ANNÉE 2014

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

En Cotutelle Internationale avec
l'Université Saint-Joseph, Liban

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Informatique

Ecole doctorale MATISSE

présentée par

Melhem El Helou

préparée à l'IRISA (UMR 6074)
Institut de Recherche en Informatique et Systèmes Aléatoires

Radio Access Technology Selection in Heterogeneous Wireless Networks

Thèse soutenue à Rennes le 28 Novembre 2014

devant le jury composé de :

Oriol Sallent
Professeur à l'Université Polytechnique de Catalogne, Espagne / rapporteur

Andrzej Duda
Professeur à Grenoble INP-Ensimag / rapporteur

Samson Lasaulce
Directeur de recherche au CNRS / examinateur

Steven Martin
Professeur à l'Université Paris-Sud / examinateur

Claude Chaudet
Maître de conférences à Telecom ParisTech / examinateur

Marc Ibrahim
Maître de conférences à l'Université Saint-Joseph de Beyrouth, Liban / co-directeur de thèse

Kinda Khawam
Maître de conférences à l'Université de Versailles Saint-Quentin-en-Yvelines / examinateur

Samer Lahoud
Maître de conférences à l'Université de Rennes 1 / co-directeur de thèse

A mes très chers parents.

Mais, comme il est écrit, ce sont des choses que l'œil n'a point vues, que l'oreille n'a point entendues, et qui ne sont point montées au cœur de l'homme, des choses que Dieu a préparées pour ceux qui l'aiment.

1 Corinthiens 2, 9.

Remerciements

Avant tout développement sur mon travail, il apparaît opportun de commencer ce manuscrit par remercier l'Eternel Dieu pour cette grâce d'être en vie et en bonne santé depuis notre naissance jusqu'à ce jour. Je remercie ma famille que j'aime infiniment: mon père Georges, ma mère Nour, ma sœur Joëlle et mon frère Mario. Ils étaient toujours à mes côtés. Je remercie mon petit neveu Emilio qui a adouci mes séjours en France. Je remercie mes amis Richard, Chafic et Ralph qui n'ont cessé de m'encourager pendant ces trois années de thèse. Je remercie aussi mes compagnons de cellule et amis Farah, Mohammad et Jean, ainsi que tous les membres de l'équipe ATNet. Avec eux, j'ai passé d'agréables moments.

Je tiens également à témoigner toute ma reconnaissance à mes deux directeurs de thèse Bernard Cousin et Dany Mezher, pour leur soutien et pour toute la confiance qu'ils m'ont accordée. Je remercie particulièrement Fadi Geara, le doyen de la faculté d'ingénierie de l'Université Saint-Joseph, qui m'a encouragé pendant tout ce périple et qui m'a accueilli au sein de la famille enseignante de l'Ecole Supérieure d'Ingénieurs de Beyrouth. Sans ces trois personnes, ma thèse n'aurait jamais vu le jour. Je remercie profondément Marc Ibrahim et Samer Lahoud, mes deux encadrants, pour leur disponibilité, leur écoute et leur soutien. Grâce à leurs conseils et orientations fructueuses, ils ont rendu cette thèse une expérience très profitable. Je tiens à remercier du fond du cœur Kinda Khawam pour l'incroyable soutien technique et moral qu'elle m'a fourni. Je remercie chaleureusement Oriol Sallent et Andrzej Duda d'avoir accepté de rapporter ma thèse et j'exprime ma profonde gratitude envers les membres du jury Samson Lasaulce, Steven Martin et Claude Chaudet. Ce fut un vrai honneur de les avoir dans mon jury de thèse.

Résumé

Introduction

La demande de réseaux sans-fil haut débit ne cesse d'augmenter. Il a été rapporté que le trafic global de données mobiles a augmenté de 81 pour cent en 2013 [Cis14]. De plus, le trafic mobile mensuel devrait dépasser 15 exaoctets en 2018, près de 10 fois plus qu'en 2013 [Cis14]. Parallèlement à cette croissance impressionnante, les opérateurs mobiles sont invités à intelligemment investir dans les infrastructures de réseau. Ils sont aussi ramenés à reconsidérer leurs modèles de tarification forfaitaire, à la recherche de retour sur investissement positif.

Pour faire face à cette énorme demande de bande passante, les réseaux de nouvelle génération reposent sur la densification des stations de base. Les cellules ont des structures hiérarchiques: macro-, micro-, pico- et femto-cellules. Toutefois, une solution rentable est d'utiliser les technologies d'accès radio (TAR) existantes. Les futurs réseaux 5G sont ainsi conçus avec la vision de l'hétérogénéité. Diverses TAR, y compris les familles 3GPP (par exemple, UMTS, HSPA et LTE) et IEEE (par exemple, WiFi et WiMAX), sont intégrées et gérées conjointement.

Améliorer l'expérience de l'utilisateur est un autre facteur clé pour les réseaux sans fil hétérogènes. Une meilleure qualité de service (*Quality of Service* ou *QoS* en anglais), une durée de vie des batteries plus longue, et des prix plus faibles résument les besoins typiques des utilisateurs [FT13]. Vu que leurs caractéristiques se complètent mutuellement, diverses TAR coopèrent pour répondre efficacement aux besoins et préférences des utilisateurs. Alors que HSPA et LTE fournissent une QoS de bout-en-bout, ils supportent parfaitement le trafic temps réel. En plus, puisqu'ils peuvent desservir de grandes surfaces, ils gèrent efficacement la mobilité des utilisateurs. Cependant, WiFi offre des débits instantanés élevés sur de petites distances, et est connu pour son efficacité énergétique et économique. Ainsi, dans les réseaux sans-fil hétérogènes, les utilisateurs sont toujours connectés au mieux [GJ03] (*Always Best Connected* en anglais): ils sont non seulement toujours connectés, mais aussi rattachés à la TAR qui répond au mieux à leurs besoins.

Dans ce contexte, la sélection de TAR est une fonction clé pour améliorer les performances du réseau et l'expérience de l'utilisateur. Elle consiste à décider quelle TAR est la plus appropriée aux mobiles. Quand l'intelligence est poussée à la périphérie du réseau, les mobiles décident de manière autonome de leur meilleur TAR. Ils cherchent à maximiser égoïstement leur utilité. Toutefois, puisque les mobiles ne disposent d'aucune information sur les conditions de charge du réseau, leurs décisions peuvent causer des dégradations de performance. En outre, déléguer les décisions au réseau optimise la performance globale, mais au prix d'une augmentation de la complexité du réseau, des charges de signalisation et de traitement. Dans cette thèse, au lieu de favoriser une de ces deux approches décisionnelles, nous proposons un cadre de décision hybride: le réseau fournit des informations pour les mobiles pour mieux décider de leur TAR. Plus précisément, les utilisateurs mobiles choisissent leur TAR en fonction de leurs besoins et préférences individuelles, ainsi que des paramètres de coût monétaire et de QoS signalés par le réseau. En ajustant convenablement les informations du réseau, les décisions des utilisateurs répondent globalement aux objectifs de l'opérateur.

Plan détaillé de la thèse et contributions

La sélection de TAR a suscité un intérêt considérable parmi les chercheurs tout au long des dernières années [WK13, PKBV11, YSN10, KKP08]. Nous exposons dans le chapitre 2 les principales méthodes décisionnelles qui ont été proposées dans la littérature. Nous les classons en approches orientées réseau et approches orientées utilisateur en fonction de qui prend les décisions, et soulignons le besoin pour des approches hybrides. En fait, pour satisfaire les objectifs de l'opérateur entre autre une utilisation efficace des ressources, les approches orientées réseau ont été adoptées. Les éléments de réseau collectent les mesures et les informations nécessaires. Ils prennent les décisions de sélection de TAR de manière transparente aux utilisateurs, afin d'optimiser les performances globales du réseau. Toutefois, et dans le but de réduire la complexité du réseau, les charges de signalisation et de traitement, les approches orientées utilisateur ont également gagné en importance. En se basant sur leurs besoins et préférences, les utilisateurs rationnels choisissent leur TAR de manière à maximiser leur propre utilité. Alors que les mobiles n'ont pas de connaissance sur les conditions de charge du réseau, les approches orientées utilisateur dégradent potentiellement les performances. Bien que les mobiles cherchent à maximiser individuellement leur utilité, leurs décisions pourraient ne pas être dans leur intérêt. Ce dilemme est connu sous le nom de *la tragédie des biens communs* [Har68].

Notre défi est alors de concevoir une méthode de sélection de TAR qui améliore conjointement la performance du réseau et l'expérience de l'utilisateur, sans pour autant augmenter excessivement les charges de signalisation et de traitement. Nous proposons dans le chapitre 3 une approche innovante de décision hybride, qui combine les avantages des approches orientées réseau et des approches orientées utilisateur. Le réseau fournit des informations pour les mobiles, sur le canal logique de communication proposé par la norme IEEE 1900.4 [Std09], pour mieux décider de leur TAR. Plus précisément, le réseau masque ses conditions de charge et se contente de diffuser des incitations de coût monétaire et de QoS, à savoir des débits minimaux garantis et des débits maximaux. Les mobiles choisissent leur TAR en fonction de leur besoins et préférences, mais aussi des paramètres de coût et de QoS signalés par le réseau. En ajustant convenablement les informations du réseau, les décisions des utilisateurs répondent globalement aux objectifs de l'opérateur et évitent les états indésirables du réseau. Notre approche permet ainsi l'auto-optimisation, un élément clé des réseaux d'auto-organisation (*Self-Organizing Networks* en anglais) [3GP10].

Les prises de décisions, côté réseau et utilisateur, sont étudiées. Quand plusieurs stations de base desservent la même région, les décisions reposent traditionnellement sur la mesure de la puissance des signaux reçus. Afin de maximiser l'expérience de l'utilisateur, nous présentons dans cette thèse une méthode de décision multicritère (MDMC) basée sur la satisfaction. Outre leurs conditions radio, les utilisateurs mobiles tiennent compte des paramètres de coût et de QoS, signalés par le réseau, pour évaluer les TAR disponibles. Des fonctions d'utilité pour les trafics inélastique, streaming et élastique ont été définies. La TAR retenue est bien celle qui maximise l'utilité attendue de l'utilisateur. En comparaison avec les solutions existantes, à savoir SAW et TOPSIS, notre algorithme satisfait au mieux les besoins de l'utilisateur (par exemple, les demandes en débit, la tolérance de coût, la classe de trafic), et évite les décisions inadéquates. Une attention particulière est ensuite portée au réseau pour s'assurer qu'il diffuse les informations décisionnelles appropriées pour améliorer l'exploitation de ses ressources radio, quand les mobiles cherchent à maximiser égoïstement leur utilité. Nous présentons deux méthodes heuristiques, à savoir la *Staircase tuning policy* et la *Slope tuning policy*, pour dériver dynamiquement quoi signaler aux mobiles. Les paramètres de QoS sont modulés en fonction des conditions de charge selon soit une fonction en escalier, soit une fonction linéaire. Pour une TAR donnée, quand le facteur de charge augmente, les incitations de QoS se réduisent pour pousser les mobiles vers les TAR les moins chargées. On se retrouve finalement avec une distribution efficace des mobiles sur les différentes TAR. Ceci conduit à des performances meilleures, des utilisateurs plus satisfaits, et des gains d'opérateur plus élevés.

Dans le chapitre 4, nous évaluons minutieusement notre approche de décision hybride. Nous considérons trois scénarios de simulation. Dans le premier, nous s'intéressons aux informations de QoS et soulignons l'importance d'offrir des incitations de QoS différenciées, mais aussi des garanties de débit aux mobiles indépendamment des conditions de charge futures du réseau. Le deuxième scénario compare notre méthode de décision multicritère basée sur la satisfaction avec d'autres algorithmes très connus dans la littérature, à savoir SAW et TOPSIS. Puisqu'elle s'intéresse aux besoins des mobiles (par exemple, les demandes en débit, la tolérance de coût, la classe de trafic), notre méthode évite les décisions surdimensionnées et sous-dimensionnées, et maximise par la suite la performance et la satisfaction des utilisateurs. Dans le troisième scénario, nous évaluons les décisions côté réseau et prouvons l'efficacité de nos deux heuristiques: la *Staircase tuning policy* et la *Slope tuning policy*. Vu que les incitations de QoS sont modulées au rythme des conditions de charge, les mobiles sont efficacement distribués sur les différentes TAR. Ceci améliore les performances globales du réseau et les satisfactions individuelles des utilisateurs.

Dans le chapitre 5, nous comparons notre approche hybride avec des approches orientées réseau, des approches orientées utilisateur et des approches hybrides. Nous mettons en évidence l'efficacité de notre solution. Elle répond aux objectifs de l'opérateur et améliore l'utilisation des ressources, mais aussi aux besoins et préférences des utilisateurs et maximise leur satisfaction.

Dans le chapitre 6, nous nous concentrons sur l'optimisation de l'information du réseau. La dérivation des paramètres de QoS est formulée comme un processus de décision semi-markovien, et les stratégies optimales sont calculées en utilisant l'algorithme de *Policy Iteration*. En outre, et puisque les paramètres du réseau ne peuvent pas être facilement obtenues, une approche par apprentissage par renforcement est introduite pour dériver quoi signaler aux mobiles. Les performances des stratégies optimales, basées sur l'apprentissage et heuristiques, comme la probabilité de blocage et le débit moyen, sont analysées. Lorsque les seuils sont pertinemment fixés, notre méthode heuristique offre des performances très proches de la solution optimale. De plus, bien que de moins bonnes performances soient observées, notre algorithme basé sur l'apprentissage a l'avantage essentiel de ne nécessiter aucun paramétrage préalable.

Le chapitre 7 conclut la thèse. Nous résumons les principales contributions, et présentons les orientations futures du travail.

Conclusion et Perspectives

Pour faire face à la croissance rapide du trafic mobile, différentes TAR sont intégrées et gérées conjointement. Dans ce contexte, cette thèse étudie le problème de sélection de TAR, une fonction clé de la gestion commune des ressources radio dans les réseaux hétérogènes. Nous avons proposé une approche hybride de décision, qui combine les avantages des approches orientées réseau et des approches orientées utilisateur. Deux problèmes de décision interdépendants sont ainsi mis en jeu. Le premier au niveau du réseau consiste à dériver, pour chaque TAR, des incitations de coût monétaire et de QoS pour aligner globalement la décision des mobiles avec les objectifs de l'opérateur. Le deuxième au niveau de l'utilisateur consiste à combiner les besoins et préférences de l'utilisateur aux informations du réseau, pour sortir une décision de sélection de TAR qui maximise l'utilité de l'utilisateur.

Nous avons évalué l'importance d'offrir des incitations différenciées, avec éventuellement des garanties de débit indépendamment des conditions de charge du réseau. Lorsque les opérateurs proposent des classes de service *Premium*, *Regular* et *Economy*, qui diffèrent par leur paramètres de coût et de QoS, nous observons des performances meilleures et des satisfactions d'utilisateur plus élevées pour les trois types de trafic étudiés (inélastique, streaming et élastique). Ainsi, quand différentes TAR sont intégrées, il est intéressant d'offrir aux mobiles une variété de choix possibles, autrement dit de ne pas fournir dans toutes les TAR les même paramètres de coût et de QoS. Par ailleurs, quand les mobiles se voient garantir des débits minimaux, les performances des sessions temps réel s'améliorent.

En outre, nous avons comparé notre approche de décision hybride avec cinq autres méthodes de sélection de TAR. En comparaison avec les approches orientées utilisateur, notre solution maximise l'utilité du réseau, définie comme la somme des débits de tous les utilisateurs, et la satisfaction moyenne de l'utilisateur. Aussi, en comparaison avec les approches orientées réseau, notre solution améliore significativement la satisfaction moyenne de l'utilisateur.

Nous avons aussi souligné l'importance de masquer les conditions de charge du réseau, et de ne signaler que certains paramètres de coût et de QoS. Notre approche hybride surperforme les méthodes non-réalistes, où les mobiles ont une connaissance parfaite des conditions de charge du réseau. Ainsi, lorsque les objectifs de l'opérateur sont implicitement intégrés dans les paramètres de QoS, les ressources radio seront mieux utilisées, et la satisfaction de l'utilisateur sera maximisée.

De plus, nous nous sommes concentrés sur l'optimisation de l'information du réseau. Pour maximiser les performances du réseau à long terme, les informations de QoS ne doivent pas uniquement tenir compte des conditions de charge courantes, mais aussi de la demande prévue. Ainsi, la dérivation des paramètres de QoS a été formulée comme un processus de décision semi-markovien et les stratégies optimales ont été résolues grâce à l'algorithme de *Policy Iteration*. Dans l'état s, le réseau décide quels paramètres de QoS il faut diffuser pour maximiser la récompense du réseau à long-terme, tout en s'alignant avec les besoins et préférences des utilisateurs. Nous avons montré comment le coût de blocage et le coefficient d'actualisation (*discount factor* en anglais) peuvent être réglés pour contrôler les objectifs d'optimisation, alors que les mobiles cherchent à maximiser leur propre utilité. Cependant, lorsque le nombre de zones, de classes de trafic et de paramètres de QoS possibles augmentent, le nombre d'états risque d'exploser. Par la suite, trouver des stratégies optimales engendrera une énorme charge de traitement. Il serait alors intéressant d'étudier des techniques de réduction pour résoudre les grands processus de décision markoviens.

En outre, vu que les paramètres du réseau ne peuvent pas être facilement obtenus, une approche par apprentissage par renforcement a également été introduite pour dériver les paramètres de QoS. Lorsque le nombre de visites de chaque paire état-action tend vers l'infini, on est théoriquement sûr d'atteindre une stratégie optimale. Cependant, en pratique et puisque les paires état-action sont très nombreuses, elles sont partiellement explorées. Ceci conduit à une stratégie satisfaisante et acceptable plutôt qu'optimale. Pour surmonter cette limitation, le Q-learning doit être mis en œuvre en utilisant un réseau de neurones. Au lieu de stocker les valeurs de Q, les réseaux neuronaux les approximent et peuvent interpoler celles des paires état-action qui n'ont pas encore été visitées.

Abstract

To cope with the rapid growth of mobile broadband traffic, various radio access technologies (*e.g.*, HSPA, LTE, WiFi, and WiMAX) are being integrated and jointly managed.

Radio Access Technology (RAT) selection, devoted to decide to what RAT mobiles should connect, is a key functionality to improve network performance and user experience. When intelligence is pushed to the network edge, mobiles make autonomous decisions regarding selection of their most appropriate RAT. They aim to selfishly maximize their utility. However, because mobiles have no information on network load conditions, their decisions may lead to performance inefficiency. Moreover, delegating decisions to the network optimizes overall performance, but at the cost of increased network complexity, signaling, and processing load. In this thesis, instead of favoring either of these decision-making approaches, we propose a hybrid decision framework: the network provides information for the mobiles to make robust RAT selections. More precisely, mobile users select their RAT depending on their individual needs and preferences, as well as on the monetary cost and QoS parameters signaled by the network. By appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states.

We first introduce our hybrid decision framework. Decision makings, on the network and user sides, are investigated. To maximize user experience, we present a satisfaction-based Multi-Criteria Decision-Making (MCDM) method. In addition to their radio conditions, mobile users consider the cost and QoS parameters, signaled by the network, to evaluate serving RATs. In comparison with existing MCDM solutions, our algorithm meets user needs (*e.g.,* traffic class, throughput demand, cost tolerance), avoiding inadequate decisions. A particular attention is then addressed to the network to make sure it broadcasts suitable decisional information, so as to better exploit its radio resources while mobiles maximize their own utility. We present two heuristic methods to dynamically derive what to signal to mobiles. While QoS parameters are modulated as a function of the load conditions, radio resources are shown to be efficiently exploited. Our hybrid approach is further compared with different RAT selection methods, highlighting its effectiveness in enhancing resource utilization and user experience.

Moreover, we focus on optimizing network information. Deriving QoS parameters is formulated as a semi-Markov decision process, and optimal policies are computed using the *Policy Iteration* algorithm. Also, and since network parameters may not be easily obtained, a reinforcement learning approach is introduced to derive what to signal to mobiles. The performances of optimal, learning-based, and heuristic policies, such as blocking probability and average throughput, are analyzed. When thresholds are pertinently set, our heuristic method provides performance very close to the optimal solution. Moreover, although lower performances are observed, our learning-based algorithm has the crucial advantage of requiring no prior parameterization.

List of Abbreviations

3GPP	Third Generation Partnership Project
5G	Fifth Generation
AAA	Authentication, Authorization, and Accounting
AHP	Analytic Hierarchy Process
BS	Base Station
BLER	Block Error Rate
CDMA	Code Division Multiple Access
CoA	Care-of-Address
CRRM	Common Radio Resource Management
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
EDGE	Enhanced Data Rates for GSM Evolution
eNode B	Evolved Node B
ETSI	European Telecommunications Standards Institute
FA	Foreign Agent
GGSN	Gateway GPRS Support Node
GRA	Grey Relational Analysis
GSM	Global System for Mobile Communications
HA	Home Agent
HiperLAN2	High Performance Radio Local Area Network Type 2
HLR/HSS	Home Location Register/Home Subscriber Server
HoA	Home Address
HSPA	High Speed Packet Access
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
LTE	Long Term Evolution

MCDM	Multi-Criteria Decision Making
MEW	Multiplicative Exponent Weighting
MIP(v6)	Mobile IP version 6
MME	Mobility Management Entity
MN	Mobile Node
MPEG-4	Moving Picture Experts Group 4
mSCTP	Mobile Stream Control Transmission Protocol
NRM	Network Reconfiguration Manager
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDN GW	Packet Data Network Gateway
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RAT	Radio Access Technology
RL	Reinforcement Learning
RNC	Radio Network Controller
RRM	Radio Resource Management
RU	Resource Unit
SAW	Simple Additive Weighting
SGW	Serving Gateway
SGSN	Serving GPRS Support Node
SIP	Session Initiation Protocol
SMDP	Semi-Markov Decision Process
SNR	Signal-to-Noise Ratio
TCP	Transmission Control Protocol
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TRM	Terminal Reconfiguration Manager
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
URI	Uniform Resource Identifier
UTRAN	Universal Terrestrial Radio Access Network
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WFQ	Weighted Fair Queueing

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Along with the rapid growth of mobile broadband traffic, different radio access technologies, including 3GPP families (e.g., UMTS, HSPA, LTE) and IEEE ones (e.g., WiFi, WiMAX), are being integrated and jointly managed. Significant standardization efforts have been invested to integrate heterogeneous RATs. Two generic approaches, namely the loose coupling and the tight coupling, have thus been introduced. This chapter briefly describes the two coupling integration approaches, and discusses the common radio resource management in heterogeneous wireless networks. The objective of the thesis and the main contributions are further presented.

1.1 Why Heterogeneous Wireless Networks?

The demand for high-quality and high-capacity radio networks is continuously increasing. It has been reported that global mobile data traffic grew by 81 percent in 2013 [Cis14]. Furthermore, monthly mobile traffic is forecast to surpass 15 exabytes by 2018, nearly 10 times more than in 2013 [Cis14]. Along with this impressive growth, mobile operators are urged to intelligently invest in network infrastructure. They may also have to reconsider their flat-rate pricing models [NKG+12], seeking positive return-on-investment.

To cope with this huge demand for capacity, next-generation networks rely on densely deployed base stations with hierarchical cell structures [Cic13] (*i.e.,* macro, micro, pico and femto cells). A cost-effective solution is to use existing radio access technologies (RATs). Upcoming 5G networks are thus being devised with the vision of heterogeneity. Various RATs, including 3GPP families (*e.g.,* UMTS, HSPA, LTE) and IEEE ones (*e.g.,* WiFi, WiMAX), are being integrated and jointly managed. An example of a heterogeneous wireless network is illustrated in Fig. 1.1.

Another key driver for heterogeneous wireless networks is to enable traffic class-aware optimal coverage, capacity, and reliability with low cost and energy consumption. Next-generation networks focus on delivering high user experience. Better quality of service (QoS), longer battery lifetime, and lower cost are typical user requirements [FT13]. Since their characteristics complement each other, various RATs cooperate to cost-efficiently meet user needs and preferences. While HSPA and LTE provide end-to-end QoS, they excellently support real-time traffic. Also, as they may cover large areas, they effectively handle user mobility. However, WiFi offers high peak rates for small ranges, and is popular for its energy and cost efficiencies. Therefore, in heterogeneous wireless networks, users are always best connected [GJ03]: they are not only always connected, but also served through the RAT that best fulfills their requirements.

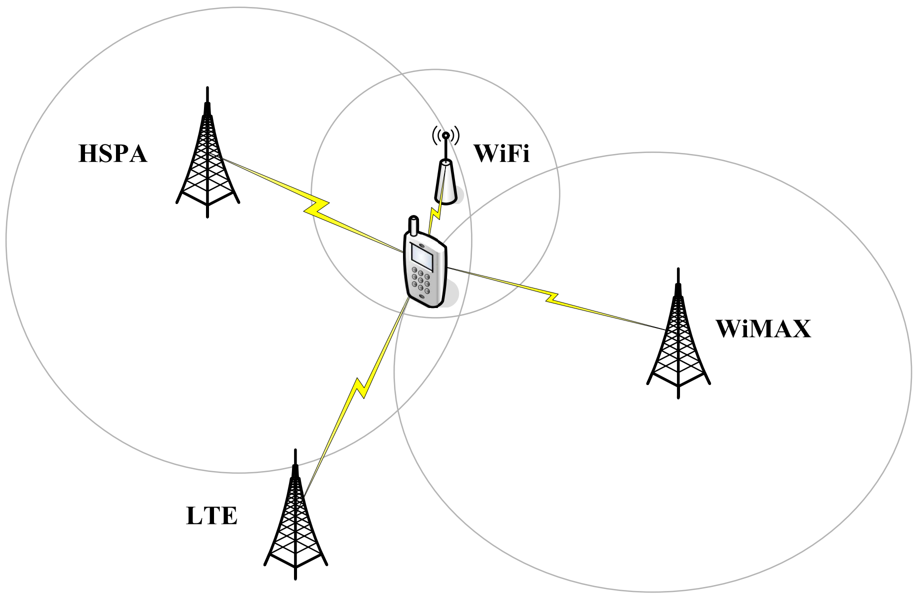

Figure 1.1: A typical heterogeneous wireless network

1.2 Integration of Heterogeneous Wireless Networks

In the recent years, as a leading indicator of the shift to heterogeneous wireless networks, mobile manufacturers have proposed multi-mode devices [ETS00]. Equipped with either multiple radio interfaces or a single reconfigurable one, multi-mode devices are able to connect, simultaneously or not, to different RATs. Concurrently, following the same trend, standardization bodies have focused their efforts to integrate both 3GPP and non-3GPP (*e.g.,* IEEE) RATs.

In this context, the European Telecommunications Standards Institute (ETSI) has presented two generic approaches, namely the loose coupling and the tight coupling, to integrate heterogeneous RATs [ETS01, LPMK05, Bea08]. Although originally conceived to interconnect HiperLAN2 and UMTS, these methods remain valid for multiple 3GPP and

IEEE RATs.

1.2.1 Loose Coupling Integration

With loose coupling integration, various RATs exist independently. They are not directly connected; instead, they are connected to the Internet. Fig. 1.2, taken with modifications from [Bea08], illustrates an example of a loosely coupled hetergeneous wireless network. HSPA, WiMAX, WiFi, and LTE data traffics are transmitted to the Internet over different core infrastructures.

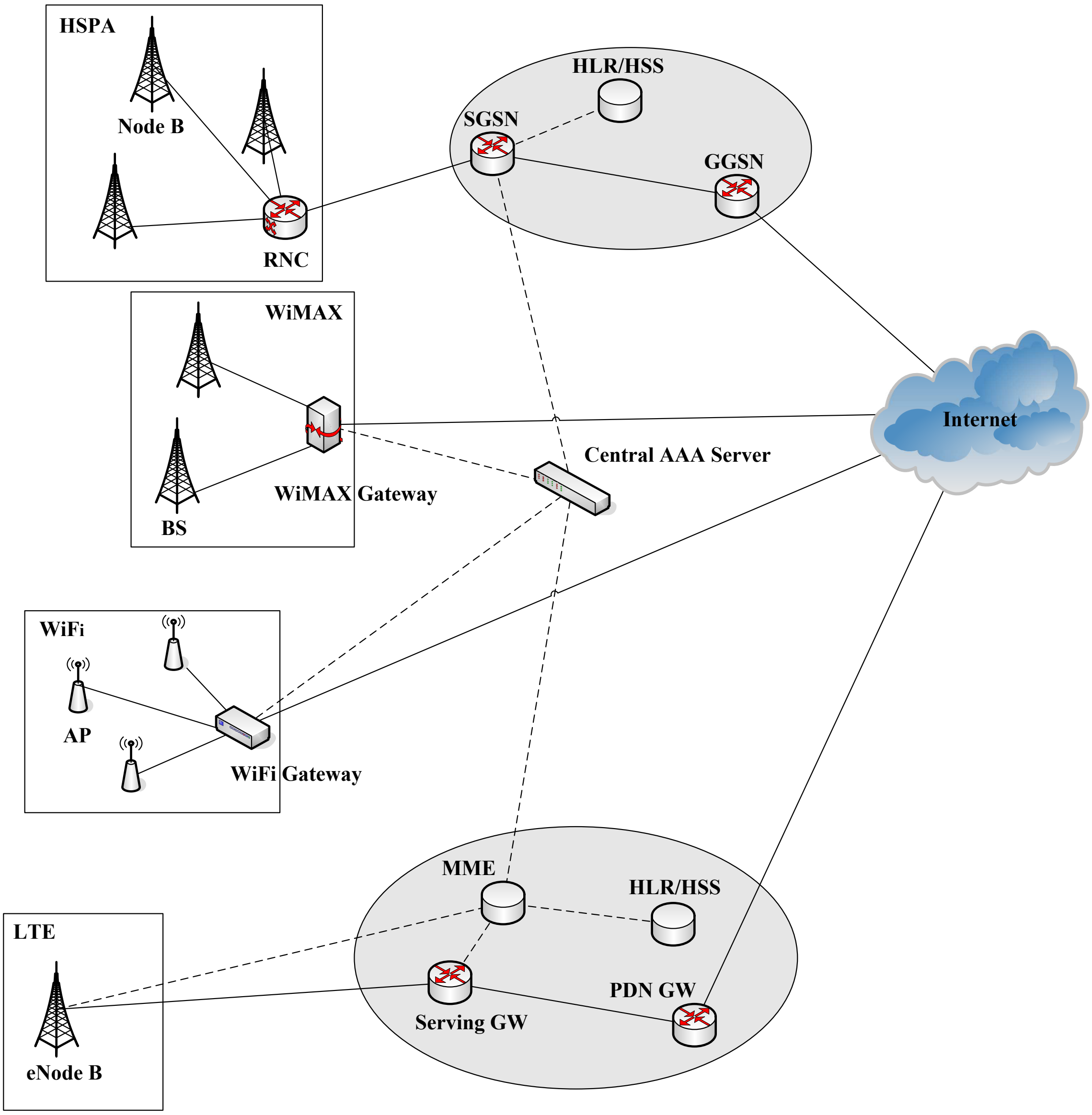

Figure 1.2: Loose coupling architecture

As serving RATs share the same Authentication, Authorization, and Accounting (AAA) server, mobile users may have a single subscription, and yet have access to several RATs. Furthermore, to handle inter-RAT mobility, border gateways usually implement Mobile IP (MIP). While moving across neighboring RATs, mobiles keep their IP address, stay connected, and maintain the ongoing communication sessions.

Loose coupling approach seems to be a short-term solution to integrate heterogeneous RATs [Bea08]. Mobile operators take advantage of multiple deployments with no major investment. However, because RATs are connected only through the Internet, it is not possible for them to easily and quickly communicate dynamic cell information (*e.g.,* cell load, interference measurements, received power level, and transmitted power level). Therefore, common radio resource management functionalities (*e.g.,* common admission and congestion control, RAT selection, inter-RAT handover, and common packet scheduling) are not efficiently provided. Usually, when RATs are loosely coupled, real-time services hardly survive during inter-RAT handovers [LPMK05].

1.2.1.1 Mobility Management

Regarded as the least common mobility denominator, MIP is far from being the only solution to provide seamless inter-RAT handovers. Mobility management can be performed at either the network layer, the transport layer, or even the application layer.

We present below some of the macro-mobility solutions. However, MIP and Session Initiation Protocol (SIP) are the two major mobility protocols.

1.2.1.1.1 Mobile IP MIP is a network layer solution to mobility management. It provides transparent handover support, including the maintenance of active Transmission Control Protocol (TCP) connections and User Datagram Protocol (UDP) port bindings.

As illustrated in Fig. 1.3, MIP introduces three functional entities: Mobile Node (MN), Home Agent (HA), and Foreign Agent (FA) [Per97]. MNs have two IP addresses: a fixed Home Address (HoA) that serves as their unique identity, and a temporary Care-of-Address (CoA) that identifies their present point of attachment, while away from their home RAT. The HA, residing on MN home RAT, and the FA, residing on MN foreign RAT, are used to bind the MN HoA to its CoA. They are in charge of packet forwarding, while mobiles roam across serving RATs.

When a MN moves to a foreign RAT, it obtains a new CoA. It then needs to inform its home RAT of its present location (*i.e.,* CoA). The HA intercepts the traffic destined to the MN, and tunnels it to the MN present point of attachment. Later, if using MIPv6, direct communications are possible between the MN and its correspondent node.

To implement MIP, operators need to introduce HA and FA entities. Moreover, when MNs are far away from their HA, they suffer from long handover delays.

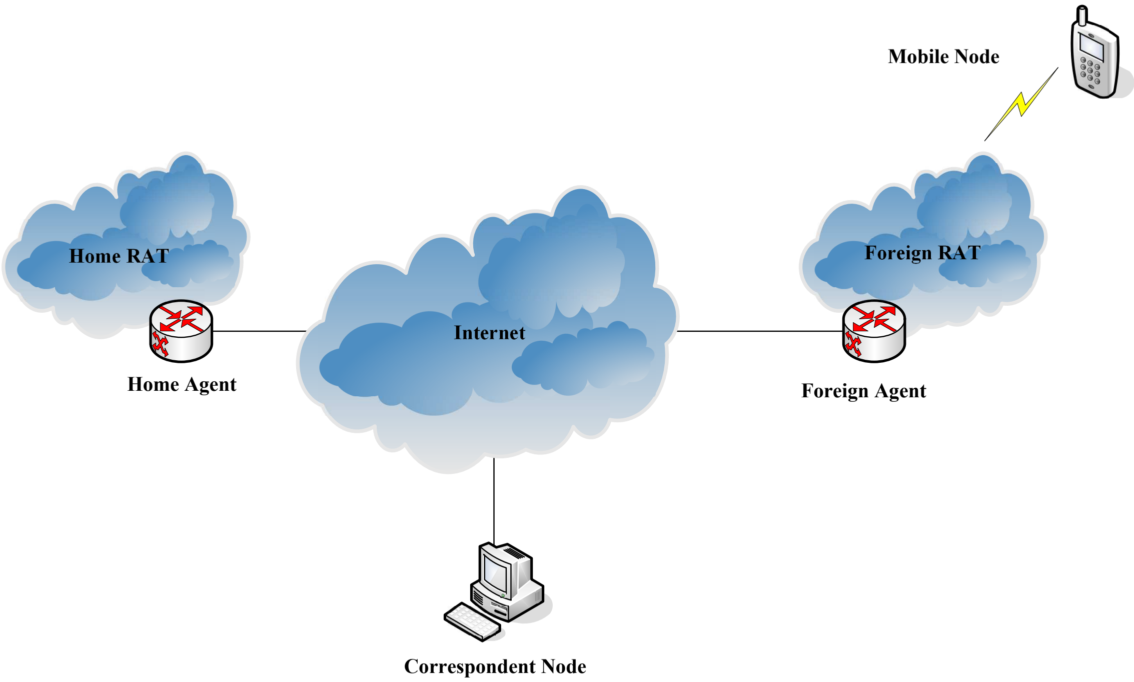

Figure 1.3: Mobile IP entities

1.2.1.1.2 Mobile Stream Control Transmission Protocol Handling mobility at the transport layer has proved to achieve better performance than at the network layer [WHB08]. Higher layers can quickly adapt to route changes. Therefore, Mobile Stream Control Transmission Protocol (mSCTP) has been introduced as a transport protocol to support inter-RAT handover [MYLR04]. It benefits from the multihoming feature and the dynamic address reconfiguration extension of SCTP. Mobiles may be configured with multiple IP addresses. As they move across various RATs, they can dynamically add, delete, and change their primary address, enabling seamless handover support.

mSCTP provides a network-independent solution to handover management: network components need not to be modified. However, as mSCTP replaces TCP, applications should use mSCTP sockets instead of TCP sockets. This practically limits the deployment of mSCTP.

1.2.1.1.3 Session Initiation Protocol SIP is an application layer solution to mobility management. It aims to keep mobility support independent of the underlying transport and network layers. SIP users are completely identified by a uniform resource identifier (URI) that is independent of their location. However, a mapping from their URI to their present IP address is established, and can be updated as mobiles roam across serving RATs.

Furthermore, SIP can be used to create, modify, and terminate two-party (unicast) and multi-party (multicast) sessions. The modifications involve changing IP addresses and ports, as well as inviting more participants.

When a mobile changes its serving RAT, it obtains a new IP address. It then needs to generate a re-invite message to its correspondent node. Therefore, packets destined to the mobile are sent to its new address.

To implement SIP, operators need to introduce SIP servers (*i.e.,* SIP proxy, registrar). Yet, as it operates at the highest level, SIP causes long handover delays [PBB+01].

1.2.2 Tight Coupling Integration

Within tight coupling integration, serving RATs are directly connected to a 3GPP core infrastructure component (*i.e.,* SGSN, GGSN, Serving GW, PDN GW). They appear as several access infrastructures to a single core network. An example of a tightly coupled heterogeneous wireless network is presented in Fig. 1.4. When HSPA is directly connected to the LTE serving gateway, non-3GPP RATs such as WiFi and WiMAX are connected to the LTE PDN gateway [3GP08].

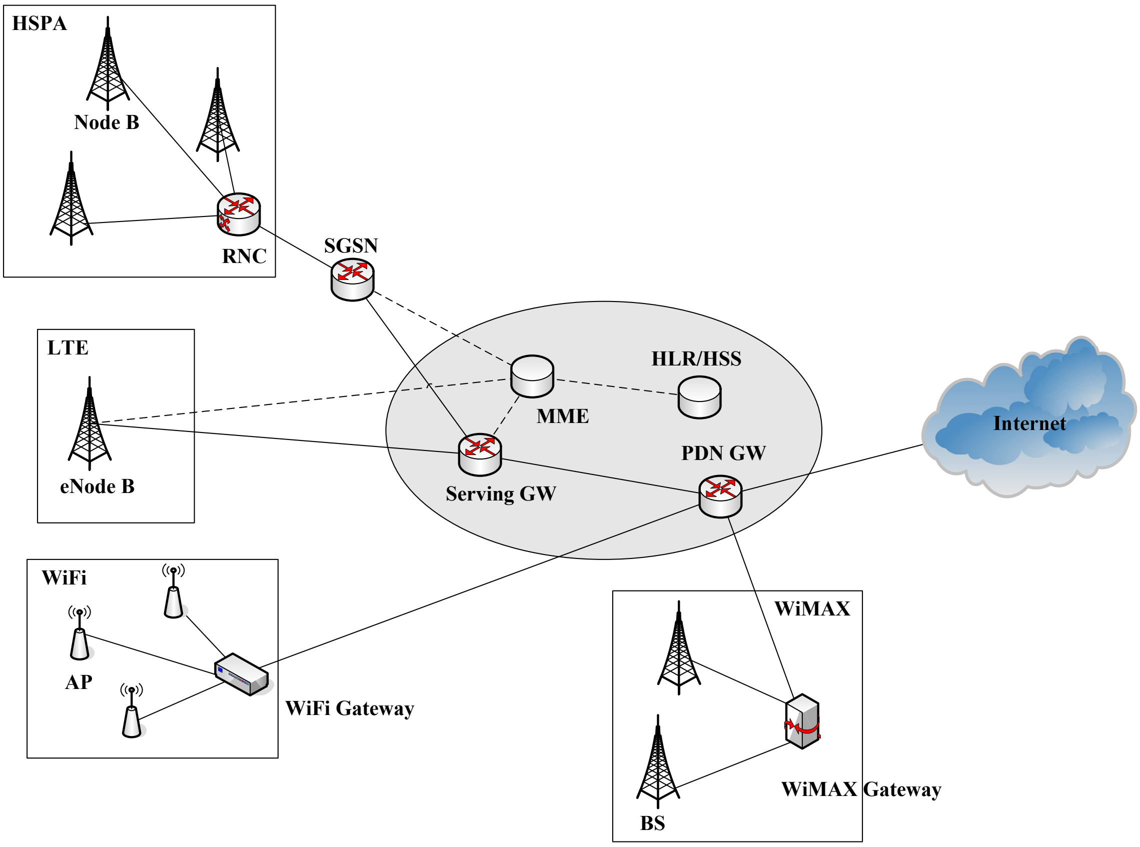

Figure 1.4: Tight coupling architecture

Serving RATs are totally transparent to the LTE core infrastructure. Border gateways hide HSPA, WiFi, and WiMAX particularities. Mobile IP, mSCTP, and SIP are no more necessary; instead, border gateways implement LTE procotols, and act as virtual LTE components. They are in charge of protocol translation and signaling exchange.

Moreover, WiFi and WiMAX can be tightly connected to an HSPA infrastructure, at either the SGSN level, the GGSN level, or even the RNC level [LPMK05, Bea08].

Tight coupling approach provides efficient common radio resource management, particularly reducing inter-RAT handover latency. Yet, by injecting WiFi, WiMAX, and HSPA data traffic into the LTE core infrastructure, the design of LTE components needs to be revisited [Bea08].

1.3 Radio Resource Management

The 3GPP Common Radio Resource Management (CRRM) functional model assumes that radio resources are divided into radio resource pools [3GP01]. Each includes a subset of radio resources managed by an RRM entity. More precisely, a radio resource pool consists of one or several cells, typically under the control of one RNC in UTRAN, or one access point controller in WiFi. Moreover, in LTE and WiMAX, either centralized or decentralized RRM are envisaged [LTE14, Ahm10]. When in centralized RRM, an additional central RRM entity is introduced, base stations exchange radio resource information and make local decisions in decentralized RRM.

Furthermore, to optimize network performance, radio resources belonging to different pools need to be jointly managed. CRRM gains, for both real-time and non-real-time traffics, have been evaluated in [THH02]. As illustrated in Fig. 1.5 that is taken from [PRGS08], CRRM entities are introduced to control local RRM entities. Centralized and decentralized CRRM are presented in [3GP01]: CRRM entities are either additional central nodes, or integrated in RRM entities.

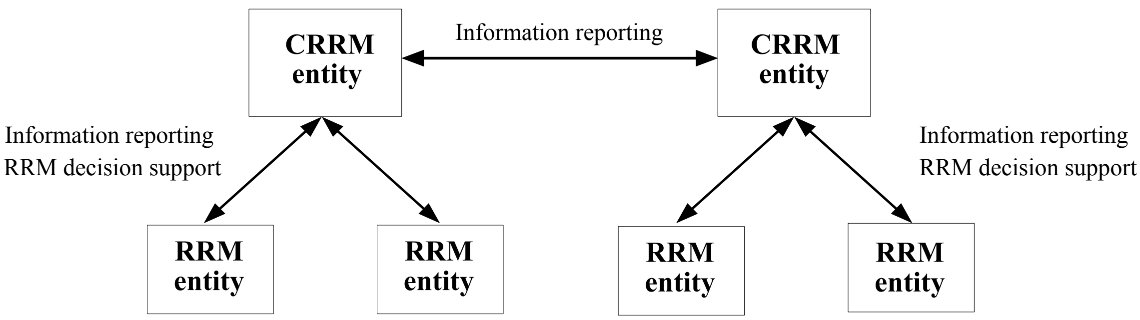

Figure 1.5: CRRM functional model

The interactions between RRM and CRRM entities support two basic functions, namely the information reporting and the RRM decision support functions. The information reporting function allows RRM entities to communicate relevant measurements and information to their controlling CRRM entity. Static cell information (*e.g.,* cell relations, capabilities and capacities) and dynamic cell information (*e.g.,* cell load, interference measurements, received power level, and transmitted power level) are reported either periodically, or when triggered by an event. Moreover, the information reporting function enables CRRM entities to share information, as represented in Fig. 1.5.

The RRM decision support function describes how RRM and CRRM entities interact to make RRM decisions. The CRRM entity can either make decisions and impose them on local RRM entities, or only advise RRM entities in their decisions. Thus, RRM functions may be splitted over RRM and CRRM entities.

Admission and congestion control, RAT selection, inter- and intra-RAT handover, packet scheduling, and power control are the main RRM functions. Depending on the degree of interaction between RRM and CRRM entities, some may be delegated to CRRM entities. Three illustrative examples are found in [PRGS08]. In the first, no RRM functions are associated to CRRM entities. However, RAT selection is performed using directed retry. In the second, only long-term functions (*i.e.,* RAT selection and inter-RAT handover) are moved to CRRM entities. In the third, long- and short-term functions (*i.e.,* admission and congestion control, RAT selection, handover, and packet scheduling) are delegated to CRRM entities. Yet, only frequent technology-dependent procedures, namely power control, remain associated to local RRM entities.

1.3.1 Packet Scheduling and RAT Selection

When resource allocation is jointly performed (*i.e.,* packet scheduling is moved to CRRM entities), traffic is splitted over many RATs. Mobiles can concurrently make use of resources belonging to different pools, as in [KIC$^+$11]. Yet, since packet scheduling is a short-term RRM function, and therefore has to be repeated at very short time intervals (in the order of milliseconds), allocating resources at the CRRM level turns out to be costly. Moreover, traffic splitting seems not to be necessary. When WiFi and WiMAX are integrated, it has been proven in [KIC$^+$11, CTG09] that elastic users should optimally be connected to a single RAT. This, however, remains true as long as mobiles are associated with their best RAT.

RAT selection, devoted to decide to what RAT mobiles connect, is a long-term RRM function. It is performed at session initiation (initial RAT selection), and eventually during session lifetime (inter-RAT handover). Yet, to maximize resource utilization, decisions should take into account information about serving RATs (*e.g.,* network load conditions). RAT selection is then ideally moved to CRRM entities, as it involves many local RRM entities. An illustrative example is when the CRRM entity collects cell load measurements from local RRM entities, and accordingly associate mobiles with the less loaded RAT.

In this thesis, we tackle the RAT selection, regarded as a key CRRM functionality. Mobiles are connected to a single RAT, and packet scheduling is locally performed in each cell. Recall that RAT selection and packet scheduling are on different time scales. Our aim is, however, to design efficient algorithms to exploit network integration. This involves answering the following questions:

- Who makes RAT selection decisions?
- How, and based on what criteria, decisions are made?
- What objectives are to be met?

1.4 Thesis Contribution

RAT selection has triggered considerable interest among researchers in the past few years [WK13, PKBV11, YSN10, KKP08]. To meet operator objectives, including efficient resource utilization, network-centric approaches have been proposed. Network elements collect necessary measurements and information. They take selection decisions transparently to end-users, in a way to optimize overall network performance. However, to reduce network complexity, signaling and processing load, mobile-terminal-centric approaches have also gained in importance. Rational users select their RAT, depending on their needs and preferences, in a way to selfishly maximize their utility. Yet, when mobiles have no information on network load conditions, mobile-terminal-centric approaches potentially lead to performance inefficiency. Although mobiles try to selfishly maximize their utility, their decisions may be in no one long-term interest. This dilemma is known as the *Tragedy of the commons* [Har68].

Our challenge is then to design a RAT selection approach, that jointly enhances network performance and user experience, while signaling and processing burden remains reduced. In the present contribution, we propose an innovative hybrid decision method, that combines benefits from both network-centric and mobile-terminal-centric approaches. The network provides information for the mobiles to make robust RAT selections. More precisely, network load conditions are masked, and only monetary cost and QoS incentives, namely minimum guaranteed throughputs and maximum throughputs, are provided. Mobiles select their RAT depending on user needs and preferences, as well as on the cost and QoS parameters signaled by the network. By appropriately tuning network information, mobile decisions are globally expected to meet operator objectives, avoiding undesirable network states. Our approach then enables self-optimization, a key feature of self-organizing networks [3GP10].

When several base stations are available, decisions are traditionally based on received-signal-strength measurements. In this thesis, so as to maximize user experience, we introduce a satisfaction-based Multi-Criteria Decision-Making (MCDM) method. In addition to their radio conditions, mobile users consider the cost and QoS parameters signaled by the network, to evaluate serving RATs. In comparison with existing MCDM solutions, our algorithm meets user needs (*e.g.,* traffic class, throughput demand, cost tolerance), avoiding inadequate decisions. A particular attention is then addressed to the network to make sure it broadcasts suitable decisional information, so as to better exploit its radio resources while mobiles maximize their own utility. We present two heuristic methods to dynamically derive what to signal to mobiles. While QoS parameters are modulated as a function of the load conditions, radio resources are shown to be efficiently exploited.

Decision makings, on the network and user sides, are investigated and evaluated separately. Our hybrid approach is then compared with multiple network-centric, mobile-terminal-centric and hybrid methods, highlighting its effectiveness in enhancing resource utilization and user experience.

Further, we focus on optimizing network information. Deriving QoS parameters is formulated as a semi-Markov decision process, and optimal policies are computed using the *Policy Iteration* algorithm. Also, and since network parameters may not be easily obtained, a reinforcement learning approach is introduced to derive what to signal to mobiles. The performances of optimal, learning-based, and heuristic policies are analyzed. When thresholds are pertinently set, our heuristic method provides performance very close to the optimal solution. Moreover, although lower performances are observed, our learning-based algorithm has the crucial advantage of requiring no prior parameterization.

1.5 Thesis Organization

The remaining of this thesis is organized as follows: RAT selection is surveyed in Chapter 2. We discuss and classify a wide range of methods, according to who makes RAT selection decisions. Chapter 3 introduces our hybrid decision approach. Decision makings, on the network and user sides, are also investigated. More precisely, our satisfaction-based multi-criteria decision-making method is presented, and two heuristic methods are proposed to dynamically tune network information.

In Chapter 4, we thoroughly evaluate our hybrid decision approach. As a matter of fact, we consider three simulation scenarios. In the first one, QoS information is investigated. We study the performance improvement achieved by providing differentiated service classes and minimum throughput guarantees to mobiles, regardless of future network load conditions. The second scenario compares our satisfaction-based multi-criteria decision-making method with other existing algorithms, namely SAW and TOPSIS. In the third scenario, we illustrate the gain from using our tuning heuristics in comparison with static network information.

Chapter 5 compares our RAT selection method with multiple network-centric, mobile-terminal-centric, and hybrid approaches. We prove the effectiveness of our solution in enhancing resource utilization and user experience.

In Chapter 6, we optimize network information using Markov decision processes. We show how to dynamically maximize long-term network reward, aligning with user preferences. Further, and since network parameters may not be easily obtained, a reinforcement learning approach is introduced to derive what to signal to mobiles. The performances of optimal, learning-based, and heuristic policies are then analyzed.

Chapter 7 concludes the thesis. We summarize the main contributions, and present future research directions.

Chapter 2

Radio Access Technology Selection

Radio Access Technology (RAT) selection, devoted to decide to what RAT mobiles connect, is a key functionality to improve network performance and user experience. When intelligence is pushed to the network edge, mobiles make autonomous decisions regarding selection of their most appropriate RAT. They aim to selfishly maximize their utility. However, because mobiles have no information on network load conditions, their decisions may lead to performance inefficiency. Moreover, delegating decisions to the network optimizes overall performance, but at the cost of increased network complexity, signaling and processing load. This chapter reviews a wide range of RAT selection methods, and classifies them according to who makes decisions. We further identify the need for efficient hybrid approaches, that jointly enhance network performance and user experience, while signaling and processing burden remains reduced.

2.1 Introduction

When several radio access technologies cover the same region, deciding to which one mobiles connect is known as the RAT selection functionality. This appears at session initiation, and during session lifetime through inter-RAT handovers. Initial RAT selection and handovers can generally be separated into three phases [KKP08, CSH$^+$01]:

- **Information Gathering:** User information (*e.g.,* user needs, preferences, and quality of service), and contextual information (e.g., radio link availability, and cell load measurements) are collected. During session lifetime, information is gathered periodically, and may trigger inter-RAT handovers. Typically, as user QoS degrades, handovers can be initiated by either the network or the mobile.

- **Decision:** At session initiation and inter-RAT handovers, decisions as to what RAT mobiles connect need to be made. They usually depend on the previously collected information.

- **Execution:** After decisions are made, messages are exchanged to (re)-establish user connectivity. When mobiles handover between serving RATs, their traffics need to be seamlessly rerouted.

User and contextual information can be gathered using the IEEE 802.21 [DLOBS+08] or the IEEE 1900.4 [Std09] standards. They provide mobile users with information on serving RATs, but do not make RAT selection decisions. Furthermore, inter-RAT handovers can be handled using MIP in loosely coupled heterogeneous wireless networks, and 3GPP mobility management protocols in tightly coupled networks. They ensure seamless service continuity. However, RAT selection decisions remain a challenging task that will be addressed throughout this thesis.

2.2 RAT Selection Criteria

RAT selections are usually based on user radio conditions (*i.e.,* received-signal-strength measurements), and resource availability. Yet, to maximize network performance and user experience, decisions need to involve additional criteria. The main are as follows:

- User QoS requirements
- Network load conditions
- Network and user energy consumption
- Operator and user preferences: cost and various subjective criteria

When multiple criteria are involved, particularly as some vary dynamically, RAT selection turns out to be a complex decision-making problem. Further, in [GAM05], RAT selection has been isomorphically mapped to a multiple choice multiple dimension knapsack problem, known to be NP-hard.

2.3 RAT Selection Objectives

Network elements and mobile users are able to make RAT selections either autonomously or collectively. Decision objectives are usually defined as utility or cost functions, and decision makers are regarded as utility maximizers or cost minimizers, respectively. In [KIC+11], RAT selection is formulated as a non-linear optimization problem. The network assigns persistent elastic users to either WiFi or WiMAX in a way to minimize global network cost, defined as the sum of individual user costs. The cost function represents user service time, that is the expected amount of time required to send a data unit. Therefore, it depends on user radio conditions and network load conditions. Further in [KIC+11], RAT selection is portrayed as a non-cooperative game. Mobile users selfishly strive to minimize their cost. They try to reach a Nash equilibrium strategy, where no mobile can decrease its cost by changing only its serving RAT. To wrap up, RAT selections are expected either to minimize decision maker costs or to maximize decision maker utilities.

Generally, utility and cost functions reflect operator interests and user experience. They describe the suitability of RAT selections with respect to one or multiple decision maker objectives. In [CKG08a], the network routes elastic users to either WiFi or HSDPA depending on user spatial distribution and network load conditions. Decisions are expected to maximize long-term average network utility, defined as the sum of individual user utilities and a blocking cost. The user utility function represents user satisfaction, and mainly depends on user throughput. The blocking cost is the penalty inflicted on the network when blocking arriving mobiles. In other words, network decisions aim to maximize the sum of user satisfactions and to minimize the user blocking probability. However, accepting more elastic users, particularly with unfavorable radio conditions, may reduce user throughputs and thereafter the sum of user satisfactions. Therefore, for example due to the CSMA/CA-based multiple access technology in WiFi, it may be better to block a user with bad radio conditions so as not to penalize all individual user throughputs. Decision maker objectives are then heterogeneous and potentially conflicting: the sum of user satisfactions and the user blocking probability need to be weighted and normalized. In [CKG08a], the user blocking probability is multiplied with a normalization and weighting factor.

Moreover, to deal with heterogeneous and conflicting objectives, multi-criteria decision-making methods are introduced [KKP08, SNW06]. Decisions are expected to maximize multi-criteria utility functions, that depend on weighted and normalized decision parameters. Various normalization techniques are implemented. For illustration, we assume that mobiles autonomously select their serving RAT, in a way to jointly maximize their throughput and minimize their monetary cost. They can choose between WiMAX and LTE, designated by W and L respectively. The throughput a user can achieve, when connected to WiMAX and LTE, is denoted by $d(W)$ and $d(L)$ respectively. Also, the cost a user pay, when connected to WiMAX and LTE, is denoted by $cost(W)$ and $cost(L)$ respectively. Mobiles estimate a utility function for each of the two serving RATs, and select the one with the highest score. However, using the Simple Additive Weighting (SAW) method, the expected utility of RAT x denoted by $U(x)$, $x \in \{W, L\}$, is defined as the weighted sum of the normalized decision parameters:

$$U(x) = w_d \cdot \hat{d}(x) + w_{cost} \cdot \widehat{cost}(x)$$

where w_p and $\hat{p}(x)$, $p \in \{d, cost\}$, respectively represent the weight and the normalized value of decision parameter p. Besides, $\hat{d}(x)$ and $\widehat{cost}(x)$ are defined as follows:

$$\hat{d}(x) = \frac{d(x)}{\max\left[d(W), d(L)\right]}$$

$$\widehat{cost}(x) = \frac{\min\left[cost(W), cost(L)\right]}{cost(x)}$$

Furthermore, using the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method, the utility function $U(x)$, $x \in \{W, L\}$, represents the relative closeness to the ideal solution, and is defined as follows:

$$U(x) = \frac{S^-(x)}{S^-(x) + S^+(x)}$$

where $S^+(x)$ and $S^-(x)$ respectively denote the distance of alternative x from the positive ideal and the negative ideal solution, and are defined as follows:

$$S^+(x) = \sqrt{\left\{w_d \cdot \left(\hat{d}(x) - \max\left[d(W), d(L)\right]\right)\right\}^2 + \left\{w_{cost} \cdot \left(\widehat{cost}(x) - \min\left[cost(W), cost(L)\right]\right)\right\}^2}$$

$$S^-(x) = \sqrt{\left\{w_d \cdot \left(\hat{d}(x) - \min\left[d(W), d(L)\right]\right)\right\}^2 + \left\{w_{cost} \cdot \left(\widehat{cost}(x) - \max\left[cost(W), cost(L)\right]\right)\right\}^2}$$

The normalized decision parameters $\hat{d}(x)$ and $\widehat{cost}(x)$ are however defined as:

$$\hat{d}(x) = \frac{d(x)}{\sqrt{d(W)^2 + d(L)^2}}$$

$$\widehat{cost}(x) = \frac{cost(x)}{\sqrt{cost(W)^2 + cost(L)^2}}$$

Many other multi-criteria decision-making methods and normalization techniques can be found in [KKP08, SNW06]. Yet, as for SAW and TOPSIS, resultant decisions exclusively depend on user preferences (*e.g.,* weights of the decision criteria), as well as on the characteristics of available alternatives. As they ignore user QoS requirements and cost tolerance, state-of-the-art methods often make inadequate decisions. To overcome this limitation, we introduce in Chapter 3 a satisfaction-based multi-criteria decision-making method. In addition to user preferences, our algorithm considers user needs (*e.g.,* traffic class, throughput demand, and cost tolerance parameter), meeting user objectives.

2.4 RAT Selection Approaches

RAT selection has triggered considerable interest among researchers in the past few years. In this section, we review some relevant work, and classify them into network-centric and mobile-terminal-centric approaches, according to who makes decisions.

2.4.1 Network-centric Approaches

Network elements, namely either centralized or decentralized CRRM entities, collect necessary measurements and information. They take selection decisions transparently to end-users, in a way to meet operator objectives. In [PRSA05], mobiles are associated with their RAT according to straightforward allocation principles. Voice GSM/EDGE (VG), Voice UMTS (VU), Indoor (IN), and VG*IN policies are presented: they associate mobiles with either GSM/EDGE or UMTS, based on their service types (*i.e.,* voice or data), and eventually on their radio conditions.

In [KIC+11, SWMG08, PK06], RAT assignment is formulated as an optimization problem. Exact and heuristic algorithms are used to derive an optimal or a near optimal solution, that optimizes global network utility or cost. In [KIC+11], the global network cost is defined as the sum of individual user service times, and depends on user radio conditions and network load conditions. In [SWMG08], the network utility represents the network revenue, and is expressed as the sum of individual user utilities. Further, the user utility is a concave, non-decreasing function of user throughput. In [PK06], the network utility accounts for user TCP throughputs, and depends on user radio conditions, TCP packet size, channel access parameters (*e.g.,* backoff window and inter-frame space in WiFi), and network load conditions. Moreover, in [LEnGSS12], RAT selection and resource allocation are simultaneously performed. The proposed CRRM algorithm considers the discrete nature of radio resources, and is thus based on integer linear programming optimization techniques. Radio resources, namely GPRS and EDGE time slots, and HSDPA codes, are distributed in a way to maximize the lowest user utility. The user utility function represents user throughput for web and email services, and the percentage of correctly transmitted video frames for real-time video services. In [GLEnSS12], based on the CEA (Constrained Equal Awards) bankruptcy rule, selection decisions try to equally satisfy mobile users: they are assigned the same amount of resources, without exceeding their individual demands.

In [ZJJ+12, ZYNT12a, ZYNT12b, SAAS10, IKT09, GPRSA08, CKG08a, CKG08b, KAK06, YK05], a Semi-Markov Decision Process (SMDP) is used to model the RAT selection problem. A set of states, actions, rewards, and transition probabilities are defined. Linear or dynamic programming algorithms are adopted to find out an optimal access policy, that maximizes long-term network reward function (*i.e.,* an expected utility calculated over an infinitely long trajectory of the Markov chain). In [CKG08a], elastic users are assigned to either WiFi or HSDPA depending on user spatial distribution, and network load conditions. WiFi access points and HSDPA base stations are co-localized, and their cells are assumed to be overlapping. The peak throughput a user can obtain, when present alone in the cell, differs depending on its geographical position. Therefore, as illustrated in Fig. 2.1, the cell is divided into r rings with homogeneous radio characteristics. Users in ring i, $i = 1, ..., r$, have a peak throughput of $\tilde{D}_i^1$ and $\tilde{D}_i^2$ when connected to WiFi and HSDPA, respectively.

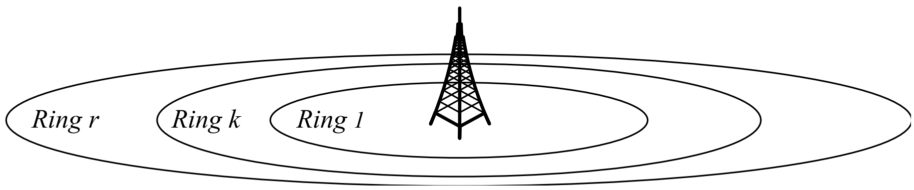

Figure 2.1: WiFi and HSDPA cell divided into r concentric rings

Network states are the $2r$-tuple $(n_1^1, ..., n_r^1, n_1^2, ...n_r^2)$, where n_i^j represents the number of mobiles in ring i that are connected to RAT j. WiFi and HSDPA are designated by $j = 1$ and $j = 2$, respectively. To ensure that all mobiles achieve an acceptable throughput, the number of mobiles that can be assigned to serving RATs is limited: $\sum_{i=1}^{r} n_i^j \leq n_{max}^j$. Moreover, decisions are state-dependent, and are expected to maximize long-term average network utility calculated per time unit. The network utility function depends on user throughputs and blocking probability. Using the *Policy Iteration* algorithm, an optimal policy is solved. It determines, for each state s, the action $a(s) = (a_1, a_2)$ to take, where a_j, $j \in \{1, 2\}$, is equal to 1 if arriving mobile is accepted in RAT j, and is null otherwise. Furthermore, in addition to network load conditions and user radio conditions, decisions in [IKT09] involve user traffic classes. State actions depend on the location and traffic class of arriving mobiles.

In [HBJG07], a fuzzy multi-criteria decision algorithm, based on simple *If X and Y then Z* rules, is presented. Individual decisions, resulting from various fuzzy rules, are aggregated to provide RAT selections. Pattern aspects (*i.e.,* fuzzy inference rules, membership functions, and their shapes) are however empirical, and rely on prior field experience. Moreover, in [GAPRS09, GAPRS08, GAPRS06], a fuzzy neural solution is introduced to jointly decide of the RAT selection and the bandwidth allocation. A reinforcement signal is generated to optimize the decision-making process: the means and the standard deviations of the input and output bell-shaped membership functions are adjusted accordingly.

As network elements gather information about individual users, namely their QoS needs, and their radio conditions in the different serving cells, network-centric approaches generally optimize resource utilization. Yet, network complexity, processing, and signaling load are drastically increased.

2.4.2 Mobile-terminal-centric Approaches

Rational users select their RAT depending on their needs and preferences, in a way to selfishly maximize their utility. Mobile-terminal-centric heuristics are proposed in [MILK12]. Distance-based, probabilistic distance-based, peak rate-based, and probabilistic peak rate-based algorithms are introduced: they indicate the probability to assign mobiles to the primary (IEEE 802.11g) and to the secondary (IEEE 802.11b) RATs, based on their distance from the two access points, or on the peak rate they can achieve when connected to these access points.

As users utility does not only depend on their own decisions, but also on the decisions of other mobiles, game theory is used as a theoretical framework to model user interactions in [AKHWC13, KIC$^+$11, IKT10, NH09, CTG09, Erc08, KAK06]. Players (*i.e.,* the individual users) try to reach a mutually agreeable solution, or equivalently a set of strategies they unlikely want to change. However, the convergence time to the equilibrium assignment seems to be long [KIC$^+$11].

In [NVACT13, CM12, FC11, WB09, BL07, SNW06, SJ05b, SJ05a, Zha04], multi-criteria decision-making methods, including SAW, TOPSIS, Multiplicative Exponent Weighting (MEW) and Grey Relational Analysis (GRA), are presented. They capture the heterogeneity of decision criteria (*e.g.,* QoS, cost, energy, and security parameters). Users with widely varying requirements gather their QoS information (*e.g.,* peak throughput when connected alone to a cell), calculate decision metrics, and select their RAT accordingly. In [Zha04, FC11, CM12], fuzzy logic is also used to deal with the imprecise information of some criteria and user preferences.

In [TFC12, DO12b, DO12a], RAT selection is formulated as a reinforcement learning problem. A set of states, actions, and rewards are defined. Mobiles iteratively learn selection decisions, through trial-and-error interaction with their environment, in a way to maximize their utility. They discover a variety of actions, and progressively favor effective ones.

As mobiles autonomously select their RAT, network operations remain reduced. Furthermore, decisions can easily involve user needs and preferences, and various mobile-terminal-related parameters. However, when mobiles do not cooperate, mobile-terminal-centric approaches potentially lead to performance inefficiency.

In Chapter 5, we investigate some network-centric and mobile-terminal-centric approaches, and compare their performance in terms of network and user utilities. While the network utility function is defined as the sum of user throughputs, the user utility function represents user satisfaction, and mainly depends on user QoS and cost parameters.

2.4.3 Incentives for Hybrid Approaches

Network-centric approaches can optimize operator objectives, but at the cost of increased network complexity, signaling and processing burden. An illustrative example is found in [IKT09], where RAT selection is formulated as a semi-Markov decision process. Mobiles are associated with either UMTS or WiMAX, based on cell load measurements, user radio conditions (*i.e.,* spatial distribution), and QoS needs (*i.e.,* traffic class, and throughput demands). Optimal decisions are derived in a way to maximize the long-term network reward. This, however, increases processing and signaling load, particularly as information about individual mobiles need to be gathered.

Moreover, mobile-terminal-centric approaches have also gained in importance. Mobile users autonomously select their RAT in a way to maximize their own utility. However, as mobiles do not cooperate, mobile-terminal-centric approaches are known for their potential inefficiency. Although mobiles strive to selfishly maximize their utility, their decisions may be in no one long-term interest. This dilemma is known as the *Tragedy of the commons* [Har68]. A simple example is found in [SJ05b, SJ05a], where Analytic Hierarchy Process (AHP) and Grey Relational Analysis (GRA) are integrated to introduce a multi-criteria decision-making method for RAT selection. When mobiles have no information on network load conditions, they use static QoS parameters (*e.g.,* peak throughput when connected alone to a cell) to evaluate serving RATs. In real networks, this obviously lead to congestion and overload conditions.

Our challenge is then to design a RAT selection approach, that jointly enhances network performance and user experience, while signaling and processing burden remains reduced. In this thesis, we propose an innovative hybrid decision method, that combines benefits from both network-centric and mobile-terminal-centric approaches. The network provides a common information for the mobiles to make robust RAT selections. Network load conditions are masked, and only monetary cost and QoS incentives to join serving RATs are provided. As radio resources may be heterogeneous in nature, such as GPRS and EDGE time slots, HSPA codes, power and allocation times, and LTE OFDMA slots, QoS incentives need to be homogenized: they are expressed as minimum guaranteed throughputs and maximum throughputs. Further, mobile users select their RAT depending on their individual needs and preferences, as well as on the cost and QoS parameters signaled by the network. By appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states.

As a matter of fact, our hybrid approach involves two inter-dependent decision-making processes. The first one, on the network side, consists in deriving appropriate network information, so as to guide user decisions in a way to meet operator objectives. The second one, where individual users combine their needs and preferences with the signaled network information, consists in selecting the RAT to be associated with, in a way to maximize user utility. Since, in their turn, user individual decisions influence the upcoming network information, the two decision makings are considered to be inter-dependent. Thus, RAT selections dynamically involve operator objectives, and user needs and preferences.

2.5 Conclusion

In this chapter, we reviewed the main RAT selection methods, and classified them into network-centric and mobile-terminal-centric approaches, according to who makes decisions. We then outlined the benefits and drawbacks of each approach. In Chapter 3, we introduce a new hybrid decision method, that:

- minimizes network complexity, signaling and processing burden: RAT selections are pushed towards the mobiles. However, a common network information assists them in their decisions.
- efficiently utilizes radio resources, despite of the non-cooperative behavior of mobile users: by appropriately tuning network information, user decisions are globally expected to meet operator objectives (*e.g.,* enhance resource utilization).

Chapter 3

A Hybrid Approach for RAT Selection

In this chapter, we tackle the RAT selection problem in heterogeneous wireless networks, and propose a hybrid decision approach. Mobile users are assisted in their decisions by the network, that broadcasts monetary cost and QoS information. Two inter-dependent decision-making problems are thus brought into play. The first one, on the network side, consists in deriving appropriate network information, so as to guide user decisions in a way to meet operator objectives. The second one, on the user side, consists in selecting the RAT to be associated with, in a way to maximize user utility. We first focus on the user side, and present a satisfaction-based multi-criteria decision-making method. Mobiles select their RAT depending on their needs and preferences, as well as on the cost and QoS parameters signaled by the network. In comparison with existing solutions, our algorithm meets user needs (e.g., traffic class, throughput demand, and cost tolerance), avoiding inadequate decisions. Further, we introduce two heuristic methods, namely the staircase and the slope tuning policies, to dynamically derive what to signal to mobiles, so as to enhance resource utilization.

3.1 Hybrid Decision Framework

3.1.1 Network Topology

We consider a heterogeneous wireless network composed of N_T RATs. The modulation and coding scheme, that can be assigned to a user connected to RAT x, differs depending on its radio conditions in the cell, more precisely on its signal-to-noise ratio denoted by SNR^x. As the number of possible modulation and coding schemes is limited, we decompose the cell into N_Z^x zones with homogeneous radio characteristics [IKT09, CKG08a, CKG08b].

Users in zone Z_k^x, $k = 1, ..., N_Z^x$, are assumed to have a signal-to-noise ratio between δ_k^x and δ_{k-1}^x, and then to use $mod^x(k)$ with $cod^x(k)$ as modulation and coding scheme:

$$(mod^x(k), cod^x(k)) = \begin{cases} \text{none} & \text{if } SNR^x(k) < \delta^x_{N_Z^x}, \\ (mod^x_{N_Z^x}, cod^x_{N_Z^x}) & \text{if } \delta^x_{N_Z^x} \leq SNR^x(k) < \delta^x_{N_Z^x-1}, \\ \cdots & \\ (mod_1^x, cod_1^x) & \text{if } \delta_1^x \leq SNR^x(k) < \delta_0^x = \infty. \end{cases} \tag{3.1.1}$$

where $\delta^x_{N_Z^x}$ is the minimum signal-to-noise ratio, that allows transmission at the lowest throughput, given a target error probability.

Furthermore, and for the sake of simplicity, users in a same zone are assumed to have the same peak throughput, realized when present alone in the cell.

In the remainder, let the N_Z^x-tuple $n^x = (n^x(k))$, for $k \in \{1, ..., N_Z^x\}$, be the state of RAT x. $n^x(k)$ represents the number of users, in zone Z_k^x, that are connected to RAT x. The state s of the heterogeneous wireless network is the concatenation of RAT x substates, for $x \in \{1, ..., N_T\}$: $s = (n^x)$, for $x \in \{1, ..., N_T\}$.

3.1.1.1 Cell Decomposition

Because of fading effects, radio conditions are time-varying. User signal-to-noise ratio can take all possible values, leading to different modulation and coding schemes. However, as RAT selections are made for a sufficiently long period of time (*e.g.,* session duration, user dwell time in the cell), users are distributed over logical zones depending on their average radio conditions, rather than on their instantaneous ones.

Another approach is found in [IKT09], where an analytical radio model, that accounts for interference, path loss, and Rayleigh fading, is used. It has been demonstrated that users need to be situated at $r_k \in [R_{k-1}^x, R_k^x[$ from their base stations, so as to have a signal-to-noise ratio between δ_k^x and δ_{k-1}^x, with at least a high probability $\mathbb{P}_{th}$. This means that the cell may be divided into concentric rings, as illustrated in Fig. 3.1, and mobiles in ring Z_k^x will use $mod^x(k)$ with $cod^x(k)$ as modulation and coding scheme, with at least a high probability $\mathbb{P}_{th}$. Further, to define the different rings, the distances R_k^x have been analytically derived, mainly as a function of δ_k^x, $\mathbb{P}_{th}$, and radio model parameters.

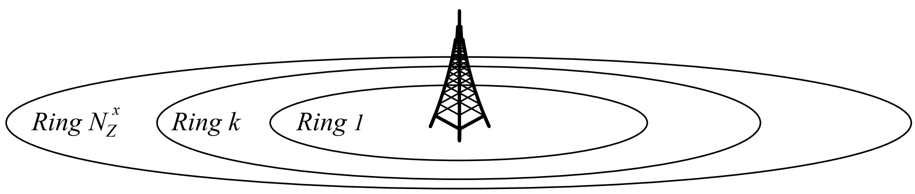

Figure 3.1: RAT x cell divided into N_Z^x concentric rings

3.1.2 Network Resources

Prior to the RAT selection process, a common admission control is assumed to be performed. New and handover sessions are admitted to the extent that joint available resources are able to meet their requirements, while not compromising the QoS level of ongoing ones. Further, after sessions are accepted, decisions are made as to what RAT they should be associated with. Robust decisions are crucial to avoid network congestion, and enhance user experience.

In RAT x, the radio resource is divided into elementary resource units (RU). Typically, in OFDM(A)-based technologies (*e.g.,* LTE and WiMAX), resource units are defined as OFDM symbols (one-dimensional allocations), or OFDMA slots (two-dimensional allocations: m subcarriers by n OFDMA symbols). However, in CDMA-based technologies (*e.g.,* HSPA), codes, power and allocation times are regarded as RUs.

In the time domain, transmissions are organized into radio frames of length T^x. At each scheduling epoch, RUs are allocated to individual users, based on a predefined scheduling algorithm. User throughputs depend on their allocated RUs (*i.e.,* their description and amount), and modulation and coding schemes. Typically, when fair time scheduling is employed, cell resources (*e.g.,* codes, power and allocation times in HSPA, OFDMA slots in LTE) are equally distributed to mobile users [THK$^+$10]. Yet, mobiles with good radio conditions (*e.g.,* cell center users) experience a higher throughput than those with bad radio conditions (*e.g.,* cell edge users).

3.1.3 Network Information

Periodically or on user request, network information is sent to all mobiles, using the logical communication channel (*i.e.,* radio enabler) proposed by the IEEE 1900.4 standard [Std09]. This logical channel allows information exchange between the Network Reconfiguration Manager (NRM) on the network side, and the Terminal Reconfiguration Manager (TRM) on the mobile-terminal side (Fig. 3.2). The purpose is to improve resource utilization and user experience in heterogeneous wireless networks.

In our work, by appropriately tuning network information, the network globally controls user decisions, in a way to meet operator objectives (*e.g.,* enhance network performance, minimize energy consumption). Network information may then be static or dynamic, so as to optimize short- or long-term network utility.

When a new or a handover session arrives, the mobile decodes network information, evaluates serving RATs, and selects the one that maximizes its own utility. As a matter of fact, selection decisions depend on user needs and preferences, as well as on the signaled network information.

The network is fully described by its state s. Yet, in our work, only monetary cost and partial QoS parameters are sent to mobiles. This reduces signaling load. Furthermore, by masking RAT load conditions, QoS information may reflect not only the current network state s, but also other network-related parameters (*e.g.,* energy consumption). For instance, QoS parameters may be tuned, so that mobile decisions are consistent with operator energy-saving objectives. This flexible design allows the network to derive cost and QoS parameters in a way to optimize a generic utility function.

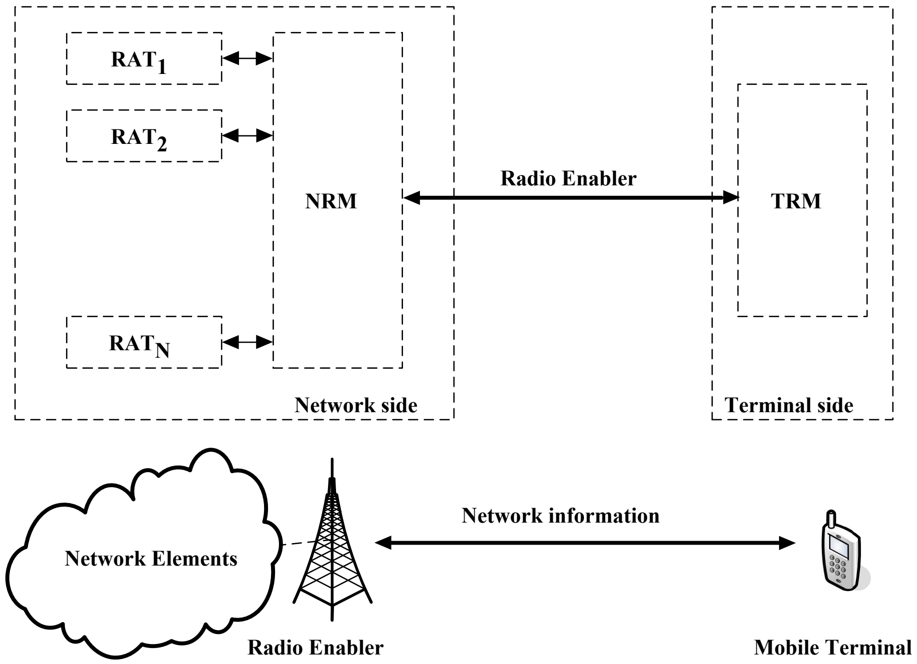

Figure 3.2: Hybrid 1900.4 network architecture

Moreover, cost and QoS parameters, signaled by the network, are seen as incentives to join serving RATs:

- Cost parameters: Because flat-rate pricing strategies waste resources [EV99], result in network congestion, and thus degrade network performance [ZYA04], they are not optimal in supporting QoS. A volume-based model is therefore proposed: mobile users are charged based on the amount of traffic they consume. In our work, *costs* are defined on a per kbyte basis.

- QoS parameters: The amount of resource units (RUs) to be allocated to future arrivals are broadcasted:

  - Mobiles are guaranteed an average minimum amount of RUs, denoted by n_{min}.
  - They also have priority to occupy up to an average maximum amount of RUs, denoted by n_{max}.

Because the smallest allocation unit (*i.e.,* RU) has different descriptions in the different RATs, there is a need to homogenize the QoS information. QoS parameters are then expressed as throughputs: d_{min} and d_{max} instead of n_{min} and n_{max}. However, as user throughputs strongly depend on their radio conditions, d_{min} and d_{max} are derived for the most robust modulation and coding scheme (*i.e.,* $mod^x_{N^x_Z}$ with $cod^x_{N^x_Z}$).

Therefore, when evaluating serving RATs, mobiles should combine their individual radio conditions with the provided QoS parameters: for that, they multiply d_{min} and d_{max} with a given modulation and coding gain, denoted by $g(M, C)$.

Although QoS parameters are provided, our decision framework is independent of local resource allocation schemes. First, the minimum guaranteed RUs, namely n_{min}, are directly granted. Then, any priority scheduling algorithm, including opportunistic schemes [Kha06, GB09, KM10, NHT12], could be adopted to share out remaining resources. Grants are, however, limited to n_{max}. Residual resources are afterwards equitably distributed: when all mobiles have received their maximum throughput, they are considered to have the same priority, leading to fair allocation.

3.1.4 RAT Selection

The network proposes one or more alternatives, that are the available RATs. For each alternative a, the network broadcasts the three parameters: $d_{min}(a)$, $d_{max}(a)$, and $cost(a)$. From the user point of view, these parameters are the decision criteria to be used to evaluate serving RATs. As in all multi-criteria decision making methods, mobiles define and compute a utility function for all of the available alternatives. This utility is obtained after normalizing and weighting the decision criteria.

In the next section, we present our Satisfaction-Based (SB) Multi-Criteria Decision-Making (MCDM) method. The particularity of our algorithm resides in the normalization step, that takes into account user needs (*i.e.,* traffic class, throughput demand, cost tolerance). By avoiding inadequate decisions, our algorithm overcomes some limitations of well-known MCDM methods.

3.2 Satisfaction-based Decision Method

3.2.1 Normalization and Traffic Classes

The normalization of the decision criteria $d_{min}(a)$, $d_{max}(a)$, and $cost(a)$ takes into consideration session traffic class, throughput demand, and cost tolerance. For traffic class c

and alternative a, $\hat{d}^c_{min}(a)$, $\hat{d}^c_{max}(a)$, and $\widehat{cost}^c(a)$ are respectively the normalized values of $d_{min}(a)$, $d_{max}(a)$, and $cost(a)$.

In our work, we define three traffic classes : inelastic, streaming, and elastic classes. Before we give the normalizing functions for each traffic class, we note that $\hat{p}^c(a), p \in \{d_{min}, d_{max}, cost\}$, can be viewed as the expected satisfaction of a class c session, with respect to criterion p, when alternative a is selected:

- Inelastic sessions ($c = I$): since designed to support constant bit rate circuit emulation services, inelastic sessions require stringent and deterministic throughput guarantees. d_{max} should have no impact on RAT selections. Besides, the satisfaction with respect to d_{min} has a step shape (Fig. 3.3). When alternative a is selected, mobiles expect to be satisfied provided that their minimum guaranteed throughput $d_{min} = d_{min}(a) \cdot g(M, C)$ is greater or equal to their fixed throughput demand R_f; otherwise, they are not satisfied.

$$\hat{d}^I_{min}(a) = \begin{cases} 0 & \text{if } d_{min}(a) \cdot g(M, C) < R_f \\ 1 & \text{if } d_{min}(a) \cdot g(M, C) \geq R_f \end{cases} \tag{3.2.1}$$

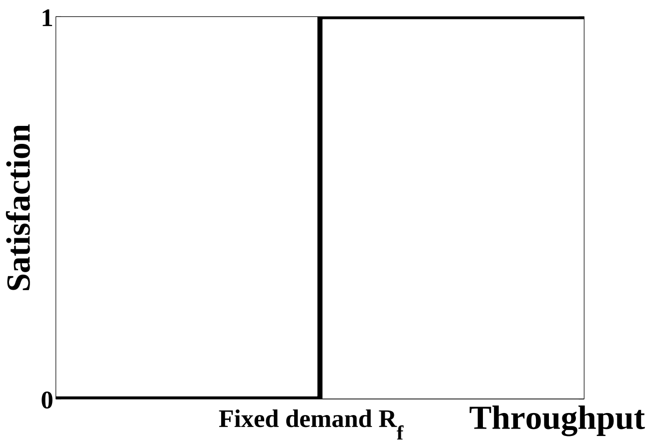

Figure 3.3: Inelastic sessions: Throughput satisfaction function

- Streaming sessions ($c = S$): since designed to support real-time variable bit rate services (*e.g.,* MPEG-4 video service), streaming sessions are fairly flexible, and usually characterized by a minimum, an average and a maximum throughput requirement. Therefore, when alternative a is selected, their expected satisfaction with respect to d_{min} and d_{max} is represented by an S-shaped function (Fig. 3.4):

$$\hat{d'}^S(a) = 1 - \exp\left(\frac{-\alpha(\frac{d'(a).g(M,C)}{R_{av}})^2}{\beta + (\frac{d'(a).g(M,C)}{R_{av}})}\right) \tag{3.2.2}$$

where $d' = \{d_{min}, d_{max}\}$.

R_{av} represents session needs: an average throughput demand. α and β are two positive constants necessary to determine the shape of the S-shaped function.

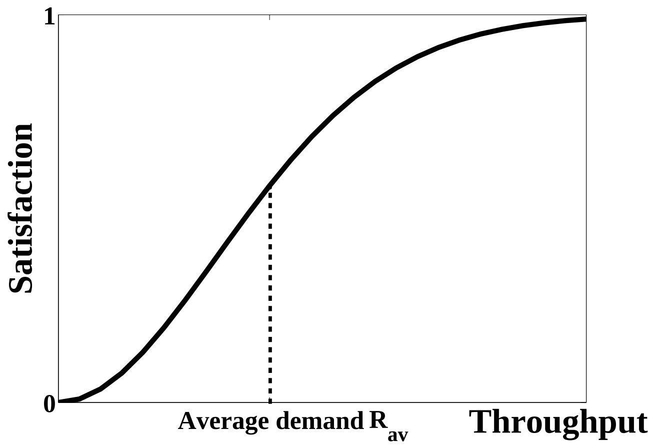

Figure 3.4: Streaming sessions: Throughput satisfaction function, ($\alpha = 9$, $\beta = 10$)

- Elastic sessions ($c = E$): since designed to support traditional data services (*e.g.,* file transfer, email and web traffic), elastic sessions typically using the TCP protocol adapt to resource availability. As they require no QoS guarantees, d_{min} has no impact on RAT selections. Moreover, the satisfaction with respect to d_{max} has a concave shape as illustrated in Fig. 3.5.

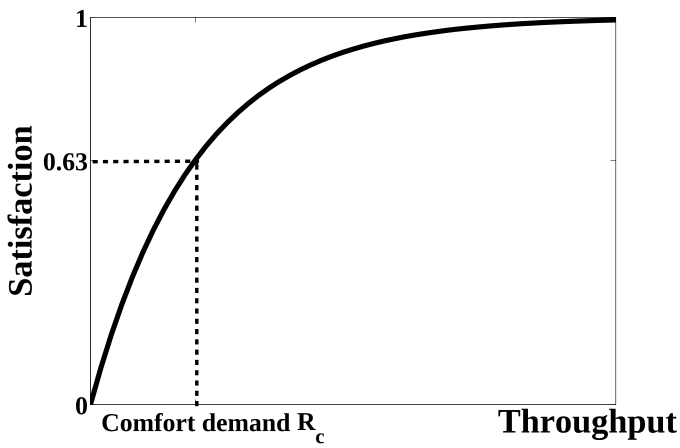

Figure 3.5: Elastic sessions: Throughput satisfaction function

User satisfaction is expected to increase slowly as its throughput exceeds its comfort throughput demand R_c (*i.e.,* the mean throughput beyond which user satisfaction exceeds 63% of maximum satisfaction):

$$\hat{d}^{E}_{max}(a) = 1 - \exp(-\frac{d_{max}(a).g(M,C)}{R_c}) \qquad (3.2.3)$$

Furthermore, the monetary cost satisfaction is modelled as a Z-shaped function for all sessions (Fig. 3.6): the slope of the satisfaction curve increases rapidly with the cost.

$$\widehat{cost}^{c}(a) = \exp(-\frac{cost(a)^2}{\lambda^c}), c \in \{I, S, E\} \qquad (3.2.4)$$

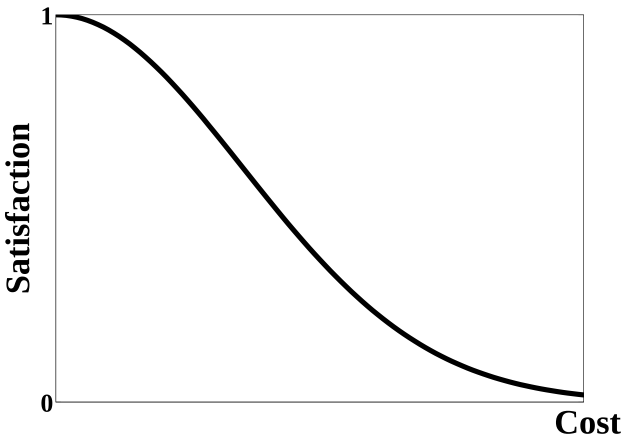

Figure 3.6: Monetary cost satisfaction function, ($\lambda^c = 25$)

λ^c represents the cost tolerance parameter: a positive constant to determine the shape of the Z-shaped function.

3.2.2 User Profile and Utility Function

The user profile defines the cost tolerance parameter and the weights to be applied to normalized criteria. More precisely, the user profile is the set of vectors ($\lambda^c, w^c_{d_{min}}, w^c_{d_{max}}$, $w^c_{cost}), c \in \{I, S, E\}$, where w^c_p is the weight of $\hat{p}^c, p \in \{d_{min}, d_{max}, cost\}$. When alternative a is selected, the expected utility of a class c session is defined as follows:

$$U^c(a) = w^c_{d_{min}} \cdot \hat{d}^c_{min}(a) + w^c_{d_{max}} \cdot \hat{d}^c_{max}(a) + w^c_{cost} \cdot \widehat{cost}^{c}(a)$$

Note that predefined user profiles (*e.g.,* cost minimizing profile, QoS maximizing profile) may be introduced. Thereby, end-users do not worry about technical details: they can use default values for the cost tolerance parameter, and the decision criteria weights.

Fig. 3.7 summarizes the decision process:

- For each alternative a, the mobile combines its radio conditions with the QoS parameters signaled by the network: it multiplies $d_{min}(a)$ and $d_{max}(a)$ with a given modulation and coding gain, to determine its perceived QoS parameters, as provided by the network.

- Then, based on user needs (*i.e.*, traffic class c, throughput demand and cost tolerance λ), it computes the normalized decision criteria: $\hat{d}^c_{min}(a)$, $\hat{d}^c_{max}(a)$ and $\widehat{cost}^c(a)$.
- Next, it combines user preferences (*i.e.*, $w^c_{d_{min}}$, $w^c_{d_{max}}$ and w^c_{cost}) with the normalized decision criteria, so as to compute the weighted normalized criteria: $w^c_{d_{min}} \cdot \hat{d}^c_{min}(a)$, $w^c_{d_{max}} \cdot \hat{d}^c_{max}(a)$ and $w^c_{cost} \cdot \widehat{cost}^c(a)$.
- Finally, it computes the utility function for each alternative a, and selects the one with the highest score.

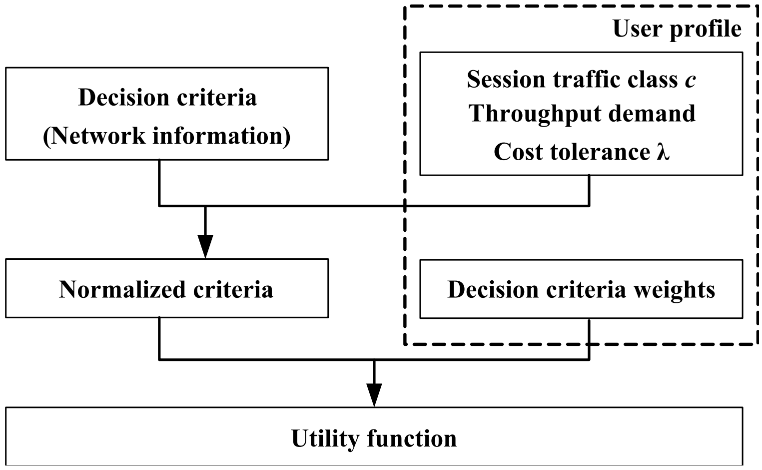

Figure 3.7: Satisfaction-based multi-criteria decision process

This decision process is performed at session initiation and possibly also during session lifetime. Mobiles decide of their serving RAT based on their individual needs and preferences, as well as on the broadcasted network information. However, they can migrate to another RAT following changes in their radio conditions. At this point, mobiles check whether their serving RAT is still their best choice, or in other words, whether it is still expected to maximize user utility. An inter-RAT handover is triggered only when another RAT can provide users with significantly higher satisfaction level. This helps to reduce unnecessary handovers (*i.e.*, ping-pong effect).

3.3 Tuning Policies

Because mobile users also rely on their needs and preferences when selecting their RAT, the network does not completely control individual decisions. Yet, by signaling appropriate decisional information, the network tries to globally guide user decisions, in a way to meet operator objectives. These may include energy savings: mobiles are pushed to some base stations, while others are switched to sleep mode so as to save energy. In our work, we assume that operators are only concerned by efficiently utilizing their radio resources: providing better network performance, higher user satisfaction, and larger operator gain.

When a RAT dominates all the others (*i.e.,* provides higher QoS parameters for the same cost, or the same QoS parameters for a lower cost), common radio resources are inefficiently utilized, causing performance degradation. In fact, mobile users would select the dominant alternative, leading to unevenly distributed traffic load. While a RAT is overcrowded, the others are almost unexploited. This inefficiency is very similar to that of the mobile-terminal-centric approaches. To avoid it, QoS parameters, signaled by the network, needs to be modulated as a function of the load conditions.

In this section, we present two heuristic methods, namely the staircase and the slope tuning policies, to dynamically derive QoS information. In order to reduce network complexity and processing load, one of the drawbacks of network-centric approaches, our policies are made simple. Yet, they help to efficiently distribute traffic load over the available RATs, and thus to better utilize radio resources.

3.3.1 Staircase Tuning Policy

The load factor represents the amount of throughput guarantees, and is defined as the ratio of the number of guaranteed allocated RUs to the total number of RUs. Fig. 3.8 illustrates how QoS parameters, namely d_{min} and d_{max} separately, are tuned as a function of the load factor using the Staircase policy. When RAT x load factor is low, the network can promise high throughput guarantees to arriving mobiles to join RAT x. The highest $d_{min}(x)$ and $d_{max}(x)$ values are signaled. However, as RAT x load factor exceeds S_1 threshold, the network needs to reduce QoS incentives in RAT x so as to avoid RAT x congestion, or in other words, to avoid resource shortage in RAT x. QoS parameters are separately decreased, following a step function. Moreover, as S_2 is reached, the network no longer provides incentives to arriving mobiles in RAT x.

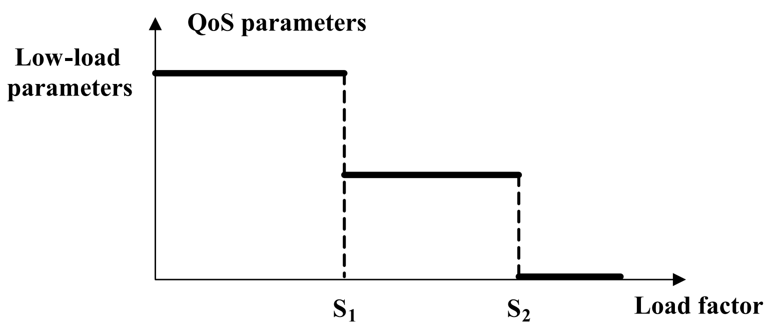

Figure 3.8: QoS parameters reduction using the Staircase policy

Usually, d_{min} and d_{max} have different values. For instance, at low load factor, $d_{min}(x)$ and $d_{max}(x)$ are equal to 1 and 1.5 Mb/s, respectively. They are respectively reduced to 0.5 and 1 Mb/s as S_1 is reached, and are both set to zero when S_2 is exceeded. Furthermore, it is worth noting that the different serving RATs can have different S_1 and S_2 values.

3.3.2 Slope Tuning Policy

As radio access technologies are progressively loaded, the Slope policy gradually tune QoS parameters as a function of the load factor (cf. Fig. 3.9). When RAT x load factor is low, the highest $d_{min}(x)$ and $d_{max}(x)$ values are signaled. Yet, when S_1 is reached, QoS parameters are linearly and separately reduced down to zero. The slope helps to better respond to traffic load fluctuations.

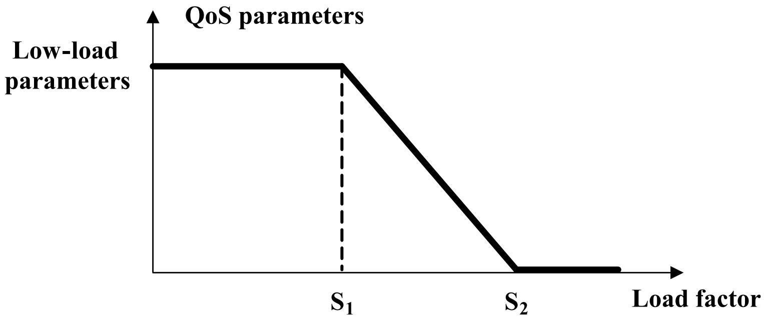

Figure 3.9: QoS parameters reduction using the Slope policy

As QoS parameters are dynamically modulated, arriving mobiles are pushed to the less loaded RATs, enhancing long-term network performance. However, using both policies, the challenge is to properly set S_1 and S_2. In the same load conditions, QoS parameters to signal strongly depend on tuning threshold values. In other words, for a given load factor, different d_{min} and d_{max} can be provided depending on S_1 and S_2, leading to different user decisions. The impact of S_1 and S_2, on network and user utilities, are further discussed in Chapter 5.

3.4 Conclusion

In this chapter, we proposed a new hybrid decision approach for RAT selection in heterogeneous wireless networks. Using the logical communication channel proposed by the IEEE 1900.4 standard, the network provides information for the mobiles to make robust RAT selections. More precisely, mobile users select their RAT depending on their needs and preferences, as well as on the cost and QoS parameters signaled by the network. By appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states. We first presented a satisfaction-based multi-criteria decision-making method, so that mobiles can efficiently evaluate serving RATs. Then, we introduced two heuristic methods, namely the staircase and the slope tuning policies, to derive QoS parameters as a function of the load conditions. They follow a linear decreasing (slope) or a staircase function.

In Chapter 4, our satisfaction-based multi-criteria decision-making method, and our tuning heuristics are thoroughly and separately evaluated.

Chapter 4

Performance Evaluation of Our Hybrid Approach

We introduced, in Chapter 3, a hybrid decision approach for RAT selection in heterogeneous wireless networks. Mobile users select their RAT depending on their needs and preferences, as well as on the cost and QoS parameters signaled by the network. We also presented a satisfaction-based multi-criteria decision-making method to evaluate serving RATs, and two heuristic policies to dynamically derive network information. In the present chapter, we thoroughly evaluate our RAT selection method. We consider three simulation scenarios. In the first one, we focus on network information, and investigate the effect of providing mobiles with differentiated services and throughput guarantees. The second scenario compares our satisfaction-based multi-criteria decision-making method with other existing algorithms, namely SAW and TOPSIS. In the third scenario, we illustrate the gain from using our tuning heuristics, in comparison with static network information.

4.1 System Model

We consider a heterogeneous wireless network composed of N_T generic OFDM(A)-based RATs. RAT x capacity is fixed to C^x. The radio resource is divided into N_{RU}^x resource units (*i.e.,* OFDM symbols or OFDMA slots). In the time domain, transmissions are further organized into radio frames of length T^x.

At each scheduling epoch, resource units are allocated to individual users, based on their priority and current needs (*i.e.,* amount of traffic waiting for transmission). Before any scheduling is applied, the minimum guaranteed RUs are directly granted. The Weighted Fair Queuing (WFQ) is then adopted to share out remaining resources. However, grants are limited to n_{max}. Session weights, in WFQ schedulers, are based on the cost users pay per unit of traffic. Residual resources are afterwards equitably distributed, according to the Round Robin service discipline. In fact, as long as resources are not fully committed, mobiles are allocated more than their guaranteed throughputs. Further, to avoid wasting resources, they can even have more than their maximum throughput announed by the network.

Because network information may be dynamically tuned, typically as a function of the load conditions, all mobiles do not necessarily perceive the same cost and QoS parameters at the time of selection. This affects their decision makings. In our work, we suppose that mobiles arrive sequentially. The total number of users is limited to N_{total}; it sets the traffic load. Their sojourn time is considered to be much greater in comparison with the simulation time $T_{simulation}$. Consequently, the network dynamics will progressively slow down until a pseudo-stationnary regime is attained, where all measurements are performed. To improve the statistical significance of the results, simulations are repeated 500 times, and performance metrics are averaged.

After they arrive, mobiles randomly select a user profile (cf. Table 4.1). As a matter of fact, they initiate either an inelastic, a streaming, or an elastic session, and determine their cost tolerance parameter λ and the weights $w_{d_{min}}$, $w_{d_{max}}$, and w_{cost} they apply to normalized decision criteria. In Table 4.1, the weights of the decision criteria are normalized such that they sum up to 1 for each user profile. Further, mobiles decode current cost and QoS information, evaluate their expected satisfaction levels, and rank the different alternatives. The needs of inelastic and streaming sessions are respectively expressed as fixed (*i.e.,* R_f), and average long-term throughput (*i.e.,* R_{av}). We assume that the set of possible throughput demands is given by $D = \{0.5, 1, 1.5, 2\}$ Mb/s.

Profile No.	**Traffic class**	λ	$w_{d_{min}}$	$w_{d_{max}}$	w_{cost}
1	Inelastic	60	0.7	0	0.3
2	Streaming	60	14/30	7/30	0.3
3	Elastic	60	0	0.7	0.3
4	Inelastic	25	0.3	0	0.7
5	Streaming	25	0.2	0.1	0.7
6	Elastic	25	0	0.3	0.7

Table 4.1: User profiles

Inelastic and streaming traffic is packetized into small units of fixed length $L^c, c \in \{I, S\}$. Inelastic sessions generate packets according to a deterministic distribution, whereas streaming sessions generate packets according to a Poisson process. These packets are segmented into blocks sized to fit one RU. In our work, we fix delay constraints for the latter session types. A maximum delay requirement of $\Delta^c, c \in \{I, S\}$ is fixed. Since resources are limited, some packets may miss their deadline. They will be dropped as they are no longer useful.

Furthermore, the needs of elastic sessions are expressed as comfort throughput (*i.e.,* R_c). We suppose that the set of possible comfort throughputs is given by $C = \{0.75, 1.25\}$ Mb/s. Inelastic and streaming sessions uniformly choose one of the possible throughput demands, regardless of the user cost tolerance parameter. Yet, we assume in the following that the comfort throughput of elastic sessions is related to the user willingness to pay, and thus imposed by the user profile.

To provide a detailed performance evaluation, three simulation scenarios are considered. In the first one, we focus on network information, and assess the effect of providing mobiles with differentiated service classes and throughput guarantees. The second scenario compares our satisfaction-based multi-criteria decision-making method with other existing algorithms, namely SAW and TOPSIS. In the third scenario, we illustrate the gain from using our tuning policies in comparison with static network information.

4.1.1 Scenario 1: QoS Information

In this first scenario, we are interested in the performance improvement achieved by providing differentiated service classes, and minimum throughput guarantees to mobile users, regardless of future network load conditions.

We consider a realistic and cost-effective deployment, where N_T RATs are co-localized: the same base station site is used, leading to cell overlapping. For the sake of simplicity, all users are assumed to belong to the same zone Z_k: they all have the same modulation and coding schemes, and thus exploit in the same manner their allocated grants. General simulation parameters are listed in Table 4.2.

Parameters	**Values**
N_T	3
$C^x, x = 1, ..., N_T$	35 Mb/s
$N_{RU}^x, x = 1, ..., N_T$	700
$T^x, x = 1, ..., N_T$	10 ms
$T_{simulation}$	300 s
$L^c, c = I, S$	125 bytes
$\Delta^c, c = I, S$	100 ms

Table 4.2: Simulation parameters for the first and second scenarios

To evaluate long-term network performance, five major key performance indicators are defined: throughput, mean waiting delay and packet drop probability (for inelastic and streaming sessions), user satisfaction, and operator gain. In our work, the waiting delay represents the time that a packet spends in the queue before being transmitted, and the packet drop probability represents the probability that a packet is rejected due to exceeding its deadline.

4.1.1.1 Service differentiation

To examine the impact of service differentiation on global network performance, we compare the following two situations:

- *Situation 1: Differentiated services network.* Radio access technologies provide differentiated service classes, namely, Premium, Regular and Economy. They differ in their QoS and cost parameters. A QoS-aware pricing scheme should then be adopted: mobiles are charged based on their priority. Otherwise, all sessions would select the premium service class, and our differentiated services model would lose its interest.

- *Situation 2: Mono-service network.* Radio access technologies provide a unique service class, namely Regular plus.

QoS parameters as perceived by mobile users, namely their d_{min} and d_{max}, and cost parameters are depicted in Table 4.3. They are assumed fixed, and do not change as the RAT load changes, except when the RAT is no longer able to provide arriving mobiles with the initial QoS parameters.

Service class	d_{min} **(Mb/s)**	d_{max} **(Mb/s)**	*Cost* **(unit/kB)**
Premium	1.5	2	6
Regular	1	1.5	4
Economy	0.5	1	2
Regular Plus	1	2	4

Table 4.3: Scenario 1: Static QoS and cost parameters

As inelastic sessions are inflexible in their requirements, selection decisions need to meet their fixed throughput demands. When the RAT is highly loaded, the resource scheduler is no more able to provide them with more than their minimum guaranteed throughputs, eventually leading to performance degradation. So as to enhance their QoS level, typically at high traffic load, mobiles should be provided with high enough throughput guarantees, or equivalently with high enough priority. Regardless of the user profile, selection decisions, when differentiated services are provided, are reported in Table 4.4.

Throughput Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**
Premium			✓	✓
Regular		✓		
Economy	✓			

Table 4.4: Satisfaction-based decisions for inelastic sessions

Fig. 4.1 and 4.2 respectively show the mean waiting delay and the packet drop probability, as a function of the total number of arrivals. When differentiated services are provided, throughput-intensive sessions select the Premium service class with the highest priority, leading to a shorter delay, a lower drop probability and subsequently a better QoS level.

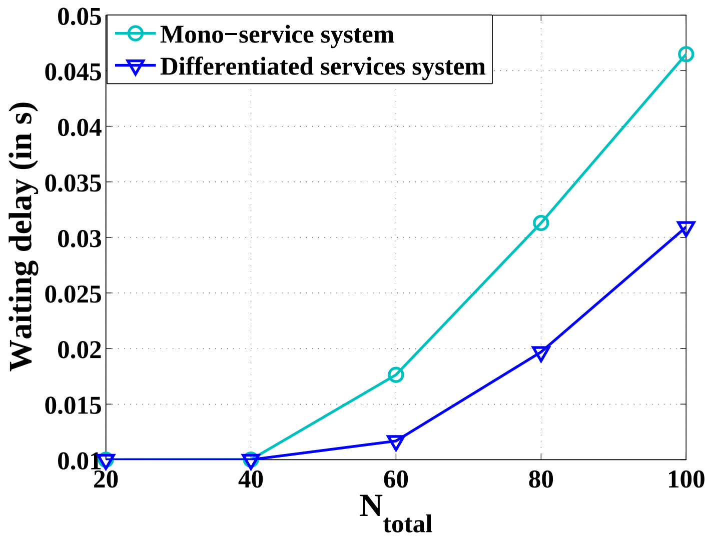

Figure 4.1: Scenario 1: Mean waiting delay for inelastic sessions

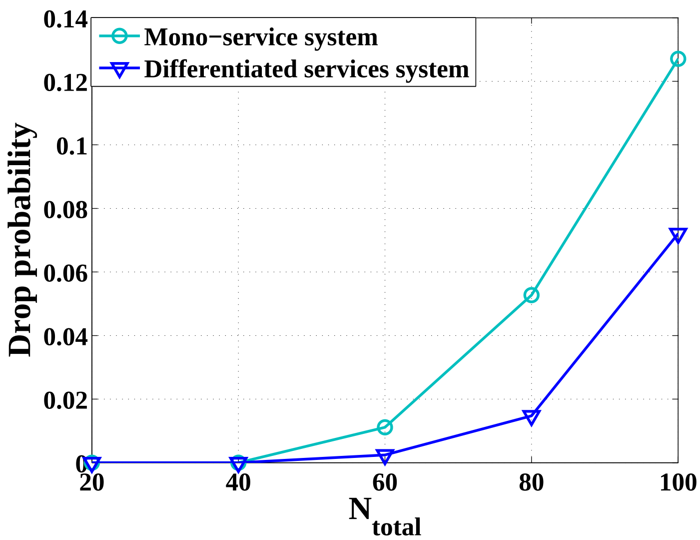

Figure 4.2: Scenario 1: Packet drop probability for inelastic sessions

We depict in Fig. 4.3 the average user satisfaction. We notice that, at low traffic load, user satisfaction is higher when a unique service class is provided. The Regular plus service class fulfills strict QoS requirements, while charging mobiles on average with lower cost. Yet, when the network gets loaded, throughput-intensive sessions see their performance degraded. The Regular plus service class is no more able to meet their inflexible throughput demands, thus strongly decreasing the average user satisfaction. However, when differentiated services are provided, throughput-intensive sessions always opt for the Premium service class, and then enjoy higher throughput guarantees. This leads to a larger overall satisfaction, as illustrated in Fig. 4.3.

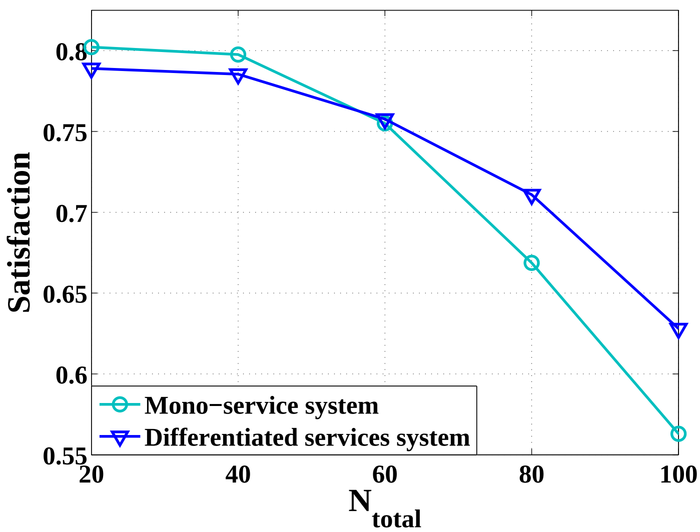

Figure 4.3: Scenario 1: User satisfaction for inelastic sessions

Furthermore, since streaming sessions are fairly flexible, mobiles may be less restrictive in their choices. Based on their preferences, users may actually look for fair enough content quality (average long-term throughput), high content quality (higher throughput), or even poor content quality (lower throughput). Selection decisions are put forward in Tables 4.5 and 4.6.

Throughput Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**
Premium		✓	✓	✓
Regular	✓			
Economy				

Table 4.5: Satisfaction-based decisions for streaming sessions: users are ready to pay for better performance

Throughput Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**
Premium				✓
Regular			✓	
Economy	✓	✓		

Table 4.6: Satisfaction-based decisions for streaming sessions: users seek to save up money

The mean waiting delay and the packet drop probability are respectively illustrated in Fig. 4.4 and 4.5. When differentiated services are provided, better performances are observed mainly at medium traffic load: demanding sessions could be provided with higher throughput guarantees (*i.e.*, with the Premium service class), and even low-priority sessions may be granted more than their minimum guaranteed throughputs.

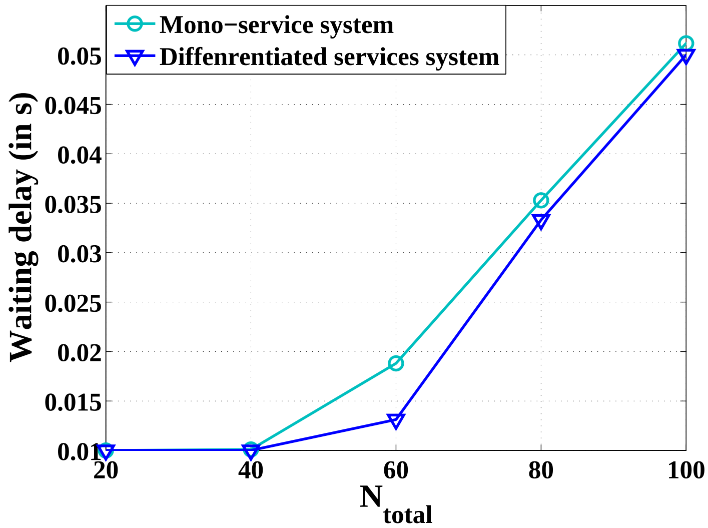

Figure 4.4: Scenario 1: Mean waiting delay for streaming sessions

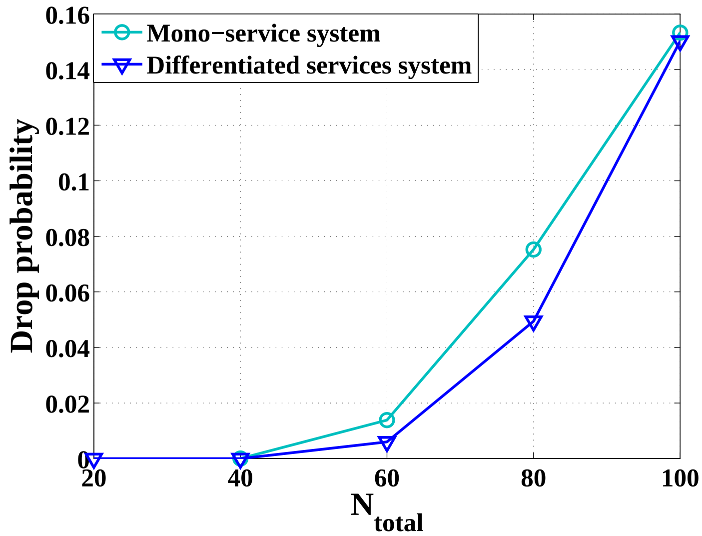

Figure 4.5: Scenario 1: Packet drop probability for streaming sessions

However, when the network gets loaded, mobiles that seek to save up money, and thus have on average lower throughput guarantees, suffer from poor performance. Further, mobiles that are ready to pay are always provided with high enough throughput guarantees, and consequently have better QoS than when a unique service class is offered. Therefore, at high traffic load, performances are on average very close. Streaming sessions that are ready to pay offset the performance degradation of those that seek to save up money.

Besides, user satisfaction is constantly higher when differentiated services are provided (Fig. 4.6). In contrast to inelastic sessions, users that seek to save up money sacrifice within limits their service quality (*i.e.*, select a cheaper service class), thus leading to a higher overall satisfaction, typically at low traffic load.

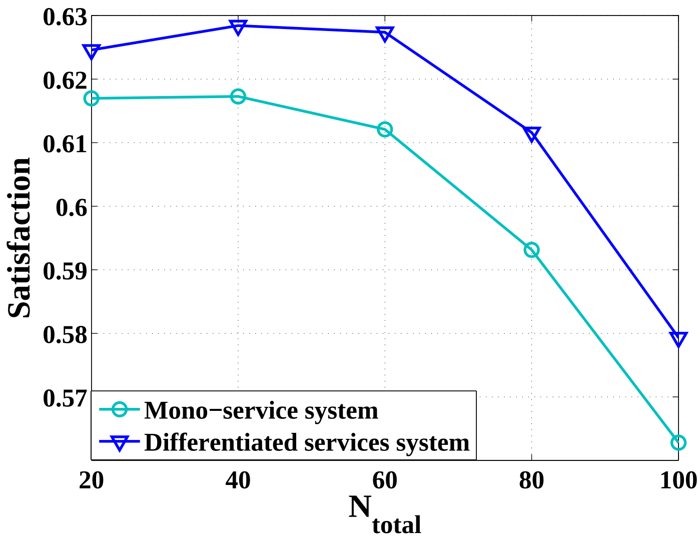

Figure 4.6: Scenario 1: User satisfaction for streaming sessions

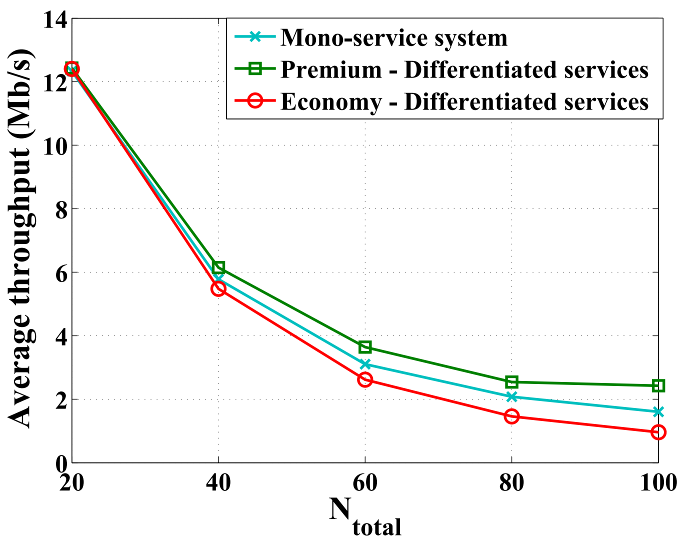

Figure 4.7: Scenario 1: Average throughput for elastic sessions

Because elastic sessions have no QoS needs, selection decisions exclusively depend on user preferences. Mobiles, that are ready to pay, select the Premium service class, and enjoy the highest throughput. However, those who seek to save up money select the Economy class, and have the lowest throughput. Furthermore, when a unique service class is provided, all sessions have similar priorities, leading to similar throughputs, as shown in Fig. 4.7.

As they are associated with the service class that best meet their preferences, elastic sessions have significantly higher satisfaction (Fig. 4.8), when differentiated services are provided.

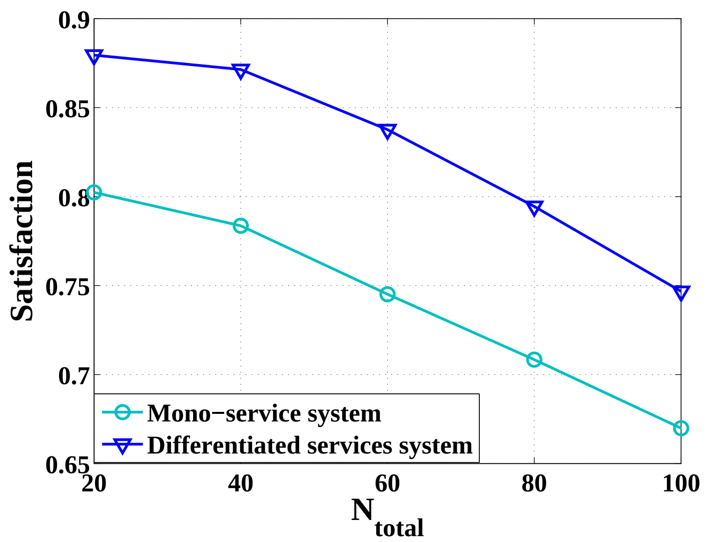

Figure 4.8: Scenario 1: User satisfaction for elastic sessions

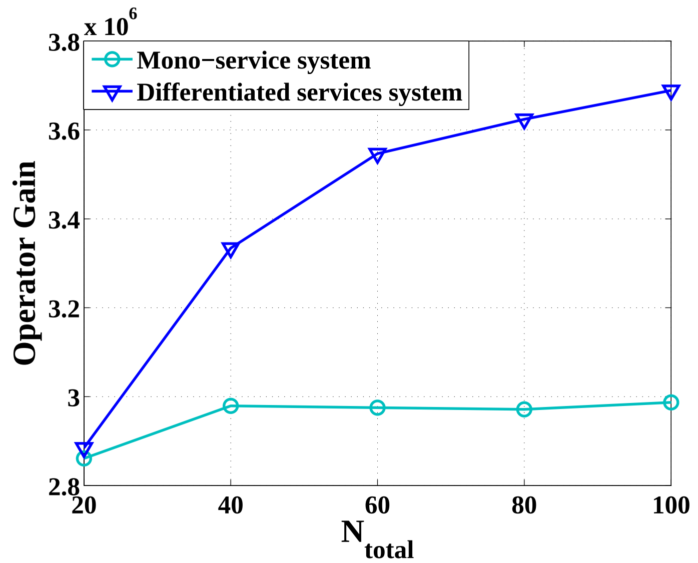

Figure 4.9: Scenario 1: Operator gain

Moreover, when differentiated services are proposed, the operator gain is maximized, as depicted in Fig. 4.9. Also, although mobiles pay on average more, they have a significantly higher satisfaction (Fig. 4.10). As a matter of fact, when differentiated service classes are provided, mobiles avoid undersized and oversized decisions, and select the service class that best meets user needs and preferences.

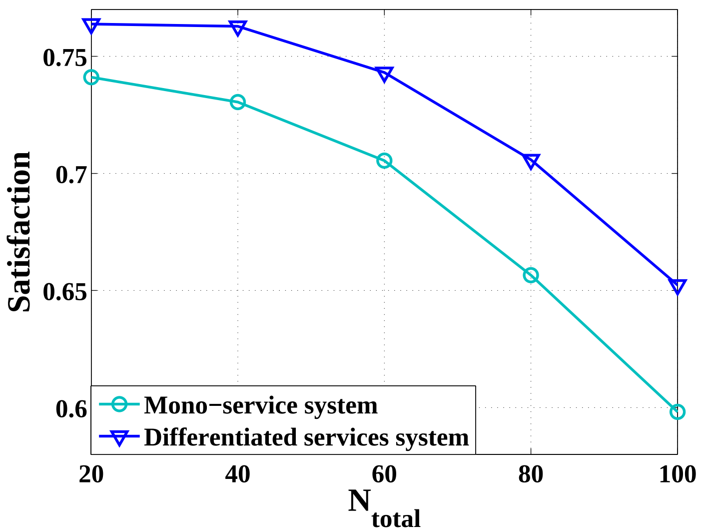

Figure 4.10: Scenario 1: User satisfaction

4.1.1.2 Throughput guarantees

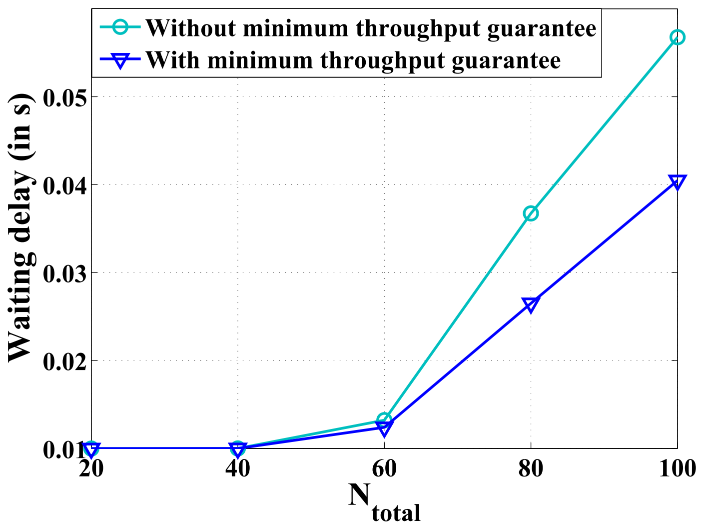

Figure 4.11: Scenario 1: Mean waiting delay for real-time sessions

We also discuss the impact of throughput guarantees on the performance of real-time sessions. When real-time sessions (*i.e.,* inelastic and streaming sessions) are provided with minimum throughput guarantees (*i.e.,* $d_{min} \neq 0$) regardless of future load conditions, they have a shorter delay (Fig. 4.11), a lower drop probability (Fig. 4.12), and thus a better QoS level. As real-time sessions are always provided with, at least, their minimum guaranteed RUs, their performances are enhanced, particularly when RATs get loaded.

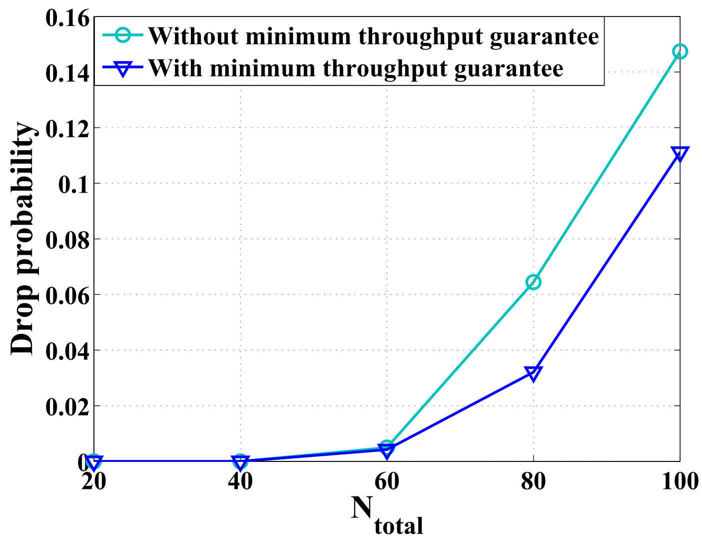

Figure 4.12: Scenario 1: Packet drop probability for real-time sessions

4.1.2 Scenario 2: Multi-Criteria Decision-Making Methods

In this second scenario, we compare our Satisfaction-Based (SB) multi-criteria decision-making method with the well-known SAW [SNW06] and TOPSIS [FC11, CM12] algorithms. As in the first scenario, hybrid cells include N_T co-localized RATs. Since we mainly focus on the decision makings, and for the sake of simplicity, all mobiles are supposed to belong to the same zone Z_k. Thus, they are assumed to have the same peak rate. General simulation parameters are depicted in Table 4.2.

Each RAT proposes three different service classes, namely Premium, Regular and Economy. QoS and cost parameters, as perceived by mobile users, are depicted in Table 4.7. Once again, they are supposed fixed and do not change as the RAT load changes, except when the RAT is no longer able to provide future arrivals with the initial QoS parameters.

Before we discuss simulation results, let us recall the SAW and TOPSIS methods. When normalizing decision criteria $d_{min}(a)$, $d_{max}(a)$, and $cost(a)$, SAW and TOPSIS ignore user needs (*i.e.,* traffic class, throughput demand, cost tolerance), and exclusively depend on available alternatives. We note $\mathcal{A}$ the set of available alternatives and $\tilde{a}$ any element that

Service class	d_{min} **(Mb/s)**	d_{max} **(Mb/s)**	$cost$ **(unit/kB)**
Premium	1.5	2	6
Regular	1	1.5	4
Economy	0.5	1	2

Table 4.7: Scenario 2: Static QoS and cost parameters

belongs to $\mathcal{A}$.

4.1.2.1 Simple Additive Weighting (SAW)

For alternative a, the normalizing functions regardless of the session traffic class c are:

$$\hat{d}'(a) = \frac{d'(a) \cdot g(M, C)}{\max_{\tilde{a} \in A} d'(\tilde{a}) \cdot g(M, C)} \tag{4.1.1}$$

where $d' = \{d_{min}, d_{max}\}$, and

$$\widehat{cost}(a) = \frac{\min_{\tilde{a} \in A} cost(\tilde{a})}{cost(a)} \tag{4.1.2}$$

The utility function of a class c session for alternative a is defined by :

$$U^c(a) = w^c_{d_{min}} \cdot \hat{d}_{min}(a) + w^c_{d_{max}} \cdot \hat{d}_{max}(a) + w^c_{cost} \cdot \widehat{cost}(a)$$

Mobiles actually select the alternative with the highest score (*i.e.,* utility function).

4.1.2.2 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

For alternative a, the normalizing functions regardless of the session traffic class c are:

$$\hat{d}'(a) = \frac{d'(a) \cdot g(M, C)}{\sqrt{\sum_{\tilde{a} \in A} (d'(\tilde{a}) \cdot g(M, C))^2}} \tag{4.1.3}$$

where $d' = \{d_{min}, d_{max}\}$, and

$$\widehat{cost}(a) = \frac{cost(a)}{\sqrt{\sum_{\tilde{a} \in A} (cost(\tilde{a}))^2}} \tag{4.1.4}$$

The positive and the negative ideal solutions, respectively denoted by a^+ and a^-, are then determined as follows:

$$a^+ = (d^+_{min}, d^+_{max}, cost^+) = (\max_{\tilde{a} \in A} \hat{d}_{min}(\tilde{a}), \max_{\tilde{a} \in A} \hat{d}_{max}(\tilde{a}), \min_{\tilde{a} \in A} \widehat{cost}(\tilde{a})) \tag{4.1.5}$$

$$a^- = (d^-_{min}, d^-_{max}, cost^-) = (\min_{\tilde{a}\in A} \hat{d}_{min}(\tilde{a}), \min_{\tilde{a}\in A} \hat{d}_{max}(\tilde{a}), \max_{\tilde{a}\in A} \widehat{cost}(\tilde{a})) \tag{4.1.6}$$

These ideal solutions do not necessarily exist: a^+ and a^- are defined as virtual alternatives with respectively the best and the worst decision criteria values.

The distance of alternative a from the positive ideal and the negative ideal solution, respectively denoted by $S^+(a)$ $S^-(a)$, are furthermore computed as:

$$S^+(a) = \sqrt{[w^c_{d_{min}}(\hat{d}_{min}(a) - d^+_{min})]^2 + [w^c_{d_{max}}(\hat{d}_{max}(a) - d^+_{max})]^2 + [w^c_{cost}(\widehat{cost}(a) - cost^+)]^2} \tag{4.1.7}$$

$$S^-(a) = \sqrt{[w^c_{d_{min}}(\hat{d}_{min}(a) - d^-_{min})]^2 + [w^c_{d_{max}}(\hat{d}_{max}(a) - d^-_{max})]^2 + [w^c_{cost}(\widehat{cost}(a) - cost^-)]^2} \tag{4.1.8}$$

The relative closeness (*i.e.,* utility function) is however defined as:

$$C(a) = \frac{S^-(a)}{S^-(a) + S^+(a)} \tag{4.1.9}$$

Mobiles select the alternative with the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution, or equivalently the alternative with the highest relative closeness.

Because they ignore user needs, SAW and TOPSIS often lead to undersized and oversized decisions. When selections are independent of session throughput demands, users with a demand of 2 Mb/s make the exactly same decisions as those with a demand of 0.5 Mb/s. As a matter of fact, their decisions exclusively depend on user preferences (*i.e.,* weights of the decision criteria), as well as on the characteristics of available alternatives. On the one hand, when users seek to save up money, they always opt for the Economy service class (*i.e.,* their best trade-off between QoS and cost parameters). As a consequence, the performance of throughput-intensive sessions are dramatically degraded. On the other hand, when they are ready to pay for better performance, they always select the Premium service class. Consequently, sessions with relatively low throughput demand will uselessly pay more: premium guarantees may not improve their performance in comparison with regular or economy ones.

Yet, our proposed Satisfaction-Based (SB) algorithm provides the best performance for the best cost. On the one hand, when session needs are stringent and inflexible, a high enough priority service class is selected, thus enhancing user performance. On the other hand, when higher throughput guarantees do not improve session performance, SB leads to a low enough priority service class, thus charging mobile users with lower cost. So as to make the comparison more fair, *enhanced SAW and TOPSIS* are used: they only explore feasible alternatives. When user throughput demand is greater than the provided d_{max}, the alternative opted for is considered to be infeasible, and thus rejected. This will prevent SAW and TOPSIS from making some undersized decisions. However, as discussed in the following paragraph, our proposed method continues to outperform them.

4.1.2.3 Comparison results

So as to enhance network performance, and as stated above, enhanced SAW and TOPSIS only explore feasible alternatives. Yet, they continue to lead to some undersized, but mostly oversized alternatives. For inelastic sessions, selection decisions, according to the different multi-criteria decision-making methods, are reported in Tables 4.8 and 4.9.

Decision Method	**SAW/TOPSIS**				**SB**			
Session Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**	**0.5**	**1**	**1.5**	**2**
Premium	✓	✓	✓	✓			✓	✓
Regular						✓		
Economy					✓			

Table 4.8: Decisions for inelastic sessions: users are ready to pay for better performance

Decision Method	**SAW/TOPSIS**				**SB**			
Session Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**	**0.5**	**1**	**1.5**	**2**
Premium				✓			✓	✓
Regular			✓			✓		
Economy	✓	✓			✓			

Table 4.9: Decisions for inelastic sessions: users seek to save up money

When users are ready to pay for better performance, SAW and TOPSIS always single out the Premium service class. Intuitively, and since inelastic session needs are fixed, this decision is oversized for 0.5 and 1 Mb/s sessions. As SB respectively opts for the Economy and the Regular service classes, QoS requirements are always perfectly satisfied, while charging mobile users with lower cost.

Also, when users seek to save up money, enhanced SAW and TOPSIS lead to the Economy service class for 1 Mb/s sessions, and to the Regular service class for 1.5 Mb/s sessions. These decisions are undersized. When the RAT is highly loaded, fixed QoS requirements are not satisfied, thus dramatically degrading session performances.

Fig. 4.13 and 4.14 respectively show the mean waiting delay and the packet drop probability, as a function of the total number of arrivals. Since it avoids undersized decisions, SB provides a shorter delay, a lower drop probability, and subsequently a better overall QoS level.

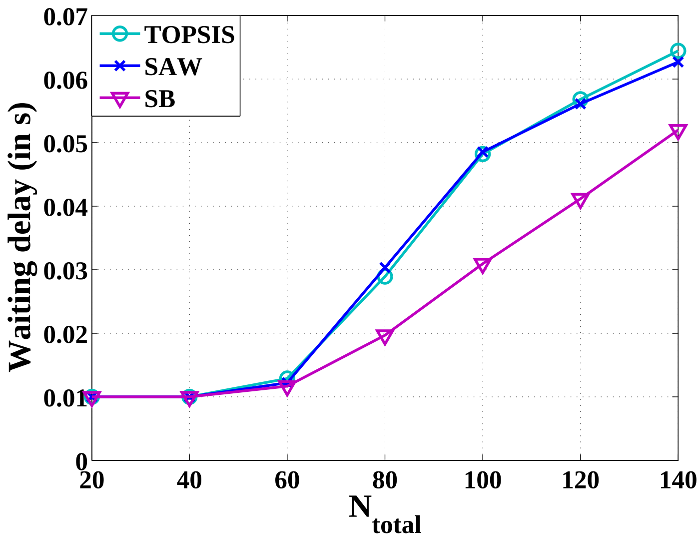

Figure 4.13: Scenario 2: Mean waiting delay for inelastic sessions

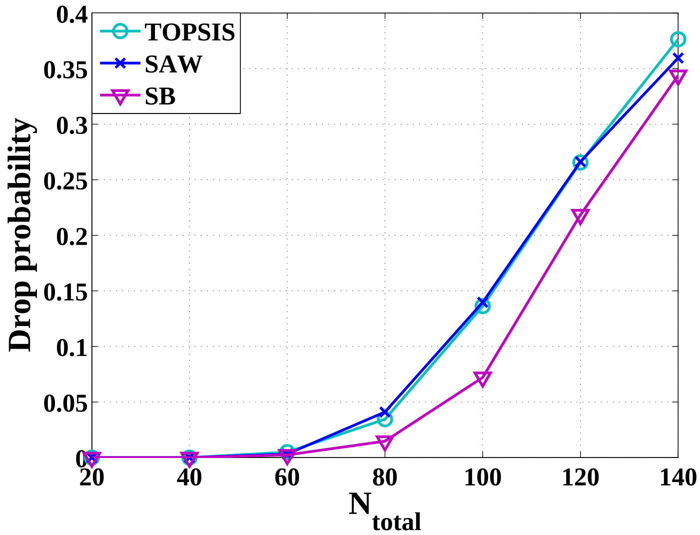

Figure 4.14: Scenario 2: Packet drop probability for inelastic sessions

We depict in Fig. 4.15 the average user satisfaction. We notice that, at low traffic load, enhanced SAW and TOPSIS provide higher satisfaction. First, undersized decisions are able to fulfill strict QoS requirements, while charging mobile users with lower cost. Second, although oversized decisions decrease user satisfaction, the reduction is not significant enough to offset the impact of undersized decisions. In other words, at low traffic load, undersized decisions considerably increase user satisfaction, because the corresponding users seek to save up money. Their QoS needs are perfectly met, while paying less. However, oversized decisions do not significantly decrease user satisfaction, because users in question are originally ready to pay. We further note that, when traffic load is moderate, SB brings the largest satisfaction, since it always meets the strict QoS requirements. As a matter of fact, using SAW and TOPSIS, undersized decisions are no more able to meet user needs, when traffic load is relatively high.

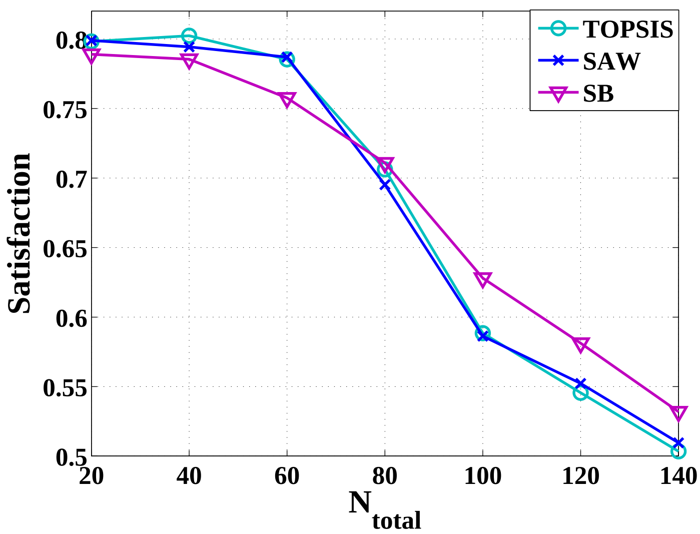

Figure 4.15: Scenario 2: User satisfaction for inelastic sessions

For streaming sessions, selection decisions are put forward in Tables 4.10 and 4.11.

Decision Method	**SAW/TOPSIS**				**SB**			
Session Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**	**0.5**	**1**	**1.5**	**2**
Premium	✓	✓	✓	✓		✓	✓	✓
Regular					✓			
Economy								

Table 4.10: Decisions for streaming sessions: users are ready to pay for better performance

Decision Method	**SAW/TOPSIS/SB**			
Session Needs (Mb/s)	**0.5**	**1**	**1.5**	**2**
Premium				✓
Regular			✓	
Economy	✓	✓		

Table 4.11: Decisions for streaming sessions: users seek to save up money

When users are ready to pay for better performance, for 0.5 Mb/s sessions, SAW and

TOPSIS lead to the Premium service class, and SB to the Regular one. SAW and TOPSIS decisions are oversized. The Regular service class actually provides users with twice their average long-term throughput.

The mean waiting delay and the packet drop probability are respectively depicted in Fig. 4.16 and 4.17. Since all methods provide the same QoS level, the Premium service class proves to be oversized for 0.5 Mb/s sessions. In comparison with SB, no performance improvement is observed. Therefore, on average, SB charges less and carries out higher user satisfaction (Fig. 4.18).

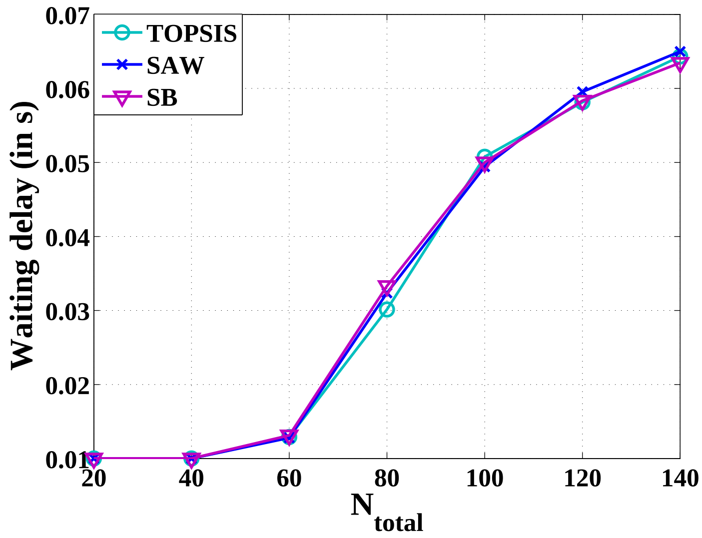

Figure 4.16: Scenario 2: Mean waiting delay for streaming sessions

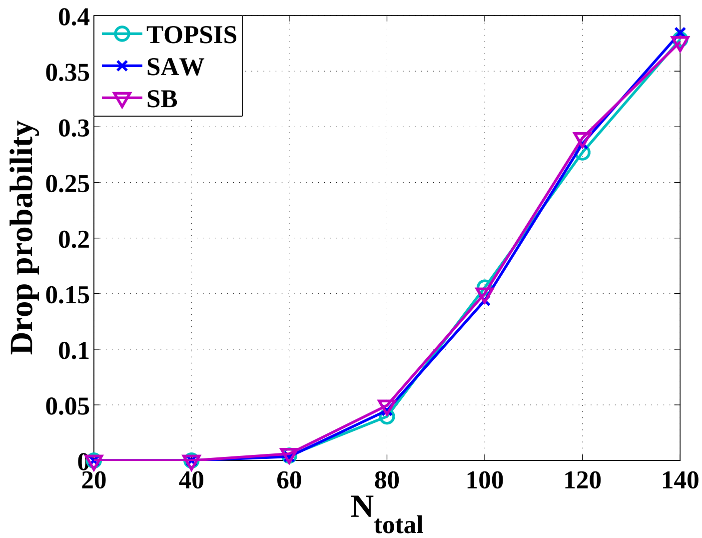

Figure 4.17: Scenario 2: Packet drop probability for streaming sessions

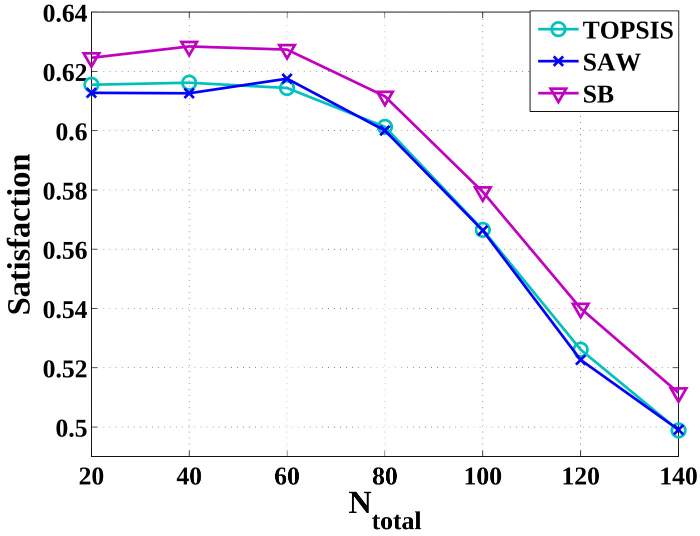

Figure 4.18: Scenario 2: User satisfaction for streaming sessions

Because elastic sessions accomodate with available bandwidth, undersized and oversized decisions do not technically exist. When SB takes into account user comfort throughput, it may theoretically reach different solutions from SAW and TOPSIS. Yet, given our simulation model and parameters, they practically all lead to the same decisions, providing the same user satisfaction (cf. Fig. 4.19).

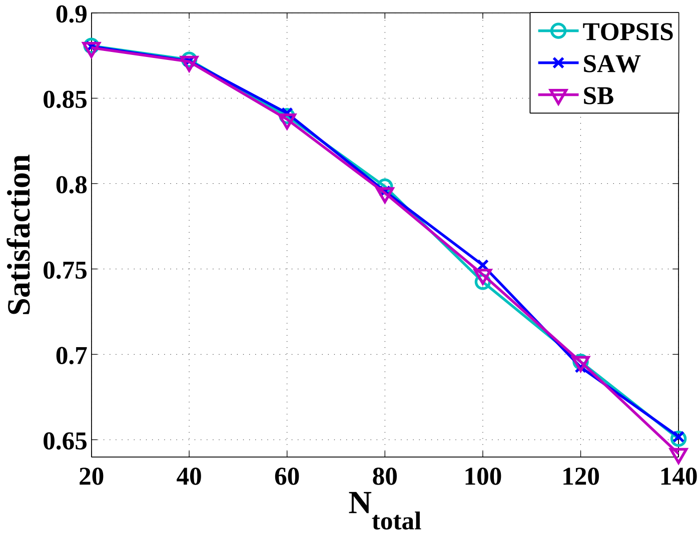

Figure 4.19: Scenario 2: User satisfaction for elastic sessions

When users are ready to pay for better performance, they systematically select the Premium service class. Nevertheless, when they seek to save up money, they choose the Economy one. As illustrated in Fig. 4.20, Premium sessions enjoy higher throughputs than Economy ones.

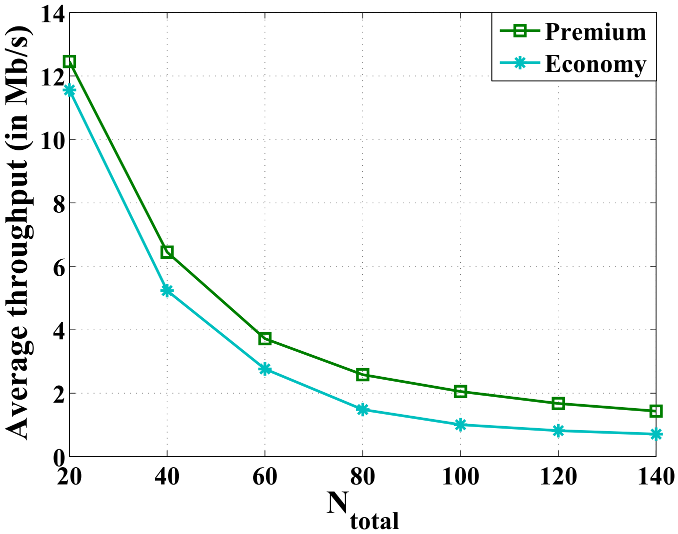

Figure 4.20: Scenario 2: Average throughput for elastic sessions

The comfort metric is defined as the ratio of the perceived throughput to the comfort throughput. Although Premium sessions have higher throughputs, their comfort metric is similar to the Economy ones except at low traffic load (cf. Fig. 4.21). Thereby, our solution ensures fairness with respect to different comfort throughputs.

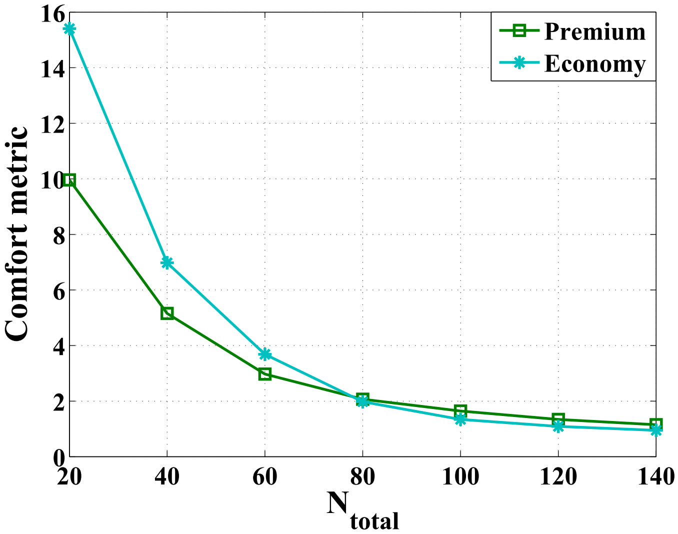

Figure 4.21: Scenario 2: Mean comfort metric for elastic sessions

Furthermore, when a RAT is no longer able to guarantee to future arrivals the initial QoS parameters, network information is modified. As they have lower throughput guarantees for the same initial monetary cost, new arrivals are considered to be disadvantaged. We depict in Fig. 4.22 the Disadvantaged Sessions Rate, denoted by DSR, and defined as the number of disadvantaged sessions over the total number of on-going sessions. Since it avoids oversized decisions, SB brings the lowest DSR. At high traffic load, higher QoS guarantees are provided respectively with SB, SAW and TOPSIS.

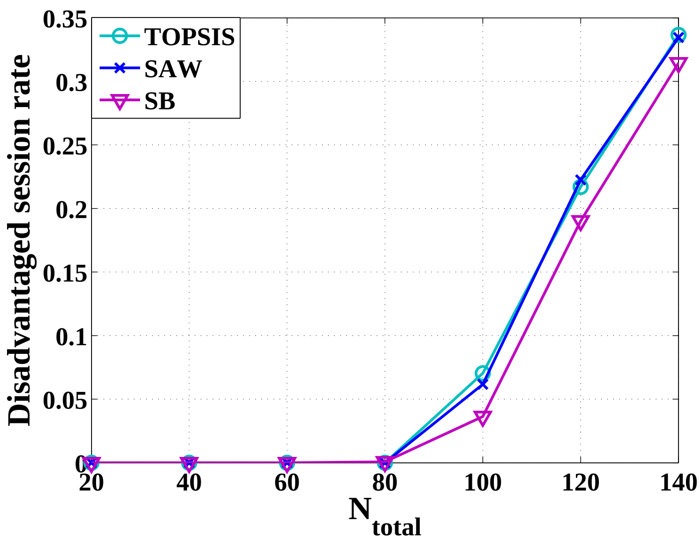

Figure 4.22: Scenario 2: Disadvantaged session rate

To wrap up, SB avoids undersized decisions, best meets QoS requirements and brings the best performance. By eliminating infeasible alternatives, enhanced SAW and TOPSIS bring similar performance as SB, for streaming and elastic sessions. However, SB considerably outperforms them for inelastic sessions, where QoS requirements are stringent and inflexible.

Also, by evading oversized decisions typically for inelastic and streaming sessions, SB charges on average less than enhanced SAW and TOPSIS. Thereby, SB leads to better performance, lower cost and therefore higher user satisfaction.

4.1.3 Scenario 3: Tuning Policies

In this third scenario, we illustrate the gain from using our tuning policies in comparison with static network information. When a RAT dominates all the others (*i.e.,* provides higher QoS parameters for the same cost or the same QoS parameters for a lower cost), QoS information are either modulated as a function of the load conditions using the staircase or the slope tuning policies, or maintained fixed leading to performance inefficiency. Recall that, prior to the RAT selection process, a common admission control is assumed to be performed. General simulation parameters are however listed in Table 4.12.

Parameters	**Values**
N_T	2
$C^x, x = 1, ..., N_T$	70 Mb/s
$N^x_{RU}, x = 1, ..., N_T$	700
$T^x, x = 1, ..., N_T$	10 ms
$T_{simulation}$	300 s
$L^c, c = I, S$	125 byte
$\Delta^c, c = I, S$	100 ms

Table 4.12: Simulation parameters for the third scenario

Each RAT is assumed to propose three different service classes, namely Premium, Regular and Economy. All RATs are supposed to initially signal the same QoS and cost parameters listed in Table 4.13.

Service class	d_{min} **(Mb/s)**	d_{max} **(Mb/s)**	$Cost$ **(unit/kB)**
Premium	1	1.35	6
Regular	0.7	1	4
Economy	0.35	0.7	2

Table 4.13: Scenario 3: Initial QoS and cost parameters

We further assume that mobiles randomly select a set of modulation and coding gains. These multiplicative factors reflect the user radio conditions in the different technologies, and are supposed to remain constant in time. Two sets of gains are considered and reported in Table 4.14. They typically illustrate the network topology of Fig. 4.23.

Set No.	**RAT 1**	**RAT 2**
1	1.5	1.5
2	2	1

Table 4.14: Modulation and coding gains

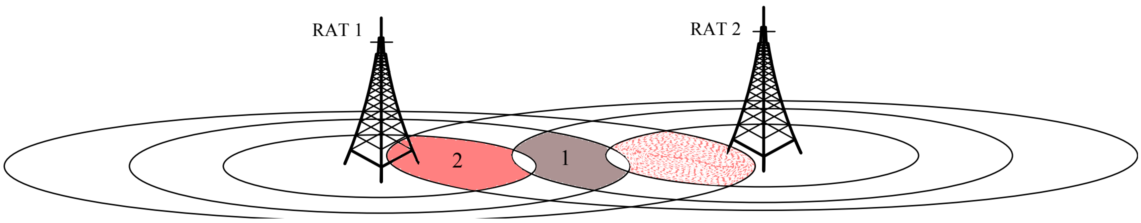

Figure 4.23: Scenario 3: A possible network topology

When the two RATs provide the same QoS parameters, users that are associated with set no. 2 would select RAT 1. They expect to have better radio conditions, and thus to perceive higher throughputs in RAT 1. All other alternatives, proposed by RAT 2, are subsequently dominated. Also, users that are associated with set no. 1 randomly join their RAT, since they expect to perceive similar throughputs in the two available RATs. This situation leads to unevenly distributed traffic load. However, when network information is dynamically modulated, according to the staircase or to the slope tuning policies, QoS parameters are tuned in a way to globally drive future arrivals to the less loaded RAT: loaded RATs provide lower QoS parameters, thus pushing future users to less loaded RATs. When staircase policy is adopted, reduced QoS parameters are presented in Table 4.15.

Service class	d_{min} **(Mb/s)**	d_{max} **(Mb/s)**
Premium	0.5	0.7
Regular	0.35	0.5
Economy	0.2	0.5

Table 4.15: Reduced QoS parameters for the staircase tuning policy

Other scenarios may also lead to unevenly distributed traffic load. For instance, when mobiles have the same modulation and coding schemes, a RAT is preferred if it initially broadcasts higher QoS parameters for the same cost, or the same QoS parameters for a lower cost. While static information absolutely leads to performance inefficiency, dynamic tuning helps to better distribute mobile users over the available RATs, and thus to efficiently utilize radio resources.

When using the staircase or the slope tuning policies, we assume that S_1 and S_2 are respectively set to 0.5 and 0.9 times the RAT capacity. Before S_1, the network provides constant QoS parameters. After S_2, QoS incentives are no longer provided to future arrivals: the network keeps a margin of about 10% of the RAT capacity to provide ongoing sessions with more than their minimum guaranteed throughputs. These parameters will be thoroughly studied in Chapter 5.

Results have shown the same trend for different simulation scenarios and parameters. Typically, we came to exactly the same conclusions with different modulation and coding gains, initial network information, network model parameters, tuning thresholds (*i.e.,* S_1 and S_2), and also when a unique service class is provided.

Because real-time (RT) sessions (*i.e.,* inelastic and streaming sessions) require tight delay constraints, access technologies should meet their throughput demands. However, users with a demand of 2 Mb/s may suffer: even the Premium guarantees may be lower than their throughput demand. When the RAT is highly loaded, the resource scheduler will not be able to provide them with more than their minimum guaranteed throughputs, thus leading to packet loss. So as to reduce the packet drop probability, we should avoid that a RAT gets overloaded long before the others. Load balancing should then be achieved.

Fig. 4.24 and 4.25 respectively show the mean waiting delay and the packet drop probability, as a function of the total number of arrivals.

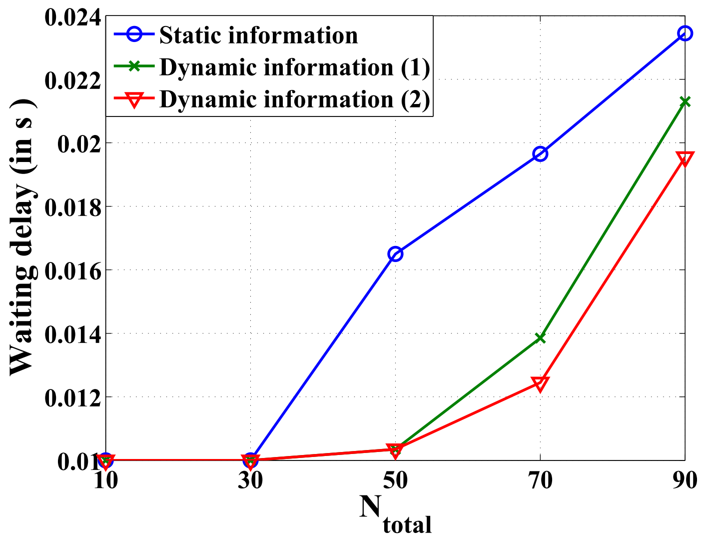

Figure 4.24: Scenario 3: Mean waiting delay for real-time sessions

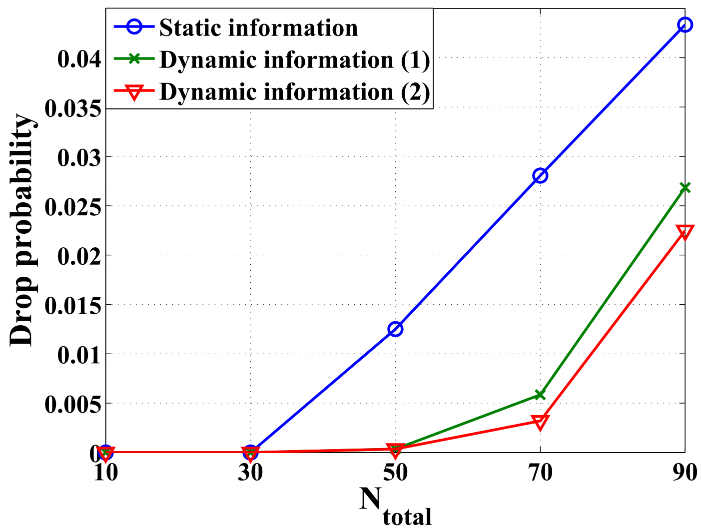

Figure 4.25: Scenario 3: Packet drop probability for real-time sessions

When the slope tuning policy denoted as Dynamic information (2) is adopted, it best responds to traffic load fluctuations, and thus provides a shorter delay, a lower drop probability and subsequently a better overall QoS level. Besides, the staircase tuning policy denoted as Dynamic information (1) is disadvantageous when all RATs have exceeded their S_1: while load conditions are critical, RAT 1 is once again privileged until the operator guarantees exceed S_2 (*i.e.,* until RAT 1 no longer provides QoS guarantees to future arrivals). Yet, the performance of real-time sessions are always significantly enhanced in comparison with the static scenario, denoted as Static information.

Moreoever, when sessions are better distributed over the two RATs, they will be allocated on average more RUs. Typically, when QoS parameters are tuned as a function of the load conditions, elastic sessions experience higher throughput and subsequently higher comfort metric, as shown in Fig. 4.26. However, at low traffic load (since tuning policies are not yet triggered) and at high traffic load (since all RATs become similarly occupied regardless of the tuning policy), performance enhancement is not that significant for elastic sessions.

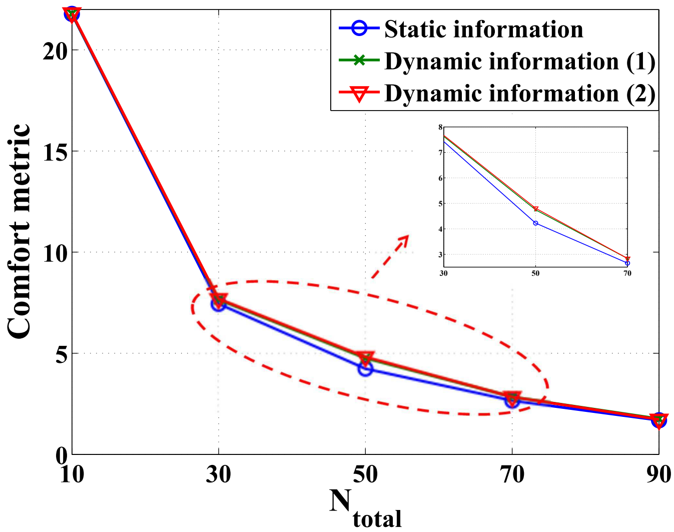

Figure 4.26: Scenario 3: Mean comfort metric for elastic sessions

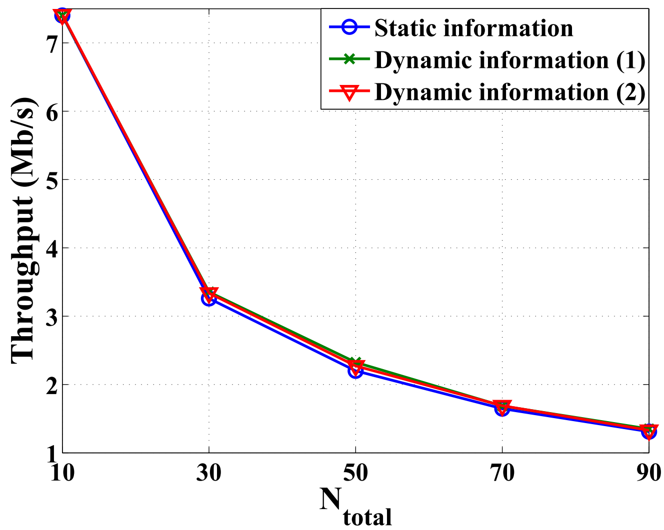

Figure 4.27: Scenario 3: Average throughput

Furthermore, when tuning policies are triggered, QoS parameters are reduced. To benefit from the same initial throughput guarantees, mobile users may have to select a higher priority service class, and thus pay more. Also, because fewer real-time packets are dropped (cf. Fig. 4.25) and more elastic packets are served (cf. Fig. 4.26), users consume on average a larger amount of traffic (Fig. 4.27), and once again pay more. We illustrate in Fig. 4.28 the average operator gain. When operators dynamically intervene, they gain more.

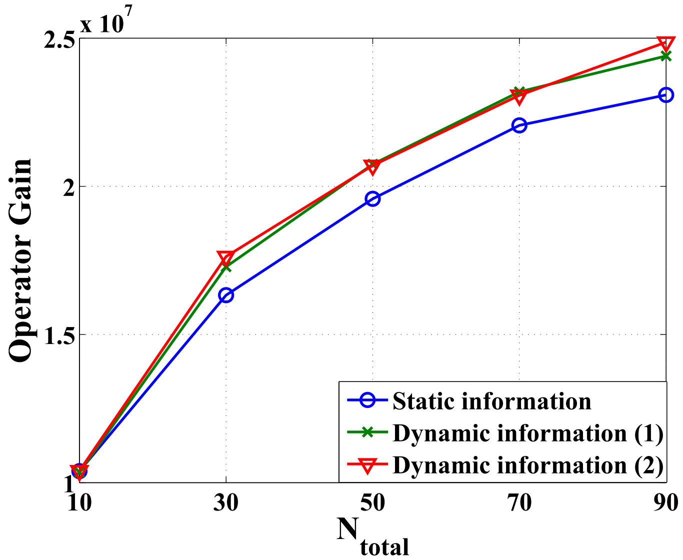

Figure 4.28: Scenario 3: Operator gain

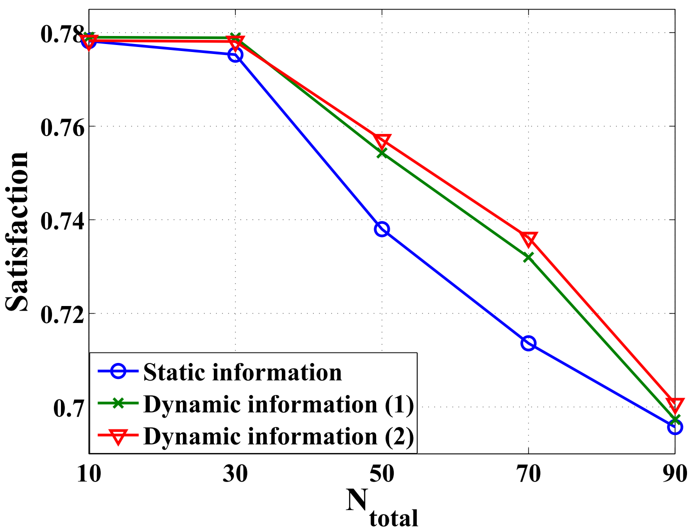

Figure 4.29: Scenario 3: User satisfaction

We depict in Fig. 4.29 the average user satisfaction. Although mobiles may pay more, we notice a higher satisfaction when tuning policies are implemented. Higher costs are then justified, since users benefit from significantly better performance. At low traffic load, tuning policies are not yet triggered. Equivalent performance, costs and subsequently satisfactions are intuitively observed. However, at very high traffic load, the performance gain over the static scheme begins to reduce; henceforth, it slightly offsets the cost considerations, leading to low discrepancy among user satisfaction.

To conclude, in comparison with the static scheme, performance results show that our tuning policies enhance network performance, provide larger operator gain and higher user satisfaction. Since it best responds to traffic load fluctuations, the slope tuning policy has proved to be an efficient strategy that enhances resource utilization. Further, in Chapter 6, we formulate tuning policies as a Semi-Markov Decision Process (SMDP), and derive optimal solutions.

4.2 Concluding Remarks

In this chapter, we evaluated our hybrid decision approach. We separately investigated decision makings, on the network and user sides. Below, we outline the main conclusions:

- When operators propose differentiated services, better network performance, higher user satisfaction, and larger operator gain can be observed. Therefore, when heterogeneous RATs are integrated, it is always beneficial if all do not provide the same QoS and cost incentives, giving mobiles a variety of possible choices.
- When mobiles are provided with minimum throughput guarantees, regardless of future network load conditions, real-time sessions see their performance enhanced.
- In comparison with well-known multi-criteria decision-making methods, namely enhanced SAW and TOPSIS, our satisfaction-based algorithm meets user needs (*e.g.,* traffic class, throughput demand, and cost tolerance), avoiding oversized and undersized decisions.
- When QoS parameters are modulated as a function of network load conditions, radio resources can be efficiently exploited. As a matter of fact, when QoS parameters are tuned according to our staircase or slope policies, better performance, higher user satisfaction, and larger operator gain are obtained, in comparison with static network information.

Chapter 5

Comparison of Our Hybrid Approach With Different Methods

In Chapter 4, our multi-criteria decision-making method, and our tuning heuristics were separately evaluated. In this chapter, we first focus on tuning thresholds, namely S_1 and S_2, and investigate their impact on network and user utilities. When QoS parameters are dynamically tuned according to the slope policy, streaming and elastic sessions are examined individually. Further, we compare our hybrid decision approach with different network-centric, mobile-terminal-centric, and hybrid methods. Peak rate maximization, Average rate maximization, Satisfaction-based using peak rate, Satisfaction-based using average rate, and exhaustive search methods are considered. Simulation results prove the effectiveness of our solution in enhancing resource utilization and user experience.

5.1 System Model

For illustration, we consider a heterogeneous wireless network composed of Mobile WiMAX and LTE RATs. They are supposed to utilize a channel bandwidth of 5 and 10 MHz respectively. Although our solution adapts to different deployment scenarios, we focus on a realistic and cost effective one, where the two RATs base stations are co-localized. The intersection of their respective zones leads to N_Z heterogeneous zones.

For the sake of simplicity, the cell is assumed divided into two zones (*i.e.,* $N_Z = 2$). While users with good radio conditions are considered adopting the (64 - QAM, 3/4) modulation and coding scheme, users with bad radio conditions are supposed to employ the (16 - QAM, 1/2) one. Their peak rates are reported in Table 5.1.

Prior to the RAT selection process, a common admission control is assumed to be performed. Further, radio resources are allocated using fair time scheduling. Yet, when mobiles select their RAT using our hybrid method, they are first provided with their mini-

RAT	**64-QAM: 3/4**	**16-QAM: 1/2**
Mobile WiMAX (5 MHz)	16.6 Mb/s	7.4 Mb/s
LTE (10 MHz)	33.5 Mb/s	14.9 Mb/s

Table 5.1: Peak rates in Mobile WiMAX and LTE

mum guaranteed throughput, given by d_{min}. Then, fair time scheduling is used to provide them with up to their maximum throughput, given by d_{max}. As long as resources are not fully committed, remaining resources are equitably distributed. Moreover, after all mobiles have received their maximum throughput, they equitably share residual resources.

Streaming and elastic sessions are individually considered in simulations. Mobiles are randomly ready either to pay for better performance, or to sacrifice within limits their service quality seeking to save up money. When user decisions need to be evaluated, or typically when their perceived satisfaction is to be computed, a set of cost tolerance parameter and QoS and cost weights is used according to user preferences (cf. Table 5.2).

Set No.	λ	w_{QoS}	w_{cost}
1	60	0.7	0.3
2	45	0.3	0.7

Table 5.2: Cost tolerance parameter and QoS and cost weights

We assume that streaming sessions have an average long-term throughput of 1 Mb/s. So as to improve their content quality, they can furthermore benefit from throughputs up to 1.5 Mb/s (*i.e.,* R_{av} = 1 Mb/s and R_{max} = 1.5 Mb/s). We depict in Table 5.3 the cost tolerance parameter λ and the weights of the decision criteria $w_{d_{min}}$, $w_{d_{max}}$ and w_{cost}, used in our hybrid approach. When profile no. 1 is assigned to users that are ready to pay for better performance, profile no. 2 is attributed to those that seek to save up money.

Profile No.	λ	$w_{d_{min}}$	$w_{d_{max}}$	w_{cost}
1	60	14/30	7/30	0.3
2	45	0.2	0.1	0.7

Table 5.3: User profiles for streaming sessions

Profile No.	λ	$w_{d_{min}}$	$w_{d_{max}}$	w_{cost}	R_c **(Mb/s)**
1	60	0	0.7	0.3	1.25
2	45	0	0.3	0.7	0.75

Table 5.4: User profiles for elastic sessions

Besides, the comfort throughput of elastic sessions, denoted by R_c, is assumed related to the user willingness to pay, and thereafter imposed by the user profile (cf. Table 5.4). Typically, when users are ready to pay for better performance, they have a comfort throughput of 1.25 Mb/s. Yet, when they seek to save up money, they are content with a comfort throughput of 0.75 Mb/s.

We report in Table 5.5 the QoS and cost parameters signaled by the network at low load factor, when using our hybrid method. We recall that the load factor represents the amount of throughput guarantees, and is defined as the ratio of the number of guaranteed allocated RUs to the total number of RUs. As RATs get loaded, d_{min} and d_{max} are linearly and separately reduced down to zero (*i.e.,* dynamically tuned according to the slope tuning policy). However, when different thresholds (*i.e.,* S_1 and S_2) are considered, different QoS parameters may be signaled for the same load conditions. This may lead to different decision makings depending on S_1 and S_2. Consequently, and before we compare our hybrid approach with other RAT selection methods, let us study the effect of S_1 and S_2 thresholds on network performance and user satisfaction.

RAT	d_{min} **(Mb/s)**	d_{max} **(Mb/s)**	*cost* **(unit/kB)**
Mobile WiMAX	1	1.5	4
LTE	1.5	2	6

Table 5.5: Initial QoS and cost parameters

To evaluate selection decisions, network and user utilities are introduced. The network utility reflects operator objectives: it is defined as the network total throughput. Furthermore, the user utility reflects the average user satisfaction: it depends on their needs and preferences, and thus take into account both QoS and cost considerations.

5.2 Effect of S_1 and S_2

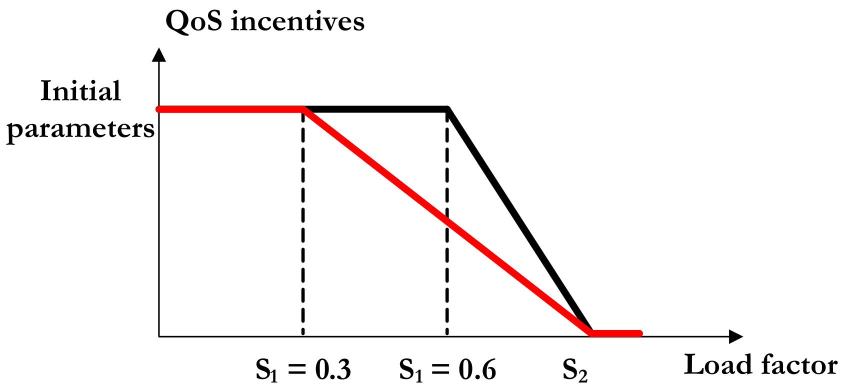

Figure 5.1: S_1 effect on signaled QoS parameters

We illustrate in Fig. 5.1 the effect of S_1 on signaled QoS parameters. The lower S_1 is, the earlier d_{min} and d_{max} get reduced, pushing more mobiles to less loaded RATs. Yet, the higher S_1, the steeper the slope. The decay rate of the QoS parameters actually increases with S_1.

Moreover, figure 5.2 depicts the effect of S_2 on signaled QoS parameters. The lower S_2, the steeper the decrease of d_{min} and d_{max}. Tuning becomes then more sensitive to load conditions. In other words, the lower S_2 is, the lower the QoS parameters are for the same load conditions, pushing more mobiles to less loaded RATs.

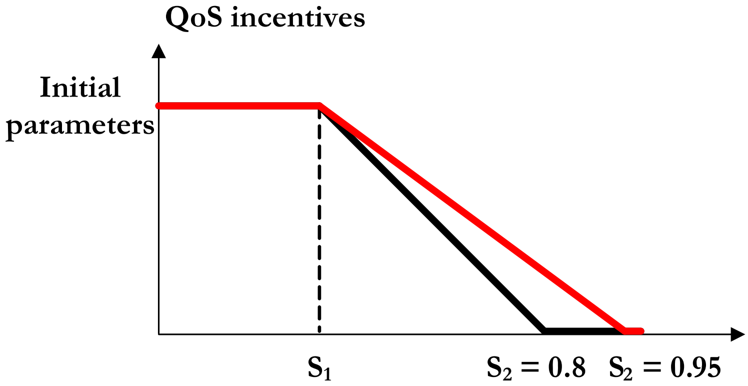

Figure 5.2: S_2 effect on signaled QoS parameters

5.2.1 Streaming Sessions

We first fix S_2 to 0.9 and vary S_1, so as to study its effect on network performance and user satisfaction.

We respectively show in Fig. 5.3 and 5.4 the network utility and the average user utility, as a function of the total throughput demand defined as the sum of user maximum throughput demands (*i.e.,* sum of user R_{max}). At very low traffic load, regardless of S_1, initial QoS parameters are broadcasted. Consequently, mobile WiMAX is generally preferred: it perfectly meets user QoS needs, while charging them less. Only users, with bad radio conditions, that are ready to pay would select the LTE technology. Equivalent decision makings are then observed for different S_1 values, leading to similar network and user utilities.

As WiMAX gets loaded, its broadcasted QoS parameters start to be reduced, pushing more arrivals to LTE. When different S_1 are examined, mobiles are differently distributed over the two RATs. Typically, when S_1 is fixed to 0.3, users are encouraged to join LTE much earlier than when S_1 is fixed to 0.6. As a result, at low and medium traffic load, the lower S_1 is, the more users join LTE and thus pay more. Similarly, the higher S_1 is, the more users continue to prefer mobile WiMAX competing for the same common resources. Yet, as shown in Fig. 5.3, mobiles can still achieve throughputs up to their R_{max} even for

$S_1 = 0.6$.

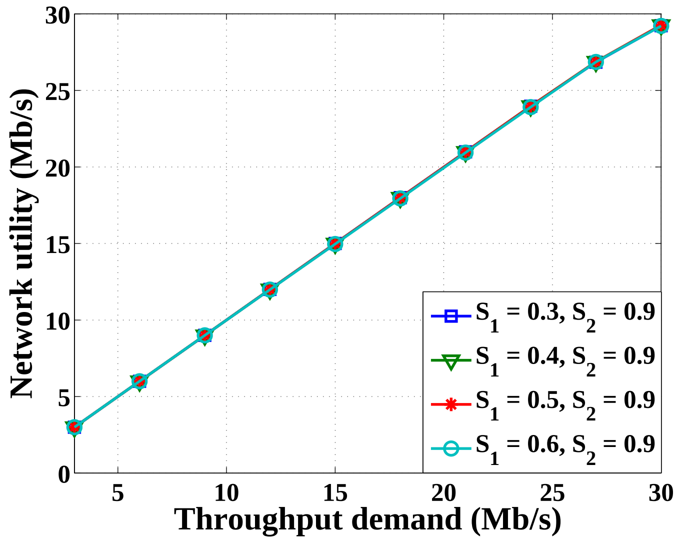

Figure 5.3: S_1 effect on network utility for streaming sessions

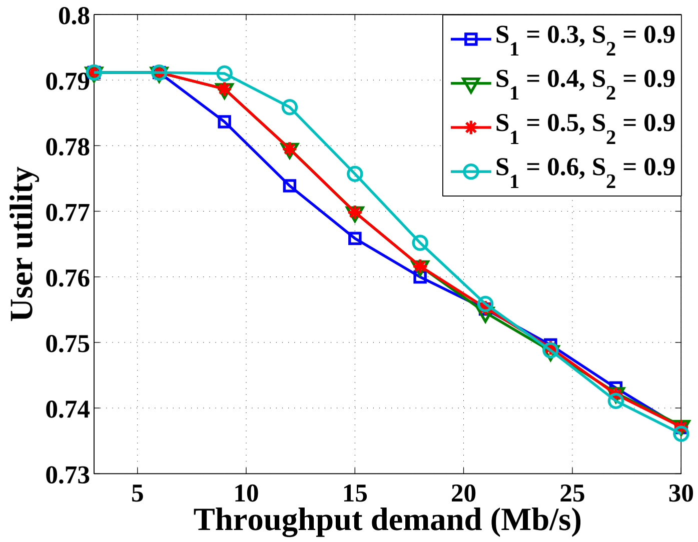

Figure 5.4: S_1 effect on user utility for streaming sessions

Actually, since their throughput demands are limited, no performance difference is observable for streaming sessions depending on S_1 (cf. Fig. 5.3). Even for $S_1 = 0.6$, at low and medium traffic load, where more users join mobile WiMAX in comparison with other cases, the network total throughput can still follow the throughput demand increase. Yet, since less users join LTE and pay more, users experience the highest satisfaction when $S_1 = 0.6$ (cf. Fig. 5.4). However, at high traffic load, the proportion of users that are associated with LTE significantly increases for high S_1 values. While the QoS parameters signaled by the WiMAX technology are being roughly reduced (high decay rate), more and more mobiles join LTE. Therefore, in the long term, the average proportion of users that are connected to LTE becomes quite similar, regardless of S_1 values. This leads to fairly close network and user utilities at high traffic load.

Furthermore, we fix S_1 to 0.6 and vary S_2, so as to study its impact on network performance and user satisfaction.

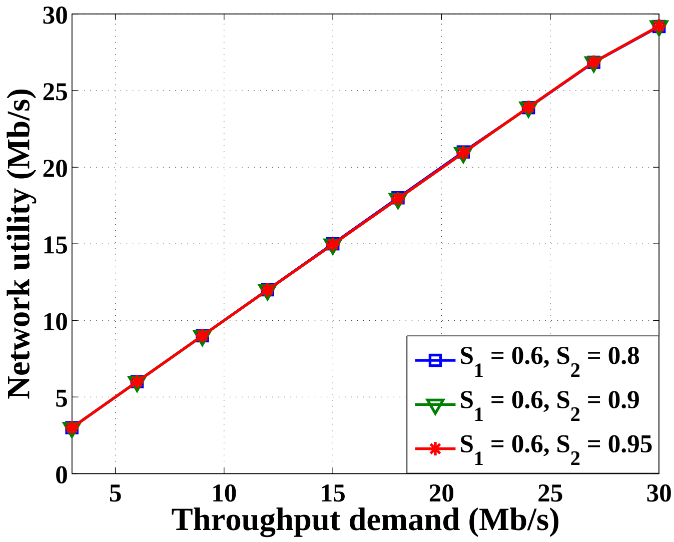

Figure 5.5: S_2 effect on network utility for streaming sessions

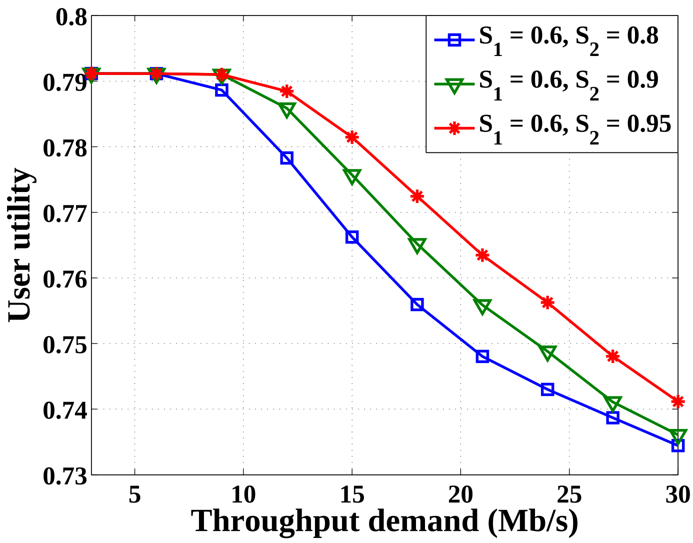

Figure 5.6: S_2 effect on user utility for streaming sessions

Following the same reasoning, the lower S_2 is, the more users are pushed to LTE. However, unlike for S_1, even when the total throughput demand is about 30 Mb/s, the proportion of users that are connected to LTE remains higher for lower S_2 values. As a matter of fact, the higher S_2 is, the longer can WiMAX provides attracting QoS guarantees for users. This leads to higher satisfaction (cf. Fig. 5.6), seeing that users perceive similar performance (cf. Fig. 5.5).

5.2.2 Elastic Sessions

Here again, we first fix S_2 to 0.9 and vary S_1, to study its impact on network performance and user satisfaction.

Fig. 5.7 and 5.8 respectively illustrate the network utility and the average user utility, as a function of the total number of users, denoted by N_{total}. The lower S_1 is, the more efficiently mobiles are distributed over the two RATs. Typically, when $S_1 = 0.3$, broadcasted QoS parameters start to be reduced much earlier in comparison with other cases. As a result, more users particularly with good radio conditions join LTE, thus enhancing resource utilization.

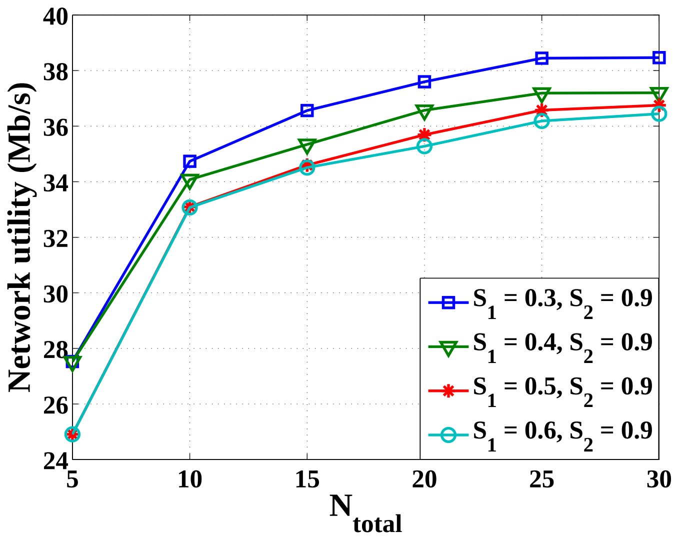

Figure 5.7: S_1 effect on network utility for elastic sessions

As a matter of fact, as tuning starts earlier, even mobiles with good radio conditions, that are typically ready to pay (*i.e.*, having a comfort throughput of 1.25 Mb/s), start earlier to join LTE. Consequently, and since elastic sessions adapt to resource availability, the network total throughput (*i.e.*, the network utility) is improved as shown in Fig. 5.7.

At low and medium traffic load, when S_1 is fixed to 0.3, more users particularly with good radio conditions join LTE in comparison with other cases. This better exploits LTE resources, enhancing network utility. Since less users are connected to WiMAX, and more users including those with good radio conditions join LTE, users have on average better performance. Yet, as they pay on average more (more users are connected to LTE), users perceive equivalent satisfaction regardless of S_1 values (cf. Fig 5.8).

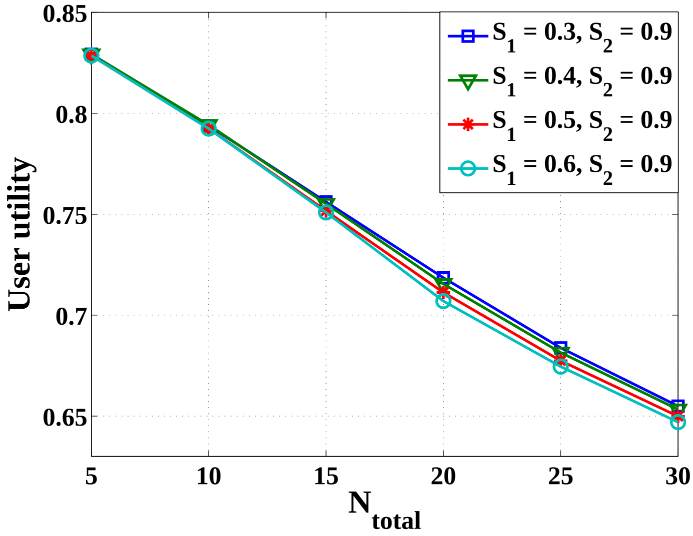

Figure 5.8: S_1 effect on user utility for elastic sessions

As N_{total} increases, the lower S_1 is, the higher is the average proportion of users with good radio conditions that are connected to LTE. This leads to continuously higher network utility. Moreover, and since in the long term the average proportion of users that are connected to LTE becomes close regardless of S_1 values, users perceive higher satisfaction for lower S_1 values.

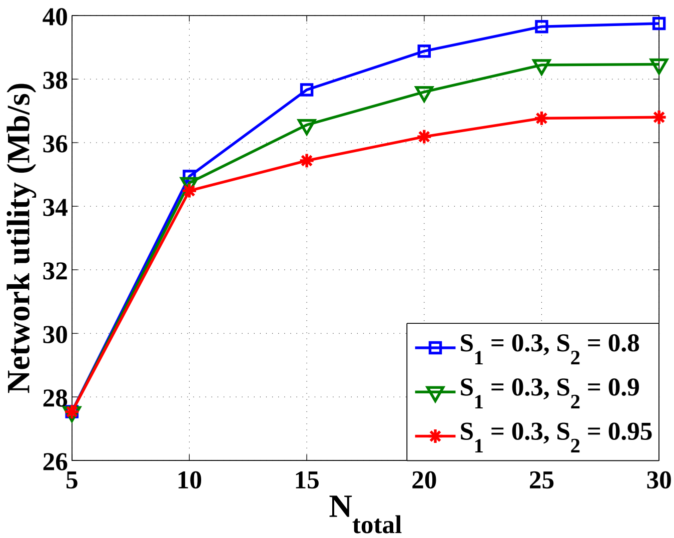

Figure 5.9: S_2 effect on network utility for elastic sessions

Hereafter, we fix S_1 to 0.3 and vary S_2, so as to study its effect on network performance and user satisfaction. Following the same reasoning, the lower S_2 is, the more users particularly with good radio conditions join LTE leading to higher network utility (cf. Fig. 5.9). Further, as for streaming sessions, the higher S_2 is, the more users join WiMAX even for $N_{total} = 30$. As a consequence, for different S_2 values, cost considerations offset performance improvement leading to close user satisfaction (cf. Fig. 5.10).

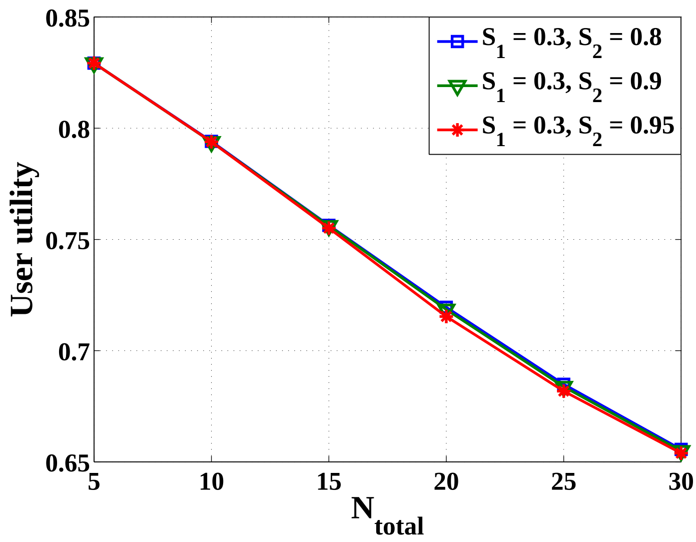

Figure 5.10: S_2 effect on network utility for elastic sessions

To conclude, we demonstrated the network ability to globally control the decisions of streaming and elastic sessions. User decisions strongly depend on how the network derives its cost and QoS parameters, and thereafter on S_1 and S_2 tuning thresholds. For the same load conditions, different threshold values have lead to different network and user utilities. Moreover, we showed that S_1 and S_2 should be set depending on session traffic classes. When a common admission control is assumed to be performed prior to RAT selections, high and low threshold values are adapted for streaming and elastic sessions, respectively.

5.3 Comparison With Multiple RAT Selection Methods

In what follows, we compare six different RAT selection methods, including our hybrid decision approach:

- Peak rate maximization: Mobile users have no information on the global network state. Based on their radio conditions, they select the RAT that offers them the best peak rate.

- Average rate maximization: Mobiles are assumed to know the exact number of users that are connected to available RATs. Assuming that fair time scheduling is employed, they select the RAT that offers them the best throughput, at the time of selection. Their estimated throughput in RAT x, denoted by $\overline{D^x}$, is computed as:

$$\overline{D^x} = \frac{D^x}{1 + N^x} \tag{5.3.1}$$

where D^x represents the user peak rate when connected to RAT x, and N^x represents the number of users that are connected to RAT x at the time of selection.

- Satisfaction-based using peak rate (SB - PR): Using their peak rates, mobiles adopt the Satisfaction-based multi-criteria decision-making method to select their best RAT. In order to evaluate serving RATs, the provided QoS parameters, in Eq. 3.2.2 and 3.2.3, are replaced with the peak rate that mobiles can achieve when connected to these RATs.

- Satisfaction-based using average rate (SB - AR): Mobiles use the Satisfaction-based multi-criteria decision-making method to select the RAT that maximizes their expected utility. In Eq. 3.2.2 and 3.2.3, the provided QoS parameters are replaced with the estimated average throughput that mobiles can obtain.

- Exhaustive search: The network considers all possible associations involving all users. It finally selects the combination that optimizes its own utility. Actually, it assigns mobiles with either WiMAX or LTE in a way to maximize the network total throughput. This is known to be the optimal method with respect to operator objectives: it leads to the highest network utility.

- Our hybrid approach: The network periodically sends decisional information (*i.e.*, cost and QoS parameters) to assist mobile users in their decisions. A RAT is considered to be low-loaded when its load factor is below S_1. Initial d_{min} and d_{max} are then signaled (cf. Table 5.5). Yet, when its load factor exceeds S_2, a RAT is considered to be highly loaded, providing no QoS guarantees.

When using the peak rate maximization and the SB - PR methods, mobiles select their RAT without any network assistance. Decisions are then mobile-terminal-centric. However, when employing the average rate maximization and the SB - AR methods, load conditions signaled by the network assist mobile users in their decisions. The latter two methods are thus considered to be hybrid. Finally, when adopting the exhaustive search method, decisions are network-centric, since they are made by the network transparently to end-users.

Because in practice telecom operators will not reveal neither the exact numbers of users that are connected to their RATs nor the scheduling algorithm they adopt, the average rate maximization and the SB - AR methods are not realistic. Yet, they serve as a means to illustrate the gain from masking network load conditions and only signaling cost and some QoS parameters, so as to enhance resource utilization.

5.3.1 Streaming Sessions

Fig. 5.11 and 5.12 respectively show the network utility and the average user utility, as a function of the total throughput demand.

The network utility, defined as the network total throughput, generally increases with the total throughput demand. Yet, when a RAT gets overloaded, its total throughput stagnates and no longer increases with additional throughput demand.

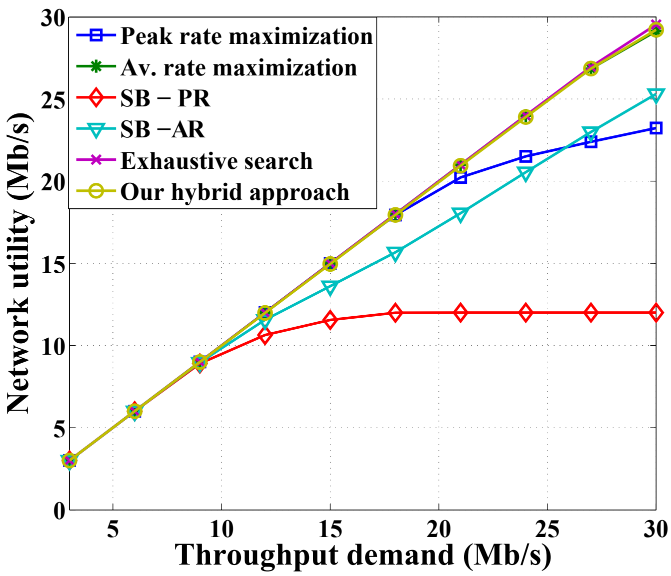

Figure 5.11: Network utility for streaming sessions

When the SB - PR method is used, all users select the mobile WiMAX technology (*i.e.,* Mobile WiMAX is their best trade-off between cost and QoS decision criteria). Regardless of user preferences and radio conditions, mobile WiMAX is expected to provide mobile users with the highest utility. Since mobiles use their peak rate in estimating their utility, their decisions do not depend on network load conditions. As a result, mobiles continue to select the WiMAX technology even when it gets overloaded.

At low traffic load, mobile WiMAX can meet user QoS needs, while charging them less. When users benefit from throughputs up to their R_{max} and pay less, they have the highest utility (*i.e.,* satisfaction). However, when WiMAX gets loaded, it becomes no longer able to fulfill user QoS needs. Typically, at medium and high traffic load, WiMAX becomes saturated leading to a significant decrease of the user throughput below R_{av} (cf. Fig. 5.11). As a consequence, user satisfaction will also dramatically decrease (cf. Fig. 5.12).

Furthermore, when the peak rate maximization method is adopted, all users select the LTE technology. Independently of their modulation and coding schemes, mobiles can achieve the best peak rate when connected to LTE. Here again, their decisions do not change with network load conditions. As a consequence, at high traffic load, user throughput goes below R_{max}. Yet, it continues to be greater than R_{av}.

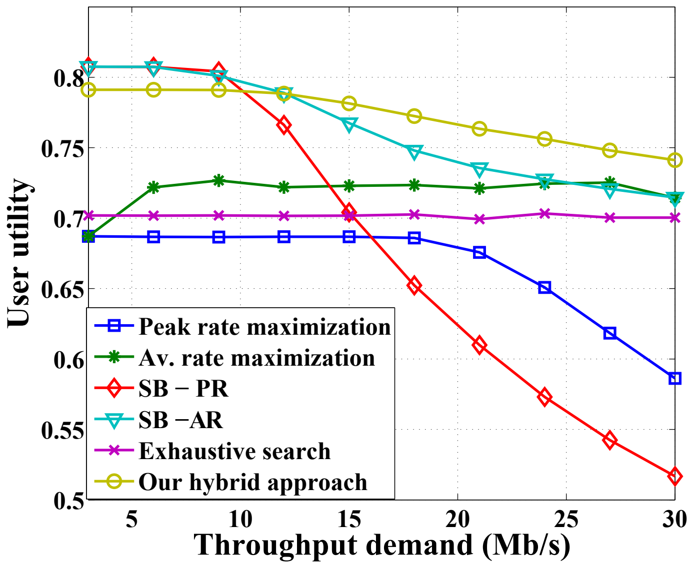

Figure 5.12: User utility for streaming sessions

Further, since LTE charges more than WiMAX does, mobile users experience the lowest satisfaction level at low traffic load. Actually, when all RAT selection schemes meet user QoS needs, the peak rate maximization method assign all users to the LTE technology, thus charging them more. At high traffic load, because user throughput decreases, their experienced utility also diminishes.

Moreover, when the SB - AR method is employed, users combine their needs and preferences with network load conditions to select their best RAT. As a consequence, at low traffic load and regardless of their radio conditions, all users select the mobile WiMAX technology: their QoS needs are perfeclty met while paying less. This leads to the highest user utility, as in the case of the SB - PR method. However, when the mobile WiMAX gets loaded, users may start to join LTE according to their radio conditions and preferences (*i.e.,* their willingness to pay for better performance). Based on their modulation and coding scheme, as well as on their cost tolerance parameter and decision criteria weights (cf. Table 5.2), users estimate the utility they can obtain in both available RATs. They then select the RAT with the highest expected utility. In fact, users with bad radio conditions that are ready to pay for better performance are the first to start to join LTE. Besides, users with good radio conditions that seek to save up money are the last to join LTE.

Consequently, since users are not proportionally distributed over the two RATs, mobile WiMAX gets overloaded before LTE. Thus, the growth rate of the network utility decreases as the total throughput demand increases (cf. Fig. 5.11). This means that the average user throughput decreases. Yet, it remains greater than R_{av}. When some users start joining LTE and so pay more while others, connected to WiMAX, start perceiving lower throughputs, the average user satisfaction also decreases as the total throughput demand increases (cf. Fig. 5.12).

Furthermore, our hybrid approach and the average rate maximization method perfectly meet user QoS needs, even at high traffic load. Their network utility, as depicted in Fig. 5.11, is very close to that of the exhaustive search method, known to be the optimal one with respect to resource utilization. Yet, as shown in Fig. 5.12, our hybrid approach provides the highest user utility.

On the one hand, when the average rate maximization method is used, mobiles select the RAT that offers them the best throughput. Therefore, load balancing is achieved: Mobile WiMAX and LTE are similarly occupied with respect to their maximum capacity. As a result, the network utility can likely follow the throughput demand increase. On the other hand, when our hybrid approach is employed, the network modulates the broadcasted QoS parameters as a function of its load conditions. It tries to push future arrivals to less loaded RATs, thus enhancing resource utilization. By integrating their needs and preferences, mobiles can avoid oversized decisions, and so improve their perceived satisfaction. Typically, at low traffic load, when both RATs can perfectly meet user QoS needs, mobile WiMAX will be preferred since it charges less. This explains why, when using our hybrid method, user utility is constantly higher than when adopting the average rate maximization method. The latter ignores user preferences (*i.e.,* its willingness to pay for better performance or to save up money) and mainly deals with load balancing. However, because the proportion of users that are connected to the LTE technology is almost constant and the user throughput is always close to R_{max}, user utility hardly changes as a function of the total throughput demand. On the other side, when using our hybrid method, since the proportion of users that are connected to LTE increases with the total throughput demand, the average user utility decreases, since LTE charges more than WiMAX. Yet, it always remains greater than that of the average rate maximization method.

Moreover, when using the exhaustive search method, the network involves all users at each decision epoch: it considers all possible combinations and selects the one that maximizes its own utility. Since user needs and preferences are ignored, and RATs are not statistically similarly occupied, this network-centric method provides the lowest user utility amongst the average rate maximization method and our hybrid approach. As a matter of fact, the network seeks to optimize its own utility, regardless of user preferences. In other words, when different combinations lead to the same network utility, they are assumed equivalent. The one that better distributes mobiles over the two RATs has no priority, since it does not improve the network utility defined as the network total throughput. As a result, the proportion of users that are connected to LTE is statistically higher than those of the average rate maximization method and our hybrid method, leading to lower user satisfaction.

To conclude, so as to illustrate the gain from masking network load conditions and only signaling cost and some QoS parameters, we compare our hybrid approach with the SB - AR one. Actually, when using our hybrid method, we can push users to LTE long before WiMAX really gets overloaded. By reducing the broadcasted QoS parameters in WiMAX, even with $S_1 = 0.6$ and $S_2 = 0.95$, future arrivals are encouraged to join LTE much earlier than the SB - AR scenario. Thereby, sessions are better distributed over the two RATs, leading to higher network utility as shown in Fig. 5.11.

At low traffic load, both methods perfectly meet user QoS needs. Yet, since the proportion of users that are connected to the most expensive RAT (*i.e.,* LTE) is higher when our hybrid approach is used, user satisfaction is lower than that of the SB - AR method. However, at high throughput demand, because future arrivals start to join LTE much earlier than the SB - AR case, WiMAX is on average less loaded when using our hybrid approach. As a consequence, WiMAX can better serve its on-going sessions, leading to higher user throughput. Therefore, although mobiles may pay more (*i.e.,* the proportion of users that are connected to LTE is higher), they experience significantly better performance leading to higher satisfaction (Fig. 5.12). After all, by dynamically tuning QoS parameters, the network enhances resource utilization while mobiles maximize their satisfaction (cf. Fig. 5.12).

5.3.2 Elastic Sessions

We respectively depict in Fig. 5.13 and 5.14 the network utility and the average user utility, as a function of the total number of users denoted by N_{total}.

When connected alone to a RAT, an elastic session can occupy all of the available resources. However, when several sessions are present, they all share these resources. As a result, the network utility, defined as the network total throughput, do not usually change as a function of the total number of users N_{total} (cf. Fig 5.13). Yet, the average user throughput is reduced.

As in the case of streaming sessions, when the SB - PR method is used, all users are connected to mobile WiMAX regardless of the network load conditions. As shown in

Fig. 5.13, the network total throughput (*i.e.*, the network utility) is close to 12 Mb/s independently of N_{total}: it actually corresponds to the weighted average total throughput, taking into account users with both good and bad radio conditions. However, the average user throughput linearly decreases with N_{total}, leading to a significant decrease of the user satisfaction (cf. Fig. 5.14).

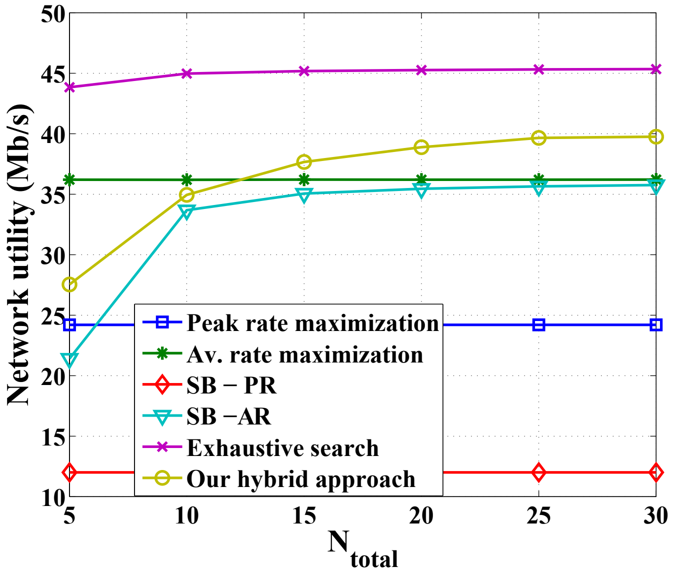

Figure 5.13: Network utility for elastic sessions

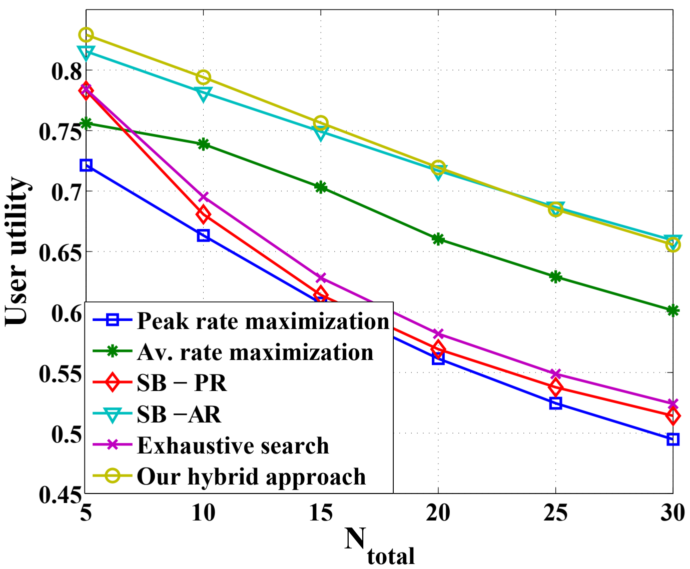

Figure 5.14: User utility for elastic sessions

Moreover, when the peak rate maximization method is adopted, all users select LTE.

The network utility is then, on average, higher than that of the SB - PR method. As a consequence, user throughput is also higher. But, since all users are connected to the most expensive RAT (*i.e.,* LTE), the satisfaction improvement with respect to the perceived throughput criterion fails to offset the satisfaction decrease with respect to the cost criterion. This leads to a lower user satisfaction in comparison with the SB - PR case (cf. Fig. 5.14).

Furthermore, when the exhaustive search method is employed, optimal resource utilization is achieved as shown in Fig. 5.13. Yet, the average user utility is not that interesting. First, when assigning mobiles to the available RATs, this network-centric method do not consider user preferences. It actually ignores user willingness to pay for better performance or to save up money, and only seeks to maximize the network total throughput. Second, in order to better exploit the available resources, only few users with good radio conditions may be assigned to LTE. The majority, with bad and also good radio conditions, will be connected to mobile WiMAX, all competing for the same resources. As a result, few users connected to LTE will have excellent throughputs, that far outweigh their R_c. The others will experience relatively low throughputs, that may be well below their R_c. This association optimizes the network total throughput, but not the user satisfaction (cf. Fig. 5.14).

In comparison with the exhaustive search method, mobiles are better distributed over the two RATs, when the average rate maximization method is adopted. In fact, users select the RAT that offers them the best throughput, leading to load balancing as in the streaming case. As a result, mobiles with equivalent radio conditions will have close throughputs regardless of their access technology. Since even users with bad radio conditions may be connected to LTE, the network utility is on average lower than that of the exhaustive search method, known to be the optimal one. However, because on average perceived throughputs better meet user needs (*i.e.,* their R_c), the user utility is significantly higher than that of the exhaustive search approach.

Moreover, when the SB - AR method is used, mobile users combine their needs and preferences with the network load conditions, so as to select their best RAT. At low traffic load (typically for $N_{total} = 5$), more users select the mobile WiMAX technology in comparison with the average rate maximization method. When WiMAX can meet user needs very well, it charges them less. Occasionally, based on the current load conditions, a user with bad radio conditions, that is ready to pay for better performance, would select the LTE technology. As N_{total} increases, more users including those with good radio conditions start to join LTE, leading to higher network utility. The latter remain almost constant at medium and high load conditions. On average, it is slightly lower than that of the average rate maximization method. Yet, since selection decisions take into account user needs and preferences, typically their cost considerations, the user utility is significantly better than that of the average rate maximization method.

Lastly, by masking network load conditions and only signaling some cost and QoS parameters, our hybrid approach drives user decisions in a way to enhance resource utilization. At low traffic load, more users typically those with bad radio conditions, that are ready to pay, select LTE. This leads to a higher network utility in comparison with the SB - AR method where, as explained before, users may occasionally join LTE (cf. Fig. 5.13). As a result, and although users pay on average more, they experience higher satisfaction since they have quite better throughput.

As N_{total} increases, QoS parameters are reduced with $S_1 = 0.3$ and $S_2 = 0.8$. As a consequence, future arrivals are encouraged to join LTE much earlier than the SB - AR case. However, users with good radio conditions that seek to save up money are the last to start joining LTE. In comparison with the SB - AR method, most users that are connected to WiMAX have good radio conditions, and more users with either good and bad radio conditions are connected to LTE. This leads to higher network total throughput, as shown in Fig. 5.13. Yet, the user utility is pretty close to that of the SB - AR scenario, since users having better performance pay on average more.

To wrap up, in comparison with different RAT selection schemes, including network-centric, hybrid and mobile-terminal-centric approaches, simulation results prove the efficiency of our hybrid approach in enhancing resource utilization and maximizing user satisfaction. In the streaming sessions scenario, it optimizes the network total throughput and maximizes the average user utility, except at low traffic load where the non-realistic SB - AR method provides higher user satisfaction. Also, in the elastic sessions scenario, our hybrid approach significantly enhances resource utilization and maximizes user utilities, in comparison with various hybrid and mobile-terminal-centric methods. Furthermore, compared with the exhaustive search method, known to be the optimal one with respect to resource utilization, our hybrid approach provides significantly higher user satisfaction.

5.4 Concluding Remarks

In this chapter, we further investigated our tuning heuristics, and studied the impact of S_1 and S_2 thresholds on network and user utilities. Simulation results showed that user decisions strongly depend on network information, and thereafter on S_1 and S_2 values. Moreover, we compared our hybrid decision approach with multiple network-centric, mobile-terminal-centric, and hybrid methods.

When users do not cooperate neither with each other nor with the network, they have no information on the global network state. As a result, their selection decisions may be in no one long-term interest, leading to performance inefficiency. Moreover, when network elements take selection decisions transparently to end-users, resource utilization is optimized. Yet, user needs and preferences are not efficiently met, leading to relatively low user satisfaction. However, when our hybrid approach is used, the network partially cooperates with mobiles assisting them in their decisions. As a matter of fact, the network masks its load conditions, and only signals cost and some QoS parameters. This decisional information guides user decisions in a way to enhance resource utilization. Besides, as user needs and preferences are also involved, selection decisions maximize user satisfaction.

We proved as well the efficiency of masking network load conditions, and only signaling cost and some QoS parameters, in enhancing resource utilization and user satisfaction. In fact, our hybrid approach outperforms non-realistic methods, where mobiles have a perfect knowledge of the network state (*i.e.,* number of users connected to available RATs). So, when operator objectives are implicitly integrated within signaled QoS parameters, radio resources are better utilized, and user satisfaction is maximized.

Finally, to conclude, compared with various hybrid and mobile-terminal-centric methods, our hybrid approach maximizes the network total throughput and the average user satisfaction. Also, compared with the optimal exhaustive search method, our approach provides significantly higher user utility.

Chapter 6

Optimizing Network Information for RAT Selection

The basic idea of our hybrid decision approach was first presented in Chapter 3, where heuristic policies are introduced to tune network information as a function of the load conditions. Simulations considered static scenarios, where mobiles are assumed to arrive sequentially, and to stay long connected to their serving RAT. A common admission control is assumed to be performed, thus limiting the total number of arrivals. In the present chapter, deriving network information is formulated as a Semi-Markov Decision Process. We first define network states, actions, state dynamics and rewards. An optimal policy (i.e., network information to signal in each state) is derived through the Policy Iteration algorithm, in a way to dynamically optimize long-term network reward. User dynamics, namely user arrivals and departures, are taken into account. Morevoer, transitions between network states depend not only on network actions, user arrival and departure rates, but also on user needs, preferences and decison-making algorithms. When all these parameters can not be easily obtained in constantly varying networks, a reinforcement learning approach is further presented to derive network information. The performances of optimal, learning-based and heuristic policies are analyzed. When tuning thresholds are pertinently set, our heuristic method provides very close performance to the optimal one. Moreover, although lower performances are observed, our learning-based algorithm has the crucial advantage of requiring no prior parameterization.

6.1 Introduction

We introduced, in Chapter 3, a hybrid RAT selection approach. The network provides information for the mobiles to make robust decisions. More precisely, mobile users select their RAT depending on their needs and preferences, as well as on the monetary cost and

QoS parameters signaled by the network. By appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states. We also presented two heuristic methods, namely the staircase and the slope tuning policies, to derive network information as a function of the load conditions. Simulations considered static scenarios, where mobiles are assumed to arrive sequentially, and to stay long connected to their serving RAT. A common admission control is assumed to be performed, thus limiting the total number of arrivals. The network dynamics will then progressively slow down until a pseudo-stationnary regime is attained, where all measurements were performed.

However, to maximize long-term network performance, network information should depend not only on current load conditions, but also on expected future demands. Deriving network information is then formulated as a Semi-Markov Decision Process (SMDP) [Put94]. The aim is to dynamically meet operator objectives, while mobiles maximize their own utility. Simulations consider dynamic scenarios, where user arrivals and departures are taken into account. Also, when network parameters are not perfectly known, a reinforcement learning approach is introduced to derive what to signal to mobiles. The network learns user needs, preferences and decision-making algorithms through interacting with them. Among the different existing reinforcement learning (RL) algorithms, we select the Q-learning method for its simplicity.

Furthermore, and as discussed in Chapter 2, SMDP and Q-learning have been widely employed in RAT selection. In [ZYNT12b, ZYNT12a, ZJJ+12, SAAS10, IKT09, CKG08a, CKG08b], RAT selection is modeled as a semi-Markov decision process. The network finds an optimal policy that maximizes its long-term reward, without aligning with user preferences. Also, in [TFC12, DO12b, DO12a], mobiles learn selection decisions through trial-and-error interaction with their dynamic environment. Yet, because of the non-cooperative behavior of mobile users, their performance may be degraded. In this chapter, SMDP and Q-learning are used in a hybrid decision approach. They enable the network to derive information for the mobiles to make decisions.

6.2 Network Model

6.2.1 Network Topology

Consider a heterogeneous wireless network composed of two OFDM(A)-based radio access technologies. Let x_1 and x_2 designate the two serving RATs within the network. Although our method adapts to different deployment scenarios, we focus on a realistic and cost effective one, where the two RATs base stations are co-localized. The modulation and coding scheme, that can be assigned to a user connected to RAT x, differs depending on its radio conditions in the cell. As the number of possible modulation and coding schemes is limited, we decompose the cell into N_Z zones with homogeneous radio characteristics. Users in zone Z_k, $k = 1, ..., N_Z$, employ $mod^x(k)$ with $cod^x(k)$ as modulation and coding scheme, if connected to RAT x. Moreover, and for the sake of simplicity, users in a same zone are assumed to have the same peak throughput, realized when connected alone to the cell.

6.2.2 Network Resources

The radio resource is divided into time-frequency resource units (RUs). Users in zone Z_k can transmit up to $b^x(k)$ bits per resource unit, when connected to RAT x:

$$b^x(k) = N_s^x \cdot N_f^x \cdot log_2[mod^x(k)] \cdot cod^x(k) \cdot (1 - BLER) \tag{6.2.1}$$

where N_s^x and N_f^x respectively denote the number of OFDM symbols and subcarriers per RU, and $BLER$ the block error rate obtained as a function of the user signal-to-noise ratio. At decision epochs, because RAT selections are made for a sufficiently long period of time (*e.g.,* session duration, user dwell time in the cell), mobiles are interested in their average radio conditions, rather than in their instantaneous ones.

In the time dimension, resources are organized into frames of length T^x. When RAT x allocates N_{RU} resource units per frame to a user in zone Z_k, its average throughput d is given by:

$$d = \frac{N_{RU} \cdot b^x(k)}{T^x} \tag{6.2.2}$$

6.2.3 Traffic Model

Users belong to N_C traffic classes. In our work, we focus on both streaming ($c = 1$) and elastic ($c = 2$) traffic classes. Class c arrivals, in zone Z_k, follow a Poisson process of rate $\Lambda(k, c)$. We assume that streaming sessions have an average long-term throughput of R_{av}. Yet, to improve their content quality, they can benefit from throughputs up to R_{max}. Their duration is considered to be exponentially distributed with a mean of $1/\mu_1$.

Moreover, elastic sessions adapt to resource availability. Their needs are expressed as comfort throughput denoted by R_c, and their size is assumed to be exponentially distributed with a mean of L bytes. However, in addition to their size, their service rate μ_2 also depends on their average throughputs.

6.3 Hybrid Decision Framework

6.3.1 Network Information

Periodically or upon user request, network information is sent to all mobiles using the logical communication channel (*i.e.,* radio enabler) proposed by the IEEE 1900.4 standard [Std09]. In our work, depending on network information, user needs and preferences, mobiles make final decisions regarding selection of their most appropriate RAT. However, by appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states.

We recall that, for RAT x, the network broadcasts partial QoS parameters, namely $d_{min}(x)$ and $d_{max}(x)$, and the cost to pay per amount of traffic, namely $cost(x)$. More precisely, mobiles are guaranteed an average minimum throughput $d_{min}(x)$, and have priority to be allocated up to an average maximum throughput $d_{max}(x)$. As $d_{min}(x)$ and $d_{max}(x)$ are derived for a generic user with the most robust modulation and coding scheme, individual users need to deduce their own QoS parameters. For that, mobiles in zone Z_k multiply the QoS parameters, signaled by the network, with their modulation and coding gain, denoted by $g(k)$.

6.3.2 RAT Selection

Using the satisfaction-based multi-criteria decision-making method we have introduced in Chapter 3, mobiles compute a utility function for each of the available RATs, and select the one with the highest score. This utility depends on user radio conditions, needs and preferences (*i.e.,* traffic class, throughput demand, QoS-maximizing or cost-minimizing preferences), as well as on the cost and QoS information sent by the network.

In our work, when $cost(x)$ is maintained fixed, $d_{min}(x)$ and $d_{max}(x)$ are dynamically tuned trying to globally control user decisions. Let N_I^x be the number of possible ($d_{min}(x)$, $d_{max}(x)$) couples, that may be signaled to incite mobile users to join RAT x. In the next section, selecting the ($d_{min}(x)$, $d_{max}(x)$) couple to be broadcasted, for each RAT x, is formulated as a Semi-Markov Decision Process (SMDP). The goal is to dynamically optimize the long-term discounted network reward, while mobiles maximize their own utility.

6.4 Semi-Markov Decision Process

At each user arrival or departure, signaled network information may have to vary. In this section, the SMDP is used to dynamically decide of the QoS parameters in a way that optimizes the long-term network reward. We first start by defining network states, actions, state dynamics and rewards. Next, using the *Policy Iteration* algorithm, we compute the optimal solution.

6.4.1 Network States

For $\{k = 1, ..., N_Z, c = 1, ..., N_C, i = 1, ..., N_I^x\}$, we define a state of RAT x to be the ($N_Z \times N_C \times N_I^x$)-tuple $n^x(t)$:

$$n^x(t) = (n^x(k, c, i, t)),$$

where $n^x(k, c, i, t)$ is a stochastic process representing the number of class c users in zone Z_k, that have joined RAT x with the i^{th} ($d_{min}(x)$, $d_{max}(x)$) couple, at time t. In the remaining, we omit t as we assume stationarity.

To protect ongoing sessions, an admission control policy is applied: new arrivals may join RAT x, with the i^{th} ($d_{min}(x)$, $d_{max}(x)$) couple, to the extent that RAT x available resources are enough to meet their d_{min}, while not compromising the QoS guarantees of ongoing sessions. Consequently, the set of admissible states in RAT x is:

$$\mathcal{N}_a^x = \left\{ n^x \in \mathbb{N}^{N_Z \times N_C \times N_I^x} \mid \sum_{k=1}^{N_Z} \sum_{c=1}^{N_C} \sum_{i=1}^{N_I^x} n^x(k, c, i) \cdot N_{min}^x(i) \leq N_{total}^x \right\} \tag{6.4.1}$$

where $N_{min}^x(i)$ is the number of RUs necessary to guarantee the d_{min} of the i^{th} QoS parameters couple, and N_{total}^x is the total number of RUs used for data transmission in RAT x.

Let the ($N_Z \times N_C \times N_I^{x_1} + N_Z \times N_C \times N_I^{x_2}$)-tuple $s = (n^{x_1}, n^{x_2})$ be the state of the heterogeneous network, defined as the concatenation of RAT x_1 and RAT x_2 substates. The state space $\mathcal{S}$ of the network is then defined as:

$$\mathcal{S} = \{s = (n^{x_1}, n^{x_2}) \mid n^{x_1} \in \mathcal{N}_a^{x_1}, n^{x_1} \in \mathcal{N}_a^{x_2}\}$$

6.4.2 Network Actions

In each state, an action is taken by the network: QoS incentives to join serving RATs are derived. An action a is the quadruple defined by $a = (d_{min}(x), d_{max}(x))$, $x \in \{x_1, x_2\}$, where $d_{min}(x)$ and $d_{max}(x)$ represent the QoS parameters of RAT x, for the most robust modulation and coding scheme. Based on their needs (*e.g.,* traffic class, throughput demand) and preferences, as well as on their modulation and coding scheme (*i.e.,* geographical position), users act differently upon these actions.

Obviously, $N_I^{x_1} \cdot N_I^{x_2}$ actions are possible. However, given a state $s = (n^{x_1}, n^{x_2})$, not all actions are feasible. We then denote by $\mathcal{A}$ the set of all possible actions, and by $\mathcal{A}(s) \subset \mathcal{A}$ the subset of feasible actions in state s.

When both RATs provide no QoS incentives (*i.e.*, $d_{min}(x_1) = d_{max}(x_1) = d_{min}(x_2) = d_{max}(x_2) = 0$), action a is blocking and new arrivals are rejected.

6.4.3 State Dynamics

As the network does not completely control individual decisions, transitions between network states do not only depend on network actions, user arrival and departure rates, but also on user needs and preferences. Consequently, the decision making on the mobile side, using a multi-criteria decision-making method, has a probabilistic impact on the transition rates.

Let $p^x(k,c,a)$ represent the probability that class c users in zone Z_k select RAT x, when action a is adopted. As action a may be blocking, $p^{x_1}(k,c,a) + p^{x_2}(k,c,a)$, $\forall k, c$, is not necessarily equal to one: it can be either zero or one. Transition rates $T(s,s',a)$ between states $s = (n^{x_1}, n^{x_2})$ and s' are then expressed as:

$$\begin{cases} \Lambda(k,c)\, p^{x_1}(k,c,a) & \text{if } s' = (n^{x_1} + e^{x_1}(k,c,i), n^{x_2}) \\ \Lambda(k,c)\, p^{x_2}(k,c,a) & \text{if } s' = (n^{x_1}, n^{x_2} + e^{x_2}(k,c,i)) \\ n^{x_1}(k,c,i)\, \mu_c^{x_1}(s) & \text{if } s' = (n^{x_1} - e^{x_1}(k,c,i), n^{x_2}) \\ n^{x_2}(k,c,i)\, \mu_c^{x_2}(s) & \text{if } s' = (n^{x_1}, n^{x_2} - e^{x_2}(k,c,i)) \\ 0 & \text{Otherwise} \end{cases} \tag{6.4.2}$$

where $e^x(k,c,i)$ is defined as a $(N_Z \times N_C \times N_I^x)$-tuple containing all zeros except for the $(k,c,i)^{th}$ element, that is equal to one, and new arrivals join RAT x with the i^{th} QoS parameters couple proposed by action a. Hence, for example, when a class c user in zone Z_k joins RAT x_1, with the i^{th} QoS parameters couple, the network moves to state $s' = (n^{x_1} + e^{x_1}(k,c,i), n^{x_2})$.

The state dynamics can equivalently be characterized by the state transition probabilities $p(s,s',a)$ of the embedded chain:

$$p(s,s',a) = T(s,s',a) \cdot \ \tau(s,a) \tag{6.4.3}$$

where $\tau(s,a)$ is the expected sojourn time for each state-action pair, defined as follows:

$$\left\{ \sum_x \sum_k \sum_c [\Lambda(k,c) p^x(k,c,a) + \sum_i n^x(k,c,i) \mu_c^x(s)] \right\}^{-1} \tag{6.4.4}$$

6.4.4 Network Reward

To formulate optimization objectives, let $r(s,a)$ denote the permanence reward earned by the network in state s, when action a is adopted. Unlike the impulsive reward, received upon transitions, the permanence reward represents the benefit and penalty continuously received by the network whilst in state s (*i.e.,* it is actually defined on a per unit time basis). In our work, we express $r(s,a)$ as the sum of a network utility $N(s,a)$ and a blocking term $B(s,a)$:

$$r(s,a) = N(s,a) + B(s,a) \tag{6.4.5}$$

The network utility is given by:

$$N(s,a) = \sum_x \sum_k \sum_c \sum_i n^x(k,c,i) d^x(k,c,i) \tag{6.4.6}$$

where $d^x(k,c,i)$ represents the average throughput of class c users in zone Z_k, that have joined RAT x with the i^{th} ($d_{min}(x)$, $d_{max}(x)$) couple. In fact, mobiles are first provided with their minimum guaranteed throughput given by $d_{min} \cdot g(k)$. Then, fair time scheduling is used to provide them with up to their maximum throughput given by $d_{max} \cdot g(k)$. Remaining resources may afterwards be equitably shared (*i.e.,* after receiving their maximum throughput, all mobiles have the same priority leading to fair time scheduling).

Furthermore, the blocking term reflects the penalty of rejecting future arrivals. $B(s,a)$ is thus proportional to the arrival rates in blocking states, and is expressed as follows:

$$B(s,a) = -\, b \cdot \sum_k \sum_c \Lambda(k,c)(1 - \sum_x p^x(k,c,a)) \tag{6.4.7}$$

where b is the cost per unit time inflicted on the network for blocking a new arrival.

6.4.5 Uniformization

In our work, we make use of the *Policy Iteration* algorithm to solve the SMDP problem (*i.e.,* to determine the action the network takes in each state). A stage of uniformization is thus required. The continuous-time Markov chain is transformed into its discrete equivalent.

Time is first discretized into intervals of constant duration τ, that is smaller than the expected sojourn time in any state: $0 \leq \tau < \tau(s,a)$, $\forall s \in \mathcal{S}$.

Transition probabilities are then modified as follows:

$$\begin{cases} \bar{p}(s,s',a) = p(s,s',a)\frac{\tau}{\tau(s,a)} & \text{for } s' \neq s \\ \bar{p}(s,s',a) = 1 - \sum_{s' \neq s} \bar{p}(s,s',a) & \text{Otherwise} \end{cases} \tag{6.4.8}$$

where $\bar{p}(s,s',a)$ represents the probability that the network moves from state s to s' within τ, when action a is adopted.

Moreover, the reward is also modified as follows: $\bar{r}(s,a) = r(s,a)\tau$, where $\bar{r}(s,a)$ is the reward earned for a time τ.

6.4.6 Policy Iteration Algorithm

A policy π is a mapping from $\mathcal{S}$ to $\mathcal{A}$. $\pi(s)$ represents the action to take in state s. Let $H_\pi(s) = s, s_1, s_2, ..., s_n, ...$ be a trajectory of the Markov chain, when policy π is adopted. The long-term discounted reward $dr(H_\pi(s))$ of state s is the discounted sum of the rewards earned on that trajectory (that starts from s), and is expressed as follows:

$$\bar{r}(s,\pi(s)) + \psi\bar{r}(s_1,\pi(s_1)) + ... + \psi^n\bar{r}(s_n,\pi(s_n)) + ...$$

where ψ is the discounting factor ($0 < \psi < 1$). In our work, we set the value function of state s, denoted by $V_\pi(s)$, as the expected value of $dr(H_\pi(s))$ over all possible trajectories.

Our goal is to find an optimal policy π_{opt}, that maximizes the expected long-term discounted reward of each state:

$$V_{\pi_{opt}}(s) \geq V_\pi(s), \qquad \forall s, \pi$$

We therefore use the following *Policy Iteration* algorithm:

- Step 0 (Initialization): We choose an arbitrary policy π.
- Step 1 (Value Determination): Given the current policy π, we solve the following system of linear equations to calculate the discounted value function V_π of all states:

$$V_\pi(s) = \bar{r}(s,\pi(s)) + \psi \sum_{s' \in \mathcal{S}} \bar{p}(s,s',\pi(s))V_\pi(s')$$

- Step 2 (Policy Improvement): When any improvement is possible, we update the current policy π. For each s $\in \mathcal{S}$, we find:

$$\hat{\pi}(s) = \arg \max_{a \in \mathcal{A}(s)} \left\{ \bar{r}(s,a) + \psi \sum_{s' \in \mathcal{S}} \bar{p}(s,s',a)V_\pi(s') \right\}$$

- Step 3 (Convergence test): If $\hat{\pi} = \pi$, the algorithm is stopped with $\pi_{opt} = \pi$. Otherwise, we set π to $\hat{\pi}$, and go to step 1.

6.5 Reinforcement Learning

In the previous section, knowing $r(s, a)$ and $p(s, s', a)$, an optimal policy π_{opt} is solved through the *Policy Iteration* algorithm. The transition probability function $p(s, s', a)$ depends on user arrival and departure rates, needs, preferences, and decision-making algorithms. However, when $p(s, s', a)$ may not be easily obtained, reinforcement learning (RL) turns out to be a good fit to derive network information. The network does not estimate user behavior, but rather learns what action to take by trial-and-error. Among the different existing RL algorithms, we select Q-learning [WD92] for its simplicity. Although originally used to solve Markov decision processes, Q-learning may be applied with slight modifications to semi-Markov decision processes [Rya02].

6.5.1 SMDP Q-learning Algorithm

The network interacts with its environment over a sequence of discrete time-steps $(t, t+1, t+2, ...)$, trying to learn what QoS parameters to signal. These time-steps refer to time intervals of fixed duration τ. The *quality* function of state-action pair $(s, \pi(s))$, denoted by $Q_\pi(s, \pi(s))$, is defined as the expected long-term discounted reward of state s, using policy π. Our aim is to find an optimal policy π_{opt}, that maximizes the *quality* function of each state s, also referred to as its Q-value:

$$\pi_{opt}(s) = \arg \max_{a \in \mathcal{A}(s)} Q_\pi(s, a), \qquad \forall s, \pi$$

Without knowledge of $p(s, s', a)$, the network, also referred to as the agent, iteratively learns optimal Q-values. At discrete time-steps, when the network state has changed, the network action terminates. QoS parameters to be signaled may have to vary. Unlike in Markov decision processes, where all actions are assumed to take constant time to complete, actions in our work can span several time-steps. They are said to be temporally-abstract. At time-step t, when state-action pair (s, a) is visited (*i.e.,* when the network in state s selects and performs action a), the network earns reward R, and ends in state s' at $t + k$. The Q-value of state-action pair (s, a) is then updated as follows:

$$Q(s, a) \leftarrow Q(s, a) + \rho\Big(R + \psi^k \max_{a' \in \mathcal{A}} \{Q(s', a')\} - Q(s, a)\Big) \tag{6.5.1}$$

where ρ is the learning rate $(0 < \rho < 1)$, that determines to what extent the learned Q-value will override the old one. When $\rho = 0$, the network does not learn. When $\rho = 1$, the network considers only the most recent Q-value. R is the discounted accumulation of all single-step rewards r_τ, received while executing action a for a time τ, and is given by:

$$R = \sum_{i=0}^{k-1} \psi^i \; r_\tau$$

Moreover, it has been proved that, while the number of visits of each state-action-pair is sufficiently large, and ρ is reduced to zero over time, $Q(s, a)$ is guaranteed to converge to $Q_{\pi_{opt}}(s, a)$ [WD92].

6.5.2 Exploration and Exploitation

At decision epochs, the network decides, randomly or based on previously learned Q-values, what QoS parameters to signal. To receive high reward, the network may prefer actions it has tried in the past and found effective. This is known as the exploitation mode. Yet, to discover effective ones, the network needs to try actions it has not selected before. It may then randomly select one of the possible actions, aiming to enhance its future decisions. This is known as the exploration mode. Since Q-learning is an online iterative learning algorithm, exploration and exploitation should be simultaneously performed. The agent must discover a variety of actions, and progressively favor effective ones. However, to estimate reliable Q-values, actions need to be sufficiently tested.

In our work, we adopt an ϵ-greedy exploration-exploitation policy. At decision epochs, the network in state s explores with probability $\epsilon(s)$, and exploits stored Q-values with probability $1 - \epsilon(s)$. To enhance long-term network performance, exploring is never stopped, but rather reduced over time. We define $\beta(s, a)$ to be the number of visits of state-action pair (s, a) up to current time-step, and choose $\epsilon(s)$ to be as follows:

$$\epsilon(s) = \frac{1}{ln(\sum_{a \in \mathcal{A}} \beta(s, a) + 3)} \tag{6.5.2}$$

$\epsilon(s)$ then belongs to $[0, 1]$, and has a logarithmic decay. Furthermore, for $Q(s, a)$ to converge to optimal Q-values, we set ρ to be a state-action pair varying over time:

$$\rho(s, a) = \frac{1}{\sqrt{\beta(s, a) + 3}}$$

Algorithm 1 describes our SMDP Q-learning algorithm for deriving network information. We summarize below the main steps. Q-values are first set to zero. The network state is randomly initialized. Once in state s, depending on $\epsilon(s)$, exploration or exploitation is executed. In exploration mode, the network randomly selects and performs action a. However, in exploitation mode, it opts for the action with the maximum Q-value:

$a = \max_a Q(s,a)$. After, at each time-step, the network state is observed. While the network is in state s, action a is maintained, and the discounted accumulation of single-step rewards R is updated. Yet, if it is in state s' (*i.e.,* the network state has changed), action a is terminated, and $Q(s,a)$ is updated according to equation 6.5.1. This is repeated until the end of the learning period.

```
Initialize
  • Q-values: Q(s,a) ← 0, ∀s ∈ S and a ∈ A
  • Number of state-action visits: β(s,a) ← 0, ∀s ∈ S and a ∈ A
  • Time-step: t ← 0
repeat
    Observe state s
    if exploration then
        choose action a at random
    else
        choose a = max_a Q(s,a)
    end
    β(s,a) ← β(s,a) + 1
    Update ε(s) according to equation 6.5.2
    R ← 0
    k ← 0
    while the network is in state s do
        Perform a
        Wait for a fixed duration τ
        Observe reward r_τ
        R ← R + ψ^k r_τ
        k ← k + 1
    end
    Observe state s'
    Update Q(s,a) according to equation 6.5.1
    s ← s'
    t ← t + k
until End of the learning period;
```

Algorithm 1: SMDP Q-learning

6.6 Performance Results

For illustration, we consider a heterogeneous wireless network composed of mobile WiMAX and LTE, respectively designated by W and L. For simplicity, users are of two types: those with good radio conditions (*i.e.,* cell-center users) and those with bad radio conditions (*i.e.,* cell-edge users). Their peak rates, when connected alone to mobile WiMAX and LTE cells, are depicted in Table 6.1. Further, class c arrivals are assumed to be uniformly distributed over the two zones, and to follow a Poisson process of rate $\Lambda_c = \Lambda$ (*i.e.*, $\Lambda(k, c) = \Lambda/N_Z$, $\forall k, c$).

RAT	**k = 1**	**k = 2**
Mobile WiMAX (3 MHz)	9.9 Mb/s	4.4 Mb/s
LTE (5 MHz)	16.6 Mb/s	7.4 Mb/s

Table 6.1: Peak rates in Mobile WiMAX and LTE

Moreover, for streaming sessions, we suppose that $R_{av} = 1$ Mb/s, $R_{max} = 1.5$ Mb/s, and $1/\mu_1 = 45$ s. For elastic sessions, we consider that $L = 5$ Mbytes, and R_C is fixed to either 1.25 or 0.75 Mb/s, depending on the QoS-maximizing or cost-minimizing preferences of mobile users. For network information, we assume that $cost(W) = 4$, $cost(L) = 6$, $N_I^W = N_I^L = 3$, $I^W = \{(0,0), (0.5,1), (1,1.5)\}$ Mb/s, and $I^L = \{(0,0), (0.75,1.25), (1.5,2)\}$ Mb/s.

The probabilities $p^x(k, c, a)$ are calculated according to the satisfaction-based multi-criteria decision-making method, we have introduced in Chapter 3. They mainly depend on user preferences, traffic class and throughput demand. Note that half of the users are ready to pay for better performances.

For comparison purposes, we also investigate the staircase tuning policy. We recall that load factors are defined as the ratios of the number of guaranteed allocated RUs to the total number of RUs. The highest QoS parameters are first signaled. Next, when a RAT load factor exceeds S_1 threshold, QoS parameters are reduced following a step function (cf. Fig. 6.1). However, if S_2 is reached, QoS incentives are no longer provided. QoS parameters to signal in RAT x, depending on the load factor ϕ^x, are reported in Table 6.2.

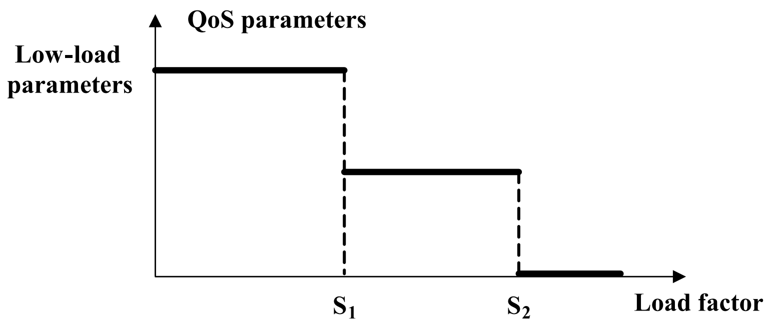

Figure 6.1: QoS parameters reduction using the Staircase policy

Before we discuss performance results, we remind in Table 6.3 some notations, useful in what follows.

QoS parameters	$\phi^x < S_1$	$S_1 \leq \phi^x \leq S_2$	$\phi^x > S_2$
$d_{min}(W)$	1 Mb/s	0.5 Mb/s	0
$d_{max}(W)$	1.5 Mb/s	1 Mb/s	0
$d_{min}(L)$	1.5 Mb/s	0.75 Mb/s	0
$d_{max}(L)$	2 Mb/s	1.25 Mb/s	0

Table 6.2: QoS parameters depending on the load factor ϕ^x

Parameters	Notation
Tuning thresholds of the staircase policy	S_1, S_2
Discount factor	ψ
Cell arrival rate	Λ
Blocking cost	b
Blocking term (penalty term)	B
Duration of learning periods	T
Duration of time-steps	τ

Table 6.3: Summary of notations

6.6.1 Staircase Policy

Using the staircase policy, we study the impact of S_1 and S_2 thresholds on network performance. Fig. 6.2 and 6.3 respectively show the average network throughput and the blocking probability, as a function of the cell arrival rate Λ, for different threshold values.

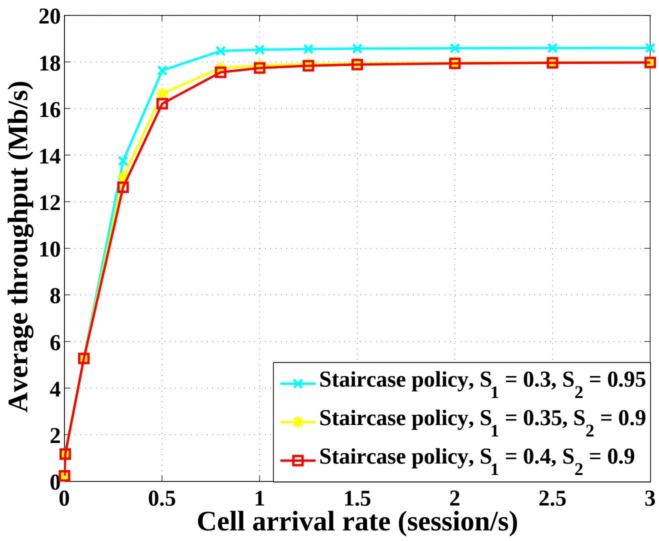

Figure 6.2: Staircase policies: Network throughput

For fixed S_1, the higher S_2 the more mobiles are admitted. Yet, higher S_2 thresholds limit user throughputs to their guaranteed ones. Besides, for fixed S_2, the lower S_1 the less mobiles benefit from the largest QoS guarantees, but much more are admitted with reduced

QoS parameters. Therefore, the average number of simultaneous sessions increases.

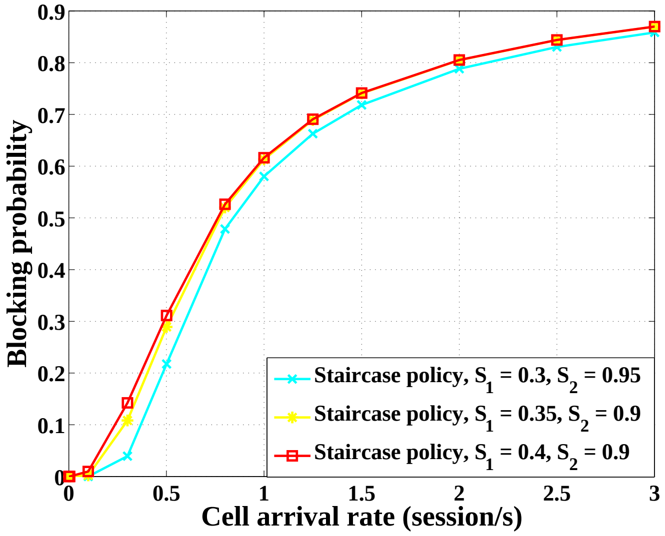

Figure 6.3: Staircase policies: Blocking probability

Obviously, the more mobiles are admitted for a fixed cell arrival rate (*i.e.,* the lower S_1 or the higher S_2), the lower the blocking probability. Also, the network total throughput augments. Typically, streaming sessions have limited throughput demands, and hence the more mobiles are admitted the larger the network throughput will potentially be.

6.6.2 Optimal Policy

The optimal policy, solved through the *Policy Iteration* algorithm, and the staircase policy are compared. Using the optimal policy, we study the impact of the blocking cost b, and the discount factor ψ on network performance.

6.6.2.1 Impact of the blocking cost

We start by inspecting the impact of the blocking cost b on network performance. So as to enlarge the number of states involved in the value function, the discount factor ψ is fixed at 0.99.

Fig. 6.4 illustrates the average reward as a function of the cell arrival rate Λ, for different blocking costs. When b is null, the reward function is reduced to the network utility representing the sum of user throughputs. Otherwise, it also includes a penalty term, that is proportional to the blocking cost b and to the cell arrival rate.

At low arrival rate, no blocking occurs leading to similar rewards regardless of b. The reward function, reduced to the network total throughput, then increases with the cell arrival rate. Yet, as the latter increases further, or equivalently, when the average number of simultaneous sessions augments, network resources are always nearly exhausted, and not enough are left to cope with future arrivals. Therefore, the blocking probability (*i.e.,* the long-term fraction of time spent in blocking states) also increases. Moreover, and since the penalty term is proportional to the cell arrival rate, the reward function received by the network whilst in a blocking state is as reduced as the arrival rate is increased. For all these reasons, the average reward decreases more when the cell arrival rate increases, except for b equals zero. In fact, when b is null, the average reward stagnates at high arrival rate. It represents the long-term sum of user throughputs. Otherwise, the average reward obviously decreases with increasing blocking costs. We further note that the optimal policy always outperforms the staircase one. However, when S_1 and S_2 are respectively set to 0.3 and 0.95, the staircase policy provides higher network reward in comparison with the case when $S_1 = 0.35$ and $S_2 = 0.85$, denoted as Staircase policy (2).

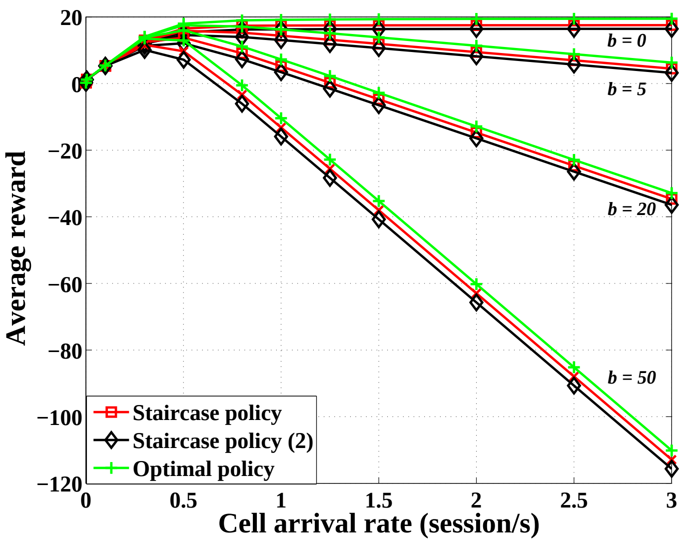

Figure 6.4: Impact of b on network reward

Moreoever, the higher b the more the network avoids blocking actions, even if at the expense of the network utility. Also, the lower b, the more the network tries to maximize its total throughput, even if leading to more blocking states. We, respectively, depict in figures 6.5 and 6.6 the network total throughput and the percentage in number of blocking states, as a function of the cell arrival rate. The optimal policy is illustrated for different values of b. Particularly, when b is zero, the network total throughput, but also the percentage of blocking states, are maximized. Therefore, the blocking cost b may be tuned to control optimization objectives. Further, when $S_1 = 0.3$ and $S_2 = 0.95$, the staircase policy achieves a higher throughput in comparison with when S_1 and S_2 are respectively set to 0.35 and 0.85. As a matter of fact, when these thresholds are carefully chosen, the staircase policy provides quite similar performances as the optimal one ($b = 50$). They both effectively avoid blocking actions and guide user decisions. In the remaining, we only consider the case where $S_1 = 0.3$ and $S_2 = 0.95$.

It is worth noting that for a given b, when the cell arrival rate is different, the state dynamics and penalty terms are also different. This may lead to dissimilar optimal policies. Thus, and as shown in Fig. 6.6, the percentage in number of blocking states first increases with the cell arrival rate. Then, when the latter increases further, for b different from zero, this percentage decreases as the penalty term becomes relatively very significant.

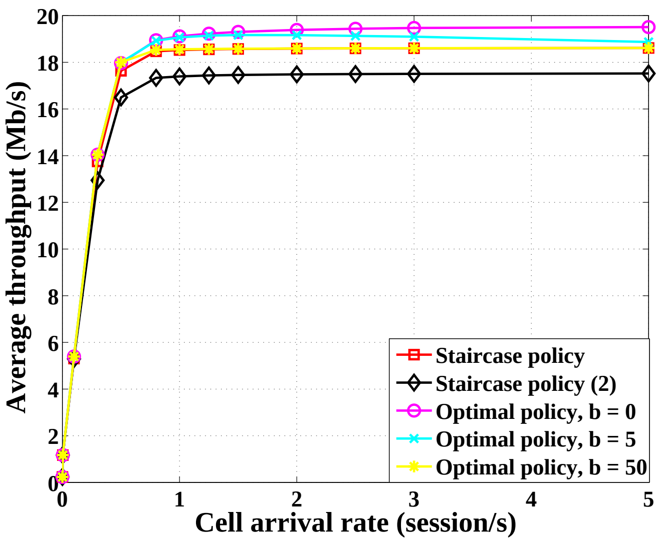

Figure 6.5: Impact of b on network throughput

Moreover, the blocking probability $\mathbb{P}_b$ depends not only on the number of blocking states, but mostly on the stationary distribution achieved by the different policies (*i.e.,* on the long-term fraction of time spent in the different states). In the following, to efficiently analyze the impact of the blocking cost on $\mathbb{P}_b$, we separately consider streaming and elastic sessions.

The service time of elastic sessions depends both on their size assumed to be exponentially distributed with a mean of 5 Mbytes, and on their perceived throughputs. As shown before, the lower b, the higher the network total throughput leading to lower average service times. When the optimal policy is adopted (*i.e.,* the actions are fixed to the optimal ones), the SMDP may be reduced to a Markov chain, where departure rates increase with decreasing blocking costs. As a result, for a given cell arrival rate, the lower b, the lower the long-term number of simultaneous sessions. This also means that, although the lower b the higher the percentage of blocking states, the long-term fraction of time spent in these states is reduced as b is low. Accordingly, the lower b, the lower $\mathbb{P}_b$ for elastic sessions as illustrated in Fig. 6.7.

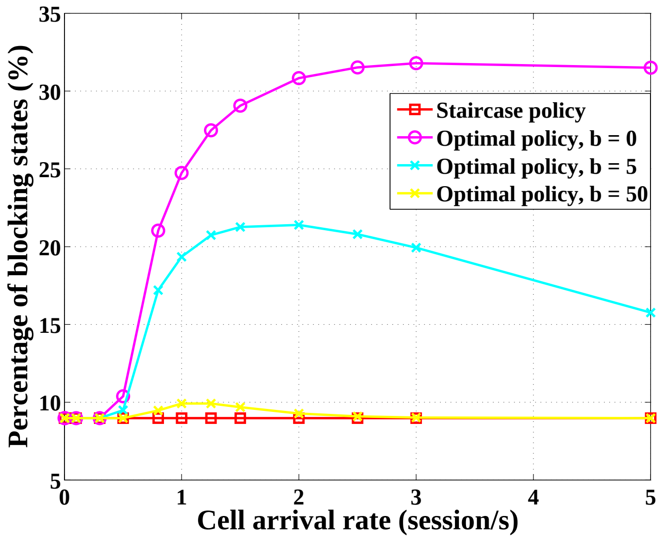

Figure 6.6: Impact of b on the percentage of blocking states

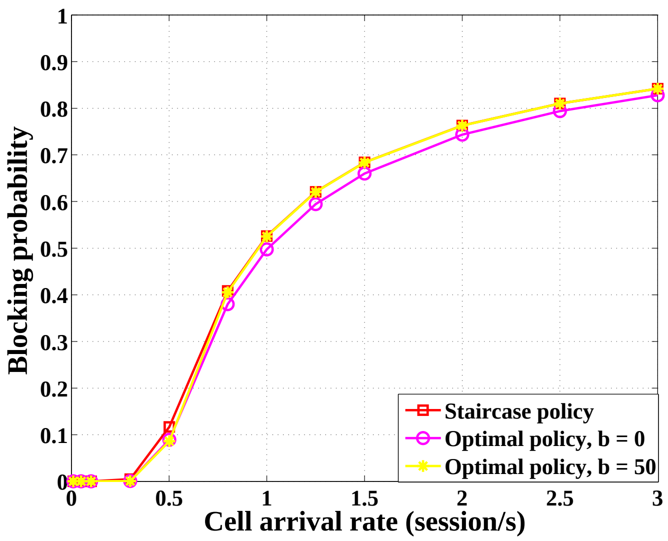

Figure 6.7: Impact of b on blocking probability for elastic sessions

Nevertheless, the service time of streaming sessions exclusively depends on their duration, considered to be exponentially distributed with a mean of 45 s. Thereby, maximizing the network total throughput will not reduce their average service times. Consequently, as the number of blocking states increases with decreasing b, the blocking probability for streaming sessions also increases (cf. Fig. 6.8). The long-term fraction of time spent in blocking states will actually be higher. Here again, for both traffic classes, the performance of the staircase policy, with carefully chosen S_1 and S_2 thresholds, is comparable to the optimal one ($b = 50$).

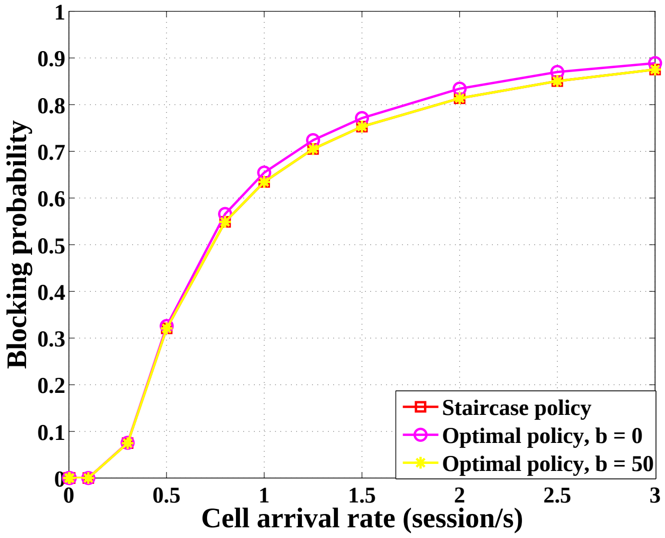

Figure 6.8: Impact of b on blocking probability for streaming sessions

6.6.2.2 Impact of the discount factor

In this paragraph, we investigate the impact of the discount factor ψ on network performance. When the blocking cost b is set to zero, the network reward is reduced to the sum of user throughputs.

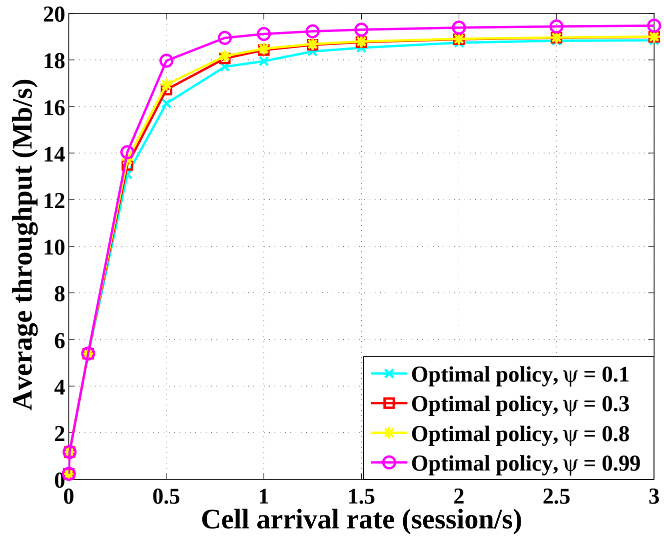

Figure 6.9: Impact of ψ on network throughput

Fig. 6.9 and 6.10 respectively illustrate the network total throughput and the blocking probability as a function of the cell arrival rate, for different ψ values. Recall that the higher ψ, the larger the number of states involved in the value function. Also, next states contribute more to the expected long-term network reward as ψ gets higher. The discount factor ψ can thus be tuned to control the optimization scope. Typically, higher ψ values imply more long-run optimization, leading to higher throughput and lower blocking probability.

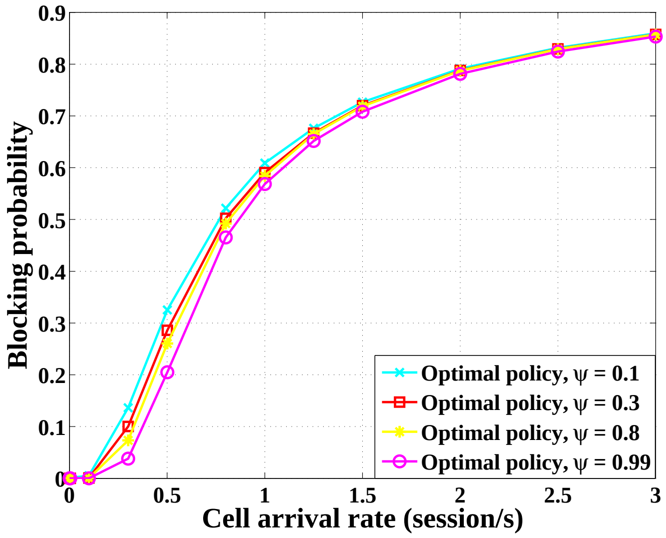

Figure 6.10: Impact of ψ on blocking probability

Further, we note that the network total throughput at low arrival rate and the blocking probability at high arrival rate are obviously quite similar, regardless of the discount factor.

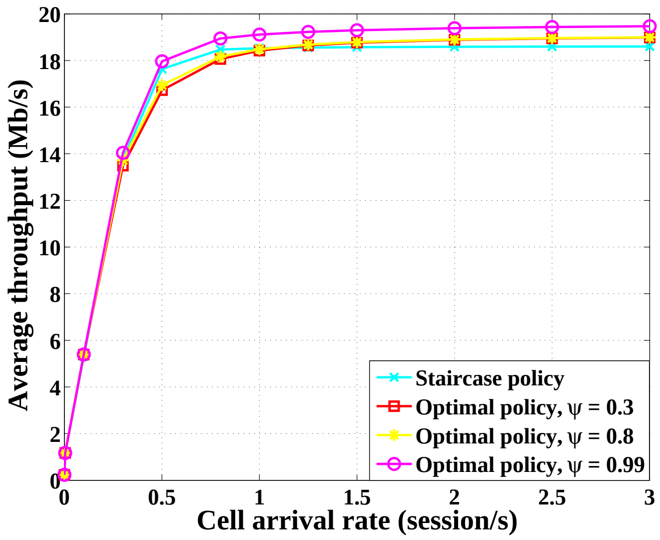

Figure 6.11: Optimal vs. staircase policies: network throughput

Fig. 6.11 and 6.12 compare the optimal policy with the staircase one. On the one hand, we notice that, at low arrival rate (typically below 1), the staircase policy outperforms the optimal one with $\psi = 0.3$ and $\psi = 0.8$. This means that the intuitive and low-complexity staircase policy efficiently guides user decision at low arrival rate. Yet, to maximize network performance, the number of states that are involved in the value function should be large enough. This can be seen with $\psi = 0.99$. On the other hand, when the cell arrival rate increases, taking into account next states becomes more relevant. In fact, when the network is expected to approach its saturation, deriving QoS parameters considering future arrivals enhances long-term network performance. Also, reducing QoS parameters in all serving RATs, following the staircase policy, proves to provide close performance to the optimal one (cf. Fig. 6.11 and 6.12).

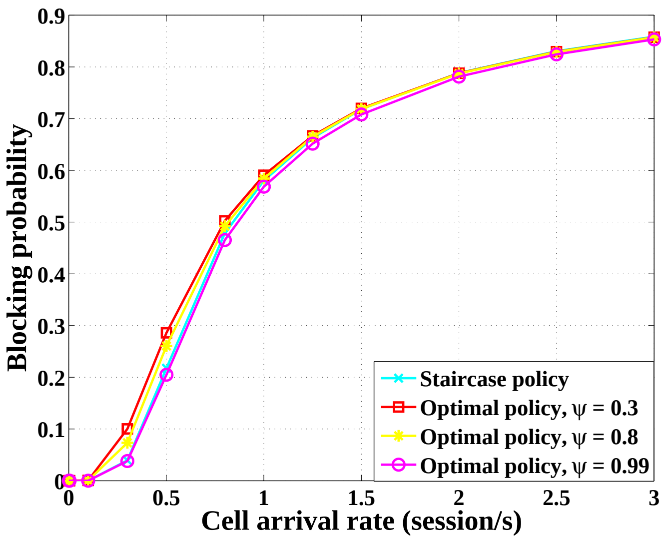

Figure 6.12: Optimal vs. staircase policies: blocking probability

6.6.3 Learning-based Policy

In what follows, the learning-based ($\psi = 0.99$) policy, the optimal ($b = 0, \psi = 0.99$) policy and the staircase policy are compared. Using the Q-learning algorithm, the agent interacts with its environment over a sequence of $T = 100000$ and $T = 250000$ time-steps, of fixed duration $\tau = 0.5$ s. Performance metrics are then averaged over 20 learning periods.

Fig. 6.13 and 6.14 respectively show the network total throughput and the blocking probability, as a function of the cell arrival rate. The optimal solution, solved using the *Policy Iteration* algorithm, provides an upper bound on the network total throughput. It also brings the lowest blocking probability, and consequently the best network performance. However, the optimal policy suffers from high computational complexity. For a fixed discount factor, the *Policy Iteration* algorithm is shown to run in at most $\frac{N_s^2(N_a-1)}{1-\psi} \cdot \log(\frac{N_s^2}{1-\psi})$

iterations, where N_s is the number of states, N_a the number of actions, and ψ the fixed discount factor [Ye11].

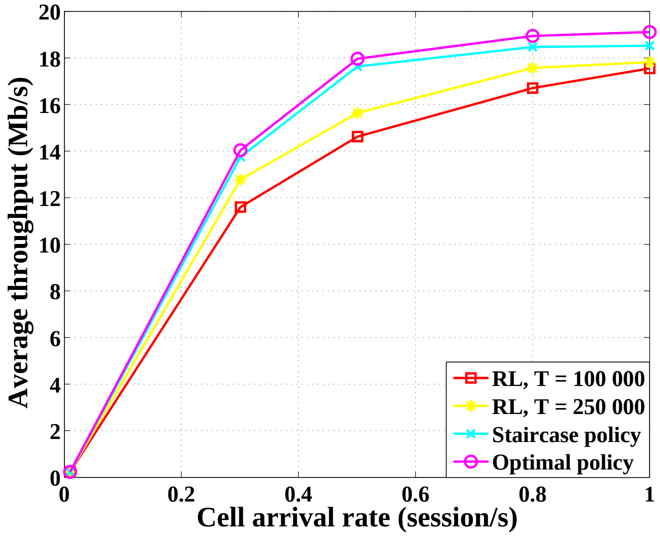

Figure 6.13: RL-based vs. optimal vs. staircase policies: network throughput

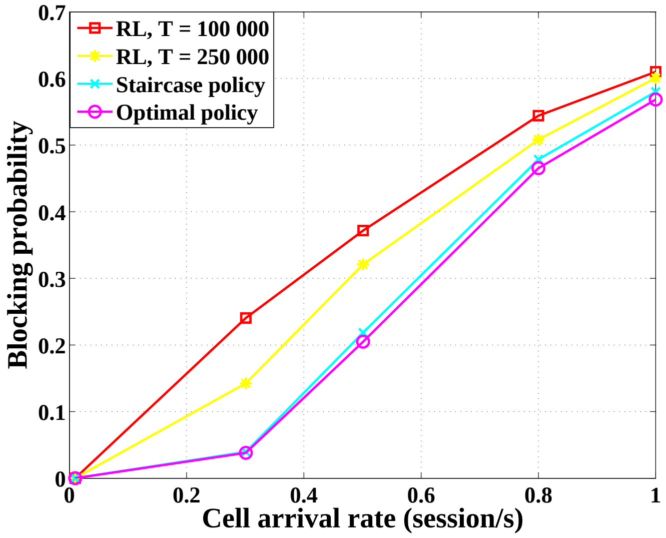

Figure 6.14: RL-based vs. optimal vs. staircase policies: blocking probability

Moreover, and as discussed before, the staircase policy provides very close performance to the optimal one despite its low complexity. Yet, a practical challenge is to efficiently set S_1 and S_2 values. When our heuristic requires no knowledge on network parameters, its performance strongly depends on the choice of the tuning thresholds.

Furthermore, unlike the optimal and the heuristic solutions, the learning-based one needs no parameterization. Theoretically, after an infinite learning period, our Q-learning algorithm converges to the optimal solution. In our work, we stop learning after a realistic duration of $T = 100000$ and $T = 250000$ time-steps. Better performances are obviously observed when $T = 250000$, in comparison with when $T = 100000$. Yet, when learning periods are voluntary limited, both the optimal and the heuristic policies outperform the learning-based ones.

6.7 Conclusion

In Chapter 3, we proposed a hybrid RAT selection approach, aiming to jointly enhance network performance and user experience. As a matter of fact, the network provides information for the mobiles to make final decisions, regarding selection of their most appropriate RAT. In this chapter, deriving network information was formulated as a semi-Markov decision process, and optimal policies were solved through the *Policy Iteration* algorithm. We showed how the blocking cost b and the discount factor ψ may be tuned to control optimization objectives, aligning with user needs and preferences. Note that user mobility can be integrated into our SMDP model. When user dwell time in zones is exponentially distributed, transitions between network states happen with an additional rate, due to user mobility.

Furthermore, we have introduced a RL-based algorithm to determine what to signal to mobiles. The performances of optimal, learning-based and staircase policies were analyzed. When S_1 and S_2 thresholds are pertinently chosen, our low-complexity heuristic provides close performance to the optimal solution. Moreover, although lower performances are observed, our learning-based algorithm has the crucial advantage of requiring no prior parameterization.

Chapter 7

General Conclusion

This chapter concludes this thesis report. We summarize the main contributions, and give the future research directions that stem from this work.

7.1 Summary of Contributions

This thesis has investigated Radio Access Technology (RAT) selection. Our work is placed in the context of heterogeneous wireless networks, where various RATs covering the same region are being integrated and jointly managed. One of the main motivations behind heterogeneous wireless networks is to cope, in a cost-efficient way, with the rapid growth of mobile broadband traffic. Another motivation is to deliver high user experience, as the different serving RATs complement each other in their characteristics.

RAT selection, devoted to decide to what RAT mobiles connect, is a key common radio resource management functionality to improve network performance and user experience. When intelligence is pushed to the network edge, mobiles make autonomous decisions regarding selection of their most appropriate RAT. They aim to selfishly maximize their utility. Yet, because mobiles have no information on network load conditions, their decisions may cause performance degradation. Moreover, delegating decisions to the network optimizes overall performance, but at the cost of increased network complexity, signaling and processing load.

Our challenge is however to design a RAT selection approach, that jointly enhances network performance and user experience, while signaling and processing burden remains reduced.

In this thesis, combining benefits from both network-centric and mobile-terminal-centric approaches, we proposed a new hybrid decision method. Using the logical communication channel (*i.e.,* radio enabler) proposed by the IEEE 1900.4 standard, the network provides information for the mobiles to make robust RAT selections. More precisely, mobile users select their RAT depending on their needs and preferences, as well as on the monetary cost and QoS parameters signaled by the network. By appropriately tuning network information, user decisions are globally expected to meet operator objectives, avoiding undesirable network states. Our approach thus enables self-optimization, a key feature of self-organizing networks.

When several base stations are available, decisions are traditionally based on received-signal-strength measurements. In our work, so as to maximize user experience, we introduced a satisfaction-based Multi-Criteria Decision-Making (MCDM) method. Mobiles compute a utility function for each of the available RATs, and select the one with the highest score. This utility however depends on user radio conditions, needs and preferences (*i.e.,* traffic class, throughput demand, QoS-maximizing or cost-minimizing preferences), as well as on the cost and QoS information sent by the network. Utility functions for inelastic, streaming, and elastic traffic classes were detailed. In comparison with existing MCDM solutions, namely SAW and TOPSIS, our algorithm meets user needs (*e.g.,* traffic class, throughput demand, cost tolerance), avoiding oversized and undersized decisions.

Furthermore, we investigated network information. While cost parameters are maintained fixed, QoS parameters are dynamically tuned trying to globally control user decisions. We presented two heuristic methods, namely the staircase and the slope tuning policies, to derive QoS information as a function of network load conditions. They follow a linear decreasing (slope) or a staircase function, and proved to efficiently exploit radio resources while mobiles maximize their own utility. As QoS parameters vary with load conditions, mobiles are effectively distributed over the different serving RATs, leading to better performance, higher user satisfaction, and larger operator gain.

Also, we studied the impact of providing mobiles with differentiated services and throughput guarantees. When operators propose Premium, Regular, and Economy service classes, that differ in their cost and QoS parameters, better network performance, higher user satisfaction, and larger operator gain can be observed. Therefore, while heterogeneous RATs are integrated, it is always beneficial if all do not offer the same QoS and cost incentives, giving mobiles a variety of possible choices. Moreover, when mobiles are provided with minimum throughput guarantees, regardless of future network load conditions, real-time sessions see their performance enhanced.

Further, we compared our hybrid decision approach with different network-centric, mobile-terminal-centric, and hybrid methods. Peak rate maximization, Average rate maximization, Satisfaction-based using peak rate, Satisfaction-based using average rate, and exhaustive search methods were considered. We highlighted the effectiveness of our solution in enhancing resource utilization and user experience. As a matter of fact, compared with mobile-terminal-centric and hybrid methods, our decision approach maximizes the network utility, defined as the network total throughput, and the average user satisfaction. Also, compared with the optimal exhaustive search method, our solution provides significantly higher user satisfaction.

We assessed as well the gain from masking network load conditions, and only signaling cost and some QoS parameters. Our hybrid approach outperforms non-realistic methods, where mobiles have a perfect knowledge of network load conditions. So, when operator objectives are implicitly involved within signaled QoS parameters, radio resources are better utilized, and user satisfaction is maximized.

Moreover, we focused on optimizing network information. Deriving QoS parameters was formulated as a semi-Markov decision process, and optimal policies were solved through the *Policy Iteration* algorithm. The aim is to dynamically optimize the long-term network reward, while mobiles maximize their own utility. We showed how the blocking cost b and the discount factor ψ may be tuned to control optimization objectives, aligning with user needs and preferences. User mobility can be further integrated into our SMDP model. Also, and since network parameters may not be easily obtained, a reinforcement learning approach was introduced to derive what to signal to mobiles. The performances of optimal, learning-based, and heuristic policies were analyzed. When tuning thresholds are pertinently chosen, our low-complexity heuristic provides close performance to the optimal one. Moreover, although learning-based tuning achieves lower performance, it does not need to know network parameters.

7.2 Future Directions

To optimize long-term network performance, QoS information needs to depend not only on present load conditions, but also on expected future demands. Thus, in our thesis, deriving QoS parameters was formulated as a semi-Markov decision process. In state s, d_{min} and d_{max} are decided in a way to dynamically maximize the long-term network reward, aligning with user needs and preferences.

Nevertheless, when the number of zones, traffic classes, and possible QoS parameters increase, the number of states becomes huge. This leads to a heavy computational load to find optimal policies. It would then be interesting to investigate reducing techniques to solve large MDP problems.

Furthermore, as network parameters can not be easily obtained, a reinforcement learning approach was also introduced to derive QoS parameters. When the number of visits of each state-action-pair is infinite, the network is theoretically guaranteed to reach an optimal policy. However, practically as the state-action pairs are huge in number, they are partially explored, leading to a satisfying policy. To handle this limitation, Q-learning needs to be implemented using a neural network. Instead of storing Q-values, neural networks approximate them, and can interpolate those of state-action pairs that have not been visited.

Moreover, our hybrid decision approach fits within the larger framework of self-organizing networks. Under overload conditions, QoS and cost parameters are tuned in a way to enhance resource utilization. It would be interesting to go further in the self-optimization mechanisms. We can investigate parameter tuning under interference conditions.

List of Publications

[HILK14] Melhem El Helou, Marc Ibrahim, Samer Lahoud, and Kinda Khawam, *Optimizing Network Information for Radio Access Technology Selection*, Proc. IEEE Symposium on Computers and Communications (ISCC), June 2014

[HLIK13b] Melhem El Helou, Samer Lahoud, Marc Ibrahim, and Kinda Khawam, *Satisfaction-based Radio Access Technology Selection in Heterogeneous Wireless Networks*, Proc. IEEE IFIP Wireless Days Conference (WD), November 2013

[HILK13] Melhem El Helou, Marc Ibrahim, Samer Lahoud, and Kinda Khawam, *Radio Access Selection Approaches in Heterogeneous Wireless Networks*, Proc. IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), October 2013

[HLIK13a] Melhem El Helou, Samer Lahoud, Marc Ibrahim, and Kinda Khawam, *A Hybrid Approach for Radio Access Technology Selection in Heterogeneous Wireless Networks*, Proc. European Wireless Conference (EW), April 2013

Bibliography

[3GP01] 3GPP TR 25.881, *Improvement of RRM across RNS and RNS/BSS*, December 2001.

[3GP08] 3GPP TR 23.882, *3GPP System Architecture Evolution: Report on Technical Options and Conclusions (Release 8)*, September 2008.

[3GP10] 3GPP TS 32.521, *Telecommunication Management; Self-Organizing Networks (SON) Policy Network Resource Model (NRM) Integration Reference Point (IRP); Requirements*, March 2010.

[Ahm10] Sassan Ahmadi, *WiMAX Network Architecture*, Mobile WiMAX, Academic Press, 2010, pp. 33 – 60.

[AKHWC13] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and Mung Chiang, *RAT Selection Games in HetNets*, Proc. IEEE Conference on Computer Communications (INFOCOM), April 2013.

[Bea08] Ronald Beaubrun, *Integration of Heterogeneous Wireless Access Networks*, Heterogeneous Wireless Access Networks: Architectures and Protocols (Ekram Hossain, ed.), Springer, 2008, pp. 1 – 18.

[BL07] Farooq Bari and Victor C.M. Leung, *Automated Network Selection in a Heterogeneous Wireless Network Environment*, IEEE Networks **21** (2007), no. 1, 34 – 40.

[Cic13] Claudio Cicconetti, *5G Radio Network Architecture*, March 2013.

[Cis14] Cisco, *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018*, February 2014.

[CKG08a] M. Coupechoux, J.-M. Kelif, and P. Godlewski, *Network controlled joint radio resource management for heterogeneous networks*, Proc. IEEE Vehicular Technology Conference (VTC Spring), May 2008.

[CKG08b] M. Coupechoux, J.-M. Kelif, and P. Godlewski, *SMDP Approach for JRRM Analysis in Heterogeneous Networks*, Proc. European Wireless Conference (EW), June 2008.

[CM12] Ioannis Chamodrakas and Drakoulis Martakos, *A Utility-Based Fuzzy TOPSIS Method for Energy Efficient Network Selection in Heterogeneous Wireless Networks*, Applied Soft Computing **12** (2012), no. 7, 1929 – 1938.

[CSH+01] P. M L Chan, R.E. Sheriff, Y.F. Hu, P. Conforto, and C. Tocci, *Mobility Management Incorporating Fuzzy Logic for a Heterogeneous IP Environment*, IEEE Communications Magazine **39** (2001), no. 12, 42 – 51.

[CTG09] P. Coucheney, C. Touati, and B. Gaujal, *Fair and Efficient User-Network Association Algorithm for Multi-Technology Wireless Networks*, Proc. IEEE Conference on Computer Communications (INFOCOM), April 2009.

[DLOBS+08] A De La Oliva, A Banchs, I Soto, T. Melia, and A Vidal, *An overview of ieee 802.21: Media-independent handover services*, IEEE Wireless Communications **15** (2008), no. 4, 96 – 103.

[DO12a] C. Dhahri and T. Ohtsuki, *Learning-Based Cell Selection Method for Femtocell Networks*, Proc. IEEE Vehicular Technology Conference (VTC Spring), May 2012.

[DO12b] C. Dhahri and T. Ohtsuki, *Q-learning Cell Selection for Femtocell Networks: Single- and Multi-user Case*, Proc. IEEE Global Communications Conference (GLOBECOM), December 2012.

[Erc08] O. Ercetin, *Association Games in IEEE 802.11 Wireless Local Area Networks*, IEEE Transactions on Wireless Communications **7** (2008), no. 12, 5136 – 5143.

[ETS00] ETSI TR 121 910, *Multi-Mode User Equipment (UE) Issues - Categories, Principles and Procedures*, July 2000.

[ETS01] ETSI TR 101 957, *Requirements and Architectures for Interworking between HIPERLAN/2 and 3rd Generation Cellular Systems*, 2001.

[EV99] R. Edell and P. Varaiya, *Providing Internet Access: What We Learn from INDEX*, IEEE Network **13** (1999), no. 5, 18 – 25.

[FC11] O.E. Falowo and H.A. Chan, *RAT Selection for Multiple Calls in Heterogeneous Wireless Networks Using Modified TOPSIS Group Decision-Making Technique*, Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), September 2011.

[FT13] Mikael Fallgren and Bogdan Timus, *Scenarios, Requirements and KPIs for 5G Mobile and Wireless System*, May 2013, Deliverable D1.1: ICT-317669.

[GAM05] V. Gazis, N. Alonistioti, and L. Merakos, *Toward a Generic "Always Best Connected" Capability in Integrated WLAN/UMTS Cellular Mobile Networks (and Beyond)*, IEEE Wireless Communications **12** (2005), no. 3, 20 – 29.

[GAPRS06] Lorenza Giupponi, Ramon Agustí, Jordi Pérez-Romero, and Oriol Sallent, *A Framework for JRRM with Resource Reservation and Multiservice Provisioning in Heterogeneous Networks*, Mobile Networks and Applications **11** (2006), no. 6, 825 – 846.

[GAPRS08] L. Giupponi, R. Agusti, J. Perez-Romero, and O. Sallent, *A Novel Approach for Joint Radio Resource Management Based on Fuzzy Neural Methodology*, IEEE Transactions on Vehicular Technology **57** (2008), no. 3, 1789 – 1805.

[GAPRS09] L. Giupponi, R. Agusti, J. Perez-Romero, and O. Sallent, *Fuzzy Neural Control for Economic-Driven Radio Resource Management in Beyond 3G Networks*, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews **39** (2009), no. 2, 170 – 189.

[GB09] Cédric Gueguen and Sébastien Baey, *A Fair Opportunistic Access Scheme for Multiuser OFDM Wireless Networks*, EURASIP Journal on Wireless Communications and Networking (2009) (2009).

[GJ03] E. Gustafsson and A. Jonsson, *Always Best Connected*, IEEE Wireless Communications **10** (2003), no. 1, 49 – 55.

[GLEnSS12] J. Gozalvez, M. C. Lucas-Estañ, and J. Sanchez-Soriano, *Joint Radio Resource Management for Heterogeneous Wireless Systems*, Wireless Networks **18** (2012), no. 4, 443 – 455.

[GPRSA08] X. Gelabert, J. Perez-Romero, O. Sallent, and R. Agusti, *A Markovian Approach to Radio Access Technology Selection in Heterogeneous Multiaccess/Multiservice Wireless Networks*, IEEE Transactions on Mobile Computing **7** (2008), no. 10, 1257 – 1270.

[Har68] G. Hardin, *The Tragedy of the Commons*, Science Journal (1968).

[HBJG07] S. Horrich, S. Ben Jamaa, and P. Godlewski, *Adaptive Vertical Mobility Decision in Heterogeneous Networks*, Proc. International Conference on Wireless and Mobile Communications (ICWMC), March 2007.

[HILK13] Melhem El Helou, Marc Ibrahim, Samer Lahoud, and Kinda Khawam, *Radio Access Selection Approaches in Heterogeneous Wireless Networks*, Proc. IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), October 2013.

[HILK14] Melhem El Helou, Marc Ibrahim, Samer Lahoud, and Kinda Khawam, *Optimizing Network Information for Radio Access Technology Selection*, Proc. IEEE Symposium on Computers and Communications (ISCC), June 2014.

[HLIK13a] Melhem El Helou, Samer Lahoud, Marc Ibrahim, and Kinda Khawam, *A Hybrid Approach for Radio Access Technology Selection in Heterogeneous Wireless Networks*, Proc. European Wireless Conference (EW), April 2013.

[HLIK13b] Melhem El Helou, Samer Lahoud, Marc Ibrahim, and Kinda Khawam, *Satisfaction-based Radio Access Technology Selection in Heterogeneous Wireless Networks*, Proc. IEEE IFIP Wireless Days Conference (WD), November 2013.

[IKT09] M. Ibrahim, K. Khawam, and S. Tohme, *Network-Centric Joint Radio Resource Policy in Heterogeneous WiMAX-UMTS Networks for Streaming and Elastic traffic*, Proc. IEEE Wireless Communications and Networking Conference (WCNC), April 2009.

[IKT10] M. Ibrahim, K. Khawam, and S. Tohme, *Congestion Games for Distributed Radio Access Selection in Broadband Networks*, Proc. IEEE Global Communications Conference (GLOBECOM), December 2010.

[KAK06] Dinesh Kumar, Eitan Altman, and Jean-Marc Kelif, *User-Network Association in a WLAN-UMTS Hybrid Cell: Global & Individual Optimality*, Rapport de recherche RR-5961, INRIA, 2006.

[Kha06] Kinda Khawam, *The Modified Proportional Fair Scheduler*, Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), September 2006.

[KIC+11] K. Khawam, M. Ibrahim, J. Cohen, S. Lahoud, and S. Tohme, *Individual vs. Global Radio Resource Management in a Hybrid Broadband Network*, Proc. IEEE International Conference on Communications (ICC), June 2011.

[KKP08] Meriem Kassar, Brigitte Kervella, and Guy Pujolle, *An Overview of Vertical Handover Decision Strategies in Heterogeneous Wireless Networks*, Computer Communications **31** (2008), no. 10, 2607 – 2620.

[KM10] Kinda Khawam and Dana Marinca, *Size-based Proportional Fair Scheduling*, Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), September 2010.

[LEnGSS12] M. C. Lucas-Estañ, J. Gozalvez, and J. Sanchez-Soriano, *Integer Linear Programming Optimization of Joint RRM Policies for Heterogeneous Wireless Systems*, Computer Networks **56** (2012), no. 1, 112 – 126.

[LPMK05] G. Lampropoulos, N. Passas, L. Merakos, and A Kaloxylos, *Handover Management Architectures in Integrated WLAN/Cellular Networks*, IEEE Communications Surveys and Tutorials **7** (2005), no. 4, 30 – 44.

[LTE14] *Multi-cell Radio Resource Management: Centralized or Decentralized?*, Nokia Solutions and Networks, January 2014.

[MILK12] F. Moety, M. Ibrahim, S. Lahoud, and K. Khawam, *Distributed Heuristic Algorithms for RAT Selection in Wireless Heterogeneous Networks*, Proc. IEEE Wireless Communications and Networking Conference (WCNC), April 2012.

[MYLR04] Li Ma, F. Yu, V.C.M. Leung, and T. Randhawa, *A New Method to Support UMTS/WLAN Vertical Handover Using SCTP*, IEEE Wireless Communications Magazine **11** (2004), no. 4, 44 – 51.

[NH09] D. Niyato and E. Hossain, *Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach*, IEEE Transactions on Vehicular Technology **58** (2009), no. 4, 2008 – 2017.

[NHT12] Rola Naja, Melhem El Helou, and Samir Tohmé, *WiMAX Double Movable Boundary Scheme in the Vehicle to Infrastructure Communication Scenario*, Wireless Personal Communications **67** (2012), no. 2, 387 – 413.

[NKG+12] Marco Nicosia, Roland Klemann, Kate Griffin, Stuart Taylor, Bernhard Demuth, Jaak Defour, Richard Medcalf, Thomas Renger, and Praveen Datta, *Rethinking Flat Rate Pricing for Broadband Services*, Cisco, July 2012.

[NVACT13] Quoc-Thinh Nguyen-Vuong, N. Agoulmine, E.H. Cherkaoui, and L. Toni, *Multicriteria Optimization of Access Selection to Improve the Quality of Experience in Heterogeneous Wireless Access Networks*, IEEE Transactions on Vehicular Technology **62** (2013), no. 4, 1785 – 1800.

[PBB+01] P.A Pangalos, K. Boukis, L. Burness, A Brookland, C. Beauchamps, and AH. Aghvami, *End-to-End SIP Based Real Time Application Adaptation*

During Unplanned Vertical Handovers, Proc. IEEE Global Communications Conference (GLOBECOM), November 2001.

[Per97] Charles E. Perkins, *Mobile IP: Design Principles and Practices*, Prentice Hall, 1997.

[PK06] K. Premkumar and A. Kumar, *Optimum Association of Mobile Wireless Devices with a WLAN-3G Access Network*, Proc. IEEE International Conference on Communications (ICC), June 2006.

[PKBV11] Kandaraj Piamrat, Adlen Ksentini, Jean-Marie Bonnin, and César Viho, *Radio resource management in emerging heterogeneous wireless networks*, Computer Communications **34** (2011), no. 9, 1066 – 1076, Special Issue: Next Generation Networks Service Management.

[PRGS08] J. Pérez-Romero, X. Gelabert, and O. Sallent, *Radio Resource Management for Heterogeneous Wireless Access Networks*, Heterogeneous Wireless Access Networks: Architectures and Protocols (Ekram Hossain, ed.), Springer, 2008, pp. 133 – 165.

[PRSA05] J. Pérez-Romero, O. Sallent, and R. Agustí, *Policy-based Initial RAT Selection Algorithms in Heterogeneous Networks*, Proc. IFIP International Conference on Mobile and Wireless Communications Networks (MWCN), September 2005.

[Put94] Martin L. Puterman, *Markov Decision Processes*, John Wiley, 1994.

[Rya02] Malcolm Ross Kinsella Ryan, *Hierarchical Reinforcement Learning: A Hybrid Approach*, 2002.

[SAAS10] Jatinder Pal Singh, Tansu Alpcan, Piyush Agrawal, and Varun Sharma, *A Markov Decision Process based Flow Assignment Framework for Heterogeneous Network Access*, Wireless Network **16** (2010), no. 2, 481 – 495.

[SJ05a] Qingyang Song and A. Jamalipour, *A Network Selection Mechanism for Next Generation Networks*, Proc. IEEE International Conference on Communications (ICC), May 2005.

[SJ05b] Qingyang Song and A. Jamalipour, *Network Selection in an Integrated Wireless LAN and UMTS Environment Using Mathematical Modeling and Computing Techniques*, IEEE Wireless Communications **12** (2005), no. 3, 42 – 48.

[SNW06] E. Stevens-Navarro and V.W.S. Wong, *Comparison Between Vertical Handoff Decision Algorithms for Heterogeneous Wireless Networks*, IEEE Vehicular Technology Conference (VTC Spring), May 2006.

[Std09] *IEEE Standard for Architectural Building Blocks Enabling Network-Device Distributed Decision Making for Optimized Radio Resource Usage in Heterogeneous Wireless Access Networks*, IEEE Std 1900.4-2009 (2009).

[SWMG08] Aimin Sang, Xiaodong Wang, Mohammad Madihian, and Richard D. Gitlin, *Coordinated Load Balancing, Handoff/Cell-Site Selection, and Scheduling in Multi-Cell Packet Data Systems*, Wireless Networks **14** (2008), no. 1, 103 – 120.

[TFC12] H. Tabrizi, G. Farhadi, and J. Cioffi, *Dynamic Handoff Decision in Heterogeneous Wireless Systems: Q-learning Approach*, Proc. IEEE International Conference on Communications (ICC), June 2012.

[THH02] A Tolli, P. Hakalin, and H. Holma, *Performance Evaluation of Common Radio Resource Management (CRRM)*, Proc. IEEE International Conference on Communications (ICC), April 2002.

[THK+10] Antti Toskala, Harri Holma, Troels Kolding, Preben Mogensen, Klaus Pedersen, and Jussi Reunanen, *High-Speed Downlink Packet Access*, WCDMA FOR UMTS - HSPA Evolution and LTE, Fifth Edition (Harri Holma and Antti Toskala, eds.), Wiley, 2010, pp. 353 – 389.

[WB09] Lusheng Wang and D. Binet, *Mobility-Based Network Selection Scheme in Heterogeneous Wireless Networks*, Proc. IEEE Vehicular Technology Conference (VTC Spring), April 2009.

[WD92] ChristopherJ.C.H. Watkins and Peter Dayan, *Technical Note: Q-Learning*, Machine Learning **8** (1992), no. 3-4, 279–292.

[WHB08] Damon Wischik, Mark Handley, and Marcelo Bagnulo Braun, *The Resource Pooling Principle*, ACM SIGCOMM Computer Communication Review **38** (2008), no. 5, 47 – 52.

[WK13] Lusheng Wang and G.-S.G.S. Kuo, *Mathematical Modeling for Network Selection in Heterogeneous Wireless Networks – A Tutorial*, IEEE Communications Surveys Tutorials **15** (2013), no. 1, 271 – 292.

[Ye11] Yinyu Ye, *The Simplex and Policy-Iteration Methods Are Strongly Polynomial for the Markov Decision Problem with a Fixed Discount Rate*, Mathematics of Operations Research **36** (2011), no. 4, 593 – 603.

[YK05] F. Yu and V. Krishnamurthy, *Efficient Radio Resource Management in Integrated WLAN/CDMA Mobile Networks*, Telecommunication Systems **30** (2005), no. 1-3, 177 – 192.

[YSN10] Xiaohuan Yan, Y. Ahmet Sekercioglu, and Sathya Narayanan, *A Survey of Vertical Handover Decision Algorithms in Fourth Generation Heterogeneous Wireless Networks*, Computer Networks **54** (2010), no. 11, 1848 – 1863.

[Zha04] Wenhui Zhang, *Handover Decision Using Fuzzy MADM in Heterogeneous Networks*, Proc. IEEE Wireless Communications and Networking Conference (WCNC), March 2004.

[ZJJ+12] Xinran Zhang, Hao Jin, Xiaodong Ji, Yong Li, and Mugen Peng, *A separate-SMDP Approximation Technique for RRM in Heterogeneous Wireless Networks*, Proc. IEEE Wireless Communications and Networking Conference (WCNC), April 2012.

[ZYA04] Fengxia Zhang, Yonghe Yan, and A. Ahmad, *Pricing for Efficient Usage in Wired and Wireless Networks*, Proc. International Telecommunications Network Strategy and Planning Symposium (NETWORKS), June 2004.

[ZYNT12a] Li Zhu, F. Richard Yu, Bin Ning, and Tao Tang, *Handoff Management in Communication-Based Train Control Networks Using Stream Control Transmission Protocol and IEEE 802.11p WLANs*, EURASIP Journal on Wireless Communications and Networking **2012** (2012), no. 1, 211 – 226.

[ZYNT12b] Li Zhu, F.R. Yu, Bin Ning, and Tao Tang, *Cross-Layer Handoff Design in MIMO-Enabled WLANs for Communication-Based Train Control (CBTC) Systems*, IEEE Journal on Selected Areas in Communications **30** (2012), no. 4, 719 – 728.