AN OPTIMAL INITIAL RADIO ACCESS TECHNOLOGY SELECTION METHOD FOR HETEROGENEOUS WIRELESS NETWORKS

by

Md Mizanur Rahman
BSc, BUET, Dhaka, Bangladesh, 2000

A thesis
presented to Ryerson University
in partial fulfilment of the
requirements for the degree of
Master of Science
in the Program of
Computer Science

Toronto, Ontario, Canada, 2013

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

MD MIZANUR RAHMAN

AN OPTIMAL INITIAL RADIO ACCESS TECHNOLOGY SELECTION METHOD FOR HETEROGENEOUS WIRELESS NETWORKS

Md Mizanur Rahman

MSc, Computer Science, Ryerson University, 2013

ABSTRACT

In Heterogeneous Wireless Networks, different overlapped Radio Access Technologies (RATs) can coexist with each other in the same geographical area. In such environment, a challenge is to select in which available RATs a user can be connected upon making an incoming service request. In this thesis, this challenge is investigated by proposing a Joint Call Admission Control (JCAC) -based approach that uses the framework of Semi-Markov Decision Process for initial RAT selection in two co-located wireless networks supporting two different service classes. The optimization problem involves the design of a cost function that weights the blocking cost and the energy consumption cost. The JCAC optimal policy is derived using the Value Iteration Algorithm. Simulations results show that the system capacity is maximized while selecting the less energy consuming RAT.

Acknowledgments

My utmost gratitude goes to my supervisor, Dr. Isaac Woungang for accepting me as his Masters student and guiding me with patience and encouragement. He has been abundantly helpful and has assisted me in numerous ways. Without his continual support and thoughtful mentoring, this thesis would not be possible.

Moreover, my appreciation goes to Dr. Glaucio H. S. Carvalho and my friends in our lab who were there when I needed help.

I dedicate this thesis to my family: my wife, my son and my parents who supported me unconditionally and have been the source of motivation and inspiration for me and made my time more enjoyable.

To my family for their kindness and encouragement.

Contents

List of Figures

List of Tables

List of Abbreviations

- **3GPP:** Third Generation Partnership Project
- **4G:** Fourth Generation Telecommunication
- **AHP:** Analytical Hierarchy Process
- **AP:** Access Point
- **CAC:** Call Admission Control
- **CDMA:** Code Division Multiple Access
- **CRRM:** Common Radio Resource Management
- **FDD:** Frequency-division Duplexing
- **FDMA:** Frequency Division Multiple Access
- **GPRS:** General packet radio service
- **GSM:** Global System for Mobile Communication
- **HetNet:** Heterogeneous Wireless Network
- **JCAC:** Joint Call Admission Control

- **JRRM:** Joint Radio Resource Management
- **MDP:** Markov Decision Process
- **OFDM:** Orthogonal frequency-division multiplexing
- **QoS:** Quality of Service
- **RAT:** Radio Access Technology
- **RRM:** Radio Resource Management
- **SAE:** System Architecture Evolution
- **SMDP:** Semi-Markov Decision Process
- **TDD:** Time-division Duplexing
- **TDMA:** Time Division Multiple Access
- **UE:** User Equipment
- **UMTS:** Universal Mobile Telecommunications System
- **WiMax:** Worldwide Interoperability for Microwave Access
- **WLAN:** Wireless Local Area Network

Chapter 1

Introduction

1.1 Context and Research Problem

The vision of Heterogeneous Wireless Networks (HetNets) is that of a new type of wireless networks where anyone can communicate with anyone else, anywhere and anytime, or can use any service of any network operator, through any network of any service provider in the most efficient and optimal way [1].

A HetNet integrates two or more different wireless networks, each having its own characteristics in terms of coverage, Quality of Service (QoS) assurance, implementation and operation costs, etc, into one common network. This means that usually the integrated networks provide overlap coverage in the same wireless service areas, allowing users to enjoy a great variety of innovative services based on their demands in a cost-efficient manner.

An illustration of a Heterogeneous Wireless Networks (HetNets) is shown in Fig. 1.1.

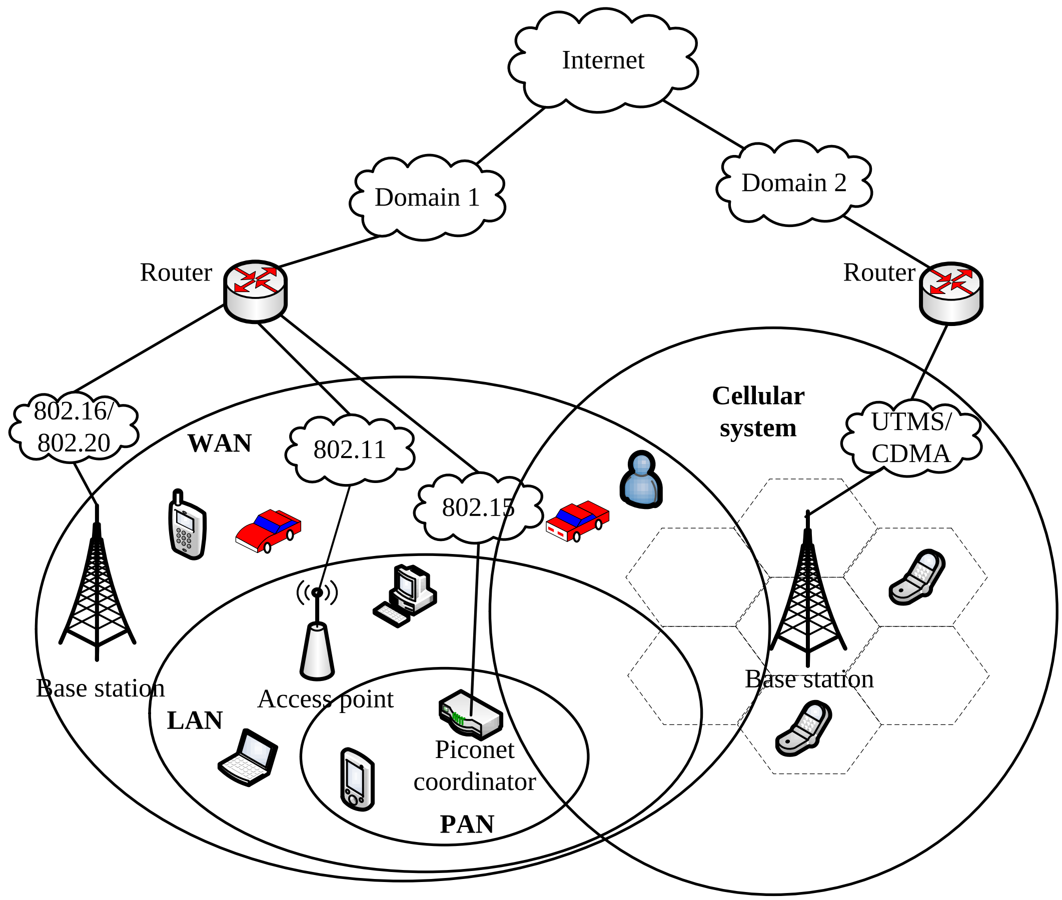


Figure 1.1: An illustration of a HetNet roaming scenario

Typically, different radio access technologies (RATs) are expected to coexist in the HetNets architecture, with the goal to provide ubiquitous access with high rates for mobile users through multi-mode terminals[2]. Indeed, these terminals have more than one radio interfaces, each enabling access to different access technologies. With such capabilities, terminals can initiate connectivity through the technology that most closely matches the users or application requirements.

From an operational point of view, HetNets (like any other mobile network) should follow the System Architecture Evolution (SAE) standards of the 3GPP release 8[3], and the design of any HetNet mechanism should follow these standards, in particular, the design of a suitable RAT discovery and selection mechanism for HetNets.

Following the SAE standards, the network discovery by user equipment (UE) should be performed in an efficient manner, i.e. without constant scanning of the other RATs in the HetNets architecture. That way, there will be some energy consumption savings at the UE side. One way to achieve this is to ensure that the information about the availability of other RATs is readily available including the location of the target UE [3]. On the other hand, the best RAT to be connected to should be selected by the UE based on criteria such as type of device, traffic load in the available RATs, usage rules and restrictions in the user subscription, to name a few.

This thesis advocates the idea that rather than concentrating the decision of network discovery and RAT selection in the hand of an UE, it would be more efficient to have an external agent to perform this decision. Such external agent is here referred to as a Radio Resource Management (RRM) framework, which will not only provide ubiquitous coverage in the HetNets, but also will support various applications and services, including the coordination of the various RATs present in the HetNets architecture. The RRM should not only decide whether an incoming service request can be accepted or rejected, but should also decide which of the available RATs is best suited to accommodate the incoming service request. Two benchmarking initial RRM frameworks that have been designed for HetNets are the Common RRM (CRRM) and the Joint RRM (JRRM)[4].

As part of the JRRM framework is the Call Admission Control (CAC) mechanism, which can be used to address the aforementioned issue since it defines (in principle) how the radio resources or wireless channels have to be efficiently shared among the incoming service requests.

Traditional CAC mechanisms for wireless networks are not suitable for network with heterogeneous entities such as HetNets and a suitable CAC protocol design within JRRM is yet to be implemented. Therefore, new CAC solutions in the form of a Joint Call Admission Control (JCAC) consisting of two functions: (1) one to decide whether an incoming service request should be accepted or rejected (blocked); and (2) one to select in which of the available RATs an incoming service request has to be accommodated; is always desirable.

1.2 Approach

We propose an optimal JCAC scheme for initial RAT selection in two co-located wireless networks (as shown in Fig. 1.2), which supports two different service classes (calls).

To meet the JCAC goals, a cost function is proposed that weights two criteria: a local cost referred to as blocking cost function, which takes into account the priority of each service class in each RAT and reflects the overloaded RAT; and an energy consumption cost which is meant to measure the battery power savings in the network.

The following components constitute our JCAC approach:

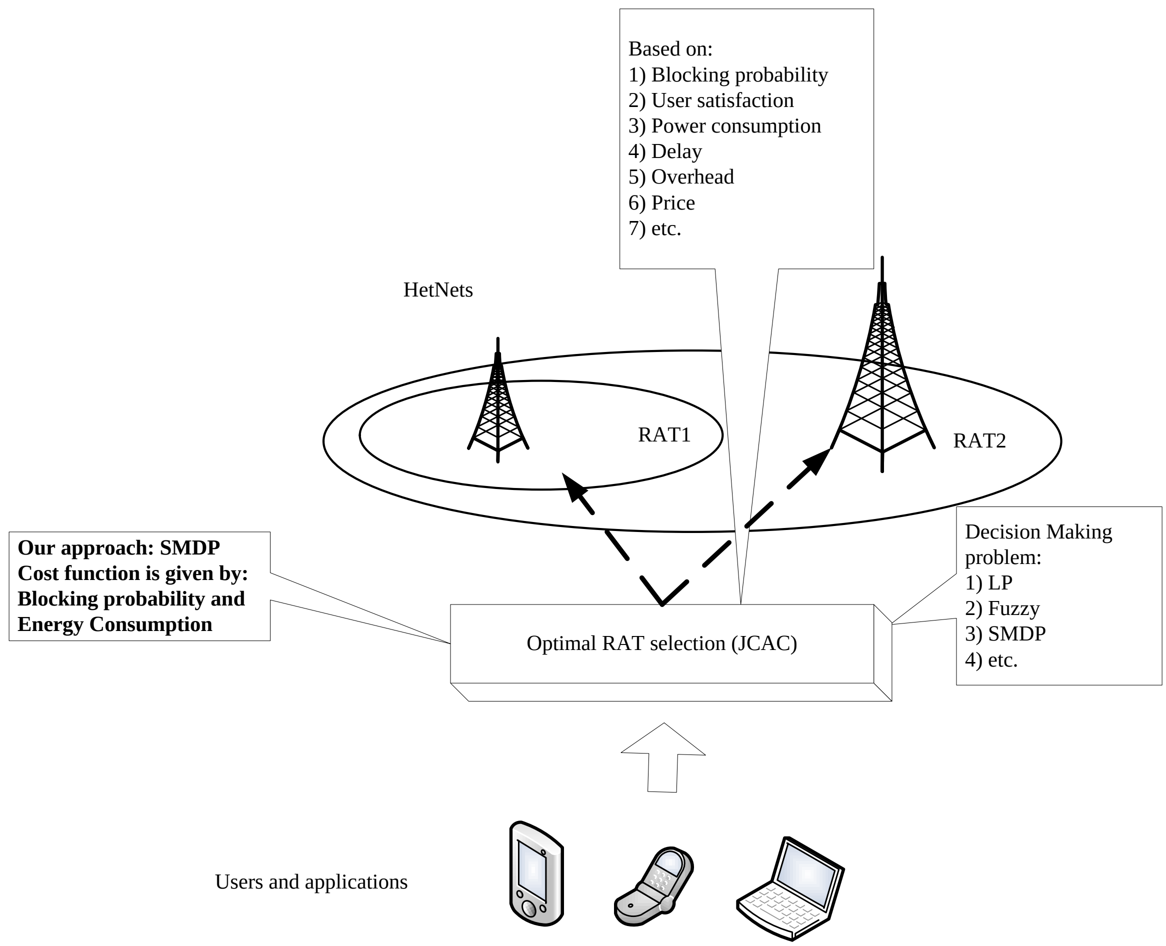


Figure 1.2: An illustration of our approach

- We use the framework of Semi-Markov Decision Process (SMDP) to formulate the JCAC optimization problem.
- We use the value iteration algorithm to compute the optimal JCAC policy.

1.3 Thesis Contributions

This thesis focuses on addressing the JCAC problem in HetNets. Our main contributions are twofold:

- We have formulated the JCAC optimization problem using SMDP, and derived the JCAC optimal policy using the value iteration algorithm.
- We have analyzed the structure of the derived optimal policy (by simulation) using an initial HetNet architecture involving two types of service classes from two co-located types of RATs.

1.4 Thesis Organization

The remainder of this thesis is organized as follows.

In Chapter 2, we describe what is HetNets, RRM, JRRM in HetNets and overview some representative related works.

In Chapter 3, we introduce the system and traffic assumption and present the formulation of the optimization problem using the SMDP approach. The value iteration algorithm is also described.

In Chapter 4, we provide the performance metrics and the simulation results under different system configurations. We also analyze the optimal policy derived from the formulated model.

In Chapter 5, we conclude our work.

Chapter 2

Background and Related Work

Even though the RAT selection problem in HetNets has attracted a lot of attention in the recent years, to our knowledge, there is no comprehensive survey on this topic, in particular, the topic of energy saving-based techniques for RAT selection in HetNets.

The objective of this chapter is to shed the light on the most relevant aspects of the RAT selection problem. Representative JRRM-based frameworks that have been proposed so far (to our knowledge) for HetNets are discussed.

2.1 Heterogeneous Wireless Networks

Heterogeneous Wireless Networks (HetNets) are typically composed of a conglomeration of multiples wireless networks and technologies such as Wifi, Bluetooth, Zigbee, WiMax, IEEE802 WLANs, cellular and mobile technologies called radio access technologies/networks (RATs), that are expected to operate together in a complementary manner. As an example, the IEEE 802.11 WLANs and 3G cellular systems can operate as a composite heterogeneous wireless network to provide higher bandwidth services over a wider geographic area.

The concept of HetNets compared to that of homogeneous networks was discussed in [5]. This concept was motivated by the desired to combine several advantages that are offered through the features of each homogeneous network, for instance, the widespread coverage feature offered by WiMax and cellular technologies, the low cost and high bandwidth derived from using WiFi, Zigbee or Bluetooh, the possibility of extending the users selection technology and radio when working with various wireless applications, to name a few.

In an HetNet environment, each wireless access network has its own characteristics such as capacity, access technology, security, power consumption, delay, coverage, access cost, to name a few [6]. An interesting feature of HetNets is the fact that some wireless access networks are overlaid by others in such a way that a multi-layer structure or a hierarchical cellular mobile network is naturally built. This architecture of overlay networks can be suitably explored to match multiple design purposes such as:

- Boosting the system capacity: By combining the capacity of each individual wireless access network within the intended coverage region, the whole system may support many more users. This fact leads to the reduction in blocking/dropping probabilities of new/handoff calls by offering alternative access points during overload situations [7].

- Increasing the system coverage. For instance, by combining WiMAX and cellular mobile networks, a large geographical area can be covered.

- Enhancing the user satisfaction. Given the differences in technologies and data rates, each wireless network can be employed to satisfy a specific target. For example, often, WLANs or picocells have been utilized to furnish the access in hot spot areas such as airports, restaurants, shopping centers, while cellular mobile networks have been used for ensuring the users mobility.

- Offering different access costs for end-users. In practice, some end-users are willing to pay high prices for wireless access given their social position, economic situation, job, or necessity. However, the major portion of users would like to be connected with the access network that provides the lowest cost and an appropriated QoS. Aware of this fact, the MNOs have designed the market strategies that are appealing to all classes of end-users, aiming at increasing their profits by holding and attracting new end-users [8]. Therefore, pricing strategies arise as one of the most important design criteria in HetNets. Thus, given the mix of access networks with different prices, an end-user is able to decide among the available options which network better fits its pocket and expectations.

In order to make full use of HetNets and enjoy the advantages that these networks offer, a mobile user must be equipped with a multi-interface device that is able to sense and connect with the access network that matches his/her personal expectations and the requirements of his/her applications.

Elementary to the operation of HetNet is the existence of modern multi-mode wireless terminals [2]. These terminals have at least one UMTS radio access mode

(FDD and/or TDD) and they support one or more other 2G RATs (e.g. GSM, cdmaOne, GPRS, etc). On the other hand, single-mode terminals are those that can support one type of RAT. This multi-mode feature enables them to access different access technologies. With such a capability, terminals can initiate connectivity through the technology that most closely matches the users or applications requirements. For operators of 4G wireless networks, when a service request comes to the system, it can direct the request to a particular network that best suits the user's requirements and/or that complies with the status of different networks. Moreover, the admission load can be balanced between the different networks by the 4G system.

2.2 Radio Resource Management in HetNets

The problem of Radio Resource Management (RRM) can be roughly defined as the problem of assigning a server for a incoming service request as long as it does not violate the service provisioning of the other ongoing users and there is enough capacity to accept it.

Traditionally, in homogeneous wireless networks, the RRM functionalities are addressed with just one network. In overlay homogeneous cellular mobile networks, each layer operates independently of the others, then the same practice is applied and the RRM algorithms have only to handle the incoming service requests offered in each layer. For instance, the call admission control (CAC) in homogeneous wireless systems has only to decide whether a user is accepted or rejected. However, the actual scenario in which heterogeneous wireless access networks cover a geographical area imposes new design paradigms to the RRM algorithms. For example, due to the offered traffic load variations in space and time, some cells located somewhere in the covered region experiment overloads situations, while others, on the other hand, are quite idle. As a consequence of this practice, wireless resources are poorly utilized and the offered traffic load is deficiently carried out.

To overcome these drawbacks, RRM frameworks such as the JRRM [9] and CRRM [1] have been proposed. The main advantage of these frameworks is the fact that they have the whole vision of all layers and they can cope with the following tasks:

- Deciding whether an incoming service request should be accepted or blocked.
- Selecting in which of the cells an incoming service request has to be accommodated.

The first objective of JRRM relies in the idea behind the homogeneous wireless systems, but the second takes advantage of the fact that one of the available wireless access networks can be chosen as long as it better suit the users expectations in terms of cost, delay, data rate, to name a few. The JRRM functionalities [10] can be grouped into three main procedures: resource monitoring, decision making, and decision enforcement (as shown in Fig. 2.1).

- Resource monitoring: This step keeps track of the resources by gathering information from both users and networks sides and forward this information to the decision making step. This step is executed at two different occasions in the framework, one after the arrival of a device/user into the networks and the other after the device/user has been connected to the networks.

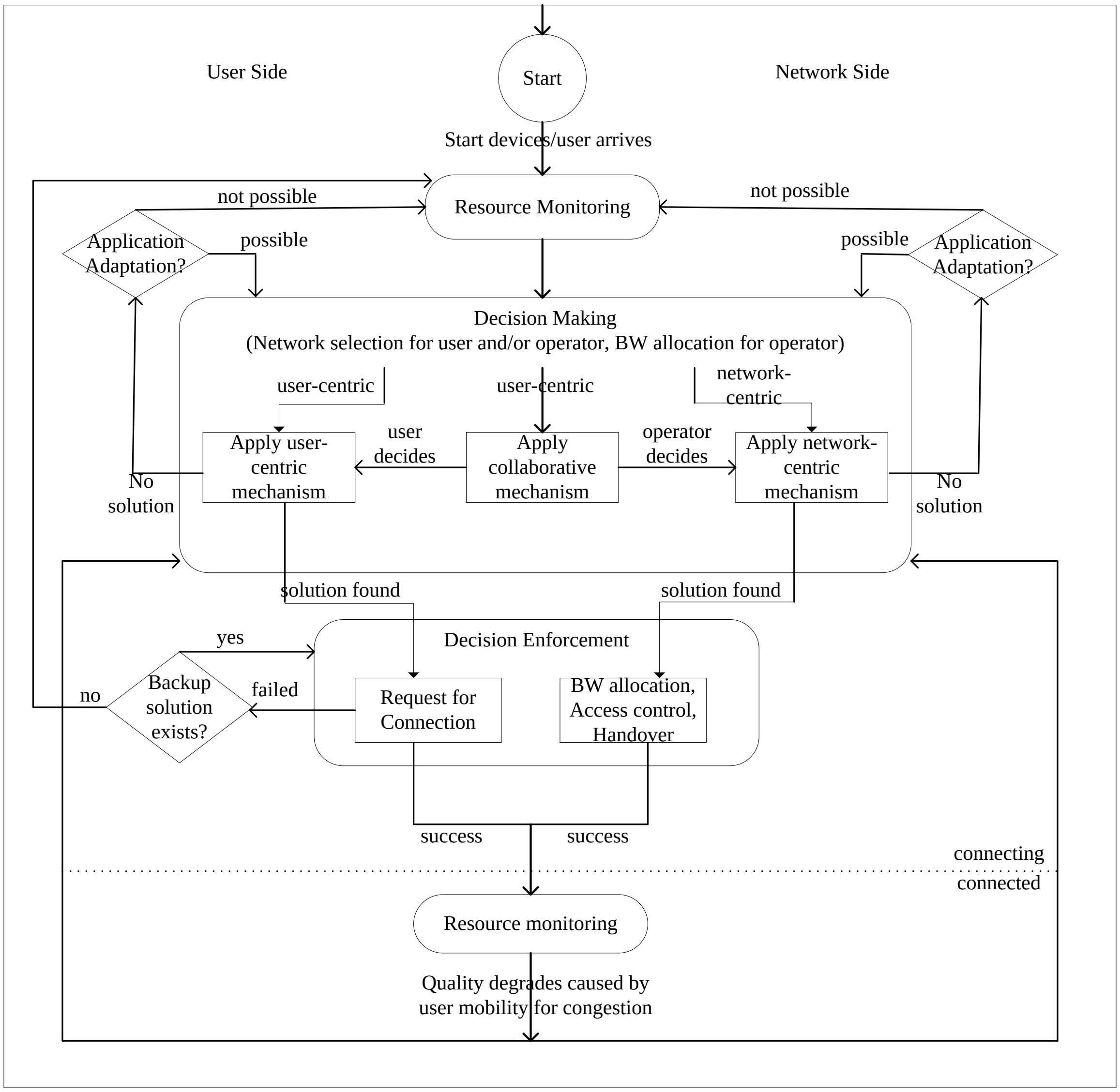


Figure 2.1: Radio Resource Management in HetNets

- Decision making: Two types of decision are taken for the connection of a device/user, one is selecting a particular network for the connection and the other is allocating the bandwidth for that connection either from a network or by distributing the bandwidth from multiple networks.

  There are three approaches for making the decision in the RRM framework according to who will benefit from it:

- Network-centric approach: In this approach, decisions are made at the network side considering both the profit for the network operator and user's requirements. But the main concern here is how to make optimal uses of the network bandwidths.
- User-centric approach: This approach mostly deals with network selection. Decisions are made at the user's terminal only considering the user's profit and not the network load balancing or any other user's profit. Therefore, the network can easily be congested and hereby consume high energy if the connection is rejected by the operator.
- Collaborative approach: In this approach, both network operators and users participate in the resources allocation by compromising the profit between users and network operators that gives low connection rejection.

• Decision enforcement: This step executes the decision adopted at the decision making step.

2.3 Related Work

Seamless integration of heterogeneous wireless networks is one of the most important requirements for the future deployment of new wireless technologies such as 4G mobile systems, 3G systems, WLANs, Bluetooth, and ultrawideband. Indeed, using dense or sparse HetNets architectures or a combination of both [11], it is expected that multiple RATs will be available in various locations in the targeted network, and the user device or networks will be able to decide on the best access to these RATs to achieve their goals. To do so, managing the resources in the network while keeping in mind the QoS from a users perspective is one of the key challenges [11].

Our focus here is on designing a RAT selection mechanism that can contribute towards a better radio resource management in HetNets. Few recent works [12, 13, 14, 15, 16, 17] dealing with this type of objective are described as follows.

In [12], Lopez-Benitez et al. developed some techniques referred to as common radio resource management (CRRM) to distribute heterogeneous traffic among the available RATs, while taking into account the radio resources available at each RAT. Their framework algorithms is shown to achieve appropriate user/service QoS levels based on some decision criteria used for determining the most suitable user-to-RAT assignment. However, the authors did not elaborate on how their proposed schemes would handle the application scenarios where the RAT selection decisions are based on the channel quality conditions. In such cases, there is no guarantee that the users connection to the RAT can be maintained.

In [13], Atanasovski et al. studied the problem of convergence of wireless access networks and reported on an architecture for resource management in HetNets (called RIWCoS). This architecture relies on the concept of Media Independent Handover Function (MIHF) user, a key component of the IEEE 802.21 framework. Although the RIWCoS framework is shown by simulations to be promising in handling emergency situations such as disaster management, its applicability requires that both terminal and network side resource management modules be implemented, under various assumptions and prerequisites.

In [14], Suleiman et al. investigated the joint radio resource management problem in HetNets and proposed a solution that considers the features of component access technologies such as the occupancy level in individual access networks, load balancing, nature of individual access networks, variability of network resources according to traffic conditions, asymmetry of access networks overlap, to name a few. A prototype implementation of their framework is presented and validated. However, this model does not consider the diversity of access networks, nor the users mobility patterns and the impact of overlapping of different access networks.

In [15], Kajioka et al. also investigated the problem of resource management in HetNets. Their proposal is an adaptive resource allocation scheme in which each node determines by itself the wireless network resources to be assigned to every applications that it support. This is achieve through designing an attractor composition model that describes the global activity shared among the entities in HetNets. However, the proposed scheme was not tested on real simulation testbed using real scenario applications.

In [16], Pei et al. studied the problem of radio resource management and network selection in HetNets composed of CDMA and WLANs networks. In the case of CDMA, radio-resource is achieved through solving an optimization problem that consider the inter cell interference levels as criterion for maximizing the total network welfare under the CDMA resource-usage constraints. For the WLAN part, radio resources is achieved through solving an optimization problem that maximizes the aggregate social welfare of the WLAN under the WLAN resource-usage constraints. In their proposed scheme, the method used to balance the load among mobile nodes was not disclosed.

In [17], Ngo and Le-Ngoc introduced two distributed resource allocation methods to optimally allocate subcarriers and power in an OFDMA-based cognitive radio ad hoc network. These methods are designed using the Lagrangian dual optimization where the throughput (or energy efficiency) is maximized subject to constraints such as the tolerable interference at the primary network level, the enforcement on the lower and upper bounds on the number of sub channels that each individual unlicensed users may occupy.

In [18], Giupponi et al. proposed a JRRM framework as a way for achieving an efficient usage of a joint pool of resources belonging to different RATs. They considered three RRM functions namely: the RAT and cell selection function which handles the beginning of each session by selecting the RAT and cell to the mobile, the bit rate allocation function which allocates a suitable bandwidth for each RAT and accepted user, and the admission control function which decides whether a service request should be accepted or rejected.

In [19], Olabisi et al. proposed an adaptive bandwidth management and joint call admission control scheme for HetNets. Which is made of three components. The Joint call admission controller which is meant to handle the call admission decision as well as distributing the traffic load uniformly among the available RATs; threshold-based bandwidth reservation function that maintains a lower handoff dropping probability, and a bandwidth adaptation controller function that executes the bandwidth adaptation when a call arrival or departure event has arisen. Their proposed adapting bandwidth mechanism gives equal priority to all calls when randomly selected calls have been downgraded or upgraded, which is bad for some call performance. In addition, their method does not consider any user or service satisfaction in its design.

In [20], Pérez-Romero et al. proposed a policy based RAT selection algorithm, where a function selects an initial RAT from a set of available RATs based on a set of different inputs such as service class, traffic load in each RAT, UE features, node mobility speed, etc. In order to avoid blocking possibilities when there is capacity available in other RATs, complex policies were proposed by combining basic policies in which the output is prioritized for a list of RATs.

In [21], Falowo et al. proposed a dynamic RAT selection algorithm for assigning a multimode terminal with a single call (or group of calls) to the most suitable RAT in the HetNets. To select the desired RAT, it rates the available RATs using a multi-criteria group decision-making technique which itself is based on a modified fuzzy TOPSIS technique by specifying a set of calls, available RATs, call priorities, and criteria weights. Their proposed algorithm is shown to save battery power consumption by selecting only one RAT at a time for multiple calls from a multimode terminal. However, it fails to treat each call when multiple calls originated from a multimode terminal are admitted into different RATs.

In [22], Haldaret et al. proposed a cross-layer architectural framework for network and channel selection in heterogeneous cognitive wireless network. The proposed framework classifies the user's application based on Analytic Hierarchy

Process algorithm then categorizes channels within an operating spectrum by using a probabilistic recurrence relation. Based on these assets, a suitable channel is selected within the spectrum as well as a proper network according to its user needs. The proposed scheme is shown to drop the blocking probability of the user's applications compared to the greedy and FCFS schemes but it requires a higher number of handoffs compared to other schemes.

In [23], Stevens et al. proposed a vertical handoff decision algorithm based on the Markov decision process for HetNets. The MDP is used to determine the optimal policy that maximizes the expected total reward per connection, where different link reward functions are assigned to network-based QoS parameters such as bandwidth and delay. However, the MDP model is formulated by using discrete-time, which means that the decision is taken at fixed time interval. This contrasts with where continuous-time MDP model is needed for taking the decision whenever any change happens in the networks.

In order to boost up the total capacity and provide better QoS in HetNets, a RAT selection procedure was proposed in [24]. The proposed approach uses a fuzzy logic algorithm to balance the load in the multiservice HWNs. First, a fuzzy logic controller is used to transform the remaining bandwidth of 3G network and WLAN, and a fuzzy engine gives a fuzzy output based on the if-then fuzzy rules. Finally through defuzzification, the fuzzy output is transformed to select a network which represents the probability for the user to choose 3G network or WLAN. However, a data user has to wait to access a network until sufficient bandwidth is released by any of the constituent network when the user is rejected.

In [25], Lucas et al. proposed an enhanced JRRM technique that simultaneously finds out for each user an adequate combination of RATs and a number of distinct radio resources within such RATs. Moreover, by considering the current network load, the proposed scheme selects the best RAT for the new incoming call and equally satisfies all users by realizing a user fairness policy. However, for high system load, some low priority users are eliminated from the radio resource distribution process in order to satisfy the minimum QOS level to all active users in the system.

In [26], Porjazoski et al. proposed a RAT selection approach for choosing both the new incoming call and ongoing calls (handover calls) based on service type, user mobility and network load. A two-dimensional Markov chain is used to analyze the performance of the proposed approach showing that and it outperforms existing single or two criteria RAT selection approaches.

In [27], Mohamed et al. evaluated various weighted algorithms to determine the appropriate weights for different criteria (i.e. velocity, user preferences, QoS) of multi attribute decision making algorithms for selecting the best access network in HetNets. However, in their approach, there is no methodical way to choose the weights for the different criteria that are involved in their design.

In [28], Zhu et al. introduced an immune optimization algorithm for solving joint call admission control (JCAC) problem in HetNets. This algorithm optimizes both the user's preference and traffic load distribution by using the concept of dynamic pricing. The proposed model evaluates alternative RATs for each arriving call based on a set of selection criteria (namely: data rate, service price etc) which are weighted according to the user's preference. However, only the dynamic pricing technique can still lead the system unbalanced traffic load.

In [29], Si et al. proposed an optimal network selection method for HetNets, which is based on a multimedia distortion as an application layer QoS and network access price. An optimal selection policy is used to assign indices to candidate networks by formulating the problem as a stochastic optimization problem that selects the lowest index network. A system reward function is used to achieve the optimization goal.

Most of the above described schemes do not utilize the advantages of the multi-criteria nature of the RAT selection that can give better performance than single criterion algorithms due to the flexible and complementary nature of the different criteria.

Considering only one criterion in the RAT selection is not sufficient to provide a good solution and usually leads to undesirable situations. Unlike previous works, our approach considers a cost function which accounts for two optimization criteria based on different weights setting: the blocking cost-which reflects the overloaded RAT and the energy consumption cost. Most of the literature use MDP for formulating the problem where the decisions are made only in fixed epochs. However, in this work, the system dynamic is governed by arrivals and departures, which are modelled by exponential distributions. This way, the times between the decision epochs are random (rather than fixed). The mathematical tool used to analyze such type of stochastic problem is the SMDP.

Chapter 3

Methodology

As discussed in Chapter 1 and Chapter 2, the problem of RAT selection in HetNets is still a challenge.

In this Chapter, we describe our novel solution to this problem, which consists of the design of a new SMDP-based model for JCAC and the value iteration algorithm for optimal RAT selection based on this model.

3.1 Traffic Model

We consider a HetNet consisting of 2 co-located RATs in which the j^{th} RAT ($j = 1, 2$) has N_j radio resources. Each incoming service (call) is served by allocating the required amount of resources from one of the available RAT. In general, a unit of radio resource depends on how the radio interface is implemented as well as which RAT technology (FDMA, TDMA, CDMA or OFDM)is used. In addition, the system capacity can be implemented by its effective or equivalent bandwidth [30, 31], no matter which multiple access technologies are used for the radio interface.

In HetNets environment, when an incoming service connection requests an access to the network, the optimal RAT selection method has to decide not only if it will be accepted, but also in which RAT it should be accepted. The HetNets supports K classes of service connections where each class is categorized by its bandwidth requirement, arrival distribution, and channel holding time. Here, we consider two types of service connections (calls),i.e. $i = 1, 2$. We also assume that the i^{th} service connection comes in according to a Poisson process with parameter λ_i and it requires b_i radio resources. The channel holding time (let connection duration + residence time) is assumed to follow an exponential distribution with mean rate equal to μ_i. Finally, the traffic intensity is defined as $\rho_i = \lambda_i/\mu_i$.

3.2 Formulation of the Optimization Problem

In order to model and solve the optimal control problem (that of selecting the RAT in HetNets), we rely on the SMDP framework.

We first design the new SMDP model and the cost function for our traffic model. Then, we introduce the data transformation method used to obtain a discrete-time based Markov Decision Process (MDP) model. After this we use value iteration algorithm [32] in the transformed model to determine the optimal JCAC policy.

A SMDP model is determined by five components, namely, the state space, the decision epochs and the actions, the expected time until the next decision epoch, the state transition probabilities, and the cost function.

3.2.1 States of the SMDP

The states of the SMDP is a five-tuple.

$$S = (n_{11}, n_{21}, n_{12}, n_{22}, e) \tag{3.1}$$

and with the following constraints associated to each RAT:

$$0 \leq n_{11} \leq \lceil N_1/b_1 \rceil$$

$$0 \leq n_{21} \leq \lceil N_1/b_2 \rceil$$

$$0 \leq n_{12} \leq \lceil N_2/b_1 \rceil$$

$$0 \leq n_{22} \leq \lceil N_2/b_2 \rceil$$

$$e = [0\ 1\ 2]^T$$

where M^T denotes the transpose of matrix M, n_{ij} is the number of calls of type i connection in RAT j, N_j is the capacity of RAT j, b_i is the bandwidth required by the type i connection(call) and $e = 0$ is the departure of connection and $e = 1$ is the arrival of connection of the type 1, and $e = 2$ is that of type 2 .

3.2.2 Decision Epochs and Actions

There are three possible actions for our JCAC policy, namely,

- Block (B): Meaning that the incoming call is blocked by the JCAC.
- Accepted into RAT1 (AR1): Meaning that the incoming call is accepted into RAT1 by the JCAC.
- Accepted into RAT2 (AR2): Meaning that the incoming call is accepted into RAT2 by the JCAC.

In each state $x \in S$, the controller can choose one of the following possible actions upon arrival of a call:

$$A(x) = \begin{cases} B, & e = 0, 1, 2 \\ AR1, & e = 1, 2 \text{ and } b_{(i=(1,2))} + b_1 n_{11} + b_2 n_{21} \leq N_1 \\ AR2, & e = 1, 2 \text{ and } b_{(i=(1,2))} + b_1 n_{12} + b_2 n_{22} \leq N_2 \end{cases} \tag{3.2}$$

3.2.3 Expected Time Until the Next Decision Epoch

If the system is in the state $x \in S$ and the action $a \in A(x)$ is chosen, then the expected time until the next decision epoch is given by:

$$\tau(x, a) = \frac{1}{\lambda_1 + \lambda_2 + n_{11}\mu_1 + n_{21}\mu_2 + n_{12}\mu_1 + n_{22}\mu_2} \tag{3.3}$$

3.2.4 Transition Probabilities

The transition probabilities among the system states are meant to specify the state dynamic. Let $p(x, y, a)$ be the probability that at next decision epoch, system will be in state $y \in S$ if action $a \in A(x)$ is chosen in state x. And $\tau(x, a)$ be the expected time until the next decision epoch if action $a \in A(x)$ is chosen in state x. The transition probabilities are then obtained as:

$$p(x, y, a) = \begin{cases} \lambda_1 \tau(x, a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 1) \Rightarrow y = x, & \text{if} \quad a = B \in A(x) \\ \lambda_1 \tau(x, a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 1) \Rightarrow y = (n_{11} + 1, n_{21}, n_{12}, n_{22}, e), & \text{if} \quad a = AR1 \in A(x) \\ \lambda_1 \tau(x, a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 1) \Rightarrow y = (n_{11}, n_{21}, n_{12} + 1, n_{22}, e), & \text{if} \quad a = AR2 \in A(x) \end{cases} \tag{3.4}$$

in case of arrival of type-1 call.

$$p(x,y,a) = \begin{cases} \lambda_2\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 2) \Rightarrow y = x, & \text{if} \quad a = B \in A(x) \\ \lambda_2\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 2) \Rightarrow y = (n_{11}, n_{21}+1, n_{12}, n_{22}, e), & \text{if} \quad a = AR1 \in A(x) \\ \lambda_2\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 2) \Rightarrow y = (n_{11}, n_{21}, n_{12}, n_{22}+1, e), & \text{if} \quad a = AR2 \in A(x) \end{cases} \tag{3.5}$$

in case of arrival of type-2 call.

$$p(x,y,a) = \begin{cases} n_{11}\mu_1\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 0) \Rightarrow y = (n_{11}-1, n_{21}, n_{12}, n_{22}, e), & \text{if} \quad a = B \in A(x) \\ n_{21}\mu_2\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 0) \Rightarrow y = (n_{11}, n_{21}-1, n_{12}, n_{22}, e), & \text{if} \quad a = B \in A(x) \\ n_{12}\mu_1\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 0) \Rightarrow y = (n_{11}, n_{21}, n_{12}-1, n_{22}, e), & \text{if} \quad a = B \in A(x) \\ n_{22}\mu_2\tau(x,a),\ x = (n_{11}, n_{21}, n_{12}, n_{22}, 0) \Rightarrow y = (n_{11}, n_{21}, n_{12}, n_{22}-1, e), & \text{if} \quad a = B \in A(x) \end{cases} \tag{3.6}$$

in case of departures of calls.

3.2.5 Cost Function

If the system is in the state $x \in S$ and the action $a \in A(x)$ is chosen, the admission control incurs based on the following cost function:

$$C(x,a) = \omega_1 g_{bc}(x,a) + \omega_2 g_{ec}(x,a), \tag{3.7}$$

where ω_1 is the weight associated to the blocking cost function $g_{bc}(x,a)$ and ω_2 is the weight associated to the energy consumption cost $g_{ec}(x,a)$. These costs can be viewed by the mobile network operators as a way to set the relative importance of each system objective. We assume that $\omega_1 + \omega_2 = 1$.

The blocking cost function $g_{bc}(x,a)$ is a fixed cost incurred whenever an incoming service (call) is blocked by the system. It is defined as:

$$g_{bc}(x,a) = \begin{cases} BC_i, \; e = 1,2 \qquad \text{and} \qquad a = B \\ \qquad\qquad 0, \quad \text{otherwise} \end{cases} \tag{3.8}$$

where BC_i is the blocking cost of the i^{th} service class.

Energy consumption means the energy required to operate a Base Station or the Access Point. The energy consumption cost function $g_{ec}(x,a)$ is defined by:

$$g_{ec}(x,a) = \begin{cases} \frac{E_j}{\max_j(E_j)}, \; e = 1,2 \qquad \text{and} \qquad a = AR_j \\ \qquad\qquad 0, \quad \text{otherwise} \end{cases} \tag{3.9}$$

where E_j is the energy consumed by the j^{th} RAT. It should be noted that in Equation 3.9, the energy consumption is normalized so as to have its value less than or equal to 1.

3.2.6 JCAC Policy

A JCAC policy is an n-tuple of a vector specifying for each state of the MDP the action to be selected in that state. Here, we consider a stationary and deterministic policy, i.e. the policy does not change in time and in a given state, the policy specifies a single action (with probability 1). Note that for a Markov decision model with finite state space and finite action sets, there exists an optimal policy which is stationary and deterministic. Such a policy is an application from S to A, which associates at each state x an action in $A(x)$:

$$\forall x \in S, R_x \in A(x) \tag{3.10}$$

It should be noted that for the derivation of the performance parameters for a given policy, the SMDP model with transition probabilities $p(x, y, R_x)$ is a traditional continuous time Markov chain.

3.2.7 Data-Transformation and Value Iteration Algorithm

To get the optimal JCAC policy, we need to convert the continuous time SMDP model into a discrete time MDP model such that for each stationary policy, the average cost per time unit in the discrete-time Markov model is the same as in the semi-Markov model. This approach is referred to as data-transformation method. After this step, the value iteration algorithm can be used in the transformed model to get the optimal policy. The data-transformation method [32] is described as follows. Let $x, y \in S$ and the action $a \in A(x)$, choose a number τ such that:

$$0 < \tau \leq \min_{x,a} \tau(x, a) \tag{3.11}$$

then perform the following transformations:

$$\begin{aligned}
\overline{S} &= S \\
\overline{A}(x) &= A(x),\ x \in \overline{S} \\
\overline{C}(x, a) &= \tfrac{C(x,a)}{\tau(x,a)},\ x \in \overline{S}\ and\ a \in \overline{A}(x) \\
\overline{p}(x, y, a) &= \begin{cases} \frac{\tau}{\tau(x,a)} p(x, y, a),\ x \neq y, x \in \overline{S}\ and\ a \in \overline{A}(x) \\ \frac{\tau}{\tau(x,a)} p(x, y, a) + [1 - \frac{\tau}{\tau(x,a)}],\ x = y, x \in \overline{S}\ and\ a \in \overline{A}(x) \end{cases}
\end{aligned} \tag{3.12}$$

where notation $\bar{i}$ means the converted component. After turning the continuous time SMDP model into a discrete time MDP model, the value iteration algorithm is used to obtain the optimal JCAC policy $R(n)$ whose average cost function is given by $g_i(R(n))$ such that:

$$0 \leq \frac{g_i(R(n)) - g^*}{g^*} \leq \epsilon \tag{3.13}$$

where g^* denotes the minimal average cost per time unit.

The pseudo-code of the value iteration algorithm is described in Algorithm 1.

Algorithm 1 Value-Iteration Algorithm

Step 0: *Initialization*

Choose $V_0(i)$ *such that* $0 \leq V_0(i) \leq min_a\{c(i,a)/\tau(i,a)\}$ *for all* i.

Choose a number τ *where* $0 < \tau < min_{i,a}\tau(i,a)$.

Let $n = 1$

Step 1: *Value-iteration step*

Compute the function $V_n(i), i \in I$ *using*

$$V_n(i) = \min_{a \in A(i)}[\frac{c(i,a)}{\tau(i,a)} + \frac{\tau}{\tau(i,a)} \sum_{j \in I} p(i,j,a)V_{n-1}(j) + (1 - \frac{\tau}{\tau(i,a)})V_{n-1}(i)] \tag{3.14}$$

Let $R(n)$ *be a stationary policy whose actions minimize the right-hand side of Equation* 3.14.

Step 2: *Compute the bounds* m_n *on the minimal cost using*

$$m_n = min_{j \in I}\{V_n(i) - V_{n-1}(i)\}, \;\; M_n = max_{j \in I}\{V_n(i) - V_{n-1}(i)\} \tag{3.15}$$

Step 3: *Stopping condition. The algorithm is stopped when policy* $R(n)$ *is obtained such that* $0 \leq (M_n - m_n) \leq \epsilon m_n$ where ϵ *is a prescribed accuracy number. Otherwise, go to step 4.*

Step 4: *Continue*

$n = n + 1$ *and go to step 1.*

Chapter 4

Performance Evaluation

4.1 Simulation Tool and Parameters

Due to the state-space explosion problem as shown in the Equation 3.1, the discrete time Markov model can only be solved for small values of N_j. Hence we have implemented a simulation model, and obtained numerical results for the cases where $N_1 = 20$ and $N_2 = 10$.

The simulation model is an event-driven system written in Borland C++ 5. Simulations for each system configuration run for sufficiently small precision value, $1.0e - 12$, in Equation 3.13 for Value Iteration Algorithm.

In order to illustrate the performances and the optimal structure of our proposed optimal RAT selection policies, the two co-located networks, RAT1 and RAT2 will be used as representatives of GSM and UMTS technologies. We have also consider two types of service classes, class-1 and class-2, for each RAT. The fixed simulation parameters are captured in Table 4.1.

Table 4.1: Fixed parameters values for simulation

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
N_1	20 *channels*	μ_1	1/120s (*voice*)	E_1	3802W	BC_1	1.0
N_2	10 *channels*	μ_2	1/120s (*voice*)	E_2	300W	BC_2	0.8

4.2 Performance Measurement

In this Section, we define the performance measurements used to evaluate the system performance. The mean carried traffic is computed by the following equation

$$O_e^a = \sum_{x\in S; e=1,2; a=AR1,AR2\in A(x)} \left(\sum_{j=1}^{2} \lambda_j + \sum_{j=1}^{2} \sum_{i=1}^{2} n_{ij}\mu_i \right) \pi_x \tag{4.1}$$

where $\pi_x; \forall x \in S$ is the continuous time Markov chain steady state probability distribution under the optimal policy.

The probability of arrival of a new type i service connection seeking admission into a RAT is blocked is called new connection blocking probability of service class i. From the viewpoint of a network operator, the connection blocking probability should be as low as possible such that more connections are accommodated in the wireless system and radio resources are efficiently utilized. Thus, given O_e^a , we can derive new connection blocking probability of service class i^{th} by using the following equation.

$$Pb_i = 1 - \frac{O_i^a}{\lambda_i} \tag{4.2}$$

The bandwidth utilization is defined as the ratio between the mean number of occupied channels and the total number of channels. The utilization of the j^{th} RAT is computed by the following equation.

$$U_j = \frac{1}{N_j} \sum_{x \in S; a \in A(x); \forall i; \forall j; n_{ij} > 0} b_i n_{ij} \pi_x \tag{4.3}$$

4.3 Simulation Scenarios

The performance of the system is evaluated under the following three different scenarios showed in Table 4.2.

Table 4.2: System Scenarios

Scenario I				Scenario II				Scenario III			
Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
b_1	2 *channels*	ω_1	0.8	ρ_1	5	ω_1	0.8	b_1	2 *channels*	ρ_1	5
b_2	1 *channel*	ω_2	0.2	ρ_2	3	ω_2	0.2	b_2	1 *channel*	ρ_2	3

- Scenario-I is for varying traffic intensity of class-1 and class-2 calls that gives a comprehensive view of the expected load on a HetNet. The traffic intensity describes the number of call requests received by the fixed network elements, in a unit area element during a time interval.

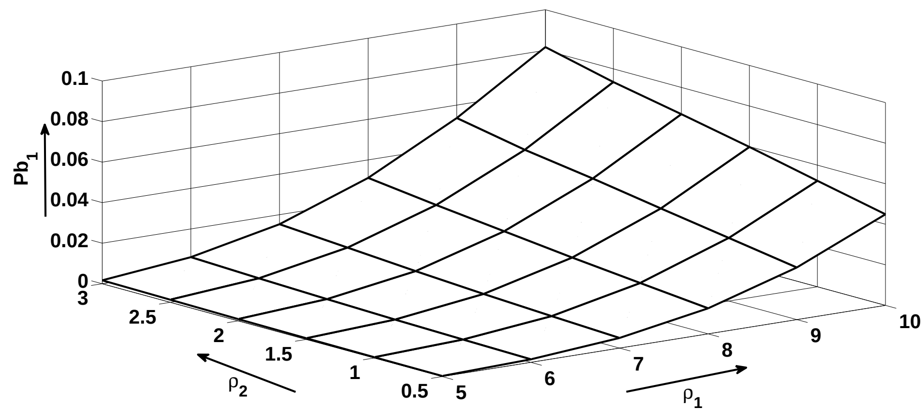
- Scenario-II is for varying required bandwidths for class-1 and class-2 calls. This scenario is to illustrate how varying bandwidth requirements are satisfied under limited resources and the impact on the blocking probability.

- Scenario-III is for varying weight for energy consumption cost in the total cost. This scenario is to illustrate the relative importance of each objective cost function on the the system performance.

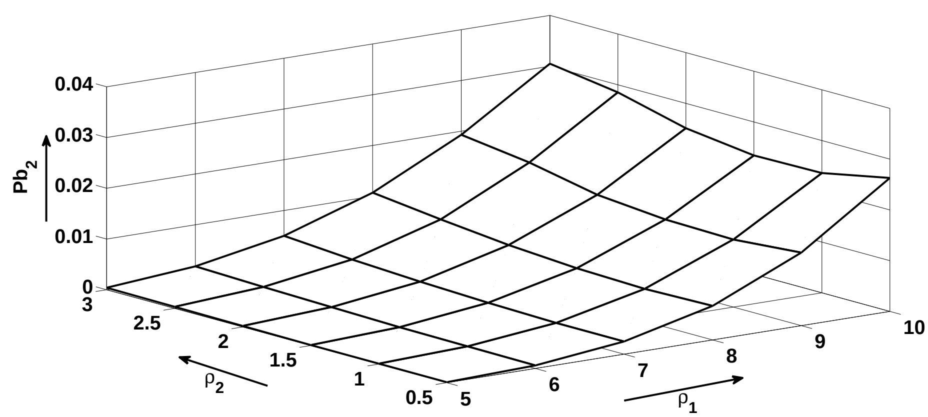
4.4 Simulation Results

We evaluate the performance of the system under the three above mentioned scenarios.

4.4.1 Analysis of Results for Scenario I



(a)

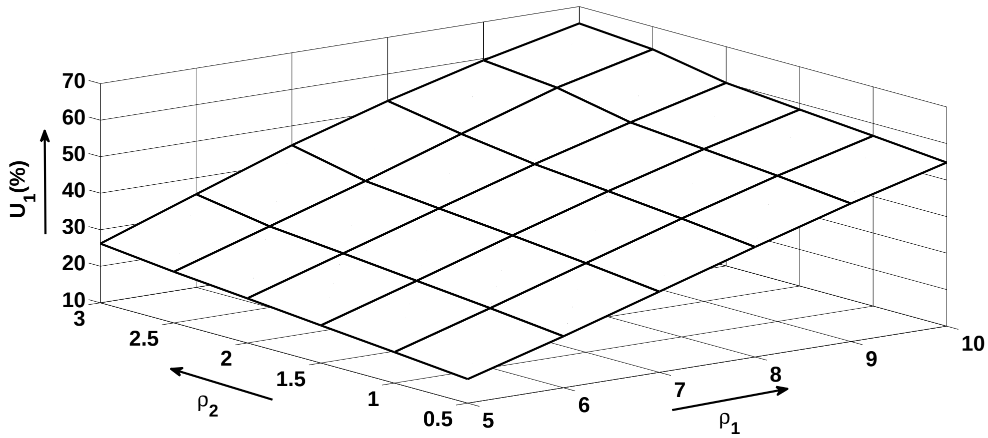


(b)

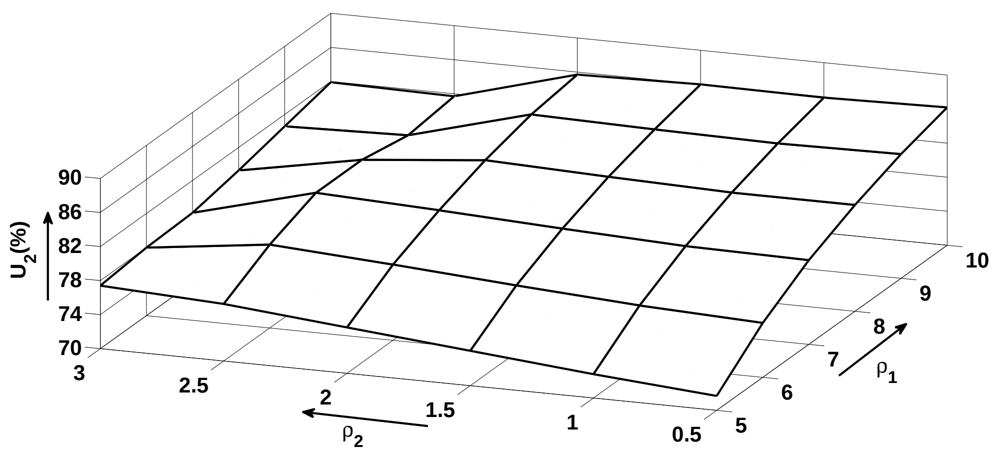
Figure 4.1: Scenario I: Blocking probability versus call intensities (a) blocking probability for class-1 call and (b) blocking probability for class-2.

In Figure 4.1, it can be observed that the higher the call intensities, the higher the blocking probabilities. This is attributed to the fact that the system capacity is fixed. It is also observed that class-2 call intensity does not greatly impact on the blocking probabilities. This is due to the fact that the system blocks less class-2 calls due to its lower call intensity. On the other hand, the blocking probabilities of both calls sharply increase when class-1 call intensity is high (from 8 to 10), meaning that the system has less channels to accept new calls in a fast way. Figure 4.1 (a) shows that the optimal policy blocks more class-1 calls compared to class-2 calls in figure 4.1 (b). This is attributed to the fact that class-1 calls require more bandwidth which make RATs consume more energy.

Figure 4.2 (a) and (b) show that when the values of ρ_1 and ρ_2 are smaller, the initial RAT selection policy decides that it is better to accept more calls in less energy consuming RAT (i.e RAT2), in order to save the overall energy. For this reason, the utilization of RAT2 is far better than that of RAT1 at this step. But when both call intensities are getting high, in order to tackle the traffic volume, the optimal policy starts taking more calls in RAT1 which makes its utilization high. Figure 4.2 (b) shows that the RAT2 utilization goes down and figure 4.2 (a)shows that the RAT1 utilization goes up when both call intensities are high. This is attributed to the fact that the optimal policy accepts more calls in RAT1 due to its higher capacity.



(a)



(b)

Figure 4.2: Scenario I: RAT utilization versus call intensities (a) RAT1 Utilization and (b) RAT2 Utilization.

In Figure 4.3, it can be observed that the optimal cost increases as the traffic intensities increase. This is attributed to the fact that the blocking probabilities of both calls increase because of the fact that the wireless channel is not free to accept new incoming call in a fast way.

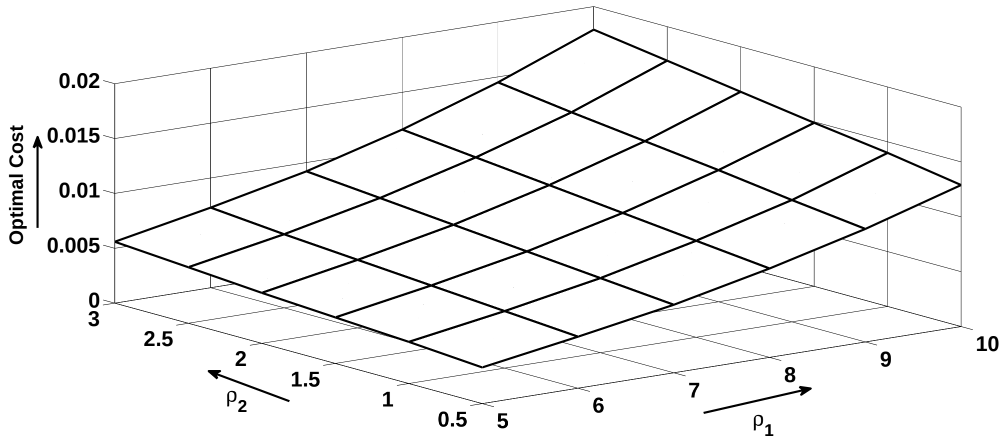
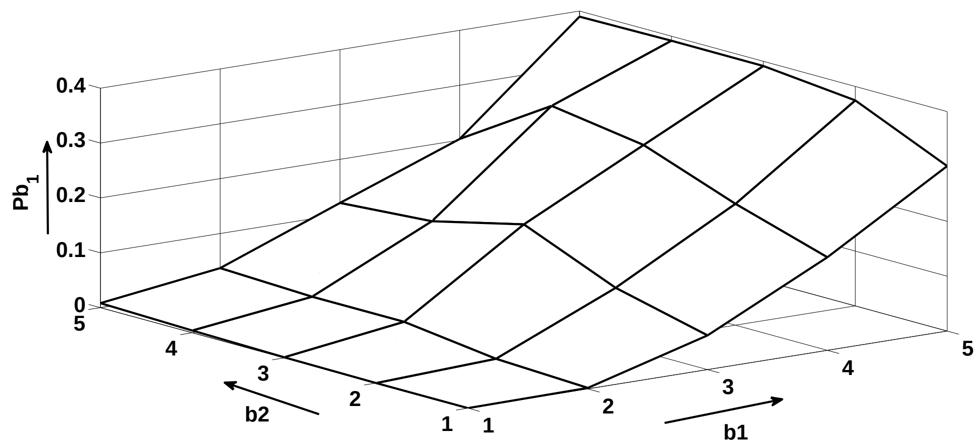


Figure 4.3: Scenario I: Optimal cost versus call intensities.

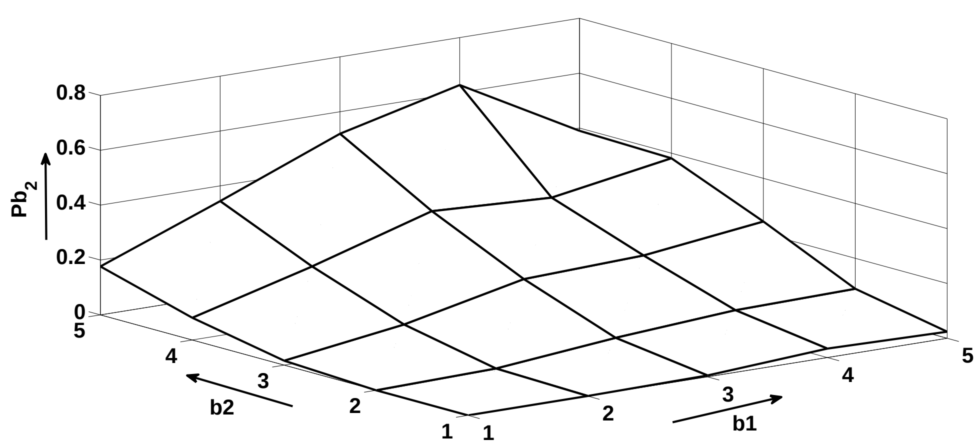
4.4.2 Analysis of Results for Scenario II

In Figure 4.4, it can be observed that when the bandwidths of each call increases, there will be less channels available for admitting new calls, therefore blocking probabilities of each call is high compare to figure 4.1.

Figures 4.4 (a)and (b) also show that both class-1 and class-2 blocking probabilities are low and do not depend on each other's bandwidth when their bandwidths are low, i.e b_1 and b_2 = 1 or 2. But when the required bandwidth of class-2 call is set at a high value (5) and the bandwidth of class-1 call grows fast, the blocking probability of class-2 call in figure 4.4 (b) goes high sharply compared to class-1 blocking probability in figure 4.4 (a). The reason behind this trend is that class-2 call is a lower priority call (i.e blocking cost is equal to 0.8) compared to class-1 call (i.e blocking cost is equal to 1).



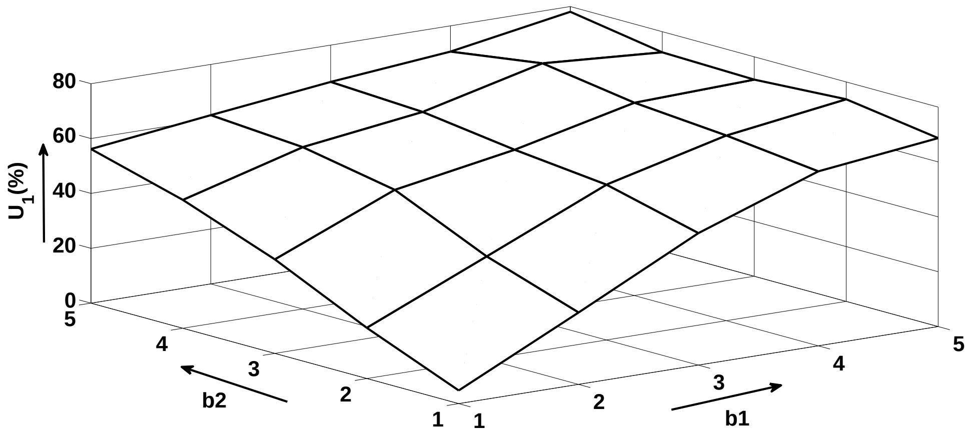
(a)



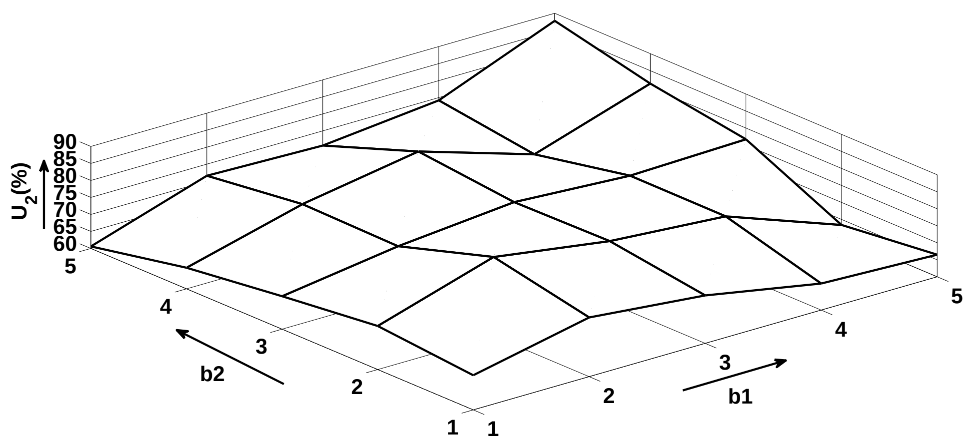
(b)

Figure 4.4: Scenario II: Blocking probability versus required bandwidths (a) blocking probability for class-1 call and (b) blocking probability for class-2 call.

In Figure 4.5, it can be observed that the utilization of RAT2 is far better than that of RAT1, in particular when the required bandwidth of each call is low. This is due to the fact that RAT2 consumes less energy compared to RAT1, which forces the optimal policy to choose RAT2 instead of RAT1. But the optimal policy gradually



(a)



(b)

Figure 4.5: Scenario II: RAT utilization versus required bandwidths (a) RAT1 Utilization and (b) RAT2 Utilization.

accepts more and more calls into RAT1 as the required bandwidth of each call increases. This is attributed to the fact that due to its less capacity (10 channels), RAT2 is unable to accept more high bandwidth calls, therefore RAT1 utilization is high in Figure 4.5 (a). On the other hand, Figure 4.5 (b) shows that the utilization of

RAT2 goes up and down frequently as it has less capacity, which leaves more or less unused channels depending on the multiple of class-1 and class-2 calls assigned into it.

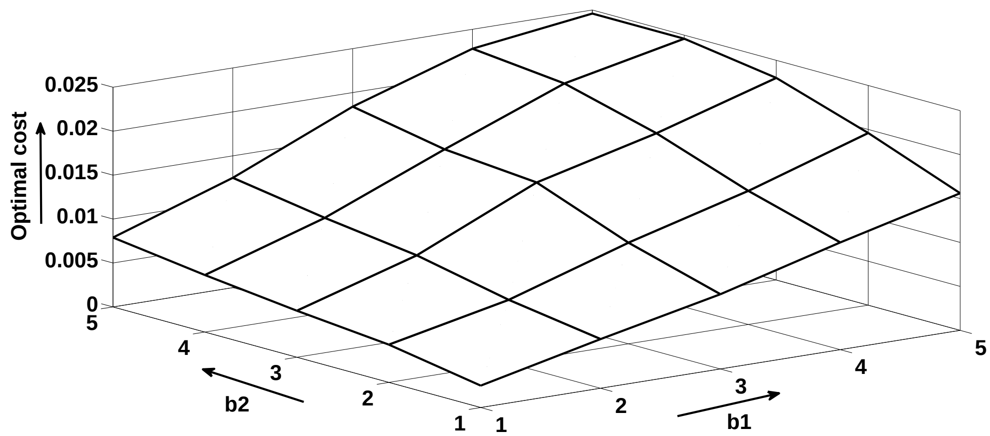


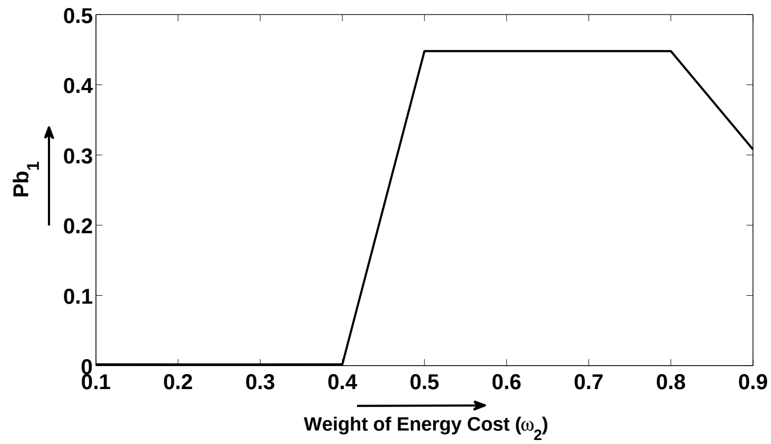
Figure 4.6: Scenario II: Optimal cost versus required bandwidths.

In Figure 4.6, it is observed that the optimal policy cost increases as the required bandwidth increases for both calls. This is attributed to the fact that high bandwidth calls increases the blocking probability. Figure 4.6 also shows that the optimal cost increases sharply when the blocking probability of both calls in figure 4.4 (a)and (b) increases rapidly for the case of higher bandwidths calls, thereby there are less channels remaining on the system for accepting new calls fast.

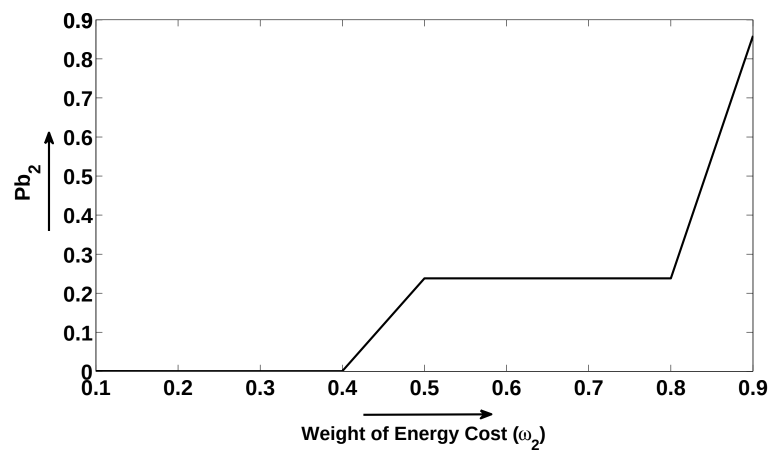
4.4.3 Analysis of Results for Scenario III

In Figure 4.7, it is observed that the optimal policy equally accepts both type of incoming calls (class-1 and class-2) when the weight of energy cost (ω_2) is set from 0.1 to 0.4. However, when we give more emphasis on energy efficiency, i.e

$\omega_2 \geq 0.5$, the optimal policy starts to reject both type of incoming calls in order to save energy.



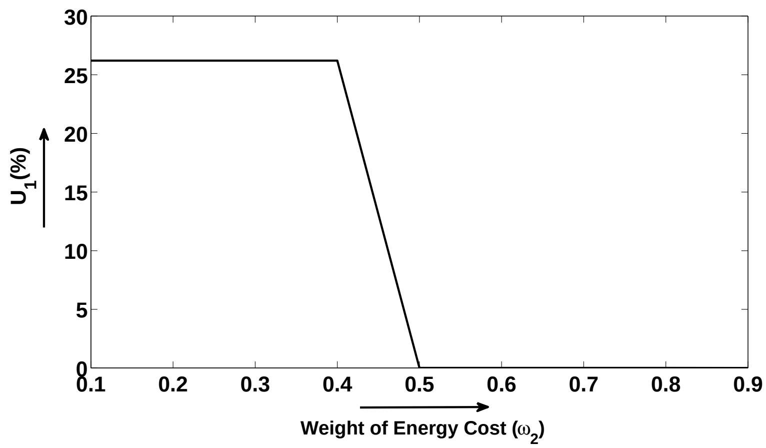
(a)



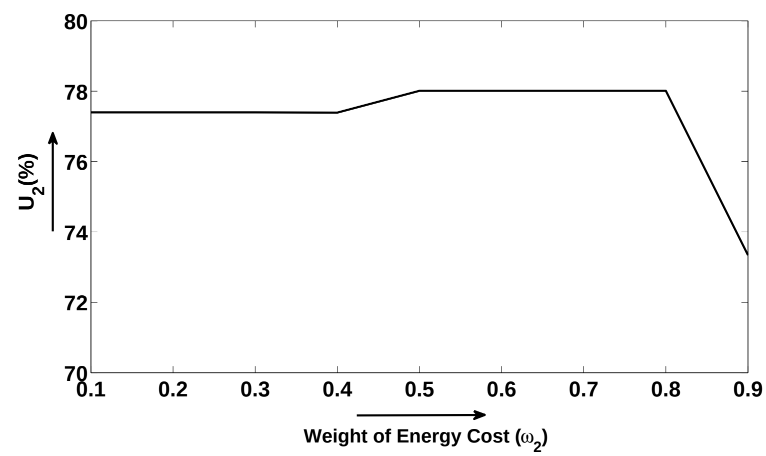
(b)

Figure 4.7: Scenario III: Blocking probability versus weight of energy consumption cost in total cost (a) blocking probability for class-1 call and (b) blocking probability for class-2 call.

It can also be observed from the Figures 4.7 (a) and (b) that more class-1 calls are blocked by the optimal policy. The reason behind this trend is that the amount of bandwidth required by class-1. However, for $\omega_2 = 0.9$, Figure 4.7 reports that the class-1 call acceptance rate is higher than the class-2 call acceptance rate by the RATs.



(a)



(b)

Figure 4.8: Scenario III: RAT utilization versus weight of energy consumption cost in total cost (a) RAT1 Utilization and (b) RAT2 Utilization.

Figure 4.8 shows that the optimal policy utilizes more channels of RAT2 (77%)

than that of RAT1 (26%). This is due to the fact that RAT1 requires far more power than RAT2 does for operating, resulting to energy consumption saving. When $0.5 \leq \omega_2 \geq 0.8$, RAT1 channel utilization reaches almost 0% and RAT2 channel utilization increases slightly as more calls are carried by it that were supposed to be served by RAT1. In Figure 4.8 (b), it can also be observed that the channel utilization of RAT2 decreases sharply when ω_2 is set to its maximum value.

Figure 4.9 shows that the optimal cost increases when ω_2 is set to the low value (0.1) to mid value (0.5). This is due to the fact that the system accepts more calls, but starts decreasing when ω_2 is set to a high value because the system starts blocking more calls from this level of weight in order to save the energy consumption.

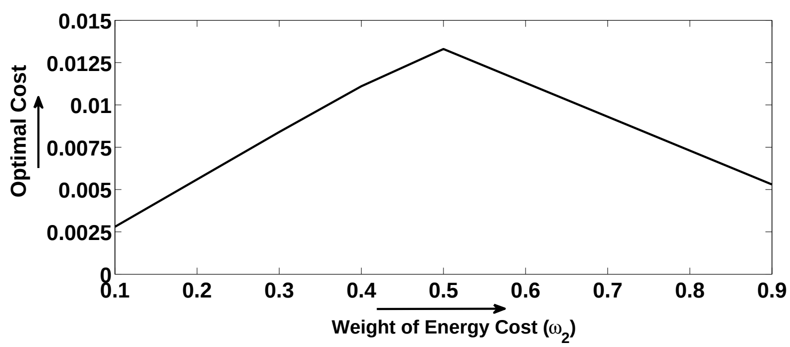


Figure 4.9: Scenario III: Optimal cost versus weight for energy consumption cost in total cost.

4.5 Analysis of the Optimal Structure of Initial RAT Selection Policy

This section provides some results illustrating the behaviour our proposed optimal RAT selection policies under the system configuration shown in Table4.3.

The focus is on how the policy allocates different calls over the existing RATs under different loads (existing number of class-1 and class-2 calls in each RAT) in order to get optimality.

Table 4.3: System Configuration for Optimal Structure

Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
N_1	20 *channels*	μ_1	1/120s (*voice*)	b_1	2 *channels*	ρ_1	5	ω_1	0.8
N_2	10 *channels*	μ_2	1/120s (*voice*)	b_2	1 *channel*	ρ_2	3	ω_2	0.2

4.5.1 Analysis of Class-1 Call Accepted by the Optimal JCAC:

We adopt the following convention (in Tables 4.4 to 4.14) to analyze the structure of the optimal policy for class-1 call:

- '+' denotes class-1 call accepted into RAT1.
- '*' denotes class-1 call accepted into RAT2.
- 'x,y' means class-1 call usually accepted into RAT2. But, RAT1 starts taking class-1 call when number of class-1 calls in RAT1 is equal to x or higher and the number of class-2 calls in RAT1 is equal to y or lower, i.e. (x, y) or (x+1, y-1).
- 'B' denotes blocking of a call.

The results are captured in Tables 4.4 to table 4.14, it can be observed that when the number of class-2 call in RAT2 is 0, regardless of the RAT1 resource occupancy, the optimal JCAC decides to accept class-1 call in RAT2 (*) until RAT2 resources are fully occupied.

Table 4.4 shows that when the number of class-2 call in RAT2 is greater than zero, and the radio resource occupancy of RAT1 is low (note that RAT1 load is less than 10 out of 20), the optimal JCAC decides to accept class-1 call in RAT1 (+).

Table 4.4: When RAT1 channel load is less than 10

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	+	+	+	+	+	+	+	+	+	+
	1	*	+	+	+	+	+	+	+	+		
	2	*	+	+	+	+	+	+				
	3	*	+	+	+	+						
	4	*	+	+								
	5	+										

From Tables 4.5 to 4.9, it can be observed that when the radio resource occupancy of RAT1 is moderate (i.e RAT1 load is from 10 to 14), the optimal JCAC decides to accept class-1 call in RAT2(*) but it again accept class-1 call in RAT1(+) when total number of calls (class-1 calls plus class-2 calls) in RAT1 becomes low.

Table 4.5: When RAT1 channel load is 10

No of class-1 calls in RAT2	No of class-2 calls in RAT2: 0	1	2	3	4	5	6	7	8	9	10
0	*	2,6	+	+	+	+	+	+	+	+	+
1	*	+	+	+	+	+	+	+	+		
2	*	+	+	+	+	+	+				
3	*	+	+	+	+						
4	*	+	+								
5	+										

Table 4.6: When RAT1 channel load is 11

No of class-1 calls in RAT2	No of class-2 calls in RAT2: 0	1	2	3	4	5	6	7	8	9	10
0	*	4,3	3,5	1,9	+	+	+	+	+	+	+
1	*	3,5	2,7	+	+	+	+	+	+		
2	*	2,7	+	+	+	+	+				
3	*	+	+	+	+						
4	*	+	+								
5	+										

Table 4.7: When RAT1 channel load is 12

No of class-1 calls in RAT2 \ No of class-2 calls in RAT2	0	1	2	3	4	5	6	7	8	9	10
0	*	5,2	5,2	3,6	2,8	+	+	+	+	+	+
1	*	5,2	4,4	2,8	1,10	+	+	+	+		
2	*	2,8	2,8	+	+	+	+				
3	*	2,8	2,8	+	+						
4	*	+	+								
5	+										

Table 4.8: When RAT1 channel load is 13

No of class-1 calls in RAT2 \ No of class-2 calls in RAT2	0	1	2	3	4	5	6	7	8	9	10
0	*	6,1	6,1	5,3	4,5	2,9	1,11	+	+	+	+
1	*	6,1	6,1	4,5	3,7	+	+	+	+		
2	*	5,3	5,3	2,9	2,9	+	+				
3	*	4,5	4,5	+	+						
4	*	+	+								
5	+										

Table 4.9: When RAT1 channel load is 14

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	7,0	7,0	5,4	5,4	3,8	3,8	+	+	+	+
	1	*	7,0	7,0	5,4	5,4	2,10	2,10	+	+		
	2	*	6,2	6,2	4,6	4,6	+	+				
	3	*	5,4	6,2	+	+						
	4	*	+	+								
	5	+										

Table 4.10 to 4.13 show that when the radio resource occupancy of RAT1 is high (i.e RAT1 load is from 15 to 18), the optimal JCAC decides to accept class-1 call in RAT2(*) until RAT2 resources are fully occupied, but it also gradually accepts class-1 call in RAT1(+) when total number of calls (i.e class-1 calls plus class-2 calls) in RAT1 become low.

Table 4.10: When RAT1 channel load is 15

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	*	*	7,1	6,3	5,5	5,5	2,1	2,11	+	+
	1	*	*	*	6,3	6,3	4,7	4,7	+	+		
	2	*	7,1	7,1	5,5	5,5	+	+				
	3	*	7,1	7,1	+	+						
	4	*	+	+								
	5	+										

Table 4.11: When RAT1 channel load is 16

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	8,0	8,0	7,2	7,2	6,4	6,4	4,8	4,8	+	+
	1	*	8,0	8,0	7,2	7,2	5,6	5,6	+	+		
	2	*	8,0	8,0	6,4	6,4	+	+				
	3	*	8,0	8,0	+	+						
	4	*	+	+								
	5	+										

Table 4.12: When RAT1 channel load is 17

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	*	*	8,1	8,1	7,3	6,5	5,7	5,7	+	+
	1	*	*	*	8,1	7,3	6,5	6,5	+	+		
	2	*	*	8,1	7,5	7,5	+	+				
	3	*	8,1	8,1	+	+						
	4	*	+	+								
	5	+										

Table 4.13: When RAT1 channel load is 18

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	*	9,0	9,0	8,2	7,4	7,4	6,6	5,8	+	+
	1	*	9,0	9,0	8,2	8,2	7,4	7,4	+	+		
	2	*	9,0	9,0	8,2	8,2	+	+				
	3	*	9,0	9,0	+	+						
	4	*	+	+								
	5	+										

It is also observed in Table 4.14 that when there is no more occupancy to accept class-1 call in RAT1(i.e when RAT1 load is from 19 to 20), the optimal JCAC decides to accept class-1 call in RAT2(*) until RAT2 resources are fully occupied. The JCAC blocked (B) the class-1 call when there is no more occupancy to accept class-1 call in either RAT1 or RAT2.

Table 4.14: When RAT1 channel load is greater than 18

		No of class-2 calls in RAT2										
		0	1	2	3	4	5	6	7	8	9	10
No of class-1 calls in RAT2	0	*	*	*	*	*	*	*	*	*	B	B
	1	*	*	*	*	*	*	*	B	B		
	2	*	*	*	*	*	B	B				
	3	*	*	*	B	B						
	4	*	B	B								
	5	B										

4.6 Summary

In this chapter, we have considered three different scenarios to show the usability of our proposed network selection algorithm. The most promising part of scenario III is finding a rule that strives to maximize the system capacity while selecting the network that consumes less energy. Simulation results of scenario III is also showing that the optimal policy selects the less energy consuming RAT often when more weight (50% or more) is given for energy consumption cost in the total cost function. Moreover,we have analyzed the structure of the optimal policy that shows when and which call goes to which RAT in order to get optimality. Thus, the simulation results validate our proposed SMDP based RAT selection scheme.

Chapter 5

Conclusion

We have proposed an optimization model based on the Semi-Markov Decision Process (SMDP) framework for the problem of selecting the initial Radio Access Technology (RAT) in co-located wireless networks. Our optimal initial RAT selection method considers a cost function that involves a blocking cost and an energy consumption cost associated with different weights to support the optimal JCAC decision. For the studied scenario with two co-located wireless networks, our simulation results demonstrate that variations in the weights of blocking cost and energy consumption cost can greatly impact both the system capacity and the network energy consumption.

For the scenarios investigated in this thesis, our SMDP model generated 27,283 pairs of state-action for the two co-located wireless networks. As future work, we believe that our proposed SMDP-based model can be extended to support more sophisticated HetNets architectures (i.e. with more than two co-located wireless networks and several service classes), which of course will involve a huge number of state actions that should be deal with. Another challenge will be the study of inter-RAT handover for such architecture.

Bibliography

[1] M. Alkhawlani, F. Al-Akwaa, and A. Mohsen, "Intelligent joint admission control for next generation wireless networks," *International Journal of Advanced Computer Science and Applications*, vol. 3, no. 4, pp. 50–55, 2012.

[2] "3GPP, 3rd Generation Partnership Project; Technical Specification Group (TSG) Terminals; Multi-mode UE issues Categories, principles and procedures (Release 1999)," 2012. [Online]. Available: http://www.3gpp.org

[3] J. Sachs and M. Olsson, "Access network discovery and selection in the evolved 3gpp multi-access system architecture," *European Transactions on Telecommunications*, vol. 21, no. 6, pp. 544–557, 2010.

[4] M. Alkhawlani and A. Ayesh, "Access network selection based on fuzzy logic and genetic algorithms," *Advances in Artificial Intelligence*, vol. 8, no. 1, p. 1, 2008.

[5] R. Bendlin, V. Chandrasekhar, R. Chen, A. Ekpenyong, and E. Onggosanusi, "From homogeneous to heterogeneous networks: A 3gpp long term evolution rel. 8/9 case study," in *IEEE 45th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, USA, March 23-25,2011, pp. 1–5.

[6] J. Peng, H. Xian, X. Zhang, and Z. Li, "Context-aware vertical handoff decision scheme in heterogeneous wireless networks," in *IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom),*, Changsha, China, November 16-18, 2011, pp. 1375–1380.

[7] X. Wu, B. Murherjee, and D. Ghosal, "Hierarchical architectures in the third-generation cellular network," *IEEE on Wireless Communications,*, vol. 11, no. 3, pp. 62–71, 2004.

[8] A. Guerrero-Ibáñez, J. Contreras-Castillo, A. Barba, and A. Reyes, "A qos-based dynamic pricing approach for services provisioning in heterogeneous wireless access networks," *Pervasive and Mobile Computing*, vol. 7, no. 5, pp. 569–583, 2011.

[9] O. Falowo and H. Chan, "Joint call admission control algorithms: Requirements, approaches, and design considerations," *Computer Communications*, vol. 31, no. 6, pp. 1200–1217, 2008.

[10] K. Piamrat, A. Ksentini, J. Bonnin, and C. Viho, "Radio resource management in emerging heterogeneous wireless networks," *Computer Communications*, vol. 34, no. 9, pp. 1066–1076, 2011.

[11] J. Ernst, "Energy-efficient next-generation wireless communications," *Handbook of Green Information and Communication Systems*, p. 371, 2012.

[12] M. López-Benítez and J. Gozalvez, "Common radio resource management al-

gorithms for multimedia heterogeneous wireless networks," *IEEE Transactions on Mobile Computing,*, vol. 10, no. 9, pp. 1201–1213, 2011.

[13] V. Atanasovski, V. Rakovic, and L. Gavrilovska, "Efficient resource management in future heterogeneous wireless networks: The riwcos approach," in *IEEE Military Communications Conference (MILCOM), 2010*, 2010, pp. 2286–2291.

[14] K. Suleiman, H. Chan, and M. Dlodlo, "Issues in designing joint radio resource management for heterogeneous wireless networks," in *IEEE 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM),Kunming, China, August 810*, 2011, pp. 1–5.

[15] S. Kajioka, N. Wakamiya, and M. Murata, "Autonomous and adaptive resource allocation among multiple nodes and multiple applications in heterogeneous wireless networks," *Journal of Computer and System Sciences*, 2011.

[16] X. Pei, T. Jiang, D. Qu, G. Zhu, and J. Liu, "Radio-resource management and access-control mechanism based on a novel economic model in heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 3047–3056, 2010.

[17] D. Ngo and T. Le-Ngoc, "Distributed resource allocation for cognitive radio networks with spectrum-sharing constraints," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3436–3449, 2011.

[18] L. Giupponi, R. Agusti, J. Pérez-Romero, and O. Sallent, "A novel joint radio

resource management approach with reinforcement learning mechanisms," in *24th IEEE International Conference on Performance, Computing, and Communications , Phoenix, AZ, USA, April 7-9, 2005*, pp. 621–626.

[19] O. Falowo and H. Chan, "Adaptive bandwidth management and joint call admission control to enhance system utilization and qos in heterogeneous wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 3, p. 2, 2007.

[20] J. Pérez-Romero, O. Sallent, and R. Agustí, "Policy-based initial rat selection algorithms in heterogeneous networks," in *The 7th International Conference on Mobile and Wireless Communications Network (MWCN), Marrakesh, Morocco, September 19-21, 2005*.

[21] O. Falowo and H. Anthony Chan, "Dynamic rat selection for multiple calls in heterogeneous wireless networks using group decision-making technique," *Computer Networks*, 2011.

[22] K. Haldar, C. Ghosh, and D. Agrawal, "Dynamic spectrum access and network selection in heterogeneous cognitive wireless networks," *Pervasive and Mobile Computing*, 2012.

[23] E. Stevens-Navarro, Y. Lin, and V. Wong, "An mdp-based vertical handoff decision algorithm for heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology,*, vol. 57, no. 2, pp. 1243–1254, 2008.

[24] S. Jin, W. Xuanli, and S. Xuejun, "Load balancing algorithm with multi-service

in heterogeneous wireless networks," in *IEEE 6th International ICST Conference on Communications and Networking,Harbin, China, August 17-19, 2011,*, pp. 703–707.

[25] M. Lucas-Estañ, J. Gozalvez, and J. Sanchez-Soriano, "Integer linear programming optimization of joint rrm policies for heterogeneous wireless systems," *Computer Networks*, vol. 56, no. 1, pp. 112–126, 2012.

[26] M. Porjazoski and B. Popovski, "Radio access technology selection algorithm for heterogeneous wireless networks based on service type, user mobility and network load," in *10th International Conference on Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS),Nis, Serbia, October 5-8, 2011*, vol. 2, pp. 475–478.

[27] L. Mohamed, C. Leghris, and A. Abdellah, "A survey and comparison study on weighting algorithms for access network selection," in *IEEE 9th Annual Conference on Wireless On-demand Network Systems and Services (WONS), Courmayeur, Italy, January 9-11, 2012*, pp. 35–38.

[28] S. Zhu, F. Liu, Y. Qi, Z. Chai, and J. Wu, "Immune optimization algorithm for solving joint call admission control problem in next-generation wireless network," *Engineering Applications of Artificial Intelligence*, 2012.

[29] P. Si, H. Ji, and F. Yu, "Optimal network selection in heterogeneous wireless multimedia networks," *Wireless Networks*, vol. 16, no. 5, pp. 1277–1288, 2010.

[30] G. Kesidis, J. Walrand, and C. Chang, "Effective bandwidths for multiclass

markov fluids and other atm sources," *IEEE/ACM Transactions on Networking (TON)*, vol. 1, no. 4, pp. 424–428, 1993.

[31] N. Nasser and H. Hassanein, "Dynamic threshold-based call admission framework for prioritized multimedia traffic in wireless cellular networks," in *IEEE Conference on Global Telecommunications (GLOBECOM'04), Dallas, USA, 2004*, vol. 2, pp. 644–649.

[32] H. Tijms, *A first course in stochastic models*. Wiley, $ISBN-13:978-0-12-398650-4$, 2003.