has two advantages: *(i)* effectively improves the capacity of the end-user terminals to select the best access network, and *(ii)* helps the operators to optimize the use of their wireless resources.

Terminal-controlled mobility management framework: In this contribution we investigate the issues related to handover management of mobile terminals moving across heterogeneous wireless packet networks. Most of today's mobile terminals are equipped with multiple wireless radio interfaces and have a limited battery lifetime. Thus, seamless mobility and power utilization efficiency become two important aspects of the handover management. We propose in this work a user-centric network selection, a power-saving interface management and an adaptive handover initiation solution at the terminal side to support seamless terminal-initiated and terminal-controlled vertical handover. The proposed access network selection is situation-aware and application-aware to suit different communication contexts. It enables terminals to select the most suitable access network according to various access network characteristics. Multiple wireless interfaces of a terminal device are handled in both idle and active communication modes to optimize the power consumption. We also address an adaptive handover initiation scheme to assist the service continuity and maintain the QoS at the connection-level. Overall, using simulations and analytical studies, the proposed terminal-controlled mobility management framework proved to be suitable and efficient in achieving its objectives: i.e. maximizing the users' profitability and/or improving its perceived QoS.

1.2.2 Improving TCP Behavior in Wireless Packet Networks: WLAN and MANET Case Studies

Transport Control Protocol (TCP) is a transport protocol that aims at ensuring high reliability by guaranteeing the reception of data packets. Today, it is the most commonly-used reliable transport protocol in the Internet and it is supported by almost all Internet applications. However, TCP was designed primarily for wired networks to address network congestion, which is the main cause for data packet loss in these networks. Conversely, wireless packet networks are characterized by various other packet loss causes. These are due to the intrinsic characteristics of wireless channels (e.g. signal fading, interference, obstacles, and environment effects) that might obstruct the proper reception of data packets at the other end. These other packet loss causes may lead TCP to be inefficient when used in wireless packet networks. Indeed, this one may reduce its throughput unnecessarily after a packet loss, thinking that this one is due to congestion. So, there is a compelling need to adapt TCP behavior in order to avoid such reactions and improve the resource usage. In this context, we propose innovative solutions to improve TCP behavior in WLANs and MANETs. These contributions are summarized in the following.

Improving TCP Behavior in WLANs: Most existing TCP variants cannot distinguish between different packet loss causes within wireless packet networks. In this contribution, we are interested in improving TCP behavior when it experiences a short loss of the 802.11 signal, leading to segment losses and triggering inappropriately TCP congestion control mechanisms. The set of measurements we realized, in a common wireless environment with signal losses due to mobility or interferences, highlighted that there is a lack of interactions between the distinct MAC and TCP loss-recovery mechanisms. We also showed the clear interest of adapting the 802.11-MAC-layer *Retry Limit* parameter in the case of signal losses due to distance or obstacles (mobility). Thus, a first level Loss Differentiation Algorithm (LDA) acting at the MAC layer is proposed to improve TCP performances in the case of segment losses due to mobility. Hence, for a signal failure, the MAC layer reacts consequently by dynamically adapting the *Retry Limit* parameter. This adaptation allows avoiding a costly end-to-end TCP loss-recovery. Segment losses due to interferences are differentiated from those due to congestions throughout the use of a second level LDA. This latter is a cross-Layer LDA acting at the TCP layer but using a specific 802.11 parameter, the *AckFailureCount*, to realize the targeted loss differentiation. The solution integrating both LDA schemes associated to the legacy TCP implementation (i.e. *TCP New Reno*) proved to be efficient and complete in differentiating losses due to mobility, interferences or congestion, and reacting adequately to these losses. This allowed a significant improvement of the wireless resource usage by TCP in WLANs.

Improving TCP Behavior in MANETs: In order to overcome the performance limitation of TCP when used in MANETs, the aim of our work is twofold. First, a complete performance evaluation study of the different existing TCP variants over MANETs is achieved. This study allows identifying the potential improvement possibilities to enhance TCP efficiency in MANETs. Second, we propose a new TCP variant that optimizes the performance of TCP in MANETs through its ability to distinguish among, and efficiently deal with, different data packet loss situations encountered within MANETs. Indeed, the complete performance study we performed pointed out that in MANETs, TCP may suffer from three different packet loss causes: congestions, wireless-channel-related losses and losses related to link-failures. The multiple LDA schemes proposed in the literature had not been designed to cope adequately with these three packet loss causes. Indeed, these LDA schemes had been optimized to data networks where the wireless link is only the last hop. This means that they might be inadequate for multi-hop wireless packet networks such as MANETs. In order to adapt TCP behavior to MANET characteristics, we propose a new TCP variant called TCP-WELCOME (Wireless Environment, Link losses, and COngestion packet loss ModEls). TCP-WELCOME aims to: *(i)* distinguish the packet loss cause by coupling loss and delay information, then, *(ii)* trigger the appropriate packet loss recovery according to the identified loss cause. The performance evaluation, through both simulations and experimental tests, shows that TCP-WELCOME improves both energy consumption and throughput.

Furthermore, TCP-WELCOME does not change the standard as it can operate with existing TCP variants.

1.2.3 Mastering QoS in Wireless Packet Networks: WLAN and BWAN Case Studies

Next generation wireless packet networks are expected to support a multitude of applications including bandwidth-intensive multimedia services such as networked games, video on demand, wireless TV as well as audio and videoconferencing. These applications necessitate a certain level of QoS in order to operate properly. Conversely, the characteristics of wireless links make it challenging to offer such QoS. Indeed, the channel resources needed to support such multitude of applications remain limited, and therefore their proper management is necessary to ensure the required QoS. This situation remains valid in all kind of wireless packet networks. In this context, proposing innovative solutions that allows mastering the QoS in wireless packet networks constitutes an important part of our work. More precisely, our contributions explore how an efficient resource management solution can be designed in the case of two specific wireless packet networks: 802.11e WLANs as well as Relay-assisted BWANs. These contributions are summarized in the following.

Mastering the QoS in 802.11e WLANs: In the 802.11e EDCA (Enhanced Distributed Channel Access) standard, there is no guaranty in terms of throughput and delay assurance for real-time and multimedia services. Before the network gets saturated, there is no QoS problem. The problem arises once the network starts to reach saturation and a high number of flows share the limited channel resources. All the solutions that only aim to enhance the performance of the EDCA mechanism cannot resolve the performance degradation problem once the channel becomes saturated. Hence, an efficient admission control scheme in EDCA is the key to guarantee the QoS required by real-time and multimedia services in WLANs. However, when reviewing the literature in this domain, it is clear that QoS control and resource management in 802.11e WLANs to support real-time and multimedia services such as voice and video still remains an open problem. Thus, our contribution consists on filling this gap by proposing a model-based admission control algorithm that is located within the QoS Access Point (QAP). To do so, an accurate analytical model is used to predict the QoS metrics that can be achieved once a new flow is introduced in the WLAN. Based on this prediction and on the QoS constraints of already accepted (active) flows as well as of the new flow, the QAP takes its decision of admitting or rejecting the new flow. The proposed admission control scheme is fully compatible with the legacy 802.11e EDCA MAC protocol. It constitutes the missing brick in order to allow mastering the wireless resource usage in 802.11e WLANs.

Mastering the QoS in relay-assisted BWANs: Broadband Wireless Access Networks (BWAN) making use of the relaying technology, constitute one of the most promising solutions thanks to the enhancements of the system capacity and coverage they provide. However, some special features, like link-asymmetry[1], may lower its positive aspects. So, in order to fully profit from this potential gain, we argue on the fact that an efficient scheduling scheme is needed for relay stations. Indeed, such scheduling scheme is currently the missing brick in order to make an efficient use of BWAN resources. The link-asymmetry problem has various effects on the efficiency of the wireless resource usage depending on whether only one relay station or multiple relay stations are used in the system. Hence, when a single-relay system is experiencing link-asymmetry, the packets to be forwarded to mobile stations may need to be buffered in the relay station and an efficient scheduling scheme is needed there to serve them. On the other hand, when using multiple relay stations in the system, the packets need to be efficiently distributed over the different available relay stations and then served accordingly minimizing the effect of link-asymmetry. The efficiency in both situations consists on the reduction of the potential resource wastage, denoted as channel-holes, due to link-asymmetry. So, in order to reach such efficiency, we propose two novel scheduling schemes, the channel-hole-based scheduling and the channel-hole-based cooperative

[1] Relay-assisted networks make use of three kinds of stations Base Stations (BS), Relay Stations (RS) and Mobile Stations (MS). Hence, a communication from a BS to an MS may be relayed by the RS when the direct link is not possible (i.e. MS outside the BS coverage). We say that there is a link-asymmetry when the channel states of the BS-RS and RS-MS links are opposite (i.e. one is experiencing bad conditions and the other good conditions).

scheduling algorithms, targeting single-relay and multiple-relay systems, respectively. These are built with the aim to improve the wireless resource usage in relay-assisted BWANs.

1.2.4 Vehicular Packet Networks: Geo-localized Communications in Urban Areas

Inter-Vehicle Communications (IVC) are attracting considerable attention from the research community and the automotive industry, where it is beneficial in providing Intelligent Transportation Systems (ITS) as well as assistance services for drivers and passengers. In this landscape, Vehicular Packet Networks are emerging as a novel category of wireless packet networks, spontaneously formed between moving vehicles equipped with wireless interfaces. These can be seen as a specific MANET use case with some distinguished characteristics such as high mobility, potentially large scale, and network partitioning. However, these distinguished characteristics introduce several challenges that can greatly impact the performances of vehicular packet networks if not dealt with adequately. In this context, our main goal is to propose new solutions that efficiently adapt to vehicular packet networks characteristics and applications within an urban environment. The ultimate objective of these solutions is to improve the wireless resource usage in such networks. They are summarized in the following.

Geo-Routing in Vehicular Packet Networks: Multi-hop data delivery between vehicles is an important aspect for the support of IVC-based applications. Although, data dissemination and routing have been extensively addressed, many unique characteristics of vehicular packet networks together with the diversity in promising applications offer newer research challenges. In order to cope with these challenges, we propose GyTAR (improved Greedy Traffic Aware Routing protocol), an intersection-based geographical routing protocol, capable to find robust and optimal routes within urban environments. The main principle behind GyTAR is the dynamic and in-sequence selection of intersections through which data packets are routed to reach their destination. The intersections are chosen considering parameters such as the remaining distance to the destination and the variation in vehicular traffic. This latter is obtained through IFTIS (Infrastructure-Free Traffic Information System), a new protocol we designed for vehicular traffic estimation, while getting rid of any infrastructure need. Furthermore, data forwarding between intersections in GyTAR adopts an improved greedy 'carry-and-forward' mechanism. The evaluation of GyTAR shows significant performance improvement in comparison to other existing routing approaches leading to a more efficient use of the vehicular network resources.

Geo-localized Virtual Infrastructure for Dissemination in Vehicular Packet Networks: One of the main challenges restraining the deployment of ITS applications is the inefficiency of data dissemination in Vehicular Packet Networks. Several ITS applications, such as accident warning and parking lot advertisements, are based on this multi-hop communication mode. Hence, supporting future large-scale vehicular networks is expected to require a combination of fixed roadside infrastructure (e.g. Road Side Units, RSU) and mobile in-vehicle technologies (e.g. On Board Units, OBU). The need for an infrastructure, however, considerably decreases the deployment area of ITS applications. In this context, we propose a self-organizing mechanism to emulate a geo-localized virtual infrastructure (GVI). This latter is emulated by a bounded-size subset of vehicles currently populating the geographic region where the virtual infrastructure is to be deployed. The GVI is designed in order to help the efficient support of dissemination-based applications in vehicular packet networks without the need to install a costly infrastructure. Analytical and simulation results show that the proposed GVI mechanism can periodically disseminate data within an intersection area, efficiently utilize the limited bandwidth, and ensure high delivery ratio.

1.3 DOCUMENT ORGANIZATION

This document contains two parts. A first part dedicated to research activities and a second part composed of a detailed curriculum vitae. The research activities part is divided into six chapters. After having introduced our main contributions in this chapter, Chapter 2 focuses on our results related to the improvement of the wireless resource usage at the connection-level. Then, the following three chapters are dedicated to our contributions related to the improvement of the wireless resource usage at the packet-level. We first discuss the improvement of TCP behavior in wireless packet networks in Chapter

3. Then, Chapter 4 gives a detailed description of our results related to resource management solutions allowing mastering the QoS in WLANs and BWANs. Discussion of our contributions related to the challenging vehicular packet networks follows in Chapter 5. Finally, Chapter 6 summarizes our conclusions and details our main perspectives.

Chapter 2. Mobility over Heterogeneous Wireless Packet Networks

Following the explosive growth of the Internet during the last two decades, the current unprecedented expansion of wireless technology promises an even greater effect on how people communicate, interact and enjoy their entertainment. The growing advances in research and development of wireless communication technologies along with the increasing capabilities of electronic devices are driving an evolution towards ubiquitous services to mobile users. Wireless networks become increasingly interoperable with each other and with the high-speed wired networks. This reflects a paradigm shift towards new generations of mobile networks where seamless mobility across heterogeneous networks becomes fundamental.

In the following, we explore how the mobility management can be achieved and improved under a user-centric terminal-controlled approach. We will first remind the context of our research as well as the challenges that faces terminal-controlled mobility-management in a heterogeneous wireless context. This will be followed by the description of our two contributions and the results we obtained in this domain. These contributions concern utility-based access network selection and terminal-controlled mobility management framework. This chapter ends with a brief conclusion summarizing our research experience in this domain.

2.1 Research Context

Advances in wireless communication systems and handheld devices are driving an evolution towards ubiquitous and seamless service delivery across multiple wireless access systems. Mobile users will be Always Best Connected (ABC), anywhere and at anytime, to diverse access technologies such as Global System for Mobile Communications (GSM), General Packet Radio Service (GPRS), Universal Mobile Telecommunications Systems (UMTS), Wireless Local Area Network (WLAN), and Worldwide Interoperability for Microwave Access (WiMAX). Future users will need this diversity and the interworking between access systems in order to maximize their profitability and/or improve their perceived Quality of Service (QoS).

Users are expected to be always best connected through different available access networks when they move from one place to another (at home, in the office, on the bus, on the train, in the shopping mall, in the cafe...). For example, a videoconference can transparently switch from an enterprise Wireless Local Area Network (WLAN) to the traditional cellular environment when driving home and to the fixed home network when arrived. In fact, users can access and maintain a seamless connectivity anywhere, anytime via any access technology owned by any operator to use any available service. Handovers between the technologies should be transparent to users, allowing a simplified and seamless on-the-move experience. In summary, *"seamless mobility is predicated on enabling a user to accomplish his or her tasks without regard to technology, type of media, or device, facilitating freedom of movement while maintaining continuity of applications experience"* [1].

In parallel to this evolution, the telecom market is facing a migration from network centricity towards user centricity. In the current network-centric approach, operators keep tight control over users so that their network is used to its greatest potential. End-users can only influence their preferences in a limited way. In the context of deregulated telecom market, the network-controlled handover management exhibits some serious limitations in the service continuity maintenance while handing over between two different network domains operated by different network operators. Among these limitations, we can find complex issues such as security context transfer and data switching management. We believe that a user-centric vision will be a mandatory evolution trend as it represents the most efficient way to ensure an ABC service.

In the user-centric vision, users will have greater control and will be able to select the access network with which they are mostly satisfied. Their terminals are in this strong position because they have access to information on their own capabilities and can store user preferences. Most importantly, they have knowledge of both serving and neighboring access networks. The terminals will then be able to trigger the handover at the right instant to achieve seamless access network switching.

The target of our research, presented in this chapter, is to study the feasibility of a terminal-controlled mobility management in presence of heterogeneous wireless access networks. The final objective of this research is to show that using such approach we are able to improve the users' perceived Quality of Service and/or maximize their profitability.

2.2 Main Challenges

Mobile systems increasingly become an inseparable part of our daily lives in various branches of living (e.g., work, education, entertainment, health care, commerce...). In parallel to that, people are looking for a life that is more enriched and cultural, more flexible and diversified, more comfortable and safe, and more personal and convenient. Clearly, people expect that the next generation of mobile communication systems will provide something more than just "faster speed". Facing up to the ubiquitous access service, users will plan to take advantage of the providers' competition as well as the diversity of available communication resources. Users are always attracted by the easy-to-use multi-modal natural human interfaces like voice and gestures but also are willing to control and customize their communication environment. As mobile terminals are evolving towards being more intelligent and more powerful, they can aid users to handle the control without any expertise. One part of this control resides in the mobility management with the aim to maximize their profitability and/or improve their perceived Quality of Service (QoS). In this work, we explore how the mobility management can be achieved and improved under the user-centric terminal-controlled approach.

In heterogeneous networks, the interworking between different access technologies can be distinguished into tight, loose and very loose couplings [2–7]. The tight-coupling approach requires a major modification to the access network architecture. Such modification is problematic as the deployed infrastructure is usually not under the control of a single operator. On the contrary, the loose-coupling does not require changes to the existing infrastructure. The common points in most of the tight and loose coupling solutions are: the interconnected access networks' operators have a close agreement among themselves and mobility management is network-controlled. Users cannot maintain the connectivity while handing over between two uncoordinated access networks even if they have access right to both networks. The current solutions do not provide any means for users to introduce their preferences in the handover management. Contrary to the tight and loose coupling approaches, we propose to go further by using a very loose coupling solution, in which mobility management becomes a third party service. Thus, we will be able to give the user the ability to personalize their mobility service. The handover can be seamlessly achieved by the control of the mobile terminal. This is what we mean by terminal-controlled mobility management.

In order to build such very-loose coupling solution, three challenges needs to be addressed: network selection, interface management, and handover management (i.e. final decision and execution):

- *Network Selection:* Network selection is a key element of the handover procedure. It includes the handover decision and drives the handover execution. A variety of access network characteristics have been considered and identified as potential network selection criteria [8–16]. Characteristics include link quality, availability, throughput, network load, file transfer delay, reliability, power consumption, bandwidth, cost of service and mobile terminal's velocity. Selection schemes consider subsets of these criteria in their decision-making strategy. In network selection, satisfying all criteria may be very difficult as some criteria may conflict. One user may prefer the cheapest access network while another may prefer the access network providing the highest performance. The utility of an access network, computed from the preference rating relationship among the considered criteria and their instantaneous measured values, is clearly needed as the selection metric. Basically, the terminal will select the access network providing the highest utility.

- *Interface Management:* In heterogeneous environments, vertical handovers between different technologies can be performed using either one single reconfigurable radio interface [2, 17] or using multiple radio interfaces [3–7]. The vertical handover using multiple interfaces has already been commercially available through the Unlicensed Mobile Access (UMA) [4] solution and will be widely used in a near future. While focusing on the vertical handover using multiple radio interfaces, one of the most relevant issues is the high power consumption of these radio interfaces. For example, according to [18], if a WiFi interface is added to a handset, two third of the battery lifetime is reduced. The power consumption efficiency is thus a must since the battery capacity is limited in any portable device. So, we argue on the necessity to explicitly consider power-saving interface management as a step within the handover procedure. It should be addressed in accordance with network selection and handover execution.

- *Handover Management:* seamless inter-operator/inter-system mobility is a mandatory requirement. This should be done without modifying the handover execution mechanism to comply with a terminal-controlled mobility management approach. However, it is necessary to triggers the handover process in such a way that the network switching is transparent to the user. The handover execution delay needs to be taken into account to calculate the most appropriate instant to trigger it. This needs to be performed while coupling the handover management process with the above mentioned power-saving interfaces management as well as user-centric network selection.

Thus, in the following, first we endeavor to design a pragmatic environment-aware and performance-aware network selection at the terminal side. Then, we describe the terminal-controlled mobility framework that we propose. This one incorporates our proposed network selection and also addresses both interface and handover management while considering the limited battery capacity of portable devices.

2.3 Utility-Based Access Network Selection

2.3.1 Research Context

ABC is a fundamental challenge for next generation heterogeneous wireless packet networks. The ABC concept refers to being not only always connected but also connected through the best available device and access technology at all times. This is important not only in providing end-users with the most suitable access network but also in providing operators with the highest spectrum utilization and revenue. New intelligent selection mechanisms are therefore needed to handle the complexity of the seamless handover and to select the best available access network that satisfies QoS requirements at the lowest cost and energy use. This is the key to being "Always Best Connected".

In this vision, network selection becomes a key element of the handover procedure. Network selection includes the handover decision and drives the handover execution. In traditional homogeneous networks, network selection is based only on factors of signal quality from serving and neighboring access nodes, like Received Signal Strength (RSS) or Signal to Inference plus Noise Ratio (SINR). But in heterogeneous networks with universal access facilities, the selection process becomes more complex because different access technologies usually provide different characteristics (QoS support, billing schemes, reliability degree...). Network selection becomes a multi-criteria decision-making problem that involves a number of parameters and complex trade-offs between conflicting criteria. As explained earlier, a variety of access network characteristics have been considered and identified as potential network selection criteria in the literature [8–16]. In this contribution, we do not discuss which subset of criteria is suitable for network selection; rather, we focus on how criteria are used to make the right decision.

The complexity of access network selection is recognized as an NP-hard optimization problem [19]. There is no optimal solution since each user has his own preferences. Satisfying all criteria can prove difficult as some criteria may conflict. One user may prefer the cheapest access network while another may prefer the access network providing the highest performance. In fact, user preferences become a means to overcome the complexity of the decision-making process. They establish a rating relationship among a set of criteria and a degree of significance for each criterion. More precisely, each preference has a relative weight that users assign to each criterion depending on their requirements. Once the criteria are identified and the preferences are fixed, we need a method to compare candidate networks in order to identify the most suitable one. Usually, the decision will be based on perceived utility. This utility is also used to deduce the network operator's payoff in the radio resource allocation game. Utility is a key metric in network selection and resource allocation.

Multi-criteria selection is a classic problem in economics and in many other fields. In network selection, one popular solution is a scoring method that quantifies the score (suitability level, value, worth) of a particular network [20] [21]. In general, the score of access network *i* is computed as:

$$U_i = \sum_j w_j \times f_j(x_{ij}) \qquad (1)$$

where x_{ij} is the value of criterion *j* in access network *i*, w_j is the preference weight of criterion *j* ($\sum_j c_j = 1$), and $f_j(.)$ is a normalized function.

The normalized function is introduced to express different characteristics of different units with a comparable numerical representation. Normalized functions take various forms: a logarithm form was used in [8], an exponential form was proposed in [9] and a linear piecewise form was studied in [11, 15, 22].

In recent years, a utility-based microeconomics model has been applied to power control in wireless cellular systems [23], to radio resources management in wired and wireless networks [24–27], and to network selection strategy [10–13]. In this model, utility refers to the level of usability that a user derives from a given product, therefore reflecting customer decision experiences [28]. In access network selection and radio resource management, it measures the users' satisfaction level corresponding to a set of characteristics of an access network, including the allocated resource parameters. The normalized function $f_j(.)$ is called a single-criterion utility function and total score U_i is a multi-criteria (aggregate) utility. In fact, the score (cost) of a particular access network is itself simply a utility. In addition to the logarithm, exponential and piecewise-linear utility forms mentioned previously, a sigmoid function has also been used to model single-criterion utility [12, 23–27].

Besides a weighted sum of all criterion utilities, an *acceptance probability* was also used as a decision metric. Acceptance has been defined as an outcome variable in the psychological process that users go through in making decisions. The concept of acceptance probability in radio resource management and access network selection was introduced in [29] and reused in [12, 25, 27, 30, 31]. Acceptance probability means that a user may choose whether to accept (or select) an access network based on its intrinsic characteristics and on the amount of resources dedicated to him. Generally, taking into account the user acceptance probability, a network operator plans his resource allocation in order to maximize revenue. As a microeconomics concept, acceptance probability is based on user utility and the price that the user is willing to pay for the connectivity service. Hence, the acceptance probability is approximately equivalent to an aggregate utility.

As described here, many different single-criterion and aggregate utility functions exist. An obvious question is whether some of them are suitable for modeling user behavior in the uncertain wireless radio environment. We will demonstrate in the following that a new utility model, that we propose, not only allows users to select the best access network but also helps operators to optimize their resource allocation and enhance their revenues.

The foreseen utility-based network selection is among our major contributions. Our final objective is to propose new single-criterion and multi-criteria utility forms, to best capture user satisfaction and sensitivity in varying access network characteristics. As explained above, this is the first step towards improving the wireless resource usage from both the users' and network operators' perspectives.

2.3.2 Contributions

Basic utility theory was developed by Von Neumann and Morgenstern [32]. Subsequently, the theory has been further explained and considerably developed. In microeconomics, utility means the ability of a product or a service to satisfy a human need. An associated term is utility function: the utility derived by a consumer from a product or a service. Different consumers with different user preferences (tastes) will have different utility values for the same product. This means that individual preferences should be taken into account when evaluating the utility. The concept of utility applies to both single-criterion (attribute, characteristic) and multi-criteria consequences. In the following we summarize our analysis of single-criterion and multi-criteria utility forms as well as the identified limitations we found for access network selection based on these forms. We also describe the new utility forms that we propose to overcome the identified limitations.

2.3.2.a Towards a new single-criterion utility form for access network selection

Several works in the literature have addressed different forms for the utility function, mainly step function, piecewise, logarithmic, exponential, and sigmoidal. Yet there is no consensus about a suitable form of the utility function to model user satisfaction. Before proposing a new form of utility function, we present an overview of those that already exist.

When evaluating the utility of an access network, we distinguish between *upward* and *downward* criteria. A criterion is classified as upward if its utility is an increasing function of its value. Upward criteria include parameters such as allocated bandwidth, throughput, reliability degree, and RSS. Conversely, the utility of a downward criterion decreases in function of its value. Downward criteria include parameters such as network usage cost, energy consumption, bit error rate, transfer delay, and handover frequency.

Let us first consider a utility function $u(x_i)$ of an upward quality-related parameter x_i, $0 \leq x_i < \infty$. In general, we can consider x_i as an amount of resources that an access network can allocate to the user. Every parameter will always have an upper and a lower limit due to technological constraints and the user's requirements (i.e. $x_\alpha \leq x_i < x_\beta$.). First, the utility function should be twice differentiable on interval $[x_\alpha\ ,\ x_\beta]$. This reflects the fact that the utility level should not change drastically given a very small change in the value of a criterion (product's characteristic) and the marginal utility should be regular. Second, the utility function is a non-decreasing function of x_i. That is, the more resources allocated to the user, the higher the utility [16]. However, the improvement of the utility disappears when the allocated resources reach a certain threshold and the upper level of user satisfaction is obtained. In fact, it obeys the law of diminishing marginal utility, i.e. $\lim_{x_i \to x_\beta} u'(x_i) = 0$. The effect of diminishing marginal utility implies the concavity of $u(x_i)$ for x_i greater than a given value. Similarly, when quality-related parameter x_i goes below a certain threshold and the utility comes close to zero, user behavior is indifferent to the decrease of x_i. In other words, the decrease in utility is negligible according to the decrease of the allocated resource if the latter is still less than a certain threshold x_α. This implies the convexity of $u(x_i)$ for x_i less than a given value. Though this latter condition is reasonable in wireless networks, it has not been considered in any of the existing utility-based network management solutions.

To sum up, the properties stipulating the form of the utility function for an upward criterion are: *(i)* twice differentiability, *(ii)* non-decreasing function of x_i, *(iii)* concavity for x_i greater than a given value, and *(iv)* convexity for x_i less than a given value. Conversely, we can easily deduce the properties that should have the utility function of a downward criterion x_j, denoted as $v(x_j)$. These properties are: *(i)* twice differentiability, *(ii)* decreasing in function of x_j, *(iii)* concavity for x_j lower than a given value, and *(iv)* convexity for x_j greater than another given value. In this case, we can also easily deduce the form of $v(x_j)$ as $v(x_j) = 1 - u(x_j)$ where $u(x_j)$ is the utility function defined above for an upward criterion.

While investigating the existing forms of utility function in the literature [8, 9, 10, 12, 13, 23, 27, 31, 33, 34] (cf. Figure 2.1), we realize that only the sigmoidal (S-shaped) functions can satisfy all the necessary conditions (twice differentiability, increasing function, concavity and convexity) of a utility function in our context of network selection. These functions are:

$$u_1(x) = \frac{1}{1+e^{\zeta(x_m - x)}} \qquad (\zeta, x_m > 0) \qquad (2)$$

$$u_2(x) = \frac{(x/x_m)^{\zeta}}{1+(x/x_m)^{\zeta}} \qquad (x_m > 0, \zeta \geq 2) \qquad (3)$$

In Figure 2.1, we show only utility forms for an upward criterion. As stated earlier, if $u(x_j)$ is suitable for an upward criterion utility, $(1 - u(x_j))$ is suitable for a downward criterion utility.

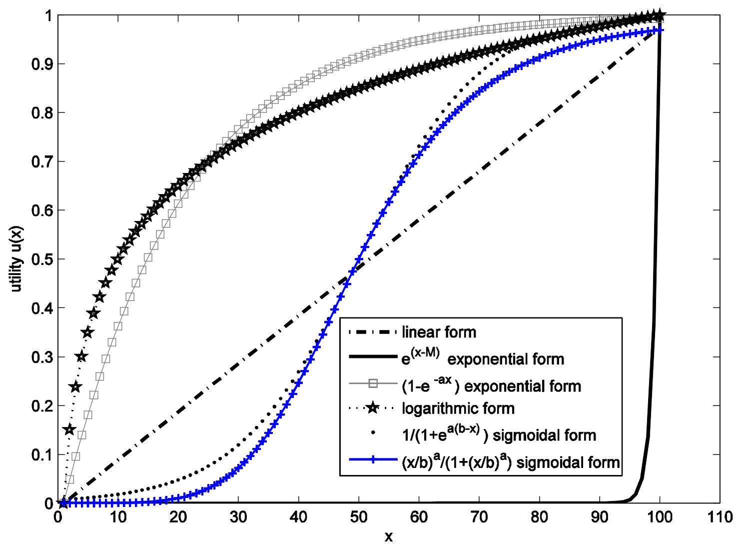


Figure 2.1. Illustration of different utility function forms

So, according to the survey we realized on existing single-criterion utility functions, the sigmoidal form is the one that is suitable for modeling utility of each network selection criterion. However, tuning the parameters (e.g. ζ and x_m) to suit the technological and user constraints (i.e. lower limit x_α and upper limit x_β for each criterion), as well as user sensitivity, is challenging. Observing the sigmoid functions $u_1(x)$ and $u_2(x)$ given in (2) and (3), we see that $u_1(x_m) = u_2(x_m) = 0.5$. The value x_m corresponds to the threshold between the satisfied and unsatisfied areas of a specific parameter. The values of x_m and ζ determine the center and the steepness of the utility curve respectively. The parameter ζ makes it possible to model user sensitivity to variation in access network characteristics. Note that x_m is user-specific and not necessarily the median of the interval $[x_\alpha , x_\beta]$.

Thus, in addition to the four properties identified above, the sigmoidal utility function needs to be redesigned to satisfy the following conditions:

$$u(x) = 0 \qquad \forall x \leq x_\alpha \qquad (4)$$

$$u(x) = 1 \qquad \forall x \geq x_\beta \qquad (5)$$

$$u(x_m) = 0.5 \qquad \text{for a given } x_m \qquad (6)$$

Furthermore, the utility function should retain a steepness parameter so as to model user sensitivity. The steepness is a free parameter that may be set differently for different criteria. In fact, we want to design a continuous function that varies from *0* to *1*: equal to *0* for x less than x_α and equal to *1* for x greater than x_β. This function should be equal to *0.5* for any given x_m. It should satisfy the necessary conditions on twice differentiability, convexity and concavity. We use a sigmoid piecewise function: one convex piece for $x_\alpha \le x \le x_m$ and one concave piece for $x_m \le x \le x_\beta$. The two pieces should be designed to be continuous and differentiable at the point x_m. To do so, the two steepness parameters can be harmonized, knowing that each piece function has a free steepness parameter. So, we choose the sigmoid form as in (3) and adapt it to take into account the identified requirements. The resulting form is the following:

$$u(x) = \begin{cases} 0 & x < x_\alpha \\ \dfrac{(\frac{x - x_\alpha}{x_m - x_\alpha})^\zeta}{1 + (\frac{x - x_\alpha}{x_m - x_\alpha})^\zeta} & x_\alpha \le x \le x_m \\ 1 - \dfrac{(\frac{x_\beta - x}{x_\beta - x_m})^\gamma}{1 + (\frac{x_\beta - x}{x_\beta - x_m})^\gamma} & x_m \le x \le x_\beta \\ 1 & x_\beta < x \end{cases} \qquad (7)$$

where $\gamma = \dfrac{\zeta(x_\beta - x_m)}{x_m - x_\alpha}$ and $\zeta \ge \max\left\{\dfrac{2(x_m - x_\alpha)}{x_\beta - x_m}, 2\right\}$. ζ and γ are the tuned steepness parameters.

One should also note that we proved mathematically that this form, which we propose, satisfies the twice differentiability, concavity, convexity and also the conditions identified in (4), (5) and (6). One should also note that (7) represents the form of the utility function for an upward criterion. In the case of a downward criterion, again the form of the utility function is $v(x) = 1 - u(x)$ where $u(x)$ follows (7).

This single criterion utility form is not intended to be unique. Indeed, in addition to this form and its monotonic transformations, there may exist other utility function forms that can satisfy all the requirements (twice differentiability, convexity, concavity). Such functions, if they exist, should be carefully designed in order to match the identified properties that are needed by access network selection.

2.3.2.b Towards a new multi-criteria utility form for access network selection

As previously mentioned, access network selection in heterogeneous networks is based on multiple criteria. Let us first investigate existing multi-criteria utility functions used in access network selection and radio resource management; namely, the additive aggregate utility and the acceptance probability.

Additive aggregate utility

A common approach to computing the aggregate multi-criteria utility of an access network is described as follows:

$$U(x) = \sum_{i=1}^{n} w_i \times u_i(x_i) \qquad where \sum_{i=1}^{n} w_i = 1 \qquad (8)$$

where x is the vector of n considered criteria and w_i are the user preferences. This approach is referred to as an additive utility. The utility-based network selection schemes, addressed in [8, 9, 16, 21] and references therein, have used this additive utility approach.

Very similar to a classic scoring method, additive utility offers an easy and accessible way to aggregate different elementary utilities. It also allows users to introduce their preferences for different criteria. But although it is widely used and has some advantages, the additive utility also has serious limitations. A fundamental issue is whether the multi-criteria utility function can be separated into independent parts where u_i, the utility of criterion i, does not depend on the value of the other criteria. In this case, the elementary utilities $u_i(x_i)$ can be simply added to produce the aggregate utility. Unfortunately, the criteria are not always independent. An example is where an access network provides good utility for all selection criteria but one. The simple numerical study in Table 2.1 illustrates such a case. Access network *B* provides good utility for all selection criteria except network load (i.e. the access network is overloaded and its utility is close to zero). Under those circumstances, connecting to this network is not useful. However, the additive multi-criteria utility leads to the selection of this access network. This limitation is characterized by the compensation possibility between different access network selection criteria, which does not capture the interdependence between the considered criteria. Generally, a suitable aggregate utility should have a small value when the utility value of a single-criterion is in the neighborhood of zero.

Table 2.1. Case Study: Additive Multi-Criteria Utility.

Utility	w_i	Network *A*	Network *B*
u (cost)	*1/3*	*0.5*	*0.8*
u (QoS)	*1/3*	*0.5*	*0.8*
u (load)	*1/3*	*0.5*	*0.05*
Total Utility		*0.5*	*0.55*

Acceptance probability

As previously mentioned, acceptance probability has been widely used in radio resource management to measure the probability that a user is satisfied with the perceived utility u given the price p. Acceptance probability is modeled by:

$$A(u, p) = 1 - \exp(-C \times u^{\mu} \times p^{\varepsilon}) \qquad (8)$$

where $\mu > 0$ and $\varepsilon > 0$ control the user sensitivity to utility and price, and C is a positive constant representing the satisfaction reference value. This model was used like an aggregate utility-based network selection in [12]. Here also the question is whether the acceptance probability adequately models aggregate utility and user acceptance probability.

Acceptance probability can overcome the limitations of the additive utility (7). However, it has three limitations when measuring the user's satisfaction. The first visible limitation is the zero price effect. An access network whose price is zero (e.g. free public WiFi) will always be selected even if it offers extremely poor connectivity. This is economically valid in general; but it should not be a factor in access network selection. The limitation is explained by the fact that acceptance probability does not take into account the future service degradation penalty. The second limitation is that utility *u* is computed for only one criterion (e.g. allocated bandwidth). In multi-criteria network selection, the utility should include all characteristics except for price *p*. The solution is to define an overall utility as either the product over a set of elementary utilities or the weighted average over a set of elementary utilities. The third limitation is that, even if an overall utility is used (e.g. the one proposed in [31]) and the zero price effect is removed, acceptance probability provides no means of introducing the preference weights w_i as an additive aggregate utility approach would do. So, these three limitations make the function proposed in (8) not suitable for network selection.

Let us also remind that initially, the concept of acceptance probability was proposed for both radio resource management and access network selection [35]. Hence, for radio resource management, taking into account the user acceptance probability a network operator plans his resource allocation in order to maximize revenue. One should also note that the three above mentioned limitations still remain in the case of radio resource management by the operator. Indeed, these limitations cause an error in the estimation of user behavior and provide no way to consider the diversity of user preferences. Thus the acceptance probability is therefore not appropriate for modeling user acceptance neither.

New multi-criteria utility function

A basic question is whether a criterion can be completely compensated by another criterion or by a set of other criteria. In other words, the nullity of a specific elementary utility does not lead to an elimination of this access network in the selection process. Generally, when the user sets a non-zero preference weight for a criterion, it means that he considers this criterion in his evaluation. If its utility is zero (i.e. its value is below x_α for an upward criterion or above x_β for a downward criterion) or close-to-zero, the corresponding access network does not satisfy the technical or user constraints. Logically, this access network should not be selected: the non-zero preference criteria are not independent of each other. As a result, the aggregate utility should reflect the interdependence among the considered criteria. Therefore, we design a new multi-criteria utility form that satisfies the following requirements: *(i)* it should increase when the elementary utility increases; *(ii)* it should be an increasing function of upward criteria and a decreasing function of downward criteria; *(iii)* it should eliminate the access networks having a zero elementary utility in the decision-making process; *(iv)* it should be able to downgrade the rank of the access network that has a close-to-zero elementary utility; *(v)* it should be equal to *1* if all elementary utilities are equal to *1* (i.e. all criteria satisfy the user's expectation). Finally, it is imperative to include the user preference weights of different criteria in the aggregate utility form to take into account the user preferences prioritization.

Given a network selection criteria vector *x* and an associated preference vector *w*, the suitable multi-criteria utility function that we proposed is formulated as:

$$U(x) = \prod_{i=1}^{n} [u_i(x_i)]^{w_i} \qquad \text{where } \sum_{i=1}^{n} w_i = 1 \qquad (10)$$

where *n* is the size of vector *x*, w_i is the preference weight for criterion *i*, and $u_i(x_i)$ is the elementary utility of criterion *i* that follows the utility form proposed in (7).

We proved mathematically that this form, that we propose, satisfies the above mentioned requirements of an aggregate utility function. We also proved that the proposed multiplicative utility form can also be used to properly model the user's acceptance probability to be used by network operator for radio resource management. Indeed, we were able to demonstrate mathematically that it avoids the three limitations identified above.

2.3.3 Summary of Results

To conclude, we can summarize our first contribution regarding the improvement of resource usage in the case of mobility over heterogeneous packet networks as follows. We proposed a complete utility theory for modeling single-criterion utility and multi-criteria utility in the context of wireless access network selection and radio resource management. The theory is based on a classic economic theory adapted to the behaviors of mobile end-users. The limitations of existing utility models were highlighted. Single-criterion and multi-criteria utility forms, with the ability to satisfy all the utility properties and to address the identified limitations, were proposed. We showed that the proposed model can also be used as a user' acceptance probability metric for the network operators' radio resource management.

After demonstrating mathematically the benefits of our single-criterion utility and multi-criteria utility, we used both numerical analysis and simulations to show the suitability and the effectiveness of the proposed utility theory in the context of access network selection and radio resource management. The numerical analysis allowed emphasizing the usefulness of both proposed single-criterion and multi-criteria utility function in comparison to the existing ones. This is realized through a simple case study where two criteria (bandwidth and price) are used highlighting the quantitative properties (i.e. interdependency vs. compensation between parameters) of our proposed utility functions. It is followed by a simulation study that targets to assess these results using more realistic case studies and five selection criteria (price, power consumption gain, maximum achievable data rate, network load and packet error rate). This later showed the user's benefit as well as the network operator benefit obtained throughout the use of the proposed functions. Thus, we were able to show that our proposed single-criterion and multi-criteria utility function are not only useful for users' access network selection but also useful for operators' resource management.

This work was one of the focuses of the Ph.D. dissertation carried out by Quoc-Thinh Nguyen-Vuong, University of Evry val d'Essonne, and defended on July 2, 2008 (see papers [P12][P41]). This work is also one of the main results of our fructuous colaboration with him developed in the frame of the ITEA2 SUMO project (June 2005 – December 2007).

2.4 TERMINAL-CONTROLLED MOBILITY MANAGEMENT FRAMEWORK.

2.4.1 Research Context

In parallel to the evolution towards converged heterogeneous networks, the telecom market is facing a migration from network centricity towards user centricity. In the current network-centric approach, operators keep tight control over users so that their network is used to its greatest potential. End-users can only influence their preferences in a limited way. In the context of deregulated telecom market, the network-controlled handover management exhibits some serious limitations in the service continuity maintenance while handing over between two different network domains operated by different operators. Among these limitations, we can find complex issues such as security context transfer and data switching management. We believe that a user-centric vision will be a mandatory evolution trend in all-IP heterogeneous wireless packet networks as it represents the most efficient way to ensure an ABC service. In this vision, users will have greater control and will be able to select the access network with which they are mostly satisfied. The users are in this strong position because their terminal can access to information on device capabilities and user preferences, and, most importantly, to knowledge of both serving and neighboring access networks. Thus, network selection is an important step of the terminal-controlled mobility management scheme. However, other challenges need to be dealt with in order to allow terminals to trigger the handover at the right instant to achieve seamless handovers and optimize terminal resources.

The mobility management solution allows mobile users to freely handover between uncoordinated available access networks. The mobility management for tight-coupling schemes is based on the existing cellular mobility solutions whereas the one for loose-coupling schemes is based on Mobile IP (MIP) [36]. Most of the current solutions require agreements between the operators, who own the different interworked access networks, to be established. The mobility management remains network-controlled. The user cannot maintain on-going sessions while handing over between two access networks belonging to two non-collaborating operators even if he has subscribed to these two networks. To overcome this issue, there is a need to design a fully terminal-controlled mobility management scheme where different access networks may be completely independent. The users are assumed to have access to these different networks involved in the handover with help of a universal Subscriber Identity Module (SIM) card or a multi-homing contract for instance. The terminal-controlled handover will thus be built on the top of this very loose-coupling interworking architecture. Seamless handover execution in such context is one of the challenges to handle.

In heterogeneous environments, vertical handovers between different technologies can be performed either using one Software-Defined Radio (SDR) interface [2, 17] or using multiple radio interfaces [3–7]. Here, we consider the case where multimode terminals are equipped with multiple wireless access interfaces. Such devices are available in the market today. The vertical handover using multiple interfaces has also already been commercially available through the UMA [4] solution and is already used in some commercial offers by network operators. However, such solutions require a careful design to minimize side effects of additional electronic devices in the terminal. One of the most relevant issues is the high power consumption of multiple radio interfaces. The power consumption efficiency is a *must* as the battery capacity is limited in portable devices. Since only the terminal can be aware of its remaining battery capacity, it is clear that only the terminal can handle the power-saving mobility management. The handover should be initiated in an adaptive manner to optimize the power consumption and to guarantee uninterrupted services. This is another challenge to deal with in the mobility management process.

Thus, proposing a new seamless mobility management framework, which does not require changes to existing network infrastructure, need to be proposed. The characteristics of this framework are: terminal-controlled, user-centricity, power-awareness, and uninterrupted-service.

The foreseen terminal-controlled mobility management framework for heterogeneous wireless packet networks is among our major contributions. Our final objective is to propose a complete mobility management solution allowing mobile users to freely handover between uncoordinated available access networks. This solution consists of a policy-based power-saving interface management scheme coupled with a user-centric network selection solution, a handover decision and triggering algorithm and a handover execution process. These are performed in an aim to improve the wireless resource usage in presence of heterogeneous wireless packet networks.

2.4.2 Contributions

The telecommunication market is influenced by three main drivers: users, network operators and service providers. In the user-centric vision, service providers are independent of the network operators. The network infrastructure can be divided into two parts: the access network and the converged core network. The former represents the current and future network operator domains while the latter is an external network independent of the operator domains (e.g. the Internet). Our mobility framework does not require changes in the access network operator and the service provider parts. Only the terminal and the converged core network are concerned.

It is widely accepted that next-generation mobile networks will be purely IP-based. Therefore service continuity can be achieved using MIP with help of the Home Agent (HA) deployed in the converged core network. The HA, which is not located in a particular access network, can thus provide the mobility management as a service for the mobile users. The mobility management is thus seen as an independent third party service. Such integration of access networks can be referred to as a very loose-coupling interworking since the corresponding network operators may be completely independent. The coupling point is located far from the radio interface. The proposed interworking architecture is depicted in Figure 2.2. Using MIP (v4 or v6) does not affect our terminal-controlled solution since the HA is located in the IP converged core network. In an IPv4-based network, we even do not need to implement

Foreign Agent (FA) entities in a particular access network since there exist solutions for MIPv4 management without it. The access networks thus remain unchanged in the very loose coupling architecture.

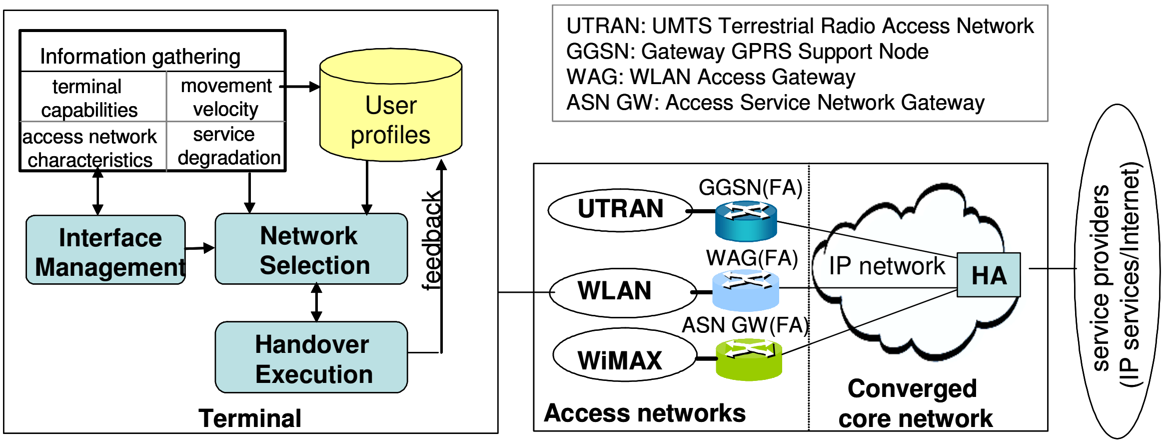


Figure 2.2. Very loose coupling interworking architecture

Figure 2.2 also depicts the different functionalities that the terminal needs to rely on. Indeed, since the terminal-controlled mobility relies mainly on the terminal rather than the network, the terminal should be smart enough to make the right decision. As stated earlier, we also consider that the intelligent terminal is equipped with multiple radio interfaces in order to access different access network technologies. To simplify the explanation hereafter, the terminal is assumed to have three radio interfaces: WLAN, WiMAX and UMTS. One should note, however, that the solutions described hereafter remain valid for any number of interfaces and underlying access technologies.

Based on the information gathering functionality (i.e. collecting available information from the surrounding access networks) and user profile (i.e. user's identities for different access networks and subscribed services, user preferences, and mobility policy), the terminal is able to perform the four tasks required by the terminal-controlled mobility management: managing the radio interfaces, selecting the most appropriate access network for handover, identifying the need of handover (handover triggering and decision), and handover execution. These four functionalities are the pillars of our proposed framework. They are described in the following.

Power-saving interface management:

Along with the emergence of multi-purpose terminals, the gap between the energy requirement and a terminal's battery capacity has progressively widened. The main source of power consumption of a portable device is related to the Radio Frequency (RF) component [39]. Therefore, we propose a policy-based interface management aiming at improving power-consumption efficiency. Precisely, we focus on the state management of the available interfaces at the terminal.

With the exception of being in the turn-off state, a wireless radio interface continually consumes energy. Obviously, the power consumed during active state is much more than in standby state. One common solution is to set the interfaces to standby state if there is no communication. However, the mobile devices spend most of the time in idle mode and consequently the standby interfaces still consume a significant portion of power. Therefore, the first attempt to achieve power consumption efficiency is to turn off the non-cellular interfaces during idle mode. Otherwise, only one interface is active at a time for communication except during vertical handover periods. This is motivated by the fact that non-cellular interfaces normally consume much more energy than cellular ones. Furthermore, cellular coverage is ubiquitous. Following this idea, the details of the policy-based interface management we propose are described in Table 2.2.

Table 2.2. Interface management proposed policy rules

Rules	Conditions	Actions
Rule 1	No on-going communication occurs.	- Turn WLAN and WiMAX interfaces off and standby UMTS interface.
Rule 2	There is an incoming communication.	- UMTS interface switches to active for handling the arriving communication. - WLAN/WiMAX interfaces switches to standby state.
Rule 3	- User initiates an application session. - Or user manually turns on the WLAN/WiMAX interface to search for available associated access networks in order to initiate communication using the appropriate one.	- Preferred/selected interface is activated. - Other interfaces switch to standby for possible handover.
Rule 4	During communication the remaining battery lifetime (*L*) using WLAN/WiMAX is less than a predefined threshold. Actually, *L* is computed as a ratio between the remaining battery power and the power consumption rate of all running applications as well as the considered radio interface.	- WLAN/WiMAX interface in standby or active state is powered off. - Network selection is triggered to select another interface and access network prior to turning off the serving interface.
Rule 5	During communication the handover frequency (*F*) using WLAN/WiMAX is greater than a predefined threshold. Indeed, *F* is defined as a ratio between the mobile's velocity and the typical cell diameter.	- *The same actions as in Rule 4.*

So, based on the gathered information as well as communication status, the interface management will decide to turn on/standby/turn off certain radio interfaces to optimize the terminal operation. It thus represents a constraint for access network selection.

Proposed Network Selection

Network selection and handover decision allow selecting the best access network and initiate the handover towards it. They are based on the gathered information, the user profiles and the interface management enforcement. To reach its objective, three main aspects of network selection are to be addressed: user preference configuration, a set of reasonable network selection parameters, and a decision-making process.

User preferences are a good means to overcome the complexity of the multi-criteria access network selection. They establish a rating relationship among criteria. Each weight is proportional to a degree of significance (e.g. high, medium, low, and ignored) of each criterion in the network selection strategy. In our configuration, we propose to formulate user preferences under the form of a matrix of three dimensions where each element, α_{klm}, is the user preference associated to criterion m, when the running application belongs to the class l and the current situation profile is k. Indeed, as there are different QoS requirements for different application classes, user preferences configuration should first take into account running applications. Applications can be simply grouped into 2 classes: real-time (voice, streaming) and non real-time services (data downloading, web browsing). An application classification into conversational, streaming, inter-active and background services [39] may also be envisioned. If multiple applications belonging to different classes are simultaneously running, the preferences according to the highest priority class are used. Secondly, with the same available access network technologies, the user decision may be different according to its situation. User situation profiles include home, in office, handset, indoors, outdoors, meeting, in car, etc. Users can change the situation profile to better suit the context and environment. We argue that the situation profile have to taken into account in the network selection and the user preferences configuration. Indeed, user preferences depend on different user situation profiles. For instance, when the user is at home, the network selection and interface management do not need to care about the power-saving criteria. The home situation profile is based on the location where the terminal's battery is charged. This is also made possible as in the future mobile devices will have capabilities to identify automatically the situation profile based on the Access Network Identifier (ANI)-based location, mobile velocity, and information collected by sensors integrated on devices [40]. So, users will thus be able to configure their preferences according to the class of application and the situation profile. A number of simplifications as well as default values and guidelines can even be imagined to help non-expert users to get their user preferences well configured.

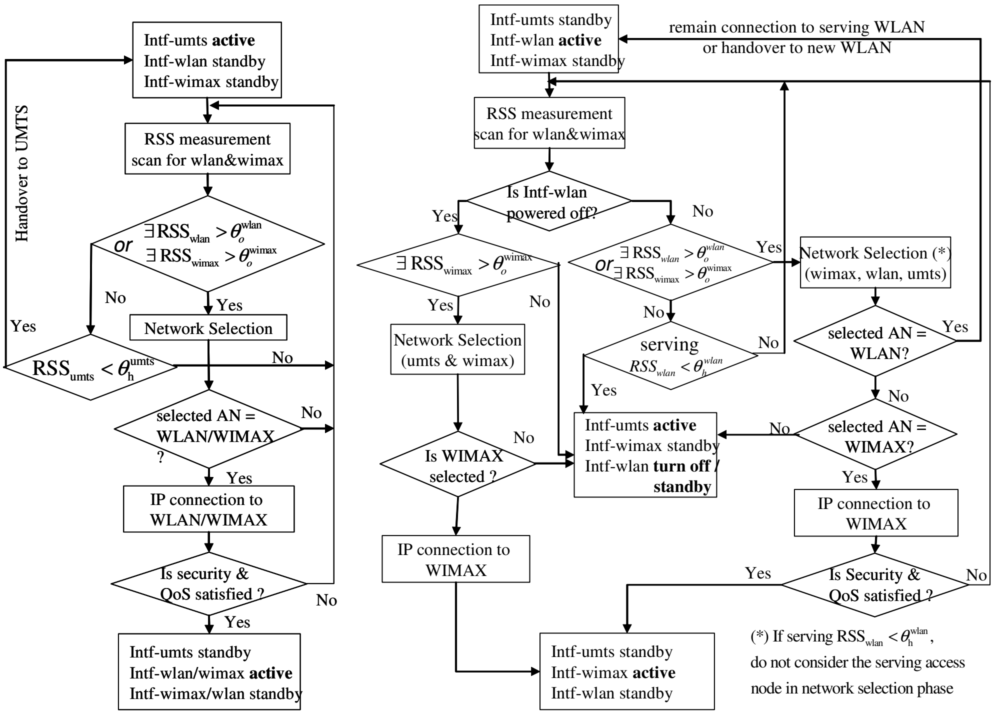
Most of current network selection solutions [8, 41] are based on dynamic QoS information such as the physical throughput or the access delay, which is challenging to obtain. Moreover, these solutions do not take into account the existence of different network administrative domains where the operators may not provide access network information to an unauthorized terminal. Furthermore, to collect such information, either the terminal must set up an IP connection with the candidate access networks without initial authentication or operators must enhance their networks to broadcast it. Both approaches are beyond the terminal capacity and necessitate the collaboration of all access networks. In our framework, we suggest using only information that the terminal can measure or estimate without need of IP connectivity as well as the information provisioned by the network. These are already numerous and significant for network selection. They include access network identity (ANI), link quality (SNR, SINR or RSS), mobile terminal's velocity, cost, energy consumption rate, battery lifetime and access network load.

Hence, after identifying the network selection criteria as well as addressing the user preferences configuration and the interface management, the network selection itself can be triggered and decision-making process launched. This latter is obviously based on our proposed single-criterion and multi-criteria utility functions proposed for network selection (cf. Section 2.3) and with respect to the given user- and technology-related parameters x_m, x_α, and x_β. Network selection is applied to a set of pre-selected access networks (i.e. black-listed access networks or access networks not supporting the required services will be eliminated from the candidate list).

Handover triggering and decision

The proposed handover decision algorithm is based on the previously presented network selection and interface management mechanisms. Hence, a vertical handover is initiated if *(1)* the user equipment (UE) prefers to handover to a higher score access network, or *(2)* the serving interface must be powered off due to the interface management, or *(3)* the current radio link is about to drop. For this latter case, if the signal strength of the serving access node drops below a handover threshold θ_h, the UE initiates the network selection algorithm to select a more suitable access network for handover. θ_h should be adaptive in order to ensure seamless handovers. Without loss of generality, we proposed an approach to compute this threshold when handovers occur between UMTS, WLAN, and WiMAX systems.

Also, we proposed detailed algorithms for vertical handover from a UMTS to a WLAN/WiMAX as well as from WLAN to UMTS/WiMAX. These are depicted respectively in Figure 2.3(a) and Figure 2.3(b). The vertical handover decision from WiMAX to UMTS/WLAN is similar to that from WLAN to UMTS/WiMAX. If we replace the term WLAN by WiMAX and vice versa in Figure 2.3(b), we obtain the corresponding algorithm.



(a) UMTS → WLAN/WiMAX handover (b) WLAN → WiMAX/UMTS handover

Figure 2.3. Handover algorithms for active UMTS and WLAN interfaces

Handover execution

The final step of the process is the enforcement of the handover decision. For that, we propose to use the MIP mechanism and multiple Care-of-Address (CoA) registration solutions [42, 43]. Using MIP does not affect our terminal-controlled solution since the HA is located in the converged core network. The converged core network is realistically assumed to be IP-based. Conceptually, the MIPv4/MIPv6 standard does not allow a terminal to register multiple CoAs bound to a single home address. According to [42, 43], a new identification called the Binding Unique Identification (BID) is used to associate with each binding cache entry to accommodate multiple binding registrations. This is useful while using multi-radio interface terminals and handing-over heterogeneous wireless packet networks. Thus, in our case, the handover execution process we propose is performed as depicted in Figure 2.4.

When the UE is powered on, it searches for the available Radio Access Networks (RAN) and retrieves an associated CoA for each radio interface. The UE registers each of its CoAs with the HA. Conventionally, the CoA associated to the UMTS radio interface is set to primary CoA. During the idle communication mode, only the primary UMTS binding cache is updated when the UE changes its network domain. Therefore, the HA will use the primary UMTS CoA to page the UE in case of an incoming call. When the UE is communicating via UMTS and the target access network, according to the network selection, is WLAN/WiMAX then the vertical handover from UMTS to WLAN/WiMAX is triggered. The exchanged message sequence is depicted in Figure 2.4(a). First, the UE sets up the connection and authenticates with the target network. The foreign network authenticates the UE by exchanging the information with the Home Authentication, Authorization and Accounting (H-AAA) server. Afterwards, the UE acquires the corresponding CoA and sends a MIP registration with a primary WLAN/WiMAX CoA option to the HA and the Corresponding Node (CN). Once the MIP registration reaches the HA and the CN, these latter set the new WLAN/WiMAX CoA as primary CoA, return the MIP registration reply and then use the new CoA path to forward/send data to the UE. After the handover is complete, the UMTS interface remains active for a time period to receive the in-flight packets on the old path and finally switches to a standby state.

The handover procedure from WLAN/WiMAX to UMTS is illustrated in Figure 4.4(b). When the UE is communicating via WLAN/WiMAX, the UMTS interface is in standby state and the UMTS CoA in the HA and the CN are still maintained. If the association between the UE and the GGSN, known as Packet Data Protocol (PDP) context, is not timer-expired, the UE only needs to send the MIP registration to the HA and its CN to perform the handover execution. If the PDP context is already released, the UE needs to reactivate the PDP context to establish the authentication and billing information with the UMTS network. The in-flight packets destined for the UE via the old path are received as long as the UE is still covered by the old WLAN/WiMAX access network.

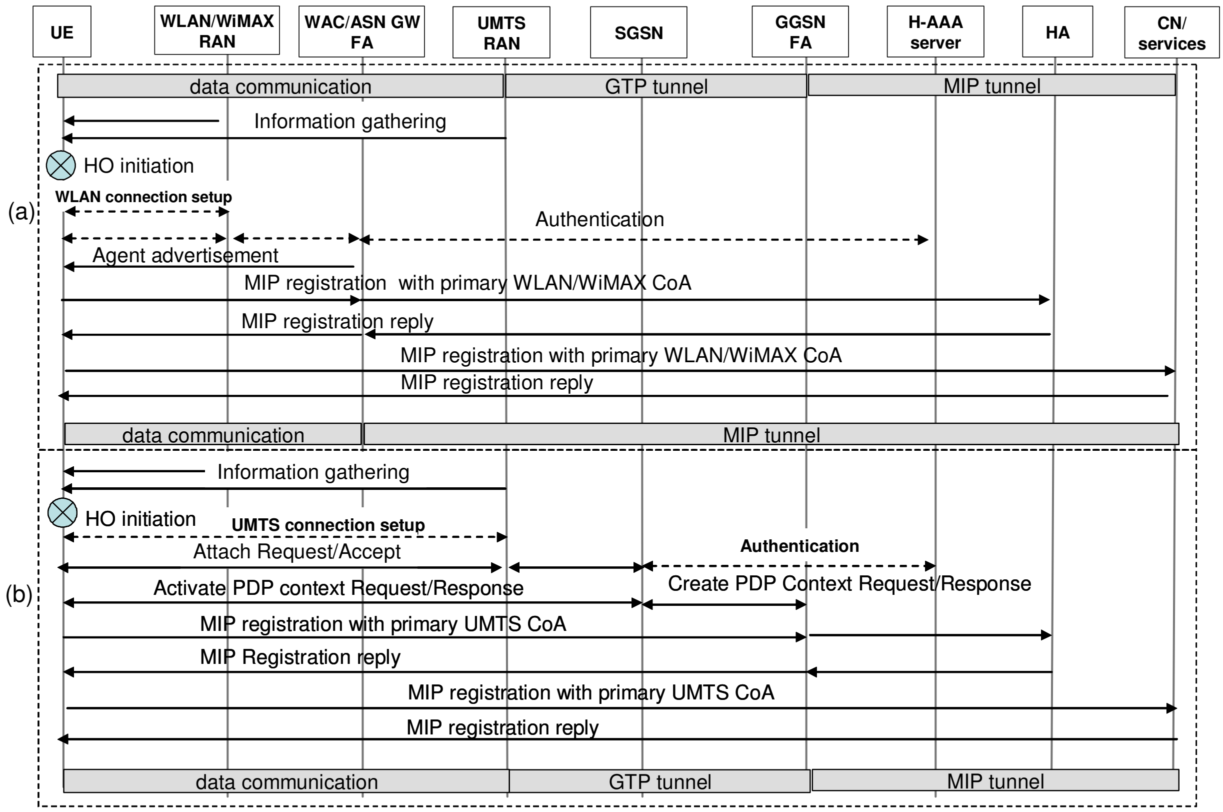


Figure 2.4. Handover procedure between UMTS and WLAN/WiMAX

2.4.3 Summary of Results

To conclude, we can summarize our second contribution regarding the improvement of resource usage in case of mobility over heterogeneous packet networks as follows. We proposed a fully terminal-controlled mobility solution across heterogeneous networks without architectural changes in network operators' infrastructure. We described a user-centric approach for vertical handover management. Our solution includes user-centric network selection, power-saving interface management and adaptive handover initiation algorithm at the terminal to support seamless terminal-initiated and terminal-controlled vertical handover. In the network selection scheme, we proposed to employ our novel multiplicative aggregate utility approach (cf. Section 2.3) to best evaluate the candidate access networks. The network selection triggering conditions have been specified to guarantee seamless handovers. The proposed access network selection is situation profile-based and application-aware (context-aware) to suit different communication contexts. It enables terminals to select the most suitable access network according to various access network characteristics. Furthermore, we proposed the power-saving interface management policy to optimize the use of terminal's battery for devices equipped with multiple radio interfaces. It serves as a complement to the network selection scheme by turning off inappropriate interfaces according to the terminal velocity and the remaining battery lifetime. The interface management solution optimizes the power consumption of today's battery-limited portable devices to extend the user communication time. Last but not least, we also tackled adaptive handover initiation to assist the service continuity. The solution is realistic and not much complex to implement in current mobile devices and networks. To conclude, we can say that this work is a step towards developing future terminals with enhanced capabilities to support users during their mobility.

In order to highlight the benefits of the proposed network selection and power-saving interface management solutions, we performed a set of simulations of all the component of our framework. We considered simulation scenarios in which a user holds a terminal equipped with 3 radio interfaces: UMTS/HSDPA, WiFi and WiMAX. At each instant the user is able to choose from three available access networks: UMTS, WiFi and WiMAX and perform terminal-controlled vertical handover to the selected access network. The selection is based on the five selection criteria mentioned previously: cost, power consumption gain, maximum achievable data rate, network load and signal quality. The values of these criteria were realistically generated over time. We also used realistic values for the energy consumption rate values of UMTS, WiMAX and WiFi interfaces. Using our simulations, we had been able to show the advantages of the application-aware and situation-aware user preferences configuration. We also showed the achieved power consumption efficiency thanks to the power-saving interface management. Again, seamless handover over a very loose coupling interworking was also proved to be possible thanks to the help of the adaptive handover initiation threshold. Finally, we compared our proposed network selection scheme with the more recent and performing one from the literature [15]. This performance comparison confirmed the efficiency of our network selection solution.

This work was one of the focuses of the Ph.D. dissertation carried out by Quoc-Thinh Nguyen-Vuong, University of Evry val d'Essonne, and defended on July 2, 2008 (see papers [P12][P16][P67]). This work is also one of the main results of our fructuous collaboration with him developed in the frame of the ITEA2 SUMO project (June 2005 – December 2007).

2.5 CONCLUSION

The work presented here was realized in the frame of collaboration with Quoc-Thinh Nguyen-Vuong during his Ph.D. work at University of Evry val d'Essonne (September 2005 – July 2008). It was undertaken in the frame of the ITEA2 SUMO project (June 2005 – December 2007) to which we had an active participation.

The obtained results are numerous. We started from the fact that the end-users are not anymore passive - they are starting to design their own services and sharing them in communities. They are increasingly demanding higher quality and more personalized services. We thus found out that the best way to offer customized services in the context of mobility is to allow users to influence and even to control the access network choice, the handover decision and the handover preparation. So, we explored how the mobility management can be achieved and improved under a user-centric terminal-controlled approach. As a preliminary step towards that, we first discussed and showed how the utility theory could be adapted for the purposes of access network selection. Thus, we proposed a novel utility-based network selection that overcomes the limitation of the existing utility forms in this context. We showed that our proposed single-criterion and multi-criteria utility functions are useful for both users' access network selection and operators' resource management. Having showed that, we built a complete terminal-controlled handover management framework around this user-centric network selection (i.e. based on our novel utility functions), a policy-based power-saving interface management scheme, a handover decision and triggering algorithm, as well as a handover execution process. We demonstrated the effectiveness of the different facets of our contributions using an exhaustive performance study. To do so, we used mathematical proofs, numerical analysis and simulations. Overall, these showed that using our solutions, we are able to improve the wireless resource usage in a heterogeneous wireless packet networks context from the user's and eventually the operator's perspective.

Chapter 3. Improving TCP Behavior in Wireless Packet Networks

Transport Control Protocol (TCP) is the most commonly-used reliable transport protocol in the Internet. Today, it is supported by almost all Internet applications. However, TCP does not always perform optimally according to the network environment in which it is used.

In the following, we will first remind the context of our research as well as the challenges that faces TCP in wireless infrastructure and infrastructure-less networks. This will be followed by the description of our three contributions and the research results we obtained in this domain. These contributions concern TCP Behavior improvement in both Wireless Local Area Networks (WLANs) and Mobile Ad hoc Networks (MANETs). This chapter ends with a brief conclusion summarizing our research experience in this domain.

3.1 Research Context

TCP was originally developed and optimized for wired infrastructure networks. Its congestion control algorithm and the data flow mechanism it employs make it performing well in such networks. Through controlling the number of data packets transmitted over the connection and taking into consideration the advertised data to be received at the receiver side, TCP has the ability to fairly manage the utilization of the available bandwidth. The success of TCP within wired networks, in addition to the wide variety of applications that are compatible with it, made TCP the premier transport protocol to be used within almost all the deployed data networks. However, taking into consideration the differences between wired and wireless environments, we expect that TCP performance would differ when applied within each of them. Many researches [44–46] were done recently to help understand the performance of TCP within wireless (infrastructure and infrastructure-less) networks. These researches reveal the problems that can be found within wireless networks and do not exist within wired ones, such as wireless channel errors and link failure problems. These problems represent new challenges to TCP since it was not originally developed to deal with the underlying issues. In fact, TCP was developed to solve the problem of data packet losses due to congestion within the network, and originally is a congestion-control-oriented protocol. This was commonsense since the main cause of data packet losses within wired networks is congestion.

Within wireless (infrastructure and infrastructure-less) networks, there are many causes of data packet losses other than congestions. This makes them a highly challenging environment for TCP as it was not designed to deal with the new data packet loss causes, which are due to the wireless environment. In order to overcome these issues, one solution was to integrate *Loss Differentiation Algorithms* (LDA) [47, 48] to TCP. These algorithms allows differentiating loss types either by using successive measurements of packet inter-arrival times or by using successive measurements of round-trip times (*RTT*), and then makes TCP reacts adequately.

The target of our research, presented in this chapter, is to study the suitability and effectiveness of existing Loss Differentiation and Recovery Algorithms in Wireless Local Area Networks as well as Mobile Ad hoc Networks. Then, according to the identified drawbacks, proposing a new set of Loss Differentiation and Recovery Algorithms that allows improving the resource usage while TCP is used within such wireless packet networks.

3.2 Main Challenges

Thanks to the emergence of the 802.11 standards [49–53], a new set of wireless packet network architectures appeared during the last decade, namely Wireless Local Area Networks (WLAN) and

Mobile Ad hoc Networks (MANETs). Despite the attractive applications that they allow, the features of these new wireless packet network architectures introduce several challenges that the research community must deal with. One of these challenges focuses on the performance of TCP when used in such environments. Indeed, even with its wide deployment today, TCP still generates an important research activity aiming at analyzing its behavior and improving its resource usage especially in the wireless realm.

When looking at the 802.11 wireless network environment, performance problems are still abundant and well known. Among the principal ones, we find:

- The low noise immunity or the sensitivity to the interferences (the 2.4 GHz bandwidth is shared with other 802.11 channels or other communications like Bluetooth);
- The throughput reduction (Auto Rate Fallback) or even the signal losses due to the distance or to fixed and mobile obstacles (walls, furniture, pedestrians, cars …);
- The average throughput decrease for all the stations if only one of them has a degraded nominal throughput [54];
- An important overhead (LLC, MAC and PHY headers; control traffic induced by the CSMA/CA access method …).

These performance problems do not only affect the throughput at the link layer but can also influence the layers above (Network, Transport and Application layers) resulting in an underutilization of the available bandwidth. This is especially true for elastic flows as TCP behavior can be highly influenced. This negative influence remains true regardless of the used wireless network architecture, WLAN or MANET.

In the wireless infrastructure-less networks, or MANETs, the problems are even more challenging. Indeed, the 802.11 performance problems cited above are decoupled by the multi-hop wireless context. Moreover, the features of MANETs introduce several new challenges for TCP that must be studied carefully. Among them, we can quote:

- *Nodes Mobility*: Mobility may induce link breakage and route failure between two neighboring nodes, as one mobile node moves out of the other's transmission range. Link breakage in turn causes burst packet losses. If not dealt with appropriately, this burst packet losses combined to the fact that route recovery may be long, may lead to a harmful reaction of TCP.
- *Energy Efficiency:* As power is limited at mobile nodes, any successful scheme must be designed to be energy efficient. TCP is not an exception to this rule. Energy efficiency is critical for prolonging network lifetime and thus avoiding link breakage that has a negative influence on TCP performance.

Handling efficiently these new challenges (channel errors, signal losses, nodes mobility, and energy efficiency) is the key in order to improve TCP resource usage in such wireless packet networks: namely WLANs and MANETs. In that context, many issues still need to be dealt with. Proposing innovative solutions to cope with some of the underlying challenges constitutes an important part of our work. Some of our selected research results are presented in the following sections.

3.3 A CROSS-LAYER APPROACH TO IMPROVE TCP BEHAVIOR IN 802.11 WLANS

3.3.1 Research Context

Wireless local area networks, making use of 802.11 (a, b, g, or n) [49–53] equipments witness a ubiquitous deployment, particularly in SOHO (*Small Office Home Office*) environments or in public points of distribution (*Hot Spot*), offering an Internet access to wandering users. Due to various and unpredictable reasons, the performances of TCP in 802.11 networks are not always as sufficient as the current applications require.

Several solutions have been proposed in the literature in order to enhance TCP performances in wireless networks [47, 48, 55-61]. These solutions act either at the Data Link Layer trying to improve the MAC or LLC operations, or the TCP layer by introducing loss recovery schemes at the sender or access point sides. Among these solutions, we have:

- Two kind of solutions are proposed at the Data Link layer to improve the TCP behavior in wireless networks: The one improving the LLC sub-layer operation [55] by queuing frames during short signal losses and the one using the *Automatic Repeat reQuest* (ARQ) Protocol along with *Forward Error Correction* [56] to resolve segment losses at the link layer. Both kind of solutions show performance improvement while handling the channel error situations they had been designed for. However, they both violate the legacy 802.11 standard and require important changes in 802.11 equipments.

- Many proposals have been suggested to adapt TCP operations to the wireless context. These proposals consist in either optimizing the behavior of existing TCP versions and creating new variants, like TCP *WestWood* [57] or TCP *Jersey* [58], or in the use of *Loss Differentiation Algorithms* (LDA) [47, 48]. They both try to tune the TCP congestion control parameters according to the estimated available bandwidth (by the proposed TCP variant) or the identified loss-type (by the LDA scheme). However, these solutions are built upon a set of assumptions that are not always valid in practice. Furthermore, all these solutions are proposed in the general context of wireless networks and do not take into account the MAC layer specificities and the interactions with it. More specifically, these mechanisms are not optimized for 802.11 wireless networks where a first level of loss recovery is already carried out at the MAC layer.

For the majority of these proposals, performance improvements are based on the distinction between signal losses and channel errors on the one hand and TCP congestions on the other hand. From what we noticed, these solutions are either not optimized for a usage in 802.11 WLANs or not conforming to the 802.11 standard. One of the characteristics of these solutions is that they deal only with one layer of the protocol stack and the interactions with other layers have often been neglected. From our point of view, a solution utilizing, more finely and in a coordinated way, the functionalities of both concerned layers (i.e. Transport and Data Link layers), would give better performances without having to modify the 802.11 standard. We strongly believe that such cross-layer solution is one of the trails to follow in order to improve WLAN resource usage by elastic flows.

The foreseen WLAN resource usage improvement when TCP is used is among our major contributions. Our final objective is to propose a complete cross layer solution in order to realize this performance improvement.

3.3.2 Contributions

A retransmission at the MAC layer indicates that a data or a control frame has been corrupted or not been received within a relatively low recovery time compared to TCP timeout (typically 1ms for MAC and 500ms for TCP). At the MAC layer, a retransmission occurs when the 802.11 acknowledgment is not received by the transmitter within the specified delay. For each retransmission, a counter is incremented until a certain threshold, named *Retry Limit*, is reached (in fact two counters and two thresholds are used according to the size of the frames). Beyond this statically configured threshold, the frame is dropped. For a connection using TCP, coherence between layers should lead to a MAC layer recovery before TCP is alerted when the segment loss is due to bad channel conditions. Otherwise, TCP will consider this segment loss as a congestion, which induces a reduction of the Congestion Window (*CWND*) and a decrease in the global throughput.

In order to better understand the issues leading to TCP performance degradation in WLAN, we first carried out a set of experimental measurements to study the two distinct MAC and TCP loss-recovery levels when a failure in the 802.11 channel inducing the loss of TCP segments occurs. To facilitate the analysis of the MAC and TCP layers behavior, two types of channel degradations commonly encountered in such environment are introduced in our experimental setting: *(i)* signal failures (caused in practice by the displacement of the pedestrian user at the cover limit of its access point or by other pedestrians/obstacles moving between the access point and the station), or *(ii)* interferences (generated by other transmissions using the same frequency band). The obtained experimental results show that in both scenarios a high number of segment losses is experienced leading to an important overhead (between 20 to 30%) caused by the MAC retransmissions for the critical periods (i.e. channel degradation periods). More importantly, we notice that the triggered MAC retransmissions are not effective. In the first degradation case, TCP is alerted before the restoration of the channel, which leads to reducing the TCP's congestion window uselessly. In the second degradation case, the MAC layer loss-recovery is insufficient to prevent some segments from being affected by the interferences, causing TCP retransmissions.

From our experimental measurements, we noticed that both situations led clearly to a fall in TCP performances. Overall, the limitations illustrated by our study show the lack of interactions between the MAC and TCP layers leading to inefficient loss recovery. Hence, we argue on the necessity of a cross-layer design to improve TCP performances in 802.11 networks. Therefore, in order to improve the resource usage by TCP in WLANs, the idea of such cross-layer solution is to dynamically adapt, according to the characteristics of the experienced channel degradation, the parameters of both MAC and TCP layers.

In our way towards a complete cross-layer solution allowing a dynamic adaptation of MAC and TCP parameters according to the characteristics of signal losses, we first targeted to show whether a higher *Retry Limit* threshold may solve the above mentioned problems or not. This idea proved to be effective in some cases. Indeed, the realized simulation study shows that in case of signal losses due to distance or obstacles, the increase of the *Retry Limit* value improves the performances substantially without causing a significant overhead. However, this increase is not effective when a TCP flow is disturbed by interferences. Also, in the case of TCP congestions, i.e. segment losses due to a traffic increase on the shared wireless link, the *Retry Limit* increase is not justified and even not suitable as there is a clear risk to extend the effect of the wireless channel saturation. Thus, we can distinguish three different events for segment losses in WLANs: congestions, signal failures due to distance or obstacles, and interferences due to other communications in the same frequency band. Each of these events should then be handled accordingly.

Hence, and as a second step of our work, we proposed two complementary Loss Differentiation Algorithms in order to classify the different segment-loss causes and to trigger the adequate loss-recovery. According to the characteristics of the various causes of segment loss (displacement of the wireless station or obstacles moving between the station and its AP; interferences caused by other transmissions in the same frequency band; or congestion due to increased traffic conditions), these two complementary Loss Differentiation Algorithms act respectively at the MAC and TCP layers. They both use MAC layer parameters to identify the loss cause and then adapt the MAC and TCP loss-recovery algorithms accordingly.

The two loss differentiation (and recovery) algorithms we proposed as well as their interactions with the legacy TCP implementation (i.e. TCP New Reno) are summarized in the following.

3.3.2.a MAC-Layer LDA

The first proposed LDA scheme, implemented at the MAC layer, is used to identify losses due to wireless link failures when the distance between the wireless station and the Access Point (AP) increases or when obstacles move temporarily between the station and its AP. When these link failures occur, a dynamic adaptation of the MAC *Retry Limit* parameter is realized. Indeed, as stated above, an increase in the *Retry Limit* value is appropriate to recover packet losses due to signal failures. Note that these link failures are very frequent in 802.11 networks and can be easily forecasted using the *Signal to Noise Ratio*

(*SNR*) parameter available at the MAC layer. This parameter is indeed more suitable to use compared to Transport-layer information, such as successive values of *RTT* or packet inter-arrival times, which are simply not measurable when the channel is unavailable.

In the MAC 802.11 frames, the "signal" field specifies the current throughput used to transmit the following data (*Data Rate*). The value stored within the "signal" field depends on the measured power received by the station before its transmission and is thus proportional to the SNR. In addition, this throughput indication is related to the *Auto Rate Fallback* (ARF) procedure implemented by all the 802.11b, g or n card manufacturers. Let us recall that this procedure automatically reduces the throughput when the distance increases causing a drop in the SNR. The proposed loss differentiation algorithm is based on the simple fact that if the SNR (or *Data Rate*) is maximal, then the probability that the segment loss is due to a signal failure caused by the distance and not to TCP congestion is very weak. Then, following this statement, a dynamic *Retry Limit* adaptation, according to the *Data Rate* given at the 802.11 MAC layer, is performed. The increase of the *Retry Limit* threshold is linear and progressive (a default value of 6 is successively added) to avoid congesting the channel unnecessarily when this latter is used by other transmissions. Algorithm 3.1 depicts the proposed LDA for the case of the 802.11g MAC layer.

Algorithm 3.1. MAC-layer LDA.

```
if (DataRate ≥ 12Mbps)                              // station is closed to AP
then
    RetryLimit = 6                                  // default value
else if (DataRate ≥ 6Mbps)                          // possible signal failure
    then
    RetryLimit = 12                                 // begin to enlarge transmission window
  else if (DataRate ≤ 6Mbps)                        // probability of failure is max.
      RetryLimit = 18                               // continue to enlarge window
      if (new segment) && (last segment dropped)    // new TCP segment and last MAC retry failed
      then
         RetryLimit = RetryLimit + 6*NSSL           //enlarge again window
      end if                                        // NSSL: Number of successive segment losses
  end if
end if
```

Note that the *Retry Limit* increase is bounded by three events:

- The arrival of the MAC acknowledgment for a retransmitted segment;
- The TCP transmission window is emptied;
- The RTO is reached without the channel being restored.

When one of these events occurs, the *Retry Limit* is reset to its initial value. This help to avoid occupying the channel unnecessarily. In addition, one should note that, during the *RL* increase phase, the fairness with other MAC traffics on the same channel is preserved thanks to the 802.11 backoff algorithm which increases exponentially the time between two retransmissions according to the index of this latter. Also, note that the other MAC traffics are not directly concerned by the *Retry Limit* increase because the *Data Rate* field carried by the frames is specific to each wireless station and its transmission conditions (distance, obstacles …).

3.3.2.b Cross-Layer LDA

The second LDA scheme acts at the TCP layer and is based on the use of a MAC layer parameter: the *AckFailureCount*. This parameter represents the number of MAC-layer retransmissions for each frame. The objective of this second LDA scheme is complementary to the MAC-level one. It aims to distinguish congestions from transient segment losses due to interferences caused by other close transmissions using the same frequency band. We argue that, using a cross-layer approach (i.e. taking into account the 802.11 specificities) as an alternative to conventional LDA schemes (i.e. using TCP-layer information only) may lead to a more accurate loss differentiation.

The idea of our alternative algorithm is to count the number of MAC retransmissions for each of the *n* segments composing the current TCP window when the TCP layer is alerted by the reception of three *Duplicate Acknowledgements*. As described in Algorithm 3.2, if for one of these segments at least, the number of MAC retransmissions (*RetryCount*) is equal to the threshold (*Retry Limit*), we consider that the loss is due to interferences (*LDA_Estimator = 1*) and not to TCP congestion (*LDA_Estimator = 0*). Indeed, in the case of congestion, the surplus of segments is eliminated from the queue of the concerned node and MAC retransmissions are supposedly not used. Inversely, in case of persistent interferences, the segment is dropped by the MAC layer after reaching the *Retry Limit* threshold. This algorithm assumes that for all the unacknowledged TCP segments, the value of *RetryCount* is stored. The *ACKFailureCount* counter available in the 802.11 *Management Information Base* (MIB) [49] gives the number of times that an expected ACK is not received and consequently the value of *RetryCount*.

Algorithm 3.2. Cross-Layer LDA.

```
if (3 dup. ack.) then                          // loss indication in TCP NewReno algo.
    LDA_Estimator = 0                          // initial value for congestion
    for (i = 0 ; i ≤ n ; i ++)                 // for all the not acknowledged segments
        if (RetryCount = RetryLimit) then      // segment is dropped, probably a short loss
            LDA_Estimator = 1                  // set value for interferences…
        end if
    end for
end if
```

Note that while the TCP sender is not a wireless host and that the TCP flow is forwarded to the wireless receiver through an AP, an additional stage is necessary. The *LDA_Estimator* is first set at the AP's MAC layer. Then, this latter informs the TCP sender by setting consequently the *ELN* (*Explicit Loss Notification*) bit of the TCP header in the ACK segments (i.e. *ELN*=*LDA_Estimator*=1 in case of interferences). The loss differentiation is finally made at the TCP sender when receiving three duplicated ACKs. One should note that the modification of the AP's firmware in this case is minimal.

Then, when the source detects a segment loss by the reception of three duplicate acknowledgements[2], the cross-layer LDA is questioned to know the cause of the segment loss:

- If the loss is classified as due to congestion (*LDA_Estimator = 0*), a normal TCP New Reno reaction is triggered and *cwnd* is halved.

[2] We assume here that the case where a segment loss is followed by a TCP retransmission after timeout is treated by the MAC-layer LDA.

- If the loss is classified as due to interferences (*LDA_Estimator = 1*), TCP parameters are kept unchanged. This allows the source to achieve higher transmission rates in the event of short successive signal losses, compared to the blind reduction of the throughput performed by the legacy operations of TCP.

The proposed TCP New Reno adaptation is summarized in Algorithm 3.3.

Algorithm 3.3. TCP NewReno adaptation.

```
if (3 dup. ack.) then
    if (LDA_Estimator = 1) then            // loss classified as due to interference
        no changes in ssthresh and cwnd    // cwnd not reduced
        no change in RTO                   // RTO not incremented
    else                                   // loss classified as due to congestion
        ssthresh = cwnd/2                  // usual New Reno operations
        cwnd = ssthresh
        RTO = 2xRTO
    end if
end if
```

3.3.2.c *Complete Cross-Layer Solution*

Figure 3.1 provides an outline of all segment-loss reasons on an 802.11 wireless link and summarizes the interactions among the two proposed LDA schemes.

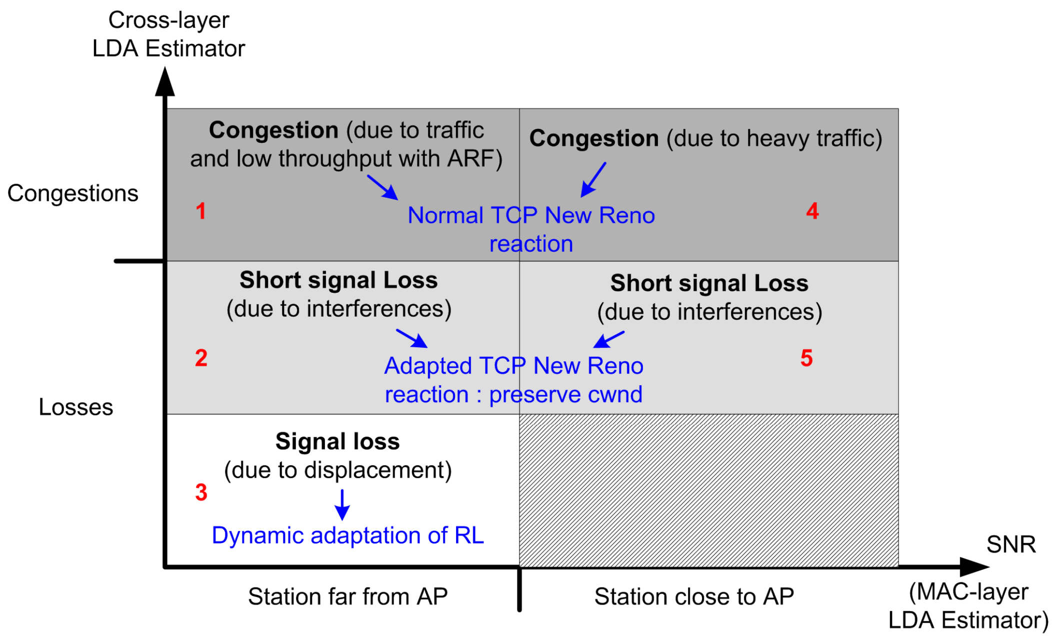


Figure 3.1. Loss-recovery for different loss cases

- The differentiation of cases 4 and 5 is carried out by the cross-layer LDA. In these cases, the TCP New Reno adaptation is in charge of improving the TCP behavior in the event of interferences.

- The MAC-layer LDA based on *SNR* is used to differentiate cases 1, 2 and 3 from cases 4 and 5. For case 3, the *Retry Limit* adaptation is used to improve the performance significantly. Note however that when the station is far away from its AP, the loss differentiation algorithms do not make possible the distinction of case 3 from cases 1 and 2. This is however not constraining. Indeed, we showed that the evolution of TCP throughput, when the *Signal Rate* is reduced to its minimum (i.e. 6 Mbps for 802.11g, for instance), is slightly influenced by the intervention of one LDA or the other. Finally, one should note that case 1 is distinguished from case 2 by the cross-layer LDA.

3.3.3 Summary of Results

To conclude, we can summarize our first contribution regarding the improvement of TCP behavior in wireless packet networks as follows. In this contribution, we first highlighted throughout measurements the lack of interactions between the MAC and TCP loss-recovery levels in the event of 802.11 signal failures. Then, we showed that acting at the MAC layer by using a higher *Retry Limit* threshold may be very useful in case of wireless link failures. Finally and most importantly, based on the above mentioned results, we proposed two Loss Differentiation Algorithms acting respectively at the MAC and TCP layers and both using MAC layer parameters. Depending on the distinguished segment loss cause (i.e. congestions, signal failures due to distance or obstacles, or interferences due to other communications in the same frequency band), adaptations of the MAC and TCP loss-recovery algorithms had been proposed.

In order to show the performance improvement of the two proposed LDA schemes as well as the complete cross-layer solution built upon them, we realized a set of simulations using NS-2 [63]. More specifically, our simulation study target was to demonstrate the effectiveness of the two LDA separately as well as their complementarities. Unsurprisingly, the first LDA scheme acting at the MAC layer showed interesting results by overcoming the losses due to distance or obstacles. The *SNR* parameter proved to be a good estimator for wireless link failures. On the other hand, thanks to its use of MAC layer parameters, the second LDA scheme acting at the TCP layer showed to improve the efficiency of the loss differentiation by 20 to 40% compared to other LDA schemes from the literature (namely the *Vegas Predictor* [64], the *Flip Flop* filter [65] and *Spike* [47]). Hence, 100% of the losses are correctly classified in the case of congestions. In this case, our simulations confirmed our assumption showing that there are almost no MAC retransmissions and that the segments are dropped before being treated by the MAC layer. Regarding interferences, some segment-losses are badly classified but the accuracy of our scheme remains higher than 90% for all studied interference situations. The complete cross-layer solution, making use of both LDA schemes jointly with an adapted version of TCP New Reno, had also been evaluated confirming the significant performance improvement as well as the effective interactions between the MAC and TCP loss-recovery levels.

This work was one of the focuses of the Ph.D. dissertation carried out by Stéphane Lohier, University Pierre and Marie Currie – Paris 6, and defended on June 19, 2006 (see papers [P14][P15][P46][P47][P49][P54][P56]).

3.4 TCP COMPUTATIONAL AND COMMUNICATION ENERGY COST IN MANETS

3.4.1 Research Context

In order to identify TCP's performance limitations and thus be able to address them, it is important to study its behavior, categorize its performance metrics and quantify them in each of the different environments where TCP can be used. These environments include also MANETs. Indeed, as stated earlier in this chapter, TCP performance is expected to be considerably affected when implemented within wireless multi-hop ad-hoc networks due to their specific characteristics (e.g. mobility and battery-dependence).

Being battery operated, ad-hoc network nodes have to be energy conservative. In the past, many research projects have been studying TCP performance in terms of energy consumption and average throughput within wireless mobile networks [45, 66, 67]. Energy consumption studies were however only partial. In [45] for instance, the authors studied the energy consumption and throughput of three

TCP variants (TCP Reno, TCP New Reno, and TCP SACK) through test-bed experiments. They evaluated TCP total energy consumption and subtracted the idle energy consumption of the TCP nodes. However, the authors applied random Round Trip Time (RTT) delays and random packet losses in their experiments. They also did not separate the computational and the communication cost parts. Studying TCP using random values does not really reflect the behavior of neither wireless ad hoc network environment nor nodes' mobility. The authors in [46] evaluated the energy consumption of TCP over wireless links with and without the use of the Selective ACKnowledgement (SACK) option. The energy consumption was obtained by measuring the time needed for the laptop's battery to get discharged rather than by direct measurement means. The study did not include other TCP variants and concentrated only on the impact of the SACK option. In [68], the authors studied the performance of different TCP variants (TCP Tahoe, TCP Reno, TCP New Reno, and TCP SACK) within static wireless ad hoc networks taking into account different ad hoc routing protocols (DSDV, DSR, and AODV). Using simulations, they evaluated both TCP throughput and communication energy consumption. However, the performance in case of link failure due to nodes' battery depletion or mobility was not studied. Moreover, none of these works [45, 67, 68] considered the computational energy cost in their studies (i.e. only the communication energy cost was taken into account). The computational energy cost of the TCP implementation was studied in [69]. The study was conducted through test-bed experiments in which random RTT delays and packet losses were applied. Here, again, using random values does not represent the actual behavior of wireless connections, since both RTT delays and data packet losses are correlated. Additionally, the authors in [69] did not evaluate the impact of the TCP loss recovery mechanisms on the computational energy cost but rather they focused on the energy costs induced by the chosen operating system and hardware platform.

None of the above mentioned studies considered evaluating TCP communication energy cost and computational energy cost in the same work. Also, using random values to represent RTT values or data packet losses do not accurately reflect the characteristics of ad hoc networks since losses and delays are correlated. We also note that the above studies did not investigate the mobility effect in mobile ad hoc networks (MANETs) environments. Therefore, we argue about the need to complement the above mentioned studies by realizing a clear and complete comparative study for the most commonly-used TCP variants (New Reno [62], SACK [70], Vegas [71], and Westwood [57])[3] undergoing different realistic data packet loss situations and RTT delays. To be complete, such study has to incorporate both computational and communication energy consumption of TCP. The ultimate goal of such study is to understand the impact of the different TCP loss recovery mechanisms on TCP performance within MANETs. Hence, the obtained conclusions can be used to derive design guidelines for new TCP enhancements suited for MANETs.

This complete and detailed measurement of TCP's energy consumption, resulting from both communication and computational energy costs, is among our major contributions. Our final objective is to produce a reference study using a hybrid approach, combining simulations and a realistic test-bed configuration, in order to reach the above mentioned objectives.

3.4.2 Contributions

In this work, we target the analysis of an important performance metric within one increasingly-important environment in which TCP is to be used. More precisely, we are interested in studying the energy cost of TCP when used in wireless multi-hop ad-hoc network environments, also known as Mobile Ad hoc Networks (MANETs). The major motivation behind this study resides in the fact that mobile devices are battery-operated and it is crucial to optimize their energy consumption in order to increase their batteries' lifetime. Prior to any improvement, there is a need to better understand how and where energy is consumed in the communication pipeline.

[3] We mention that both TCP Tahoe and TCP Reno are not studied in this work since they are obsolete even within wired networks.

The energy consumption of a node can be represented by two major parts: *(i)* the computational energy cost, and *(ii)* the communication energy cost. The computational energy cost of TCP is the energy spent within the node's CPU to perform the various copy operations, compute checksums, respond to timeouts or triple duplicate acknowledgments, adjust timers, and perform all other book keeping operations. This is the cost linked to the execution of the different TCP congestion control algorithms (Slow-Start, Fast Retransmit/Fast Recovery, and Congestion Avoidance). The communication energy cost is the energy consumed in the transmission and reception of TCP segments over the wireless links along the route between the source and the destination.

The undertaken work is a complete performance study of different TCP variants, within MANETs, in terms of energy consumption. The four major TCP variants, namely TCP New Reno, TCP SACK, TCP Vegas and TCP Westwood are considered. In order to measure both the computational and communication energy cost while executing their different congestion control algorithms, we implemented different data packet loss models (congestion, interference, link loss, and signal loss) and followed the following methodology:

- First, TCP's communication energy cost is evaluated through simulations using NS-2. This is performed when facing the above mentioned realistic data packet loss scenarios (network congestion, interference, link failures, and signal losses), which are obtained in a MANET setting.

- Then, we measured the node-level computational energy consumption, due to TCP operations, using a realistic test-bed configuration. This configuration introduces the effect of a real wireless multi-hop and mobile ad-hoc network environment (i.e. realistic data packet delays and losses). Such effects are introduced using a MANET delay and a packet-loss emulation tool, which we designed and implemented, called SEDLANE - Simple Emulation of Delays and Losses for Ad-hoc Networks Environment).

- Finally, the total energy cost of TCP is obtained by incorporating into NS-2 the measured computational energy cost of TCP's congestion control algorithms (Slow-Start, Fast Retransmit/ Fast Recovery, and Congestion Avoidance). Indeed, the NS-2 energy model does not include this feature. It applies only the communication energy cost.

Thus, after having obtained the communication energy consumption of the most commonly-used TCP variants when facing the different data loss situations that we identified as being common in MANETs (i.e. network congestion, interference, link failures, and signal losses), the following step is to complete this study by including TCP's computational energy cost. The first step towards building our test-bed configuration, as well as the associated methodology to measure the computational energy cost, is to design a tool that allows generating different realistic data packet loss situations and RTT delays that accurately reflect the characteristics of MANETs. For that purpose, we designed and implemented the "Simple Emulation of Delays and Losses for Ad-hoc Networks Environment (SEDLANE)" tool.

Figure3.2. The principle of SEDLANE operation.

Since it is easier to control the network environment using a simulation tool, we profit from this advantage by generating the desired scenarios within MANETs using NS-2 [63] as a network simulation tool. Then, we inject the trace files into Dummynet [72] as a traffic shaping tool to apply the same effects on a real traffic scenario, as if we construct a real multi-hop wireless ad hoc network environment. In this way, we obtain a Simple Emulation of Delays and Losses for Ad Hoc Networks Environment (SEDLANE) as depicted in Figure 3.2. SEDLANE allows emulating a multi-hop wireless ad hoc network of any scale in a virtual way and with any mobility scenario without the need of physically moving the ad hoc nodes. This tool enables us to test the performance of network protocols and applications (transport layer and above) at the end points of any mobile ad hoc network (emulated by SEDLANE). The final aim is to test the end-to-end nodes' performance that cannot be obtained through simulations, such as hardware-related parameters (e.g. computational energy cost at end points). SEDLANE allows evaluating the tested applications when running in real situations. The entire wireless ad hoc network environment is emulated using only one machine (e.g. a particular machine installed in the middle of the two communicating end points; or even one of the communication end-points that can also play this role). SEDLANE emulates network nodes' mobility, ad hoc routing protocol effect, and TCP connection throughput through RTT values (delays) and data packet losses within the network.

Once SEDLANE designed, implemented and proved to reproduce accurately the targeted MANET data loss and RTT delay situations, we integrated it into the measurement methodology used in [69], which we extended for the purpose of our study. Thus, thanks to the resulting implemented test-bed, depicted in Figure 3.3, we were able to measure finely the computational energy consumption of the main TCP functions (Slow-Start, Fast Retransmit/Fast Recovery and Congestion Avoidance) as well as compare the computational energy cost of the most commonly-used TCP variants (New Reno, SACK, Vegas, and Westwood). The former was thus incorporated into the NS-2 energy model allowing a comprehensive evaluation of the total energy consumption (both communication and computational) of these TCP variants using simulations.

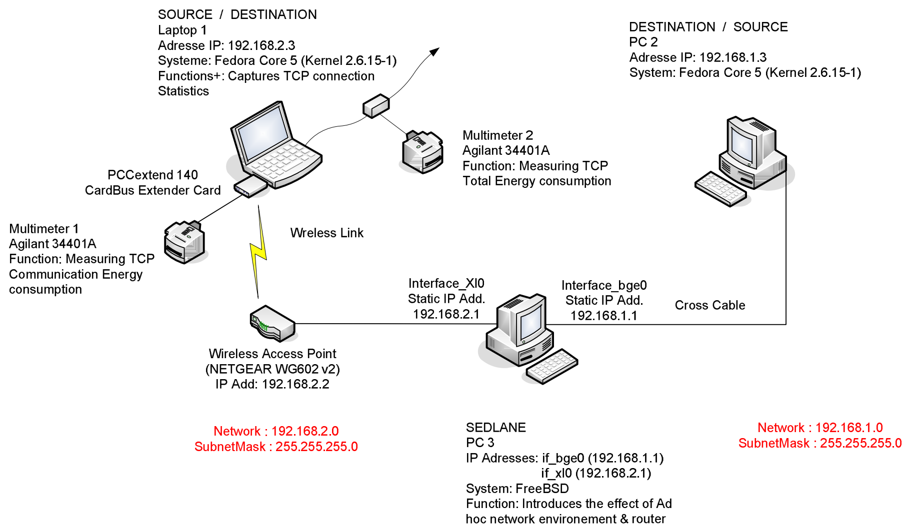


Figure 3.3. TCP Energy Consumption Measurements Test-bed.

Using NS-2, SEDLANE as well as the computational energy measurement test-bed, we had been able to conduct a complete study incorporating both, the computational and the communication energy consumption of TCP. This allowed reaching our ultimate goal which is to understand the impact of the different TCP loss recovery mechanisms on TCP performance within MANETs.

3.4.3 Summary of Results

To conclude, we can summarize our second contribution regarding the improvement of TCP behavior in wireless packet networks as follows. We conducted a complete study of the energy efficiency of four TCP variants (TCP New Reno, TCP SACK, TCP Vegas and TCP Westwood) within MANETs. We started by identifying the different data packet loss situations that TCP may confront within such networks: *(i)* network congestion, *(ii)* interference, *(iii)* link failure, and *(iv)* signal loss. Our study concerned both TCP computational energy and communication energy costs. The TCP's communication energy cost results were obtained using NS-2, while the computational energy cost results were obtained through a hybrid approach (i.e. using simulation results to configure a realistic test-bed and then perform accurate experiments).

In order to evaluate and measure one of the performance parameters targeted by our study, the computational energy cost of TCP, we established a measurement test-bed built around an emulation tool (SEDLANE) that we designed and implemented. SEDLANE uses both data losses and RTT delay values as network performance parameters in order to emulate the targeted connection. SEDLANE implements many calculation algorithms that extract the necessary data information from NS-2 trace files, and calculate the connection parameters used to configure the communication channels applied to manipulate the flow of data packets over the emulation test-bed. Before using SEDLANE in our measurement study, we first evaluated its accuracy. The validation results confirmed that SEDLANE is capable of emulating the different network parameters and having an accurate emulation of network performance. The results also showed that SEDLANE can emulate MANET scenarios and reproduce the same performances in terms of delay and data packet losses as in NS-2. This remains true regardless of the MANET features (such as mobility rates, ad hoc routing protocols, etc.). SEDLANE constituted the first important result of this contribution.

Unsurprisingly, the obtained results from our complete comparative study showed that current TCP variants cannot cope with all data packet loss situations found within MANETs. TCP behavior needs to be enhanced in order to handle the different loss causes other than congestion (link failure, signal loss, and interference). When looking at the reaction of TCP when faced with non-congestion loss types, we noticed high degradation in performance leading to waste of a scarce MANET resource: the nodes' available energy (batteries). Thanks to our complete performance study, we were able to show clearly that the existing TCP variants, which were originally developed for wired networks, do not always behave optimally when confronted with the different data packet loss causes that characterize MANETs. Some show good performance in certain cases and bad performance in others. It is worth to note that, when these variants were developed, caring about the nodes' energy consumption was not a big concern. Also, the variants that were developed with loss differentiation capabilities, as TCP Westwood, to enhance TCP performance within wireless networks do not take into consideration other data packet loss situations that would be faced within MANETs, such as link failures. We found also that the complexity of the implemented congestion control algorithms of TCP variant and their failure to cope adequately with certain loss causes may lead to unnecessary energy wastage at the node's level. Overall, we identified some tracks to follow in order to create a novel TCP variant that is energy-efficient in MANETs. Knowing where the most TCP energy consumption is spent is the main key to improve TCP functions and performance within MANETs. Such novel TCP variant constitutes the objective of the following contribution.

This work was one of the focuses of the Ph.D. dissertation carried out by Alaa Seddik-Ghaleb, University of Evry val d'Essonne, and defended on March 30, 2009 (see papers [P10][P35][P44][P48]). This work is also one of the main results of a fructuous cooperation with Orange Labs in the frame of a 3-year (June 2005 – May 2008) Research Grant (External Research Contract).

3.5 COUPLING LOSS AND DELAY DIFFERENTIATION TO IMPROVE TCP BEHAVIOR WITHIN MANETS

3.5.1 Research Context

Mobile Ad hoc Networks are differentiated from traditional wired networks by the multitude of packet loss situations they are subjected to. This is due to the intrinsic characteristics of wireless channels (e.g. signal fading, interference, obstacles, and environment effects) that might obstruct the proper reception of data packets at the other end. Moreover, in some cases, these vulnerabilities of wireless channels can result in a complete link failure. Although link failure is of low probability in wired networks, since physical cables constitute the data transmission media, it is rather common in MANETs (due to nodes' mobility, battery depletion or obstacles).

TCP is a transport protocol that aims at ensuring high reliability by guaranteeing the reception of data packets. Again, TCP was designed primarily for wired networks to address network congestions, which is the main cause for data packet loss in these networks. Therefore, other types of data packet loss encountered in MANETs are prone to misinterpretation by TCP, which, in turn, will lead to TCP performance degradation. Indeed, our previously obtained results, which are summarized in §3.4, show clearly that none of the current major TCP variants (New Reno, SACK, Vegas, and Westwood) is well adapted to deal with all data packet loss situations that can be encountered within MANETs. The results show that the performance of TCP degrades significantly within MANETs. Moreover, some of the studied TCP variants perform well in certain cases while they perform badly in other cases.

The results of our previous work (cf. §3.4.) show that the ability of some TCP variants (e.g. TCP Westwood and TCP Vegas) to distinguish among congestion-induced and wireless-related data losses leads to an improved performance in some cases. However, TCP variants that incorporate a Loss Differentiation Algorithm, either implicitly or explicitly, do not consider all types of data packet losses that can be encountered within MANETs. In fact, they consider congestion-induced and wireless-channel-related losses only. While in MANETs, we have other data packet loss causes such as link failure that might be due to nodes' mobility or nodes' battery depletion. Hence, we conclude that each data packet loss situation necessitates a specific and adequate recovery reaction in order to optimize TCP performances.

The work done in this domain [73–80] addresses the problems of TCP within wireless infrastructure networks where only the last hop is the wireless channel. Most of this work takes into consideration how to distinguish between wireless-channel-related and congestion-related data packet losses. Within wireless infrastructure networks, a link failure affects only the end user (the concerned node), while within wireless infrastructure-less networks, such as MANETs, a link failure affects the whole network as each node might forward data packets to other network nodes. From this perspective comes the importance of dealing with link-failure-induced losses, since it is a very common situation within MANETs. Thus, it is necessary to develop a new TCP variant that is able to deal with the most common MANET-related data packet loss causes. The key features that such new TCP variant should have are:

1. Distinguish among the different data packet loss causes: *(i)* network congestion induced losses, *(ii)* wireless channel related losses (interferences and signal losses), and *(iii)* link failure induced losses.

2. Recover from data packet losses by dealing with each identified data packet loss cause accordingly: stopping data transmission when necessary or adjusting the TCP performance parameters (CWND and RTO) when needed. Such recovery should take into consideration network resources optimization through: *(i)* minimizing radio communication energy consumption, *(ii)* minimizing total TCP's node energy consumption, and *(iii)* maximizing TCP throughput.

Improving TCP behavior within MANETs is among our major contributions. Our final objective is to present a new variant of TCP that is most adapted to the nature of Mobile Ad hoc Networks and their most common data packet loss causes.

3.5.2 Contributions

According to the conclusions gained from our previous performance study (cf. §3.4), we developed TCP-WELCOME, which stands for "TCP variant for Wireless Environment, Link-losses and COngestion packet loss ModEls", a new TCP variant for MANETs. TCP-WELCOME is developed with the aim to optimize TCP performance from the following points of view: maximizing the utilization of the available bandwidth (throughput), while minimizing the nodes' overall (communication and computational) energy consumption.

In fact, the problem of TCP within MANETs comes from its inability to distinguish between the different data packet loss causes within the network. This often leads to an aggressive reaction from TCP when facing a data packet loss that is not due to congestion. Indeed, dealing with any data packet loss as if it was due to congestion results in resource waste both at the network and nodes' levels. This waste is represented by a high energy consumption (as discussed in §3.4) but also by a low bandwidth utilization. Therefore, in order to enhance its performance within MANETs, it is obvious that TCP should be able to identify each data packet loss cause and react accordingly, triggering the most suitable loss recovery action that optimizes both the network and node's resources.

TCP-WELCOME is designed to distinguish among and recover from the following data packet loss situations that we identified: *(i)* link failure, *(ii)* wireless channel related errors and *(iii)* congestion. It includes a Loss Differentiation Algorithm (LDA) that is able to distinguish among the above mentioned data packet loss situations, and a Loss Recovery Algorithm (LRA) that is capable to recover adequately from each of these situations. It is an end-to-end sender side modification to the legacy TCP (i.e. TCP New Reno) in which our proposed LDA and LRA schemes are introduced.

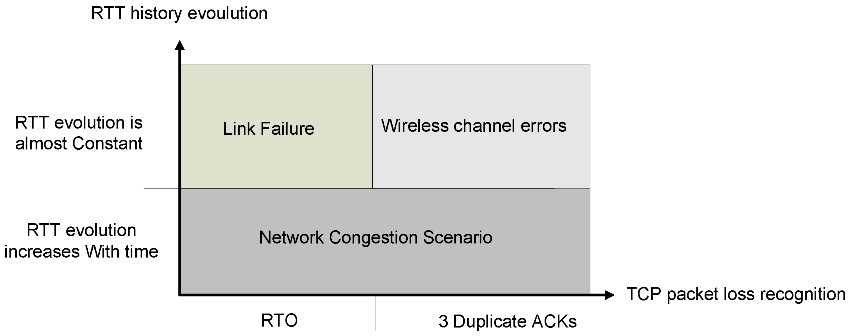


Figure 3.4. TCP-WELCOME - Proposed Loss Differentiation Algorithm

As depicted by Figure 3.4, the main idea upon which is built our LDA scheme is based on observing the history of Round Trip Time (*RTT*) samples evolution within the network and the way in which TCP identifies the data packet loss (Three *Duplicate Acknowledgments* vs. *Retransmission TimeOut* expiration). It is the first scheme that proposes coupling both parameters in order to classify the cause of packet losses. Indeed, usually in the literature, only one of these parameters is used. Once the data packet loss cause is identified, an appropriate reaction is taken by our proposed LRA scheme. Figure 3.5 depicts the different reactions promoted by our LRA in order to enhance the performance of TCP when confronted with each of the three above mentioned packet loss situations. Our LRA allows also adjusting TCP parameters (*RTO* and *CWND*) according to the identified packet loss cause.

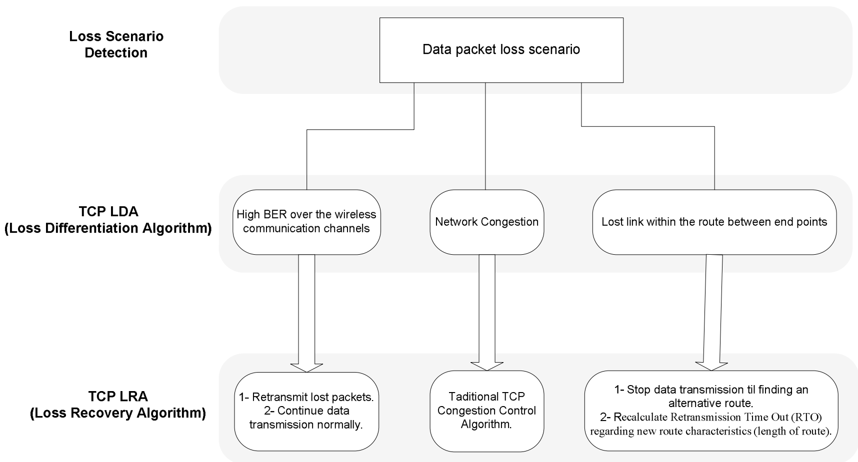


Figure 3.5. TCP-WELCOME - Proposed LDA and LRA Schemes

The loss differentiation and loss recovery rules used in TCP-WELCOME for each considered packet loss cause are summarized in the following.

3.5.2.a Wireless Channel Related Losses

A data packet loss is classified as due to wireless channel errors, experienced within one of the links over the communication path, if the evolution of *RTT* samples over the connection is not highly fluctuating (i.e., staying around an average value) and the data packet loss is identified through three *Duplicate Acknowledgements*. Indeed, our various experiments showed that when the connection is experiencing wireless-channel-related losses, both queuing and propagation delays stay almost constant. In this case, *RTT* samples over the connection are not expected to experience high fluctuations with time. Additionally, when there is a valid route between the source and the destination, despite the presence of wireless channel errors, the source can always receive acknowledgements from the destination. The corruption of acknowledgements is of a lesser probability since the acknowledgement packet size is relatively small. This is why we have *RTT* samples not highly fluctuating and also the ability to receive three *Duplicate Acknowledgements*. We thus associated this pattern to the wireless channel related losses.

When a packet loss is classified as due to wireless channel errors, there is no need for TCP to adapt its parameters. Indeed, this is not necessary since we are using the same route and no congestion is sensed over this one. So, the loss-recovery rule is to perform packet retransmission but no decrease (or increase) in the Congestion Window (*CWND*) size or the Retransmission TimeOut (*RTO*) value is triggered.

3.5.2.b Link Failure Related Losses

When a link failure occurs along the route towards the destination, all the packets of the concerned TCP connection are lost in a burst. In this case, the TCP sender does not receive duplicate acknowledgments. Instead, it recognizes data packet losses through *RTO* expiration. Usually, this situation is interpreted by the legacy TCP as a strong congestion situation. In our case, we suggest to verify the evolution of the *RTT* samples of the already acknowledged packet. If this evolution is relatively constant, then the loss is classified as due to a link failure and not to congestion. Indeed, in the congestion case we expect a regular increase in the *RTT* values before packet losses are experienced.

Typically in MANETs, the routing protocol tries to recover the link loss by finding a new route towards the destination. Our observation is that both propagation and queuing delays through this new route are often different compared to those of the old route. This is obvious since the load and the length of the new discovered route is more likely to be different from those of the old route. So, the TCP parameters (*CWND* and *RTO*) used through the old route are not valid anymore and need to be adapted.

The ration between the *RTT* values over both routes is the indicator that allows us to estimate this difference. Let RTT_{old} be the delay over the lost old route and RTT_{new} be the delay over the new recovered route. In TCP-WELCOME we propose to adapt the *RTO* and *CWND* values as follows (i.e. after route recovery):

$$RTO_{new} = \left(\frac{RTT_{new}}{RTT_{old}} \right) RTO_{old} \tag{1}$$

$$CWND_{new} = \left(\frac{RTT_{old}}{RTT_{new}} \right) CWND_{old} \tag{2}$$

3.5.2.c Network Congestion Related Losses

In case of congestion, the queuing delay at the concerned network nodes increases gradually. Indeed, in this case, their buffers are more and more filled with time. So, if a packet-loss is experienced and the evolution of *RTT* samples at the sender side is increasing progressively, then, the loss is classified as due to network congestion. This is performed regardless of how TCP recognizes the data packet losses: three *Duplicate Acknowledgments* reception or *RTO* expiration. In this case, the TCP reaction is kept unchanged (i.e. normal TCP New Reno behavior is triggered).

3.5.3 Summary of Results

To conclude, we can summarize our last contribution regarding the improvement of TCP behavior in wireless packet networks as follows. We proposed TCP-WELCOME, a new TCP variant that is suitable for MANETs. Unlike other TCP variants, it uses a Loss Differentiation Algorithm (LDA) that is able to identify accurately the three common data packet loss causes within such environments: network congestion, wireless channel related errors, and link failures. To do so, our LDA couples delay and loss information. TCP-WELCOME adopts also a new Loss Recovery Algorithm (LRA) that reacts efficiently to each identified data packet loss cause with the most appropriate action - i.e. by adapting adequately TCP congestion-control parameters (RTO, CWND, and SSThreshold). The design guidelines behind TCP-WELCOME had been derived thanks to the conclusion drown from the complete comparative study of the most commonly-used TCP variants when used in MANETs. This comparative study has been performed earlier and is presented above (cf. §3.4).

In order to show the performance improvement of TCP-WELCOME, we implemented it in both Linux Kernel and the Network Simulator version-2 (NS-2). We used the same evaluation methodology we designed for our previous contribution regarding "*TCP Computational and Communication Energy Cost in MANETs*" (cf. §3.4). However, this time not only the energy cost was evaluated but also the throughput. Hence, we compared TCP-WELCOME performances with the four most-commonly used TCP variants under different data packet loss scenarios (congestion, interference, link failure, and signal loss). This comparative study showed that both TCP average throughput and total energy consumption have been significantly improved. We also showed that TCP-WELCOME outperforms other TCP variants in most cases thanks to its ability to identify correctly the data packet loss cause through its LDA and to take the most appropriate action to recover from data packet losses through its LRA.

This work was one of the focuses of the Ph.D. dissertation carried out by Alaa Seddik-Ghaleb, University of Evry val d'Essonne, and defended on March 30, 2009 (see papers [P37]). This work is also one of the main results of a fructuous cooperation with Orange Labs in the frame of a 3-year (June 2005 – May 2008) Research Grant (External Research Contract).

3.6 CONCLUSION

The work presented here was realized in the frame of two Ph.D. dissertations for which I was the main supervisor: Alaa Seddik-Ghaleb (September 2005 – March 2009) and Stéphane Lohier (June 2004 – June 2006). It was also partly supported by a 3-year (June 2005 – May 2008) Research Grant (External Research Contract) from Orange Labs.

The obtained results are numerous. We started from the well-known fact that TCP does not always perform optimally depending on the network environment in which it is used. More precisely, we focused on two wireless network environments: Wireless Local Area Networks and Mobile Ad hoc Networks. For each of these network environments, our research started by setting-up experimental test-beds that allowed us to quantify the performance degradation experienced by TCP within such environments and also to identify some trails for performance improvements. Our literature review in this domain showed that the existing solutions clearly do not respond to the drawbacks that we identified. This is what motivated the design of the set of mechanisms presented in this chapter. We found out that using cross-layer design in the case of WLAN and coupling loss and delay information in the case of MANETs are the good trails to follow in order to handle the identified drawbacks, and thus to improve wireless resource usage by TCP in such networks. Thus, we demonstrated the effectiveness of these approaches using an exhaustive performance study. In both cases, we compared our approach to other major solutions from the literature showing the clear performance improvement we obtained.

Chapter 4. Mastering QoS in Wireless Packet Networks: WLAN and BWAN Case Studies

The next generation wireless networks are expected to support a multitude of applications ranging from data services, such as e-mail, fax, and stock-market information, to bandwidth-intensive multimedia services, such as networked games, video on demand, wireless TV as well as audio and videoconferencing. The characteristics of wireless links, as well as the desire to maintain a certain level of QoS while having the freedom to be connected anywhere and anytime, offer significant challenges for resource management. Indeed, the available channel resources to support such multitude of applications remain limited, and therefore its proper management is necessary to ensure the required QoS. This situation remains valid in all kind of wireless packet networks.

In the following, we explore how such a resource management solution can be designed in the case of *Wireless Local Area Networks* (*WLANs*) as well as *Broadband Wireless Access Networks* (*BWANs*). We will first remind the context of our research as well as the challenges that faces resource management in such networks. This will be followed by the description of our three contributions and the results we obtained in this domain. These contributions concerns the study of efficient model-based admission control for 802.11e WLANs, as well as improved scheduling for relay-assisted OFDMA-based BWANs such as 802.16j and 802.16m networks. This chapter ends with a brief conclusion summarizing our research experience in this domain.

4.1 Research Context

The increasing availability of wireless broadband networks has accelerated the widespread use of multimedia services. Wireless technologies have been evolving towards high data rates and high Quality of Service (QoS) support to ensure high-quality multimedia services such as wireless TV, mobile gaming, video on demand, audio and videoconferencing, and so on. The capabilities of portable devices continue to proliferate to accommodate future wireless multimedia services. However, ensuring an efficient support of such services over wireless packet networks remains a challenging task. In such networks, the flows of the different running multimedia applications share the available wireless resources. Hence, at each new flow arrival, the existing flows may loose a certain degree of their already achieved performance. Therefore, a resource management solution, controlling the share of the wireless resources and improving their use, is required to master the offered QoS whatever the wireless packet network technology.

In this chapter, we target two wireless network contexts, namely 802.11e Wireless Local Area Networks and Relay-assisted Broadband Wireless Access Networks, in which we target to propose adequate resource management solutions. In 802.11e WLANs, we will see that the main QoS issues come from the fact that in such networks there is no guarantee in terms of throughput and delay assurance for real time and multimedia services. Thus, an efficient admission control scheme, to be associated to the Enhanced Distributed Channel Access (EDCA) MAC protocol, is needed in order to control the number of active flows in a given WLAN setting and thus to allow offering the missing guarantee. Conversely, in Relay-assisted BWANs, the main QoS issue comes from the fact that most of the research in this domain considers that there is no need for queuing the packets to be relayed within the Relay Stations. We will see that this is not always true and that an adequate scheduling is required at the Relay Station to allow ensuring a certain level of QoS in such networks.

The target of our research, presented in this chapter, is to study the suitability and effectiveness of existing resource management solutions in 802.11e Wireless Local Area Networks as well as Relay-assisted Broadband Wireless Access Networks. Then, according to the identified drawbacks, to propose a new set of resource management solutions that allows improving the wireless resource usage within such wireless packet networks.

4.2 MAIN CHALLENGES

From the above discussion, it is clear that a resource management solution, controlling the share of the wireless resources and improving their use, is required to master the offered QoS within wireless packet networks. It is obvious that such a resource management solution have to match the specificities of the wireless packet network technology in which it is used.

If we look at IEEE 802.11 WLANs, the increasing demand for real-time multimedia application support led to the specification of a new standard amendment: IEEE 802.11e. This standard amendment aims to support QoS by providing differentiated classes of service in the MAC layer so that it can deliver time-critical multimedia traffic, in addition to traditional data packets. However, despite these enhancements of the MAC layer, this new standard still cannot guarantee the required QoS for real-time multimedia applications such as voice and video. For this reason, and as it stands, the IEEE 802.11e WLAN standard cannot be directly used as a framework for multimedia applications. In fact, when the network is not fully saturated, the service differentiation works properly and gives the multimedia traffic the opportunity to be served with a good QoS. However, a problem arises when the network starts to reach the saturation state. In this case, all traffic flows suffer from high collision rate, and the QoS for multimedia applications is degraded. To solve this problem and to extend the capability of 802.11e WLAN to deliver multimedia traffic with success, an efficient admission control must be used jointly with the 802.11e standard. To be able to guaranty the QoS requirements in terms of required bandwidth and tolerated delay, this algorithm must have as input parameters: the accurate values of the achievable throughput and access delay. One effective way to obtain such input parameters is to use an analytical model of the protocol. In this case, the admission control algorithm is denoted as model-based. We thus argue that an accurate model-based admission control algorithm is the missing brick for providing the ability to manage effectively the wireless resources in 802.11e WLANs.

Alternatively, while looking to the recent advances in Broadband Wireless Access Networks, we notice the increasing use of relay technologies. Indeed, the relay-assisted Orthogonal Frequency-Division Multiple Access (OFDMA) cellular system is a promising architecture for future wireless communication systems. Cooperative relaying allows to a source node (the *Base Station* or a *Mobile Station*) to use the antenna of one of its neighbors (the *Relay Station*) to communicate with a destination node (a *Mobile Station* or the *Base Station*) when the direct communication channel is poor. By doing so, such wireless packet network architecture promises to provide significant enhancements of the system capacity and coverage. However, if not handled adequately, some special features of the relaying technology may lower its positive aspects. Indeed, as link qualities are time varying, it may happen that a non-optimal coordination among the Base Station, the Relay Stations and the Mobile Stations lead to resource wastage. So, in order to improve this coordination, an adequate scheduling is to be implemented in the Base Station and the Relay Stations. However, it is often assumed in the literature that the packets transiting by the Relay Station are immediately relayed to their destination without the need of scheduling. This is not realistic when the link quality towards the destination is temporarily bad. We thus argue that an adequate scheduling at the Relay Station is the missing brick for providing the ability to manage effectively the wireless resources in Relay-assisted BWANs.

Thus, in the following, we first endeavor to develop and validate a complete analytical model that highlights the exact behavior of the 802.11e EDCA MAC protocol. Then, based on this complete analytical model, we propose an efficient model-based admission control scheme allowing the efficient control of the wireless resource usage in 802.11e WLANs. Finally, and as a second part of this chapter, we propose adequate scheduling algorithms to be used in relay stations and aiming to improve the wireless resource usage in Relay-assisted BWANs.

4.3 ANALYTICAL MODELING OF THE 802.11E MAC

4.3.1 Research Context

In the recent years, a paramount number of studies appeared in the literature investigating the performance of 802.11 Distributed Coordination Function (DCF) and 802.11e EDCA protocol. This performance has been explored by means of not only simulations but also analytical models, with the aim to either predict analytically performance metrics or to understand the protocol behavior. Almost all models for EDCA extend the Bianchi's two-dimensional Markov chain model [81] that was proposed for 802.11 DCF.

Two types of analytical models exist in the literature: the first one is based on the discrete Markov chain model, while the second one is based on the mean value analysis. At the level of our knowledge, there is no accuracy comparison between these two approaches. But, it is clear that the first approach gains a great success in the field of 802.11 modeling. The second approach, despite its low computation complexity, omits system details and necessitates a high number of assumptions. Alternatively, the analytical models proposed can be classified into two groups: the first one includes models under network saturation conditions and the second one is related to those applicable to all network conditions. The saturation assumption enables queuing dynamics to be neglected and avoids the need for detailed modeling of traffic characteristics. So, the saturation assumption greatly simplifies the modeling. On the other hand, although performance evaluation under saturation conditions provides the fundamental bounds on system throughput and delay, it cannot reveal the best working scenarios. It was proved by simulation, as well as by older models, that the maximum capacity of IEEE 802.11 [82, 83] and 802.11e [84, 85] can only be achieved in the non-saturated case. So, a framework capable of analyzing and predicting the performance under both saturation and non-saturation conditions can be very helpful in leading to first better understanding the EDCA mechanism behavior, and at a latter extent facilitating the design of an accurate admission control algorithm.

Table 4.1. Comparison of Major EDCA Models.

	CW	AIFS	TXOP	RL	B stop	Int. Coll.	Th	D
[86]	✓	~	×	✓	×	×	✓	✓
[87]	✓	~	×	×	×	×	✓	×
[88]	✓	×	×	✓	~	×	✓	✓
[89]	✓	~	×	×	×	×	✓	✓
[90]	✓	~	×	×	×	×	✓	✓
[91]	✓	~	×	✓	×	×	✓	×
[92]	✓	~	~	✓	✓	×	✓	×
[93]	✓	~	×	✓	✓	×	✓	×
[94]	✓	~	×	✓	×	~	✓	×
[95]	✓	~	×	✓	✓	~	✓	✓
[96]	✓	~	×	✓	✓	✓	✓	✓
[97]	✓	~	×	✓	✓	✓	✓	×
[98]	✓	~	×	✓	✓	✓	✓	×
[99]	✓	~	×	✓	✓	✓	✓	✓
[100]	✓	~	×	✓	✓	✓	✓	✓
[101]	✓	~	×	✓	✓	✓	✓	✓

✓=exists ×=does not exist ~=exists but not accurately

In order to verify whether such a framework already exists in the literature, we performed an extensive study of the value added and the limitations of the major already-existing models for 802.11e EDCA. This extensive study compared the features of the most important models from the literature [86–101]. One should note that other models exist in the literature and are not discussed here. In fact, these other models are small derivations of the ones we chose to include in our study. Table 4.1 illustrates the summarized comparative study of the main features of the major existing models. This comparative study can be very helpful to understand how EDCA modeling evolved. Also, it shows the missing features to be completed in order to obtain a more accurate analytical model. Such accurate model is needed for the design of an accurate admission control algorithm in which it will be the basis. The performed comparison is based on: the EDCA parameters that are modeled (CW: Contention Window, AIFS: Arbitrary InterFrame Spacing, and TXOPLimit: Transmission Opportunity Limit), the baseline features of EDCA that are considered in the model (RL: Retry Limit, B stop: Backoff freeze, and Int. coll.: Internal collision), and the performance metrics that are calculated through it (Th: Throughput, and D: Access delay).

From this table, we can draw two main conclusions:

- First, even though almost all EDCA models take into account AIFS and CW differentiation, they unfortunately neglect the use of different TXOPLimit. The TXOPLimit differentiation parameter is very important as has been shown in the analysis done in [102]. It was demonstrated that it has a noticeable effect on the global performance.

- Second, almost all EDCA models do not accurately consider the modified backoff countdown operation in the 802.11e EDCA. This one should cooperate with the AIFS countdown operation. After each frozen counter in the backoff, the channel must be sensed idle during the whole AIFS period before the backoff countdown procedure starts again. Many of the previously presented models do not reiterate correctly the AIFS countdown before unfreezing the backoff counter. Actually, these models, which are only two-dimensional Markov chains (one for the backoff stage and one for the backoff counter), could not cover this co-operation process because of this limitation. Therefore, a three-dimensional Markov chain is minimally required.

As a conclusion, all the previous works address only partially the challenges for modeling EDCA. Such partial model may be sufficient if the target is to analyze the effect or the performance characteristics of one parameter taken apart. However, when an accurate analysis of the access delay and achievable throughput is needed or when an analytical model-based admission control algorithm need to be designed, it is then clear that a more complete and accurate model is required.

The foreseen analytical model for 802.11e EDCA is among our major contributions. Our final objective is to propose a complete analytical model that highlights the exact behavior of the EDCA MAC protocol [103]. Hence, the requirements of such new model are as follow:

1. It should consider the TXOPLimit differentiation parameter accurately. It should highlight the impact of the three differentiation parameters (AIFS, CW and TXOP) on the performance of the system.

2. It should consider the AIFS and backoff co-operation process in EDCA as defined by the standard.

3. It should capture all QoS specific features for EDCA as described in the standard [103] and model these features correctly. All the aspects of the protocol such as the AIFS countdown procedure, the backoff suspension, the virtual collision between the ACs of the same station, the external collision with other stations, the post collision, the post backoff and the retry limit should be present in the model.

4. Furthermore, it should handle general traffic conditions and should not be limited to the saturation conditions assumption.

As explained above, the resulting analytical model is the first step towards designing an accurate admission control algorithm and thus mastering the 802.11e wireless resource usage.

4.3.2 Contributions

One of the ultimate goals of 802.11e analytical modeling research is to obtain a model that provides the most accurate prediction of the performance metrics, such as achievable throughputs and access delays, in order to be used in an efficient admission control algorithm. As explained earlier, the currently proposed 802.11e EDCA models, even though numerous, do not reach this accuracy due to the several simplifications that have been done. As will be shown in the following, it is possible to propose an analytical model that is characterized by a high accuracy level and a low computational overhead. In the following, we summarize our major results in this domain: the proposed Markov chain model as well as the major related formulas.

4.3.2.a Markov Chain Model

To be capable of introducing all EDCA features in our model, our approach consists of following the states in which an Access Category, or AC, (Traffic Class in 802.11e EDCA terminology) transits during its transmission cycle. This approach is the key that makes our model complete and therefore differentiates it from existing models. In our model, each state represents an AC in a time slot (802.11e EDCA suggests the use of four ACs per station). The states that an AC can occupy at a randomly chosen time slot are grouped into seven periods representing the EDCA operations [103]: *Idle period, AIFS period, backoff period, frozen period, collision period, post collision period and transmission period.*

The proposed Markov chain is a discrete four dimensional one drawn for each AC. The first dimension *p(t)* indicates the period in which the AC is at time *t*. The second dimension *s(t)* represents the backoff stage, and the third one *b(t)* denotes the value of the backoff counter. Finally, the fourth dimension *r(t)* indicates the remaining time to leave the current period. So, *(p(t),s(t),b(t),r(t))* is a discrete time Markov chain based on the assumption that the collision probability p_c, the channel busy probability p_b and the empty queue probability p_e are constant.

At each time slot, the state of each AC is determined by (i,j,k,d); i = I stands for Idle, A for AIFS, F for Frozen, B for Backoff, C for Collision, PC for Post Collision and T for Transmission. $j = 0,1,2\ldots m$; m being the retry limit; k is uniformly chosen from $[0,w_j]$ where w_j depends on the backoff stage and satisfies $w_{j+1} = 2w_j +1$ when $w_j < w_{max}$ $(w_0=w_{min})$, and $w_{j+1} = w_j$ when $w_j = w_{max}$. Finally, d, the time remaining to leave the period, depends on the value of i (i.e. on the period itself). The entire Markov chain is traced in Figure 4.1. The content of the bloc concerning the frozen and AIFS periods is drawn in Figure 4.2.

The transition probabilities are based on p_c (collision probability), p_b (channel busy probability) and p_e (probability of empty queue). To resolve the Markov chain system, we consider $P_{i,j,k,d}$ the steady state probability of (i,j,k,d), then we first assume $P_{B,0,0,0}$ is known and we calculate the other state probabilities with respect to it. Finally, to get $P_{B,0,0,0}$, we resolve the normalization condition: $\sum_{i,j,k,d} P_{i,j,k,d} = 1$.

Equation (1) gives a relation between $P_{B,0,0,0}$, p_c, p_b, p_e and other system parameters (namely EDCA and network parameters). So, to resolve the system, we just need to compute p_c, p_b and p_e.

$$P_{B,0,0,0} = \begin{bmatrix} \dfrac{(1-p_e)p_c(1-p_c{}^m)}{1-p_c}\left(1+[T_c]+\dfrac{1}{p_b}\left(\dfrac{1}{(1-p_b)^{[T_{pc}]}}-1\right)\right)+\dfrac{1}{2(1-p_b)}\left([N]p_b+\dfrac{1}{(1-p_b)^{[A]}}\right)\left(\sum_{j=1}^{m} p_c{}^j w_j(1-p_e)+w_0\right) \\ +1+[W]p_e+\dfrac{1-(1-p_b)^{[A]}}{p_b(1-p_b)^{[A]}}+[T_s](1-p_c^{m+1}(1-p_e)-p_c p_e) \end{bmatrix}^{-1} \quad (1)$$

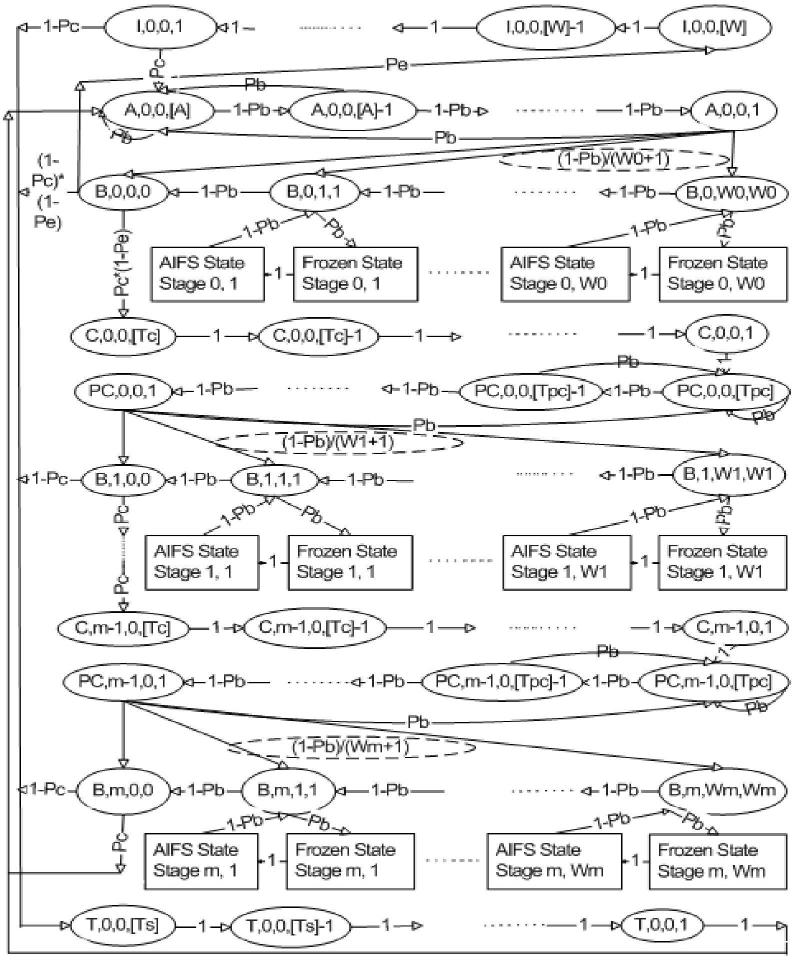


Figure 4.1. Markov chain of the model.

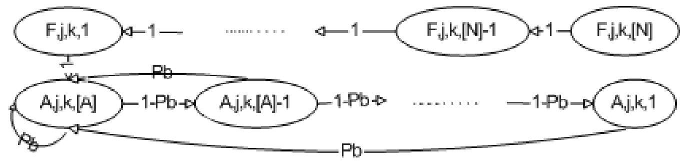


Figure 4.2. Frozen and AIFS blocs (j, k); 0≤j≤m, 1≤k≤wj.

A- Computation of p_{ci} and p_b

Let us first see how a collision may occur, and when the channel is considered as busy. Our idea is to firstly get the value from the viewpoint of an AC, then from the viewpoint of a station. Moreover, we differentiate between virtual (i.e. collisions among ACs of the same station) and external (i.e. collisions of frames from different stations) collisions. We assume a fixed number of active stations *M* in the same radio range. In each station, all ACs are active. Let τ_i be the probability for an AC_i attempt to access the channel in a random time slot. According to our model:

$$\tau_i = P_{I,0,0,1} + \sum_{j=0}^{m} P_{B,j,0,0} = \left((1 + p_{ei}) + \frac{p_{ci}(1 - p_{ci}^m)(1 - p_{ei})}{1 - p_{ci}} \right) P_{B,0,0,0} \tag{2}$$

From the viewpoint of a station, the probability τ that it accesses the channel is the probability that at least one of its ACs tries to access the channel:

$$\tau = 1 - \prod_{i=0}^{3} (1 - \tau_i) \tag{3}$$

The internal collision probability for AC_i is the probability that at least one of the higher priority ACs tries to access the channel simultaneously with it, so $P_{c\mathrm{int}_i} = 1 - \prod_{j>i}(1-\tau_j)$, while the external collision probability is the probability that at least one of the other (*M-1*) stations tries to access the channel at the same time slot, $P_{cext_i} = 1-(1-\tau)^{M-1}$.

So, the total collision probability is given by:

$$P_{ci} = 1-(1-\tau)^{M-1}\prod_{j>i}(1-\tau_j) \tag{4}$$

For the busy channel probability, p_b, let first compute v_i, the probability for the channel to be occupied by an AC_i. Thus, v_i corresponds to the probability for AC_i to be in transmission or external collision state:

$$v_i = \left([T_{si}]\left(1-p_{ci}^{m+1}(1-p_{ei})-p_{ci}p_{ei}\right)+[T_{ci}](1-p_{ei})\frac{p_{ci}-p_{ci}^{m+1}}{1-p_{ci}}p_{cext_i}\right)P_{B,0,0,0} \tag{5}$$

The probability for the channel to be occupied by a station is the probability for it to be occupied by one and only one AC from this station, so:

$$v = \sum_{i=0}^{3} v_i \prod_{j\neq i}(1-v_j) \tag{6}$$

Finally, the channel is considered as idle if none of the stations is using it. Thus:

$$P_b = 1-(1-v)^M \tag{7}$$

B- Computation of p_{ei} and $[W_i]$

To calculate p_{ei}, the probability that the AC_i queue is empty, and $[W_i]$, the mean waiting time during the Idle period, we assume a Poisson arrival process of rate λ (packet/s).

For this consideration, $\rho_i = \lambda_i D_i$ is the probability that the queue contains data to be handled, therefore the probability that the queue is empty is:

$$p_{ei} = 1-\rho_i = 1-\lambda_i D_i \tag{8}$$

where D_i is the mean service time, which is here the mean access delay of AC_i. For the average idle time, it is obvious that an AC transits to idle state after the completion of the last transmission and the post backoff, while the queue remains empty. So, the average waiting time depends on these times as well as on the packet inter-arrival time. It is thus obtained using the following equation:

$$[W_i] = 1/\lambda_i - D_i - [T_s] - T_{PB} \tag{9}$$

where $1/\lambda_i$ is the packet inter-arrival time, and T_{PB} is the time needed to achieve the post-backoff procedure.

4.3.2.b Throughput Analysis

Using the model obtained above, we can easily derive the achievable throughput for each AC as follows:

$$S_i = p_{ei}S_{insat} + (1-p_{ei})S_{isat} \tag{10}$$

where S_{insat} is the non-saturated throughput and S_{isat} is the saturation throughput.

$$S_{insat} = \frac{M p_{ei} E[P] \rho_i / 1 - \rho_i}{T_{PBi} + [W_i] + p_{ci} D_i + (1 - p_{ci})} \tag{11}$$

$$S_{isat} = \frac{p_{si} E[P] N_{TXOP_i}}{(1 - p_b) + p_b \sum_{j=0}^{3} p_{sj} T_{si} + p_b \left(1 - \sum_{j=0}^{3} p_{sj}\right) T_c} \tag{12}$$

In this last equation: $p_{si} = M \sum_{d=1}^{[T_s]} p_{T,0,0,d} (1 - v)^{M-1} \prod_{j>i} (1 - v_j)$ $N_{TXOPi} = \left[\frac{TXOP_i}{T_{s1} + SIFS} \right]$, $T_{si} = \mathrm{N}_{\mathrm{TXOPi}} (T_{s1} + SIFS)$, E[P] is the average packet payload size, T_{s1} is the transmission time for one frame, and T_c is the collision time.

The effect of the TXOP parameter appears in our model in the computation of the transmission time T_{si}. In fact, during a TXOPLimit, a station may be allowed to transmit multiple data frames from the same AC with a Short Inter Frame Space (SIFS) between an ACK and the subsequent data frame, as illustrated in Figure 4.3. A zero value for TXOPLimit indicates that a single frame may be transmitted. T_{si} computation takes this into consideration.

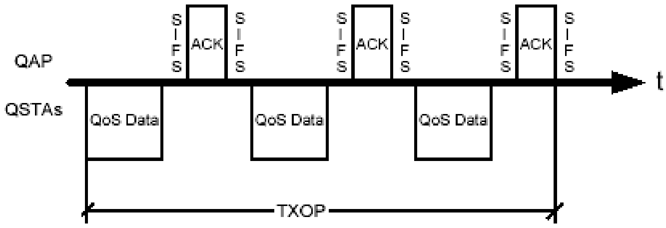


Figure 4.3. Transmission during TXOPLimit

4.3.2.c Access delay analysis

We define $D_{i,j,k,d}$ as the delay needed to go from the state (i,j,k,d) to the successful transmission state. The fourth dimension of our Markov chain allows us to compute recursively the delays for all Markov chain states. The final equation that gives the access delay in saturation and non-saturation conditions is therefore equal to:

$$D_i = (p_{ei} p_{ci} + 1 - p_{ei}) D_{A,0,0,[A]} + p_{ei} (1 - p_{ci}) \tag{13}$$

where $D_{A,0,0,[A]}$ is the access delay of the first state of the Markov chain.

4.3.2.d Model resolution

The analytical model proposed is a non-linear equation system ((1), (4), (7) and (8)). It can be resolved by means of numerical methods. Once resolved, transition probabilities will be known. These probabilities are then used in the throughput and delay computation using two simple equations: (10) and (13). Equations (10) and (13) represent the close-form expressions developed to simplify the computational costs.

Hence, the computational complexity of our model resides only in the non-linear system resolution. This contains 5 nested "for loops": one for p_b and four for p_c (i.e. one for each AC). Of course, if the program tries all possible values for p_b and p_{ci}, it will lead to a high complexity. In order to reduce the computational complexity of our model resolution, we suggest here the use of a simple method consisting of searching the value that leads to the smallest estimation error in a given interval, and then we re-calculate around this value till we get the solution. This method, illustrated in Algorithm 4.1, greatly reduces the calculation time and gets the exact solution with a very small estimation error (~10^{-6}) and in a short period of time.

Algorithm 4.1. Optimised model resolution.

i=0.1, p=0.5, b=0.5,c(i)=0.5/ initialize step and interval */*

while *(estimation errors for p_b and p_{ci} > specified error)* ***do***

    while *(i > 0.00001)* ***do***

      for *p_b=b-p : i : b+p* ***do*** */* each time do 10 iterations*/*

        calculate p_b that leads to smallest error given $p_c(i)$

      end for

      b=pb, p=p/10, i=i/10 / starts from the new value*/*

    end while

  / Here the best value of p_b is given*

  for *i=1:4* ***do*** */* do the same thing for each $p_c(i)$*/*

    while *(i > 0.00001)* ***do***

      for *pc(i)=c(i)-p : i : c(i)+p* ***do***

        calculate pc(i) that leads to smallest error given pb

      end for

      c(i) = pc(i), p = p/10, i = i/10

    end while

  end for

  i=0.1, p=0.5

end while

Calculate throughput and delay using p_b and p_{ci}

4.3.3 Summary of Results

To conclude, we can summarize our first contribution regarding the mastering of the quality of service in 802.11e wireless packet networks as follows. In this first contribution, we proposed a four-dimensional Markov chain model for EDCA, as described in the final standard [103]. Our model allows us to compute the available throughput and mean access delay within each AC whatever are the traffic conditions, ranging from extremely non-saturated to extremely saturated network. The model proposed is the only one in the literature that models all the features of EDCA. Indeed, our work started by the fact that there is a compelling need to have an analytical model that provides the most accurate prediction of the performance metrics, such as throughputs and delays, in order to be used in an efficient admission control algorithm. Our literature review showed that all the currently proposed 802.11e EDCA models, even though numerous, do not reach this accuracy. This is due to the big number of simplifications that they contain. To fill this gap, we developed a new and complete analytical model that: *(i)* reflects the correct functioning of EDCA, *(ii)* includes all the 802.11e EDCA differentiation parameters, *(iii)* takes into account all the features of the protocol, and *(iv)* can be applied to all network conditions, ranging from non-saturation to saturation conditions. Additionally, this model is developed in order to be used in an admission control procedure, so it was designed to have a low complexity and an acceptable response time.

The proposed model is validated by means of both calculations and extensive simulations. To decide about the accuracy of our model, we compared the numerical results to simulation results (obtained using NS-2 [63] in which the TKN's EDCA implementation [104] is added). The performed comparison showed a nearly-exact match between the analytical and the simulation curves proving that our model gives a very good estimation of the performance metrics for any of the configurations of EDCA parameters. When analyzing finely the accuracy of our model, we found that our model is extremely accurate when the system is in the non-saturation region. It also gives a very good accuracy when the system is in the saturation region. Within the transition region (i.e. right before reaching the saturation), we noticed an estimation error between the analytical and the simulation results. This is due to two effects: the queue model assumption that we introduced as well as our interpolation approach to compute the available throughput and mean access delay. This error stays however low and does not really affect the global accuracy.

This work was one of the focuses of the Ph.D. dissertation carried out by Nada Chendeb-Taher, University of Evry val d'Essonne, and defended on March 31, 2009 (see papers [P36][P42][P65]). This work is one of the main results of a fructuous cooperation with the Lebanese University at Tripoli, as Nada's Ph.D. thesis had been co-directed by Dr. Bachar El Hassan from the Lebanese University at Tripoli and me as main supervisor. The results of this research also benefited to the ITEA2 HDTVnext project (April 2008 – October 2010) in which we participated.

4.4 Admission Control for Multimedia Applications over 802.11e WLANs

4.4.1 Research Context

In the 802.11e EDCA standard, there is no guaranty in terms of throughput and delay assurance for real time and multimedia services. Before the network gets saturated, there is no QoS problem. The problem arises once the network starts to reach saturation and a high number of flows share the limited channel resources. At each new flow arrival, the existing flows loose a certain degree of their already achieved performance. This is due to the fact that channel resources have to be distributed among different active flows according to their priority. This means that real time and multimedia services which are unable to adapt their flows to this resource limitation cannot be supported correctly. Thus, it is clear that a resource management solution, controlling the number of active flows in a given *Wireless Local Area Network* (*WLAN*) settings is required. More precisely, there is a compelling need for an efficient admission control mechanism capable of maintaining the QoS required by multimedia applications.

Many proposals for an 802.11e EDCA admission control scheme exist in the literature [105–119]. We mainly distinguish those based on measurements, those based on analytical models, and those based on an hybrid approach (i.e. combining measurements and analytical models). However, even though each approach has its benefits, the proposed schemes in the literature suffer from several limitations.

The idea behind the ***Measurement-Based Admission Control*** [105–109] approach is that the QAP (QoS Access Point) measures regularly (i.e. at each measurement interval) the channel conditions (i.e. its occupation). Based on the values obtained, it makes its decision to accept or reject upcoming flows. Such a simple approach has two advantages:

1. No complex numerical computation is needed before making a decision. Simple computation of additional loads that can be generated by the activation of the new flow in the network is used. These additional loads are then compared to the measured residual capacity of the channel.

2. Measurements of current channel conditions are more precise than when these are obtained using numerical estimations.

However, the Measurement-Based Admission Control approach suffers from multiple limitations which can be summarized by:

1. The QAP cannot consider the real QoS requirements in terms of throughput and delay as a decision criteria. In fact, the measurements cannot be mapped to QoS metrics and there are no

means to verify that the QoS guarantees, of the currently running flows and the newly accepted flow, will be fulfilled.

2. The measurements can only give the channel utilization conditions. They cannot give the achievable values of throughput and access delay.

3. It is very hard to choose the measurement interval value. This value must be sufficiently high to reflect the steady state functioning regime, and at the same time, it must be sufficiently low to reflect any change in the channel conditions. This tradeoff is difficult to achieve and there is no precise solution explaining what the best measurement interval value is.

4. It is often necessary to use signaling between the QAP and the QoS STAtions (QSTAs) to share measured information.

Conversly, the idea behind ***Analytical-Model-Based Admission Control*** [110–116] approach is that the QAP uses numerical computation to predict the performance metrics if a new incoming flow is accepted. A decision whether to accept or reject this new flow is thus made. The schemes following the Analytical-Model-Based Admission Control have many advantages. Among them, we can cite the following:

1. The decision and the processing are made only within the QAP without the need to gather information from the QSTAs. Hence, there is no need for signaling that increases the control load in a limited resources network.

2. The decision is based on the real QoS requirements of the applications. In fact, the QAP uses the predicted performance metrics if the new flow is admitted. Based on these predictions and on the QoS requirements of the new flow and the already active flows, the QAP decide to reject or admit the new flow.

3. The QoS metrics are predicted analytically without the need to introduce the new flow.

However, using currently existing analytical models, this is obtained either at a price of a high response time when the used analytical model is computationally complex, or at the price of the accuracy when the used analytical model is simplified introducing severe approximations. In this second case, the obtained analytical model does not reflect correctly the 802.11e EDCA protocol behavior.

The third category of admission control schemes [117–119] is the ***Hybrid (Measurement and Model-Based) Admission Control*** approach. In this approach, the QAP achieve at each measurement interval the measurements of some defined variables such as busy channel probability and collision probabilities. Then, it uses the measured values in the analytical model to predict the performance metrics and make the decisions. These mechanisms couple the advantages of the two above mechanisms. They however keep some limitations. One of their major limitations is that they are based on the actual measured values of the channel conditions (busy channel probability and collision probabilities). These values are updated using Exponential Weighted Averaging (EWA) and Exponential Weighted Moving Averaging (EWMA) techniques to predict the performance metrics that may result from the activation of the new flow. This may of course lead to wrong predictions. These mechanisms cannot have the real measurements values without really activating the new flow.

From these, we can say that to map the real QoS requirements of real-time and multimedia applications to decision criterion, a measurement-based admission control cannot be efficient. A hybrid admission control based simultaneously on measurements and analytical model can be used, however, we have to sacrifice a certain level of prediction precision caused by the weight of old measurement values on the decisions. Finally, a model-based admission control can be interesting only if the analytical model is characterized by a good accuracy and by a low numerical computation overhead. We showed in our previous contribution (cf. §4.3), that it is possible to design a complete analytical model for 802.11e EDCA that has these characteristics.

The foreseen efficient admission control scheme for 802.11e WLANs is among our major contributions. Our final objective is to propose an admission control mechanism that is based on our analytical model (cf. §4.3). This later was developed with two main objectives: *(i)* to provide a sufficient degree of accuracy and precision and *(ii)* to have a low computational overhead. It is thus suitable for the usage in an admission control scheme allowing an efficient support of real time and multimedia services such as voice and video. The resulting model-based admission control scheme will thus allow controlling the efficient use of the wireless resource in 802.11e WLANs.

4.4.2 Contributions

As shown in the previous contribution (cf. §4.3), an effort were consecrated to develop an accurate analytical model that predicts on the best the correct values of the achievable performance metrics. Likewise, we tried to simplify and decrease on the best the numerical computation overhead and therefore the response time. This was obtained by the use of a resolution algorithm faster than the classic algorithm. All of these efforts were performed with the aim to get round of the two possible limitations of model-based admission control schemes: the accuracy and the computational complexity.

Here, we propose a model-based admission control algorithm that is located within the QAP. It uses our accurate analytical model to predict the QoS metrics that can be achieved once a new flow is introduced in the WLAN. Based on this prediction and on the QoS constraints of already accepted (active) flows as well as of the new flow, the QAP takes its decision of admitting or rejecting the new flow.

Contrarily to other model-based admission control algorithm that use only the access delay [113, 114], only the achievable throughput [115], or are based on a model assuming saturation conditions only [110–112] in the decision making process, our used analytical model is applicable to all functioning regions: i.e. saturation, non-saturation and transition regions. Also, in our solution, we use the two metrics, the achievable throughput and the access delay, as decision criterion. In fact, although real time services are delay sensitive, and have a strict requirement in terms of access delay, the use of the two performance metrics as decision criterion means that the proposed admission control scheme can be applied at the same time to video applications which have high constraints in terms of throughput assurance, and to voice applications which are sensitive to both criteria.

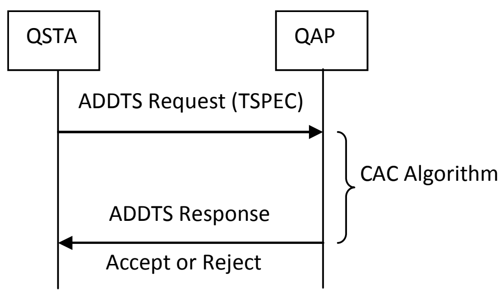


Figure 4.4. Admission control procedure in 802.11e EDCA.

The proposed admission control scheme is fully compatible with the legacy 802.11e EDCA MAC protocol. Figure 4.4 describes the simple signaling exchanges between a QSTA requesting the admission of a new flow and the QAP performing the admission control process and responding by admittance or rejection. So, each QSTA, sends an *ADDTS request (ADD Traffic Stream Request)* for each new flow belonging to a specified AC to the QAP while specifying its QoS requirements (in our case, we are interested to the mean achievable throughput and the mean access delay). If the request is accepted, the QAP responds with a positive *ADDTS response (ADD Traffic Stream Response)* and the flow becomes active and starts the channel contention procedure. Once the transmission is achieved, a *DELETS (Delete*

Traffic Stream) message is sent by the QSTA to the QAP aiming to inform this later that it can delete this flow from the set of active flow.

At the reception of the *ADDTS Request*, the QAP extracts the specifications of the new flow from the *Traffic Specification* (*TSPEC*) field (AC, frame size, mean arrival rate and tolerated delay). Using the analytical model, it predicts the new system probabilities while considering this new flow active. Then, using these probabilities, the QAP compute the achievable throughput and the access delay for the already admitted flows as well as for this new flow. If the QoS requirements can be met, the new flow is admitted and an *ADDTS Response* with the response '*Accept'* is sent to the requesting station. This new flow and its QoS requirements are added to the set of active flows. In the other case, an *ADDTS Response* with the response '*Reject'* is sent to the requesting station.

Algorithm 4.2. Admission Control Algorithm within the QAP

for each *ADDTS_Request from Waiting_Flows* ***do***

F_i = New_Flow (ADDTS_Request)

AC_i = Get Access Category (F_i)

N = Nb_wireless_stations (Admitted_Flows & F_i)

MN= Nb_ACs_per_Station (Admitted_Flows & F_i)

TAR = Total_Arrival_Rate_per_AC (Admitted_Flows & F_i)

Resolve system equations (N, MN, TAR)

for each *Access Category* ***AC*** *in (VO, VI, BE, BK)* ***do***

Calculate Achievable_Throughput (AC)

Calculate Access_Delay (AC)

if *(AC ≠ AC_i)* ***then***

if *(Calculated_Throughput(AC) <* **sum** *(Requested_Throughput(Admitted_Flows(AC))* ***or*** *Calculated_Delay(AC)>* **max** *(Requested_Delay(Admitted_Flows(AC)))* ***then***

Reject (F_i)

Send ADDTS_Response(reject); ***go to ******

end if

else */* This AC is the requested flow's AC, AC_i */*

if *(Calculated_Throughput(AC) -* **sum** *(Requested_Throughput(Admitted_Flows(AC)) < Requested_Throughput(F_i)* ***or*** *Calculated_Delay(AC) >* **max** *(Requested_Delay (Admitted_Flows(AC) & F_i))* ***then***

Reject (F_i)

Send ADDTS_Response(reject); ***go to ******

end if

end if

end for */* end testing all ACs */*

Admit(F_i)

Send ADDTS_Response(accept)

Admitted_Flows = Admitted_Flows & F_i

**** /* go to the next waiting flow */*

end for

Two sets of flows have to be managed by the QAP: 'Waiting_Flows' and 'Admitted_Flows'. 'Waiting_Flows' contains the flows waiting for a response from the QAP. These are the flows that have already sent the *ADDTS Request* but not treated yet by the QAP and therefore did not received the *ADDTS Response*. 'Admitted_Flows' contains the list of already admitted flows that did not send the *DELETS* message yet. It is very importnant that the QAP keeps information concerning the already admitted flows and their specifications. This is required to check if the admittance of any new flow may violate the QoS constraints of the previously admitted flows. Once the *DELETS* message is received by the QAP, it deletes the corresponding flow from the set of 'Admitted_Flows'.

The pseudo code of the proposed admission control algorithm within the QAP is presented in Algorithm 4.2. According to this algorithm, the QAP extracts the AC of the new flow, updates the number of active stations, the number of active access categories in each QSTA and the total arrival rate of each AC while taking into consideration the new flow. These variables are used to resolve the non-linear equation system of the analytical model presented in §4.3.2.d. After the resolution, the transition probabilities of the Markov chain are known, they are used to compute the achievable throughput and access delay per AC by the use of equations (10) and (13) respectively (cf. §4.3.2.b and §4.3.2.c). These calculated values are then compared to the required values of all the flows belonging to the corresponding AC. For the throughput constraint, the sum of the required throughputs for all the flows belonging to a given AC must be less than the achievable throughput for this AC to admit the new flow. For the access delay constraint, the calculated access delay for a given AC must be less than the maximal tolerated access delay of the flows belonging to this AC. It is obvious that these conditions must be verified for the four ACs to allow admitting the new flow.

4.4.3 Summary of Results

To conclude, we can summarize our second contribution regarding the mastering of the quality of service in 802.11e wireless packet networks as follows. In this second contribution, we proposed an admission control scheme based on the analytical model presented in §4.3. Our proposed scheme benefits from the good accuracy and the low numerical computation overhead that characterize our analytical model. Contrarily to existing model-based admission control schemes, our scheme considers all functioning regions of an 802.11e WLAN (i.e. saturation, non-saturation and transition regions) and the two main QoS metrics (i.e. achievable throughput and access delay). These features make our proposed admission control scheme, the more complete and more accurate one in the literature.

In order to validate our proposed admission control scheme, we used both numerical analysis as well as simulation. Concerning the validation scenarios, we were interested to study the behavior of the admission control in different situations: with presence/absence of different application types (voice, video and best effort data services). Our objective was to check if the proposed admission control mechanism is able to maintain and guaranty the QoS requirements for real time and multimedia services such as voice and video. From the obtained results, we can affirm that the proposed admission control mechanism is capable of maintaining and guarantying the QoS for real time and multimedia services. It succeeded to avoid the network from functioning in a severe saturation state in which high performance degradation take place for all application types. Hence, the main objective of our work was achieved. We also noticed that voice applications are very sensitive to delay while video applications are more sensitive to the throughput. Indeed, for voice application the admission of any new voice station near the saturation region (i.e. transition region) increases the delay considerably. The admission control predicts this delay increase and rejects the request. For video applications, it is the bandwidth constraint violation that limits the admissibility of new video flows while the delay remains at an acceptable level. This confirms the interest of using both metrics for admission control which is not the case of other schemes in the literature. Finally, we noticed that the use of the default EDCA parameters as suggested by the 802.11e standard helps to inhibit the impact of best effort data flows on the admission of real time and multimedia services in the WLAN.

This work was one of the focuses of the Ph.D. dissertation carried out by Nada Chendeb-Taher, University of Evry val d'Essonne, and defended on March 31, 2009 (see papers [P34]). This work is also one of the main results of a fructuous cooperation with the Lebanese University at Tripoli as Nada's Ph.D. thesis had been co-directed by Dr. Bachar El Hassan from the Lebanese University at Tripoli and me as main supervisor. The results of this research also benefited to the ITEA2 HDTVnext project (April 2008 – October 2010) to which we participated.

4.5 IMPROVING THE RESOURCE USAGE IN RELAY-ASSISTED BWANS

4.5.1 Research Context

After having studied the QoS issues in 802.11 WLANs and proposed new means to master the wireless resource usage in such networks, let us now go further by studying another kind of networks: Broadband Wireless Access networks. This new kind of wireless packet networks making use of Orthogonal Frequency-Division Multiple Access (OFDMA) cellular system is getting more and more interest in both the research as well as standardization communities. Among the recent advances introduced by such systems, is the use of relay technologies. Indeed, recently relay technologies have been actively studied and considered in the standardization process of next-generation mobile broadband communication systems such as IEEE 802.16j [120] and IEEE 802.16m [121].

The relay-assisted OFDMA cellular system is a promising architecture for future wireless communication systems. The profit of introducing Relay Stations (RS) in a cellular system is the extension of signal and service coverage, the enhancement of the system's overall throughput, and the power saving compared to the sole use of Base Stations (BS). Under such relay-assisted transmissions, two stages are to be performed for cooperative communications: first, a transmission by the Base Station (BS) occurs; then transmissions between Relay Stations (RSs) and Mobile Stations (MSs) are triggered to assist the first one. Hence, the RS can forward high data rates in remote areas while keeping a low cost of infrastructure.

Such architecture offers huge potential for enhancing system capacity, coverage as well as reliability. But this potential gain in capacity and coverage is highly dependent on the scheduling scheme, a topic which draws more and more attention of the research community [122–126]. Some works on scheduling schemes can be found in [124–126], but they only focus on the scheduling in the BS and assume that all the packets transmitted from the BS to RS are immediately relayed to the destination Mobile Station (MS) without the need of scheduling at the RS. However, as the channel states of the links (BS-RS and RS-MS) vary in time independently to each other, if there is asymmetry between the channel states of the BS-RS link and RS-MS link, radio resources might be wasted. This asymmetry also leads to potential queuing in the RS and thus to the necessity of scheduling. In [122], the authors studied the above mentioned problem and proposed the use of opportunistic scheduling in both the BS and the RS. The basic idea is to allocate subchannels to the Mobile Station (MS) experiencing the best channel condition at each frame. However, due to the link asymmetry that may happen in relay networks, the opportunistic scheduling scheme is not the best choice for relay-assisted networks, although it is considered as the most promising scheduling algorithm in the non relay-assisted networks. For example, in a two-hop transmission, even the RS-MS link is good enough to be chosen according to opportunistic scheduling, it might be wasted if there are no buffered packets to be transmitted. Similarly, for the bad RS-MS link, the packets buffered in the RS might be finally dropped after a long period as this bad RS-MS link will not be allocated with subchannels due to the principle of opportunistic scheduling. In this case, the BS-RS link capacity which is used to transmit these dropped packets is also wasted as the packets did not finally reached their destinations. Considering this as well as other situations that may arise while opportunistic scheduling is used at the RS level, we think that this kind of scheduling is not effective in relay-assisted networks and thus not appropriate.

Actually, we strongly believe that if the RS adequately allocates subchannels to those MSs with suboptimal RS-MS links status but who has queued packets in the RS buffer or those MSs with suboptimal BS-RS links but with good RS-MS links, the radio resource can be utilized more effectively. [123] notices this problem, but it assumes the RS-MS links are stable for long period of times (several frames) once their status are detected to be good. This assumption is not reliable and not precise as the channel status change frequently especially when the MS is moving with high speed or the communication environment is complex, such as in presence of massive high buildings. Hence, by using a scheduling scheme based on such assumption, it is still possible to result in radio resource wasting due to link asymmetry. So, in an aim to improve the wireless resource usage, there is a need for a new scheduling algorithm to be used in the RS. This algorithm should allow reducing the amount of wasted subchannel resources due to the BS-RS and RS-MS link asymmetry. We denote the measure of wasted resources due to link asymmetry by "*channel-hole*". More precisely, we define the subchannels allocated to the BS-RS link or RS-MS link but without packets transmission as *channel holes*. Thus, the *channel-hole* has the same unit as subchannels as it actually represents the subchannels that are not fully used in the system. By trying to reduce the number of *channel-holes* (i.e. radio resource waste situation in presence of link asymmetry), such a new scheduling algorithm is expected to be more efficient than the one proposed in [123].

Significant improvement of the system performance is expected when using such channel-hole-based scheduling algorithm instead of opportunistic scheduling in the RS. However, the proposed scheduling algorithm should also take into account the case where multiple RSs are used in the cellular system. Indeed, the impact of link asymmetry on the scheduling is decoupled when multiple-relays are used. For example, let us consider the case where two RSs are used in a cellular system in which the BS-RS links are independent but the RS-MS links end with the same MS. In the BS-RS stage, if the first RS has better link status than the second one, the BS will distribute the packets to the first RS according to the opportunistic scheduling policy. However, in the RS-MS stage, if the link status is the reverse case, (i.e. the second RS has better link status with the concerned MS than the first one) the second RS-MS link will be wasted as the packets are handled by the first RS which stores them in its buffer waiting for better link conditions to transmit them. We call this situation as "*intersect link-asymmetry*" problem. So, unlike the single relay case where packets to different MSs are buffered in the same RS, in the multiple relay case the packets to different MSs might be distributed to different RSs. Such distribution should be made taking into account this "*intersect link-asymmetry*" problem in order to optimize the radio resource usage and avoid resource wastage in the multiple-relay cellular system. The *intersect link-asymmetry* problem seems to be a serious problem in resource management in cellular systems using multiple relays. However, this one is often neglected in the literature [127–131]. Taking into account the *intersect link-asymmetry* problem while designing a channel-hole-based cooperative scheduling algorithm adapted to multiple-relay systems will thus allow minimizing the radio resource waste in such systems.

The foreseen channel-hole-based scheduling and channel-hole-based cooperative scheduling algorithms, for single-relay and multiple relay systems respectively, are among our major contributions. Our final objective is to propose novel scheduling schemes that minimize the amount of subchannels allocated to the BS-RS links or RS-MS links but without packets transmissions (i.e. *channel-holes*). This will be first performed for the case of cellular systems using a single RS. Then, the case where multiple RSs are used and the presence of the *intersect link-asymmetry* problem will be considered. Such scheduling algorithms aim to improve the wireless resource usage in OFDMA-based Wireless Broadband Access Networks making use of relaying technologies.

4.5.2 Contributions

As explained above, the *channel-hole* problem has to be treated differently in a single-relay system and a multiple relay system. Indeed, in this second case, an additional issue needs to be taken into account: namely, the *intersect link-asymmetry* problem. So, in the following we propose two scheduling algorithms targeting the single-relay and multiple-relays cases, respectively.

4.5.2.a Channel-hole scheduling in single-relay systems

Here, we consider a single cell downlink transmission in the manner of OFDMA, with one BS and one RS in the cell as depicted in Figure 4.5. The frame structure is illustrated in Fig 4.6. In the frequency domain, the downlink channel is divided into N orthogonal subchannels each of which is constituted of a group of adjacent subcarriers. In the time domain, each frame is divided into BS subframe and RS subframe. In the BS subframe only the BS can transmit data to the RS and MSs, while in the RS subframe the BS and the RS can transmit data to MSs. Hence, the RS can receive data only in the BS subframe and transmit data only in the RS subframe, while the BS can transmit data in both subframes. It is also assumed that the RS operate in decode-and-forward mode and cannot transmit and receive data simultaneously.

Moreover, we assume the wireless channel is time-varying and frequency-selective but it is assumed to be flat within a subchannel and to be unchanged during a frame period. So, Adaptive Modulation and Coding (AMC) is performed on per frame basis, and MSs feedback to the BS or to the RS their Channel State Information (CSI) containing their per-subchannel Signal-to-Noise-Ratio (SNR) in each frame.

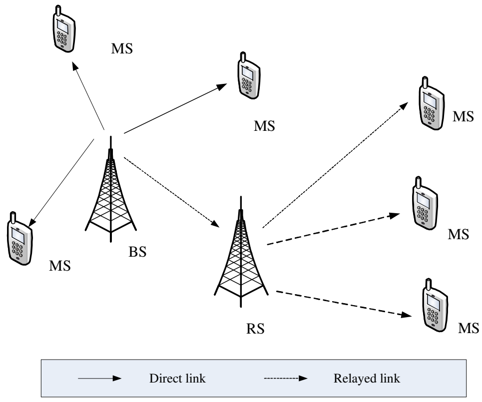


Figure 4.5. Single-relay system.

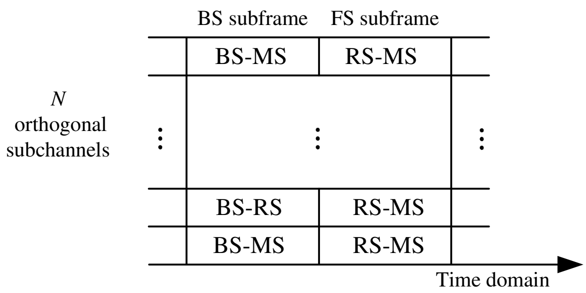


Figure 4.6. Frame structure for the single-relay system.

We define the following sets:

$\mathrm{K} = \{k_0, k_1\}$: set of indicies for BS and RS

$\mathrm{M} = \{1, 2, \cdots, M\}$: set of indicies for MSs

$\mathrm{N} = \{1, 2, \cdots, N\}$: set of indicies for subchunnels

$\Gamma = \{1, 2\}$: set of indicies for subframes

For K, "k_0" represents the BS and "k_1" represents the RS. For Γ, "1" represents the BS subframe and "2" represents the RS subframe. During a frame, all the data are delivered to the destination (RS or MS) through a logical link (*s*,*d*), where $s \in \mathrm{K}$ and $d \in k_1 \cup \mathrm{M}$. Different combinations of *s* and *d* constitute the aggregate X, i.e. (*s*,*d*) ϵ X.

Since each subchannel is assigned to each subframe, there is no intracell interference. During the *t*-th frame, we define the subchannel assignment indicator:

$$d_t(l,n,\tau)=\begin{cases}1, \text{ if subchannel } n \text{ is assigned to the } l-\text{th path at subframe } \tau\\ 0, \text{ otherwise}\end{cases} \quad (1)$$

where $n \in \mathrm{N}$, $l \in \mathrm{X}$ and $\tau \in \Gamma$.

We suppose the power is equally distributed on each subchannel, which gives:

$$P_\tau(n)=\frac{P_{tot}}{N} \quad (2)$$

where P_{tot} is the total transmission power of the system.

The noise in the system is defined as:

$$\sigma^2 = N_0\frac{B}{N} \quad (3)$$

where B is the bandwidth and N_0 is the power spectrum density of the noise.

We also define the subchannel gain of the l-th link on the subchannel n as $h_t(l,n,\tau)$, where $n \in \mathrm{N}$, $l \in \mathrm{X}$ and $\tau \in \Gamma$. So, during the t-th frame, the instantaneous capacity of the l-th link on the subchannel n can be written as:

$$c_t(l,n,\tau)=\frac{B}{N}\log_2\left(1+\frac{P_\tau(n)\left|h_t(l,n,\tau)\right|^2}{\sigma^2}\right) \quad (4)$$

Hence, during the t-th frame, the capacity of the l-th link can be calculated as:

$$R_t(l)=\sum_{n=1}^{N}\sum_{\tau=1}^{2}d_t(l,n,\tau)c_t(l,n,\tau) \quad (5)$$

Based on the above analysis, we propose an improved scheduling algorithm which is called channel-hole based scheduling. We assume here the existence of a fixed or dynamic path selection that decides whether the relayed link or the direct link should be selected for a transmission. Once the path is selected for each MS, the selected links are kept until the long-term average SNR of the MS change significantly. The path selection is performed periodically by assessing the long-term channel statistics.

The scheduling scheme we propose is described in Algorithm 4.3. In this algorithm, we first estimate the instantaneous link capacities and sort them in descending order. We assume here that the channel rate is flat within a frame and unchanged during it. The sorting list contains the BS-MS links (i.e. for those links for which we selected the direct link transmission), BS-RS links and RS-MS links. As we mainly consider the scheduling scheme in the RS, the scheduling in the BS can implement any. Here we use the opportunistic scheduling in the BS subframe, i.e. the BS always chooses the best links for its transmissions, which includes BS-MS direct transmissions and BS-RS relayed transmission. Each subchannel is assigned to only one link in each subframe. Actually, the allocations are also performed according to the path selection which determines the number of direct links and relayed links in each frame. Our focus is to improve the performance of relay links. Thus here we simply first satisfy the direct links and then use the left subchannels to support the relay transmissions.

Algorithm 4.3. Channel-hole based scheduling algorithm for single-relay systems

```
t=1
while t>0 do
    Estimate the instantaneous capacity of the l-th link on the subchannel n according to the
    equation (5) based on CSI and sort them in descending order ;
    if BS subframe then
        at BS
        Choose the best N BS-RS and BS-MS links according to the sorting list;
        For i=1 to N do
            if i is a direct link then
                transmit the packets;
            Else
                if i is a relayed link then
                    label and buffer the packet according to the destination MS_i;
                end if
            end if
        end for
    end if
    if RS subframe then
        at RS
        Temporarily allocate subchannels to the best y RS-MS links;
        For i=1 to M do
            if exist RS_que(MS_i) > 0  then
                Flag_channel-hole = 1;
            end if
        end for
        if Flag_channel-hole = 1 then
            for i=1 to M do
                if RS_que(MS_i)=0 and the i-th RS-MS link is allocated with subchannel k then
                    Re-allocate subchannel to a sub-optimal RS-MS link;
                end if
            end for
        end if
        Forward the packets;
        at BS
        Transmit the packets using the (N-y) subchannels;
    end if
    t++;
end while
```

In the BS subframe, the packets sent from the BS and buffered in the RS will constitute the RS queue and they are classified and accumulated according to the destination MS, that is *RS_que*(MS_i), where $j = 1, 2, \cdots, M$. Then, in the RS subframe, the direct link is given the higher priority and will be allocated (*N-y*) subchannels while the remaining *y* subchannels are used for the RS-MS links. The RS starts by temporarily allocating the subchannel to the top *y* RS-MS links that have good link statuses. Then, the RS check whether this allocation leads to channel-hole problems. The principle is if the *i*-th RS-MS link is allocated with subchannels by the RS and the *i*-th MS queue length is zero in the meanwhile, there will be a channel-hole. It is notable that if all the $RS\_que(MS_i)$ equals to zero, it means no packets in the RS queue. In this case, even it belongs to the channel-hole concept according to the definition, it is actually not because there are no packets in the buffer and no more subchannel become needed. We use $Flag_{channel\text{-}hole} = 1$ to denote the existence of a channel hole. When $Flag_{channel\text{-}hole} = 1$, it means that a channel-hole problem is experienced and that re-allocation is needed. When $Flag_{channel\text{-}hole}$ equals to 1, the RS choose sub-optimal RS-MS links to use the channel-hole resources. The RS check the sorting list in the (*M-y*) remaining links and find the links whose RS queue length is bigger than zero, which means the links that have packets in the buffer and needs subchannel resources to forward them. Our scheduling algorithm will thus re-allocate the subchannel to them. One should finally note that when no channel-hole is experienced ($Flag_{channel\text{-}hole} = 0$) after the first allocation to the RS-MS links, the subchannel surplus is given to the direct BS-MS links.

4.5.2.b Channel-hole cooperative scheduling in multiple-relay systems

After having defined the channel-hole scheduling scheme for the single relay case let us now consider the multiple relay case. In this case, we consider a single-cell downlink OFDMA system with one BS, *R* RSs and *M* MSs, as illustrated in the left part of Figure 4.7. For the MSs under the coverage area of the BS, this latter is able to communicate with them directly. For the MSs that are out of the BS range, the RSs at the edge of the BS's range cooperate to realize the communication between the BS and these MSs. This way of operation is also called non-transparent relay mode [126] in IEEE 802.16j networks. Similarly to the single-relay case, during each frame of relaying, two transmission stages are to be performed: first a BS transmission occurs in a BS-RS stage then RSs transmissions are triggered in an RS-MS stage.

In a relay-assisted OFDMA system, the entire frequency band is divided into a number of subchannels and each subchannel can be flexibly allocated to the BS-RS links or the RS-MS links based on some strategies. These strategies are not our focus in this work. Here we assume that N_1 subchannels are used for the BS-RS links and N_2 subchannels are used for the RS-MS links. We also assume that each subchannel is allowed to be used only once during one relay transmission, which includes one BS-RS and one RS-MS transmission. So, we assume that we have a total of *N* subchannels available with $N = N_1 + N_2$.

In the right part of Figure 4.7, we describe an identical queue model for the relay system that will allow facilitating the description of our scheduling algorithm for multiple-relay systems. Actually, the packets in the BS can be grouped according to their destination. So, we assume there are *M* "virtual queues" in the BS to buffer these grouped packets. These virtual queues are denoted vq_1, …, vq_i, …, vq_M where $i = 1, …, M$,. Thus, there are $M * R$ potential BS-RS links to be allocated with subchannels in the BS-RS stage. Then, during the RS-MS stage, the *R* RSs, denoted RS_1, …, RS_r, …, RS_R where $r = 1, …, R$, cooperate to aid transmitting packets to the corresponding MSs when these are out of the BS transmission range. This collaboration takes the form of a distributed scheduling with which the $M * R$ potential RS-MS links have the opportunity to obtain subchannels during the RS-MS stage.

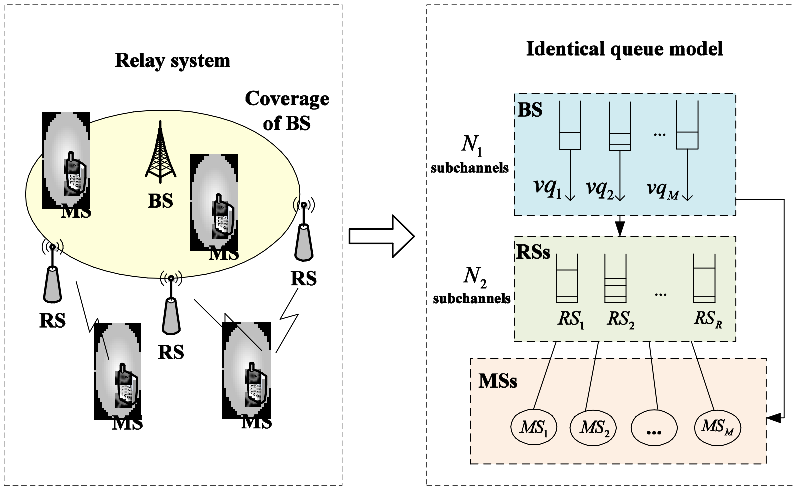


Figure 4.7. The considered relay system and its identical queue model.

As explained earlier, in the case where two or more RSs allow communicating with the same MS, it may happen that an "*intersect link-asymmetry*" problem appears. In the case of two RSs allowing serving the same MS, this problem is defined as follow: the first BS-RS link has a better channel quality than the second BS-RS link, and at the same time the respective RS-MS links have the reverse quality (i.e. the first RS-MS link has bad channel conditions while the second one is of good quality). In this case, if we use opportunistic scheduling at the BS by taking into account only the BS-RS links qualities (i.e. without taking into account the RS-MS links qualities), the first BS-RS link is selected for data packet transmission. This decision leads to wasting the second RS-MS link resources while the data packets are handled by the first RS. Indeed, as it has good channel conditions, the second RS-MS link will be allocated subchannels by the opportunistic scheduling while the first one is obliged to store the data packet in its buffer waiting for better link conditions to transmit them. In some cases, such situations may lead to resource wastage. In order to overcome this, we first need to precisely measure the potential resource wastage. So, here also we use the "*channel hole*" measure, i.e. subchannels allocated to the BS-RS links or RS-MS links but without packet transmissions, as the basis to propose an effective scheduling algorithm in multiple-relay systems. This later is required to reduce the number of "*channel-holes*" that may be due to the "*intersect asymmetry link*" problem. Thus, based on these two concepts, namely the "*channel hole*" measure and the "*intersect link-asymmetry*" problem, we propose in this section, our channel-hole-based cooperative scheduling algorithm as outlined in the Algorithm 4.4.

In the Algorithm 4.4, $Queue_i$ is the queue length of the virtual queue which the link correspond to. $Queue_i = 0$ means that a channel hole possibly exists for the i-th link. This means that if this link is finally allocated with subchannels, a channel hole definitely exist leading to resource wastage. Our proposed algorithm aims to avoid the above case to happen. In the algorithm, the BS and the RSs choose the best N_1 or N_2 links for the BS-RS stage and RS-MS stage, respectively, in order to allocate subchannels to them. If the selected links has no queued packets in their virtual queue, the subchannels are then allocated to the second-best links; if the second-best links also has no queued packets to forward, the subchannels will be allocated to the third-best links. This process will continue until all the links are visited. In step 4, the modified queue status will be used as the judgment for the potential channel holes in the next frame.

It is notable that N_1 and N_2 are probably different in different frames as we can have different strategies of subcarrier combination to form subchannels. Such strategies are not in our focus interest in this work. The BS can perform the scheduling according to the N_1 and N_2 values and then transmit the command information in the packet headers at the beginning of each frame.

Algorithm 4.4. Channel-hole based scheduling algorithm for multiple-relay systems

Step 1: *At the beginning of each time-frame t, the BS updates the instantaneous capacity estimation for all links involved in the BS-RS and RS-MS stages using CSI (i.e. according to the equation (5)); then sorts them in descending order.*

Step 2: *the BS chooses the best N_1 links for the BS-RS stage and the best N_2 links for the RS-MS stage as the candidate for subchannels allocation, then execute the following program :*

for *k = 1 to 2* ***do***

  $j = N_k$;

  for *$i = 1$ to N_k* ***do***

    while *$Queue_j = 0$* ***do***

      if *$Queue_i = 0$* ***then***

        j++;

        Allocate i-th subchannel to j-th link;

      else

        Allocate i-th subchannel to i-th link;

      end if

    end while

  end for

end for

Step 3: *Use the resulting subchannel allocation to transmit data in both the BS-RS and RS-MS stages.*

Step 4: *After the transmission within one frame, the BS and RS modify the record of their virtual queue status (i.e. empty vs. not empty).*

4.5.3 Summary of Results

To conclude, we can summarize our third contribution regarding the mastering of the quality of service in relay-assisted broadband wireless access networks as follows. In this third contribution, we first discussed the special features of relay-assisted networks, and disclosed the shortcomings of opportunistic scheduling when it is used in relay stations. According to this, we defined a new concept, the "*channel-hole*", which is a measure of the wasted subchannel resources caused by the asymmetry between the BS-RS and RS-MS links. Based on this measure, we proposed two novel scheduling schemes for single-relay systems and multiple-relay systems, respectively. In the single relay case, our new scheduling scheme to be used in the RS, which is called channel-hole based scheduling, is designed with the aim to minimize the wireless resource wastage due to channel-holes. As it is needed by our scheduling algorithm, we also derived analytically the link capacity expression. In the multiple relay case, the proposed channel-hole based cooperative scheduling have as objective to minimize the channel-holes due to the intersect link-asymmetry problem. Indeed, this later seems to be one of the most serious problems in resource management in cellular systems using multiple relays. However, we noticed that this problem is often neglected in the literature. These two contributions are a first step towards improving the wireless resource usage in relay-assisted broadband wireless access networks such as IEEE 802.16j or IEEE 802.16m.

In order to validate our two proposed scheduling schemes, we used both numerical analysis as well as simulations. In the single-relay case we compared our channel-hole scheduling to opportunistic scheduling. Simulations showed that by allowing the proposed scheduling at the RS, the amount of wasted subchannels is significantly lower to those wasted while adapting the opportunistic scheduling at the RS. This leads to an important improvement of the system performances in terms of resource utilization and buffer occupancy. In the multiple-relay case, we built a two-dimensional Markov chain in order to model the system, analyze the average queue length and the average packet delay when our channel-hole cooperative scheduling algorithm is used. Simulations were used to validate and reinforce our theoretical analysis showing the behavior of our scheduling algorithm when the packet arrival rate increases. Also, using simulations we compared our proposed scheduling algorithm to other algorithms from the literature. Although it is considered as the most promising scheduling technology for next generation networks in many literatures, our comparison study showed that our proposed algorithm performs better in terms of average queue length and average packet delay than opportunistic scheduling. In addition, we observed that our proposed scheduling algorithm also outperforms a resource management solution that is only based on a relay selection scheme even if this one is ideal.

This work was one of the focuses of the Post-doctoral fellowship carried out by Yan Li, under my supervision at ENSIIE, from October 2009 to September 2010. This research is still in-progress and several papers describing our first results are currently under review. This work is also one of the main results we obtained in the frame of the ITEA2 HDTVnext project (April 2008 – October 2010).

4.6 CONCLUSION

The work presented here was realized in the frame of one Ph.D. dissertation, Nada Chendeb-Taher (September 2005 – March 2009), and one post-doctroral fellowship, Yan Li (October 2009 – September 2010). I was the main supervisor for both, while Nada's Ph.D. thesis was also co-supervised by Dr. Bachar El Hassan from the Lebanese University at Tripoli. The work presented here was also undertaken in the frame of the ITEA2 HDTVNext project (April 2008 – October 2010), project to which we had an active participation.

The obtained results are numerous. We started from the fact that there is a need for a complete analytical model that highlights the exact behavior of the 802.11e EDCA MAC protocol. Indeed, our study of the existing analytical models for 802.11e EDCA MAC protocol clearly showed that these suffer from many limitations making them incomplete and inaccurate. Hence, such existing models used in admission control algorithms will necessarily lead to either non-optimal or wrong decisions. Thus, to fill in this gap, we developed a new and complete analytical model that: *(i)* provide a sufficient degree of accuracy and precision, and *(ii)* can be applied to all network conditions, going from non-saturation to saturation conditions. Additionally, as this model was developed in order to be used for admission control purposes, it was designed to have a low complexity and an acceptable response time. Then, after having validated the proposed analytical model, we used it in order to build an efficient model-based admission control algorithm for 802.11e WLANs. Indeed, our literature review of existing admission control algorithms for 802.11e WLANs showed that developing such an admission control algorithm based on an analytical model that is characterized by a good accuracy and by a low numerical computation overhead can lead to effective results. Our validation campaign allowed confirming this expectation.

Alternatively, while studying resource management in Broadband Wireless Access Networks, we identified one of the performance limitations in such networks, namely scheduling in relay nodes. Indeed, in Relay-assisted BWANs, using a single-relay or multiple-relays per cell, it is often considered that there is no need for scheduling in relay stations and that the packets arriving at the relay station are forwarded immediately towards their destination. This immediate forwarding is not always possible, especially when the BS-RS and RS-MS links are asymmetric. Link asymmetry may have several impacts in Relay-assisted BWANs leading to important resource wastage if not taken into account by an adequate scheduling to be implemented in the relay stations. Thanks to a measure of these wasted resources, denoted as channel-holes, we were able to design a new set of scheduling schemes that are suitable for single-relay and multiple relay systems, respectively. These schemes, which had been designed with the aim to reduce the number of channel-holes, proved to be more efficient in terms of wireless resource usage improvement compared to other existing schemes.

To conclude we can say that resource management can be seen in different manners depending on the wireless packet network technology in which it is implemented. In 802.11e WLANs the identified missing resource management component was the efficient admission control algorithm while in Relay-assisted BWANs this one was the adequate scheduling in Relay Stations. We showed that using our proposed resource management solutions, we are able to improve the wireless resource usage in both studied wireless packet network contexts.

Chapter 5. Vehicular Packet Networks: Geo-localized Communications in Urban Areas

Inter-Vehicle Communication (IVC) is attracting a considerable attention from the research community and the automotive industry, where it is beneficial in providing Intelligent Transportation Systems (ITS) as well as assistance services for drivers and passengers. In this context, Vehicular Packet Networks are emerging as a novel category of wireless networks, spontaneously formed among moving vehicles and road infrastructure both equipped with wireless interfaces. The distinguished characteristics of vehicular packet networks (such as high mobility, potentially large scale, and network partitioning) introduce several challenges, which can greatly impact the future deployment of these networks.

In the following, we focus on inter-vehicle communication in urban environments. Our main goal is to propose new routing and dissemination algorithms, which efficiently adapts to vehicular packet networks characteristics and applications in order to improve the wireless resource usage in such environments. We will first remind the context of our research as well as the challenges that faces vehicular packet networks in urban environments. This will be followed by the description of our three contributions and the results we obtained in this domain. These contributions concern road-traffic density estimation using Vehicular Ad hoc Networks (VANETs), geo-localized traffic-aware routing and data dissemination based on a geo-localized virtual infrastructure. This chapter ends with a brief conclusion summarizing our research experience in this domain.

5.1 Research Context

Road safety has been an important concern in the world over the past few years since millions of people die every year because of car accidents and many more are injured. Current statistics show that road traffic accidents in the Member States of the European Union annually claim about 39000 lives and leave more than 1.7 million people injured, representing an estimated cost of 160 billion €uros [132].

Automated highway systems and Intelligent Transportation Systems (ITS) were introduced to accelerate the development and use of intelligent integrated safety systems that use information and communication technologies as an intelligent solution in order to increase road safety and reduce the number of accidents in our future roads. In contrast, as mobile wireless devices became an essential part of our lives, and the ubiquitous ‘anywhere, anytime’ connectivity concept is gaining attraction, access to dedicated comfort services from vehicles is in great demand. The proliferation of cooperated system approach for ITS and the focus on Information and Communications Technologies (ICT) services on the one hand as well as the growing number of communication infrastructure-enabled vehicles on the other hand has opened up new business models and key market segments for many stakeholders in the ITS-market.

Vehicular Packet Networks are a cornerstone of the envisioned Intelligent Transportation Systems (ITS). By enabling vehicles to communicate with each other via Inter-Vehicle Communication (IVC) as well as with roadside base stations via Roadside-to-Vehicle Communication (RVC), vehicular packet networks could contribute to safer, more comfortable and highly efficient roads. The opportunities and areas of applications of vehicular packet networks are growing rapidly, with many vehicle manufacturers and private institutes actively supporting research and development in this field. The integration with on-board sensor systems and the progressive diffusion of on-board localization systems (i.e. Global Positioning System or GPS) make vehicular packet networks suitable for the development of active safety applications, including collision and warning systems, driver assistance and intelligent traffic management systems. On the other hand, inter-vehicular communication (IVC) also fuels the vast opportunities in online vehicle entertainment (such as gaming or file sharing), and enables the integration with Internet services and applications [133].

In the near future, it is expected that several of the above mentioned ITS applications and systems will be deployed in city environment, both for car-to-car communication services and value-added infrastructure-based ITS services. To guarantee efficiency to different applications, several important issues have to be tackled. Among the most critical one we can quote efficient data dissemination and routing. Indeed, most of the ITS applications and systems rely for data delivery using multi-hop communications, forming what we call a Vehicular Ad hoc Network (VANET). Efficient data dissemination and routing in VANET are highly challenging tasks due to the inherent characteristics of vehicular networks [134], which are highly mobile and disconnected in nature. Indeed, the network density pattern is governed by time and location and is highly unpredictable especially in city environments. This may greatly affects the performances of multi-hop communications in VANETs.

The target of our research, presented in this chapter, is to propose an efficient routing strategy as well as an intelligent information dissemination approach to be used in a city-scale. The final objective of this research is to show that it is possible to make use efficiently of road-traffic density information in order to improve the performances of vehicular packet networks.

5.2 MAIN CHALLENGES

With the advances in wireless communication technologies, the concept of networked-car has received immense attention all over the world. This increasing importance has been recognized by major car manufacturers, governmental organizations, and academic community. In USA, the Federal Communications Commission (FCC) has allocated 75 MHz of spectrum at 5.9 GHz for Dedicated Short-Range Communications (DSRC) [135]. In recent years, several research initiatives including VII [136], NoW [137][137], CVIS [138][138], COMeSafety [139], C2CCC [140] among others, are being investigated, targeting to accomplish the dream of networked-car and successful implementation of vehicular networks as a major step towards the realization of Intelligent Transportation Systems (ITS). As a result, an increasing number of car manufacturers are equipping vehicles with on-board computing and wireless communication devices as well as in-car sensors and GPS systems. This realized in anticipation of the deployment of large-scale vehicular networks.

Along with the recent developments in vehicular communications, a number of attractive applications, unique for the vehicular setting, have emerged [141]. This spectrum ranges from active safety or safety of life applications to infotainment services. To achieve this, Vehicular Ad hoc Networks (VANETs) have received considerable attention in recent times. For example, a VANET can be used for issuing driver alerts during specific events like potential traffic jams, hazardous road conditions (slippery road warning) or accidents (to avoid multi-car collisions). Apart from road safety applications, VANETs are also useful for several other communication instances in the ITS scenarios. Among these comfort applications, we may quote i) info-mobility (weather information, gas station or restaurants location, city leisure information, movie trailer downloads, tourist information, etc.), ii) mobile e-commerce (advertisements or announcements of sales information), iii) infotainment and interactive services (internet access, distributed games, chats, music downloads, etc.). Thus, multi-hop data delivery between vehicles, both unicast and geo-cast, is an important aspect for the support of such VANET-based applications.

Although, data dissemination and routing have been extensively addressed in wireless packet networks, many unique characteristics of VANETs together with the diversity in promising applications offer newer research challenges. Indeed, in comparison to other wireless packet networks, VANETs come with unique attractive features [142]:

- *Unlimited transmission power:* mobile device power issues are usually not a significant constraint in vehicular networks in comparison to the case of Mobile Ad hoc Networks (MANETs) or Wireless Sensor Networks (WSN), since the node itself can provide continuous power for computing and communication.

- *Higher computational capability*: operating vehicles can afford significant computing, communication and sensing capabilities.

- *Predictable mobility:* unlike general MANETs, where it is hard to predict the nodes' mobility patterns, vehicles can have very predictable movement that is (usually) limited to roadways. Roadway information is often available from positioning systems and map-based technologies such as GPS. Given the average speed, current speed, and road trajectory, the future position of a vehicle can thus be (easily) predicted.

However, to bring its potency to fruition, VANETs have also to cope with formidable challenges [143] that include:

- *Potentially large scale:* unlike most of mobile ad hoc networks studied in the literature that usually assume a limited area, VANETs can in principle extend over the entire road network and include a large number of participants.

- *High mobility:* The environment in which vehicular networks operate is extremely dynamic, and includes extreme configurations: in highways, relative speed of up to 300 km/h may occur, while density of nodes may be 1 or 2 vehicles per kilometer in lightly loaded roads.

Partitioned network: VANETs are frequently partitioned networks. Indeed, the dynamic nature of road traffic may result in large inter-vehicle gaps in sparsely populated scenarios, and in several isolated clusters of nodes.

- *Network topology and connectivity:* the topology is very different from that in a mobile ad hoc network as vehicles are moving and changing their position constantly and very rapidly, making the topology and the connectivity very dynamic. The network topology changes frequently as the links between nodes connect and disconnect very often [144]. The degree to which the network is connected is highly dependent on two factors: the range of wireless links and the fraction of participating vehicles, since only a fraction of vehicles on the road could be equipped with wireless interfaces.

One should also note that these challenges are even decoupled in city environments. Indeed, in such environments the road network is more complex, the vehicles have often pause times during their journey (e.g. traffic-lights), and they select their destination randomly.

Taking into account these challenges as well as city-environment characteristics, temporary disconnection in VANETs seems to be clearly unavoidable. It is thereby of imminent practical interest to consider the vehicular traffic density in any newly proposed multi-hop communication protocol in order to maximize the benefits of such protocol. Thus, in the following, we first propose a completely distributed and infrastructure-free mechanism for city road density estimation. Then, and based on such traffic information system, we propose a novel intersection-based geographical routing protocol, capable to find robust and optimal routes within urban environments. Finally, in order to help the efficient support of dissemination-based applications, a self-organizing mechanism to emulate a geo-localized virtual infrastructure is proposed. This one is intended to be deployed in the intersections having an acceptable level of vehicular density.

5.3 TRAFFIC DENSITY ESTIMATION USING VANETS

5.3.1 Research Context

Vehicular networks are the major ingredient of the envisioned Intelligent Transportation Systems (ITS) concept. Road traffic information processing is another important component of ITS which is currently attracting wider research focus. This has widespread applications in the ITS context but not only. Indeed, as explained earlier we strongly believe that multi-hop communications protocols can benefit from the vehicular traffic density information in order to improve their performances. We strongly believe that, in addition to being useful for real-time traffic congestion warning, efficient traffic density information estimation can also be used as a critical metric in a VANET routing protocol for determining vehicular data routing paths.

Traditional ITS traffic information systems are based on a centralized structure. This centralized structure is focused around a traffic management centre that collects data from the street network, via sensing devices, and processes them. The resulting traffic information is made available to the drivers via broadcast service or alternatively on demand via cellular phones. The centralized approaches are dependent on fixed infrastructure which demands public investments from government agencies or other relevant operators to build, maintain and manage such infrastructure: a large number of sensors are needed to be deployed in order to monitor the traffic situation. The traffic information service is then limited to streets where sensors are integrated. Besides, centralized designs have the disadvantage of being rigid, difficult to maintain and upgrade, require substantial computing/communications capabilities, and are susceptible to catastrophic events (sabotage or system failures). Moreover, such systems are characterized by long reaction times and thus are not useable by the applications requiring reliable decision making based on accurate and prompt road traffic awareness.

To overcome the disadvantages of centralized schemes, initiatives towards decentralizing traffic information systems began to appear. In the work described in [145], the authors propose a simple time-dependent solution for road congestion estimation. The solution is based on an opposite stream vehicle communication approach where each vehicle exchanges information about the average entry-exit times (travel times) of segments with their neighbors in the opposite streams. This updated traversal times, which are assumed to be stored at the street segments, could help the driver as descriptive network information to re-evaluate current routes and to possibly produce new routes. To avoid the redundancy, it is strictly considered that only vehicles traversing in opposite directions will exchange information (the vehicles closely moving in the same directions will maintain similar information). A second viable decentralized approach named SOTIS, which stands for Self-Organizing Traffic Information System, was proposed in [146]. The traffic condition is evaluated in terms of average velocity for traffic segments. Then, the vehicles (SOTIS equipped) periodically exchange and update their current position and traffic information (average velocity), in order to inform the surrounding vehicles about their individual knowledge. As another car receives this information, it is compared with a ‘knowledge base’ to establish whether the information is more accurate/updated than the existing information. Entries in the ‘knowledge base’ are associated with geographic coordinates and are combined with a digital ‘local map’. They can be displayed as warnings/indicators on an in-car display, for instance. Another decentralized solution is proposed as part of the e-Road project [147], which aims to achieve a scalable communication infrastructure for Inter-Vehicle Communications. As part of this project, they propose a system called TrafficView [148], which aims at disseminating and gathering information about vehicles in the road. Using such a system, the vehicle driver will be aware of the road traffic, which helps in driving through foggy weather or for determining optimal route in a long distance trip. The system works based on frequent message broadcast based on the information provided by the GPS receiver (like location, speed, current time and direction of the vehicle). The received neighborhood information is validated to avoid conflicting information and is then moved and merged with existing validated dataset. Periodically, the system displays the data from the validated set. The basic idea behind the information dissemination in TrafficView system is data aggregation.

The decentralized mechanisms for traffic density estimation presented above do not provide a scalable and accurate solution for dynamic traffic density estimation as required by vehicular networking applications. The associated time lag for the update information and inaccuracy of information in case of sparse networks renders the work presented in [145] unsuitable for many vehicular networking applications. Moreover, there is no collaborative processing of information in SOTIS that takes place between vehicles, and there is no attempt to identify abstract traffic events by interpreting the sensor data. The aggregation techniques exploited in TrafficView does not explicitly site the issues related to information exchange and processing. The success of data aggregation methods depends on the frequency of the update messages and could saturate the IEEE 802.11 capabilities, thereby imposing different problems in the system. Moreover, both SOTIS and TrafficView are designed for highways. They are not adapted to city environments, and in particular to downtown areas. From all these, it is clear that a better adapted mechanism for road density estimation is still needed. This one should provide an improved accuracy and promptness of delivery, where scalability is a primary constraint so that it does not saturate the system capacity.

The foreseen road density estimation mechanism is among our major contributions. Our final objective is to propose a completely distributed and infrastructure-free mechanism for vehicular traffic density estimation. The proposed solution has to be adaptive and scalable and targets city environments. Particularly, it should be adapted to both: *(i)* integration to real-time traffic congestion warning systems, and *(ii)* serving as a critical metric for determining optimal vehicular data routing paths in Vehicular Ad Hoc Networks. Indeed, making available such a critical metric is the first step towards improving the performances of the VANET routing process.

5.3.2 Contributions

The accuracy, promptness and scalability are the three constraints that we have identified for traffic density estimation. In order to fulfill these constraints, we propose a new completely distributed and infrastructure-free mechanism for the estimation of traffic density in city-roads. Our proposed mechanism, called IFTIS for "Infrastructure-Free Traffic Information System", is built around a set of realistic assumptions. We assume that each vehicle participating in the cooperative road traffic information processing knows its own geographic position using a Global Position System (GPS). Moreover, we consider that such vehicles are equipped with digital maps to determine which road they are in. Therefore, the available traffic information can be visualized and stored based on road identifiers in the digital map. We also assume that each vehicle is equipped with at least one wireless transceiver. Finally, each vehicle is required to maintain a neighbor table in which position, velocity and direction of each neighboring vehicle are recorded. This table is built and updated thanks to the periodic exchange of *Hello* packets by all vehicles. Additionally, the vehicles also record the identity of the road segment (section of street between two intersections) in which their neighbors are traveling (represented as road ID).

IFTIS is based on the distributed exchange and maintenance of traffic information between vehicles traversing a road segment. More precisely, the vehicles are arranged into location-based groups. For that, each road segment is dissected into small fixed area cells, each defining a group. Note that the cell size depends on the transmission range of vehicles (around 250 m) and the coordinates of the cell center gives the cell a unique identifier (ID). Cells, and hence groups, overlap in such a way that any vehicle moving from one cell to the next belongs at least to one group. Within each cell, a small area around the cell center (cf. the hatched area in Fig. 5.1) is designated to be the leader zone. Among vehicles within this zone, the closest vehicle to the cell center is considered as the group leader for a given duration. Note that the overlapping zone is so small that it is not possible that a vehicle is considered to be group leader of two adjacent cells.

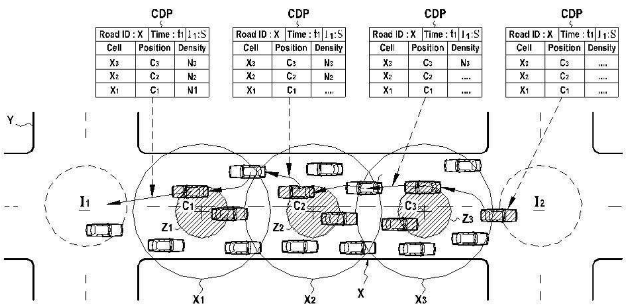


Figure 5.1. Relaying local density information between groups.

Cells Data Packet (CDP)		
Road ID	Time	
Cell ID	Cell 's Center (Position)	Cell's Density

Figure 5.2. CDP message format.

In order to compute the road density information, IFTIS uses Cell Density Packets (CDPs). A CDP is generated at each intersection and is sent back towards the other intersection. As illustrated in Figure 5.2, each CDP contains fields identifying the road ID, transmission time[4], and a list of route anchors (i.e. position of cell centers). When initiating the CDP, a vehicle records the road ID, the transmission time and a list of anchors through which the packet has to pass while traveling to the other intersection, and then, sends the packet in the backward direction (see Figure 5.1). Local density information (i.e. the number of its neighbors which belong to the corresponding road segment[5]) is computed by each group leader at the reception of the CDP packet and relayed towards the destination intersection using this packet. The CDP gathers all local density information of a given road (i.e., number of vehicles within each of its cells).

Algorithm 5.1. Forwarding CDP Packet.

Notation:

CDP*: the CDP packet to forward*

fv*: the forwarding vehicle,*

Ibegin*: intersection at the beginning of the road section*

Iend*: intersection at the end of the road section*

Ni1 *:number of vehicles on cell i moving from* ***Ibegin*** *to* ***Iend***

Ni2 *:number of vehicles on cell i moving from* ***Iend*** *to* ***Ibegin***

Algorithm:

A vehicle ***fv*** *receives the* ***CDP*** *packet:*
if fv *is* **not** *around* ***Ibegin***
then
  if fv *is a* ***group leader*** *// fv is in the red zone of Cell I*
  then
    //Update ***CDP***
    Fill the density of cell i Ni = Ni1 + Ni2;
    NextAnchor = center of cell i+1;
  End if
  //Select next Hop & forward ***CDP***
  fv *selects neighbors (N) moving towards* ***Ibegin***
  if *∃****v*** *∈ N closer to NextAnchor*
  then
    fv *forwards* ***CDP*** *to* ***v***
  else *// fv is the closest vehicle to NextAnchor*
    Store ***CDP*** *and carry it*
  End if
else *//the CDP reaches Ibegin*
  Broadcast ***CDP*** *around* ***Ibegin***
End if

[4] Note that all the vehicles are synchronized by GPS.

[5] This information is already available in the neighbors table of the elected group leader.

In order to forward CDP packets towards the destination intersection through the set of anchor points, IFTIS doesn't require any advanced routing process. A simple greedy strategy can be used for that (i.e. sending the packet to the closest vehicle to the next anchor point among its neighbors). Hence, once a CDP is generated at an intersection, it is forwarded towards the first anchor on the basis of such greedy strategy. Once the message is received by a group leader (i.e. the closest vehicle to the anchor point), this later updates it by including the density of the corresponding cell and then forwards it towards the next anchor. This is repeated until the CDP is completed while arriving to the destination intersection. This is illustrated in Figure 5.1 where the CDP packet is generated by the vehicle leaving road X. This packet is then updated by the vehicles around cells' centers (C_1, C_2, and C_3) as it is traveling from one cell to another.

When the last anchor (i.e. the destination intersection) is reached, the CDP contains the cell-density information of all the traffic groups in the road. Then, it is propagated to vehicles which are around the intersection so that all vehicles traversing through the intersection will receive it. These vehicles analyze the packet content and calculate the density for the respective road segment from which the CDP was received. Algorithm 5.1 illustrates the entire IFTIS algorithm.

Last but not least, one should note that in order to control the IFTIS overhead and make our mechanism adaptive to the traffic evolution, the CDP is generated by previous active group leaders once they reach a road intersection. In other words, only a vehicle which has already updated a CDP message will generate a new CDP before leaving the road. This has two benefits. First, it controls the generation of CDP messages, avoiding scalability and overhead issues. Second, it is adapted to the dynamics of the vehicular traffic within the road segment. Indeed, when the vehicular traffic decreases, the traffic density changes quickly. In this case, the speed of group leader vehicles increases, which in turn increase the frequency of CDP packet generation. Conversely, when the vehicular traffic increases the traffic density changes slowly. In this case, the speed of group leader vehicles decreases, which in turn decreases the frequency of CDP packet generation.

5.3.3 Summary of Results

To conclude, we can summarize our first contribution regarding the improvement of wireless resource usage in vehicular packet networks as follows. In this first contribution, we proposed IFTIS, a completely distributed and infrastructure-free mechanism to determine the vehicular traffic density in city environments. IFTIS is a scalable and adaptive mechanism that makes efficient use of the vehicles traversing the intersections to optimally manage and drive the traffic density estimation process. The estimated road traffic density information is useful for several ITS-related applications. Particularly, the proposed mechanism is suitable for integration to real-time traffic congestion warning systems, leveraging on the proposed distributed mechanism that provides updated traffic information to drivers. It may also be used as a critical metric for determining optimal vehicular data routing paths in Vehicular Ad Hoc Networks (VANETs). Checking the possibility to make such critical metric available was our main motivation in this work.

In order to highlight the benefits of IFTIS, we performed a set of simulations using Qualnet [149] as simulation tool and VanetMobiSim [150] as a realistic vehicular mobility model. We analyzed two performance features of IFTIS: the accuracy to estimate the traffic density by comparing the estimated values to real values, and the promptness of information delivery by analysing the end-to-end delay between the source and destination intersections. The performance analysis of IFTIS depicted a good level of accuracy (i.e. 5% maximum error for a high vehicular density and 8.3% for a low vehicular density) and promptness (i.e. low CDP packet transit delay of few hundreds of milliseconds on average with some rare peaks of few seconds). The analysis, conducted for different density values indicated that IFTIS can scale well enough to adapt to changing traffic conditions. We also showed that while introducing traffic perturbations, such as traffic lights, this has only a limited impact on traffic estimation accuracy.

This work is one of the main results of a fructuous cooperation with Orange Labs in the frame of a CIFRE Convention. It was one of the focuses of the Ph.D. dissertation carried out by Moez Jerbi, University of Evry val d'Essonne, and defended on November 6, 2008 (see papers [P11][P43]).

5.4 GEO-ROUTING IN VANETS

5.4.1 Research Context

As explained earlier, most of the ITS applications and systems rely for data delivery on multi-hop communications, forming what we call a Vehicular Ad hoc Network (VANET). More precisely, most of the distributed infotainment applications and user services (web browsing, chat, file sharing, games, sale information advertisements, available parking lot announcements, etc.) assume the existence of an efficient data routing upon which they are built. This routing protocol ensures the user connectivity in the vehicular environment, allows service continuity and the possible extension to the wired network.

Topology-based and position-based routing are two strategies of data forwarding commonly used for multi-hop wireless networks. Topology-based protocols use the information of available network links for packet transmission and every node needs to build and maintain a routing table. This can be challenging in highly dynamic settings such as vehicular networks. In contrast, position-based protocols assume that every node is aware of its location, the location of neighboring nodes and also the location of the destination node. The increasing availability of GPS-equipped vehicles makes position-based routing a convenient routing strategy for vehicular networks as compared to topology-based approaches. However, the position-based protocols developed for MANET (Mobile Ad hoc Networks) may not be directly applied to vehicular environments owing to the unique vehicular network characteristics [151]. Several variants of position-based routing have been proposed for data forwarding in vehicular networks. These can be grouped into three classes according to the forwarding strategy they use: 1) restricted directional flooding, 2) hierarchical forwarding, and 3) greedy forwarding.

- The first class of protocols, using restricted directional flooding, is more about dissemination than explicit unicast routing. Many broadcast-based protocols based on this forwarding strategy have been proposed so far [152–156]. Among these protocols, MDDV (Mobility centric Data Dissemination algorithm for Vehicular networks) [153][153] exploits geographic forwarding to the destination region, favoring those paths where vehicle density is higher. In MDDV, messages are carried by head vehicles, i.e., best positioned vehicles towards the destination with respect to their neighbors. UMB (Urban Multi-hop Protocol) [152][152] is another restricted directional flooding protocol. Its specificity is that it considers the road topology. Basically, while forwarding messages, each node selects the farthest node using location information, in order to reduce the number of hops. At each intersection, it assumes a special fixed station called repeater to deliver the message to different directions. In [154], the authors suggest the use of a probabilistic propagation function in order to forward messages which assign higher weights to preferable locations inside the intended destination area.

- The second class of position-based routing protocols makes use of hierarchical approaches for data forwarding. This class of forwarding strategies relies on a protocol hierarchy, where different protocol-levels are used in different stages of the forwarding process. The terminodes project [157][157] implemented a hierarchical routing framework incorporating two routing classes, one for nodes in the surrounding area (i.e. terminodes local routing) and one for routing over larger distances (i.e. terminodes remote routing). Depending on the routing hierarchy, the protocol implements two forwarding strategies, namely GPF (Geodesic Packet Forwarding) and AGPF (Anchored Geodesic Packet Forwarding) schemes.

- Unlike the broadcast-based protocols and hierarchical approaches, the third class deals with routing protocols adopting greedy forwarding [158–162]. With greedy forwarding, a node forwards a packet to a neighbor that is located closer to the destination. If this forwarding strategy fails, since there may be situations in which there is no node closer to the destination than the forwarding node, recovery strategies are triggered to deal with this failure.

The third class of position-based routing making use of the greedy forwarding strategy, also called geographical routing approach, seems to be the best adapted to the dynamic nature of large scale ad hoc networks such as VANETs. However, it is not directly applicable to VANETs. Indeed, we note that existing geographic routing like GPSR [163][158] are often based on a simple greedy forwarding concept (closest vehicle to the destination) without taking into account the urban environment characteristics. This may lead to poor signal reception due to radio obstacles such as high-rise buildings.

Fortunately, in vehicular settings, the availability of navigation systems makes it possible to exploit maps and traffic information to guide the diffusion and forwarding of messages. Recent approaches examine this information to “plan” the best route to reach the destination and then use source- or trajectory-based routing [164][164] to forward messages along the desired trajectory. For example, GVGrid [158][158] is a QoS-based VANET routing protocol which exploits geographical information. It divides a geographical area into grids and forwards packets along the roads crossing different grids. However, it assumes a dense network, which does not always hold true in VANETs. The work in [159][159] computes the sequence of intersections that must be traversed by each packet to reach its destination. This information is then included in the packet in the form of geographic source routing. A-STAR (Anchor-based Street and Traffic Aware Routing) [160][160] is also an intersection-based routing scheme designed specifically for VANETs in a city environment. It features the novel use of city bus route information to identify anchor paths of higher connectivity so that more packets can be delivered to their destinations successfully. Another recent example of vehicular routing that exploits the availability of map information is discussed in [161][161]. This routing protocol, aimed at sparsely connected vehicular networks, uses a store and forward technique and approaches the destination by selecting the direction with the lowest estimated delay to the destination. The forwarding algorithm selects the next hop by choosing either the neighbor that is the nearest to the destination (which may lead to routing loops), or a neighbor that is approaching the target location. SADV (Static-node assisted Adaptive data Dissemination protocol for Vehicular networks) [162][162] gives insightful results in terms of data delivery performance under low vehicle density. However, this is obtained with the help of static nodes deployed at intersections. These static nodes are used for storing the packets till there are vehicles within the communication range over a best path for forwarding the packets and thus wait for the best delivery paths to become available.

Most of these protocols do not take into account the vehicular traffic, which means that such algorithms may fail in case they try to forward a packet along streets where no vehicles are moving. Such streets should be considered as ‘broken links’ in the topology. Moreover, a packet can be received by a node that has no neighbors nearer to the receiver than the node itself. In this case, the problem of a packet having reached a local maximum arises. These problems can be overcome to some extent knowing the real topology, by opting to use only streets where vehicular traffic exists. In addition, in [159][159] and [160][160], forwarding a packet between two successive intersections is done on the basis of a simple greedy forwarding mechanism. This classic greedy approach works well since it is independent of topological changes but it suffers from inaccurate neighbor tables since it does not consider the vehicle direction and velocity. The simple knowledge of node’s position could not be sufficient in a network with frequent topological changes, such as in a VANET (nodes already gone). Moreover, it may be possible to lose some good candidate nodes to forward the packets. In such a situation, it is utmost important to guarantee a high stability for the links and therefore robustness of the routing protocol. So, in order to provide a solution to the above-mentioned problems, there is a need for a new intersection-based geographical routing protocol capable to find robust routes within city environments.

The foreseen geographical routing protocol for VANETs is among our major contributions. Our final objective is to propose a position-based traffic aware routing algorithm that uses an improved greedy forwarding strategy and that is adapted to city environments. Particularly, the main objective of such routing algorithm is to provide better throughput and lower routing overhead by adopting an efficient carry-and-forward scheme, exploiting the vehicular traffic density and efficient intersection selection schemes. This is performed in an aim to improve the wireless resource usage by unicast traffic in vehicular packet networks.

5.4.2 Contributions

In the following, we focus on the design of a robust routing protocol taking into account the characteristics of urban environments, the inherent characteristics of vehicular networks as well as the network density (i.e. estimated vehicular traffic density). The resulting novel geographic routing protocol for urban VANET is called GyTAR: *improved Greedy Traffic Aware Routing* protocol. GyTAR is an intersection-based geographical routing protocol that utilizes the vehicular traffic density and the road-topology to efficiently relay data in the network.

GyTAR considers that each vehicle in the network knows its own position and speed using GPS and can determine the position of their neighboring intersections through pre-loaded digital maps, which provides a street-level map. The presence of such kind of maps is a valid assumption since vehicles are becoming increasingly equipped with on-board navigation systems. Furthermore, a sending node needs to know the current geographical position of the destination in order to make a routing decision. This is achieved, in real time, thanks to a location service, which can be made available, for example, using city-scale Wireless Sensor Networks (WSN) [165][165]. We also assume that each vehicle is required to maintain a neighbor table where the position, velocity and direction of each neighboring vehicle are recorded. This table is built and updated thanks to the periodic exchange of *Hello* packets by all vehicles. Finally, we do not impose any restriction to the communication model, and GyTAR is applicable to both completely ad hoc (mobile sources and destinations, e.g., gaming on-the-move applications) and hybrid (fixed destination, e.g., a service hotspot) communications.

In such a context, GyTAR is designed to work optimally within urban environments in order to find robust routes. The main objective of GyTAR is to provide better throughput and lower routing overhead. To reach this objective, it is organized into three mechanisms: *(i)* a completely decentralized scheme for the estimation of the vehicular traffic density in city-roads, *(ii)* a mechanism for the dynamic selection of the intersections through which packets are forwarded to reach their destination, and *(iii)* an improved greedy forwarding mechanism to be used to forward packets between two intersections. Hence, using GyTAR, packets will move successively closer towards the destination along the streets where there are enough vehicles providing connectivity. We showed on our previous contribution (cf. §5.3) that IFTIS can be used in order to accurately and promptly estimate the vehicular traffic density in city-roads. Hence, we concentrate in the following on explaining the two other schemes upon which GyTAR is built: the efficient intersection selection scheme and the improved carry-and-forward scheme.

5.4.2.a Intersection Selection

Similar to position-based source routing, GyTAR adopts an anchor-based routing approach with street awareness. Thus, data packets are routed between vehicles, following the street map topology. However, unlike GSR [159] and A-STAR [160], where the sending node statically computes a sequence of intersections the packet has to traverse in order to reach the destination, intermediate intersections in GyTAR are chosen dynamically and in sequence, considering both the variation in the vehicular traffic density and the distance to destination. Partial successive computation of the path has a threefold advantage: *(i)* the size of packet header is fixed; *(ii)* the computation of subsequent anchors is done exploiting more updated information about vehicular traffic distribution; and *(iii)* subsequent anchors can be computed exploiting updated information about the current position of the destination.

When selecting the next destination intersection, a node (the sending vehicle or an intermediate vehicle in an intersection) looks for the position of the neighboring intersections using the map. A score is attributed to each intersection considering the traffic density and the curvemetric distance[6] to the destination. The best destination intersection (i.e., the intersection with the highest score) is the geographically closest intersection to the destination vehicle having the highest vehicular traffic. Thus, using this real time traffic aware approach, the determined end-to-end route will be the one with higher connectivity.

[6] This term describes the distance measured when following the geometric shape of the road.

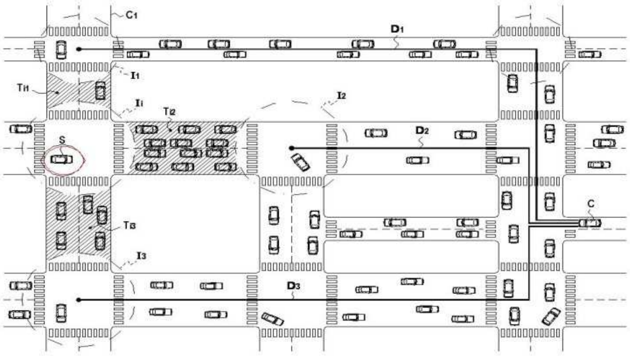


Figure 5.3. Selecting intersections in GyTAR.

Figure 5.3 shows an example of how the next intersection is selected. In this scenario, once vehicle **S** (located in the intersection **Ii**) receives a packet, it computes the score of each neighboring intersection. Considering its curvemetric distance to the destination (**D2**) and the traffic density (**Ti2**), intersection (**I2**) will have the highest score. It is then chosen as the next anchor.

To formally estimate the score of an intersection, we define the following notations:

- *J*: the next candidate intersection;
- *I*: the curent intersection;
- D_j: the curvemetric distance from the candidate intersection *J* to the destination;
- D_i: the curvemetric distance from the current intersection to the destination;
- $D_p = D_j/D_i$ (D_p determines the closeness of the candidate intersection to the destination point);
- Between intersection *I* and intersection *J*:
  - N_i : number of vehicles within cell i;
  - N_{con}: constant representing the ideal connectivity degree we can have within a cell;
  - N_c : number of cells[7] between *I* and *J;*
  - N_{avg}: average number of vehicles per cell: $N_{avg} = \frac{1}{N_c} x \sum_{i=1}^{Nc} N_i$ (1);
  - σ : standard deviation of cell densities N_i (it indicates how much variation there is away from the N_{avg}): $\sigma = \sqrt{\left(\frac{1}{N_c} x \left(\sum_{i=1}^{Nc} (N_i - N_{avg})^2\right)\right)}$ (2);
- α, β: used as weighting factors for the distance and vehicular traffic respectively (with $\alpha + \beta = 1$).

Hence,

$$Score(J) = \alpha \cdot f(D_j) + \beta \cdot g(T_j) = \alpha \times [1 - D_p] + \beta \times \left[\min\left(\frac{1}{\sigma + 1} * \frac{N_{avg}}{N_{con}}\right), 1\right] \quad (3)$$

As we can see, this equation is based on two factors:

- The first factor D_j is a measure of the distance to the destination in road length. Shorter distances to the destination are preferred. To calculate the score distance, we proposed the following function:

[7] The cell size is determined based on the wireless transmission range of vehicles.

$f(D_j) = \lfloor 1 - D_p \rfloor$ where D_p determines the closeness of the candidate intersection to the destination point. Hence, the closer the potential intersection *j* is, the lower the parameter D_p is, and the higher the score distance is.

Note that when the candidate intersection corresponds to the final destination intersection, D_p is equal to zero, which corresponds to the highest score distance we could have ($f(D_j)=1$).

- The second factor T_j is a measure of the traffic density between the current intersection and the potential intersection *j*. Well balanced streets with higher density are preferred. One possible function to calculate the score density is:

$$\mathrm{g(Tj)} = \left\lceil \min\left(\frac{1}{\sigma+1} * \frac{N_{avg}}{N_{con}}\right), 1 \right\rceil \qquad (4)$$

As we can see, the score density depends on three parameters (N_{avg}, σ and N_{con}). Indeed, for a given street, the traffic density estimation module (i.e. IFTIS – cf. §5.3) provides the density N_i of each cell. The cells density is then used to calculate N_{avg} (average number of vehicles per cell) using equation (1) and σ (standard deviation of cells density) using equation (2). Whereas N_{con} is a constant which represents the ideal connectivity degree we can have within a cell to ensure a good end-to-end connectivity. In order to compute the N_{con} value, we performed an analytical study which showed that the ideal connectivity degree is obtained while maximizing the following function (i.e. N_{con} is equal to the *ρ* value that maximizes the following probability):

$$P(\rho, R, m) = \prod_{1}^{m-1} F(R, \rho) = \left[1 - \exp(-2.\rho.R)\right]^m \qquad (4)$$

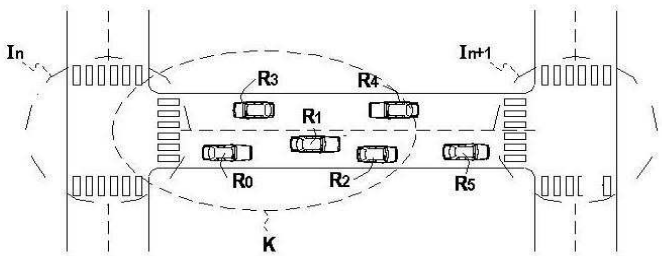
where *R* is the radio communication range, L the length of the considered road segment, and *m* is defined as the integer value of [*L/R*].

Thus, (N_{avg} /N_{con}) determines how high the whole street density is, while σ indicates if the street is well balanced or not. Hence, by multiplying (N_{avg} /N_{con}) by (1/ σ +1), we penalize streets with a large standard deviation since this corresponds to scenarios where we have gaps within the street (isolated clusters of vehicles).

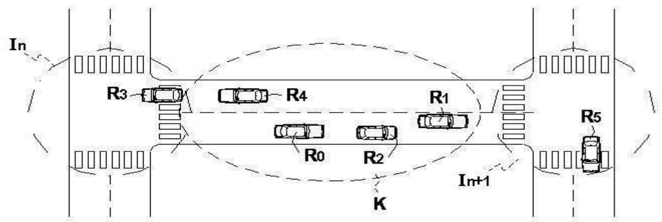
Note that like $f(D_j)$, $g(T_j)$ should provide values less than 1. This is why we used the *min* function.

5.4.2.b Forwarding data between two intersections

Once the next intersection selected, there is a need for a mechanism that allows forwarding the data packet from the current to the next intersection. To do so, we propose to use an improved greedy strategy. For that, all data packets are marked by the location of the next intersection. Each vehicle maintains a neighbor table in which the velocity information of each neighbor vehicle is recorded. This table is updated through *Hello* messages exchanged periodically by all vehicles. Thus, when a data packet is received, the forwarding vehicle forecast the current position of each neighbor using the corresponding recorded information (velocity, direction and the latest known position), and then selects the next hop neighbor (i.e. the one forecasted as being the closest to the next intersection). Note that in the case where there are two possibilities (i.e. two vehicles at the same distance to the next intersection), the forwarding vehicle picks up one, randomly.



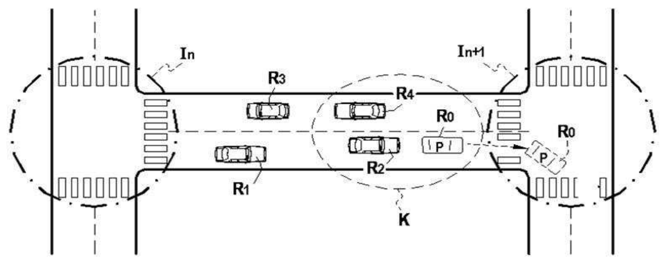
(a) Configuration at $t_1 < t_2$



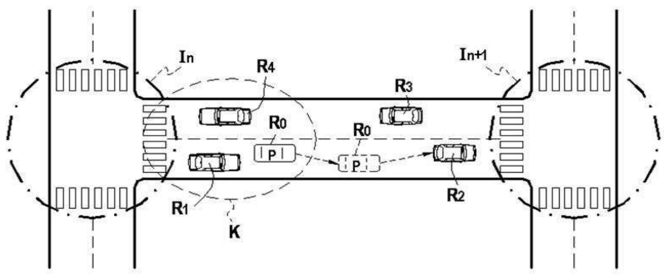
(b) Configuration at current time t_2

Figure 5.4. Forwarding data between two intersections using an improved greedy strategy.

The proposed improved greedy forwarding approach is illustrated in Figure 5.4. In this example, vehicle (R1), which is moving in the same direction as the forwarding vehicle with a speed greater than vehicle (R2), will be chosen as next hop at time $t=t_2$. Indeed, at $t= t_2$, vehicle (R1) is forecasted to be the closest vehicle to the next intersection (cf. Figure 5.4(a)). However, while using a legacy greedy forwarding approach (i.e. without using prediction), the forwarding vehicle would choose vehicle (R4) as the next hop instead of vehicle (R1) since it was the closest to the destination intersection at time ($t_1 < t_2$) (i.e. the last time the neighbor table was updated - cf. Figure 5.4(b)). Note that most of the existing greedy-based routing protocols do not use the prediction and consequently might lose some good candidates to forward data packets.



(a) Forwarding vehicle (R0) carrying data packet to the next intersection.



(b) Forwarding vehicle carrying data packet to the first vehicle reaching its radio range.

Figure 5.5. Recovery strategy used in a local optimum.

Despite the improved greedy routing strategy, the risk remains that a packet gets stuck in a local optimum (i.e. the forwarding vehicle might be the closest to the next intersection). Hence, a recovery strategy is required. The recovery strategy adopted by GyTAR is based on the idea of 'carry-and-forward' [166]: the forwarding vehicle of the packet in a recovery mode will carry the packet until the next intersection (see Figure 5.5(a)) or until another vehicle, closer to the destination intersection, enters its transmission range (see Figure 5.5(b)).

5.4.3 Summary of Results

To conclude, we can summarize our second contribution regarding the improvement of wireless resource usage in vehicular packet networks as follows. In this contribution, a robust intersection-based geographic routing protocol is discussed. This protocol is targeting the efficient relaying of data for vehicle-to-vehicle communications and vehicular internet access, while operating within urban environments. The GyTAR protocol efficiently utilizes the unique characteristics of vehicular environments like the highly dynamic vehicular traffic, road traffic density as well as the road topology in making routing and forwarding decisions. The selection of intermediate intersections among road segments is performed dynamically and in-sequence based on the scores attributed to each intersection. The scores are determined based on the dynamic traffic density information and the curvemetric distance to the destination. The traffic density information for intersection score calculation is obtained thanks to IFTIS, our decentralized mechanism that proved estimating nearly accurate vehicular traffic along city roads, with very low percentage of error (cf. §5.3).

In order to show the performance improvements of GyTAR and its various features, we performed a set of simulations using Qualnet [149] as simulation tool and VanetMobiSim [150] as a realistic vehicular mobility model. Our simulation study compared four protocols. Two versions of GyTAR that we implemented: B-GyTAR (Basic GyTAR without local recovery, i.e., a packet is simply dropped when it encounters a local optimum situation), and GyTAR with local recovery; A version of the position-based vehicular routing protocol GSR [159] (which more closely resembles the nature of our algorithm) that we also implemented since there is not any publicly available implementation of the protocol; And the Location Aided Routing (LAR) Protocol [167]. Simulation results show that GyTAR performs better in terms of throughput, delay and routing overhead compared to the other protocols (LAR and GSR) proposed for vehicular networks. While comparing GyTAR to B-GyTAR, our simulation results also showed that the delay increase due to the local recovery feature of GyTAR is fully compensated by the important increase in the throughput. Finally, we looked for optimal values for the weighting factors (α,β) of the traffic density and distance information components in the intersection scores. We evaluated their impact and analyzed their sensitivity. This allowed showing that there is a good balance between the distance and traffic-density parameters in almost all cases. To sum up, we can say that the robust intersection selection and the improved greedy carry-and-forward scheme with recovery, suggests that GyTAR should be able to provide stable communication while maintaining higher throughput and lower delays for vehicular routing in urban environments. Thus, by using GyTAR for unicast routing, we can reach our objective to improve the wireless resource usage in vehicular packet networks.

This work is one of the main results of a fructuous cooperation with Orange Labs in the frame of a CIFRE Convention. It was one of the focuses of the Ph.D. dissertation carried out by Moez Jerbi, University of Evry val d'Essonne, and defended on November 6, 2008 (see papers [P11][P13][P27][P45][P68]).

5.5 GEO-LOCALIZED VIRTUAL INFRASTRUCTURE FOR DISSEMINATION IN VANETS

5.5.1 Research Context

As explained earlier the opportunities and areas of applications of vehicular packet networks are growing rapidly. Vehicular packet networks are expected to be used by a wide range of applications ranging from active safety applications (including collision and warning systems, driver assistance applications and intelligent traffic management systems) to online vehicle entertainment (such as gaming, file sharing, or integration with Internet services and applications [133]). Many of these applications rely on distributing data, e. g., on the current traffic situation, on free parking lots…. Often, data needs to be distributed over long distances, for example to allow a driver to choose between different arterial roads when driving into the city center. Typically, such applications are based on some form of proactive information dissemination in an ad hoc manner - i.e. by forming Vehicular Ad hoc Networks (VANETs). Proactive information dissemination is, however, a difficult task due to the highly dynamic nature of VANETs. Indeed, VANETs are characterized by their frequent fragmentation into disconnected clusters that merge and disintegrate dynamically [144]. In addition, the results presented in [168] clearly show that during the rollout of VANET technology, some kind of support is needed. Otherwise, many envisioned applications are unlikely to work until a large fraction of vehicles participate.

One of the largely accepted solutions towards efficient data dissemination in vehicular packet networks is by exploiting a combination of fixed roadside infrastructure (e.g. Road Side Units, RSU) and mobile in-vehicle technologies (e.g. On Board Units, OBU). For example, in [169], roadside base stations are used to bridge network partitions in vehicular networks. A car already informed of an accident forwards the alert when passing by a roadside base station. Subsequently, the base-station forwards the message to other base-stations located in the alert zone. Each of the informed stations periodically broadcasts the alert to inform passing vehicles. Another recent example of broadcasting protocol specifically designed for vehicular packet networks with infrastructure support is the Urban Multi-hop Broadcast (UMB) protocol presented in [152]. UMB gives insightful results in terms of successful delivery rate. However, this is obtained with the help of repeaters at the road intersections. The need for an infrastructure considerably decreases the deployment area of UMB-based networks as UMB fails to handle intersections without a repeater. So, while such infrastructure-based approaches may work well, they may prove costly as they require the installation of new infrastructures on road network, especially if the area to be covered is large. The need for an infrastructure will considerably decrease the deployment area of VANET applications. Thus, in order to overcome this issue in certain areas, one idea is to design a self-organizing mechanism to emulate such infrastructure. The resulting geo-localized virtual infrastructure (GVI) will thus be represented by a bounded-size subset of vehicles populating the concerned geographic regions. Hence, a vehicle that enters the geographic region of a GVI attempts to participate in the mechanism; a vehicle that leaves the geographic region ceases to emulate the GVI. If all the vehicles leave a GVI's region, then the GVI fails; if vehicles return, then the GVI restarts...

The foreseen geo-localized virtual infrastructure for data dissemination in VANETs is among our major contributions. Our final objective is to design a self-organizing mechanism that emulates a geo-localized virtual infrastructure (GVI) and study its feasibility. Particularly, the main objective of such mechanism is twofold: approaching the performance of a real infrastructure while avoiding the cost of installing it. This is performed in an aim to improve the wireless resource usage by dissemination-based application in vehicular packet networks without the need to install a costly infrastructure.

5.5.2 Contributions

Before discussing the design of the Geo-localized Virtual Infrastructure (GVI), let us first discuss where such mechanism may be instantiated. Indeed, a critical question that arises is where to position the GVI, in order to allow for a best-possible support of VANETs. This depends on which environment the GVI will perform and for which application. As we are dealing with city environments, an intersection sounds suitable as geographic region because of its better line-of-sight and also because it is a high traffic density area. Hence, the proposed GVI mechanism can periodically disseminate the data within a signalized (traffic lights) intersection area, controlled in fixed-time and operated in a range of conditions extending from under-saturated to highly saturated. Thus, it can be used to keep information alive around specific geographical areas (i.e. nearby accident warnings, advertisements and announcements, available parking lot at a parking place, etc.). It can also be used as a solution for the infrastructure dependence problem of some existing dissemination protocols like UMB [152]. One should also note that the GVI mechanism can be preferably instantiated in intersections with an acceptable level of car density (like in downtowns and highly used roads). In intersections implying low car densities, we can either decide deploying GVI or a static repeater. These situations being rare in metropolitan areas, the implied consequences remain reduced.

The GVI mechanism consists on electing vehicles that will perpetuate information broadcasting within an intersection area. To do so, the GVI is composed of two phases: *(i)* selecting the vehicles that are able to reach the broadcast area (i.e. a small area around the intersection center, where an elected vehicle could perform a local broadcast); then, *(ii)* among the selected vehicles, electing the local broadcaster which will perform a local single-hop broadcast once it reaches the broadcast area (i.e. the intersection center).

In the first phase (i.e. selection of candidate vehicles), among the vehicles which are around the intersection, only those which are within the notification area are designated as potential participants to the GVI mechanism. They are selected as candidates only if they are able to reach the intersection center. The considered notification area is a region around the intersection starting at TR/2 before and extending to TR/2 beyond the intersection where TR is the transmission range of a vehicle. Figure 5.6 illustrates the candidate vehicles selection where vehicles {A, B, C, D, E, F, G, H} could participate to the GVI mechanism since they are located within the notification area and only vehicles {A, B, D, F} are selected as candidates because they are moving towards the broadcast area.

Then, during the second phase (i.e. local broadcaster election), each vehicle selected as candidate vehicle starts by computing the time period Δ needed to reach the intersection center by considering its geographical location, direction and speed. According to this time period, it computes a weight $P(\Delta)$. This one has to be maximal when the expected delay matches the desirable broadcast cycle time T of the GVI and it decreases when we are far from T. One possible function for computing the weight is given by (1) (σ is a constant) but other functions (e.g. triangle) can also be considered.

$$P(\Delta) = 1/\left(\sigma.\sqrt{2\Pi}\right) \times \exp\left(-1/2\left(\left(\Delta - T\right)/\sigma\right)^2\right) \qquad (1)$$

After the weight calculation, a waiting time $WT(P) = MaxW\ (1 - P/P_{max})$ is assigned to each candidate vehicle. The candidate vehicle with the highest weight P will have the shortest waiting time WT to broadcast a short informative message telling other candidate vehicles that it has been elected as the local broadcaster.

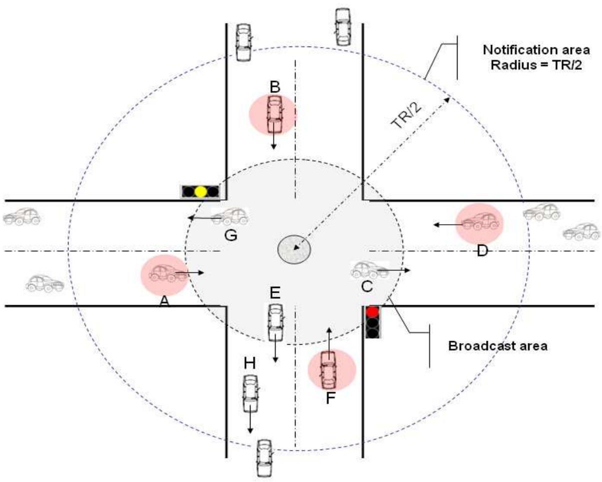


Figure 5.6. Selecting vehicles candidates in the GVI mechanism.

One may also note that the probability of having a collision between two informative messages is weak. This is due to two reasons, the length of these messages and the number of vehicles that may compute similar weights P. In the unlikely event of a collision among two broadcasted messages, the

GVI will have multiple elected nodes which will perform the local broadcast while arriving at the intersection center instead of one. So, such collisions will not break the GVI (i.e. no dramatic effect).

The reason to choose the intersection region starting at TR/2 before the intersection is that the elected vehicle has to inform the other candidate vehicles. In the worst case, the elected vehicle is TR/2 away from the intersection and it can cover the points up to TR/2 away at the other side of the intersection.

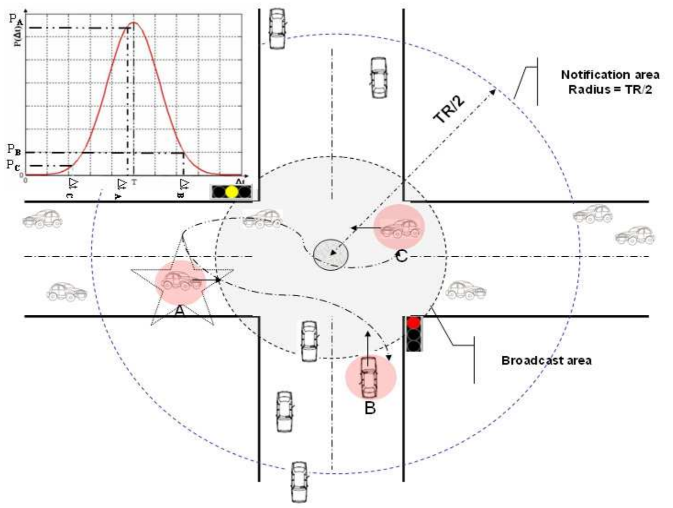


Figure 5.7. Electing the local broadcaster in the GVI mechanism.

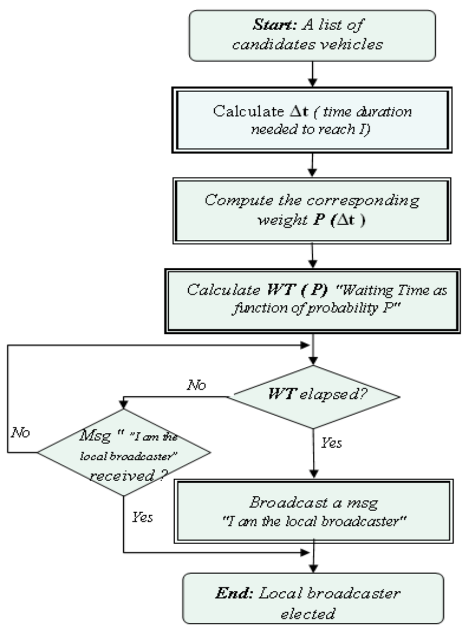


Figure 5.8. Electing the local broadcaster in GVI mechanism.

An example of vehicle election is illustrated in Figure 5.7. In this example, candidate nodes, vehicles A, B and C, compute the time period Δ to reach I (i.e. the intersection center) considering their position, direction and speed. B will have a long time period Δ_B since it is stopped at the traffic light. C has a very short time period Δ_C since it is very close to I. A requires a time period Δ_A very close to the broadcast cycle time *T*. Consequently, A will have the highest value of *P* and the shortest *WT*(*P*). A will be the first to send a message to vehicles B and C informing them that it has been elected to perform the local broadcast once it reaches the broadcast area around the intersection center I. Once vehicles within the transmission range of the elected vehicle receive the broadcasted message, they will participate in the election of the next local broadcaster following the same process. Note that the elected vehicle has always the closest time duration to *T*. Hence, we can ensure that GVI will perform a periodic local broadcast. Figure 5.8 illustrates the whole process of vehicle election in the GVI mechanism.

5.5.3 Summary of Results

To conclude, we can summarize our third contribution regarding the improvement of wireless resource usage in vehicular packet networks as follows. In this contribution, we presented an elegant solution for building a Geo-localized Virtual Infrastructure using inter-vehicle communications. GVI is a self-organizing mechanism that allows the vehicles populating the concerned geographic regions to virtually play the role of a repeater in order to disseminate localized information. The proposed mechanism has various potential applications ranging from safety to convenience applications, solving by the way the infrastructure dependence problem of some existing dissemination protocols. GVI is designed in an aim to improve the wireless resource usage by dissemination-based application in vehicular packet networks without the need to install a costly infrastructure.

In order to analyze the performance of GVI and thus show that such mechanism is realistic, we used both discrete event simulations and theoretical analysis. More precisely, the objective of our performance evaluation was twofold: *(i)* to investigate the impact of traffic properties and system parameters (broadcast time-cycle, *T*) on performance criteria, and *(ii)* to analyze the probability that the geo-localized virtual infrastructure breaks because of a low traffic density. We thus showed that by changing the broadcast time-cycle, *T*, a good trade-off can be achieved between the probability to inform a vehicle (that is a measure of quality of service) and the number of copies of the same message received by a vehicle (that is a measure of cost to provide the service). We also showed through analytical modeling that the probability that the GVI breaks (i.e. the probability to fail during the election of the next local broadcaster because of a low traffic density) is negligible when the average traffic arrival rate is higher or equal to 0.2 vehicles/seconds. Overall, our analytical and simulation results showed that the proposed GVI mechanism can periodically disseminate data within an intersection area, efficiently utilize the limited bandwidth and ensure high delivery ratio.

This work is another important result of the fructuous cooperation with Orange Labs in the frame of a CIFRE Convention. It was one of the focuses of the Ph.D. dissertation carried out by Moez Jerbi, University of Evry val d'Essonne, and defended on November 6, 2008 (see papers [P38][P39]).

5.6 CONCLUSION

The work presented here was realized in the frame of the Ph.D. dissertation of Moez Jerbi at University of Evry val d'Essonne (November 2005 – November 2008). This Ph.D. work was held in the frame of a CIFRE convention with Orange Labs (France Telecom Group) and for which I was the secondary supervisor. This work is thus also part of the fructuous cooperation that I have with Orange Labs since 2005.

The obtained results are numerous. We started from the fact that it is of imminent practical interest to consider the vehicular traffic density in any newly proposed multi-hop communication protocol in order to maximize its benefits. Indeed, while analyzing the vehicular networks characteristics, especially in city-environment, temporary disconnection seems to be clearly unavoidable. So, as a preliminary step towards improving the wireless resource usage in vehicular packet networks, we started by designing and validating a completely distributed and infrastructure-free mechanism for city road density estimation. The performance evaluation of this mechanism, called IFTIS, proved that it is able to estimate nearly accurate vehicular traffic along city roads, with very low percentage of error. Then, we used this traffic density estimation to help determining optimal vehicular data routing paths in VANETs. Hence, based on this critical metric, we proposed a novel intersection-based geographical routing protocol, called GyTAR, which proved to be capable to find robust and optimal routes within urban environments. Its performance comparison with other competing protocols from the literature showed clearly that GyTAR gives better results in terms of throughput, delay and routing overhead.

Our analysis of ITS applications showed that these rely for data delivery on two multi-hop communications modes: unicast routing and dissemination. In order to help the efficient support of dissemination-based applications, our third contribution in this topic was to propose a self-organizing mechanism that emulates a geo-localized virtual infrastructure, called GVI. The idea of this mechanism is to replace the costly infrastructure in the areas that are well populated (i.e. areas in which the average traffic density is within an acceptable level). Indeed, previous works showed that the performances of VANET dissemination protocols can be highly improved in presence of fixed infrastructure in strategic places, i.e. the intersections. We showed through discrete event simulations and analytical studies that GVI, our proposed mechanism, can realistically replace such infrastructure as it is able to periodically disseminate the data within the area in which it is deployed, efficiently utilize the limited bandwidth and ensure high delivery ratio.

To conclude we can say that, overall, we showed that using our routing and dissemination solutions, we are able to improve the wireless resource usage in vehicular packet networks compared to what is proposed in the literature. Furthermore, this improvement is obtained using pure VANET settings (i.e. without the help of any costly infrastructure).

Chapter 6. Conclusion and Perspectives

6.1 Concluding Remarks

Our research started from the fact that the Internet is constantly evolving towards offering an ever growing number of networked services going from data-intensive to QoS-demanding multimedia services. These services are assumed to be offered anytime, anywhere, using any terminal, and through any kind of networks. However, the current Internet architecture was not designed to efficiently deal with this variety of supported services, available networks, and diversity of user devices. Among the issues that need to be handled in order to support such diversity, performance degradation is among the most important ones. So, it is crucial to rethink the architectures and protocols with the aim to annihilate the performance bottlenecks. Thus, a wider variety of services can be provided, diverse user terminals can be supported, and a multitude of networks can be used.

Future data-intensive and QoS-demanding multimedia services impose to both, networks and terminals, a certain number of constraints in order to operate correctly. These constraints, mainly performance ones, can be expressed in terms of bandwidth, delay, jitter, packet loss rates as well as terminal's energy consumption. Moreover, these services are expected to be available throughout various novel wireless packet network technologies such as Wireless Local Area Networks, Broadband Wireless Access Networks, Mobile Ad hoc Networks as well as their diverse use cases. So, the development of mechanisms that allow using efficiently the wireless and terminal resources in such a realm is mandatory in order to wipe out the performance bottlenecks and be able to offer guarantees to the above mentioned constraints.

Proposing mechanisms leading to an efficient usage of the scarce wireless and terminal resources constituted the target of our research. Different facets of this problem have been addressed and solved. These facets relate to the improvement of the resource usage at both packet and connection levels.

To meet our objective of improving the resource usage at the connection-level, the question we answered is: *how to maintain the experienced QoS while the users are willing to move transparently over heterogeneous wireless packet networks?* Indeed, the users are expecting more and more communication capabilities while using their mobiles. They are willing to access a broad range of data-intensive and multimedia services without regard to the technology, the access network type, or the device. Their only concern is to access these services while having a: *(i)* complete freedom of movement, *(ii)* service continuity with an improved perceived QoS, and *(iii)* most competitive cost according to their personal requirements. From a networking perspective, this entails a need to rethink completely the mobility management process in order to make an efficient use of both the terminal and wireless resources in such a heterogeneous environment. The response to this question was the key driver of our work. To achieve this, our process was first to discuss and adapt the theory (i.e. *utility theory*) to be used in order to match the personal requirements of the user while choosing the best access network. Then, around this theory, we built a complete terminal-controlled handover management framework with the target to efficiently answer the above mentioned user concerns. The proposed solutions proved to be adequate to improve the wireless resource usage in heterogeneous wireless packet network environment from both the user's and operator's perspectives.

At the packet-level, the improvement of resource usage translates into the following question: *how to improve the QoS experienced by the packet flows generated by data and multimedia services?* The challenges behind this question are numerous. Indeed, answering this question relates on both, the type of service we are dealing with (*elastic* vs. *QoS-demanding*) as well as the wireless packet network technology or architecture in which the service is used (*WLANs*, *BWANs*, *MANETs*,…). Hence, we identified: *(i)* for elastic traffic, TCP behavior as the major performance bottleneck to deal with in wireless packet networks; *(ii)* for QoS-demanding traffic, multiple missing resource management bricks that needs to be apprehended; and *(iii)* for challenging networking environments, such as vehicular networks in urban areas, providing efficient multi-hop communication solutions is the main concern. Giving a complete and comprehensive answer to each of these issues, with the aim to get an improved usage of the wireless (and sometimes terminal) resources, was thus our primary target. To achieve this, we proposed several solutions by acting at the transport, network and link layers. All the proposed solutions proved to be clearly more effective than concurrent solutions.

6.2 PERSPECTIVES

This document enumerated a number of contributions that we conducted in the frame of our research targeting the improvement of the resource usage in wireless packet networks. Many other research challenges in this topic still need to be resolved aiming for the continuous expansion of the telecommunication and Internet sectors. As an open perspective of our works, we aim at investigating new trails for making the wireless resource usage more and more efficient.

6.2.1 Improving the Resource Usage at the Packet-level

Among the new trails we are investigating, we can cite the use of two novel approaches to improve the resource usage at the packet-level: *cooperative relaying* and *network coding*. Indeed, we think that these approaches can be of primary interest in order to improve the reliability of data delivery as well as bandwidth utilization of multi-hop wireless networks.

The expected improvements can be even more remarkable in challenging networking environments such as Vehicular Networks as well as Wireless Sensor Networks (WSN). These two networking environments offer a huge research potential due to their characteristics. While vehicular networks are characterized by a large range of medium-to-high velocities, predictable mobility patterns, frequent network partitioning, and challenging communication contexts, to name a few; WSNs are characterized by their restricted resources (energy, memory, processor) and the random quality of the communication medium. Both experience extreme communication conditions due to these characteristics. The potential of wireless resource usage improvement using *cooperative relaying* and/or *network coding* is thus immense.

One of the most interesting alternatives to enhance the channel capacity is to use cooperative relaying. By using such a technique, a source node *S* can profit from the availability in its vicinity of several antennas (i.e. the communication interfaces of its neighbors) to enhance its communication channel when this one is poor. Hence, if the direct link between a source *S* and a destination *D* is experiencing bad channel conditions, *S* may use some of its neighbors (i.e. those experiencing better channel conditions with *D*) to relay its packets towards *D*. Proposing practical (i.e. protocol design) and theoretical (i.e. analytical models) solutions to exploit efficiently this spatial diversity can greatly improve the wireless resource usage. The idea is to design protocols and analytical tools that allow enhancing the communication channel while taking into account the characteristics of the targeted wireless environment. More precisely, we will consider realizing this for both challenging networking environments cited above. In vehicular networks, such solutions should handle the rapid evolution of the spatial diversity due to the characteristics of such networks. For wireless sensor networks, the interactions with low-power listening techniques at the MAC layer as well as the interaction with the routing process to enhance it are two primary ideas that we are following.

Alternatively, the network coding theory has a significant impact on the way an information network is considered, whether at the practical level (i.e. protocol design) or the theoretical level (i.e. graph and information theories). Indeed, the well established forwarding schemes based on the *store-and-forward* transportation model are more to be insufficient in many kinds of communication networks, and especially in multi-hop wireless networks. Indeed, it had been shown in the literature that the inherent shared and broadcast nature of the radio channel is not well used. Hence, if well designed a solution based on network coding can lead to significant increase in the performance of wireless networks. In fact, the basis of a network-coding-based communication system is to employ coding mechanisms at intermediate nodes, in order to achieve bandwidth optimality. These mechanisms may vary from a solution to another, but they consist essentially of a linear combination of two or more messages in one output message using random coding vectors of equal length. The objective here will be to use the huge potential offered by this new approach in order to get a substantial additional improvement of the wireless resource usage compared to what it is possible today when a classical forwarding approach is used. More precisely, the idea is to design the adequate mechanisms that will allow achieving bandwidth optimality for both challenging networking environments cited above. In vehicular networks these mechanisms have to cope with the disruptive vehicular environment. They have also to take advantage from the predictability of mobility patterns and the road topology characteristics. For wireless sensor networks, node capabilities (power-saving operation mode, low processing and low memory capabilities) are the constraint to be taken into account. In both cases, network coding mechanisms can be designed in order to improve the routing, dissemination and multicasting, as well as the end-to-end communication robustness within these challenging environments.

Combining both approaches, network coding and cooperative relaying, is also an option to allow even better performance enhancement possibilities. Indeed, we can see from their respective definitions that these two approaches are complementary. Furthermore, they are both acting at different layers. While network coding mechanisms are usually used either at the network or the physical layer, the cooperative relaying protocols that we are targeting are MAC layer ones. This combination is expected to give wireless coverage enhancement and capacity augmentation at the one-hop level thanks to cooperative relaying, while network coding will allow achieving bandwidth optimality at either the two-hop level (i.e. using *Physical-layer Network Coding*) or multi-hop level (i.e. using *Packet-level Network Coding*). We strongly believe that making both approaches collaborating tightly and in a cross-layer manner can allow going further in the improvement of the resource usage in multi-hop wireless networks such as vehicular networks and WSNs.

The opportunities presented here represent some of the major short and medium-term perspectives we propose to investigate in order to continue improving the resource usage at the packet-level.

6.2.2 Improving the Resource Usage at the Connection-level

In addition to the above mentioned packet-level issues, vehicular networks also engenders multiple new challenges at the connection-level. Indeed, in order to offer a cost-effective always-best service to vehicular users, we shall consider the co-existence of the numerous network technologies that are being developed and deployed (some of them are already deployed and the others are expected to come to a deployment phase in the near future). Coming to a full interworking among all cohabiting technologies (LTE-Advanced, 802.16, 802.11p/WAVE ...) necessitates the resolution of multiple technological bolts. This is especially challenging while taking into account the interworking with the upcoming multi-hop wireless vehicular networks combining Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications as a mean to access to a converged core network. Mixing all available wireless technologies (LTE-Advanced, 802.16, 802.11p/WAVE ...) as well as combining V2V and V2I communications will thus be the basis of a "*self-organized multi-technologies vehicular network*". In order to be made available, such "*self-organized multi-technologies vehicular network*" needs:

- *Efficient Network Selection:* the coexistence of multiple wireless and mobile technologies necessitate that mobile devices, at each instant, answers the question about which network technology to use. Current initiatives in that domain concern mainly the selection among WLAN access, WiMAX access and cellular access using access points and base stations. Multi-hop wireless networks are not considered at all. So, there is a need to propose an efficient network selection that takes into account the different possible use cases that may happen in the "self-organized multi-technologies vehicular network" realm (i.e. network selection among different communications paradigms, namely multi-hop or cellular, and different wireless technologies such as LTE-Advanced, 802.16, and 802.11p/WAVE). The objective here is still, depending on the application, to choose the network that maximizes the user's profitability and/or improves its perceived QoS.

- *Mobility Management:* the efficient network selection is performed at each instant. So, handovers are most likely to happen among the different communication paradigms or wireless technologies due to the presence of a better communication facility or the imminence of a network disconnection. An efficient mobility management solution that targets to avoid end-to-end service disruption is thus to be provided. In the frame of the envisioned self-organized multi-technologies vehicular network, such a mobility management is more complex to be realized than in the case of heterogeneous wireless packet networks (i.e. as considered in Chapter 2). Indeed, the multi-hop nature of some involved networks as well as the characteristics of the vehicular networks (i.e. large range of medium-to-high velocities, predictable mobility patterns, frequent network partitioning, and exigent communication contexts) makes it challenging to offer such an efficient mobility management solution while avoiding service disruptions.

- *Hybrid Routing over Heterogeneous Networks:* self-organization in a multi-technologies vehicular network engenders multiple problems at the routing level. These issues are due to the heterogeneous nature of the involved communications paradigms and wireless technologies as well as the specificities of each of them. One of the major challenges will then be to solve these issues. More precisely, there is a need to propose a hybrid routing framework along with a set of dedicated routing algorithms that are compliant with this routing framework as well as with the communication paradigms they target.

The issues presented above represent some of the challenges that need to be solved, at short and medium-term, in order to allow going further in the mobility management with the aim to reinforce the connection-level quality of service in an extreme but realistic use case: Intelligent Transportation Systems. Indeed, in such systems there is a clear necessity to use all the available communication means in order to build advanced services and make them available anytime, anywhere, with the best available quality and at the lowest possible cost.

Chapter 7. Bibliography

[1] J. Strassner, “Seamless service mobility: How autonomic networking and communications will shape the future of next generation services,” *1st International Workshop on Seamless Service Mobility, SSMO’07*, Marrakech, Morocco, July 2007.

[2] L. Harju and J. Nurmi, “A baseband receiver architecture for UMTS/WLAN interworking applications,” *9th International Symposium on Computers and Communication, ISCC’04,* Alexandria, Egypt, Jun-July 2004.

[3] 3GPP, “3GPP system to Wireless Local Area Network interworking,” *TS23.234 v7.2.0*, 2006.

[4] 3GPP, “Radio Access Network; Generic Access Network (GAN); stage 2, Rel.8, Technical Specification,” *TS43.318 v8.1.0*, 2008.

[5] J. Liu, K. Kazaura and M. Matsumoto, “A low latency inter-system handover scheme for multiple interfaces terminal,” *6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, London, UK, May 2004.

[6] S.-L. Tsao and C. Lin, “Design and evaluation of UMTS-WLAN interworking strategies,” *56th IEEE Vehicular Technology Conference, VTC-Fall’02*, Vancouver, Canada, September 2002.

[7] L. Gras, Q.T. Nguyen-Vuong, Y. Ghamri-Doudane, N. Agoulmine, M. Kassar, B. Kervella and G. Pujolle, “Terminal Mobility in Mobile and Wireless Realm: tight, loose vs. very loose coupling,” *1st International Workshop on Seamless Service Mobility, SSMO’07*, Marrakech, Morocco, July 2007.

[8] H. J. Wang, R. H. Katz, and J. Giese, “Policy-enabled handoffs across heterogeneous wireless networks,” *IEEE Workshop on Mobile Computing Systems and Applications, WMCSA’99*, New-Orleans, Louisiana, February 1999.

[9] L. Chen, T. Sun, B. Chen, V. Rajendran, and M. Gerla, “A smart decision model for vertical handoff,” 4th ANWIRE International Workshop on Wireless Internet and Reconfigurability, ANWIRE 2004 (in conjunction with IFIP Networking'04), Athens, Greece, May 2004.

[10] O. Ormond, G. Muntean, and J. Murphy, “Utility-based intelligent network selection in beyond 3G systems,” *IEEE International Conference on Communications, ICC’06*, Istanbul, Turkey, June 2006.

[11] O. Ormond, G. Muntean, and J. Murphy,, “Economic model for cost effective network selection strategy in service oriented heterogeneous wireless network environment,” *10th IEEE/IFIP Network Operations and Management Symposium, NOMS'06*, Vancouver, Canada, April 2006.

[12] J. Chen, K. Yu, Y. Ji, and P. Zhang, “Non-cooperative distributed network resource allocation in heterogeneous wireless data networks,” *15th IST Mobile & Wireless Communications Summit, Mobile Summit’06*, Myconos, Greece, June 2006.

[13] H. Chan, P. Fan, and Z. Cao, “A utility-based network selection scheme for multiple services in heterogeneous networks,” *Intl. Conf. on Wireless Networks, Communications and Mobile Computing, WNCMC’05*, Maui, HI, USA, June 2005.

[14] W. Zhang, J. Jaehnert, K. Dolzer, “Design and Evaluation of a Handover Decision Strategy for 4th Generation Mobile Networks,” *57th IEEE Vehicular Technology Conference, VTC-Spring’03*, Jeju Island, Korea, April 2003.

[15] W. Shen and Q-A Zeng, “Cost-Function-based Network Selection Strategy in Integrated Wireless and Mobile Networks,” *21st International Conference on Advanced Information Networking and Applications Workshops, AINAW’07*, Niagara Falls, Canada, May 2007.

[16] E. Adamopoulou, K. Demestichas, A. Koutsorodi and M. Theologou, “Intelligent Access Network Selection in Heterogeneous Networks-Simulation Results,” *2nd International Symposium on Wireless Communication Systems, ISWCS2005*, Sienna, Italy, September 2005.

[17] M. Siebert et al., “Enhanced measurement procedures for vertical handover in heterogeneous wireless systems,” *14th IEEE Conf. on Personal, Indoor and Mobile Radio Communications, PIMRC’03*, Beijing, China, September 2003.

[18] E. Shil, P. Bahl, and M. J. Sinclair, “Wake on wireless: An event driven energy saving strategy for battery operated devices,” *8th ACM Annual International Conference on Mobile Computing and Networking, Mobicom’02*, Atlanta, GA, USA, September 2002.

[19] V. Gazis, N. Houssos, N. Alonistioti, and L. Merakos, “On the complexity of "Always Best Connected" in 4G mobile networks,” *58th IEEE Vehicular Technology Conference, VTC-Fall’03*, Orlando, Florida, October 2003.

[20] J. Kleijnen, “Scoring methods, multiple criteria and utility analysis,” *ACM Sigmetrics Performance Evaluation Review*, vol. 9, N° 3, pp. 45-56, Fall 1980.

[21] A. Hasswa, N. Nasser, and H. Hassanein, “Generic vertical handoff decision function for heterogeneous wireless networks,” *2nd IEEE/IFIP International Conference on Wireless and Optical Communications Networks, WOCN’05*, Dubai, UAE, March 2005.

[22] O. Ormond, G. Muntean, and J. Murphy, “Network selection decision in wireless heterogeneous networks,” *16th IEEE Intl Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC’05*, Berlin, Germany, September 2005.

[23] M. Xiao, N. Shro_, and E. Chong, “A utility-based power-control scheme in wireless cellular systems,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 210-221, April 2003.

[24] H. Lin, M. Chatterjee, S. Das, and K. Basu, “ARC: an integrated admission and rate control framework for competitive wireless CDMA data networks using noncooperative games,” *IEEE Transactions on Mobile Computing*, Vol. 4, N° 3, pp. 243-258, May-June 2005.

[25] L. Badia and M. Zorzi, “On utility-based radio resource management with and without service guarantees,” *7th ACM International Symposium on Modeling, analysis and simulation of wireless and mobile systems, MSWiM’04*, Venezia, Italy, October 2004.

[26] X. Duan, Z. Niu, and J. Zheng, “Utility optimization and fairness guarantees for multimedia traffic in the downlink of DS-CDMA systems,” *IEEE Global Telecommunications Conference, Globecom’03*, San Francisco, CA, USA, December 2003.

[27] L. Badia, M. Lindström, J. Zander, and M. Zorzi, “An economic model for the radio resource management in multimedia wireless systems,” *Elsevier Computer Communications Journal*, Vol. 27, N° 11, pp. 1056-1064, July 2004.

[28] P. Fishburn, “Utility theory for decision making,” *Wiley New York*, June 1970.

[29] L. Badia, M. Lindström, J. Zander, and M. Zorzi, “Demand and pricing effects on the radio resource allocation of multimedia communicationsystems,” *IEEE Global Telecommunications Conference, Globecom’03*, San Francisco, CA, USA, December 2003.

[30] L. Badia and M. Zorzi, “An analysis of multimedia services in next generation communication systems with QoS and revenue management,” *59th IEEE Vehicular Technology Conference,VTC-Spring’04*, Milan, Italy, May 2004.

[31] X. Bai, L. Bölöni, D. Marinescu, H. J. Siegel, R. Daley, and I.-J. Wang, “Are Utility, Price, and Satisfaction Based Resource Allocation Models Suitable for Large-Scale Distributed Systems?” *3rd International Workshop on Grid Economics and Business Models, GECON’06*, Singapore, May 2006.

[32] J. Von Neumannn and O. Morgenstern, “Theory of Games and Economics Behavior,” *Princeton Univ. Press*, Princeton, N.J., 1953.

[33] L. Giupponi, R. Agusti, J. Prez-Romero, and O. Sallent, “Towards Balancing User Satisfaction and Operator Revenue in Beyond 3G Cognitive Networks,” *15th IST Mobile & Wireless Communications Summit, Mobile Summit’06*, Myconos, Greece, June 2006.

[34] S. Pal, S. Das, and M. Chatterjee, “User-Satisfaction based Differentiated Services for Wireless Data Networks,” *IEEE International Conference on Communications, ICC’05*, Seoul, Korea, May 2005.

[35] L. A. Dasilva, “Pricing for QoS-Enabled networks: a survey,” *IEEE Communications Surveys and Tutorials*, Vol. 3, N° 2, pp. 2-8, Second Quarter 2000.

[36] C. Perkins, “IP mobility support for IPv4,” *IETF RFC 3344*, August 2002.

[37] D. Johnson, C. Perkins, and J. Arkko, “Mobility support in IPv6,” *IETF RFC 3775*, June 2004.

[38] M.P. Michael, “Energy awareness for mobile devices”, *Paper for Research Seminar on Energy Awareness*, University of Helsinki, Finland, 2005.

[39] 3GPP, "Quality of service (QoS) concept and architecture (R6)," *TS23.107 v6.3.0*, June 2005.

[40] A. Devlic and G. Jezic, "Location-Aware Information Services using User Profile Matching," *8th International Conference on Telecommunications, ConTEL'05*, Zagreb, Croatia, June 2005.

[41] Q. Song and A. Jamalipour, "A network selection mechanism for next generation networks," *IEEE International Conference on Communications, ICC'05*, Seoul, Korea, May 2005.

[42] R. Wakikawa et al., "Multiple Care of Addresses Registration", *IETF, Internet-Draft, draft-ietf-monami6-multiplecoa-00.txt*, June, 2006.

[43] G. Tsirtsis et al., "Flow Movement for Mobile IPv4", *IETF, Internet-Draft, draft-tsirtsis-mipv4-flowmove-00.txt*, August 2006.

[44] M. Zorzi and R. Rao, "Energy Efficiency of TCP in a local wireless environment", *Mobile Networks and Applications*, Vol. 6 , Issue 3, pp. 265-278, July 2001.

[45] H. Singh and S. Singh, "Energy consumption of TCP Reno, New Reno, and SACK in multi-hop wireless networks, " *ACM International Conference on Measurement and Modeling of Computer Systems, ACM SIGMETRICS'02*, Marina del Rey, CA, USA, June 2002.

[46] S. Agrawal and S. Singh, "An Experimental Study of TCP's Energy Consumption over a Wireless Link," *4th European Personal Mobile Communications Conference*, *EPMCC'01*, Vienna, Austria, February 2001.

[47] S. Cen, P. C. Cosman, and G. M. Voelker, "End-to-End differentiation of congestion and wireless losses," *IEEE/ACM Transactions on Networking (TON)*, Vol. 11, Issue 5, October 2003.

[48] S. Bregni, D. Caratti, F. Martignon, "Enhanced Loss Differentiation Algorithms for Use in TCP Sources over Heterogeneous Wireless Networks," *IEEE Global Communications Conference, Globecom'03*, San Francisco, CA, USA, December 2003.

[49] IEEE 802.11 WG, Local and metropolitan area networks-Specific requirements "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 1999.

[50] IEEE 802.11a WG, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications—Amendment 1: High-speed Physical Layer in the 5 GHz band," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 1999

[51] IEEE 802.11b WG, Local and metropolitan area networks- Supplement to 802.11-1999 "Part 11- Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher speed Physical Layer (PHY) extension in the 2.4 GHz band," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 1999

[52] IEEE 802.11g WG, Local and metropolitan area networks-Specific requirements "Part 11- Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 2003.

[53] IEEE 802.11n WG, Local and metropolitan area networks-Specific requirements "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 2009.

[54] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," *The Conference on Computer Communications - Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM'03*, San Francisco, CA, USA, March-April, 2003.

[55] M. Bottigliengo, C. Casetti, C.-F. Chiasserini, M. Meo, "Short-term Fairness for TCP Flows in *The Conference on Computer Communications - Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM'04*, Hong Kong, March, 2004.

[56] A. Chockalingam, M. Zorzi, V. Tralli, "Wireless TCP performance with link layer FEC/ARQ," *IEEE International Conference on Communications, ICC'99*, Vancouver, British Columbia, Canada, June 1999.

[57] S. Mascolo, M.Y. Sanadidi, C. Casetti, M. Gerla, and R. Wang, “TCP Westwood: End-to-End Congestion Control for Wired/Wireless Networks,” *Wireless Networks Journal*, vol. 8, pp. 467-479, September 2002.

[58] K. Xu, Y. Tian, and N. Ansari, “TCP-Jersey for Wireless IP Communications,” *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 4, May 2004.

[59] H. Balakrishnan and R. H. Katz, “Explicit Loss Notification and Wireless Web Performance,” *IEEE Global Internet Mini-Conference, Globecom’98*, Sydney, Australia, November 1998.

[60] S. Biaz and X. Wang, “Can ECN Be Used to Differentiate Congestion Losses from Wireless Losses?,” *Department of Computer Science and Software Engineering, Auburn University,* Technical Report CSSE04-04, May 2004.

[61] C. H. Ng, J. Chow, and L. Trajkovic, “Performance Evaluation of TCP over WLAN 802.11 with the Snoop Performance Enhancing Proxy,” *Opnetwork 2002*, Washington, DC, August 2002.

[62] S. Floyd, T. Henderson, Gurtov, and A., "The NewReno Modification to TCP's Fast Recovery Algorithm," *RFC 3782, IETF*, April 2004.

[63] The Network Simulation – NS-2, http://www.isi.edu/nsnam/ns/

[64] S. Biaz and N. H. Vaidya, “Distinguishing Congestion Losses from Wireless Transmission Losses: a Negative Result,” *Proc. of IEEE 7th International Conference on Computer Communications and Networks, IC3N’98*, Lafayette, LA, USA, October 1998.

[65] D. Barman and I. Matta, “Effectiveness of Loss Labeling in Improving TCP Performance in Wired/Wireless Networks,” *The 10th IEEE International Conference on Network Protocols, ICNP'02*, Paris, France, November 2002.

[66] A. Seddik-Ghaleb, Y. M Ghamri-Doudane, and S.-M. Senouci, “A Performance Study of TCP variants in terms of Energy Consumption and Average Goodput within a Static Ad Hoc Environment,” *ACM International Wireless Communications and Mobile Computing Conference, IWCMC’06*, Vancouver, Canada, July 2006.

[67] A. Seddik-Ghaleb, Y. M Ghamri-Doudane, and S.-M. Senouci, “Effect of Ad Hoc Routing Protocols on TCP Performance within MANETs,” *IEEE International Workshop on Wireless Ad-hoc and Sensor Networks, IWWAN’06*, New York, NY, USA, June 2006.

[68] V. Ramarathinam and M. A. Labrador, “Performance Analysis of TCP over Static Wireless ad hoc networks,” *15th International Conference on Parallel and Distributed Computing Systems, PDCS’02*, Cambridge, MA, USA, September 2002.

[69] B. Wang and S. Singh, “Computational energy cost of TCP,” *The Conference on Computer Communications - Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM’04*, Hong Kong, March, 2004.

[70] K. Fall and S. Floyd, “Simulation-based comparison of Tahoe, Reno, and sack TCP,” *ACM Computer Communications Review*, vol. 5, no. 3, July 1996.

[71] L. S. Brakmo, S. W. O’Malley, and Larry L. Peterson, “TCP Vegas: New Techniques for Congestion Detection and Avoidance,” *ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM 1994*, London, UK, August 1994.

[72] Dummynet, http://info.iet.unipi.it/~luigi/dummynet/

[73] L. A. Grieco and S. Mascolo, "TCP Westwood and Easy RED to Improve Fairness in High-Speed Networks," *7th IFIP/IEEE International Workshop on Protocols For High-Speed Networks, PfHSN’02*, Berlin, Germany, April 2002.

[74] S. Mascolo, C. Casetti, M. Gerla, M. Y. Sanadidi, and R. Wang, "TCP Westwood: Bandwidth estimation for enhanced transport over wireless links," *7th Annual International Conference on Mobile Computing and Networking, ICMCN’01*, Rome, Italy, July 2001.

[75] R. Wang, M. Valla, M. Y. Sanadidi, B. K. F. Ng, and M. Gerla, "Efficiency/Friendliness Tradeoffs in TCP Westwood," *7th IEEE Symposium on Computers and Communications, ISCC’02,* Taormina, Italy, July 2002.

[76] M. Mancuso, "A Novel Scheme of Loss Differentiation and Adaptive Segmentation to Enhance TCP Performance over Wireless Networks," *Technical Report, Politecnico di Milano, Dept. of Electronics and Information*, July 2002.

[77] G. Xylomenos, G.C. Polyzos, Mahonen P., and M. Saaranen, "TCP Performance Issues over Wireless Links," *IEEE Communications Magazine*, vol. 39, no. 4, pp. 52-58, April 2001.

[78] S. Lohier, Y. Ghamri-Doudane, and G. Pujolle, "Cross-Layer Loss Differentiation Algorithms to improve TCP Performances in WLANs," *11th IFIP International Conference on Personal Wireless Communications, PWC'06*, Albacete, Spain, September 2006.

[79] Y. Tobe, H. Aida, Y. X Aida, and H. Tokuda, "Detection of Congestion Signals from Relative One-Way Delay," *Transactions of Information Processing Society of Japan (IPSJ)*, Vol. 42, No. 12, pp. 3222-3230, December 2001.

[80] S. Cen, P. C. Cosman, and G. M. Voelker, "End-to-End Differentiation of Congestion and Wireless Losses," *SPIE Multimedia Computing and Networking, MMCN'02*, San Jose, CA, USA, January 2002.

[81] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535-547, Mar. 2000.

[82] D. Malone, K. Duffy, and D. J. Leith, "Modeling the 802.11 Distributed Coordination Function in Non-saturated Heterogeneous Conditions," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 159-172, Feb. 2007.

[83] F.-Y. Hung and I. Marsic, "Analysis of Non-Saturation and Saturation Performance of IEEE 802.11 DCF in the Presence of Hidden Stations," *66th IEEE Vehicular Technology Conference, VTC-Fall'07*, Baltimore, MD, USA, Sep. 2007.

[84] P. E. Engelstad and O. N. Osterbo, "Analysis of the Total Delay of IEEE 802.11e EDCA and 802.11 DCF," *IEEE International Conference on Communications, ICC'06*, Istanbul, Turkey, June 2006.

[85] W. Zhang, J. Sun, J. Liu, and H.-B. Zhang, "Performance analysis of IEEE 802.11e EDCA in wireless LANs," *Journal of Zhejiang University*, vol. 8, no. 1, pp. 18-23, Jan. 2007.

[86] X. Ling, K.-H. Liu, Y. Cheng, X. Shen, and J. W. Mark, "A Novel Performance Model for Distributed Prioritized MAC Protocols," *IEEE Global Communications Conference, Globecom'07*, Washington, DC, USA, Nov. 2007.

[87] Y. Lin and V. W. S. Wong, "Saturation throughput of IEEE 802.11e EDCA based on mean value analysis," *IEEE Wireless Communications and Networking Conference. WCNC'06*, Las vegas, NV, USA, Mar. 2006.

[88] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1506-1515, Jul. 2005.

[89] J. Hui and M. Devetsikiotis, "A Unified Model for the Performance Analysis of IEEE 802.11e EDCA," *IEEE Transactions on Communications*, vol. 53, no. 9, pp. 1498-1510, Sep. 2005.

[90] K. Xu, Q. Wang, and H. Hassanein, "Performance analysis of differentiated QoS supported by IEEE 802.11e enhanced distributed coordination function (EDCF) in WLAN," *IEEE Global Communications Conference, Globecom'03*, San Francisco, CA, USA, Dec. 2003.

[91] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 5, pp. 917-928, Jun. 2004.

[92] I. Inan, F. Keceli, and E. Ayanoglu, "Modeling the 802.11e Enhanced Distributed Channel Access Function," *IEEE Global Communications Conference, Globecom'07*, Washington, DC, USA, Nov. 2007.

[93] L. Xiong and G. Mao, "Saturated Throughput Analysis of IEEE 802.11e EDCA," *Computer Networks*, vol. 51, no. 11, pp. 3047-3068, Aug. 2007.

[94] H. Wu, X. Wang, Q. Zhang, and X. Shen, "IEEE 802.11e Enhanced Distributed Channel Access (EDCA) Throughput Analysis," *IEEE International Conference on Communications, ICC'06*, Istanbul, Turkey, June 2006.

[95] T.-C. Tsai and M.-J. Wu, "An analytical model for IEEE 802.11e EDCA," *IEEE International Conference on Communications, ICC'05*, Seoul, Korea, May 2005.

[96] Z.-n. Kong, D. H. K. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 10, pp. 2095-2106, Dec. 2004.

[97] Y. Yan and C. Pan, "An Improved Analytical Model for IEEE802.11e Enhanced Distributed Channel Access," *International Symposium on Information Technology Convergence, ISITC'07*, Jeonju, Korea, Nov. 2007.

[98] Z. Tao and S. Panwar, "Throughput and delay analysis for the IEEE 802.11e enhanced distributed channel access," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 596-603, Apr. 2006.

[99] J. W. Tantra, C. H. Foh, and A. B. Mnaouer, "Throughput and delay analysis of the IEEE 802.11e EDCA saturation," *IEEE International Conference on Communications, ICC'05*, Seoul, Korea, May 2005.

[100] P. E. Engelstad and O. N. Osterbo, "Queueing Delay Analysis of IEEE 802.11e EDCA," *3rd IFIP Annual Conference on Wireless On demand Network Systems and Services, WONS'06*, Les Ménuires, France, Jan. 2006.

[101] W. Zhang, J. Sun, J. Liu, and H.-B. Zhang, "Performance analysis of IEEE 802.11e EDCA in wireless LANs," *Journal of Zhejiang University*, vol. 8, no. 1, pp. 18-23, Jan. 2007.

[102] N. Chendeb, Y. Ghamri-Doudane, and B. El-Hassan, "Effect of Transmission Opportunity Limit on Transmission Time Modeling in 802.11e," *7th IEEE International Workshop on IP Operations and Management, IPOM'07*, San José, CA, USA, Oct.-Nov. 2007.

[103] IEEE 802.11e WG, 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 (QoS) Enhancements," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, 2005.

[104] IEEE 802.11e implementation for NS-2, http://www.tkn.tu-berlin.de/research/802.11e_ns2/

[105] S. Garg and M. Kappes, "Admission control for VoIP traffic in IEEE 802.11 networks," *IEEE Global Communications Conference, Globecom'03*, San Francisco, CA, USA, Dec. 2003.

[106] D. Gu and J. Zhang, "A new measurement-based admission control method for IEEE802.11 wireless local area networks," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC'03*, Beijing, China, Sep. 2003.

[107] L. Zhang and S. Zeadally, "HARMONICA: enhanced QoS support with admission control for IEEE 802.11 contention-based access," *10th IEEE Real-Time and Embedded Technology and Applications Symposium, RTAS'04*, Toronto, Canada, May 2004.

[108] H. Yoon, J. Kim, and D. Shin, "Dynamic admission control in IEEE 802.11e EDCA-based wireless home network," *IEEE Consumer Communications and Networking Conference, CCNC'06*, Las Vegas, NV, USA, Jan. 2006.

[109] R. G. Garroppo, L. Tavanti, S. Lucetti, and S. Giordano, "Extension and Application of the Network Utilization Characteristic Metric to IEEE 802.11e," *IEEE International Conference on Communications, ICC'07*, Glascow, Scotland, Jun. 2007.

[110] Y.-L. Kuo, C.-H. Lu, E. H. K. Wu, and G.-H. Chen, "An admission control strategy for differentiated services in IEEE 802.11," *IEEE Global Communications Conference, Globecom'03*, San Francisco, CA, USA, Dec. 2003.

[111] J. Zhu and A. O. Fapojuwo, "A new call admission control method for providing desired throughput and delay performance in IEEE802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, Vol. 6, N° 2, pp. 701-709, Feb. 2007.

[112] S. Oh, J. Shin, D. Kwak, and C. Kim, "A Novel Call Admission Control Scheme for the IEEE 802.11e EDCA," *International Conference on Advanced Computing Technologies, ICACT*, Hyderabad, India, Dec. 2008.

[113] X. Chen, H. Zhai, X. Tian, and Y. Fang, "Supporting QoS in IEEE 802.11e wireless LANs," *IEEE Transactions on Wireless Communications*, Vol. 5, N° 8, pp. 2217-2227, Aug. 2006.

[114] A. Bai, T. Skeie, and P. E. Engelstad, "A Model-Based Admission Control for 802.11e EDCA using Delay Predictions," *26th IEEE International Performance Computing and Communications Conference, IPCCC*, New Orleans, LA, USA, Apr. 2007.

[115] B. Bellata, "Flow-level QoS guarantees in IEEE 802.11e-EDCA based WLANs," *Ph.D. thesis, Pompeu Fabra Univesity*, Barcelona, Spain, Dec. 2006.

[116] B. Bellalta, C. Cano, M. Oliver, and M. Meo, "Modeling the IEEE 802.11e EDCA for MAC Parameter Optimization," *4th International Working Conference on Performance Modeling and Evaluation of Heterogeneous Networks, Hets-Nets*, Bradford, UK, Sep. 2006.

[117] D. Pong and T. Moors, "Call admission control for IEEE 802.11 contention access mechanism," *IEEE Global Communications Conference, Globecom'03*, San Francisco, CA, USA, Dec. 2003.

[118] M. El Masri, G. Juanole, and S. Abdellatif, "Hybrid Admission Control Algorithm for IEEE 802.11e EDCA: Analysis," *7th International Conference on Networking, ICN,* Cancun, Mexico, Apr. 2008.

[119] B. Bensaou, Z.-N. Kong, and D. H. K. Tsang, "A Measurement-Assisted, Model-Based Admission Control Algorithm for IEEE 802.11e," *9th International Symposium on Parallel Architectures, Algorithms, and Networks, I-SPAN*, Sydney, Australia, May 2008.

[120] IEEE 802.16j WG, Local and metropolitan area networks-Specific requirements "Part 16: Air Interface for Broadband Wireless Access Systems—Amendment 1: Multihop Relay Specification," *IEEE Standard for Information technology -- Telecommunications and information exchange between systems*, June 2009.

[121] IEEE 802.16m WG, Local and metropolitan area networks-Specific requirements "Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems—Advanced Air Interface," *Draft Amendment to IEEE Standard for Local and Metropolitan Area Networks*, Aug. 2010.

[122] K. Byung-Gook and L. Jang-Won, "Opportunistic Power Scheduling for OFDMA Cellular Networks with Scheduling at Relay Stations," *IEEE Wireless Communications & Networking Conference, WCNC'09*, Budapest, Hungary, Apr. 2009.

[123] M. Kaneko and P. Popovski, "Radio Resource Allocation Algorithm for Relay-Aided Cellular OFDMA System," *IEEE International Conference on Communications, ICC'07*, Glascow, Scotland, Jun. 2007.

[124] H. Chi-Yao and P. Ai-Chun, "3-Approximation algorithm for joint routing and link scheduling in wireless relay networks," *IEEE Transactions on Wireless Communications*, Vol. 8, N° 2, pp. 856-861, Feb. 2009.

[125] P. Tarasak and S. Sumei, "Joint Cooperative Diversity and Proportional Fair Scheduling in OFDMA Relay Systems," *68th IEEE Vehicular Technology Conference, VTC-Fall'08*, Calgary, Canada, Sep. 2008.

[126] G. Senarath et al. "Multi-hop Relay System Evaluation Methodology," *Contribution to IEEE 802.16j WG, IEEE 802.16j-06/013r3*, Feb. 2007.

[127] T. C.-Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, Vol. 25, N° 2, pp.328-339, Feb. 2007.

[128] L. You, M. Song, and J. Song, "Cross-layer optimization for fairness in OFDMA cellular networks with fixed relays," *IEEE Global Communications Conference, Globecom'08*, New-Orleans, LA, USA, Nov. 2008.

[129] K. Sundaresan and S. Rangarajan, "Efficient algorithms for leveraging spatial reuse in OFDMA relay networks," *28th IEEE Conference on Computer Communications, INFOCOM'09*, Apr. 2009.

[130] M. Kaneko and P. Popovski, "Adaptive resource allocation in cellular OFDMA system with multiple relay stations," *65th IEEE Vehicular Technology Conference, VTC-Spring'07*, Dublin, Ireland, Apr. 2007.

[131] B.-G. Kim and J.-W. Lee, "Opportunistic power scheduling for OFDMA cellular networks with scheduling at relay stations," *IEEE Wireless Communications & Networking Conference, WCNC'09*, Budapest, Hungary, Apr. 2009.

[132] Community Database on Accidents on the Roads in Europe (CARE), http://europa.eu.int/comm/transport/care/index en.htm

[133] A. Nandan, S. Das, G. Pau, M. Gerla, M.Y. Sanadidi, "Co-operative Downloading in Vehicular Ad-Hoc Wireless Networks," *2nd Annual Conference on Wireless On-demand Network Systems and Services, WONS'05*, St. Moritz, Switzerland, January 2005.

[134] I. Broustis and M. Faloutsos, "Routing in Vehicular Networks: Feasibility, Security, and Modeling Issues," *International Journal of Vehicular Technology,* Vol. 08, March 2008.

[135] DSRC (Dedicated Short Range Communications), *http://www.leearmstrong.com/DSRC/DSRCHomeset.htm*

[136] VII (Vehicle Infrastructure Initiative), *http://www.vehicle-infrastructure.org*

[137] NoW (Network-on-Wheels), *http://www.network-on-wheels.de*

[138] CVIS (Cooperative Vehicle-Infrastructure Systems), *http://www.cvisproject.org*

[139] COMeSafety (Communications for eSafety), *http://www.comesafety.org*

[140] C2C-CC (Car2Car Communication Consortium), *http://www.car-to-car.org*

[141] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET Routing on City Roads using Real-Time Vehicular Traffic Information," *IEEE Transactions on Vehicular Technology*, Vol. 58, N° 7, pp. 3609-3626, September 2009.

[142] M. Nekovee, "Sensor networks on the road: the promises and challenges of vehicular adhoc networks and vehicular grids," *Workshop on Ubiquitous Computing and e-Research, UK-UbiNet'05*, Edinburgh, UK, May 2005.

[143] J. Blum, A. Eskandarian and L. Hoffmman, "Challenges of inter-vehicle ad hoc networks," I, Vol. 5, No 4, pp. 347-351, December 2004.

[144] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "Analytical models for information propagation in vehicle-to-vehicle networks," *60th IEEE Vehicular Technology Conference, VTC-Fall'04*, Los Angeles, CA, USA, September 2004.

[145] A. K. Ziliaskopoulos and J. Zhang, "A Zero Public Infrastructure Vehicle Based Traffic Information System," *Transportation Research Board Meeting Annual Meeting*, *TRB'03*, National Academies, Washington, D.C., January 2003.

[146] L. Wischhof, A. Ebner, H. Rohling, M. Lott and R. Halfmann, "SOTIS – A Self-Organizing Traffic Information System," *57th IEEE Vehicular Technology Conference, VTC-Spring'03*, Jeju Island, Korea, April 2003.

[147] E-Road Project, *http://discolab.rutgers.edu/traffic/*

[148] T. Nadeem, S. Dashtinezhad, C. Liao and L. Iftode, "TrafficView: Traffic Data Dissemination using Car-to-Car Communication," in *ACM Mobile Computing and Communications Review (MC2R)*, Vol. 8, N° 3, pp. 6-19, July 2004.

[149] Qualnet simulator, *http://www.scalable-networks.com*

[150] J. Haerri, M. Fiore, F. Filali, C. Bonnet, "VanetMobiSim: generating realistic mobility patterns for VANETs", *ACM Int. Wrkshp on Vehicular Ad Hoc Networks, VANET'06*, Los Angeles, CA, USA, Sep. 2006.

[151] J. Blum, A. Eskandarian and L. Hoffmman, "Challenges of intervehicle ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No 4, pp. 347-351, Dec. 2004.

[152] G. Korkmaz, E. Ekici and F. Özgüner, "Urban multihop broadcast protocol for inter-vehicle communication systems," *ACM Workshop on Vehicular Ad Hoc Networks, VANET 2004,* Philadelphia, PA, USA, Oct. 2004.

[153] H. Wu, R. Fujimoto, R. Guensler, and M. Hunter, "MDDV: a mobility centric data dissemination algorithm for vehicular networks," *1st ACM Workshop on Vehicular Ad Hoc Networks, VANET'04*, Philadelphia, PA, USA, Sep. 2004.

[154] D. Sormani, G. Turconi, P. Costa, D. Frey, M. Migliavacca, and L. Mottola, "Towards Lightweight Information Dissemination in Inter-Vehicular Networks," *3rd ACM International Workshop on Vehicular Ad Hoc Networks, VANET'06,* Los Angeles, CA, USA, Sep. 2006.

[155] L. Wischhof, A. Ebner and H. Rohling, "Information Dissemination in Self-organizing Inter-vehicle Networks," *IEEE Transactions on Intelligent Transportation Systems,* Vol. 6, No. 1, pp. 90–101, Mar. 2005.

[156] A. Benslimane, "Optimized Dissemination of Alarm Messages in Vehicular Ad-hoc Networks (VANET)," *7th IEEE International Conference on High Speed Networks and Multimedia Communications, HSNMC'04*, Toulouse, France, Jun. 2004.

[157] J. P. Hubaux, T. Gross, J. Y. Le Boudec, M. Vetterli, "Towards Self-Organized Mobile Ad Hoc Networks: The Terminodes Project," *IEEE Communications Magazine*, Vol. 39, No. 1, pp. 118-124, Jan. 2001.

[158] W. Sun, H. Yamaguchi, K. Yukimasa and S. Kusumoto "GVGrid: A QoS Routing Protocol for Vehicular Ad Hoc Networks," *IEEE International Workshop on Quality of Service, IWQoS'06*, New Haven, CT, USA, Jun. 2006.

[159] C. Lochert, H. Hartenstein, J. Tian, D. Herrmann, H. Füßler, M. Mauve, "A Routing Strategy for Vehicular Ad Hoc Networks in City Environments," *IEEE Intelligent Vehicles Symposium, IV'03*, Columbus, OH, USA, Jun. 2003.

[160] B.-C. Seet, G. Liu, B.-S. Lee, C. H. Foh, K. J. Wong, K.-K. Lee, "A-STAR: A Mobile Ad Hoc Routing Strategy for Metropolis Vehicular Communications," *3rd IFIP International Conferences on Networking, NETWORKING'04*, Athens, Greece, May 2004.

[161] J. Zhao and G. Cao, "VADD: Vehicle-assisted data delivery in vehicular ad hoc networks" *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 3, May. 2008.

[162] Yong, Chen and Li, "A Static-Node Assisted Adaptive Routing Protocol in Vehicular Networks," *4th ACM international workshop on Vehicular ad hoc networks, VANET'07,* Montréal, QC, Canada, Sep. 2007.

[163] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," *6th ACM/IEEE International Annual Conference on Mobile Computing and Networking, MOBICOM'00,* Boston, MA, USA, Aug. 2000.

[164] D. Niculescu and B. Nath, "Trajectory based forwarding and its applications," *9th ACM International Annual Conference on Mobile Computing and Networking, MOBICOM'03*, San Diego, CA, USA, Sep. 2003.

[165] The CitySense Sensor Network Project, http://www.citysense.net

[166] J. Davis, A. Fagg, and B. Levine, "Wearable Computers as Packet Transport Mechanisms in Highly-Partitioned Ad-Hoc Networks," *5th International Symposium on Wearable Computing, ISWC'01*, Zurich, Switzerland, Oct. 2001.

[167] Y. Ko and N. Vaidya, "Locaton-aided routing (LAR) in mobile ad hoc networks," *4th ACM/IEEE International Annual Conference on Mobile Computing and Networking, MOBICOM'98*, Dallas, TX, USA, Aug. 1998.

[168] C. Lochert, B. Scheuermann, M. Caliskan, M. Mauve, "The feasibility of information dissemination in vehicular ad-hoc networks," *4th Annual Conference on Wireless On demand Network Systems and Services, IEEE WONS '07,* Obergurgl, Austria, Jan. 2007.

[169] J. Nzoonta and C. Borcea, "STEID: A protocol for emergency information dissemination in vehicular networks," *NJIT Technical Report*, 2006.

Chapter 8. List of Publications

8.1 Patents (3)

[P1] M. Jerbi, S.-M. Senouci, Y. Gourhant, Y. M. Ghamri-Doudane, "Procédé de diffusion de paquets de données dans un réseau de nœuds mobiles et terminal associe", 16/11/2007, INPI FR # 07 59115. International Patent Extension (Ongoing).

[P2] M. Jerbi, S.-M. Senouci, Y. Gourhant, Y. M. Ghamri-Doudane, "Procédé de diffusion de paquets de données dans un réseau de nœuds mobiles et terminal associe", 26/10/2007, INPI FR 07 06771. International Patent Extension (Ongoing).

[P3] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, "Procédé de contrôle d'émission de paquets TCP dans un réseau ad hoc, programme d'ordinateur et station émettrice/réceptrice associés", 11/05/2007, INPI FR # 07 03391. International Patent Extension (Ongoing).

8.2 Book Chapters (6)

[P4] M. Jerbi, M. Cherif, S.-M. Senouci, Y. M. Ghamri-Doudane, « Vehicular Communications Networks: Current Trends and Challenges », In "Handbook of Research on Next Generation Mobile Networks and Ubiquitous Computing", IGI-Global, to appear in 2010.

[P5] Y. M. Ghamri-Doudane, A. Munaretto, H Chaouchi, « Quality of Service in WiFi », in « Management, Control and Evolution of IP Networks », iSTE Ed., Jan. 2007.

[P6] Y. M. Ghamri-Doudane, « New Approaches for the Management and Control of IP Networks », in « Management, Control and Evolution of IP Networks », iSTE Ed., Jan. 2007.

[P7] N. Achir, Y. M. Ghamri-Doudane, M. S. P. Fonseca, « Un exemple de réseaux actifs : le projet AMARRAGE (in French) », chapitre dans le traité IC2 « Intelligence dans les réseaux », Hermès Science Ed., Jul. 2005.

[P8] Y. M. Ghamri-Doudane, A. Munaretto, H Chaouchi, « La qualité de service dans l'IEEE 802.11 (in French) », chapitre dans le traité IC2 « L'Internet Ambiant », Hermès Science Ed., Oct. 2004.

[P9] Y. M. Ghamri-Doudane, « Les nouvelles approches pour la gestion et le contrôle des réseaux IP nouvelle génération (in French) », chapitre dans le traité IC2 « Evolution du monde IP », Hermès Science Ed., Oct. 2004.

8.3 Articles in Refereed International Journals (11)

[P10] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, N. Agoulmine, « Measurement of TCP Computational and Communication Energy Cost in MANETs », to appear in Elsevier Pervasive and Mobile Computing (acceptance notification received in Jun. 2010).

[P11] M. Jerbi, S.-M. Senouci, T. Rasheed, Y. M. Ghamri-Doudane, « Towards Efficient Geographic Routing in Urban Vehicular Networks », IEEE Transactions on Vehicular Technology, Vol. 58, n°. 9, pp. 5048-5059, Nov 2009.

[P12] Q.-T. Nguyen-Vuong, N. Agoulmine, Y. M. Ghamri-Doudane, « A user-centric and context-aware solution to interface management and access network selection in heterogeneous wireless environments », Elsevier Computer Networks (ComNets) Journal, Vol. 52, n°. 18, Pages 3358-3372, Dec. 2008.

[P13] M. Jerbi, S.-M. Senouci, Y. M. Ghamri-Doudane, « Towards Efficient Routing in Vehicular Ad Hoc Networks », Ubiquitous Computing and Communication (UBICC) Journal, special issue on UBIROADS, March 2008.

[P14] S. Lohier, Y. M. Ghamri-Doudane and G. Pujolle, « Cross-Layer Design to Improve Elastic Traffic Performance in WLANs », ACM/Wiley International Journal of Network Management, May 2007.

[P15] S. Lohier, Y. M. Ghamri-Doudane and G. Pujolle, « Cross-layer loss differentiation algorithms to improve TCP Performance in WLANs », Springer Telecommunications Systems Journal , Volume 36, Numbers 1-3, pp. 61-72, November 2007.

[P16] Q.-T. Nguyen-Vuong, N. Agoulmine, Y. M. Ghamri-Doudane, « Terminal-Controlled Mobility Management in Heterogeneous Wireless Networks », IEEE Communications Magazine, Vol. 45 – n°4, pp. 122-129, April 2007.

[P17] K. Haddadou, S. Ghamri-Doudane, Y. M. Ghamri-Doudane, N. Agoulmine, « Designing Scalable On-Demand Policy-Based Resource Allocation in IP Networks », IEEE Communications Magazine, Vol. 44 – n°3, pp. 142-149, March 2006.

[P18] T. M. T. Nguyen, N. Boukhatem, Y. M. Ghamri-Doudane and G. Pujolle, « COPS-SLS: A Service Level Negotiation Protocol for the Internet », IEEE Communications Magazine, Vol. 40 – n°5, pp. 158-165, May 2002.

[P19] N. Achir, N. Agoulmine, M.S. P. Fonseca, Y. M. Ghamri-Doudane and A. Mehaoua, « Active Networking System Evaluation: A Practical Experience », Networking and Information Systems Journal, Vol. 3 - No. 3, pp. 431-448, Hermès Oxford, 2000.

[P20] K. Boussetta, Y. M. Ghamri-Doudane and A-L. Beylot, « Simulcast-based DCA scheme for unicast and multicast communications in LEO satellite networks », Networking and Information System Journal, Vol. 3 – n°5-6, pp. 853-864, Hermès Oxford, 2000.

8.4 ARTICLES IN REFEREED REGIONAL JOURNALS (2)

[P21] D. Vieira, C. A. V. Melo, A. Bezerra, Y. M. Ghamri-Doudane, N. L. S. Fonseca « A Content-Oriented Web Cache Policy under P2PVideo Distribution Systems (in Portuguese) », to appear in IEEE Latin America Transactions (acceptance notification received in Apr. 2010).

[P22] M. S. P. Fonseca, N. Agoulmine, Y. M. Ghamri-Doudane, N. Achir, and G. Pujolle, « Une nouvelle architecture de contrôle des réseaux DiffServ basée sur la technologie réseaux actifs et les politiques », In Annales des télécommunications, Vol. 59, n°5-6, May-June 2004.

8.5 MANUSCRIPTS UNDER SUBMISSION IN REFEREED INTERNATIONAL JOURNALS (3)

[P23] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, N. Agoulmine, « Coupling Loss and Delay Differentiation to Enhance TCP Performance within Wireless Multi-hop Ad-hoc Networks », under review.

[P24] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, N. Agoulmine, « Towards Real-Time and Multimedia Application Support in 802.11e EDCA: An Efficient Model-Based Admission Control Algorithm », under review.

[P25] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, N. Agoulmine, « An Accurate Analytical Model for 802.11e EDCA under Different Traffic Conditions with Contention Free Bursting », under review.

8.6 ARTICLES FOR THE LAYMAN (2)

[P26] Q.-T. Nguyen Vuong, Y. M. Ghamri-Doudane, N. Agoulmine, « Projet SUMO : Fourniture de services multimédia sans couture dans des environnements sans fil hétérogènes », revue Techniques de l'Ingénieur, Lettre « réseaux sans fil », N°10, Nov.-Dec. 2007.

[P27] M. Jerbi, S.-M. Senouci, Y. M. Ghamri-Doudane « GyTAR: protocole de routage géographique pour les réseaux de véhicules », revue Techniques de l'Ingénieur, Lettre « réseaux sans fil », N°5, Jan.-Feb. 2007.

8.7 ARTICLES IN REFEREED INTERNATIONAL CONFERENCES – FULL PAPERS (37)

[P28] M. Abdennebi, Y. M. Ghamri-Doudane, Y. Li, « Adaptive CAC with Traffic Flows Classification for IEEE 802.16 Networks », IFIP Wireless Days, WD'10, (Venice, Italy), Oct. 2010.

[P29] I. Salhi, Y. M. Ghamri-Doudane, E. Livolant, S. Lohier, « CoZi: basic Coding for better Bandwidth Utilization in ZigBee Sensor Networks », IEEE Global Communications Conference, GLOBECOM'10, (Miami, Florida, U.S.A), Dec. 2010.

[P30] I. Salhi, Y. M. Ghamri-Doudane, S. Lohier, G. Roussel, « Network Coding for Event-Centric Wireless Sensor Networks », IEEE International Conference on Communications, ICC'10, (Cape Town, South Africa), May-Jun. 2010.

[P31] A. Ben Nacef, S.-M. Senouci, Y. M. Ghamri-Doudane, A.-L. Beylot, « Enhanced relay selection decision for cooperative communication in energy constrained networks », IFIP Wireless Days, WD'09, (Paris, France), Dec. 2009.

[P32] S. Lohier, A. Rachedi, Y. M. Ghamri-Doudane, «A Cost Function for QoS-Aware Routing in Multi-tier Wireless Multimedia Sensor Networks », IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, MMNS'09, (Venice, Italy), Oct. 2009.

[P33] A. Bezerra, C. Melo, D. Vieira, Y. M. Ghamri-Doudane, N.L.S. Fonseca, « A Content-Oriented Web Cache Policy », IEEE Latin-American Conference on Communications, LatinCom'09, (Medelin, Colombia), Sept. 2009.

[P34] N. Chendeb-Taher, Y. M. Ghamri-Doudane, B. El Hassan, N. Agoulmine, « An Efficient Model-Based Admission Control Algorithm to Support Voice and Video services in 802.11e WLANs», IEEE Global Information Infrastructure Symposium, GIIS'09, (Hammamet, Tunisia), Jun. 2009.

[P35] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, "Computational Energy Cost of TCP in MANETs", ACS/IEEE International Conference on Computer Systems and Applications, AICCSA'09, (Rabat, Morocco), May 2009.

[P36] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, "A Complete and Accurate Analytical Model for 802.11e EDCA under Saturation Conditions", ACS/IEEE International Conference on Computer Systems and Applications, AICCSA'09, (Rabat, Morocco), May 2009.

[P37] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, "TCP WELCOME: TCP variant for Wireless Environment, Link losses, and COngestion packet loss ModEls", IEEE International Conference on COMmunication Systems and NETworkS, COMSNETS'09, (Bangalore, India), Jan. 2009.

[P38] M. Jerbi, A.-L. Beylot, S.-M. Senouci, Y. M. Ghamri-Doudane, "Geo-localized Virtual Infrastructure for VANETs: Design and Analysis", IEEE Global Communications Conference, GLOBECOM'08, (New Orleans, Louisiana, U.S.A), Dec. 2008.

[P39] M. Jerbi, A.-L. Beylot, S.-M. Senouci, Y. M. Ghamri-Doudane, "Geo-localized Virtual Infrastructure for VANETs: Design and Analysis", IEEE Intelligent Transportation Systems - Telecommunications, ITST'08, (Phuket, Thailand), Oct. 2008.

[P40] K. Haddadou, S. Ghamri-Doudane, Y. M. Ghamri-Doudane, N. Agoulmine, « A Dimensioning and Deployment Tool for on Demand Policy-based Resource Management System", IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, MMNS'08, (Samos Island, Greece), Sept. 2008.

[P41] Q-T- Nguyen Vuong, Y. M. Ghamri-Doudane, N. Agoulmine, "On Utility Models for Access Network Selection in Wireless Heterogeneous Networks", IEEE/IFIP Network Operations & Management Symposium, NOMS'08, (Salvador, Brazil), Apr. 2008.

[P42] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, "Effect of Transmission Opportunity Limit on Transmission Time Modeling in 802.11e", IEEE International Workshop on IP Operations and Management, IPOM'07, (San José, California, USA), Oct.- Nov. 2007.

[P43] M. Jerbi, S.-M. Senouci, T. Rasheed, Y. M. Ghamri-Doudane, «An Infrastructure-Free Traffic Information System for Vehicular Networks », IEEE International Symposium on Wireless Vehicular Communications, WiVec'07, (Baltimore, Maryland, USA), Sept.-Oct. 2007.

[P44] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, « Emulating End-to-End Losses and Delays for Ad Hoc Networks », IEEE International Conference on Communications, ICC'07, (Glasgow, Scotland), Jun. 2007.

[P45] M. Jerbi, S.-M. Senouci, Y. M. Ghamri-Doudane, R. Meraihi, « An Improved Vehicular Ad Hoc Routing Protocol for City Environments », IEEE International Conference on Communications, ICC'07, (Glasgow, Scotland), Jun. 2007.

[P46] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « Link Available Bandwidth Monitoring for QoS Routing with AODV in Ad Hoc Networks », IFIP/IEEE International Conference on Management of Multimedia Networks and Services, MMNS'06, (Dublin, Ireland), Oct. 2006.

[P47] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « Cross-Layer Loss Differentiation Algorithms to Improve TCP Performance in WLANs », IFIP Symposium on Personal and Wireless Communications, PWC'06, (Albacete, Spain), Sept. 2006.

[P48] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, « A Performance Study of TCP variants (Tahoe, Reno, New Reno, SACK, Vegas, and Westwood) in terms of Energy Consumption and Average Goodput within a Static Ad Hoc Environment », ACM International Wireless Communications and Mobile Computing Conference, IWCMC'06, (Vancouver, Canada), Jul. 2006.

[P49] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « MAC-layer Adaptation to Improve TCP Flow Performances in 802.11 Wireless Networks », IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob'06, (Montreal, Canada), Jun. 2006.

[P50] Y. M. Ghamri-Doudane, S. Senouci, N. Agoulmine, « P-SEAN: A Framework for Policy-Based Server Election in Ad hoc Networks », IEEE/IFIP Network Operations & Management Symposium, NOMS'06, (Vancouver, Canada), Apr. 2006.

[P51] K. Haddadou, S. Ghamri-Doudane, Y. M. Ghamri-Doudane, N. Agoulmine, « On Scalability of Dynamic Resource Allocation in Policy-enabled Networks: Practical and Analytical Evaluations », IEEE Global Telecommunications Conference, Globecom'05, (St. Louis, Missouri, U.S.A.), Nov.-Dec. 2005, Vol. 1, pp. 419-424, IEEE publisher.

[P52] E. Lehtihet, H. Derbel, N. Agoulmine, Y. M. Ghamri-Doudane, S. Van der Meer, « Initial Approach Toward Self-Configuration and Self-Optimization in IP Networks », IFIP/IEEE International Conference on Management of Multimedia Networks and Services, MMNS'05, (Barcelona, Spain), Oct. 2005, Lecture Notes in Computer Science, n°3754, pp. 371-382, Springer Verlag.

[P53] K. Haddadou, Y. M. Ghamri-Doudane, S. Ghamri-Doudane, N. Agoulmine, « Proactive Two-Tier Bandwidth Brokerage for on-demand Policy-based Resource Allocation in Stateless IP Networks », IFIP/IEEE International Conference on Management of Multimedia Networks and Services, MMNS'05, (Barcelona, Spain), Oct. 2005, Lecture Notes in Computer Science, n°3754, pp. 350 - 361, Springer Verlag.

[P54] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « The Benefits of a Cross-Layer Approach for TCP Performance Improvements in WLANs », IEEE workshop on Applications and Services in Wireless Networks, ASWN'05, (Paris, France), Jun. 2005.

[P55] K. Haddadou, Y. M. Ghamri-Doudane, M. Girod-Genet, A. Meddahi, L. Bernard, G. Vanwormhoudt, H. Afifi, N. Agoulmine, « Toward Feasibility and Scalability of Session Initiation and Dynamic QoS Provisioning in Policy-enabled Networks », IFIP Networking Conference, Networking'05, (Waterloo, Canada), May 2005. Lecture Notes in Computer Science, n°3462, pp. 277-288, Springer Verlag.

[P56] S. Lohier, S. Senouci, Y. M. Ghamri-Doudane, G. Pujolle, « A Reactive QoS Routing Protocol for Ad Hoc Networks », European Symposium on Ambient Intelligence, EUSAI'03, (Eindhoven, Netherlands), Nov. 2003, Lecture Notes in Computer Science, n°2875, pp. 27-41, Springer Verlag.

[P57] P. Simons, Y. M. Ghamri-Doudane, S. Senouci, F. Bachiri, A. Munaretto, J. Kaasila, G. Le Grand, I. Demeure, « Ambience Project - Find a Meeting », European Symposium on Ambient Intelligence, EUSAI'03, (Eindhoven, Netherlands), Nov. 2003, Lecture Notes in Computer Science, n°2875, pp. 3-14, Springer Verlag.

[P58] Y. M. Ghamri-Doudane, R. Naja, G. Pujolle, S. Tohme, « P3-DCF: Service Differentiation in IEEE 802.11 WLANs using Per-Packet Priorities », IEEE Semi-annual Vehicular Technology Conference, VTC'03-Fall, (Orlando, Florida), Oct. 2003, Vol. 5, pp. 3429-3433, IEEE publisher.

[P59] M. Achir, Y. M. Ghamri-Doudane, G. Pujolle, « Predictive Resource Allocation in Cellular Networks using Kalman Filters », IEEE International Conference on Communications, ICC'03, (Anchorage, Alaska), May 2003, Vol. 2, pp. 974-981, IEEE publisher.

[P60] N. Achir, N. Agoulmine, M. S. P. Fonseca, Y. M. Ghamri-Doudane, G. Pujolle, « Active Technology as an Efficient Approach to Control DiffServ Networks: the DACA Architecture », IFIP/IEEE International Conference on Management of Multimedia Networks and Services, MMNS'02, (Santa Barbara, California, U.S.A.), October 2002, Lecture Notes in Computer Science, n°2496, pp. 170-183, Springer Verlag.

[P61] Y. M. Ghamri-Doudane, A. Munaretto, N. Agoulmine, « Policy Control of Nomadic Corporate Ad hoc Networks », International Conference on Telecommunication Systems, Modeling and Analysis, ICTSM'02, (Monterey, California), Oct. 2002, Vol. 2, pp. 420-431.

[P62] N. Achir, M. S. P. Fonseca, Y. M. Ghamri-Doudane, N. Agoulmine, A. Mehaoua, « Active Networking System Evaluation: A Practical Experience », International Workshop on Mobile Multimedia Communications, MoMuC'00, (Tokyo, Japon), Oct. 2000, pp. P-27-1 to P-27-6.

[P63] K. Boussetta, Y. M. Ghamri-Doudane, A-L Beylot, « Simulcast-based DCA scheme for unicast and multicast communications in LEO satellite networks », IEEE Wireless Communications and Networking Conference, WCNC'00, (Chicago, Illinois, U.S.A.), Sept. 2000, Vol. 3, pp. 1472-1476, IEEE publisher.

[P64] F. Jaillet, Y. M. Ghamri-Doudane, M. Melkemi, A. Baskurt, « Adaptive Contour Sampling and Coding using Skeleton and Curvature », IEEE International Conference on Image Processing, ICIP'00, Vol. 2, (Vancouver, Canada), Sept. 2000, pp 899-902, IEEE Publisher.

8.8 ARTICLES IN REFEREED INTERNATIONAL WORKSHOPS – FULL PAPERS (7)

[P65] N. Chendeb-Taher, Y. M. Ghamri-Doudane, B. El Hassan, N. Agoulmine, "First Step towards an Efficient Admission Control: A Complete Analytical Model for 802.11e EDCA for Throughput and Delay Prediction", IEEE Broadband Convergence Networks Workshop, BCN'09, (New York, New York, U.S.A.), Jun. 2009.

[P66] Q-T- Nguyen Vuong, Y. M. Ghamri-Doudane, N. Agoulmine, "Novel Approach for Load Balancing in Heterogeneous Wireless Packet Networks", IEEE Broadband Convergence Networks Workshop, BCN'08, (Salvador, Brazil), Apr. 2008.

[P67] L. Gras, Q.-T. Nguyen-Vuong, Y. M. Ghamri-Doudane, N. Agoulmine, M. Kassar, B. Kervella, G. Pujolle, « Terminal Mobility in the Mobile and Wireless Realm: Tight, Loose vs. Very Loose Coupling Interworking », Seamless Service MObility Workshop, SSMO'07, (Marrakech, Morocco), Jul. 2007.

[P68] M. Jerbi, SM. Senouci, Y. M. Ghamri-Doudane, « Towards Efficient Routing in Vehicular Ad Hoc Networks », International Workshop on ITS for Ubiquitous Roads, UBIROADS'07, (Marrakech, Morocco), Jul. 2007.

[P69] V. K. Gondi, N. Agoulmine, Y. M. Ghamri-Doudane, « Mobility Management over Heterogeneous Multi-operators Networks using RII architecture », Annual Workshop of HP OpenView University Association, HP-OVUA'06, (Presqu'ile de Giens, France), May 2006.

[P70] E. Lehtihet, H. Derbel, N. Agoulmine, Y. M. Ghamri-Doudane, « Autonomic Computing: A Novel Goal-based Management Architecture and its Experimental Evaluation », HP Openview Univesity Association (HP-OVUA) - Plenary Workshop, HPOVUA'05, (Porto, Portugal), Jul. 2005.

[P71] S. Ghamri-Doudane, S. van der Meer, R. O'Connor, Y. M. Ghamri-Doudane, N. Agoulmine, « Resources Discovery and Management Using Policies in Smart Spaces », HP Openview Univesity Association - Plenary Workshop, HPOVUA'04, (Paris, France), Jun. 2004.

8.9 Articles in Refereed International Conferences – Short Papers (7)

[P72] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, "A Complete and Accurate Analytical Model for 802.11e EDCA under Different Traffic Conditions", IEEE International Conference on COMmunication Systems and NETworkS, COMSNETS'09, (Bangalore, India), Jan. 2009.

[P73] S. Lohier, G. Roussel, Y. M. Ghamri-Doudane, « Hierarchical QoS-aware Routing in Multi-tier Wireless Multimedia Sensor Networks », IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services, MMNS'07, (San José, California, USA), Oct.-Nov. 2007.

[P74] N. Chendeb, Y. M. Ghamri-Doudane, B. El Hassan, "Transmission Time Analysis and Modeling in 802.11e Contention Free Burst Mode", IEEE Global Information Infrastructure Symposium, GIIS'07, (Marrakech, Morocco), Jul. 2007.

[P75] S. Lohier, Y. M. Ghamri-Doudane, « Efficiency of Loss Differentiation Algorithms in 802.11 Wireless Networks », IFIP/IEEE International Conference on Management of Multimedia Networks and Services, MMNS'06, (Dublin, Ireland), Oct. 2006.

[P76] M. Nafaa, N. Agoulmine, Y. M. Ghamri-Doudane, "Persistent zone/peer naming in Mobile P2P Networks", Annual Workshop of HP OpenView University Association, HP-OVUA'06, (Presqu'ile de Giens, France), May 2006.

[P77] S. Ghamri-Doudane, Y. M. Ghamri-Doudane, N. Agoulmine, « Resources Discovery and Management Using Policies in 802.11 Networks », HP Openview Univesity Association - Plenary Workshop, HPOV-UA'03, (Genève, Suisse), Jul. 2003.

[P78] N. Agoulmine, N. Achir, M. S. P. Fonseca, Y. M. Ghamri-Doudane, A. Mehaoua, « Active Networking as an Efficient Approach to Deploy and Manage Multimedia Services », HP OpenView Univesity Association - Plenary Workshop, HPOV-UA'00, (Santorini, Grèce), Jun. 2000.

8.10 Articles in Refereed National Conferences (6)

Full Papers (4):

[P79] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « Adaptation de la couche MAC pour améliorer les performances des flux TCP dans les réseaux 802.11 », Rencontres Francophone sur les aspects algorithmique des télécommunications, AlgoTel'06, (Trégastel, France), May 2006.

[P80] A. Seddik-Ghaleb, Y. M. Ghamri-Doudane, S.-M. Senouci, « Études de la consommation d'énergie des versions TCP dans les réseaux ad hoc mobiles », De Nouvelles Avancés pour les Communications, DNAC'04, (Paris, France), Dec. 2004

[P81] Y. M. Ghamri-Doudane, A. Munaretto, N. Agoulmine, « Une nouvelle architecture de gestion par politiques pour les réseaux ad hoc d'entreprises nomades », Colloque Francophone sur la Gestion de Réseaux et de Services, GRES'03, (Fortaleza, Brésil), Fev. 2003, pp. 471-487.

[P82] M. S. P. Fonseca, N. Achir, Y. M. Ghamri-Doudane, N. Agoulmine, A. Mehaoua, « Avaliações de Sistemas de Redes Ativas : Uma experiência prática », 19éme Simpósio Brasileiro de Telecomunicações, SBrT'01, (Fortaleza, Brésil), Sept. 2001.

Short Papers (2):

[P83] S. Lohier, Y. M. Ghamri-Doudane, G. Pujolle, « Intérêts d'une approche inter-couche pour la QoS dans les réseaux 802.11 (papier court) », Colloque Francophone sur l'Ingénierie des Protocoles, CFIP'05, (Bordeaux, France), Mar. 2005.

[P84] Y. M. Ghamri-Doudane, S. M. Senouci, G. Pujolle, « Adaptation du modèle de contrôle par politiques aux réseaux ad hoc (Poster) », Colloque Francophone sur l'Ingénierie des Protocoles, CFIP'02, (Montréal, Canada), May 2002.

8.11 SEMINARS, INVITED TALKS, PANELS, TUTORIALS (9)

[P85] Y. M. Ghamri-Doudane, S. M. Senouci, « Geo-Routing in VANETs », Tutorial, RESCOM Summer School, (Presqu'ile de Giens, France), June 2010.

[P86] Y. M. Ghamri-Doudane, M. Jerbi, S. M. Senouci, « GyTAR: improved greedy traffic aware routing protocol for vehicular ad hoc networks in city environments », Presentation, Working-Seminar in Vehicular Technologies & Communications, (University College Dublin, Dublin, Ireland), May 2010.

[P87] Y. M. Ghamri-Doudane, M. Jerbi, S. M. Senouci, « Intersection-based Routing in Urban Vehicular Ad hoc Networks », Seminar, School of Computer Science and Informatics, University College Dublin (UCD), (Dublin, Ireland), March 2010.

[P88] Y. M. Ghamri-Doudane, M. Jerbi, S. M. Senouci, « Protocoles pour les communications dans les réseaux de véhicules en environnement urbain : Routage et GeoCast bases sur les intersections », Séminaire au Laboratoire Informatique de l'Institut Gaspard Monge (IGM-LabInfo), Université de Marne la Vallée, (Marne la Vallée, France), November 2008.

[P89] Y. M. Ghamri-Doudane, « Seamless Service Mobility Challenges and Solutions: Introduction and conclusion statements », Panel « Seamless Service Mobility Challenges and Solutions », Seamless Service MObility Workshop, SSMO'2007, (Marrakech, Morocco), July 2007.

[P90] Y. M. Ghamri-Doudane, « Proving the Scalability and Managing the Heterogeneity are the Issues », Panel « Research challenges in network management », The 17th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management, DSOM'2006, (Dublin, Ireland), Octobre 2006.

[P91] Y. M. Ghamri-Doudane, R. Naja, G. Pujolle et S. Tohmè, « P3-DCF : différenciation de service par trame dans les réseaux IEEE 802.11 », Atelier de présentation des résultats du projet ITEA AMBIENCE au groupe Thalès Communications, (Colombes, France), December 2003.

[P92] S. M. Senouci, Y. M. Ghamri-Doudane, A. Munaretto, et S. Lohier, « L'intelligence ambiante vue du point de vue réseau », DNAC - Ecole de Printemps 2003, (Tozeur, Tunisia), May 2003.

[P93] Y. M. Ghamri-Doudane, S. M. Senouci, A. Munaretto, et M. Dias-de-Amorim, « Le projet Ambience : les perspectives au niveau QoS », De Nouvelles Avancés pour les Communications, DNAC'2002, (Paris, France), December 2002.

8.12 DISSERTATIONS (3)

[P94] Y. M. Ghamri-Doudane, « Support et gestion de la qualité de service dans les réseaux sans fil », Ph.D. Dissertation, Laboratoire LIP6, Université P & M Curie - Paris 6, November 2003.

[P95] Y. M. Ghamri-Doudane, « Codage de contours par échantillonnage adaptatif », Master Dissertation, Laboratoire LIGIM, Université Claude Bernard - Lyon 1, July 1999.

[P96] Y. M. Ghamri-Doudane, « Étude comparative des deux modèles d'illumination globale : le tracé de rayons et la radiosité », Engineering Diploma Dissertation, INI, Alger, September 1998.

Part 2. Curriculum Vitae

Chapter 1. Biography

Yacine Ghamri-Doudane (ghamri@ensiie.fr) received an engineering degree in computer science from the National Institute of Computer Science (INI), Algiers, Algeria, in 1998, an M.S. degree in signal, image and speech processing from the National Institute of Applied Sciences (INSA), Lyon, France, and a Ph.D. degree in computer networks from the Pierre & Marie Curie University, Paris 6, France, in 1999 and 2003, respectively.

Yacine is currently associate professor (maître de conferences) at the Ecole Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise (ENSIIE), a major French post-graduate school located in Evry, France, and a member of the Gaspard Monge Computer Science Laboratory (LIGM – UMR 8049) at Marne-la-Vallée, France. From September 01, 2009 to August 31, 2009 he was on a secondment at the French National Centre of Scientific Research (CNRS) where he was part of the Gaspard Monge Computer Science Laboratory (LIGM – UMR 8049) at Marne-la-Vallée, France. His current research interests include Wireless Sensor Networks (WSN), Vehicular Networks, TCP and Multimedia over Wireless, QoS in WLAN/WMAN, Mobility Management in 4G Mobile Networks, Management of Wireless/Mobile Networks. Yacine holds three (3) international patents and he authored or co-authored six (6) book chapters, eleven (11) peer-reviewed international journal articles and more than 50 peer-reviewed conference papers. Since 1999, he participated or still participates to several national and European-wide research projects. Among them RNRT AMARRAGE (1999 - 2002), RNRT ARCADE (2001 - 2003), ITEA AMBIENCE (2001 - 2003), ITEA2 SUMO (2005 - 2007), ITEA2 HDTVnext (2008 - 2010), ITEA2 DiYSE (2009 - 2012), the IntelliCIS IC0806 COST Action (2009 - 2013), the WiNeMo IC0906 COST Action (2010 - 2014) and DIGITEO Envie Verte (2010 - 2013).

Since 2008, Yacine acted as an Expert for the Natural Sciences and Engineering Research Council (NSERC) of Canada by evaluating several project proposals. He also participated to several local committees within ENSIIE (School Council, School Board, and Faculty Recruiting Committees) and also other French universities (Faculty Recruiting Committees at the University of Paris 13 and at ENSEIRB-MATMECA).

As part of his professional activities linked to the computer networking research community, Yacine also acted as IEEE ICC Selected Area in Communications Symposium Co-Chair in 2009 and 2010, ACM IWCMC Vehicular Communications Technologies (VCT) Symposium Co-Chair in 2010, and IEEE/IFIP Wireless Days Wireless Multimedia Track Co-Chair in 2010. He was the TPC Co-Chair of SSMO'08 and IEEE MACE'09 workshops as well as the General Co-Chair of the VehiCom'09 workshop. He was also involved in the organizing committees of IEEE GIIS 2007, IEEE MUCS 2008, IEEE IM 2009 and IEEE GIIS 2009. Since December 2009, he is the Chair of the IEEE Communications Society (ComSoc) Technical Committee on Information Infrastructure (TCII) and the Secretary of the IEEE ComSoc Humanitarian Communications Technologies Ad hoc Committee (HCTC). He is the founding Area-Editor of the IEEE ComSoc Ad Hoc and Sensor Network Technical Committee (AHSN TC) Newsletter. He also acted or still acts as TPC member of the following IFIP, ACM or IEEE conferences (ICC, Globecom, PIMRC, WCNC, VTC, IM/NOMS, IWCMC, GIIS, WiMob, Wireless Days, CNSM, DSOM, IPOM, MMNS, ...) and workshops (WICTD, NiVi, IVCS, On-MOVE, Vehi-Mobi, MACE, CQR, MUCS, NTSEM, BcN, DANMS, UBIROADS, SSMO, …). He is a Member of IEEE.

Chapter 2. Detailed Curriculum Vitae

Name**: GHAMRI DOUDANE**
Surname: **Mohamed Yacine**
Date of Birth: **17/10/1976**

Affiliation: **LIGM (UMR 8049) and ENSIIE**

Administrative Addresses:

Teaching:

ENSIIE
1 Square de la résistance,
91025 Evry Cedex.

E-mail : ghamri@ensiie.fr
Tel.: +33.1.69.36.73.48
Fax : +33.1.69.36.73.27

Research:

Laboratoire d'Informatique Gaspard Monge
LIGM - UMR 8049
Cité Descartes, 77454 Marne-la-Vallée.

E-mail: yacine.ghamri@univ-mlv.fr
Tel.: +33.1.60.95.77.37
Fax: +33.1.60.95.75.57

Personal Address:

18-20, rue Petit
92110 Clichy.

Tel.: +33.1.74.63.47.34
Mob.: +33.6.85.93.34.69

2.1 CURRENT SITUATION

- **Associate Professor** (**Maître de conférences**) at l'Ecole Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise (ENSIIE – previously IIE-CNAM).
  Hired in Septembre 2004 – **Tenure** in Septembre 2005.
- **Researcher** at Laboratoire d'Informatique Gaspard Monge (LIGM) – Since Septembre 2009.
  Joint Research Unit of CNRS / Univ. Paris-Est Marne-la-Vallée / ESIEE / ENPC (**UMR 8049**).
- **Holder** of the Doctoral and Research Supervision Award (**PEDR**) since 2007.

2.2 EDUCATION

2003: **Ph.D.** (*Doctorat*) in Computer Science (Major: Computer Networks) from Université Pierre et Marie Curie – Paris 6.
Dissertation Title: QoS Support and Management in Wireless Networks.
Hosting Laboratory: LIP6.
Ph.D. Advisor: Pr. Guy Pujolle.
Ph.D. Defense Panel: R. Schutz (Thales - President), H. Afifi (INT - Reviewer), W. Donnelly (WIT - Reviewer), P. Jacquet (INRIA - Examiner), S. Tohmé (ENST - Examiner), N. Agoulmine (Univ. of Evry - Examiner).
Defended on: November 26, 2003.
Appreciation: With Honors (*Mention*: Très Honorable).

1999: **Master's** (*DEA*) in Signal, Image, and Speech Processing (Major: Image Compression) from Institut National des Sciences Appliquées (INSA) de Lyon.

Rank: 2nd (*Mention*: Bien).

1998: **Engineer Diploma in Computer Science** (Major: Computer Systems) from Institut National d'Informatique (INI – previously CERI), Algiers, Algeria.
Rank (Core Training): 1**st** (among ~250 students).
Rank (Major): 2**nd** (among ~90 students).
Appreciation (*Mention)*: Très Bien.

1993: "**Baccalauréat**" (Major: mathematics) from Académie d'Alger.
Appreciation (*Mention)*: A. bien.

2.3 PROFESSIONAL EXPERIENCE

2009-2010: Researcher Seconded at CNRS - Laboratoire d'Informatique Gaspard Monge (LIGM).
Joint Research Unit of CNRS / Univ. Paris-Est Marne-la-Vallée / ESIEE / ENPC (UMR 8049).

2004-2009: Researcher at Laboratoire Réseaux et Systèmes Multimédia (LRSM).
Emerging Resaerch Team (EE) of Université d'Evry val d'Essonne located at ENSIIE.

2003-2004: Lecturer (Attaché Temporaire d'Enseignement et de Recherche / ATER temps-plein) at Conservatoire National des Arts et Métiers (CNAM).

2002-2003: Lecturer (Attaché Temporaire d'Enseignement et de Recherche / ATER mi-temps) at Université d'Evry val d'Essonne.

2.4 RESEARCH VISITS

2010: Invited Researcher at University College Dublin - Irlande.
Duration: 2 weeks.

2007: Invited Researcher at University of Waterloo - Canada.
Duration: 1 week.

2005: Invited Researcher at TSSG / Waterford Institute of Technology - Irlande.
Duration: 1 week.

2003: Invited Researcher at TSSG / Waterford Institute of Technology - Irlande.
Duration: 1 month.

2002: Invited Researcher at Université du Québec À Montréal - Canada.
Duration: 4 months.

2.5 INTERNATIONAL VISITORS

2010: Pr. Mario Gerla, University of California Los Angeles (UCLA), U.S.A..
Duration: 1 month – May/June 2010.

2010: Dr. Nawel Zangar-Attia, Institut National des Sciences Appliquées et de Technologie (INSAT), Tunisia.
Duration: 2 weeks – March 2010.

2009: Dr. Javier Baliosian, Universidad de la Republica, Uruguay.
Duration: 2 weeks – November 2009.

2009: Dr. Aldri dos Santos, Universidade Federal do Paraná (UFPR), Brazil.
Duration: 1 week – July 2009.

2.6 SUPERVISION ACTIVITIES AND PH.D. PANEL PARTICIPATION

Post-docs (2):

1. Yan Li – Scheduling and Resource Management in Relay-assisted BWANs (**October 2009 – September 2010**) – *Funding:* ITEA2 HDTVnext Project.

2. Marwen Abdennebi – Adaptive Admission Control and Predictive Resource Reservation for Video Traffic in 802.16 rtPS (**October 2009 – August 2010**) – *Funding:* ITEA2 HDTVnext Project.

Research Engineer Supervision (1):

1. Philippe Valembois (100%) – Design, Implementation and Maintenance of an Autonomic Management Platform for Wireless Sensor Networks (**Since 10/2010**) – *Funding*: ITEA2 DiYSE Project.

Current Doctoral Supervisions (3):

1. Sylvain Cherrier (30%) – Autonomic Management of Multimedia Services in Internet of Things (Wireless Multimedia Sensor Networks) (**Since 10/2009**) – *Funding*: PRAG at Université Paris-Est Marne la Vallée.
2. Ahmed Ben Nacef (30%) – Cooperative Relaying in WSNs (**Since 10/2008**) – *Funding:* CIFRE Scholarship with Orange Labs.
3. Ismail Salhi (60%) – Applying Network Coding Principals to Improve WSN Performances (**Since 12/2008**) – *Funding:* ITEA2 DiYSE Project.

Past Doctoral Supervisions (4):

1. Nada Chendeb-Taher (60%) – Analytical Modeling and Resource Management in IEEE 802.11e Networks (**Ph.D. Defended on March 31, 2009**) – *Funding:* Scholarship from Agence Universitaire de la Francophonie (AUF – Co-supervision with Université Libanaise à Tripoli) + ITEA2 HDTVnext Project.
2. Alaa Seddik-Ghaleb (60%) – Study of the TCP behavior over MANETs: Performance Analysis and Joint Improvement of Throughputs and Energy Consumption (**Ph.D. Defended on March 30, 2009**) – *Funding:* Orange Labs Grant.
3. Moez Jerbi (40%) – Geographic Routing and Geo-localized Dissemination in urban VANETs (**Ph.D. Defended on November 2008**) – *Funding:* CIFRE Scholarship with Orange Labs.
4. Stéphane Lohier (60%) – TCP over IEEE 802.11 infrastructure Networks: Cross-Layer Optimization (**Ph.D. Defended on June 2006 – Ph.D. co-supervised starting June 2004**) – *Funding*: PRAG at IUT Champs sur Marne.

Collaboration with Ph.D. Students (collaborations that led to Journal publications):

1. Quoc-Thinh Nguyen-Vuong – Collaboration on Optimized Mobility Management in 4G Networks (**Ph.D. Defended on July 2008**).
2. Kamel Haddadou – Collaboration on Design and Evaluation of Policy-based per-Flow Resource Management Systems in Stateless IP Networks (**Ph.D. Defended on September 2007**).

Ph.D. Panel participation as Examiner (2):

1. Samir El-Haddad – Energy-Efficient Routing Protocols for WSNs (April 2, 2010).
2. Bastien Mainaud – Wireless Sensor Networks for Assistance to Persons: Design and Development of Reliability Mechanisms (July 22, 2010).

Master Supervision (16) – Since 10/1999:

2009-2010:

1. Nadia Haddadou (50%) – Secure Protocols for Data Harvesting, Storage and Sharing in VSNs
2. Ahlem Khlass (50%) – A Cross layer solution to Improve Network Connectivity and Capacity in Cooperative Vehicular Networks using Distributed Wireless Network Coding
3. Philippe Valembois (100%) – Proposition, implantation and evaluation (real tests) of intra-flow prioritization policy for 802.11aa WLANs
4. Ibtissem Boulanouar (50%) – Comparative Study of Middleware Solutions and Management Architectures for the Internet of Things: DiYSE Project Case Study.

2008-2009:

5. Thu Thuy Le (50%) – Cross-Layer Optimization of OFDM Broadband Wireless Access Networks with Cooperative Relaying.

6. Zakaria Telli (100%) – Proposition and evaluation (simulation) of intra-flow prioritization policy for 802.11aa WLANs.

2007-2008:

7. Lounes Baleh (100%) – Real-Time Adaptation of HD Audio/Video streams to 802.11e Available Resources.

2006-2007:

8. Aeli Tursun (100%) – Geo-localized Environment Perception in VANET/VSN.

2004-2005:

9. Yassine Chettoui (50%) – Throughput Estimation in Ad hoc Networks: A Cross-Layer Reactive Approach.

2000-2004:

10. Alaa Seddik (50%) – Comparative Study of TCP Variants' Energy Consumption in MANETs (2003/2004).
11. Salim Benayoune (100%) – A Cross-Layer Approach for Optimized Broadcasting in MANETs (2003/2004).
12. Sonia Waharte (100%) – QoS in 802.11 WLANs: Protocols and Algorithmic Aspects (2001/2002).
13. Mounir Achir (100%) – Predictive Resource Allocation in Cellular Networks (2001/2002).
14. Stéphane Lohier (50%) – QoS-Aware AODV Routing for 802.11-based MANETs (2001/2002).
15. Rima Tfaily (50%) – Active Control of DiffServ Networks (2000/2001).
16. Karim Rizkallah (50%) – Adapting the Policy-based Management Model to Ad hoc Networks (2000/2001).

Lab Projects (1 to 3 month duration) – 1st year Master, "Magistère", 1st and 2nd years engineering schools students: 21 since 10/1999.

2.7 Participation to Collaborative Research Projects

Current Projects:

Envie Verte Project: Regional (DIGITEO) project regrouping 3 academic laboratories from the Ile de France region: LIMSI (Lead), LIGM, CEDRIC/CNAM.
Duration: 36 months (September 2010 – August 2013).
Responsibilities: **Principal Investigator (PI) at LIGM.**
Scientific Participation: Autonomic Management of Multimedia Services in Internet of Things (Wireless Multimedia Sensor Networks).

WiNeMo Action: COST Action (IC0906) that has as objective to foster discussions and collaborative research activities on « *Wireless Networking for Moving Objects* ». WiNeMo regroups partners from 20 European countries.
Duration: 48 months (June 2010 – May 2014).
Responsibilities: **Principal Investigator (PI) at LIGM and ENSIIE, Member of the *Management Committee* representing France, Liaison with the COST IC0806 IntelliCIS Action.**
Scientific Participation: Optimizing the resource usage in vehicular networks (V2V/V2I, VANET).

IntelliCIS Action: COST Action (IC0806) that has as objective to foster discussions and collaborative research activities on « *Intelligent Monitoring, Control and Security of Critical Infrastructure Systems* ». IntelliCIS regroups 58 partners from 28 European countries and 2 non-European countries.
Duration: 48 months (May 2009 – April 2013).
Responsibilities: **Principal Investigator (PI) at LIGM and ENSIIE, and Substitute Member of the *Management Committee* representing France.**

Scientific Participation: Using Wireless Sensor Networks for the monitoring, the control and management of critical infrastructures.

DiYSE Project: ITEA2 project regrouping 35 European partners: Alcatel-Lucent (BE - Lead), Philips (BE), KU Leuven (BE), VUB (BE), Polytech'Mons (BE), Geosparc (BE), Kysoh (BE), WIT (IE), FeedHenry (IE), Alcatel-Lucent (FR), Thales Communications (FR), Neotiq (FR), Archos (FR), ENSIIE (FR), Telecom & Management SudParis (FR), Forthnet (GR), Mobilera (TR), Arti Teknoloji (TR), Turkcell (TR), Pozitim Teknoloji (TR), Nokia (FI), VTT (FI), Movial (FI), University of Tampere (FI), Oulu Yliopisto (FI), VIDERA (FI); Laurea (FI), finwe (FI), Rinnekoti-Säätiö (FI), Robotiker (SP), Atos Origin (SP), Universidad de Alcala (SP), Universidad Politécnica de Madrid (SP), I&IMS (SP), ESI (SP), AnswerTech (SP).
Duration: 30 months (September 2009 – February 2012).
Responsibilities: **Principal Investigator (PI) at ENSIIE.**
Scientific Participation: Our participation in the project is twofold. The first aspect deals with the improvement of the resource usage (network coding, topology control, routing) in Wireless (Multimedia) Sensor Network. The second aspect deals with the autonomic management of Wireless Sensor Networks and their Integration to the Internet (Internet of Things).

<u>Finalized Projects (from 2004 to 2010)</u>

HDTVnext Project: ITEA2 project regrouping 20 European partners: NXP Semiconductors France (FR - Lead), Activa Multimedia (SP), Barco (BE), DS2 (SP), ESI (SP), Grass Valley France (FR), Information & Image Management Systems, S.A. (SP), ENSIIE (FR), Maxisat (FI), Mobilera (TR), ON2 (FI), Pace France (FR), Philips Innovative Applications N.V. (BE), ROBOTIKER-TECNALIA (SP), Telefónica I+D (SP), Thales Communications (FR), Thomson R&D (FR), Trinnov Audio (FR), UAB (SP), VITEC Multimedia (FR).
Duration: 31 months (April 2008 – October 2010).
Responsibilities: **Principal Investigator (PI) at ENSIIE + Inter-Project Workshop Organization with a target to disseminate the projects results.**
Scientific Participation: Proposing and evaluating a new set of algorithms for the efficient support of High-Definition (HD) multimedia steams over wireless links/networks. We are more precisely interested on multimedia streams transmission with QoS (Resource Management) over the following wireless networks: 802.11e (WiFi with WMM), 802.16d (Fixed WiMAX) and 802.16j (WiMAX with Cooperative relaying).

POLYMAGE Project: DATAR/MENRT/CDC project regrouping several academic, institutional and associative partners: Ville d'Evry, INT Evry, UEVE, Vidéon, Maison de quartier des Epinettes, Fédération des VDPQ, Alice Coopératif Concept.
Duration: 12 months (2005 - 2006).
Scientific Participation: Indoor and Outdoor Wireless Video Steaming: Experimental Performance Evaluation

TCP & ad hoc Project: Research Grant (External Research Contract) from Orange Labs.
Duration: 36 months (June 2005 – May 2008).
Responsibilities: **Principal Investigator (PI) at LRSM Lab.**
Scientific Participation: Study of the TCP behavior over MANETs: Performance Analysis and Joint Improvement of Throughputs and Energy Consumption.

SUMO Project: ITEA2 project regrouping 12 European partners: Alcatel R&I (FR - Lead), Université Paris 6 (FR), UEVE / LRSM (FR), INT Evry (FR), Telenor R&D (NO), NRK (NO), Beep Science (NO), Euskaltel RTD (SP), CEIT (SP), DoNewTech Solutions (SP), Birdstep Technology ASA (NO), UNIK (NO).
Duration: 30 months (July 2005 – December 2007).
Responsibilities: **Principal Investigator (PI) at LRSM Lab + Lead of the SUMO Architecture Special Interest Group + Organization of an Open Scientific Workshop (SSMO 2007)**.
Scientific Participation: Our contribution in the project is twofold. Our first contribution is on optimizing the mobility management in 4G Networks (i.e., optimizing vertical handovers over heterogeneous wireless access networks). Our second contribution is on improving multimedia streaming over 4G networks.

Finalized Projects (from 1999 to 2004)

AMBIENCE Project: ITEA2 project regrouping 20 European partners: Philips (NL – Lead), Philips (UK), France Télécom (FR), Thales (FR), Thomson MM (FR), Vitec Multimedia (FR), Barco (BL), LIP6 (FR), University of Vienna (AT), University of Amsterdam (NL), ENST (FR), VTT Electronics (FI), Italdesign (IT), CCC (FI), Epictoid (NL), NetHawk (FI), Memodata (FR), Knowledge (GR), KU Leuven (BL), Telisma (FR).
Duration: 28 months (July 2001 – November 2003).
Responsibilities: ***co*Technical Leader of LIP6 activities**.
Scientific Participation: Service Differentiation in IEEE 802.11 WLANs.

ARCADE Project: RNRT national collaborative project regrouping 5 industrial and academic partners: LIP6 (Lead), INRIA, France Télécom, Thales, QoSMIC.
Duration: 24 months (January 2001 – December 2002).
Responsibilities: **Technical Activities Leader** (LIP6 was leading the consortium).
Scientific Participation: Definition of Architecture for SLS negotiation and IP network management using policies.

AMARRAGE Project: RNRT national collaborative project regrouping 8 industrial and academic partners: Thales (Lead), France Télécom, BoostWorks, LIP6, LORIA, ENST Paris, LAAS, L2TI – Institut Galilée.
Duration: 28 months (1999 –2002).
Scientific Participation: Definition of QoS Management Architecture using Active Policies (mixing the concepts of active networks and policy-based management).

2.8 Professional Activities and Community Services

IEEE:

- **Chair** of the IEEE Communications Society (ComSoc) **Technical Committee on Information Infrastructure** (**TCII**) – 12/2009 à 11/2011 (Secretary from 12/2007 to 11/2009).
- **Secretary** of the IEEE Communications Society (ComSoc) **Humanitarian Communication Technology Ad hoc Committee** (**HCTC**) – since 12/2009.
- **Member**: IEEE, Communications Society (ComSoc), Ad hoc and Sensor Networks Technical Committee (AHSN TC), Technical Committee on Network Operation and Management (CNOM), Autonomic Communications sub-Committee.

- **Area Editor** of the **Ad hoc and Sensor Networks Technical Committee (AHSN TC) Newsletter** – since 12/2007.

Editorial Board Membership:
- **Associate Editor** – Journal of Computer Systems, Networks, and Communications (JCSNC) – since 06/2010.

Participation to the Organizing Committee of the following conferences:
- IFIP Wireless Days 2010 (**Track Co-Chair**)
- IEEE/ACM IWCMC 2010 (**Symposium Co-Chair**)
- IEEE ICC 2010 (**Symposium Co-Chair**)
- IEEE GIIS 2009 (**Publication Chair**)
- IEEE ICC 2009 (**Symposium Co-Chair**)
- IEEE IM 2009 (**Poster Co-Chair**)
- IEEE GIIS 2007 (**Publication Chair**)

Participation to the Organizing Committee of the following workshops:
- IEEE MACE 2009 (**TPC Co-Chair** – organized in the frame of the IEEE/IFIP ManWeek'09)
- IEEE/ACM VehiCom 2009 (**General Co-Chair** – satellite workshop of IEEE/ACM IWCMC'09)
- IEEE MUCS 2008 (**Poster Chair** – satellite workshop of IEEE NOMS'08)
- IEEE SSMO 2007 (**Program Co-Chair** – satellite workshop of IEEE GIIS'07)

Participation to the Organizing Committees of the following local events:
- **Scientific committee member** of the **ResCom Summer School** (Summer School of the GDR CNRS ASR ResCom Pole) held from June 13 to 18, 2010 in Presqu'ile de Giens (France).
- Co-organizer of the **Vehicular Networks (REVE) thematic day** in the frame of **GDR CNRS ASR ResCom Pole** (October 20, 2008).
- Co-organizer of the **non-thematic spring event** of the **GDR CNRS ASR ResCom Pole** (February 7 and 8, 2008).

TPC Membre for the following events:
- Conferences:
    - IEEE ICC (2007, 2008, 2009, 2010, 2011)
    - IEEE GlobeCom (2007, 2008, 2009, 2010)
    - IEEE PIMRC (2008, 2009)
    - IEEE WCNC (2010, 2011)
    - IEEE VTC-Spring (2010, 2011)
    - IEEE VTC-Fall (2010)
    - IEEE/IFIP NOMS (2008, 2010)
    - IEEE/IFIP IM (2009)
    - IEEE GIIS (2007, 2009)
    - IFIP/IEEE Wireless Days (2008, 2009, 2010)
    - IEEE/ACM IWCMC (2007, 2008, 2009, 2010, 2011)
    - IEEE WiMob (2009)
    - IEEE/IFIP CNSM (2010)
    - IEEE/IFIP DSOM (2006, 2007,2008, 2009)
    - IEEE/IFIP MMNS (2008, 2009)
    - IEEE/IFIP IPOM (2007)
    - IEEE BCFIC (2011)
    - IEEE RIVF (2010)
    - IEEE GCC (2009)
    - IEEE CODS (2007)
    - IEEE IIT/Innovations (2007)

- IFIP/IEEE NetCon (2005, 2006)

- Workshops:
  - IFIP WICTD 2010 (// IFIP WCC 2010)
  - IEEE CQR (2008, 2009, 2010)
  - IEEE/IFIP MACE (2008, 2009, 2010)
  - IEEE Vehi-Mobi 2010 (// IEEE ICC 2010)
  - IEEE IVCS 2009 and 2010 (// IEEE CCNC 2009 et 2010)
  - IEEE NiVi 2009 (// IEEE Globecom 2009)
  - IEEE On-MOVE 2009 et 2010 (// IEEE LCN 2009 et 2010)
  - IEEE MUCS 2010 and 2011 (// IEEE PerCom 2010 et 2011)
  - IEEE NTSEM (// IEEE ICC 2011)
  - IEEE MUCS 2009 (// IEEE ICAC'09)
  - IEEE BcN 2008 and IEEE MUCS 2008 (// IEEE NOMS'08)
  - IEEE BcN 2007 and IEEE MUCS 2007 (// IEEE IM'07)
  - IEEE BcN 2006 (// IEEE NOMS'06)
  - IEEE MUCS 2006 (//Pervasive'06)
  - IEEE DANMS 2007 (// IEEE Globecom'07)
  - SSMO 2007 and UBIROADS 2007 (// IEEE GIIS'07)
- National Conferences:
  - Algotel (2006, 2007).
  - JDIR (2005)

Session Chair:

- AlgoTel'06, MMNS'06, DSOM'06, ICC'07, DSOM'07, MACE'07, NOMS'08, MUCS'08, MACE'09, WD'09, WD'10.

Reviewing and expertise:

- **Reviewer for the following journals**: IEEE Transactions on Vehicular Technologies, IEEE Transactions on Intelligent Transportation Systems, IEEE Journal on Selected Areas in Communications, Elsevier Journal on Parallel and Distributed Computing, IEEE Communications Magazine, IEEE Vehicular Technologies Magazine, Elsevier Computer Networks, Elsevier Computer Communications, Wiley Wireless Networks, International Journal of Communication Systems, Journal of Pervasive Computing and Communications, Journal of Mobile Multimedia, Journal of Network and Systems Management, Annals of Telecommunications.
- Expert for the **Natural Sciences and Engineering Research Council** (NSERC) of Canada – since 2008.

Other:

- Panelist - « **Research challenges in network management** » Panel - IEEE/IFIP DSOM 2006.
- Panel Organizer - « **Seamless Service Mobility: Challenges and Solutions** » Panel - IEEE SSMO 2007.
- Membre of the **Best Paper Award Selection Committee** - IEEE/IFIP MMNS 2006.

Community Services:

- **Member of the School Council** (renamed **Studies Council** in 2009) of ENSIIE – since November 2005.
- **Member of the Board** of ENSIIE – from December 2006 to November 2009.
- **Member of** the joint ENSIIE and University of Evry **recruiting committees** (lecturers and associate professors) - from March 2007 to September 2008.
- **Member of** ENSIIE **recruiting committee** in 2008 (Associate Professor in Computer Science - MCF 27 n. 543) and 2009 (Associate Professor in Computer Science - MCF 27 n. 550)

- **Member of** the Paris 13 University **recruiting committee** in 2009 (Associate Professor in Computer Networks - MCF 61 n. 786)
- **Member of** ENSEIRB-MATMECA **recruiting committee** in 2010 (Associate Professor in Computer Science - MCF 27 n. 1345)
- **Head of Internship affairs** at ENSIIE – 1st and 2nd years students – in 2006-2007
- **Head of Master Thesis affairs** at ENSIIE – 3rd year students – from 2007 to 2009
- **Supervision and organization** of the participation of ENSIIE students to SIANA 2005 (Semaine International des Arts Numériques et Alternatifs).

2.9 TEACHING ACTIVITIES

Since 1999, I delivered a set of courses in different higher education institutions. Between 1999 and 2004, I was first teaching assistant (vacataire) from 1999 to 2002, then as Lecturer (Attaché Temporaire d'Enseignement et de Recherche – ATER) from 2002 to 2004. In September 2004, I was appointed Associate Professor (Maître de Conférences) at ENSIIE. The following tables (in French) summarize the teaching duties I have made.

Annual Teaching Duties as Associate Professor (2004 - 2009)

Public	**Type**	**Matière**	**Volumes horaires (équ. TD)**				
			04-05	05-06	06-07	07-08	08-09
Ingénieur 3éme année	Cours	Réseaux : Routage et QoS	36 heures	36 heures	36 heures	36 heures	36 heures
Ingénieur 3éme année	TD	Réseaux : Routage et QoS	15 heures	15 heures	15 heures	15 heures	15 heures
Ingénieur 3éme année	TP	Réseaux : Routage et QoS			12 heures	12 heures	12 heures
Ingénieur 3éme année	Projets	Simulations réseaux	24 heures	24 heures	24 heures	24 heures	24 heures
Ingénieur 1ére année	TD	Systèmes Informatiques	99 heures	99 heures	99 heures	69 heures	69 heures
Ingénieur 3éme année	Suivis de stages	Réseaux et services répartis	12 heures	24 heures	18 heures	27 heures	15 heures
Ingénieur 1ére et 2éme année	Evaluations de stages	Réseaux, Web, programmation	6 heures	6 heures	6 heures	6 heures	6 heures
Formation FIP	Suivis pédagogiques	Réseaux, Web, programmation		4,5 heures	4,5 heures	4,5 heures	4,5 heures
Décharges admin.	Responsable des stages	1A et 2A			8,75 heures		
		3A				30 heures	30 heures
Total			**192 heures**	**208,5 heures**	**223,25 heures**	**211,5 heures**	**199,5 heures**

NB. In 2009-2010, I was on secondment in CNRS holding a full-time researcher position. During this academic year, I had no teaching duties. I conducted full-time research at Laboratoire d'Informatique Gaspard Monge (LIGM – UMR 8049).

Teaching Duties as Teaching Assistant and Lecturer Sorted by Topic and Level (1999 - 2004)

	3éme cycle	2éme cycle			1er cycle	Total (éq. TD)
	DEA / DESS	**Ecole d'ingénieurs**	**Maîtrise**	**Licence**	**DEUG**	
Réseaux	CM : **7h** TD : **6h** TP : SP : **6h**		CM : **4,5h** TD : **33h** TP : **96h** SP :	CM : **30h** TD : **16h** TP : **84h** SP : **54h**		**279,25h**
Programmation		CM : TD : TP : **80h** SP : **24h**		CM : TD : **29h** TP : SP :	CM : TD : **56h** TP : SP : **18h**	**167,66h**
Systèmes d'exploitations				CM : TD : **18h** TP : **36h** SP :		**42h**
Total (éq. TD)	**22,5h**	**79,66h**	**103,75h**	**215h**	**68h**	**488,91h**

(CM = Cours Magistraux ; TD : Travaux Dirigés ; TP : Travaux Pratiques ; SP = Suivi de Projets)