

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**Efficient Joint Call Admission Control and Bandwidth
Management Schemes for QoS Provisioning in
Heterogeneous Wireless Networks**

Olabisi Emmanuel FALOWO



**This thesis is submitted
for the degree of
Doctor of Philosophy in Electrical Engineering
in the Faculty of Engineering and the Built Environment
University of Cape Town**

March 2008

As the candidate's supervisor, I have approved this dissertation for submission.

Name: Professor H. Anthony Chan

Signed by candidate

Signed: signature removed Date: March 17 2008

University of Cape Town

Declaration

I declare that this thesis is my own work. Where collaboration with other people has taken place, or material generated by other researchers is included, the parties and/or materials are indicated in the acknowledgements or are explicitly stated with references as appropriate.

This work is being submitted for the Doctor of Philosophy in Electrical Engineering at the University of Cape Town. It has not been submitted to any other university for any other degree or examination.

First Names: Olabisi Emmanuel

Last Name: Falowo

Name:

Date:

To God for His Great Mercies

University of Cape Town

Abstract

Next generation wireless network (NGWN) will be heterogeneous where different radio access technologies (RATs) coexist. This coexistence of different RATs necessitates joint radio resource management (JRRM) for enhanced QoS provisioning and efficient radio resource utilization. Joint call admission control (JCAC) algorithm is one of the joint radio resource management algorithms. The basic functions of a JCAC algorithm are to decide whether or not an incoming call can be accepted into a heterogeneous wireless network, and to determine *which of the available RATs is most suitable to admit the incoming call*. The objective of a JCAC algorithm is to guarantee the QoS requirements of all accepted calls and at the same time make the best use of the available radio resources. Traditional call admission control algorithms designed for homogeneous wireless networks do not provide a single solution to address the heterogeneous architecture, which characterizes NGWN. Consequently, there is need to develop JCAC algorithms for heterogeneous wireless networks.

The thesis proposes three JCAC schemes for improving QoS and radio resource utilization, which are of primary concerns, in heterogeneous wireless networks. The first scheme combines adaptive bandwidth management and joint call admission control. The objectives of the first scheme are to enhance average system utilization, guarantee QoS requirements of all accepted calls, and reduce new call blocking probability and handoff call dropping probability in heterogeneous wireless networks. The scheme consists of three components namely: joint call admission controller, bandwidth reservation unit, and bandwidth adaptation unit. Using Markov decision process, an analytical model is developed to evaluate the performance of the proposed scheme considering three performance metrics, which are new call blocking probability, handoff call dropping probability, and system utilization. Numerical results show that the proposed scheme improves system utilization and reduces both new call blocking probability and handoff call dropping probability.

The second proposed JCAC scheme minimizes call blocking probability by determining the optimal call allocation policy among the available RATs. The scheme measures the

arrival rates of different classes of calls into the heterogeneous wireless network. Using linear programming technique, the JCAC scheme determines the call allocation policy that minimizes call-blocking probability in the heterogeneous network. Numerical results show that the proposed scheme reduces call-blocking probability in the heterogeneous wireless network.

The third algorithm addresses the problem of highly unbalanced network load in heterogeneous wireless networks where users' preferences are considered in making RAT selection decisions. Independent users' preferences for a RAT often results in unbalanced network load, which in-turn leads to poor overall radio resource utilization and high call blocking probability. To address this problem, a JCAC scheme, which incorporates dynamic pricing to balance traffic load among available RATs in a heterogeneous wireless network, is proposed. By dynamically adjusting the service price in each of the available RATs, the proposed JCAC scheme evens out the unbalanced traffic load caused by independent users' preferences. The JCAC scheme uses fuzzy multiple attribute decision making (MADM) technique to select the most appropriate RAT for each incoming call based on the user's preference. A model is developed to evaluate the overall new call blocking probability, handoff call dropping probability, and percentage load in each RAT in heterogeneous wireless networks. Performance of the proposed JCAC scheme is compared with the performance of a scheme that does not incorporate dynamic pricing. Numerical results show the proposed JCAC scheme improves traffic load distribution and reduces both new call blocking probability and handoff call dropping probability in the heterogeneous network.

The three JCAC schemes proposed in this thesis are applicable to a single operator with multiple RATs. The JCAC schemes are also applicable to different operators of RATs provided there is cooperation among the operators. The objectives of the three JCAC schemes are to enhance connection level QoS and improve radio resource utilization in heterogeneous wireless networks. Numerical results validate the effectiveness of the three schemes.

Acknowledgements

My profound gratitude goes to my supervisor, Professor H. Anthony Chan for his guidance and advice throughout the period of my PhD research at the University of Cape Town.

I would like to acknowledge the financial support from German Academic Exchange Service (DAAD). The help and support of Mr. Ventura, Professor M. E. Dlodlo, Dr E. Bejide, and other staff of the Department of Electrical Engineering, University of Cape Town are highly appreciated.

I am grateful to the members of Communication Research Group, University of Cape Town for creating a conducive research atmosphere. I also appreciate the support of Mr. S. A. Adelana, Mr Oloruntoyin, Mr and Mrs A. O. Bello, Dr S. C. Enemour, Mr. A.J. Isafiade, and Mr. B. Omolofe. I am grateful to the staff of the Department of Electrical and Electronic Engineering, Federal University of Technology Akure.

Special thanks to my parents, brothers and sisters for their constant support, love, encouragement, and sacrifice.

Finally, I am forever grateful to God for the gift of life.

<u>Declaration.....</u>	<u>iii</u>
<u>Abstract.....</u>	<u>v</u>
<u>Acknowledgements.....</u>	<u>vii</u>
<u>List of Figures.....</u>	<u>xii</u>
<u>Glossary.....</u>	<u>xix</u>
<u>1 Introduction.....</u>	<u>1</u>
1.1 Need for Joint Call Admission Control Algorithms.....	2
1.1.1 <i>Efficient Radio Resource Utilization.....</i>	2
1.1.2 <i>Consistent QoS Provisioning</i>	2
1.1.3 <i>Overall Network Stability.....</i>	3
1.1.4 <i>Enhancement of Users' Satisfaction</i>	3
1.2 Problems in the Existing JCAC Schemes.....	4
1.2.1 <i>Single-Criterion JCAC Algorithms</i>	4
1.2.2 <i>Multiple-Criteria JCAC Algorithms.....</i>	5
1.3 Research Objectives	6
1.4 Scope of Research.....	7
1.5 Contributions.....	7
1.5.1 <i>A critical survey of joint call admission control in heterogeneous wireless networks.....</i>	7
1.5.2 <i>Combination of adaptive bandwidth management and JCAC to improve system utilization and quality of service in heterogeneous wireless networks</i>	7
1.5.3 <i>Development of analytical model to evaluate new call blocking probability, handoff dropping probability, and system utilization in heterogeneous wireless networks</i>	8
1.5.4 <i>Reduction of call-blocking probability through optimal allocation of calls</i>	8
1.5.5 <i>Use of dynamic pricing to balance traffic load in multiple-criteria joint call admission control of heterogeneous wireless networks.....</i>	8
1.6 Thesis Outline	10
<u>2 Heterogeneous Wireless Networks</u>	<u>12</u>
2.1 Motivation for Heterogeneous Wireless Networks.....	12
2.1.1 <i>Limitation of a Single RAT.....</i>	12
2.1.2 <i>Evolution of Wireless Technologies.....</i>	13

2.1.3	<i>Users' Demand for Advanced Service and Complementary Features of Different RATs</i>	13
2.2	Challenges of Heterogeneous Wireless Networks	13
2.2.1	<i>Joint Call Admission Control</i>	14
2.2.2	<i>Seamless Vertical Handoff</i>	14
2.2.3	<i>Common Authentication and Authorization</i>	14
2.2.4	<i>Network Security</i>	14
2.2.5	<i>Unified Billing</i>	15
2.2.6	<i>Availability of Multimode Terminals</i>	15
2.3	Radio Resource Management in Heterogeneous Networks	16
2.4	Summary	17
<u>3</u>	<u>Review of JCAC and Bandwidth Management Schemes</u>	<u>18</u>
3.1	JCAC algorithms	18
3.1.1	<i>Requirements for JCAC algorithms</i>	18
3.1.2	<i>Classification of JCAC Algorithms</i>	21
3.1.3	<i>Multiple-Criteria JCAC Algorithms</i>	38
3.1.4	<i>JCAC Algorithms Design Considerations</i>	42
3.2	Bandwidth Management Schemes	48
3.2.1	<i>Complete Sharing</i>	49
3.2.2	<i>Complete Partitioning</i>	49
3.2.3	<i>Handoff Call Prioritization</i>	50
3.2.4	<i>Service-Class Prioritization</i>	54
3.3	Summary	55
<u>4</u>	<u>Adaptive Bandwidth Management and JCAC Scheme</u>	<u>56</u>
4.1	Introduction	56
4.2	System Model	57
4.3	Components of the AJCAC Scheme	60
4.3.1	<i>Joint Call Admission Controller</i>	60
4.3.2	<i>Threshold-Based Bandwidth Reservation Unit</i>	64
4.3.3	<i>Bandwidth Adaptation Controller</i>	65
4.4	Markov Chain Model of the AJCAC Scheme	66
4.4.1	<i>New Call Blocking Probability</i>	70
4.4.2	<i>Handoff Call Dropping Probability</i>	71
4.4.3	<i>Overall System Utilization</i>	71
4.5	Performance Evaluation	72

4.5.1	<i>Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability.....</i>	73
4.5.2	<i>Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability.....</i>	76
4.5.3	<i>Impact of varying the call arrival rate on the normalized average system utilization.....</i>	78
4.6	Summary.....	81
5	<u>Reduction of Call Blocking Probability through Optimal Allocation of Calls.....</u>	82
5.1	Introduction.....	82
5.2	System Model.....	82
5.3	Proposed Optimal JCAC Scheme.....	85
5.3.1	<i>Joint Call Admission Controller.....</i>	86
5.3.2	<i>Arrival Rate Measurement Unit.....</i>	86
5.3.3	<i>Bandwidth Reservation Unit.....</i>	86
5.3.4	<i>Optimal Policy Determination Unit.....</i>	86
5.4	Markov Chain Model.....	89
5.4.1	<i>New Call Blocking Probability.....</i>	90
5.4.2	<i>Handoff Call Dropping Probability.....</i>	91
5.5	Performance Evaluation.....	92
5.6	Summary.....	102
6	<u>Dynamic Pricing for Load balancing in a Heterogeneous Wireless Network Using a Multiple-Criteria JCAC Scheme.....</u>	103
6.1	Introduction.....	103
6.2	Users' Preferences, Load Balancing, and Pricing.....	103
6.2.1	<i>Users' Preferences.....</i>	103
6.2.2	<i>Load Balancing.....</i>	106
6.2.3	<i>Pricing.....</i>	107
6.3	System Model.....	108
6.4	Proposed Scheme.....	111
6.4.1	<i>Joint Call Admission Controller.....</i>	111
6.4.2	<i>Call Admission Rate Measurement Unit.....</i>	120
6.4.3	<i>Price Update Unit.....</i>	120
6.4.4	<i>Bandwidth Reservation Unit.....</i>	123
6.5	Markov Model.....	124
6.5.1	<i>New Call Blocking Probability.....</i>	127

6.5.2	<i>Handoff Call Dropping Probability</i>	127
6.5.3	<i>Percentage Utilization of Each RAT</i>	128
6.6	Performance Evaluation	128
6.6.1	<i>General User Scenario</i>	130
6.6.2	<i>High-Price-Sensitive User Scenario</i>	137
6.6.3	<i>Low-Price-Sensitive User Scenario</i>	140
6.7	Summary	144
7	<u>Conclusion and Future Work</u>	<u>145</u>
7.1	Summary of Contributions	145
7.1.1	<i>Adaptive Bandwidth Management and Joint Call Admission Control to Enhance System Utilization and QoS in Heterogeneous Wireless Networks</i>	145
7.1.2	<i>Reduction of Call Blocking Probability through Optimal Call Allocation Policy</i>	146
7.1.3	<i>Dynamic Pricing for Balancing Traffic Load in Heterogeneous Wireless Network Using Multiple-Criteria JCAC</i>	146
7.2	Future Work	147
7.2.1	<i>Extension of JCAC Scheme to Cover a Non Fully-Overlapping Heterogeneous Wireless Network</i>	147
7.2.2	<i>Consideration of Packet-Level QoS Metrics</i>	147
7.2.3	<i>Consideration of Other Step Size Rules</i>	148
7.2.4	<i>Combination of JCAC with JSS</i>	148
7.2.5	<i>Investigation of the Effect of Variation in Service Price on Operator' Revenue</i>	148

List of Figures

Figure 1-1. Basic functions of JCAC algorithms.....	2
Figure 2-1. JCAC in heterogeneous wireless network.....	12
Figure 2-2. Independent RRM in heterogeneous wireless networks.	16
Figure 2-3. Joint RRM in heterogeneous wireless networks.	16
Figure 3-1. Classification of JCAC algorithm based on RAT selection approaches.	21
Figure 3-2. State transition diagram for integrated UMTS-WLAN network.	31
Figure 3-3. Flow chart of utility-based JCAC [38].....	34
Figure 3-4. Fuzzy logic controller.....	36
Figure 3-5. Fuzzy logic decision strategy [21].	37
Figure 3-6. Two-RAT heterogeneous wireless network.	41
Figure 3-7. Six design consideration for JCAC algorithms.....	43
Figure 3-8. State transition diagram for complete sharing policy.	49
Figure 3-9. State transition diagram for complete partitioning policy: first partition.	50
Figure 3-10. State transition diagram for complete partitioning policy: second partition.....	50
Figure 3-11. State transition diagram for guard bandwidth scheme.	51
Figure 3-12. State transition diagram for fractional guard bandwidth policy.....	52

Figure 4-1. A two-RAT heterogeneous wireless network with co-located cells.	57
Figure 4-2. Proposed JCAC algorithm.....	61
Figure 4-3. Splitting of the arrival process in the group of co-located cells.....	63
Figure 4-4. Bandwidth reservation policy for the proposed scheme.	64
Figure 4-5. Effect of varying the call arrival rate on the new call blocking probability.	74
Figure 4-6. Effect of varying the call arrival rate on the handoff call dropping probability.	74
Figure 4-7. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls.	75
Figure 4-8. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls.	76
Figure 4-9. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls.	77
Figure 4-10. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-2 calls.	77
Figure 4-11. Impact of varying the call arrival rate on the normalized average system utilization for each class of calls.....	79
Figure 4-12. Impact of varying the call arrival rate on the normalized average system utilization for both class-1 and class-2 calls combined.	80
Figure 4-13. Effect of varying the call arrival rate on utilization of each RAT by both classes of calls combined.	81

Figure 5-1. A two-class three-RAT heterogeneous wireless network.	83
Figure 5-2. Proposed optimal JCAC scheme.	85
Figure 5-3 . Bandwidth reservation policy.	86
Figure 5-4. Effect of varying the call arrival rate on new call blocking probability of class-1 and class-2 calls ($\lambda_1=\lambda_2$).	92
Figure 5-5. Effect of varying the call arrival rate on overall new call blocking probability ($\lambda_1=\lambda_2$).	93
Figure 5-6. Effect of varying the call arrival rate on handoff call dropping probability of class-1 and class-2 calls ($\lambda_2=\lambda_1$).	94
Figure 5-7. Effect of varying the call arrival rate on overall handoff call dropping probability ($\lambda_2=\lambda_1$).	94
Figure 5-8. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls for the OJCAC scheme ($\lambda_2=\lambda_1$).	95
Figure 5-9. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls for the OJCAC scheme ($\lambda_2=\lambda_1$).	96
Figure 5-10. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls ($\lambda_2=\lambda_1$).	97
Figure 5-11. Effect of varying the call arrival rate on new call blocking probability of class-1 and class-2 calls ($\lambda_2=2\lambda_1$).	98
Figure 5-12. Effect of varying the call arrival rate on overall new call blocking probability ($\lambda_2=2\lambda_1$).	98

Figure 5-13. Effect of varying the call arrival rate on handoff call dropping probability of class-1 and class-2 calls ($\lambda_2=2\lambda_1$).	99
Figure 5-14. Effect of varying the call arrival rate on overall handoff call dropping probability ($\lambda_2=2\lambda_1$).	100
Figure 5-15. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls for the OJCAC scheme ($\lambda_2=2\lambda_1$).	100
Figure 5-16. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls for the OJCAC scheme ($\lambda_2=2\lambda_1$).	101
Figure 5-17. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls, ($\lambda_2=2\lambda_1$).	102
Figure 6-1. A two-RAT heterogeneous wireless network.	108
Figure 6-2. Proposed JCAC scheme.	111
Figure 6-3. Call request/response procedure.	112
Figure 6-4. Assignment of weight to each of the RAT selection criteria.	112
Figure 6-5. Linguistic terms to fuzzy number conversion scale.	113
Figure 6-6. Proportion of calls admitted into each RAT.....	118
Figure 6-7. Proportion of calls admitted into each RAT.....	119
Figure 6-8. Proposed Fuzzy MADM-based RAT selection method.....	119
Figure 6-9. Price update process.	121

Figure 6-10. Splitting of the arrival process.	122
Figure 6-11. Bandwidth reservation policy.....	123
Figure 6-12. Call blocking/dropping probability versus price update time ($\lambda^n = 1$, z=0.1).	130
Figure 6-13. Call blocking/dropping probability versus price update time ($\lambda^n = 3$, z=0.1).	132
Figure 6-14. Call blocking/dropping probability versus price update time ($\lambda^n = 3$, z=0.2)	133
Figure 6-15. Call blocking/dropping probability versus price update time ($\lambda^n = 3$, z=0.3).	133
Figure 6-16. Percentage of load in each RAT versus time for JCAC without dynamic pricing ($\lambda^n = 3$).....	134
Figure 6-17. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3$, z=0.1).	134
Figure 6-18. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n = 3$, z=0.1).....	135
Figure 6-19. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3$, z=0.2)... ..	135
Figure 6-20. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n = 3$, z=0.2).	136
Figure 6-21. Variation of service price with update time using linear price-update function ($\lambda^n = 3$, z=0.1).	136

Figure 6-22. Variation of service price with update time using exponential price-update function ($\lambda^n = 3, z=0.1$).	137
Figure 6-23. Call blocking/dropping probability versus price update time ($\lambda^n = 3, z=0.1$).	138
Figure 6-24. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3, z=0.1$). ..	138
Figure 6-25. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential function ($\lambda^n = 3, z=0.1$).	139
Figure 6-26. Variation of service price with update time using linear price-update function ($\lambda^n = 3, z=0.1$).	139
Figure 6-27. Variation of service price with update time using exponential price-update function ($\lambda^n = 3, z=0.1$).	140
Figure 6-28. Call blocking/dropping probability versus price update time ($\lambda^n = 3, z=0.1$).	141
Figure 6-29. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3, z=0.1$). ..	141
Figure 6-30. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n = 3, z=0.1$).	142
Figure 6-31. Variation of service price with update time using linear function ($\lambda^n = 3, z=0.1$).	143
Figure 6-32. Variation of service price with update time using exponential function ($\lambda^n = 3, z=0.2$).	143

List of Tables

Table 1-1. Relationship among JCAC algorithm requirements.	20
Table 3-1. Summary of the RAT selection approaches.	22
Table 3-2. Four scenarios for the JCAC algorithm.	27
Table 3-3. Total throughput (mbps) for the two basic policies.	29
Table 3-4. Users' preferences.	40
Table 3-5. Operators' preferences.	42
Table 3-6. Design considerations for JCAC algorithms.	47
Table 3-7. Summary of bandwidth allocation strategies for wireless networks	48
Table 4-1. Simulation parameters.	72
Table 6-1. RAT selection criteria for incoming calls.	116
Table 6-2. Initial and new prices used in the decision matrix.	118
Table 6-3. RAT selection criteria.	129

Glossary

Basic Bandwidth Unit (BBU): The unit of radio resources (such as time slots, code sequence, etc.). It is dependent on the specific technological implementation of the radio interface.

Call: In communication networks, a call is any request to use the network's radio resource for services such as voice, video, web browsing, etc.

Call Admission Control (CAC): The process of regulating the number of calls admitted into a resource-constraint network so that the QoS requirements of admitted calls are guaranteed.

Call Holding Time (CHT): The total length of time of the duration of a call.

CDMA: Acronym for Code Division Multiple Access

Cell: The geographic area encompassing the signal range from one base station.

Cell Residence Time (CRT): The amount of time during which a mobile terminal stays in a cell during a single visit.

Channel Holding Time (CHT): The total length of time that a call makes use of a channel.

Co-located Cells: Cells with fully overlapping coverage.

EDGE: Acronym for enhanced data rate for GSM evolution

EV-DO: Acronym for evolution-data optimized

FDMA: Acronym for frequency division multiple access

GERAN: Acronym for GSM EDGE radio access network

GSM: Acronym for global system for mobile communications

Handoff: The process by which an ongoing call is transferred from one base station (or access point) to another. When a mobile user is in motion, in a car for example, and travels out of range of the original cell site, the ongoing call must be passed over to the next cell site.

Heterogeneous Wireless Network: A network that comprises fully or partially overlapping sub networks, which are based on different radio access technologies.

Homogeneous Wireless Network: A network that is based on a single radio access technology e.g., WLAN, WiMAX, UMTS, CDMA2000, etc.

Horizontal Handoff Call: A handoff between base stations that are using the same type of radio access technology.

JRRM: Acronym for joint radio resource management

Multimode Terminal: A mobile user's device that can connect to two or more networks that are based on different radio access technologies.

OFDM: Acronym for orthogonal frequency division multiplexing.

RAT: Acronym for radio access technology.

RRM: Acronym for radio resource management

TDMA: Acronym for time division multiple access

UMTS: Acronym for universal mobile telecommunications system.

UTRAN: Acronym for UMTS terrestrial radio access network

Vertical Handoff: Handoff between based stations (or access points) of two networks, which are based on different radio access technologies

1 Introduction

Next generation wireless network (NGWN) will be heterogeneous where different radio access technologies (RATs) coexist. In NGWN, each radio access technology has limited and different features in capacity, coverage, security level, service cost, and QoS levels offered to subscribers. A subscriber using a multimode mobile terminal (MT) in the heterogeneous wireless network will be able to access network services through any of the available RATs.

The motivation for heterogeneous wireless networks arises from the fact that no single RAT can provide ubiquitous coverage and continuous high QoS levels across multiple smart spaces, e.g., home, office, public smart spaces, etc. [1]. Therefore, multiple networks based on different technologies are being deployed in the same geographical area. The coexistence of different RATs in the same geographical area necessitates joint radio resource management (JRRM) for enhanced QoS provisioning and efficient radio resource utilization. The concept of JRRM arises in order to efficiently manage the common pool of radio resources that are available in each of the existing RATs [2, 3]. JRRM will promote integration and interoperability across existing systems.

Joint call admission control (JCAC) algorithm is one of the JRRM algorithms. The basic function of a call admission control (CAC) algorithm is to decide whether a new or handoff call can be accepted into a resource-constrained network without violating the service commitments made to already admitted calls. Given the existing traffic information, a CAC algorithm decides whether there is sufficient radio resource to meet the quality of service requested by an incoming call. CAC has been extensively studied in wired networks, and in homogeneous wireless networks such as GSM, UMTS, WLAN, and satellite network [4-6]. Reviews of homogeneous CAC algorithms appear in [7-9]. However, unlike homogeneous CAC algorithms, JCAC algorithms for heterogeneous wireless networks do not only decide whether an incoming call can be accepted or not. They also decide *which of the available radio access networks is most suitable to accommodate the incoming call*. Fig. 1.1 illustrates the basic functions of JCAC

algorithms. The two basic functions of JCAC algorithms are closely coupled. JCAC algorithms must manage individual services and technologies while maintaining an overall view of the total resources available in the heterogeneous wireless networks.

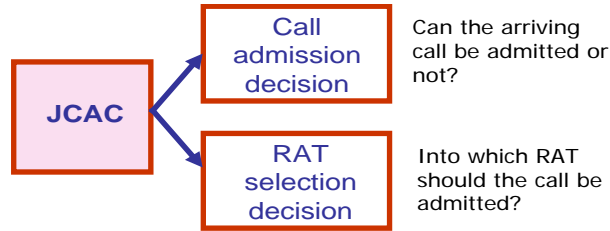


Figure 1-1. Basic functions of JCAC algorithms.

1.1 Need for Joint Call Admission Control Algorithms

The need for JCAC algorithms stems from the fact that traditional homogeneous CAC algorithms do not provide a single solution to address the heterogeneous architectures, which characterize next generation wireless networks [10]. JCAC algorithms are necessary for efficient utilization of radio resources, consistent provisioning of QoS across different RATs, overall stability of network, and enhancement of users' satisfaction. The need for JCAC algorithms is discussed in the following subsections.

1.1.1 Efficient Radio Resource Utilization

In wireless networks, radio resource is often scarce and expensive. It is the limiting factor regarding the maximum network capacity. Therefore, efficient use of radio resources has always been a primary concern in wireless communication [11]. A major objective of JCAC is to make the best use of radio resources in the available RATs in a heterogeneous wireless network while providing the requested QoS to users.

1.1.2 Consistent QoS Provisioning

Provision of guaranteed QoS for mobile subscribers is a challenging problem for next generation wireless networks. A single RAT cannot always meet the QoS requirements of all subscribers due to limited capacity and coverage. When the limited radio resource in

the air interface of a wireless network is oversubscribed by customers, network congestion occurs, and the packet-level QoS experienced by active users begins to deteriorate. In heterogeneous networks, if the packet-level QoS requirements of an incoming call cannot be met by a particular RAT, JCAC will admit the call into another RAT that can guarantee its QoS requirements. JCAC will therefore improve packet-level QoS in heterogeneous wireless networks. In a homogeneous network, a new call is blocked while a handoff call is dropped when no radio resources are available to accommodate the incoming call. However, with JCAC in a heterogeneous wireless environment, a new call that cannot be admitted into one RAT due to unavailability of radio resource can be admitted into another RAT. As a result, new-call blocking probability will be reduced. Similarly, a handoff call that cannot be admitted into a target cell in its current RAT due to unavailability of radio resource will be admitted into another RAT. Handoff call dropping probability will thereby be reduced. JCAC will therefore improve connection-level QoS in heterogeneous wireless networks.

1.1.3 Overall Network Stability

Each RAT in a heterogeneous wireless network has a maximum load capacity. If there is independent CAC in each RAT, some RATs may be overloaded whereas others are underutilized. A JCAC algorithm can be used to efficiently distribute network load across the available RATs. If a particular RAT is being overloaded, it may be necessary to handover some ongoing calls from this RAT to another RAT. Consequently, JCAC will ensure overall network stability.

1.1.4 Enhancement of Users' Satisfaction

Users have different preferences for different services. A user may prefer the cheapest RAT for a particular service whereas another user may prefer the best-QoS RAT for the same service regardless of the cost. Homogeneous CAC algorithms confine users to a particular RAT and therefore results in low users' satisfaction. JCAC in a heterogeneous network will enhance users' satisfaction by considering individual user's preferences in making call admission decisions.

1.2 Problems in the Existing JCAC Schemes

A number of JCAC algorithms have been proposed for heterogeneous wireless networks. These algorithms are broadly divided into two classes namely single-criterion and multiple-criteria JCAC algorithms. Single-criterion JCAC algorithms make RAT selection decisions based on a single criterion whereas multiple-criteria JCAC algorithms make RAT selection decision based on multiple criteria. In the following subsection, the existing JCAC algorithms and associated problems are discussed under these two broad classes.

1.2.1 Single-Criterion JCAC Algorithms

Examples of single criterion used for RAT selection are network load (load-based) [12, 13], service class (service-class-based) [14, 15], path loss (path-loss-based) [16], service cost (service-cost-based) [17], and network layer (network-layer based) [18]. Apart from load-based JCAC algorithms, a major problem with other single-criterion JCAC is that they can lead to highly unbalanced network load among available RATs in the heterogeneous wireless networks. Highly unbalanced traffic load in heterogeneous wireless networks will result in poor overall radio resource utilization and high overall call blocking/ dropping probability.

Load-based JCAC algorithms on the other hand uniformly distribute traffic loads among available RAT in heterogeneous wireless networks. Therefore load-based JCAC algorithm improves overall resource utilization and reduced call blocking/dropping probability. Gelabert *et al.* [12] studied the impact of load balancing among different RATs in heterogeneous cellular networks. However, handoff calls are not considered in the study. The algorithm deals with initial RAT selection only for new calls. Moreover, no analytical model is presented in the study to investigate connection-level QoS. Pillekeit *et al.* [13] proposed a load balancing algorithm for heterogeneous UMTS/GSM network with collocated cells. The load balancing algorithm is triggered when a certain load threshold is exceeded in order to balance the traffic load in the heterogeneous network. However, the

algorithm treats both new calls and handoff calls alike. In practice, it is necessary to keep handoff call dropping probability below new call blocking probability. Moreover, no analytical model is presented to investigate radio resource utilization and connection-level QoS in the heterogeneous wireless network.

In the previous works mentioned above, no analytical model has been developed for JCAC algorithms in order to investigate connection-level QoS parameters and system resource utilization in heterogeneous wireless networks.

1.2.2 Multiple-Criteria JCAC Algorithms

Multiple-criteria JCAC algorithms make RAT selection based on two or more criteria. Examples of multiple criteria used are (1) Signal strength, bandwidth, cost, reliability, latency, battery status, priority [19], (2) Price, bandwidth, SNR, sojourn time, handoff seamlessness, and battery consumption [20], (3) SNR, Data Rate [21], and (4) link quality, price, bandwidth [22]. NGWN will be user-centric, considering users preferences in making RAT-selection decisions. Therefore, it is assumed that these criteria are evaluated for individual users wanting to make a call, and RAT selection is based on individual user's preferences.

Chan *et al* [19] presented a RAT selection algorithm based on the concept of fuzzy multiple objective decision making (MODM). Seven example criteria are used in the algorithm namely, signal strength, bandwidth, charging model, reliability, latency, battery status and the user's preferred segment (priority). The purpose of the RAT selection algorithm is to select the most suitable RAT for a particular call based on the criteria mentioned above.

Zhang [20] proposed an approach, which uses fuzzy logic to represent the imprecise information of some RAT selection criteria. The fuzzy MADM (multiple attribute decision) method operates in two steps. The first step is to convert the imprecise fuzzy variable to crisp numbers. The second step is to use classical multiple attribute decision technique to determine the ranking order of the candidate networks. The highest-ranking RAT is then selected for the incoming call.

Wilson *et al* [21] proposed a decision strategy for making the optimal choice of wireless access networks. Fuzzy logic is used as the inference mechanism and a prototype is developed. The prototype uses two metrics from a candidate network, one metric from application requirements, and user defined criteria as input. Based on these criteria, the most suitable RAT is selected for each incoming call.

The major problem with the above algorithms is that they can lead to highly-unbalanced traffic load among different RATs in a heterogeneous wireless network. This is because users act independently and many of the users may prefer to be connected through a particular RAT. Moreover, no analytical model has been developed to study connection-level QoS (new call blocking probability and handoff call dropping probability) and radio resource utilization in the proposed multiple-criteria JCAC algorithms.

1.3 Research Objectives

Efficient radio resource utilization and QoS provisioning are major concerns in wireless networks. The main objective of this research is to develop joint call admission control and bandwidth management schemes to improve system utilization and connection-level QoS in heterogeneous wireless networks. The following three areas are investigated under this research.

1. Joint Call admission control and adaptive bandwidth management to enhance QoS and system utilization in heterogeneous wireless networks.
2. Optimal RAT selection policy to reduce call blocking probability in heterogeneous wireless networks.
3. Dynamic pricing for balancing traffic load in multiple-criteria user-centric JCAC of heterogeneous wireless networks.

1.4 Scope of Research

This research focuses on improving connection-level QoS and radio resource utilization in heterogeneous wireless networks. Packet-level QoS are not covered in this thesis although it may be investigated in the future. The scope of the work considered in the research is heterogeneous wireless networks with overlapping coverage such as heterogeneous cellular networks with co-located cells.

1.5 Contributions

The major contributions of this thesis are as follows:

1.5.1 A critical survey of joint call admission control in heterogeneous wireless networks

A critical survey of existing joint call admission control algorithms for heterogeneous wireless networks is presented. Requirements of JCAC algorithms are discussed. Eight different approaches for selecting the most appropriate radio access technology (RAT) for incoming calls in heterogeneous wireless networks are examined. JCAC algorithms are classified based on these approaches, and the advantages and disadvantages of each approach are discussed. Finally, six different design considerations are analyzed for JCAC algorithms.

1.5.2 Combination of adaptive bandwidth management and JCAC to improve system utilization and quality of service in heterogeneous wireless networks

An adaptive bandwidth management and joint call admission control scheme is proposed for heterogeneous wireless networks. The objectives of the proposed adaptive JCAC scheme are to enhance average system utilization, uniformly distribute traffic load among available RATs, guarantee QoS requirements of all accepted calls, and reduce new call blocking probability and handoff call dropping probability in heterogeneous wireless

networks. Numerical results show an improvement in average system utilization of up to 20%. Results also show an improvement in connection-level QoS.

1.5.3 Development of analytical model to evaluate new call blocking probability, handoff dropping probability, and system utilization in heterogeneous wireless networks

In the existing JCAC algorithms, no analytical model has been developed to evaluate connection-level QoS in heterogeneous wireless networks. Based on Markov decision process, an analytical model is developed for evaluating new call blocking probability, handoff call dropping probability, and system utilization in heterogeneous wireless networks.

1.5.4 Reduction of call-blocking probability through optimal allocation of calls

An optimal RAT selection JCAC scheme is proposed to reduce call blocking probability in heterogeneous wireless networks. The algorithm makes call admission decisions such that overall call-blocking probability is reduced in the heterogeneous wireless network. Optimal splitting of arrival calls is determined using linear programming optimization technique. Numerical results show that the algorithm reduces call-blocking probability in the heterogeneous wireless network.

1.5.5 Use of dynamic pricing to balance traffic load in multiple-criteria joint call admission control of heterogeneous wireless networks

Dynamic pricing is proposed to balance traffic load among available RATs in heterogeneous wireless networks where users' preferences are considered in making RAT selection decisions. Independent users' preferences in heterogeneous wireless networks often lead to highly unbalanced network load, which in-turn increases overall call blocking/ dropping probability and reduces radio resource utilization. By dynamically adjusting the service price in each of the available RATs, the proposed JCAC scheme

evens out the unbalanced traffic load caused by independent users' preferences. The performance of the proposed JCAC scheme is compared with the performance of a scheme that does not incorporate dynamic pricing. Numerical results show that the proposed JCAC scheme reduces call blocking probability and improves radio resource utilization.

These contributions are contained in the author's papers listed below.

Journal Publications

- (1) **O.E. Falowo** and H. A. Chan, "Joint Call Admission Control Algorithms: Requirements, Approaches, and Design Considerations," Elsevier Computer Communications (2007), Vol 31/6, pp. 1200-1217, DOI:10.1016/j.comcom.2007.10.044.
- (2) **O.E. Falowo** and H. A. 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, Article ID 34378, 11 Pages, 2007, DOI:10.1155/2007/34378.
- (3) **O.E. Falowo** and H.A. Chan, "Dynamic Pricing for Load Balancing in User-Centric Joint Call Admission Control of Next Generation Wireless Networks," Submitted to Elsevier Computer Communications.

Peer-Reviewed Conference Publications

- (4) **O.E. Falowo** and H.A. Chan, "Optimal Joint Radio Resource Management to Improve Connection-Level QoS in Next Generation Wireless Networks," Proceedings of the IEEE Radio and Wireless Symposium, Orlando, Florida, USA, 22-24 January, 2008.
- (5) **O.E. Falowo** and H.A. Chan, "Joint Call Admission Control Algorithm to Enhance Connection-level QoS in Heterogeneous Cellular Networks," Proceedings of the IEEE AFRICON, Country Club, Namibia , 26-28 September, 2007, ISBN: 0-7803-8606-X.
- (6) **O.E. Falowo** and H.A. Chan, "Dynamic Threshold-Based Joint Call Admission Control Scheme for Multi-Service Heterogeneous Cellular Networks," Proceedings of the 3rd International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), Shanghai, China, 21-23 September, 2007.
- (7) **O.E. Falowo** and H.A. Chan, "Analysis of Joint Call Admission Control Strategies for Heterogeneous Cellular Networks," Proceedings of the 12th IEEE Symposium on

Computers and Communications, Aveiro, Portugal, July 1-4, 2007, ISBN: 1-4244-1521-7.

- (8) **O.E. Falowo** and H.A. Chan, "Fuzzy Logic Based Call Admission Control for Next Generation Wireless Networks," Proceedings of 3rd International Symposium on Wireless Communication Systems, Valencia, Spain, 5-8 September, 2006, ISBN: 1-4244-0398-7.
- (9) **O.E. Falowo** and H.A. Chan, "Joint Call Admission Control for Next Generation Wireless Networks," IEEE Canadian Conference on Electrical and Computer Engineering, Ottawa, Canada, 7-10 May 2006, ISBN: 1-4244-0038-4.

1.6 Thesis Outline

The remainder of this thesis is organized as follows.

Chapter 2 gives a description of heterogeneous wireless networks. The motivation for heterogeneous wireless networks, challenges, and need for joint radio resource management are discussed.

Chapter 3 presents a review of joint call admission control and bandwidth management schemes in wireless networks. Six requirements for JCAC algorithms are discussed. JCAC algorithms are classified into eight groups based on RAT selection approaches. Multiple-criteria JCAC algorithm are discussed and seven design considerations for JCAC algorithms are analyzed. Lastly, bandwidth management schemes are described.

Chapter 4 proposes an adaptive bandwidth management and joint call admission control scheme for heterogeneous wireless networks. The components of the proposed scheme are described. A Markov model is developed to evaluate new call blocking probability, handoff call dropping probability, and system utilization in the heterogeneous wireless network. Performance of the proposed scheme is compared with that of a non-adaptive JCAC scheme, and numerical results are discussed.

Chapter 5 presents an optimal JCAC scheme for reducing call blocking probability in heterogeneous wireless networks. The components of the proposed scheme are discussed. Using linear programming technique, the optimal RAT selection policy is determined. A

model is developed to evaluate new call blocking probability and handoff call dropping probability in the heterogeneous wireless network. Numerical results are discussed.

Chapter 6 deals with the problem of load balancing in multiple-criteria JCAC of heterogeneous wireless networks. The relationships among users' preferences, load balancing, and pricing are discussed. A JCAC algorithm, which incorporates dynamic pricing, is proposed for load balancing in heterogeneous wireless networks. Components of the proposed scheme are discussed and a Markov model is developed for the scheme. Three scenarios namely general user scenario, high-price-sensitive user scenario, and low-price-sensitive user scenario are considered. The performance of the proposed scheme is evaluated and discussed.

Finally, chapter 7 summarizes the contributions of the thesis, and suggests areas for future study.

University of Cape Town

2 Heterogeneous Wireless Networks

It is envisioned that next generation wireless network (NGWN) will be heterogeneous where different radio access technologies coexist. In NGWN, mobile users will be able to communicate through any of the available radio access technologies (RATs) and roam from one RAT to another, using multimode terminals (MTs) [23, 24]. As shown in Fig. 2.1, a mobile terminal initiating a call in a heterogeneous wireless network can be admitted into any of the available RATs that can support the call.

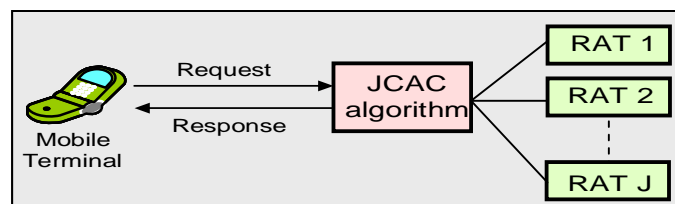


Figure 2-1. JCAC in heterogeneous wireless network.

NGWN will provide high bandwidth access anytime and anywhere for different classes of services such as voice, video, web surfing, etc. The following subsections discuss the motivation for heterogeneous wireless networks, its challenges, and its radio resource management (RRM).

2.1 Motivation for Heterogeneous Wireless Networks

The motivations for heterogeneous wireless networks are (1) limitation of a single RAT, (2) evolution of wireless technology, and (3) users' demand for advanced service and complementary features of different RATs.

2.1.1 Limitation of a Single RAT

Every RAT is limited in one or more of the following: data rate, coverage, type of services, and quality of service it can provide. A motivation for heterogeneous wireless networks arises from the fact that no single RAT can provide ubiquitous coverage and continuous high QoS levels across multiple smart spaces, e.g., home, office, public smart spaces, etc.

[1]. Consequently, multiple RATs are being deployed in the same geographical area to meet deferent needs of customers.

2.1.2 Evolution of Wireless Technologies

Researchers have come up with more and more spectrally efficient multiple access techniques and modulation schemes. Consequently, wireless networks have evolved from one generation to another. However, due to huge investment on existing radio access technologies, operators do not readily discard their existing radio access technologies when they acquire a new one. This situation has led to coexistence of multiple RATs in the same geographical area. For example many of the 3G (W-CDMA) service providers still retain their 2G (FDMA/TDMA) infrastructure. Consequently, the two or more different RATs coexist.

2.1.3 Users' Demand for Advanced Service and Complementary Features of Different RATs

There is increasing demand for different types of services by users. Some services are better provided on one access network than other. Moreover, different RATs possess different capabilities in data rate, coverage area, security level, QoS-level, etc. It is sometimes necessary to deploy two or more RATs with complimentary features in the same geographical area in order to have the combined benefits provided by the multiple RATs. An example of RATs with complementary features is integrated UMTS-WLAN network.

2.2 Challenges of Heterogeneous Wireless Networks

Integrating multiple RATs brings many challenges ranging from interworking among inherently different wireless radio access technologies to QoS provisioning. These challenges include joint call admission control, seamless vertical handoff, common authentication and authorization, network security across different RATs, unified billing, availability of multimode mobile terminal, etc. These challenges are briefly discussed in the following subsections. However, this thesis focuses on joint call admission control.

2.2.1 Joint Call Admission Control

In a heterogeneous wireless network, a joint call admission control algorithm is needed to decide whether an incoming call can be accepted or not. In addition a JCAC is needed to decide *which of the available RATs is most suitable to accommodate the incoming call*.

2.2.2 Seamless Vertical Handoff

Vertical handoff is the handoff between different types of networks [25, 26]. Seamless vertical handoff implies that the handoff procedure is transparent to upper-layer applications. Consequently, users will not be aware of any change in the network they are using, and will not be required to interact with the network to enable vertical handover. Achieving seamless vertical handoff is a major challenge in NGWN.

2.2.3 Common Authentication and Authorization

Authentication is the act of verifying a claimed identity, in the form of a pre-existing label from a mutually known namespace, as the originator of the message (message authentication) or as the channel end point [27]. Authorization is the act of determining whether a particular right can be granted to the presenter of a particular credential. This particular right can be, for example, an access to a resource [27]. One of the major issues in interworking of wireless networks is the provision of common authentication and authorization mechanism, which allow a subscriber in one RAT to have access to services in another RAT.

2.2.4 Network Security

Security is one of the technical challenges in the interworking of wireless networks. Each network has its own security mechanism, which may not be compatible with that of others. Moreover, some network provides more secured services than the other. For example, the access security features in UMTS are a superset of those provided in GSM. Some new security features are introduced in UMTS to correct the perceived weaknesses of GSM

security. It is very essential to ensure that interworking of different RATs does not compromise the security of the entire heterogeneous wireless network.

2.2.5 Unified Billing

Another challenge of heterogeneous wireless networks is common billing, which means that a subscriber will receive a single combined bill for using services offered by different network operators, and across multiple RATs. This implies that the home network of the subscriber or a third party will be responsible for collecting the billing information and aggregating them together. Two billing options are available; pre-paid and post-paid billing. In pre-paid billing, a subscriber has paid in advance. When the subscriber uses network services, the home network operator (or the third party) checks the charging information and deduct corresponding charges from the subscriber's credit. When the subscriber is out of credit the access network services are denied. In post-paid billing, a subscriber has a billing agreement with the home service provider to pay at intervals and is charged regularly for the usage within the agreed period of time.

2.2.6 Availability of Multimode Terminals

The next generation wireless network is expected to support diverse types of terminals, such as mobile phones, personal digital assistants, laptops, etc. These terminals need to possess a wide range of capabilities in order to take the full advantage of heterogeneous wireless networks. The terminals must be able to access the core network by choosing any of the available RATs. In order to achieve this purpose, terminals will have multiple access interfaces (multi-modality) or have a dynamically reconfigurable access interface [28]. It is projected that 2G/2.5G/3G triple-mode terminals will be available for most users in 2009-2010, with a penetration of up to 90 % [29]. Within the same period of time, it is also expected that 2G/2.5G/3G/WLAN quad-mode terminals will have a penetration of about 50 % [29].

2.3 Radio Resource Management in Heterogeneous Networks

In wireless networks, radio resource management algorithms are responsible for efficient utilization of the air interface resources in order to guarantee quality of service, maintain the planned coverage area, and offer high capacity. Radio resource management is the bridge between diverse RATs [30]. In heterogeneous wireless networks, radio resources can be independently managed as shown in Fig 2-2 or jointly managed as shown in Fig. 2-3.

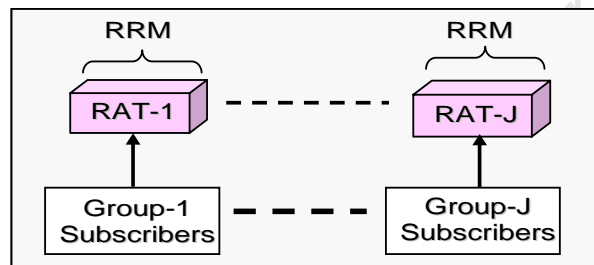


Figure 2-2. Independent RRM in heterogeneous wireless networks.

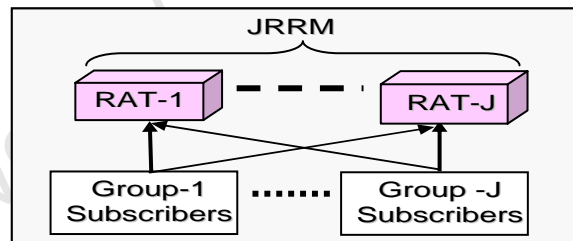


Figure 2-3. Joint RRM in heterogeneous wireless networks.

With independent RRM, each group of subscribers is confined to a single RAT, whereas with JRRM, a subscribers (using a multimode terminal) from any group can be connected through any on the available RATs that can support its class service.

The concept of JRRM arises in order to efficiently manage the common pool of radio resources that are available in each of the existing RATs [2, 31]. In wireless networks, each radio resource pool consists of resources that are available in a set of cells, typically

under the control of a radio base station or an access point. Joint RRM among multiple operators of different RATs is one way towards provisioning of low-cost mobile multimedia entertainment [32].

The coexistence of multiple RATs in the same geographical location necessitates JRRM for enhanced quality of services and efficient radio resource utilization. JRRM will enhance QoS by providing, through combination of RATs, different services that cannot be supported by individual RATs. Moreover, JRRM in heterogeneous wireless networks will enhance the use of radio resources by providing extended capacity and coverage. In addition seamless vertical handover among different RATs is only feasible if there is JRRM among the available RATs.

There are two JRRM algorithms that are of great interest, namely the joint call admission control algorithm and joint session scheduling (JSS) algorithm. Using JCAC, subscribers can be admitted into any of the available RATs in heterogeneous wireless networks. However, a call cannot be split among two or more RATs. On the other hand, joint session scheduling in heterogeneous wireless networks will enable packets of one data flow to be delivered simultaneously through multiple RATs. JSS offers higher flexibility than JCAC. However, JSS will introduce a high level of signaling overhead and complexity into the heterogeneous wireless network. This research focuses on JCAC algorithms.

2.4 Summary

This chapter gives a brief overview of heterogeneous wireless networks. The motivations for heterogeneous wireless networks are highlighted and major technical challenges are discussed. The need for joint radio resource management among available RATs in heterogeneous wireless networks is discussed. Finally, joint call admission control algorithm and joint session scheduling algorithm, which are two JRRM algorithms of great interest, are described.

3 Review of JCAC and Bandwidth Management Schemes

3.1 JCAC algorithms

In heterogeneous wireless networks, the basic functions of a JCAC algorithm are to decide the following: (1) *whether an incoming call can be admitted into a heterogeneous wireless network or not*, and (2) *which of the available RATs is most suitable to accommodate the incoming call*. It is desirable that a JCAC algorithm for heterogeneous wireless networks meet certain requirements. These requirements are discussed in the following subsections.

3.1.1 Requirements for JCAC algorithms

Six requirements for JCAC algorithms are discussed in this section. They are multi-service, efficiency, simplicity, high-execution speed, scalability, and stability.

3.1.1.1 Multi-service

Next generation wireless network will support multiple classes of services such as real-time voice and video, video and audio streaming, browsing, etc. Supporting multiple services in heterogeneous wireless networks will enhance users' satisfaction because different users have different service requirements. Moreover, supporting multiple services will increase operators' revenue. Therefore, JCAC algorithms need to support multiple services.

3.1.1.2 Efficiency

Performance of JCAC algorithms is measured in a number of ways such as radio resource utilization, new call blocking probability, handoff call dropping probability, system utilization, average delay, operators' revenue, and users' satisfaction. JCAC algorithms must be efficient in achieving the designed goals. Generally, an efficient JCAC algorithm will guarantee the QoS requirement of accepted calls and achieve high radio resource utilization.

3.1.1.3 Simplicity

Implementation cost and scalability problem require that a JCAC algorithm be as simple as possible. A simple algorithm will have a low computational overhead and therefore will not incur additional delay in the network. However, an overly simple JCAC algorithm may not achieve high radio resource utilization. For good quality of service and efficient radio resource utilization, a sophisticated JCAC algorithm is required to support multiple services, especially in a scenario where users are dynamically roaming across different access networks. Therefore, there is a tradeoff between simplicity and the efficiency of JCAC algorithms.

3.1.1.4 High- execution speed

Call admission control algorithms operate in real-time. Therefore the execution speed should be very high so that they do not cause additional delay in the network. High execution speed of JCAC algorithms will enhance QoS in the heterogeneous wireless network.

3.1.1.5 Scalability

There has been an increase in the demand for multimedia data services in recent years. This increase in demand is likely to grow at an even faster pace in the future due to advances in multimedia distribution services. Therefore, the overall system capacity of heterogeneous wireless networks must be expandable in terms of the number of subscribers supported, data rate, and geographical coverage. A JCAC algorithm must be able to accommodate increase in capacity or size of individual RATs. It must also accommodate the integration of other access networks. Moreover, exchange of a large amount of information among base stations or different RATs may exert a significant overhead cost in heterogeneous wireless networks as the size of the network increases. Therefore, JCAC algorithms must minimize the amount of information exchange in the heterogeneous networks.

3.1.1.6 Stability

It is necessary that JCAC algorithms ensure overall stability of the heterogeneous wireless network. Instability refers to a situation where certain RATs suddenly become overloaded

whereas some other RATs with overlapping coverage are underutilized. In which case, it may be necessary to move some subscribers back and forth from the suddenly overloaded RAT to the underutilized RAT. A major disadvantage of instability is increase in frequency of vertical handoff (i.e. handoff between two different RATs), thereby reducing the overall efficiency of the network. Therefore, it is desirable that a JCAC algorithm ensures overall stability of the heterogeneous wireless network.

Some of the six requirements mentioned above are closely related whereas some others are conflicting. For example, efficiency, speed, and scalability are related requirements. An efficient JCAC algorithm will have high-execution speed and be scalable. On the other hand, simplicity and efficiency are conflicting requirements. An overly simple JCAC algorithm will not be efficient. Supporting multiple services and simplicity are also conflicting requirements. Therefore, there is a need for a compromise among the conflicting requirements. Table 1-1 shows the relationship among these requirements.

Table 1-1. Relationship among JCAC algorithm requirements.

JCAC Requirement	Multiple services	Efficiency	Simplicity	High Speed	Scalability	Stability
Multiple services						
Efficiency						
Simplicity	×	×				
High Speed		✓				
Scalability		✓	✓	✓		
Stability		✓	×	✓		

Key: X represents conflicting requirements

✓ represents interrelated requirements

3.1.2 Classification of JCAC Algorithms

JCAC algorithms can be classified into two broad groups based on RAT selection approaches, which is either based on single criterion or multiple criteria. The two groups can be further classified into eight groups. This classification is depicted in Fig 3-1, summarized in Table 3-1, and subsequently described in this section.

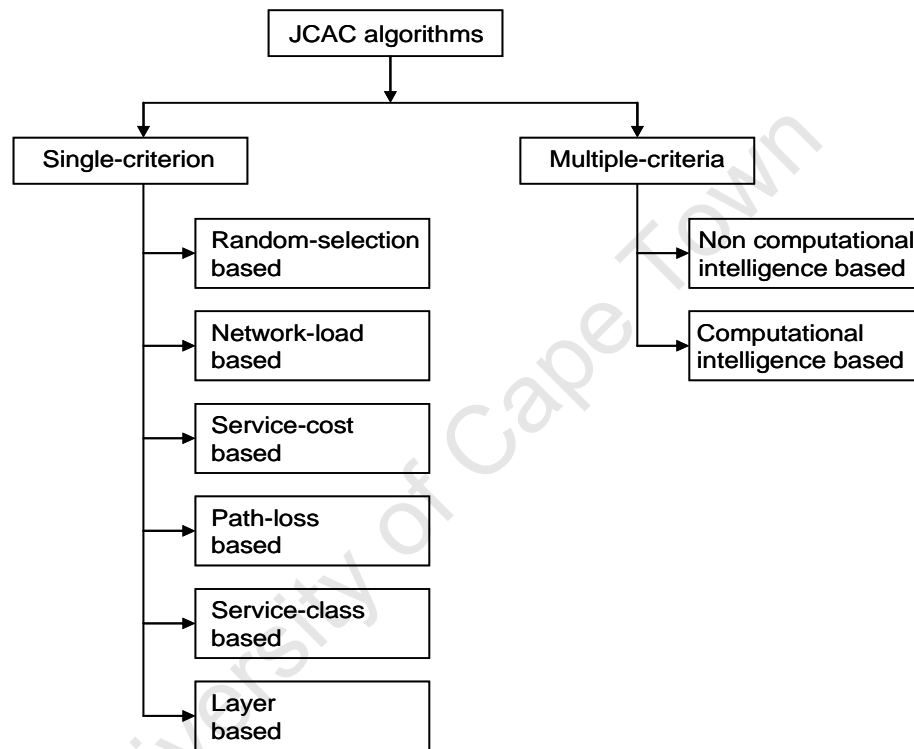


Figure 3-1. Classification of JCAC algorithm based on RAT selection approaches.

Table 3-1. Summary of the RAT selection approaches.

Approach	Main Idea	Advantage	Disadvantage
Random-selection-based	Calls are randomly admitted into any of the available RATs	Simple	Not efficient, high blocking probability
Network-load-based	Calls are admitted into the least loaded RAT in the heterogeneous network such that network load is almost the same for all the available access network	High network stability due to uniform load distribution	Network centric, hence low user satisfaction
Service-cost -based	Calls are admitted into the least expensive RAT such that the subscriber incurs the least service cost in the heterogeneous network	Reduced overall service cost	Unbalanced network load
Path-loss-based	Calls are admitted based on path-loss measurement.	Lower bit-error rate and higher throughput	High frequency of vertical handover
Service-class-based	Calls are admitted into a particular RAT based on the class of service	High QoS	Unbalanced network load
Layer-based	Calls are admitted based on layers, starting from the upper layer. If the current layer is loaded, JCAC tries the next lower layer.	Simple	Highly unbalanced network load
Non computational intelligence-based	Calls are admitted into a particular RAT based on some cost function or utility function derived from multiple criteria without the use of computational intelligence techniques	Efficient, maximize some objective function, and improve users satisfaction	High computational overhead, complex
Computational intelligence-based	Calls are admitted into a particular RAT, which is chosen by applying a computational intelligence technique (e.g., fuzzy logic) to some RAT selection criteria	Efficient, improves users satisfaction function	Complicated

3.1.2.1 Random-selection- based JCAC

In this approach, when a new or vertical handoff call arrives, one of the available RATs is randomly selected for the call. If there is no enough radio resource to accommodate the call in the selected RAT, the call is blocked or dropped. In a variant of random-selection-based JCAC, if the selected RAT cannot accommodate the call, another RAT is randomly selected. The call is blocked or dropped if none of the selected RAT can accommodate it. Random-selection-based JCAC algorithm is generally used in performance evaluation of other RAT selection algorithms by comparison. The advantage of this algorithm is that it is easy to implement. However, it has a high call blocking probability, and low radio resource utilization efficiency.

3.1.2.2 Load-based JCAC

The objective of load-based JCAC algorithms is to uniformly distribute traffic load among all the available RATs in a heterogeneous wireless network. Balancing load among multiple RATs in heterogeneous wireless networks allows for a better utilization of the radio resources [33, 34]. In heterogeneous wireless networks, traffic load can be continuously balanced as in [33], balanced at certain intervals of time as in [34], or balanced when a particular differential load threshold is reached as in [13]. Load balancing can also be forced or unforced. Forced load balancing is carried out by moving some ongoing call(s) from a highly-loaded RAT to a less-loaded RAT (call reassignment). On the other hand, unforced load balancing is carried out only during the admission of a new call or necessary vertical handoff call (i.e. a vertical handoff that occurs because an active mobile terminal is moving outside the coverage area of the current RAT). The new or vertical handoff call is then admitted into the least loaded RAT among the available RATs. A major advantage of load balancing algorithms is high network stability due to uniform load distribution. However, load balancing JCAC algorithms are network-centric.

Gelabert *et al* [12] evaluated the performance of a load-balancing RAT selection algorithm for new calls in a heterogeneous wireless network which consists of a UTRAN and a GERAN. For UMTS network, the uplink load factor is estimated as

$$\eta_{UL} = 1 - \frac{P_N}{I_{total}} \quad (3.1)$$

where P_N is the background thermal noise and I_{total} is the total received wideband power. The downlink load factor is estimated as

$$\eta_{UL} = \frac{P_{total}}{P_{maximum}} \quad (3.2)$$

where P_{total} is the total downlink transmission power and $P_{maximum}$ is the maximum Node-B transmission power. For GERAN, the uplink and downlink loads are obtained from the average amount of Time Slots (TSL) utilized by GSM/EDGE services as

$$TSL_{utilization} = \frac{\text{Used TSL in previous frame}}{\text{Available TSL for GSM and EGPRS}} \quad (3.3)$$

Seven collocated omnidirectional cells are considered for GERAN and UTRAN, and two service types: voice and interactive users are assumed. The performance metrics used are (1) throughput (Mbps) against number of users, (2) average cell load against number of users, and (3) weighted delay (seconds) against number of users. For comparison purposes, a service-based JCAC algorithm is also simulated. The service-based JCAC algorithm allocates users according to the demanded service-type.

The two algorithms (load-based and service-based JCAC) are simulated. As expected, simulation results show that the load balancing JCAC algorithm maintains approximately equal load in both GERAN and UTRAN whereas for service-based JCAC algorithm, the cell load in GERAN is up to 4.8 times that of UTRAN in the extreme case. Results also show that the load-based JCAC has a higher total aggregated throughput than that of service-based JCAC algorithm. However, the weighted delay experienced by interactive users is more for load-based JCAC algorithm than for service-based JCAC algorithm. In the extreme case, for the load-based JCAC algorithm, this delay is up to six times that of the service-based JCAC algorithm.

Murray *et al* [1] proposed a policy based RAT selection algorithm which chooses the access network that is currently least loaded for a particular class of service request (i.e. voice, www, video streaming). The performance of the proposed policy based RAT selector is investigated via simulations for EDGE and UMTS networks. Three classes of user traffic are considered; voice, www, and video streaming. The major performance metrics used are network utilization and QoS offered to the end users. QoS offered to users is measured in terms of the video frame drop rate (FDR) for video streaming sessions and the block error rate (BLER) for web sessions. The performance of the policy-based JCAC algorithm is compared with that of a random-selection based JCAC algorithm. For a scenario of 300 users in the heterogeneous wireless network, results show that the policy based RAT selection algorithm has a frame drop rate which varies from 0 to about 3.2 as the number of video sessions increases whereas the random selection algorithm has a frame drop rate which varies from 0 to about 40.2 as the number of video sessions increases. Results also show that the policy-based JCAC algorithm has a frame error rate, which varies from 0 to 4 whereas the random-selection JCAC algorithm has a frame rate which varies from 0 to 4.2. Thus the overall performance of the policy-based JCAC algorithm is better than that of the random-selection based JCAC algorithm.

Pillekeit *et al* [13] proposed a forced-based load balancing algorithm for co-located UMTS/GSM networks. The proposed JCAC algorithm does not continuously try to balance the load levels among the available RATs. The algorithm is triggered only when the differential load between the two networks is above a certain threshold. This approach has the advantage of reducing the radio resources consumed due to signaling overhead. It also reduces the frequency of vertical handoff.

Performance of the forced-base load balancing algorithm is evaluated using STEAM (Simulation Tool for the Evaluation of Algorithms in Mobile Networks), the system simulator tool of Lucent Technologies. In the simulated scenario, all cells are co-located and all mobile phones use circuit switched voice service. In GSM mode, the mobile terminals use a full rate codec with a data rate of 13 Kbps whereas in UMTS mode, they use a data rate of 12.2 Kbps (spreading factor of 128 is assumed).

For comparison purposes, a scenario where there is no joint call admission control between the two RATs is also simulated. Results show that with the load balancing algorithm in place, a gain of 8.4% of the overall traffic capacity can be achieved compared to the sum of the traffic capacity of the two disjoint systems (i.e. without JCAC).

3.1.2.3 Service-class- based JCAC

Service-class based JCAC algorithms admit calls into a particular RAT based on the class of service, such as voice, video streaming, real-time video, web browsing, etc. [14]. This approach is based on the fact that different RATs are optimized to support different classes of service. For example, GSM is designed for voice services whereas EV-DO is optimized for data services. Therefore, the algorithm admits an incoming call into a RAT that can best support the service class of the call. Service- class-based JCAC algorithms have the advantage of high packet-level QoS. However, they may lead to highly unbalanced network load.

Service-class based JCAC can be classified as rigid or flexible. Rigid service-class-based JCAC algorithm tries to admit an incoming call of a specific class into a particular RAT, if the preferred RAT for this call cannot accommodate the call, probably because there is no enough radio resource, other RATs are not acceptable. Therefore the call is blocked. On the other hand, flexible service-class-based JCAC algorithm tries to admit an incoming call of a specific class into a particular RAT. If the preferred RAT for this call cannot accommodate the call, other RATs are acceptable. Flexible service-class-based JCAC algorithm has a lower call blocking probability when compared with rigid service-class-based JCAC algorithm.

Zhan [14] investigated a service-based JCAC algorithm for heterogeneous GSM/UMTS network with overlapping coverage. It is assumed that the two RATs have the same capacity for all types of traffic. It is assumed that the two RATs have the same capacity for all types of traffic. Three classes of calls: voice, streaming, and data are considered, and each account for 40%, 30%, and 30% respectively. The preferred network for voice is GSM whereas UMTS is the prefer network for streaming and data. Four different scenarios are considered for the JCAC algorithm. These scenarios are summarized in Table 3-2.

Table 3-2. Four scenarios for the JCAC algorithm.

Scenario	Allocated RAT		
	Voice	Streaming	Data
1	GSM	UMTS	Least loaded RAT
2	GSM	Least loaded RAT	UMTS
3	Least loaded RAT	UMTS	UMTS
4	Least loaded RAT	Least loaded RAT	Least loaded RAT

Two performance metrics are used in the investigation. These metrics are (1) *loss probability against normalized arrival time* and (2) *degradation against normalized arrival time*. Results show that scenario 1 has the least loss probability followed by scenario 4 whereas scenario 3 has the highest loss probability. Results also show that 1 has the least degradation, followed by scenario 4 whereas scenario 3 has the highest degradation.

Song *et al* [35] proposed a service-based JCAC algorithm for integrated WLAN/UMTS network, considering voice and data calls. In double-coverage region, the algorithm tries to admit voice calls into UMTS network. If there is no radio resource to accommodate the call in UMTS, the call is admitted into WLAN. Otherwise the voice call is blocked. On the other hand, the JCAC algorithm tries to admit data call into WLAN. If the data call cannot be admitted into WLAN, it is blocked. The main objective of the proposed algorithm is to reduce the frequency of handoff voice calls in the heterogeneous wireless network.

Numerical analysis is carried out to evaluate the performance of the proposed JCAC algorithm, considering a case where only one WLAN access point (AP) is located in each UMTS cell. The performance of the proposed scheme is compare with that of a “WLAN-first” scheme. In the WLAN-first scheme, the JCAC algorithm first tries to admit both new

voice calls and new data calls into WLAN in the double-coverage area so as to get the benefit of possible larger bandwidth and less cost. If the calls are rejected, the algorithm will then try the cellular system for admission of the calls. Whenever the WLAN coverage is available, on-going voice and data calls are handed over to the WLAN by the “WLAN first” scheme, provided there is enough spare capacity to accommodate the handoff calls.

The performance metrics used are (1) *average number of handoff per voice call against user mobility measure* and (2) *average percentage of time served by the cellular network against user mobility measure*. Results show that under high user mobility, the proposed service-based JCAC algorithm reduces the average number of handoff per voice call to about 37% of that of the “WLAN-first” JCAC scheme. Results also show that for the proposed algorithm, the percentage of time served by the cellular network is up to 1.3 times of that of “WLAN-first” scheme.

Romero *et al* [15], proposed a service-based RAT selection policy for heterogeneous wireless networks. A heterogeneous network comprising GERAN and UTRAN is used, and a mix of voice and interactive users (e.g., www browsing) are considered. Examples of the service-based selection policies are defined in the following [15]:

- (1) VG (voice GERAN) policy: This policy has only the service class as input and allocates voice users into GERAN and other services into UTRAN.
- (2) VU (voice UTRAN) policy: This policy acts in the opposite direction to VG and allocates voice users to UTRAN and interactive users to GERAN.

The performance of the VG and the VU policies are evaluated in terms of aggregated throughput (Mbps) for 400 voice users and different numbers of www users. Results obtained are shown in Table 3-3 [15]. Table 3-3 shows that VG policy outperforms VU, revealing the suitability of allocating voice users in GERAN.

Table 3-3. Total throughput (mbps) for the two basic policies.

Number of users		VU policy		VG policy	
Voice	www	Uplink	Downlink	Uplink	Downlink
400	200	2.08	2.17	2.14	2.22
	600	2.88	3.09	2.95	3.15
	1000	3.64	3.96	3.76	4.08

Koo *et al* [36] investigated a service-based JCAC algorithm in a heterogeneous wireless network comprising a GSM/EDGE network and a WCDMA network. Two service classes: voice and data are considered. The admissible region of the considered heterogeneous system depends on how different services are allocated onto the subsystems. In the service-based JCAC algorithm, incoming calls are admitted into the subsystem where the expected relative resource consumption (relative resource cost) for the service class of the incoming call is the smallest. They also considered another JCAC (called the worst case JCAC) algorithm, in which the call assignment rule is opposite to that of the service-based JCAC algorithm. Using numerical simulations, the performance of both the service-based JCAC and the worst case JCAC algorithms are compared to that of separate operation of CAC algorithm in individual networks.

Results show that the service-based JCAC algorithm provides up to 60% Erlang capacity improvement over the separate operation of CAC algorithm in individual networks. Results also show that the worst-case JCAC algorithm still provides about 15% capacity improvement over the separate operation of the CAC algorithm in individual networks.

3.1.2.4 Path-loss-based JCAC

Path-loss based JCAC algorithms make call admission decision based on path-loss measurements taken in the cells of each RAT. These JCAC algorithms have the advantage of low bit-error rate and high throughput. However, they can cause high frequency of vertical handover.

Perez-Romero *et al* [16] proposed path-loss based JCAC algorithms for new calls (initial RAT selection) and handoff calls (vertical handoff RAT selection) in heterogeneous CDMA/TDMA network.

The initial RAT selection algorithm admits new calls based on path-loss (PL) measurements taken in the best CDMA cell, provided the terminal is in the establishment phase. Path loss is computed by measuring the received downlink power from a common control channel whose transmitted power is broadcast by the network. The measurements are averaged in periods of T seconds. If the resulting path loss is higher than a given threshold PL_{th} , TDMA network is selected. If the path loss is less than the given threshold, CDMA network is selected. If there is no radio resource to accommodate the call in the selected RAT, the other RAT will be selected. If there is no radio resource to accommodate the call in neither of the two RATs, the call will be blocked.

The vertical handoff (VHO) RAT selection algorithm selects RATs for VHO calls. The objective of the VHO RAT selection algorithm is to keep the high-path-loss users connected to FDMA/CDMA, and the low-path-loss users to CDMA, depending on how propagation conditions vary during the session lifetime. However, in order to avoid undesired ping-pong effects, leading to continuous RAT changes for users with path loss close to the threshold PL_{th} , a hysteresis margin Δ (dB) is introduced. This hysteresis margin together with a number of consecutive samples that must be taken specify the conditions that must be satisfied before the VHO decision is taken. The number of consecutive samples is M_{up} when the path loss condition is above the threshold, and M_{down} when the path loss condition is below the threshold.

3.1.2.5 Service-cost- based JCAC

Service cost-based JCAC algorithms admit incoming calls into the least expensive RAT so that the subscriber incurs the least service cost in the heterogeneous wireless network. This approach is based on the fact that service cost differs from one RAT to another. Variation in service cost can be attributed to the cost of equipment and the cost of procuring spectrum license. These algorithms have the advantage of reducing the overall service cost incurred by subscribers. However, they can lead to highly unbalanced network load.

In [38], the service cost reduction benefit of a service-cost based JCAC algorithm is evaluated for integrated WLAN/UMTS. Given that the cost of service is cheaper in WLAN than in UMTS network, the algorithm tries to admit incoming calls into WLAN as much as possible in order to reduce overall service cost incurred by the subscriber. A Markov chain model is developed to evaluate the service-cost reduction benefit of the algorithm for a single mobile user of real-time service. Fig. 3-2 shows the state transition diagram in the WLAN-UMTS network.

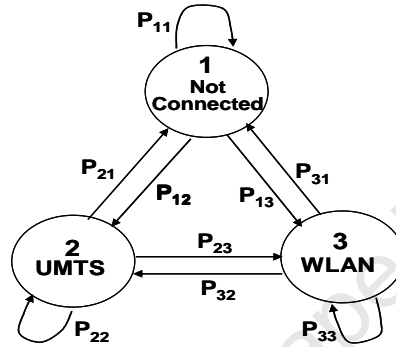


Figure 3-2. State transition diagram for integrated UMTS-WLAN network.

Performance metric used is *service cost (%) versus availability of WLAN coverage*. The total service cost C_T incurred by the user in the integrated network during time T is given by

$$C_T = \sum_{i=1}^3 C_i \cdot \tau_i \quad (3.4)$$

where C_i is the service cost per unit time in state i and τ_i is the sojourn time in state i ($i=1,2,3$). C_1 is always equal to 0.

The service cost incurred by the subscriber in the integrated WLAN-UMTS network is expressed as a percentage of the service cost that will be incurred in homogeneous UMTS network. Results show that the overall service cost incurred by the subscriber in integrated network reduces as the WLAN coverage increases. When WLAN coverage is about 60% of that of UMTS, and the service cost per unit time in WLAN is about 50% of that of UMTS, the overall service cost incurred by the subscriber in the integrated WLAN-UMTS

network is about 80% of that of the homogeneous UMTS network. However, this approach cannot be used to evaluate the service cost incurred by a subscriber using a non-real time service, in which case, the service cost incurred in each state does not necessarily depend on the sojourn time.

3.1.2.6 Layer-based JCAC

In overlaid networks, layer-based JCAC algorithm admits calls into a particular layer. If the layer cannot accommodate the call, the JCAC algorithm tries to admit the call in next available layer. Layer based algorithms are simple but can lead to highly unbalanced network load.

Ali and Pierre [18] proposed a layer-based predictive JCAC algorithm for overlaid heterogeneous wireless networks. The algorithm tries to admit an incoming call into a particular layer, say layer k . If the call is blocked in layer k due to unavailability of radio resource to accommodate the call, the algorithm seeks to admit the call in the next lower layer until it tries all the available layers. The call is finally blocked if it cannot be admitted into any of the available layers. In a similar way, the algorithm seeks to admit an incoming handoff call into a new cell within the current layer (e.g., layer k). If the handoff calls cannot be admitted into the current layer (k), the algorithm tries to admit the call into the next layer ($k-1$). The objective of the algorithm is to minimize new call blocking probability while guaranteeing a hard constraint on handoff call dropping probability. Guard bands are reserved in each layer to prioritize handoff calls over new calls. New call blocking probability is minimized by searching for optimal guard bandwidth to be reserved in each layer. Using artificial neural network technique, the algorithm predicts call traffic and mobility parameters based on aggregated history of user call sessions and cell visits. The predicted information is used to determine the number of guard channel to be reserved and the performance of the algorithm depends on the accuracy of the predicted information. Layer-based JCAC algorithm may lead to unbalanced load among the available RATs.

Performance of the proposed layer-based scheme is evaluated using a trace driven simulation of a 24 hours call and mobility traffic of voice calls from SUMATRA (Stanford

University Activity TRAcEs). A three-RAT heterogeneous wireless network consisting of GPRS, UMTS, and WLAN is considered. One GPRS cell covers 7 UMTS cells and one UMTS cell covers 7 WLAN cells. The coverage outage of WLAN is 50% and for UMTS is 75%.

Performance of the layer-based predictive JCAC scheme is compared with that of separate-layer predictive CAC scheme (i.e there is independent admission control among different RATs) in the same heterogeneous network. Results show that the proposed JCAC scheme improves the overall call blocking probability by about 97% when compared to the separate layer predictive CAC scheme. However, the proposed JCAC scheme can lead to highly-unbalanced load among the three layers (RATs) considered.

3.1.2.7 Non computational-intelligence-based JCAC

Incoming calls are admitted into a particular RAT based on some utility function or cost function derived from a number of criteria without the use of computational intelligence techniques. These algorithms are very efficient but are often complex and incur high computational overhead.

Ormond *et al* [38] proposed a utility-based algorithm that accounts for user time constraints, estimates complete file delivery time (for each available access network), and then selects the most promising access network based on consumer surplus (CS) difference. The algorithm is designed for non-real-time services. It is assumed that every user has a patience limit with a threshold value for the duration the user is willing to wait for the complete transfer of his/her data. Beyond this threshold, the user becomes dissatisfied and unwilling to pay any money for the file delivery.

In the proposed scheme, a multi-mode mobile terminal initiating a call will survey the radio interfaces and determine a list of current available access networks. The flow chart of the proposed JCAC is shown in Fig. 3-3.

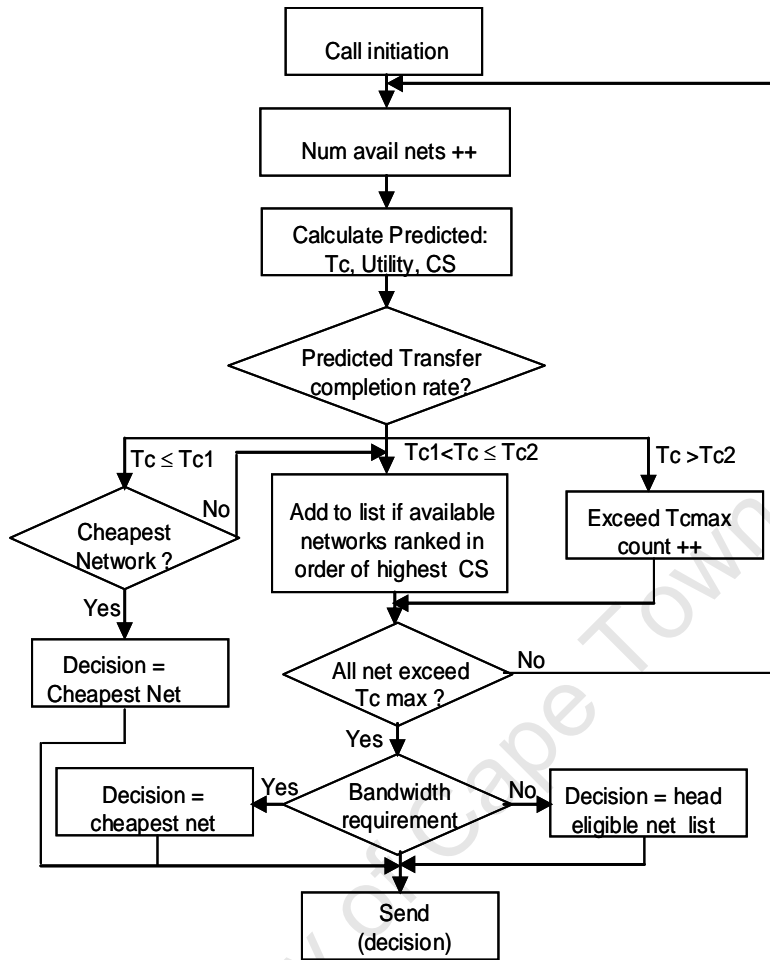


Figure 3-3. Flow chart of utility-based JCAC [38].

As shown in Fig. 3-3, the proposed consumer surplus-based JCAC algorithm first evaluates the predicted completion time (T_c), the predicted utility (U_i), and consumer surplus for each candidate network. The network with the best predicted CS, which is also predicted to meet the completion deadline, is then selected as the most suitable network for the data transaction. T_{c1} denotes the user's best expectation for transfer completion time and T_{c2} the maximum transfer completion time that a user is willing to wait. Three examples of utility functions are considered in the algorithm.

A scenario containing two partly overlapping WLANs, each with a number of terminals generating background traffic is considered. The simulation model is developed in NS2 version 2.27 with IEEE 802.11b wireless LAN parameter settings (data rate 11 Mbps).

Performance of the algorithm is evaluated in terms of *average completion time (s) against file size (KB)*, *average price per file (Cent) against file size (KB)*, and *percentage transfers over Tc2 against file size (KB)*.

Results show that the performance of the algorithm depends on the utility function employed. The differences in the performance of the algorithm for the three utility functions also increase as the file size increases. For instance, for a small file size (i.e. less than 1000 KB), the difference in percentage transfers over Tc2 among the three utility functions is less than 10. However, at 1500 KB, the difference is about 15.

Chen *et al* [22] proposed a network selection and radio resource allocation algorithm for heterogeneous wireless networks. The algorithm is based on a concept called arbitration probability which indicates the willingness of a data user to use a network's resources. Arbitration probability is calculated for each available network using relative link quality, user's satisfaction on quality of service, and monetary cost. It is assumed that each network broadcasts its access bandwidth value to users. After calculating the arbitration probability for each network, a user can decide which network is suitable to bear its service.

The performance of the proposed algorithm is evaluated using Network Simulator 2 (NS-2). In order to simplify the simulations, two WLANs (IEEE 802.11a and IEEE 802.11g) are used. Each of the WLANs has a single access point with overlapping coverage with 50 mobile users located in the coverage area. Performance of the proposed RAT selection scheme is evaluated in terms of *network revenue against simulation time*. Results shows that from 40-200 seconds of the simulation, network revenue varies from 15-23 units for the proposed arbitration scheme whereas it varies from about 7-17 units for a scenario without the proposed arbitration scheme. Thus the proposed arbitration probability based network selection scheme improves network revenue in the heterogeneous network.

3.1.2.8 Computational- intelligence-based JCAC

Computation intelligence based JCAC algorithms choose a RAT for an incoming call by applying a computational intelligence technique to some RAT selection criteria. Computational intelligence techniques commonly used are Fuzzy logic [19], Fuzzy-neural

[3], Fuzzy MADM (Multiple Attribute Decision Making) method [20, 21, 40], and genetic algorithm. Computation-intelligence-based JCAC algorithms have high efficiency, and improve users' satisfaction in the heterogeneous wireless network. However, they are complicated.

Most computational-intelligence-based JCAC algorithms incorporate fuzzy logic. A fuzzy logic controller consists of a fuzzifier, inference engine, fuzzy rule base, and a defuzzifier. These components are illustrated in Fig. 3-4.

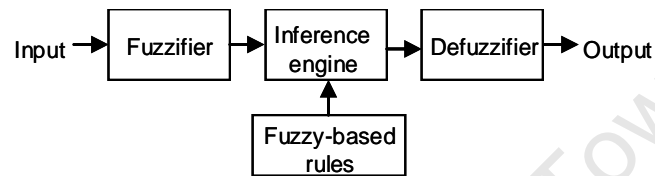


Figure 3-4. Fuzzy logic controller.

The fuzzifier translates the input numerical measurements to the corresponding linguistic values of the fuzzy sets in the input universe discourse. Examples of fuzzy variables commonly used are signal strength (SS) received by the multimode MT from the base station (or access point) in each cell, and the current load (L) in the cell. The inference engine makes use of some predefined fuzzy rules to determine, for each RAT, whether the incoming call can be admitted into the selected cell or not. The predefined rules are a series of “If then” rules. Defuzzification involves the conversion of the fuzzy outputs into crisp output. Common defuzzification methods are weighted average method, centroid method, etc.

Wilson *et al* [21] proposed a decision strategy for making the optimal choice of wireless access networks. Fuzzy logic is used as the inference mechanism and a prototype is developed. The prototype uses two metrics from a candidate network, a metric from application requirements, and user defined criteria as input. The Fuzzy-based selection strategy is shown in Fig. 3-5.

The prototype is tested using only the two metrics, one from a candidate network and the other metric from application requirements. Results obtain show a good approximation for

level of fitness. One critical issue about Fuzzy logic is definition of Fuzzy sets and rules. In general, methods based on fuzzy logic are cumbersome to use, which require much expert knowledge and user involvement in order to make decision rules [41].

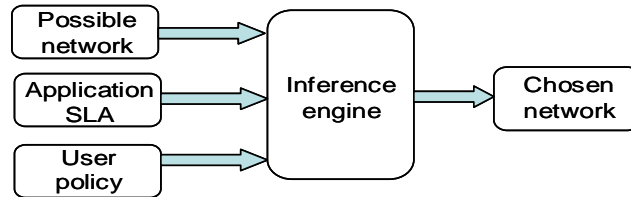


Figure 3-5. Fuzzy logic decision strategy [21].

Agusti *et al* [3] presented a methodology based on fuzzy logic and reinforcement learning mechanisms that combines technical and economical issues to provide the specific RAT and bandwidth allocations in integrated 3G-WLAN networks.

Zhang [20] proposed an approach which uses fuzzy logic to represent imprecise information of some attributes. The fuzzy MADM method operates in two steps. The first step is to convert the imprecise fuzzy variable to crisp numbers. The second step is to use classical MADM technique to determine the ranking order of the candidate networks. The highest ranking RAT is then selected for the call.

In [19], a segment selection algorithm based on the concept of fuzzy multiple objective decision making (MODM) is proposed by Chan *et al*. Seven example criteria are used in the algorithm namely, signal strength, bandwidth, charging model, reliability, latency, battery status and the user's preferred segment (priority). The purpose of the segment selection algorithm is to select the most suitable segment for a particular service class based on the criteria mentioned above.

Karabudak *et al* [10] proposed a cost function based JCAC algorithm incorporating genetic algorithms. The objective of the algorithm is to maximize wireless network utilization, meet mobile terminal QoS requirements, and reduce handoff latency. In the proposed algorithm, all the parameters that affect handoff process in each network such as signaling

(Sig), switching (Sw), bandwidth (Pw), and power consumption (Bw) are fed into a cost function. The cost function is as follows:

$$(\text{Real Cost})_N = F(\text{Sig}_N, \text{Sw}_N, \text{Pw}_N, \text{Bw}_N) \quad (3.5)$$

where $(\text{Real Cost})_N$ is the cost of handoff in RAT-N, and $\text{Sig}_N, \text{Sw}_N, \text{Pw}_N, \text{Bw}_N$ represent Sig, Sw, Pw, Bw in RAT-N respectively. Based on the cost function, an optimization problem is formulated using Continuous Time Markov Decision Process [39]. Genetic algorithms are then used to optimize the model and to make the final RAT selection decision. Simulation experiments are performed for the proposed algorithm using different handoff scenarios. For comparison purposes, a heuristic algorithm is also implemented using value iteration algorithm and linear searching for the JCAC in the same heterogeneous wireless systems. The simulation environments for the proposed JCAC algorithm, heuristic, and the other algorithms used are developed using Java, version j2sdk-1-4-1. The performance metrics used are (1) *latency against handoff events*, (2) *handoff cost against handoff events*, and (3) *admission percentage against handoff scenario*. Results show that the proposed GAC algorithm reduces latency by up to 50% of that of the heuristics-based JCAC algorithm.

3.1.3 Multiple-Criteria JCAC Algorithms

Multiple criteria JCAC algorithms make RAT selection decision based on many selection criteria. The non computational-intelligence-based JCAC algorithms and the computational-intelligence-based JCAC algorithms previously discussed under section 3.1.2.7 and section 3.1.2.8 are based on multiple criteria. The purpose of multiple-criteria JCAC algorithms is to the combine the various criteria in order to select the most suitable RAT for a new call session or handover call [41]. RAT selection criteria are described in the following. Generally, RAT selection criteria can be based on user's preference, operator' preference, or combination of both. These preferences are discussed in the following subsections.

3.1.3.1 RAT selection based on user's preferences

Next generation wireless network will be user-centric. Therefore users' preferences for a particular RAT should be considered in making call admission decisions. Users can indicate their preferences for a particular RAT when making access request, and can even dynamically change their preferences with time.

Factors that determine users' preference for a RAT are as follows:

(1) Least service cost: Service cost varies from one access network to another. As a result, a user may prefer to be connected through the cheapest available RAT so that overall service cost will be reduced.

(2) Minimum Delay: Different traffic types usually require different QoS deliveries whereas different RAT can offer different level of QoS to each of the traffic types. Some services may be better supported on a particular RAT than others. A user may prefer to be connected through the RAT which can offer the minimum delay for a particular service, even at a higher cost.

(3) Maximum data rate: Different access technologies offer different data rates. Most multimedia applications are adaptive. For example voice can be encoded at 16 kbps, 32 kbps, 64 kbps, and 128 kbps by choosing appropriate encoding mechanisms. Similarly, video applications can be made rate adaptive by using, for instance, a layered coding method. In layer coding method, the lowest layer (i.e., the base layer) contains the critical information for decoding the image sequence at its minimum visual quality. Additional layers provide increasing quality. As an illustration, if one watches a 30-minute video-clip encoded at 256 kbps and 64kps respectively. At 256 kbps, one will see better pictures with better resolution than at 64 kbps. Therefore, a user of real-time service may prefer to be connected through the RAT with the highest data rate in order to enhance service quality.

A user of non-real time service may prefer to be connected through the RAT with the highest data rate in order to reduce service delivery time.

(4) **Widest Coverage:** An active mobile subscriber moving at high speed in a heterogeneous network environment will likely experience more handover (vertical and horizontal) during its call lifetime. Vertical handover incurs delay and loss of packets. To reduce the frequency of vertical handover, the user may prefer to be connected through the available RAT with the widest coverage.

(5) **Least battery power consumption:** One of the key challenges in wireless communication is efficient use of energy stored by the batteries of mobile terminals. Efficient power utilization in mobile terminals will avoid the need for frequent batteries recharge. A user may prefer to be connected through a RAT that will minimize its energy consumption.

(6) **Highest network security:** Different access technologies offer different levels of security. A user may prefer to be connected through the RAT with the highest security.

Table 3-4 shows RAT selection criteria based on user preferences.

Table 3-4. Users' preferences.

Users' Preferences	Advantage
Least service cost	Reduces overall service cost incurred by subscribers
Minimum delay (s)	Enhances quality of service
Maximum data rate (Kbps or Mbps)	Reduces service-delivery time for non-real-time services, enhances quality of service for adaptive real-time services
Widest coverage (m or Km)	Reduces handoff frequency for highly mobile subscribers
Least battery power consumption	Increases battery lifetime and reduces recharge frequency
Highest network security	Enhances information confidentiality and integrity

3.1.3.2 RAT selection based on operators' preferences

Operator preferences can be uniform load distribution, revenue maximization, handoff call dropping minimization, or optimal radio resource utilization. These preferences are discussed in the following:

(1) Uniform load distribution: Operators may prefer to distribute the network load among the available RATs in order to avoid overload in some RATs. This approach however, may result in low user satisfaction because their preferences are not considered.

(2) Revenue maximization: Users' willingness to pay for network services varies. Some users (e.g., premium users) are willing to pay high price for high quality of service whereas some users are less willing to pay such a high price, and are less concerned about service quality. An operator's preference may be to admit certain number of users (based on the willingness to pay) into particular RATs in such a way as to maximize the overall revenue obtained from the heterogeneous network.

(3) Call blocking/dropping minimization: The number of RATs available to a mobile subscriber depends on its current location. Fig. 3-6 shows a two RAT-heterogeneous wireless network. As illustrated in Fig. 3-6, MT1 (mobile terminal 1) MT 5, and MT 6 can only be admitted into RAT 1 whereas MT2, MT 3, and MT 4 can be admitted into either of the two RATs.

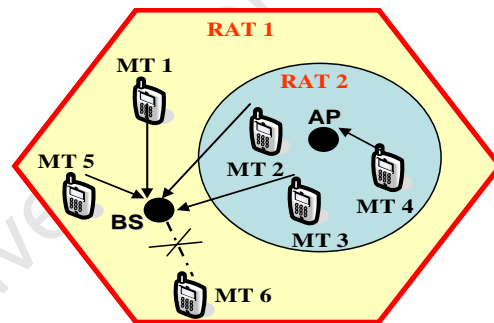


Figure 3-6. Two-RAT heterogeneous wireless network.

Assume that each of the two RATs (RAT 1 and RAT 2) can support four calls. MT1, MT2, MT3, MT5 are admitted into RAT 1 and MT 4 is admitted into RAT 2. An incoming call, MT6, is blocked in RAT 1 due to unavailability of radio resource. If MT 2 or MT 3 had been admitted into RAT 2, it would have been possible to accommodate MT6 in RAT 1. Therefore, RAT selection policy affects the overall call blocking/dropping probability. The operator preference may be to admit users into different RATs so as to minimize the overall call blocking/dropping probability in the heterogeneous wireless network.

(4) Optimal radio resource utilization: Capacity in cellular networks can be expanded by rearranging traffic (both voice and data) among different RATs [29]. Therefore, the operator's preference may be to maximize radio resource utilization by distributing different traffic among the available RATs.

Table 3-5 summarizes RAT selection criteria based on operators' preferences.

Table 3-5. Operators' preferences.

Operators' Preferences	Advantage
Uniform load distribution	Enhances overall network stability. Prevent over subscription of some RATs
Revenue maximization	Increases operator's revenue
Call blocking/dropping minimization	Enhances connection-level QoS
Optimal radio resource utilization	Improves radio resource utilization efficiency

A major challenge in the design of multiple-criteria JCAC algorithms is how to combine many selection criteria in making a RAT selection decision for an incoming call. Current approaches have incorporated non computational-intelligence-based techniques and computational-intelligence-based techniques. These approaches have already been discussed under session 3.1.2.7 and session 3.1.2.8.

3.1.4 JCAC Algorithms Design Considerations

A number of issues have to be considered in the design of JCAC algorithms. In this section, six design considerations for JCAC algorithms are analyzed. These design considerations are shown in Fig. 3-7.

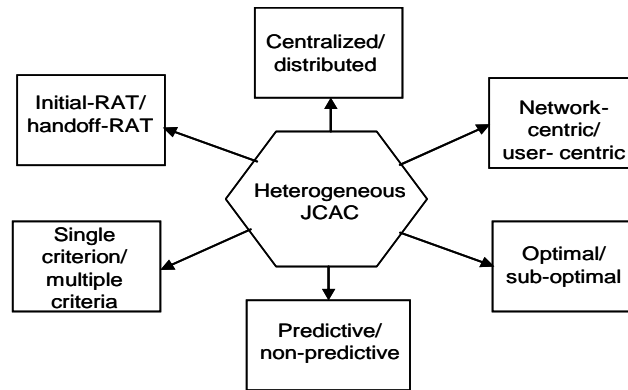


Figure 3-7. Six design consideration for JCAC algorithms.

3.1.4.1 Centralized versus Distributed

This design consideration is based on several centralized and decentralized architectures which have been proposed for implementation of radio resource management (RRM) and common radio resource management (CRRM) entities in heterogeneous wireless networks [42-46].

In a centralized JCAC algorithm, one or more central entities coordinate joint call admission decision. Centralized JCAC algorithms are more efficient than distributed JCAC. However, centralized JCAC requires very frequent interactions among the different entities in the heterogeneous network, thus leading to a high amount of signaling overhead [44]. A centralized JCAC algorithm is not scalable and can also create a bottleneck within the heterogeneous network. Moreover, a centralized approach is not fault-tolerant.

On the other hand, distributed JCAC algorithms are distributed within the heterogeneous network. Every mobile terminal does not have to communicate with a centralized entity. Consequently, signaling overhead is reduced. Another major advantage of distributed JCAC algorithm is scalability. However they are less efficient compared with the centralized JCAC. An example of a distributed JCAC algorithm appears in [34].

3.1.4.2 User-centric versus Network-centric

In user centric JCAC algorithms, user preferences are of most importance in making the choice of most suitable RAT. A multi-mode mobile terminal initiating a call will survey

the radio interfaces, and determine a list of current available access networks. Based on the user's preferences and the information obtained from the available RATs, the most suitable RAT is selected for the call. User-centric JCAC algorithms enhance users' satisfaction and are very suitable for next generation wireless network. An example of user-centric algorithms can be found in [38].

In network-centric JCAC, selection decisions are made by the network. Operators' policies are of major consideration in selecting the most suitable RAT for incoming call. Network-centric JCAC algorithms facilitate network stability and improve overall network management. However network centric JCAC algorithms are more complicated. An example of network-centric RAT selection algorithm appears in [1].

3.1.4.3 Optimal versus Sub-optimal

Optimal JCAC algorithms utilize optimization techniques such as linear programming, descent search, value iteration, and artificial intelligent techniques (such as genetic algorithms [10]) to search for some optimal value. The objective of optimizations in these algorithms is to maximize or minimize certain parameters (e.g., maximize radio resource utilization, maximize operator's revenue, minimize new call blocking probability, etc.). Optimal JCAC algorithms are very efficient but are usually very difficult to implement due to high level of complexity. Yu and Krishnamurthy [47] proposed an optimal JCAC that maximizes overall network revenue while satisfying the quality of service constraints in both the WLAN and the CDMA. Nasser [48] proposed an optimal JCAC to maximize system utilization in integrated WLAN/UMTS network. These two optimal algorithms [47, 48] are based on Semi Markov Decision Process and Linear programming optimization technique.

On the other hand, suboptimal JCAC algorithms are less efficient but are more realistic and easier to implement. Example of suboptimal JCAC algorithm can be found in [34].

3.1.4.4 Predictive versus Non-predictive

Predictive JCAC algorithms incorporate prediction techniques in selecting the most suitable RAT for a new or inter-system handoff call [18, 49]. Predicted information may be

user's mobility pattern, call holding time, network load condition, etc. Predictive JCAC algorithms are more efficient but are more prone to error. The overall performance depends on the accuracy of the predicted information.

Kafle *et al* [49] proposed an algorithm which uses prediction of user mobility pattern to select an access network that maximizes the expected value of user satisfaction in heterogeneous network environment. User satisfaction is formulated as a function of bandwidth utility and handoff latency. In the proposed algorithm, when a call request arrives, the network selection algorithm first predicts the user's mobility pattern and call holding time. It then lists all access networks that are available from the user's current location and that can serve the call request. The list is sorted in descending order of bandwidths. Starting from the network that has the largest bandwidth on the list, the probability that the user moves from the network under consideration is estimated based on the network layout and user's mobility pattern. This probability is used to estimate the effective user satisfaction. The value obtained is compared with the bandwidth utility function of the next network on the list. If the user satisfaction from the use of current network under consideration is larger than the utility function of next network, the current network is selected. Otherwise, the next network on the list is chosen and the above procedure is repeated to find an access network with optimal user satisfaction. The major disadvantage of this algorithm is the processing overhead required in keeping up-to-date information about the users in the network.

Non predictive RAT selection algorithms on the other hand do not incorporate prediction techniques. RAT selection decisions are based on the available information. They are less prone to error but not as efficient as predictive JCAC algorithms. Examples of non-predictive RAT selection algorithms can be found in [15, 34].

3.1.4.5 Single-criterion versus Multiple-criteria

Single JCAC algorithms make RAT selection based on one criterion. They are simple and easy to implement but they are not as efficient as multiple-criteria JCAC. An Example of this algorithm can be found in [15]. On the other hand, multiple criteria JCAC algorithms make RAT selection based on multiple criteria. They are more efficient but are at the

expense of higher complexity. They often incorporate fuzzy logic or fuzzy multi attribute decision making (MADM) techniques. Examples of multiple criteria JCAC can be found in [21].

3.1.4.6 Initial RAT versus Handoff RAT

Initial-RAT selection JCAC algorithms are designed for new call sessions while handoff-RAT JCAC algorithms are designed for handover call sessions. Examples of Initial RAT selection algorithm can be found in [15]. An example of inter-system handover RAT selection algorithms can be found in [20].

The six design considerations are summarized in Table 3-6.

University of Cape Town

Table 3-6. Design considerations for JCAC algorithms.

Design Consideration	Design options	Advantage	Disadvantage	Comment
Centralization	Centralized	More efficient	Complex, non-scalable, high signaling overhead, not-fault tolerant	Impractical
	Distributed	Scalable, simple	Less efficient	Commonly used
Centricity	Network-centric	High efficiency and improved overall network management	More complex	-
	User-centric	Enhance users satisfaction	Unbalanced heterogeneous network load	Preferable to subscribers
Optimality	Optimal	More efficient	Complex	More desirable, incorporates optimization techniques e.g., linear programming, genetic algorithm
	Sub optimal	More realistic, scalable	Less efficient	More realistic
Prediction	Predictive	More efficient	More error-prone	Performance depends on the accuracy of the predicted information
	Non-predictive	Simple, more accurate	Less efficient	-
Network selection criteria	Single criterion e.g., network load, call service class, etc.	Simple	Less efficient	-
	Multiple criteria	More efficient	Cumbersome	Often incorporate Fuzzy logic, Fuzzy MADM technique, etc.
Call type	Initial RAT selection	Simple	-	Selects RAT new calls
	Handoff RAT selection		More complicated	Selects RAT for vertical handoff calls

3.2 Bandwidth Management Schemes

In this section, bandwidth allocation strategies for wireless networks are reviewed. Bandwidth allocation strategies for wireless networks can be classified into four groups, which are complete sharing, complete partitioning, handoff call prioritization, and service class prioritization. This classification is summarized in Table 3-7, and subsequently described in this section.

Table 3-7. Summary of bandwidth allocation strategies for wireless networks

Bandwidth Allocation Strategy	Main Idea	Advantage	Disadvantage
Complete Sharing	An incoming call is accepted, regardless of the class/ type, as long as there is enough radio resource to accommodate it.	Implementation simplicity and high radio resource utilization	High handoff call dropping probability. No differential treatment for calls with stringent QoS requirements
Complete Partitioning	Available bandwidth is partitioned into pools and each pool is dedicated to a particular type of calls. An incoming call can only be admitted into a particular pool.	Implementation simplicity	Poor radio resource utilization
Handoff Call Prioritization	Handoff calls are given more access to radio resources than new calls. New calls may be blocked whereas handoff calls are still being admitted.	Low handoff call dropping probability	High new call blocking probability
Service-Class Prioritization	Certain classes of calls are given preferential treatment over some other classes of calls. For example, class-1 calls may be blocked whereas class-2 calls are still being admitted.	Differential treatments of calls based on QoS requirements	Implementation complexity

3.2.1 Complete Sharing

Complete sharing scheme is a first come first serve scheme and it is the simplest bandwidth allocation policy. It is a non-prioritization scheme in which new and handoff calls are treated the same way. An incoming call is accepted as long as there is enough radio resource to accommodate it. When the network gets to its maximum capacity, a new call will be blocked while a handoff call will be dropped. Two major advantages of complete sharing CAC scheme are implementation simplicity and good radio resource utilization. However, it has a high handoff call dropping probability because it does not give preference to any call. Consequently, complete sharing CAC scheme has a poor QoS performance [50]. Fig 3-8 is the state transition diagram for complete sharing scheme where $\lambda_n, \lambda_h, \mu_n$ and μ_h represent new call arrival rate, handoff call arrival rate, new call departure rate, and handoff call departure rate respectively.

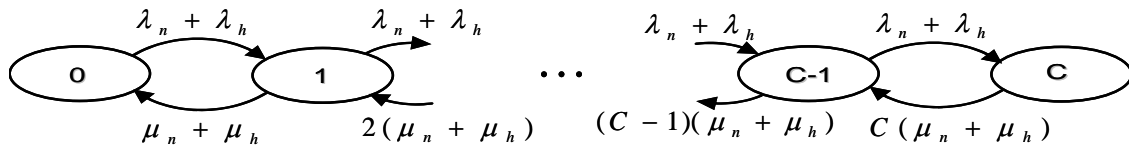


Figure 3-8. State transition diagram for complete sharing policy.

3.2.2 Complete Partitioning

In the complete partitioning CAC scheme, entire available bandwidth is partitioned into pools. Each pool is dedicated to a particular type of calls (new or handoff calls) and/or particular traffic class of calls. An incoming call is admitted if there is an available channel in the pool allocated for the type/class of the incoming call. This policy allocates a fixed bandwidth C_1 (C_2) to service s_1 (s_2) such that $C_1 + C_2 \leq C$. The acceptable states of this policy are a subset of the complete sharing case. This is a case of two independent queues, and the blocking probability is given by the well known Erlang-B formula.

Fig. 3-9 and Fig 3-10 are the state transition diagrams of a system where the available resource (C) is partitioned into two (C_1 and C_2). C_1 is used for new calls (Fig. 3-9) whereas C_2 is used for handoff calls (Fig. 3-10).

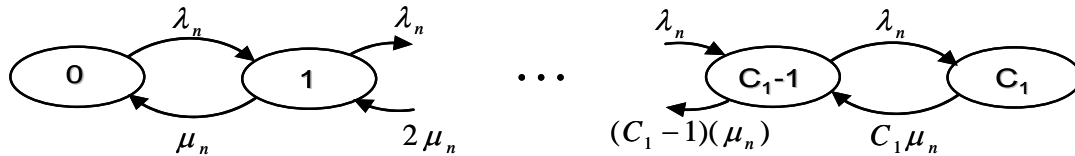


Figure 3-9. State transition diagram for complete partitioning policy: first partition.

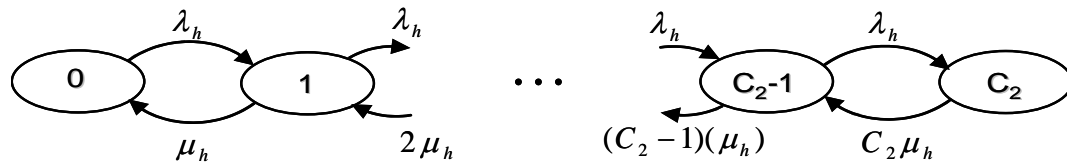


Figure 3-10. State transition diagram for complete partitioning policy: second partition.

Lai *et al* [51] proposed a fair call admission control which divides the available bandwidth into segments, and group call requests into different categories such that call requests in group-*i* can only be accepted if there is enough bandwidth in segment-*i*. Though this CAC approach is simple to implement, the major disadvantage is that it leads to poor radio resource utilization. For instance, it is possible to have unused bandwidth in one segment whereas call requests in other segments are rejected due to unavailability of bandwidth [52].

In [53], Kanter proposed a movable boundary scheme for a system which integrates voice and data. In the scheme, separate bandwidths are assigned to voice and data (i.e. complete partitioning). However, the boundary for the partition is movable and therefore can effectively deal with the traffic variation.

3.2.3 Handoff Call Prioritization

Due to users' mobility within the coverage of wireless networks, an accepted call that has not been completed in the current cell has to be transferred (handed over) to another cell. This call may not be able to get a channel in the new cell to continue its service due to limited radio resources in wireless networks. Eventually, it may be dropped. However, wireless network subscribers are more intolerant to dropping a handoff call than blocking a new call. Therefore, in order to ensure that handoff call dropping probability is kept below

a certain level, handoff calls are usually admitted with a higher priority compared with new calls. Handoff call prioritization has an advantage of low handoff call dropping probability. However, the advantage of low handoff call probability is at the expense of new call blocking probability, which is high. The concept of handoff calls prioritization in wireless networks was first introduced in the mid-1980s [54]. Since then several handoff-priority-based schemes have been proposed [55]. These schemes are reviewed in the following subsection.

3.2.3.1 Guard Channel

In this scheme, some channels (referred to as guard channels) are specifically reserved in each cell to take care of handoff calls. For example, if the total number of available channels in a single cell is C and the number of guard channels is $C - K$, a new call is accepted if the total number of channels used by ongoing calls (i.e., busy channels) is less than the threshold K , whereas a handoff call is always accepted if there is an available channel. According to this channel reservation, the threshold must be chosen such that the handoff call dropping probability is as low as possible, while the system can admit as many incoming new calls as possible. Fig. 3-11 shows the state transition diagram for a single-class service using guard bandwidth scheme.

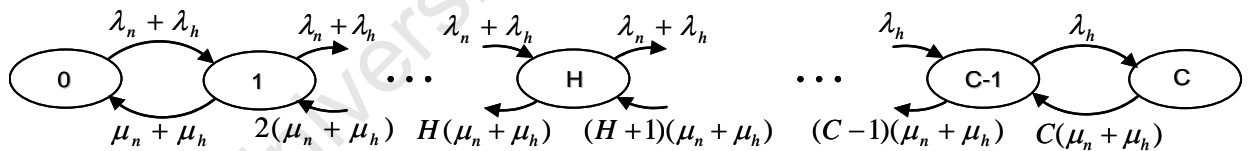


Figure 3-11. State transition diagram for guard bandwidth scheme.

Guard channel (GC) scheme can be divided into two categories: static and dynamic strategies. Hong and Rappaport [54] used a fixed GC to give preferential treatment to handoff calls considering only one service class. Rappaport and Purzynski [56] improved on the fixed GC scheme by considering multiple services. Chen *et al* [57] introduced multiple thresholds to deal with multimedia traffic with different priorities. The major advantage of static guard channel scheme is simplicity because there is no need for exchange of control information between base stations. However, static guard channel schemes have been

shown to be inefficient since they cannot adapt to quick variation in the traffic condition. This has led to the development of dynamic guard channel scheme.

Dynamic channel scheme improves the system efficiency by adaptively changing the number of channels reserved at periodic time intervals. Channel reservation is based on traffic condition and estimated handoff rate from the neighboring cells. Estimated handoff rate is derived from the number of calls in the neighboring cells, handoff history, mobility pattern of the calls, etc. Chen *et al* [58] proposed a dynamic call admission scheme for QoS priority handoff in multimedia homogeneous cellular system.

3.2.3.2 Fractional Guard Channel

In fractional guard channel scheme, handoff calls are prioritized over new calls by accepting an incoming new call with a certain probability that depends on the number of busy channels. In other words, when the number of busy channels becomes larger, the acceptance probability for a new call becomes smaller, and vice versa. This approach helps to reduce the handoff call dropping probability.

The policy has a threshold, H for limiting the acceptance of new calls. A handoff is accepted as long as there is a channel available. Before the wireless system gets to threshold, H , new calls are accepted with a probability of 1. After threshold, H , a new call is accepted with a probability of α_p where $0 \leq \alpha_p \leq 1$ and $H < p < C$. New calls are rejected when the system reaches the maximum capacity. Fig. 3-12 is the state transition diagram for fractional guard bandwidth policy.

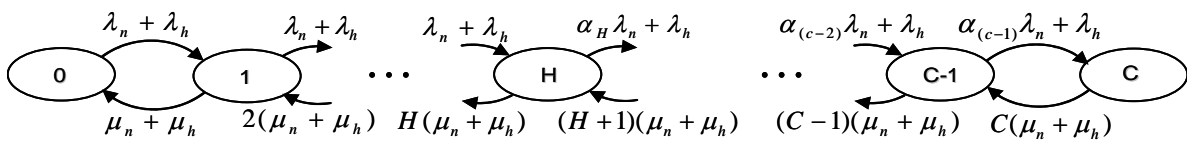


Figure 3-12. State transition diagram for fractional guard bandwidth policy.

Limited Fractional Guard Channel policy is a special case of fractional guard channel policy [59]. Handoff calls are accepted as long as there is a channel available. Before the system gets to threshold, H , new calls are accepted with a probability of 1. At threshold, H ,

a new call is accepted with a probability of α_p . After threshold, H, new calls are blocked, only handoff calls are accepted.

3.2.3.3 Queuing Priority Scheme

Queuing priority scheme accepts calls (new and handoff) whenever there are free channels. When all the channels are occupied, handoff calls are queued while new calls are blocked [60] or all incoming calls are queued with certain rearrangement in the queue. When radio resource becomes available, one or some of the calls in the handoff queue are served until there is no more resource. The remaining calls are queued until resource becomes available again. However, a call is only queued for a certain period of time. If radio resource is not available within this period, the call will be dropped.

The main disadvantage of queuing priority scheme is that it needs a lot of buffers to deal with real-time multimedia traffic. It also needs a sophisticated scheduling mechanism in order to meet the QoS requirements of delay-sensitive calls (i.e. to guarantee that the queued data will be transmitted without excessive delay) [58].

3.2.3.4 QoS Degradation Scheme

QoS degradation can either be bandwidth degradation or delay degradation. In bandwidth degradation method, calls are categorized as adaptive (degradable) and non-adaptive (non-degradable) calls. Degradable calls have flexible QoS requirements (e.g., minimum and maximum data rates). For most multimedia applications, e.g., voice over IP or video conferencing, service can be degraded temporarily as long as it is still within the pre-defined range. Bandwidth degradation reduces handoff call dropping by reducing the bandwidth of the ongoing adaptive calls during network congestion. When a handoff call arrives and there is network congestion, the system is able to free some radio resource to admit the handoff calls by degrading some of the ongoing adaptive calls.

In [61], bandwidth degradation is used to prioritize handover call requests over new call requests by temporally degrading the bandwidth of some ongoing adaptive calls. Once the total required bandwidth exceeds the cell capacity, the system reduces the bandwidth currently assigned to degradable calls in order to admit handover calls, and hence reduces

the probability of handover failures. In [62], when the network is operating at maximum capacity, bandwidth is borrowed from existing adaptive calls to admit handoff calls. This approach gives priority to handoff calls without affecting the minimum QoS requirements of on-going calls. It also results in better utilization of resources.

In delay degradation method, the amount of radio resources allocated to non-real-time (delay-tolerant) services is reduced during network congestion. When a handoff call arrives and there is no radio resource to accommodate the handoff call. Some non-real-time services are degraded to free some bandwidth, which is used to accommodate the incoming handoff call. Sen *et al* [63] presented an analysis of queue build-up (and delay degradation) for non-real-time packets in a mixed traffic scenario.

3.2.4 Service-Class Prioritization

In wireless systems which support multiple service classes, the limited bandwidth has to be shared among the multiple traffic classes. Complete sharing scheme allows the network radio resource to be shared among the various service classes without preference for any class. However, one major challenge in the design of CAC policy is to provide preferential treatment among users of different service classes while still utilizing the system resources efficiently [63]. Preferential treatments are given to certain classes of calls for the following reasons: (1) some calls (such as voice call) have stringent QoS requirements and therefore require preferential treatment. (2) Some subscribers in a particular service class are willing to pay more for better QoS. Service class prioritization scheme is more complicated than complete sharing and complete partitioning schemes.

Bartolini and Chamtac [64] proposed a model for multi-class environment that permits call transition between different classes. They also show that under some assumptions, the optimal policy has the shape of multi-priority threshold policy. Aboelaze [65] proposed a CAC algorithm that supports differentiated fairness among different service classes.

It should be noted that the bandwidth management schemes reviewed in section 3.2 have been mainly applied to homogeneous wireless networks. Therefore, it is very necessary to study bandwidth management in emerging heterogeneous wireless networks.

3.3 Summary

This chapter surveys existing JCAC algorithms for heterogeneous wireless networks and bandwidth management schemes for wireless networks. Efficient radio resource utilization and QoS provisioning are key requirements in wireless communication. Existing JCAC algorithms are broadly classified into single-criterion and multiple criteria JCAC scheme. Among single-criterion JCAC algorithms, load-based JCAC algorithm has the advantage of uniformly distributing traffic load among available RATs in a heterogeneous wireless network. Uniform distribution of traffic load allows for better utilization of radio resources and enhances QoS. Multiple criteria-JCAC algorithms combine many criteria in making RAT selection decision. However multiple criteria JCAC algorithm can lead to highly unbalanced network load in the heterogeneous wireless network due to independent users' preferences. Four bandwidth management schemes namely complete sharing, compete partitioning, handoff call prioritization, and service class prioritization schemes are reviewed. In practice it is essential to prioritize handoff calls over new calls and sometimes necessary to prioritize among different classed of calls. In chapter 4, load-base JCAC and adaptive bandwidth management are combined to enhanced radio resource utilization and connection-level QoS in heterogeneous wireless networks.

4 Adaptive Bandwidth Management and JCAC Scheme

4.1 Introduction

This chapter proposes an adaptive bandwidth management and joint call admission control (AJCAC) scheme to enhance QoS and system utilization in heterogeneous wireless networks supporting multiple classes of calls such as voice and video. The proposed AJCAC scheme is designed to simultaneously achieve the following objectives in heterogeneous wireless networks:

- (1) Distribute traffic load uniformly among available RATs to improve overall system utilization, and reduce overall call blocking/dropping probability,
- (2) Guarantee the QoS requirements of all admitted calls,
- (3) Prioritize handoff calls over new calls,
- (4) Adapt the bandwidth of ongoing calls to improve system utilization and reduce call blocking/dropping probability.

Uniform distribution of traffic load among multiple RATs in heterogeneous wireless networks allows for a better utilization of the radio resources. Balancing of traffic load prevents over subscription of RATs that are nearly filled to maximum capacity. Besides, balancing of traffic load among available RATs reduces the frequency of vertical handoff. If the traffic load is evenly distributed among overlapping cells of available RATs in a heterogeneous wireless network, it will be possible to accommodate a handoff call in a neighbouring cell belonging to the same RAT as the current cell of the call (horizontal handoff). However, if traffic load is not evenly distributed among the overlapping cells, the frequency of vertical handoff will increase.

QoS requirements of all admitted calls are guaranteed by allocating at least the minimum bandwidth needed to each of the admitted calls. Handoff calls are prioritized over new

calls by using different call rejection thresholds for new and handoff calls, and by using different bandwidth adaptation mechanism for new and handoff calls.

The contributions of this chapter are twofold. Firstly, adaptive bandwidth management is combined with JCAC to enhance connection-level QoS and system utilization in heterogeneous wireless networks. The second contribution is the development of an analytical model for the AJCAC scheme and the derivation of overall system utilization, new call blocking probability, handoff call dropping probability, and investigation of the tradeoffs between new call blocking probability and handoff call dropping probability.

4.2 System Model

The study considers a heterogeneous wireless network, which consists of J number of RATs with co-located cells, similar to [12, 13, 16]. Wireless networks such as GSM, GPRS, UMTS, EV-DO, etc., can have the same and fully overlapped coverage, which is technically feasible, and may also save installation cost [66]. Fig. 4-1 illustrates a two-RAT heterogeneous wireless network.

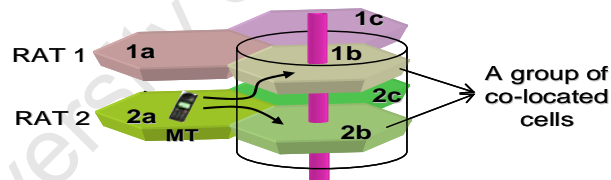


Figure 4-1. A two-RAT heterogeneous wireless network with co-located cells.

In heterogeneous wireless networks, radio resources can be independently or jointly managed. The study considers a situation where radio resources are jointly managed in the heterogeneous network and each cell in RAT- j ($j=1, \dots, J$) has a total of B_j basic bandwidth units (bbu). The physical meaning of a unit of radio resources (such as time slots, code sequence, etc.) is dependent on the specific technological implementation of the radio interface [67]. However, no matter which multiple access technology (FDMA, TDMA, CDMA, or OFDM) is used, system capacity can be interpreted in terms of effective or equivalent bandwidth [68-70]. Therefore, the bandwidth of a call refers to the number of

bbu that is adequate for guaranteeing the desired QoS for the call, which is similar to the approach used for homogeneous networks in [70-72]. It is assumed that packet-level QoS is stochastically assured by allocating at least the minimum effective bandwidth required to guarantee a given maximum probability on packet drop, delay, and jitter [73].

The approach used in this research is based on decomposing heterogeneous wireless network into groups of co-located cells. As shown in Fig. 4-1, cell 1a and cell 2a form a group of co-located cells. Similarly, cell 1b and cell 2b form another group of co-located cells, and so on. Based on the following assumption commonly made in homogeneous cellular networks, it is assumed that the types and amount of traffic are statistically the same in all cells of each RATs [70, 71, 74, 75, 76, 77]. Therefore, the types and amount of traffic are statistically the same in all groups of co-located cells.

A newly arriving call will be admitted into one of the cells in the group of co-located cells where the call is located. When a mobile subscriber using a multimode terminal and having an ongoing call is moving from one group of co-located cells to another group of co-located cells, the ongoing call must be handed over to one of the cells in the new group of co-located cells. For example (Fig. 4-1), an ongoing call can be handed over from cell 2a to cell 2b or from cell 2a to cell 1b. Note that the handover consists of both horizontal and vertical handovers. The correlation between the groups of co-located cells results from handoff connections between the cells of corresponding groups. Under this formulation, each group of co-located cells can be modeled and analyzed individually. Therefore, the study focuses on a single group of co-located cells.

The heterogeneous wireless network supports K classes of calls. Each class is characterized by bandwidth requirement, arrival distribution, and channel holding time. Each class- i call requires a discrete bandwidth value, $b_{i,w}$, where $b_{i,w}$ belongs to the set $B_i = \{b_{i,w}\}$ for $i = 1, 2, \dots, K$ and $w = 1, 2, \dots, W_i$. W_i is the number of different bandwidth values that a class- i call can be allocated. $b_{i,1}$ (also denoted as $b_{i,min}$) and b_{i,W_i} (also denoted as $b_{i,max}$) are respectively, the minimum and maximum bandwidth that can be allocated to a class- i call. Note that $b_{i,w} < b_{i,(w+1)}$ for $i = 1, 2, \dots, K$ and $w = 1, 2, \dots, (W_i - 1)$.

The requested bandwidth of an incoming class- i call is denoted by $b_{i, req}$, where $b_{i, req} \in B_i$. Let $m_{i,j}$ and $n_{i,j}$ denote respectively, the number of ongoing new class- i calls and handoff class- i calls, in RAT- j with $0 \leq c \leq m_{i,j}$ (for new calls) and $0 \leq c \leq n_{i,j}$ (for handoff calls). Let $b_{i, assigned\ c}$ denote the bandwidth assigned to call c of class- i in RAT- j in the group of co-located cells where $b_{i, assigned\ c} \in B_i$. A call c of class- i is degraded if $b_{i, assigned\ c} < b_{i, req}$ whereas the call is upgraded if $b_{i, assigned\ c} > b_{i, req}$.

If a class of calls (i.e. class- i calls) requires a fixed number of channels (i.e. constant bit rate service), it becomes a special case in our model in which $b_{i, min} = b_{i, max}$ and the set B_i has only one element. However, it will not be possible to upgrade or degrade this class of calls. Following the general assumption in cellular networks, new and handoff class- i calls arrive in the group of co-located cells according to Poisson process with rate λ_i^n and λ_i^h respectively. The call holding time (CHT) of a class- i call is assumed to follow an exponential distribution with mean $1/\mu_i$ [74, 75].

To characterize mobility, the cell residence time (CRT), i.e., the amount of time during which a mobile terminal stays in a cell (same as the time it stays in a group of co-located cells) during a single visit, is assumed to follow an exponential distribution with mean $1/h$, where the parameter h represents the call handoff rate. It is assumed that the CRT is independent of the service class.

The channel holding time is the minimum of the CHT and the CRT. Because minimum of two exponentially distributed random variables is also exponentially distributed [78], the channel holding time for new class- i calls, and for handoff class- i call, is assumed to be exponentially distributed with means $1/\mu_i^n$ and $1/\mu_i^h$ respectively.

Note that this set of assumptions has been widely used for homogeneous cellular networks in the literature, and is found to be generally applicable in the environment where the number of mobile users is larger than the number of channels [78].

4.3 Components of the AJCAC Scheme

This section describes the proposed AJCAC scheme which consists of three components namely joint call admission controller, threshold-based bandwidth reservation unit, and bandwidth adaptation (BA) controller. These components are described in the following subsections.

4.3.1 Joint Call Admission Controller

The joint call admission controller implements the JCAC algorithm. The basic function of the JCAC algorithm is to make call admission decision and uniformly distribute traffic load among all the available RATs in the network. During call setup, a multi-mode mobile terminal requesting a service sends a request to the joint call admission controller, which implements the JCAC algorithm. The service request contains the call type, service class, and bandwidth requirements. The JCAC procedure is shown in Fig. 4-2.

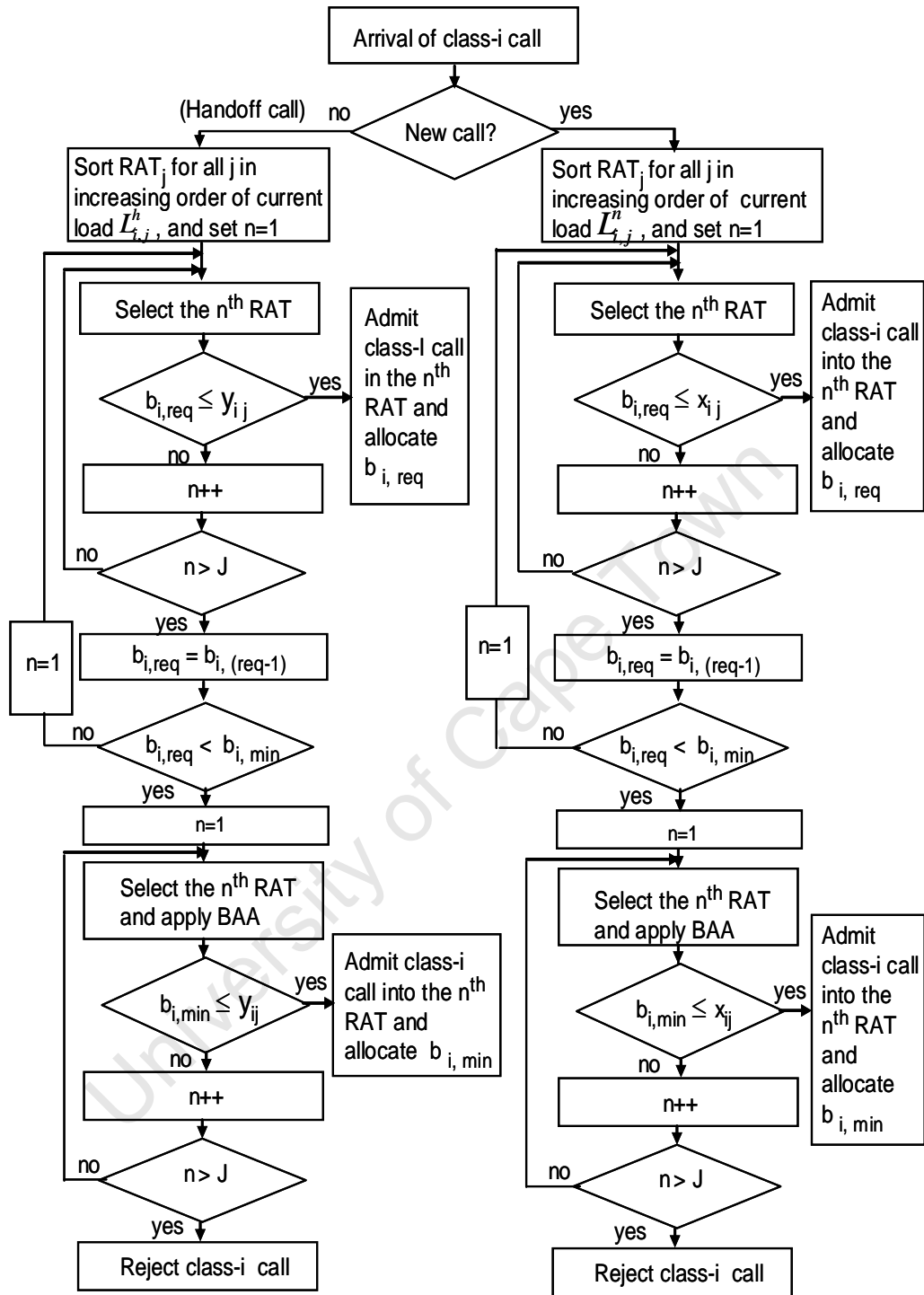


Figure 4-2. Proposed JCAC algorithm.

Whenever a call arrives, the JCAC algorithm attempts to allocate the maximum bbu for this call (i.e. set $b_{i, req} = b_{i, max}$). Thus, if the available bbu in the selected RAT is larger than or equal to $b_{i, req}$, the call will be assigned a bandwidth between $b_{i, req}$ and $b_{i, max}$. If the available bbu is less than $b_{i, req}$ but greater than or equal to $b_{i, 1}$ ($b_{i, min}$), the call will be assigned a bandwidth between $b_{i, 1}$ and $b_{i, req}$. If the available bbu in all the RATs is less than $b_{i, 1}$, BA algorithm (BAA) will be invoked to reduce the bandwidth of some ongoing call(s) in the chosen RAT. If the available bbu is still less than $b_{i, 1}$, the call will be rejected.

For new class- i calls, let $C_{i,j}^n$ denote the total bbu available in RAT- j , $\alpha_{i,j}$ the fraction of bbu available in RAT- j over the summation of bbu available in all RATs, $x_{i,j}$ the residual bbu available in RAT- j , and $L_{i,j}^n$ the current load in RAT- j . For handoff class- i calls, the corresponding values are $C_{i,j}^h, \beta_{i,j}, y_{i,j},$ and $L_{i,j}^h$. Then

$$\alpha_{i,j} = \frac{C_{i,j}^n}{\sum_{j=1}^J C_{i,j}^n} \quad \forall i, j \quad (4.1)$$

$$\sum_{j=1}^J \alpha_{i,j} = 1 \quad \forall i \quad (4.2)$$

Similarly,

$$\beta_{i,j} = \frac{C_{i,j}^h}{\sum_{j=1}^J C_{i,j}^h} \quad \forall i, j \quad (4.3)$$

$$\sum_{j=1}^J \beta_{i,j} = 1 \quad \forall i \quad (4.4)$$

When a new or handoff call arrives into a group of co-located cells, the JCAC algorithm selects the least loaded RAT available for the incoming call. The action of selecting a RAT for each arriving new or handoff call in the group of co-located cells leads to splitting of the arrival process. Fig. 4-3 illustrates the splitting of the arrival among J number of RATs in the group of co-located cells.

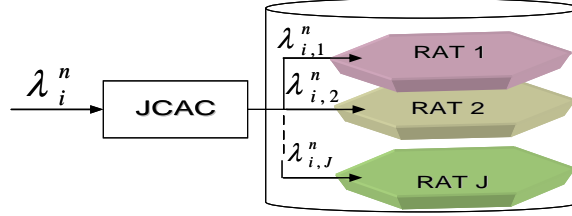


Figure 4-3. Splitting of the arrival process in the group of co-located cells.

As shown in Fig. 4-3, the arrival rate in the group of co-located cells is split among all the available RATs. Each RAT has a fraction of the arrival rate (λ_i^n). Due to uniform-load-distribution action of the JCAC algorithm, the mean arrival rates of class- i calls into each RAT in the group of collocated cells are as follows:

$$\lambda_{i,j}^n = \alpha_{i,j} \lambda_i^n \quad \forall i, j \quad (4.5)$$

$$\lambda_i^n = \sum_{j=1}^J \lambda_{i,j}^n \quad \forall i \quad (4.6)$$

Similarly

$$\lambda_{i,j}^h = \beta_{i,j} \lambda_i^h \quad \forall i, j \quad (4.7)$$

$$\lambda_i^h = \sum_{j=1}^J \lambda_{i,j}^h \quad \forall i \quad (4.8)$$

where λ_i^n and λ_i^h denote the arrival rates of new class- i calls and handoff class- i calls respectively, into the group of co-located cells. $\lambda_{i,j}^n$ and $\lambda_{i,j}^h$ denote the arrival rates of new class- i calls and handoff class- i calls, respectively, into RAT- j in the group of co-located cells

The arrival rates of a split Poisson process are also Poisson [79]. Therefore, given that the mean arrival rate of class- i calls into the group of co-located cells is Poisson, the mean arrival rates of the split class- i calls into RAT-1, RAT-2, ..., RAT- J are also Poisson.

4.3.2 Threshold-Based Bandwidth Reservation Unit

In order to maintain lower handoff dropping probability than new call blocking probability, the bandwidth reservation unit implements a bandwidth reservation policy that uses different thresholds for new and handoff calls. Fig. 4-4 shows the bandwidth reservation policy for a two-class and two-RAT system.

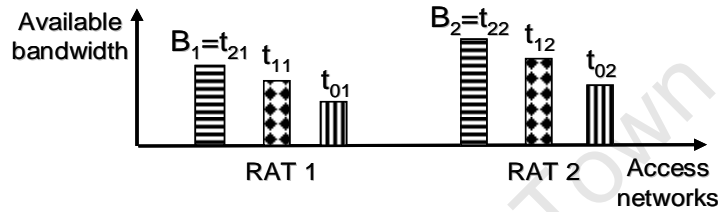


Figure 4-4. Bandwidth reservation policy for the proposed scheme.

The policy reserves bandwidth for aggregate handoff calls, thus giving them priority over new calls. The policy also prioritizes among different classes of handoff calls according to their QoS constraints by assigning a series of bandwidth thresholds $t_{1,j}$, $t_{2,j}$, ..., $t_{k,j}$, for handoff calls such that:

$$t_{0,j} \leq t_{1,j} \leq \dots \leq t_{i,j} \leq t_{(i+1),j} \dots \leq t_{k,j} = B_j \quad \forall j \quad (4.9)$$

where $t_{0,j}$ denotes the total number of bbu available for all new calls in RAT- j , and $t_{i,j}$ denotes the total number of bbu available for handoff class- i calls in RAT- j . B_j denotes the total number of bbu available in RAT- j .

Other bandwidth reservation policy such as cut-off-priority bandwidth allocation policy could be used in the proposed JCAC schemes. However, threshold-based bandwidth allocation scheme is used in the proposed scheme because it is fairer than cut-off-priority bandwidth allocation policy.

4.3.3 Bandwidth Adaptation Controller

The bandwidth adaptation controller executes the BAA which is triggered when a new call arrives or when a call is completed. Most multimedia applications are adaptive. For example voice can be encoded at 16 kbps, 32 kbps, 64 kbps, and 128 kbps by choosing appropriate encoding mechanisms. Similarly, video applications can be made rate adaptive by using for instance, a layered coding method. In layered coding method, the lowest layer (i.e., the base layer) contains the critical information for decoding the image sequence at its minimum visual quality. Additional layers provide increasing quality. All these encoded layers may be transmitted when the network is underutilized. However, when the network load is being fully utilized, only base layer(s) which contain critical information may be transmitted.

As an illustration, if one would watch a 30-minute video-clip encoded at 256 kbps and 64kps respectively. At 256 kbps, one would see better pictures with better resolution than at 64 kbps. Therefore, the bandwidth adaptation affects the quality of the real-time applications rather than the transmission time. However, the minimum requested QoS is maintained by ensuring that the bbu of the calls are not degraded below the required minimum.

In the proposed AJCAC scheme, when the system is underutilized, all arriving new and handoff class- i calls are admitted by the JCAC with the highest bandwidth level (i.e. $b_{i,max}$) for the calls. This approach increases bandwidth utilization for the heterogeneous wireless network. However, when the system is fully utilized, bandwidth adaptation controller is invoked to execute BAA.

The BAA is triggered whenever there is a call arrival event or a call departure event. The BAA performs two main procedures: downgrades and upgrades ongoing calls. The downgrading procedure is activated in the arrival epoch (i.e. when a new or handoff arrives to an overloaded group of co-located cells). BAA reduces the bandwidth of some ongoing call(s) randomly selected in the system to free just enough bbu to accommodate the incoming call. Note that an adaptive class- i call is never degraded below the minimum

bbu necessary to guarantee its QoS requirements. The upgrading procedure is activated in the departure epoch.

In the arrival epoch, the BAA downgrading procedure can be implemented in two ways. In the first implementation, only ongoing new calls can be downgraded to accommodate an incoming new call whereas both ongoing new and handoff calls can be downgraded to accommodate an incoming handoff call. This approach further prioritizes handoff calls over new calls, in addition to the prioritization obtained by using different rejection thresholds for new and handoff calls. In the second implementation, both new and handoff calls can be downgraded to accommodate an incoming new (or handoff) call. In this case, prioritization of handoff calls over new calls can only be achieved by using different rejection thresholds for new and handoff calls.

In the departure epoch, when a call departs from a RAT in the group of co-located cells, some of the ongoing call(s) randomly selected in RAT of the group of co-located cells may be upgraded by the BAA algorithm.

4.4 Markov Chain Model of the AJCAC Scheme

The AJCAC scheme can be modeled as a multi-dimensional Markov chain. The state space of the group of co-located cells can be represented by a $(2 \cdot K \cdot J)$ -dimensional vector given as:

$$\Omega = (m_{i,j}, n_{i,j} : i = 1, \dots, k, j = 1, \dots, J) \quad (4.10)$$

The non-negative integer $m_{i,j}$ denotes the number of ongoing new class- i calls in RAT- j , and the non-negative integer $n_{i,j}$ denotes the number of ongoing handoff class- i calls in RAT- j . Let S denote the state space of all admissible states of the group of co-located cells as it evolves over time. An admissible state s is a combination of the numbers of users in each class that can be supported simultaneously in the group of co-located cells while maintaining adequate QoS and meeting radio resource constraints. The state S of all admissible states is given as:

$$\begin{aligned}
S &= \{ \Omega = (m_{i,j}, n_{i,j} : i = 1, \dots, k, j = 1, \dots, J) : \\
&\sum_{i=1}^k \sum_{c=1}^{m_{i,j}} b_{i, assigned_c} \leq t_{0,j}^n \quad \forall j \wedge \\
&\sum_{c=1}^{n_{i,j}} b_{i, assigned_c} \leq t_{i,j}^h \quad \forall i, j \wedge \\
&\sum_{i=1}^k \sum_{c=1}^{m_{i,j}} b_{i, assigned_c} + \sum_{i=1}^k \sum_{c=1}^{n_{i,j}} b_{i, assigned_c} \leq B_j \quad \forall j \}
\end{aligned} \tag{4.11}$$

The constraints simply imply that the sum of the bandwidth units of all admitted class- i calls cannot be more than the total bandwidth units available for that class of calls. Given that the system is in the current state, s , for the AJCAC scheme, the state transition could be triggered by any of the following events:

1) Admission of a new class- i call into RAT- j with the successor state s_1^{+1} and transition rate $q(s, s_1^{+1})$. It follows that

$$q(s, s_1^{+1}) = \lambda_{i,j}^n, \quad (s, s_1^{+1} \in S) \tag{4.12}$$

2) Admission of a handoff class- i call into RAT- j with the successor state s_2^{+1} and transition rate $q(s, s_2^{+1})$. It follows that

$$q(s, s_2^{+1}) = \lambda_{i,j}^h, \quad (s, s_2^{+1} \in S) \tag{4.13}$$

3) Departure of a new class- i call from RAT- j with the successor state s_1^{-1} and transition rate $q(s, s_1^{-1})$. It follows that

$$q(s, s_1^{-1}) = m_{i,j} \mu_i^n, \quad (s, s_1^{-1} \in S) \tag{4.14}$$

4) Departure of a handoff class- i call from RAT- j with the successor state s_2^{-1} and transition rate $q(s, s_2^{-1})$. It follows that

$$q(s, s_2^{-1}) = n_{i,j} \mu_i^h, \quad (s, s_2^{-1} \in S) \tag{4.15}$$

where $s, s_1^{+1}, s_1^{-1}, s_2^{+1}$, and s_2^{-1} are the following matrices.

$$s = \begin{bmatrix} m_{11} \cdots m_{1j} \cdots m_{1J} \\ \vdots \\ m_{i1} \cdots m_{ij} \cdots m_{iJ} \\ \vdots \\ m_{k1} \cdots m_{kj} \cdots m_{kJ} \\ n_{11} \cdots n_{1j} \cdots n_{1J} \\ \vdots \\ n_{i1} \cdots n_{ij} \cdots n_{iJ} \\ \vdots \\ n_{k1} \cdots n_{kj} \cdots n_{kJ} \end{bmatrix}, \quad s_1^{\pm 1} = \begin{bmatrix} m_{11} \cdots m_{1j} \cdots m_{1J} \\ \vdots \\ m_{i1} \cdots m_{ij} \pm 1 \cdots m_{iJ} \\ \vdots \\ m_{k1} \cdots m_{kj} \cdots m_{kJ} \\ n_{11} \cdots n_{1j} \cdots n_{1J} \\ \vdots \\ n_{i1} \cdots n_{ij} \cdots n_{iJ} \\ \vdots \\ n_{k1} \cdots n_{kj} \cdots n_{kJ} \end{bmatrix}, \quad s_2^{\pm 1} = \begin{bmatrix} m_{11} \cdots m_{1j} \cdots m_{1J} \\ \vdots \\ m_{i1} \cdots m_{ij} \cdots m_{iJ} \\ \vdots \\ m_{k1} \cdots m_{kj} \cdots m_{kJ} \\ n_{11} \cdots n_{1j} \cdots n_{1J} \\ \vdots \\ n_{i1} \cdots n_{ij} \pm 1 \cdots n_{iJ} \\ \vdots \\ n_{k1} \cdots n_{kj} \cdots n_{kJ} \end{bmatrix} \quad (4.16)$$

The decision epochs are the arrival or departure of a new or handoff call. Joint call admission decisions are taken in the arrival epoch. Every time a new or handoff class- i call arrives in the group of co-located cells, the JCAC algorithm decides whether or not to admit the call, and in which RAT to admit it. Note that call admission decision is made only at the arrival of a call, and no call admission decision is made in the group of co-located cells when a call departs. When the system is in state s , an accept/reject decision must be made for each type of possible arrival, i.e., an arrival of a new class- i call, or the arrival of a handoff class- i call in the group of co-located cells. The following are the possible JCAC decisions in the arrival epoch.

1) Reject the class- i call (new or handoff) in the group of collocated cells, in which case the state s does not evolve.

2) Admit the class- i call into RAT- j without adapting the bandwidth of ongoing call(s) in the RAT, in which case the state s evolves.

3) Admit the class- i call into RAT- j after adapting the bandwidth of ongoing call(s) in the RAT, in which case state s evolves.

Thus, the call admission action space A can be expressed as follows:

$$A = \{ a = (a_1^n, \dots, a_k^n, a_1^h, \dots, a_k^h) : \quad (4.17) \\
a_i^n, a_i^h \in (0, \pm 1, \dots, \pm j, \pm (j + 1), \dots, \pm J), i = 1, \dots, k \}$$

where a_n^i denotes the action taken on arrival of a new class- i call within the group of co-located cells, and a_i^h denotes the action taken on arrival of a handoff class- i call from an adjacent group of co-located cells. a_i^n (or a_i^h) = 0 means reject the new class- i (or handoff class- i) call. a_i^n (or a_i^h) = +1 means accept the new class- i (or handoff class- i) call into RAT-1 without adapting the bandwidth of existing call(s). a_i^n (or a_i^h) = -1 means accept the new class- i (or handoff class- i) call into RAT-1 after adapting (degrading) the bandwidth of existing call(s). a_i^n (or a_i^h) = + j means accept the new class- i (or handoff class- i) call into RAT- j without adapting the bandwidth of existing call. a_i^n (or a_i^h) = - j means accept the new class- i (or handoff class- i) call into RAT- j after adapting (degrading) the bandwidth of existing call(s).

In the departure epoch, the bandwidth adaptation unit makes the decision to adapt (upgrade) or not to adapt the bandwidth of ongoing call(s). Thus, the call departure action space W can be expressed as follows:

$$W = \{w = (0, 1)\}$$

where $w = 0$ means do not adapt the bandwidth of the ongoing call(s) and $w = 1$ means adapt the bandwidth of ongoing call(s).

Based on its Markovian property, the proposed JCAC scheme can be model as a $(2*K*J)$ -dimensional Markov chain. Let $\rho_{new_{i,j}}$ and $\rho_{han_{i,j}}$ denote the load generated by new class- i calls and handoff class- i calls, respectively, in RAT- j . Then,

$$\rho_{new_{i,j}} = \frac{\lambda_{i,j}^n}{\mu_i^n} \quad \forall i, j \quad (4.18)$$

$$\rho_{han_{i,j}} = \frac{\lambda_{i,j}^h}{\mu_i^h} \quad \forall i, j \quad (4.19)$$

From the steady-state solution of the Markov model, performance measures of interest can be determined by summing up appropriate state probabilities. Let $P(s)$ denote the steady-state probability that the system is in state s ($s \in S$). From the detailed balance equation, $P(s)$ is obtained as:

$$P(s) = \frac{1}{G} \prod_{i=1}^k \prod_{j=1}^J \frac{(\rho_{new\ i,j})^{m_{i,j}} (\rho_{han\ i,j})^{n_{i,j}}}{m_{i,j}! n_{i,j}!} \quad \forall s \in S \quad (4.20)$$

where G is a normalization constant given by:

$$G = \sum_{s \in S} \prod_{i=1}^k \prod_{j=1}^J \frac{(\rho_{new\ i,j})^{m_{i,j}} (\rho_{han\ i,j})^{n_{i,j}}}{m_{i,j}! n_{i,j}!} \quad (4.21)$$

The performance metrics used for the proposed AJCAC scheme are new call blocking probability (NCBP), handoff call dropping probability (HCDP), and overall system utilization. These metrics are derived as follows.

4.4.1 New Call Blocking Probability

A new class- i call is blocked in the group of co-located cells if none of the available RATs has enough bbu to accommodate the new call with the minimum bandwidth requirement after degrading the ongoing new calls. Let $S_{b_i} \subset S$ denote the set of states in which a new class- i call is blocked in the group of co-located cells. It follows that:

$$S_{b_i} = \{s \in S : (b_{i, \min} + \sum_{x=1}^k m_{x,j} b_{x, \min} > t_{0,j}^n \vee b_{i, \min} + \sum_{x=1}^k m_{x,j} b_{x, \min} + \sum_{x=1}^k \sum_{c=1}^{n_{x,j}} b_{x, assigned\ c} > B_j) \forall j\} \quad (4.22)$$

Thus the new call blocking probability, P_{b_i} , for a class- i call in the group of co-located cells is given by:

$$P_{b_i} = \sum_{s \in S_{b_i}} P(s) \quad (4.23)$$

4.4.2 Handoff Call Dropping Probability

A handoff class- i call is dropped in the group of co-located cells if none of the available RATs has enough bbu to accommodate the handoff call with the minimum bandwidth requirement after degrading the ongoing new calls and handoff calls. Let $S_{di} \subset S$ denote the set of states in which a handoff class- i call is dropped in the group of co-located cells. It follows that:

$$S_{di} = \{s \in S : ((1 + n_{i,j})b_{i, \min} > t_{i,j}^h \quad \vee \quad b_{i, \min} + \sum_{x=1}^k (m_{x,j} + n_{x,j}) b_{x, \min} > B_j) \forall j\} \quad (4.24)$$

Thus the handoff call dropping probability for a class- i call, P_{d_i} , in the group of co-located cells is given by:

$$P_{d_i} = \sum_{s \in S_{d_i}} P(s) \quad (4.25)$$

4.4.3 Overall System Utilization

The average utilization of the heterogeneous wireless network can be obtained by summing up for all the admissible state s ($s \in S$), the product of the system utilization in a particular state s ($s \in S$) and the probability $P(s)$ of the system being in that state.

The average utilization of the heterogeneous wireless network by class- i calls ($U_{class-i}$) can be derived as follows:

$$U_{class-i} = \sum_{s \in S} P(s) \sum_{j=1}^J \left(\sum_{c=1}^{m_{i,j}} b_{i,assigned_c} + \sum_{c=1}^{n_{i,j}} b_{i,assigned_c} \right) \quad (4.26)$$

The average utilization of RAT- j in the heterogeneous wireless network by all calls (U_{RAT-j}) can be derived as follows:

$$U_{RAT-j} = \sum_{s \in S} P(s) \left(\sum_{i=1}^k \left(\sum_{c=1}^{m_{i,j}} b_{i,assigned_c} + \sum_{c=1}^{n_{i,j}} b_{i,assigned_c} \right) \right) \quad (4.27)$$

The average utilization of the entire heterogeneous wireless network by all calls (U) can be derived as follows:

$$U = \sum_{s \in S} P(s) \sum_{j=1}^J \sum_{i=1}^k \left(\sum_{c=1}^{m_{i,j}} b_{i,assigned_c} + \sum_{c=1}^{n_{i,j}} b_{i,assigned_c} \right) \quad (4.28)$$

4.5 Performance Evaluation

In this section, the performance of the proposed AJCAC scheme is evaluated with respect to new call blocking probability, handoff call dropping probability, and radio resource utilization. For comparison, a JCAC algorithm that does not incorporate adaptive bandwidth management is also modeled for the same heterogeneous wireless network. New call blocking probability, handoff call dropping probability, and radio resource utilization are also derived for the non-adaptive JCAC (NAJCAC) scheme. Simulation is conducted using MATLAB. The results of the proposed AJCAC scheme are compared with that of the NAJCAC. The system parameters used are shown in Table 4-1.

Table 4-1. Simulation parameters.

Parameters	Class-1 call	Class-2 call
Class- i call bbu set	{2, 3, 4}	{3, 5, 7}
Requested bbu ($b_{i, req}$)	3	5
λ_i^n	[1, 8]	[1, 8]
μ_i	0.5	0.5
Other Parameters		
B_1	B_2	$t_{0,1}$ $t_{0,2}$ $t_{1,1}$ $t_{1,2}$ $t_{2,1}$ $t_{2,2}$ h
30	60	15 30 30 60 30 60 0.5

The arrival rate of handoff class- i calls in the group of co-located cells is assumed to be proportional to the arrival rate of new class- i calls by $\lambda_i^h = (h/\mu)\lambda_i^n$ where h is the handoff rate.

4.5.1 Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability

Fig. 4-5 and Fig. 4-6 show the performance of the proposed AJCAC scheme compared with that of NAJCAC. As shown in Fig. 4-5, the NCBP of each class of calls increases with the call arrival rate. The NCBP, P_{b1} is always less than the NCBP, P_{b2} because class-2 calls require more bbu than class-1 calls. Thus a class-2 call may be blocked due to insufficient bbu to accommodate it whereas a class-1 call may still be accepted into the network. However, for both classes of calls, the NCBP for the AJCAC scheme is always less than the corresponding NCBP for the NAJCAC scheme. Note that lower NCBP of the AJCAC scheme implies that its connection-level QoS is better than that of the NAJCAC scheme. The reason why the NCBP of the AJCAC scheme is less than the NAJCAC scheme is as follows. When the total bbu allocated to new calls is being fully utilized, incoming new calls are rejected by the NAJCAC scheme whereas the AJCAC scheme adapts (degrades) the bandwidth of some of the ongoing adaptive calls to free just enough bbu to accommodate the incoming new calls. Consequently, the NCBP of the AJCAC scheme is less than that of the NAJCAC scheme. However, an adaptive class- i call is never degraded below the minimum bbu necessary to guarantee its minimum QoS requirements.

Fig. 4-6 shows a similar trend for the HCDP for each class of calls, which increases with the call arrival rate. The HCDP, P_{d1} is always less than the HCDP, P_{d2} , because class-2 calls require more bbu than class-1 calls. However, for both classes of calls, the HCDP for the AJCAC scheme is always less than the corresponding HCDP for the NAJCAC scheme. The reason why the HCDP of the AJCAC scheme is less than the NAJCAC scheme is as follows. When the System is being fully utilized, incoming handoff calls are rejected by the NAJCAC scheme whereas the AJCAC scheme adapts (degrades) the bandwidth of

some of the ongoing adaptive calls to free just enough bbu to accommodate the incoming handoff calls. Consequently, the HCDP of the AJCAC scheme is less than that of the NAJCAC scheme.

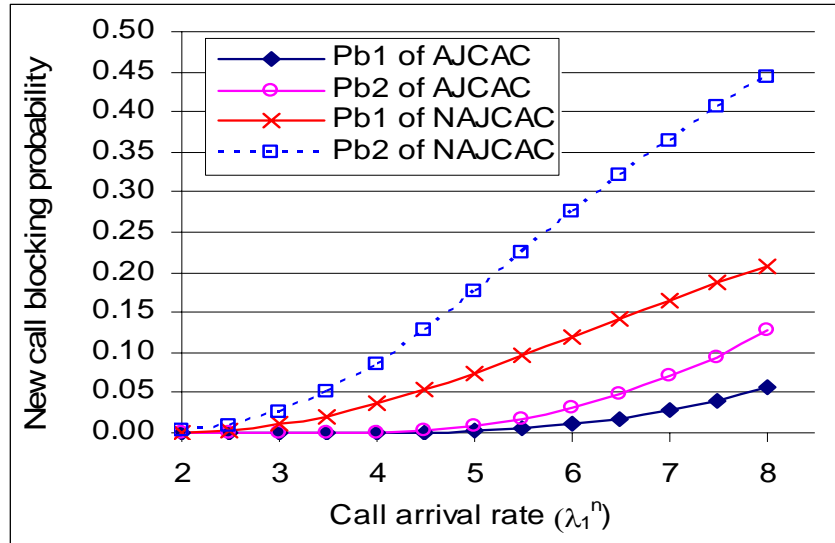


Figure 4-5. Effect of varying the call arrival rate on the new call blocking probability.

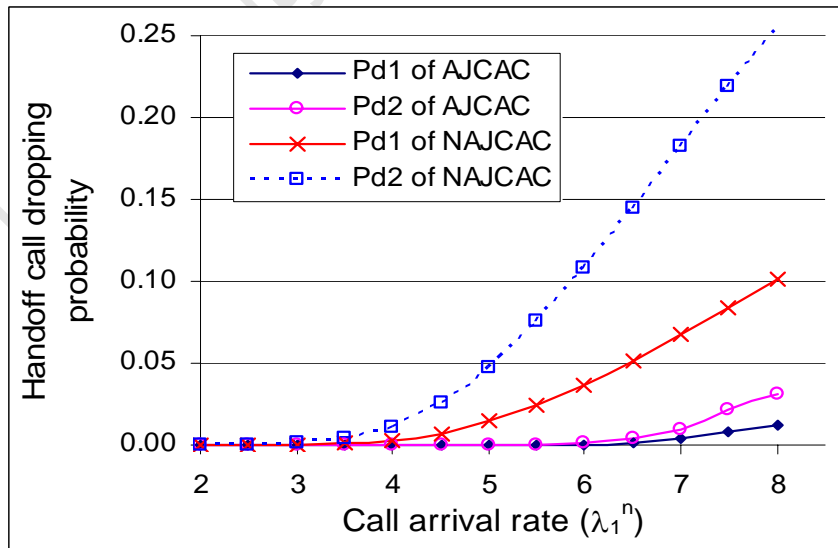


Figure 4-6. Effect of varying the call arrival rate on the handoff call dropping probability.

Fig. 4-7 and Fig. 4-8 compare NCBP and HCDP of the AJCAC scheme for class-1 and class-2 call, respectively. One of the objectives of the AJCAC scheme is to prioritize handoff calls over new calls. Fig. 4-7 shows that the HCDP, Pd1 of the AJCAC scheme is always less than the Pb1. Similarly, it can be seen in Fig. 4-8 that the HCDP, Pd2 is always less than the NCBP, Pb2. This shows that handoff calls are prioritized over new calls. This prioritization of the handoff calls over new calls is achieved by making the handoff call rejection thresholds higher than the new call rejection thresholds.

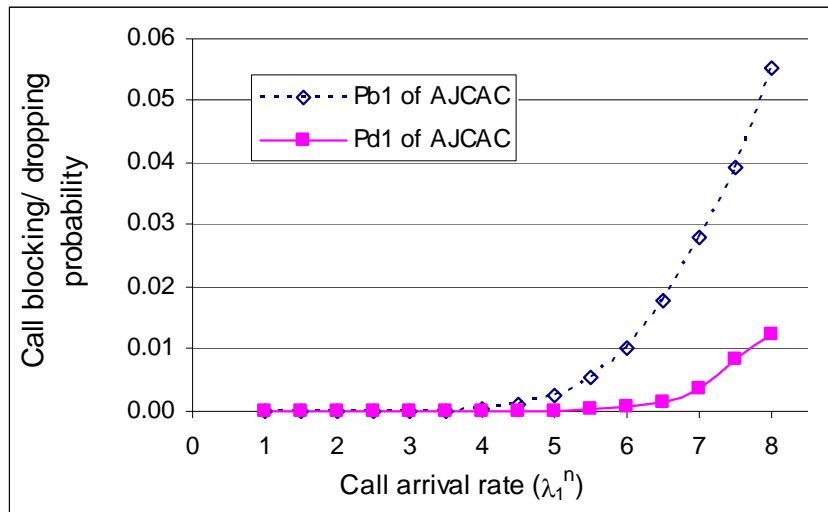


Figure 4-7. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls.

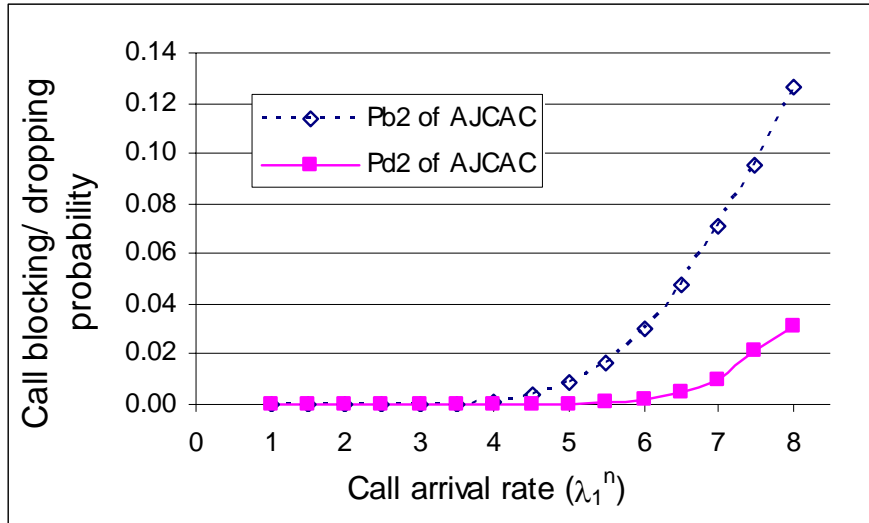


Figure 4-8. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls.

4.5.2 Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability

Fig. 4-9 and 4-10 show the effect of varying the new call rejection threshold, T_0 on the NCBP and HCDP of the AJCAC and NAJCAC schemes for class-1 calls and class-2 calls respectively. The additional system parameters used are as follows: $T_{01} = T_0$, $T_{02} = 2T_0$, $T_0 = [0, 30]$, $\lambda_1^n = \lambda_2^n = 8$. As shown in Fig. 4-8, at low threshold values, the NCPB, Pb1 for the two JCAC schemes is high whereas the HCDP, Pd1 is low. As the threshold value, T_0 increases, Pb1 decreases because new calls are given more access to the available bandwidth. On the other hand, the handoff dropping probability, Pd1 increases as a result of the higher degree of sharing between the new and the handoff calls. However, Pb1 and Pd1 of the AJCAC scheme are always less than the corresponding Pb1 and Pd1 the NAJCAC scheme.

Fig. 4-10 shows a similar trend for class-2 calls. At low threshold values, the NCPB, Pb2 for the two JCAC schemes is high whereas the HCDP, Pd2 is low. As the threshold value,

T_0 increases, Pb_2 decreases whereas handoff dropping probability, Pd_2 increases. However, Pb_2 and Pd_2 of the AJCAC scheme are always less than the corresponding Pb_2 and Pd_2 of the NAJCAC scheme.

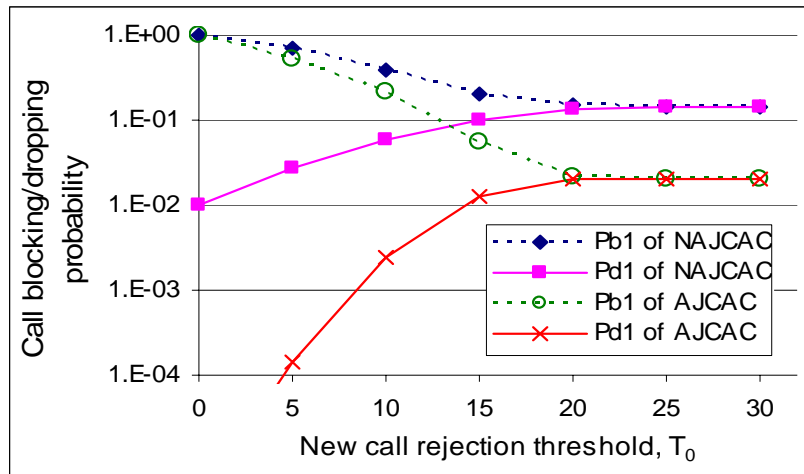


Figure 4-9. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls.

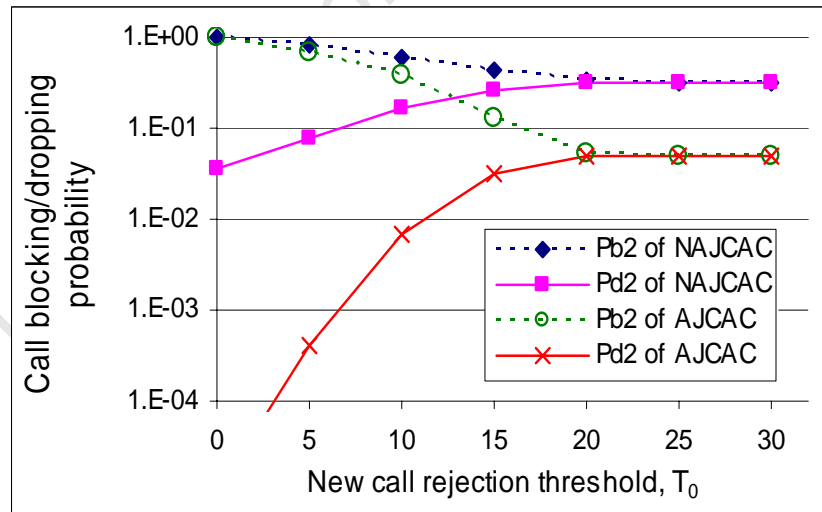


Figure 4-10. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-2 calls.

4.5.3 Impact of varying the call arrival rate on the normalized average system utilization.

Fig. 4-11 shows the normalized average system utilization of the heterogeneous wireless network for class-1 and class-2 calls. When the new call arrival rate is low, the normalized average system utilization of the AJCAC scheme by class-1 calls and class-2 calls respectively is higher than the corresponding normalized average system utilization of the NAJCAC scheme. The reason for improvement in system utilization of the AJCAC scheme over NAJCAC scheme is as follows. When the system load is low, the AJCAC scheme allocates maximum bbu to all admitted calls, thereby improving the overall system utilization whereas the NAJCAC scheme allocates just the requested bbu to all admitted calls in the same class regardless of whether the traffic load is low or high. However, when the system is operating at full capacity, the AJCAC scheme degrades the bbu of some ongoing calls and free just enough bbu to accommodate incoming new calls. Fig. 4-11 also shows that class-2 calls benefit more from the adaptive bandwidth management. The reason for this is as follows: For the NAJCAC scheme, class-2 calls are rejected as the arrival rate increases whereas class-1 calls are still admitted because they required less bbu than class-2 calls. However, for the AJCAC scheme, class-2 calls are still being admitted as well as class-1 calls as the arrival rate increases because the BAA is invoked to reduce the bandwidth of some existing calls in order to accommodate incoming class-2 calls. Therefore, class-2 calls have a better share of the radio resources. Fig. 4-11 also shows that for both AJCAC and NAJCAC schemes normalized average utilization by class-2 calls is greater than the corresponding normalized average utilization by class-1 calls because class-2 calls require more bbu than class-1 calls.

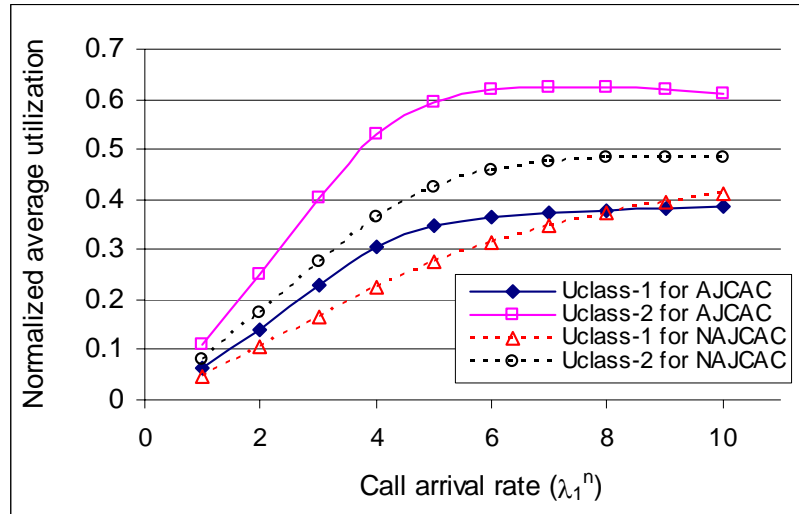


Figure 4-11. Impact of varying the call arrival rate on the normalized average system utilization for each class of calls.

Fig. 4-12 shows the normalized average system utilization of the entire heterogeneous wireless network by both classes of calls combined together. The normalized average system utilization of the AJCAC scheme is higher than the normalized average system utilization for the NAJCAC scheme. The reason for improvement in system utilization of the AJCAC scheme over NAJCAC scheme is as follows. When the system load is low, the AJCAC scheme allocates maximum bbu to all admitted calls, thereby improves the overall system utilization whereas the NAJCAC scheme allocates just the requested bbu to all admitted calls in the same class regardless of whether the traffic load is low or high. However, when the system is operating at the full capacity, the AJCAC scheme degrades the bbu of some ongoing calls and free just enough bbu to accommodate incoming new calls. Fig. 4-12 shows that the AJCAC scheme improves the system utilization by up to 20% of the NAJCAC scheme.

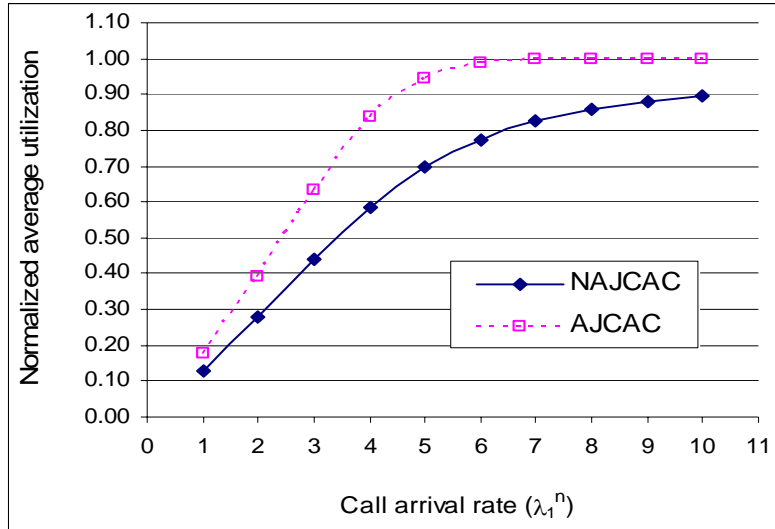


Figure 4-12. Impact of varying the call arrival rate on the normalized average system utilization for both class-1 and class-2 calls combined.

Fig. 4-13 shows the ratio of the average system utilization of each RAT to the average utilization of the entire heterogeneous wireless network by the two classes of calls combined. In other words, Fig 4-13 shows the proportion of the total resources in the heterogeneous wireless network that is utilized from each RAT. It can be seen that as the arrival rate increases the proportion of resources that is being utilized in RAT-2 is almost twice that of RAT-1. This is expected because the total capacity of RAT-2 is twice the total capacity of RAT-1. Note that one of the objectives of the schemes (both AJCAC and NAJCAC) is to uniformly distribute traffic load among the available RATs. Fig. 4-13 also shows that for each of the RATs, the normalized average system utilization of the AJCAC scheme is higher than the corresponding normalized average system utilization for the NAJCAC scheme. The reason for improvement in system utilization of the AJCAC scheme over NAJCAC scheme is as follows. When the system load is low, the AJCAC scheme allocates maximum bbu to all admitted calls, thereby improving the overall system utilization whereas the NAJCAC allocates just the requested bbu to all admitted calls in the same class regardless of whether the traffic load is low or high. However, when the system is operating at full capacity, the AJCAC scheme degrades the bbu of some ongoing calls and free just enough bbu to accommodate incoming new calls.

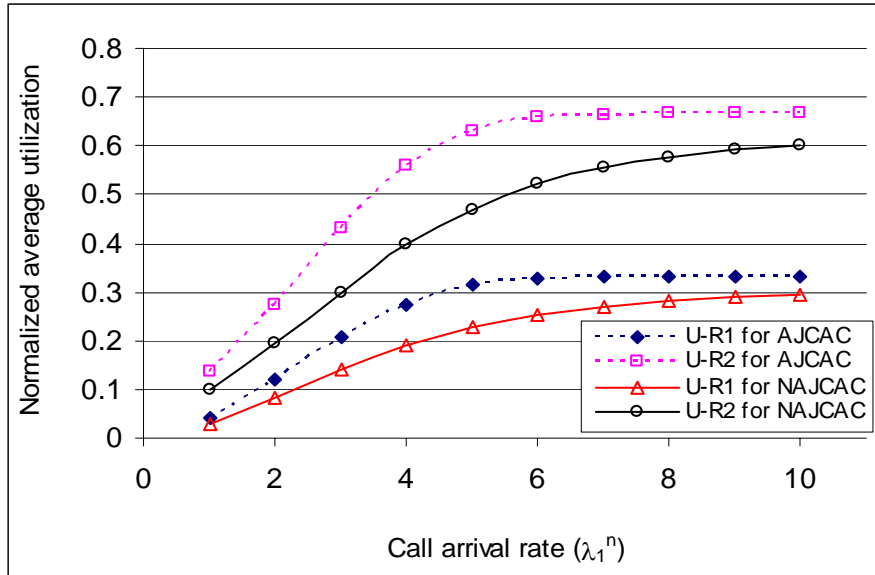


Figure 4-13. Effect of varying the call arrival rate on utilization of each RAT by both classes of calls combined.

4.6 Summary

This chapter proposes an adaptive bandwidth management and JCAC scheme to enhance overall system utilization and connection-level QoS in heterogeneous wireless networks. The adaptive JCAC scheme improves overall system utilization by adapting the bandwidth of calls based on current traffic condition and by uniformly distribute traffic load among the available RATs. The AJCAC scheme guarantees the QoS requirements of all accepted call and reduces both new call blocking probability and handoff call dropping probability in the heterogeneous wireless networks. It prioritizes handoff calls over new calls by using different call rejection thresholds for new and handoff calls. The Markov chain model makes it possible to derive new call blocking probability, handoff call dropping probability, and average system utilization for the adaptive JCAC scheme. Performance of the adaptive JCAC scheme is compared with the performance of a non adaptive JCAC scheme in the same heterogeneous wireless network. Results show that new call blocking probability and handoff call dropping probability are significantly reduced by using the adaptive load-based JCAC scheme. Moreover, the AJCAC scheme improves the system utilization by up to 20% of the NAJCAC scheme.

5 Reduction of Call Blocking Probability through Optimal Allocation of Calls

5.1 Introduction

This chapter proposes an optimal JCAC (OJCAC) scheme to reduce call blocking probability in heterogeneous wireless networks. The objectives of the OJCAC scheme are to guarantee the QoS requirements of all admitted calls, prioritize handoff calls over new calls, and reduce call blocking probability. The contributions in this chapter are twofold. The first contribution is the development of an optimal RAT selection policy to reduce call blocking probability in NGWN. The second contribution is the development of an analytical model for the OJCAC scheme, and the derivation new call blocking and handoff call dropping probability.

5.2 System Model

This study considers a heterogeneous wireless network, which comprises a set of RATs, H with co-located cells in which radio resources are jointly managed. H is given as:

$$H = \{RAT - 1, RAT - 2, \dots, RAT - J\} \quad (5.1)$$

where J is the total number of RATs in the heterogeneous wireless network. The heterogeneous wireless network supports k -classes of calls, and each RAT in set H can support certain classes of calls. Let $H_i (H_i \subseteq H)$ denote the set of RATs which can support class- i calls in the heterogeneous wireless network, and let $h_i (h_i \subseteq h)$ denote the set of indices of all RAT- j which belong to H_i , where $h = \{1, 2, \dots, J\}$. Furthermore, let $D_j (D_j \subseteq D)$ denote the set of call classes that can be supported by RAT- j ($j=1, 2, \dots, J$) where $D = \{\text{class-1}, \dots, \text{class-}k\}$. Let $d_j (d_j \subseteq d)$ denote the set of indices of all class- i calls which belong to D_j , where $d = \{1, \dots, k\}$.

The heterogeneous system considered in this chapter is different from the heterogeneous system considered in Chapter 4. In Chapter 4, it is assumed that all RATs in the

heterogeneous wireless network support all classes of calls. Therefore, a class- i call is blocked in the heterogeneous wireless network only if none of the available RATs has enough bbu to accommodate the class- i calls. In this chapter, all RATs do not necessary support all classes of calls. Thus, a specific class of calls can be blocked because all the RATs that support the class of calls are fully loaded whereas other RATs that do not support the class of calls may be underutilized. In this chapter, based on the assumption that all RATs in the heterogeneous wireless network do not necessarily support all types of calls, there is a problem of optimal allocation of calls among RAT. Fig. 5-1 shows a two-class three-RAT heterogeneous wireless network.

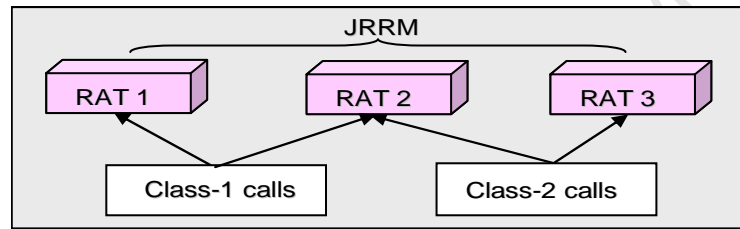


Figure 5-1. A two-class three-RAT heterogeneous wireless network.

In the example shown in Fig. 5-1, RAT-1 can support only class-1 calls, RAT-2 can support both class-1 and class-2 calls, and RAT-3 can support only class-2 calls. As shown in the Fig. 5-1, $H = \{\text{RAT-1, RAT-2, RAT-3}\}$, $H_1 = \{\text{RAT-1, RAT-2}\}$, $H_2 = \{\text{RAT-2, RAT-3}\}$, $h = \{1, 2, 3\}$, $h_1 = \{1, 2\}$, $h_2 = \{2, 3\}$. $D = \{\text{class-1, class-2}\}$, $D_1 = \{\text{class-1}\}$, $D_2 = \{\text{class-1, class-2}\}$, and $D_3 = \{\text{class-2}\}$.

The definition of classes of calls adopted in this chapter is different from that adopted in Chapter 4. In Chapter 4, one of the major objectives is to investigate the effect of bandwidth adaptation on connection level QoS and radio resource utilization. Therefore, calls belonging to the same class can be allocated different amount of basic bandwidth units (bbu) depending on the current load in the heterogeneous wireless network. However, every class of calls has a minimum and a maximum number of bbu that can be allocated to any call that fall within the class.

In this chapter, the objective of the proposed scheme is to minimize call blocking probability by optimal allocation of calls among available RATs. Therefore, the number of bbu that can be allocated to a particular class of calls is kept constant so that the reduction in call blocking probability as a result of optimal allocation of calls can be clearly observed. If the bbu allocated to calls belonging to the same class are not fixed, it will be difficult to conclude whether the reduction in call blocking probability in the proposed scheme is as a result of optimal allocation of calls or as a result of bandwidth adaptation or as a result of both.

Each cell of RAT- j ($j = 1, \dots, J$) has a total of C_j basic bandwidth units (bbu). The physical meaning of a unit of radio resources (such as time slots, code sequence, etc.) is dependent on the specific technological implementation of the radio interface [67]. However, no matter which multiple access technology (FDMA, TDMA, CDMA, or OFDMA) is used, system capacity can be interpreted in terms of effective or equivalent bandwidth [68-70]. Therefore, the bandwidth of a call represents the number of bbu that is adequate for guaranteeing the desired QoS for the call, which is similar to the approach used for homogeneous networks in [70-72].

The approach used in this research is based on decomposing heterogeneous wireless networks into groups of co-located cells. Following the general assumption which is made in homogeneous wireless networks, it is assumed that the types and amount of traffic are statistically the same in all cells of each RAT [70, 71, 74, 75, 76, 77]. Therefore, the types and amount of traffic are statistically the same in all groups of co-located cells.

The correlation between the groups of co-located cells results from handoff connections between the cells of corresponding groups. Under this formulation, each group of co-located cells can be modeled and analyzed individually. Therefore, the research focuses on a single group of co-located cells. New and handoff class- i calls arrive in the group of co-located cells according to Poisson process with rate λ_i^n and λ_i^h respectively. The call holding time (CHT) of a class- i call follows an exponential distribution with mean $1/\mu_i$ [74, 75].

To characterize mobility, the cell residence time (CRT), i.e., the amount of time during which a mobile terminal stays in a cell (same as the time it stays in a group of co-located cells) during a single visit, is assumed to follow an exponential distribution with mean $1/h$, where the parameter h represents the call handoff rate. The channel holding time is the minimum of the CHT and the CRT. The channel holding time for new and handoff class- i calls is assumed to be exponentially distributed with means $1/\mu_i^n$ and $1/\mu_i^h$ respectively. These assumptions have been widely used for homogeneous wireless networks in the literature, and are found to be generally applicable in the network where the number of users is larger than that of channels [78].

5.3 Proposed Optimal JCAC Scheme

The OJCAC scheme is illustrated in Fig. 5-2 using a two-class three-RAT heterogeneous wireless network. As shown in Fig. 5-2, the OJCAC scheme comprises four components namely the joint call admission controller, the arrival rate measurement unit, the optimal policy determination unit, and the bandwidth reservation unit. These components are described in the following subsections.

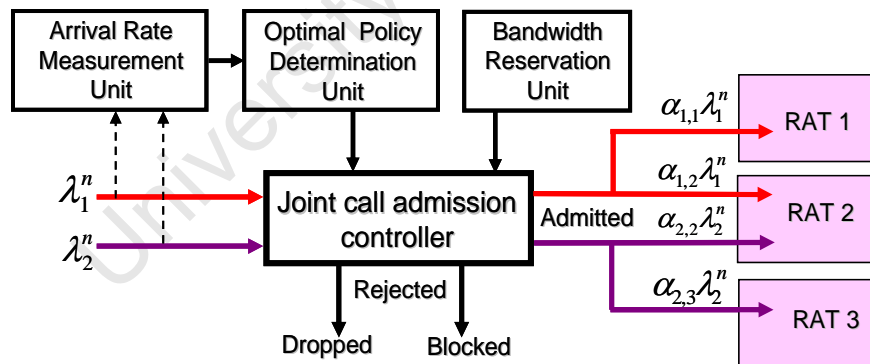


Figure 5-2. Proposed optimal JCAC scheme.

5.3.1 Joint Call Admission Controller

The joint call admission controller implements the JCAC algorithm. The basic function of the JCAC algorithm is to make call admission decisions. During a call setup, a multi-mode mobile terminal requesting a service sends a request to the joint call admission controller, which implements the JCAC algorithm. The service request contains the call type, service class, and bandwidth requirements. The JCAC algorithm makes call admission decision based on the optimal call allocation policy determine optimal call allocation unit.

5.3.2 Arrival Rate Measurement Unit

The arrival rate measurement unit measures the arrival rates of different call classes as shown in Fig. 5-2. This measurement is done periodically. The measured value of the call arrival rates for each class of call is periodically sent to the optimal call allocation unit, which determines the optimal call allocation policy.

5.3.3 Bandwidth Reservation Unit

In order to maintain a lower handoff call dropping probability than new call blocking probability, certain amount of bandwidth is exclusively reserved for handoff calls in all cells of each group of co-located cells. Fig. 5-3 shows the bandwidth reservation policy. C_j and T_{0j} are the threshold for rejecting new and handoff calls respectively in RAT- j .

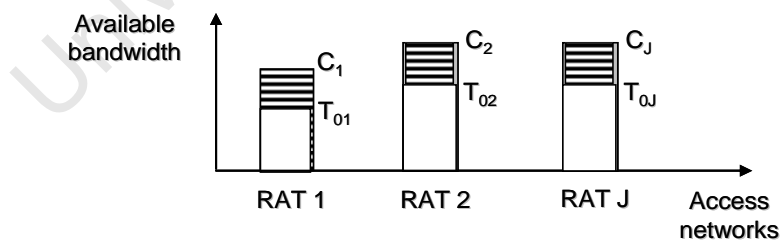


Figure 5-3. Bandwidth reservation policy.

5.3.4 Optimal Policy Determination Unit

Based on the measured mean call arrival rates, the optimal policy determination unit determines the call allocation policy (i.e. values of α_{ij} and β_{ij}) that will minimize call

blocking probability in the heterogeneous wireless network. α_{ij} (or β_{ij}) is the fraction of the new (or handoff) class- i calls admitted into RAT- j ($\text{RAT-}j \in H_i$).

5.3.4.1 Splitting of Arrival Process

When a new or handoff class- i call arrives into a group of co-located cells, the JCAC algorithm selects a RAT in set H_i for the incoming call. The action of selecting a RAT for each arriving new or handoff call in the group of co-located cells leads to splitting of the arrival process. Fig. 5-2 illustrates the splitting of the arrival process for the two-class three-RAT heterogeneous network.

Let λ_i^n and λ_i^h denote the mean arrival rates of new class- i calls and handoff class- i calls, respectively, in the group of co-located cells. Furthermore, λ_{ij}^n and λ_{ij}^h denote the mean arrival rates of new and handoff class- i calls, respectively in RAT- j . Let α_{ij} and β_{ij} denote the fraction of new and handoff class- i calls admitted into RAT- j respectively ($\text{RAT-}j \in H_i$).

As shown in Fig. 5-2, λ_1^n is split into λ_{11}^n and λ_{12}^n (i.e. $\alpha_{11}\lambda_1^n$ and $\alpha_{12}\lambda_1^n$) and λ_2^n is split into λ_{22}^n and λ_{23}^n (i.e. $\alpha_{22}\lambda_2^n$ and $\alpha_{23}\lambda_2^n$). The same procedure applies to handoff calls where λ_1^h is split into λ_{11}^h and λ_{12}^h (i.e. $\beta_{11}\lambda_1^h$ and $\beta_{12}\lambda_1^h$) and λ_2^h is split into λ_{22}^h and λ_{23}^h (i.e. $\beta_{22}\lambda_2^h$ and $\beta_{23}\lambda_2^h$).

Note that the arrival rates of a split Poisson process are also Poisson [79]. Therefore, given that the mean arrival rate of class- i calls into the group of co-located cells is Poisson, the mean arrival rates of the split class- i calls into RAT- j ($\forall \text{RAT-}j \in H_i$) are also Poisson.

5.3.4.2 Optimal RAT Selection Policy

Given any values of threshold T_{oj} for rejecting new calls in RAT- j ($j=1,\dots,J$) in a heterogeneous wireless network, there exist optimal values of α_{ij} and β_{ij} ($j=1,\dots,J$, $i \in d_j$) that minimize the overall call blocking probability in the heterogeneous wireless network.

Based on the measured values of the arrival rates of each class of new and handoff calls, the optimal policy determination unit finds the optimal values of α_{ij} and β_{ij} ($j=1, \dots, J$, $i \in d_j$). During the next period, T , these values are then used by the joint call admission control algorithm to make a RAT selection decision for each arriving call.

Let $\rho_{new_{i,j}}$ and $\rho_{han_{i,j}}$ denote the load generated by new and handoff class- i calls, respectively, in RAT- j . Then,

$$\rho_{new_{i,j}} = \frac{\lambda_{i,j}^n}{\mu_i^n} = \frac{\alpha_{i,j} \lambda_i^n}{\mu_i^n} \quad \forall i \in d_j, j = 1, \dots, J \quad (5.2)$$

$$\rho_{han_{i,j}} = \frac{\lambda_{i,j}^h}{\mu_i^h} = \frac{\beta_{i,j} \lambda_i^h}{\mu_i^h} \quad \forall i \in d_j, j = 1, \dots, J \quad (5.3)$$

Total load generated by all new calls in RAT- j , p_j^n is:

$$p_j^n = \sum_{i \in d_j} P_{new_{i,j}} \quad (5.4)$$

Total load generated by all handoff calls in RAT- j , p_j^h is:

$$p_j^h = \sum_{i \in d_j} P_{han_{i,j}} \quad (5.5)$$

The optimal RAT selection policy that minimizes overall new call blocking probability in the heterogeneous wireless network is the policy that satisfies the following conditions:

$$\frac{\rho_j^n + \rho_j^h}{T_{0j}} = \frac{\rho_{(j+1)}^n + \rho_{(j+1)}^h}{T_{0(j+1)}} \quad j = 1, 2, \dots, J - 1 \quad (5.6)$$

Equation (5.6) may not always have a feasible solution. In order to find a solution to (5.6), the optimal RAT selection problem is formulated as follows:

$$\begin{aligned}
& \text{Minimize } z = \sum_{j=1}^{J-1} y_j \\
& \text{Subject to: } \frac{\rho_j^n + \rho_j^h}{T_{0j}} = \frac{\rho_{(j+1)}^n + \rho_{(j+1)}^h}{T_{0(j+1)}} + y_j \\
& y_j \geq 0 \quad \forall j = 1, 2, \dots, J-1 \\
& \sum_{j \in h_i} \alpha_{i,j} = 1, \quad \sum_{j \in h_i} \beta_{i,j} = 1 \quad \forall i \\
& \lambda_{i,j}^h = \left(\frac{h^h}{u_i} \right) \lambda_{i,j}^n \quad \forall i \\
& \alpha_{i,j} \geq 0, \beta_{i,j} \geq 0 \quad \forall i, j \in h_i
\end{aligned} \tag{5.7}$$

From the objective function and the constraints, the values of α_{ij} and β_{ij} are obtained using linear programming.

5.4 Markov Chain Model

The optimal JCAC scheme can be modeled as a multidimensional Markov chain. The current state of the heterogeneous system is represented as follows:

$$x = (m_{i,j}, n_{i,j} : i \in \{d_j\}, j = 1, \dots, J) \tag{5.8}$$

The non-negative integers $m_{i,j}$ and $n_{i,j}$ denote respectively, the number of ongoing new and handoff class- i calls in RAT- j . Let S denote the state space of all admissible state s as it evolves over time. The state S is given as:

$$\begin{aligned}
S = \{ & x = (m_{i,j}, n_{i,j} : \forall i \in \{d_j\}, j = 1, \dots, J) | \\
& \sum_{i \in d_j} m_{i,j} \leq T_{0j} \quad \& \quad \sum_{i \in d_j} (m_{i,j} + n_{i,j}) \leq C_j \quad \forall j \}
\end{aligned} \tag{5.9}$$

The call admission decision epochs are the arrival of a new or handoff call. When the system is in state s , an accept/reject decision must be made for each type of possible arrival in the group of co-located cells.

The call admission action space A can be expressed as:

$$A = \{a = (a_1^n, \dots, a_k^n, a_1^h, \dots, a_k^h) : a_i^n, a_i^h \in \{v : v = 0 \text{ or } v \in h_i\}, \forall i\} \quad (5.10)$$

where a_i^n (or a_i^h) denotes the action taken on arrival of a new (or handoff) class- i call in the group of co-located cells. a_i^n (or a_i^h) = 0 means reject the new (or handoff) class- i call. a_i^n (or a_i^h) = j means accept the new (or handoff) class- i call into RAT- j ($j \in h_i$).

Let $P(s)$ denote the steady state probability that the system is in state s ($s \in S$). $P(s)$ is obtained as:

$$P(s) = \frac{1}{G} \prod_{j=1}^J \prod_{i \in d_j} \frac{(\rho_{new\ i,j})^{m_{i,j}}}{m_{i,j}!} \frac{(\rho_{han\ i,j})^{n_{i,j}}}{n_{i,j}!} \quad \forall s \in S \quad (5.11)$$

where G is a normalization constant given by:

$$G = \sum_{s \in S} \prod_{j=1}^J \prod_{i \in d_j} \frac{(\rho_{new\ i,j})^{m_{i,j}}}{m_{i,j}!} \frac{(\rho_{han\ i,j})^{n_{i,j}}}{n_{i,j}!} \quad (5.12)$$

5.4.1 New Call Blocking Probability

A new class- i call will be blocked in the group of co-located cells if none of the RATs in set H_i can accommodate the call. Let $S_{bi} \subset S$ denote the set of states in which a new class- i call is blocked in the group of co-located cells. It follows that:

$$\begin{aligned}
S_{b_i} &= \{s \in S : (b_i + \sum_{x \in d_j} b_x m_{x,j} > T_{0j} \vee \\
& b_i + \sum_{x \in d_j} b_x (m_{x,j} + n_{x,j}) > C_j) \quad \forall \quad j \in h_i \}
\end{aligned} \tag{5.13}$$

Thus the new blocking probability (NCBP), P_{b_i} , for a new class- i call in the group of co-located cells is given by:

$$P_{b_i} = \sum_{s \in S_{b_i}} P(s) \tag{5.14}$$

The overall new call blocking probability, P_b , for all classes of calls is given by:

$$P_b = \sum_{i=1}^K (\lambda_i / \sum_{i=1}^k \lambda_i) P_{b_i} \tag{5.15}$$

5.4.2 Handoff Call Dropping Probability

A handoff class- i call is dropped in the group of co-located cells if none of the RATs in set H_i can accommodate the call. Let $S_{d_i} \subset S$ denote the set of states in which a handoff class- i call is dropped in the group of co-located cells. It follows that:

$$S_{d_i} = \{s \in S : (b_i + \sum_{x \in d_j} b_x (m_{x,j} + n_{x,j}) > C_j \quad \forall \quad j \in h_i) \} \tag{5.16}$$

Thus the handoff call dropping probability (HCDP), P_{d_i} , for a handoff class- i call in the group of co-located cells is given by:

$$P_{d_i} = \sum_{s \in S_{d_i}} P(s) \tag{5.17}$$

The overall handoff call dropping probability, P_d , for all classes of calls is given by:

$$P_d = \sum_{i=1}^K (\lambda_i / \sum_{i=1}^k \lambda_i) P_{d_i} \tag{5.18}$$

5.5 Performance Evaluation

In this section, the performance of the OJCAC scheme is evaluated with respect to NCBP and HCDP using the two-class three-RAT heterogeneous wireless network in Fig. 5-2. For comparison, an equal-sharing JCAC scheme (called EJCAC scheme), which equally shares the radio resources of each RAT among all class- i calls ($\forall i \in d_j$) is also modeled and evaluated. The simulation is conducted using MATLAB. The following system parameters are used: $C_1=20$, $T_{01}=12$, $C_2=20$, $T_{02}=12$, $C_3=10$, $T_{03}=6$, $hh=0.5$, $b_1=1$, $b_2=2$, $\mu_1^n = \mu_2^n = 0.5$, $\lambda_1^n = [1, 6]$, $\lambda_2^n = \lambda_1^n$. Fig. 5-4 shows the effect of varying the new call arrival rate on the NCBP (Pb1 and Pb2) of the OJCAC and EJCAC schemes. As shown in Fig. 5-4, Pb1 and Pb2 increase with arrival rate for both OJCAC and EJCAC schemes. However, Pb2 of the OJCAC scheme is lower than the corresponding Pb2 of the EJCAC scheme. In the extreme case, Pb2 of the OJCAC is about 30% of the Pb2 of EJCAC scheme. The OJCAC scheme is able to reduce Pb2 by determine the optimal ratio of each class- i calls that is admitted into each RAT- j ($\forall j \in h_i$). However, the reduction in Pb2 of the OJCAC scheme is at the expense of its Pb1. It can be seen that the Pb1 of the OJCAC scheme is a little higher than the corresponding Pb1 of the EJCAC scheme.

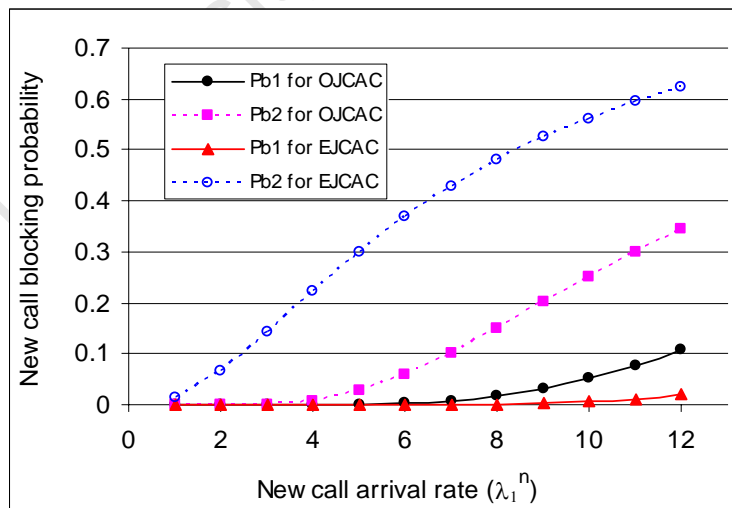


Figure 5-4. Effect of varying the call arrival rate on new call blocking probability of class-1 and class-2 calls ($\lambda_1=\lambda_2$).

Fig. 5-5 shows the overall NCBP (i.e. the Pb of class-1 and class-2 calls combined) for the OJCAC scheme and the EJCAC scheme. It can be seen that the Pb of both schemes increases with increase in arrival rate. However, the Pb of the OJCAC scheme is less than the corresponding Pb of the EJCAC scheme. Thus, the OJCAC scheme reduces the overall NCBP in the heterogeneous wireless network. In the extreme case, Fig. 5-5 shows that for the OJCAC scheme, the overall NCBP is about 17% of the overall NCBP of the EJCAC scheme.

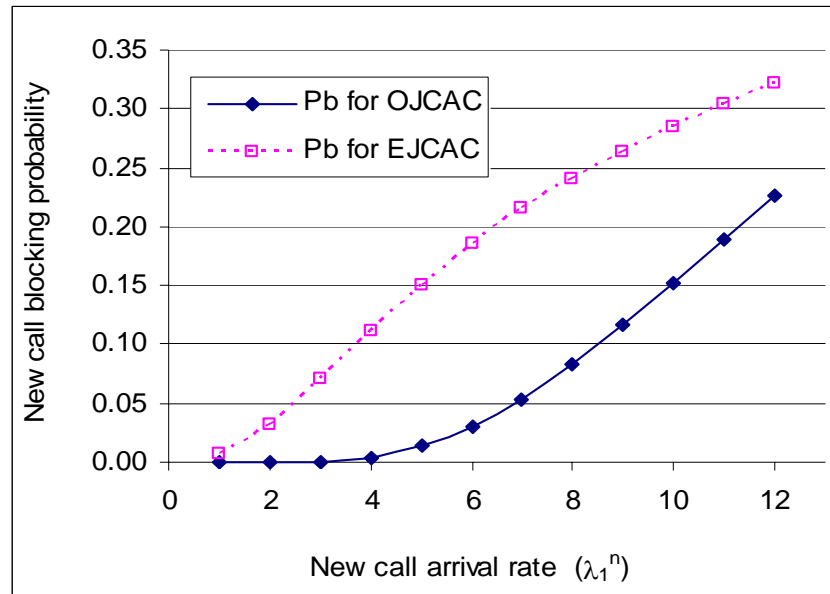


Figure 5-5. Effect of varying the call arrival rate on overall new call blocking probability ($\lambda_1=\lambda_2$).

Fig. 5-6 shows the effect of varying the new call arrival rate on the HCDP (Pd1 and Pd2) of the OJCAC and EJCAC schemes. As shown in Fig. 5-6, Pd1 and Pd2 increases with arrival rate for both OJCAC and EJCAC schemes. However, Pd2 of the OJCAC is lower than the corresponding Pd2 of the EJCAC scheme. In the extreme case, the Pd2 of the OJCAC scheme is about 10% of the Pd2 of EJCAC scheme. However, the reduction in Pd2 of the OJCAC scheme is at the expense of its Pd1. It can be seen that the Pd1 of the OJCAC scheme is a little higher than the corresponding Pd1 of the EJCAC scheme.

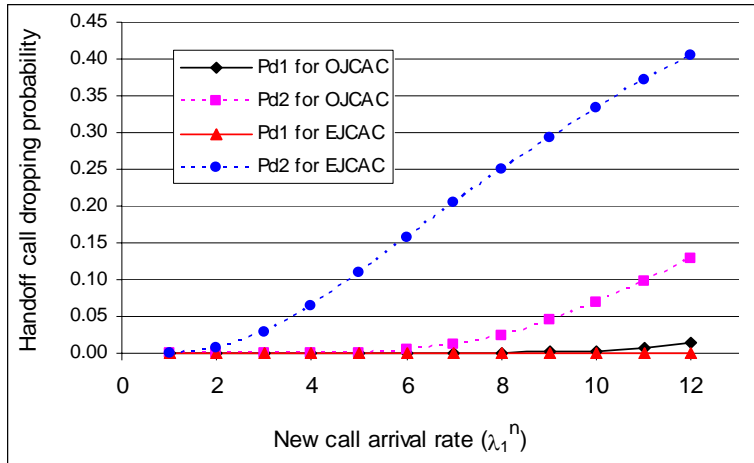


Figure 5-6. Effect of varying the call arrival rate on handoff call dropping probability of class-1 and class-2 calls ($\lambda_2=\lambda_1$).

Fig. 5-7 shows the effect of varying the new call arrival rate on the overall HCDP (i.e. the Pd of class-1 and class-2 calls combined) of the OJCAC scheme and the EJCAC scheme. It can be seen that the Pd of both schemes increases with increase in arrival rate. However, the Pd of the OJCAC scheme is less than the corresponding Pd of the EJCAC scheme. Thus the OJCAC scheme reduces the overall HCDP in the heterogeneous wireless network. In the extreme case, Fig. 5-7 shows that the OJCAC scheme reduces the overall HCDP of the calls to about 37% of the HCDP of the EJCAC scheme.

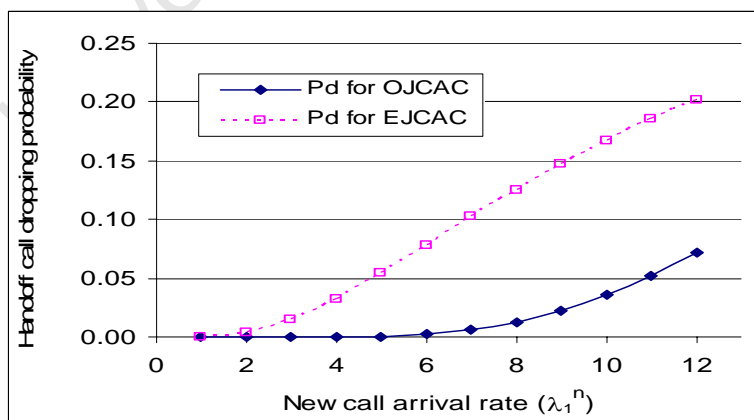


Figure 5-7. Effect of varying the call arrival rate on overall handoff call dropping probability ($\lambda_2=\lambda_1$).

One of the objectives of the OJCAC scheme is to prioritize handoff calls over new calls. Fig. 5-8 compares the P_{b1} and P_{d1} of the OJCAC scheme. It can be seen that the P_{b1} and P_{d1} of the OJCAC scheme increases with increase in arrival rate. However, P_{d1} of the OJCAC scheme is always less than the corresponding P_{b1} . This is because the OJCAC scheme prioritizes handoff calls over new calls by using for new calls, a lower call rejection threshold than that of handoff calls.

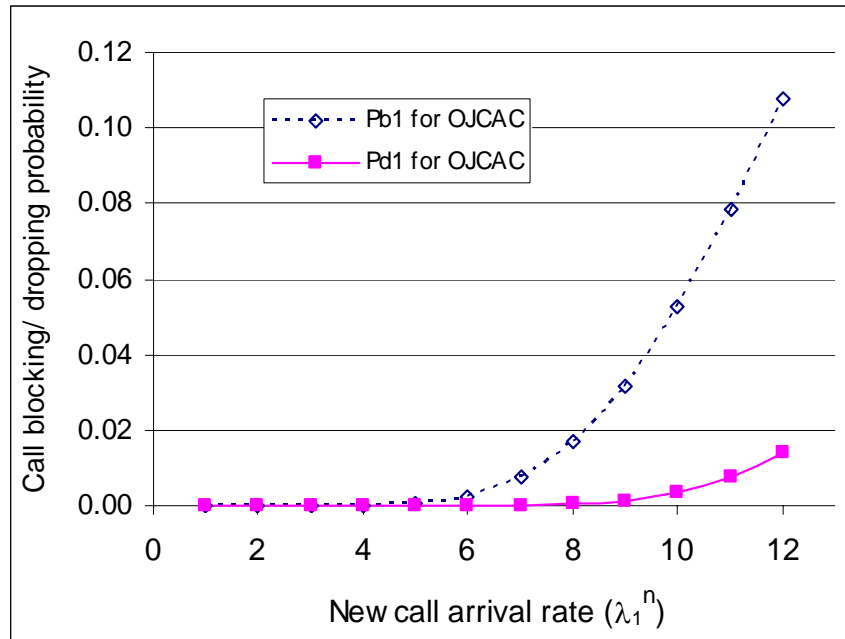


Figure 5-8. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls for the OJCAC scheme ($\lambda_2=\lambda_1$).

In a similar way, Fig. 5-9 compares the P_{b2} and P_{d2} for the OJCAC scheme. It can be seen that the P_{b2} and P_{d2} of the OJCAC scheme increases with increase in arrival rate. However, P_{d2} of the OJCAC scheme is always less than the corresponding P_{b2} . This is because the OJCAC scheme prioritizes handoff calls over new calls by using for new calls, a lower call rejection threshold than that of handoff calls.

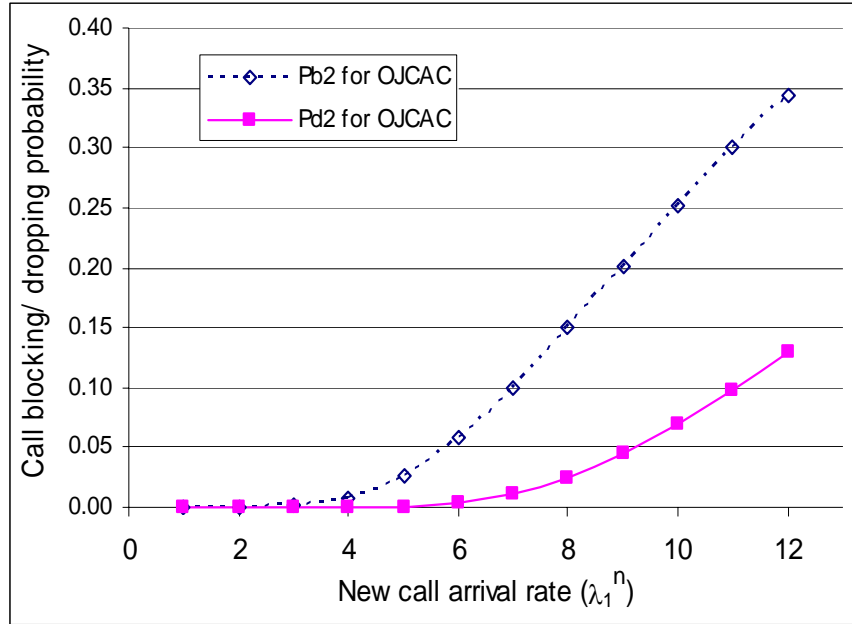


Figure 5-9. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls for the OJCAC scheme ($\lambda_2=\lambda_1$).

Fig. 5-10 shows the effect of varying the new call rejection threshold, T_0 on Pb and Pd for the OJCAC scheme. As shown in Fig. 5-10, at low threshold values, the NCBP, Pb1 and Pb2 for the OJCAC schemes are high whereas the HCDP, Pd1 and Pd2 are low. As the threshold value, T_0 increases, Pb1 and Pb2 decrease because new calls are given more access to the available bandwidth. On the other hand, Pd1 and Pd2 increase as a result of higher degree of sharing between new and handoff calls. However, Pb1 is always less than the corresponding Pb2 because class-1 calls require less bbu than class-2 calls. Thus it is possible to block a class-2 call due to unavailability of radio resources while it is still possible to admit a class-1 call. For the same reason, Pd1 is always less than the corresponding Pd2.

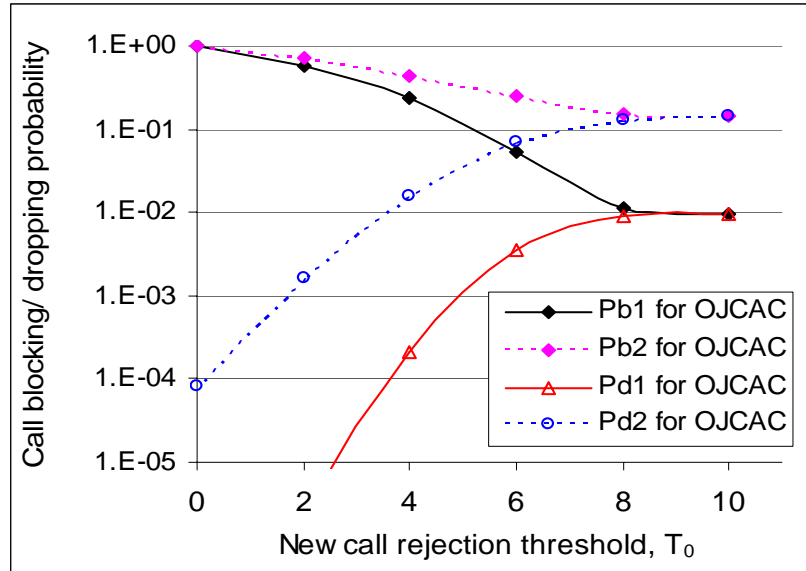


Figure 5-10. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls ($\lambda_2 = \lambda_1$).

In the following graphs, the effect of varying the ratio of λ_1^n and λ_2^n on the performance of the OJCAC scheme is investigated. The system parameters used are the same as those used in Fig. 5-4 to Fig. 5-10 except for the value of λ_2^n which is double (i.e. $\lambda_2^n = 2\lambda_1^n$).

Fig. 5-11 shows the effect of varying the new call arrival rate on the NCBP (Pb1 and Pb2) of the OJCAC and EJCAC schemes. As shown in Fig. 5-11, Pb1 and Pb2 increases with arrival rate for both OJCAC and EJCAC schemes. The reduction in Pb2 of the OJCAC scheme is at the expense of its Pb1. It can be seen that the Pb1 of the OJCAC scheme is a little higher than the corresponding Pb1 of the EJCAC. Fig 5-11 shows that with the variation in the ratio of λ_1 and λ_2 , the OJCAC scheme still outperforms the EJCAC scheme.

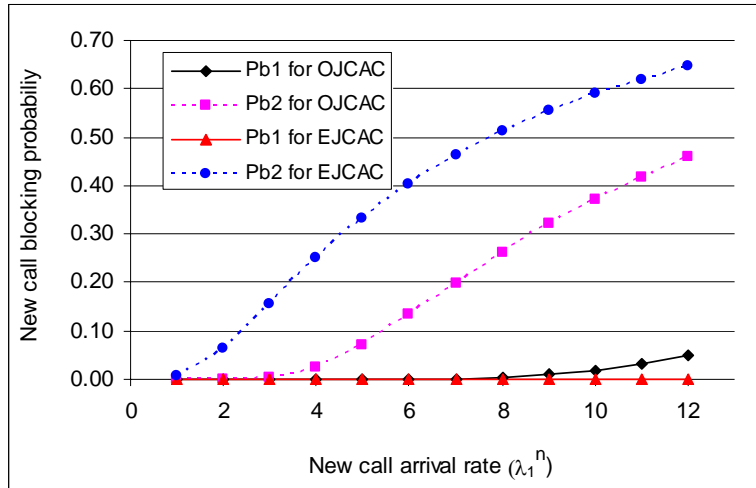


Figure 5-11. Effect of varying the call arrival rate on new call blocking probability of class-1 and class-2 calls ($\lambda_2=2\lambda_1$).

Fig. 5.12 shows the effect of varying the new call arrival rate on the overall NCBP (i.e. the Pb of class-1 and class-2 calls combined) for the OJCAC scheme and the EJCAC scheme. It can be seen that the Pb of both schemes increases with increase in arrival rate. However, the Pb of the OJCAC scheme is always less than the corresponding Pb of the EJCAC scheme. Fig 5.12 shows that even with the variation in the ratio of λ_1 and λ_2 , the OJCAC scheme reduces the overall NCBP in the heterogeneous wireless network.

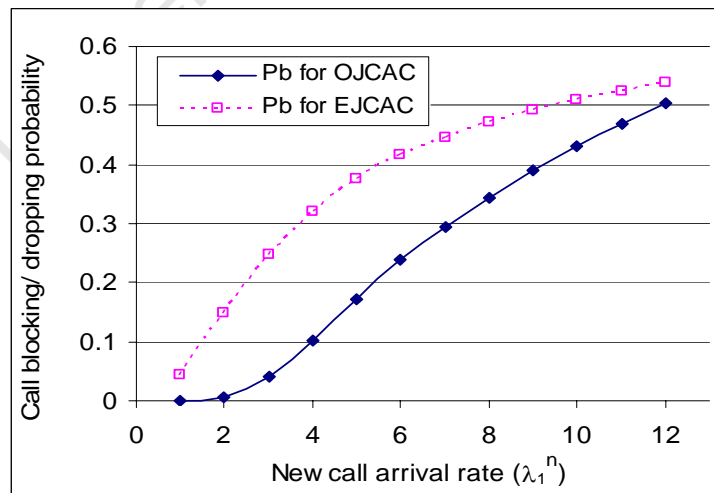


Figure 5-12. Effect of varying the call arrival rate on overall new call blocking probability ($\lambda_2=2\lambda_1$).

Fig. 5-13 shows the effect of varying the new call arrival rate on the HCDP (Pd1 and Pd2) of the OJCAC and EJCAC schemes. As shown in Fig. 5-13, Pd1 and Pd2 increase with increase in arrival rate for both OJCAC and EJCAC schemes. However, Pd2 of the OJCAC scheme is lower than the corresponding Pd2 of the EJCAC scheme. However, the reduction in Pd2 of the OJCAC scheme is at the expense of its Pd1. It can be seen that the Pd1 of the OJCAC scheme is a little higher than the corresponding Pd1 of the EJCAC scheme. Fig 5-13 shows that with the variation in the ratio of λ_1 and λ_2 , the OJCAC scheme still outperforms the EJCAC scheme.

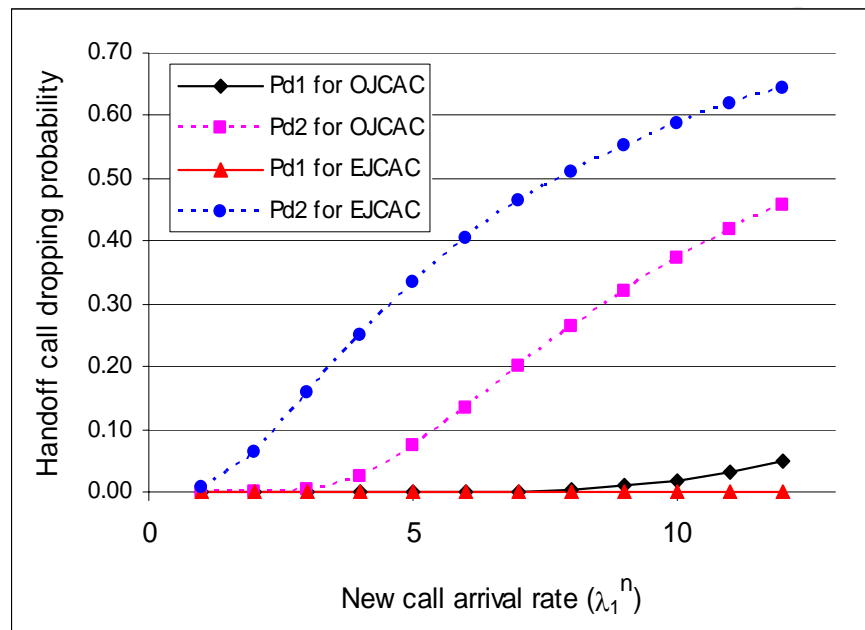


Figure 5-13. Effect of varying the call arrival rate on handoff call dropping probability of class-1 and class-2 calls ($\lambda_2=2\lambda_1$).

Fig. 5-14 shows the effect of varying the new call arrival rate on the overall HCDP (i.e. the Pd of class-1 and class-2 calls combined) for the OJCAC scheme and the EJCAC scheme. It can be seen that the Pd of both schemes increases with increase in arrival rate. However, the Pd of the OJCAC scheme is less than the corresponding Pd of the EJCAC. Thus the OJCAC reduces the overall HCDP in the heterogeneous wireless network despite the variation in the ratio of λ_1 and λ_2 .

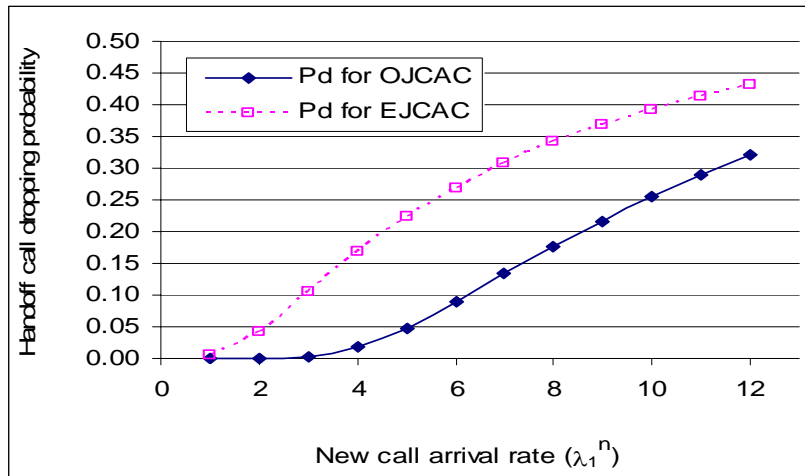


Figure 5-14. Effect of varying the call arrival rate on overall handoff call dropping probability ($\lambda_2=2\lambda_1$).

Fig. 5-15 compares the Pb1 and Pd1 for the OJCAC scheme when $\lambda_2=2\lambda_1$. It can be seen that the Pb1 and Pd1 of the OJCAC scheme increases with increase in arrival rate. However, Pd1 of the OJCAC scheme is less than the corresponding Pb1. This is because the OJCAC scheme prioritizes handoff calls over new calls by using for new calls, a lower call rejection threshold than that of handoff calls.

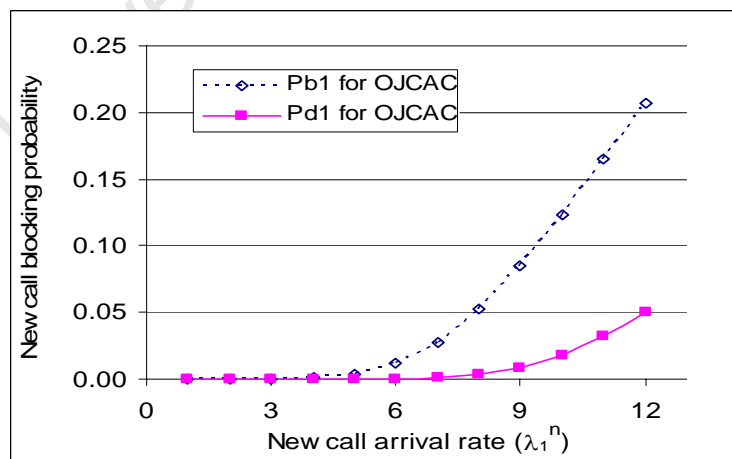


Figure 5-15. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-1 calls for the OJCAC scheme ($\lambda_2=2\lambda_1$).

In a similar way, Fig. 5-16 compares the Pb2 and Pd2 for the OJCAC scheme when $\lambda_2=2\lambda_1$. It can be seen that the Pb2 and Pd2 of the OJCAC scheme increases with increase in arrival rate. However, Pd2 of the OJCAC scheme is less than the corresponding Pb2. This is because the OJCAC scheme prioritizes handoff calls over new calls by using for new calls, a lower call rejection threshold that that of handoff calls.

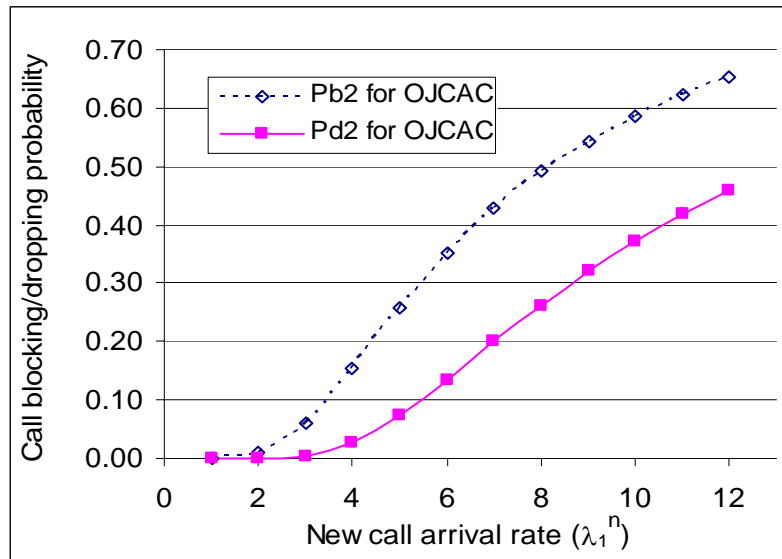


Figure 5-16. Effect of varying the call arrival rate on the new call blocking probability and handoff call dropping probability of class-2 calls for the OJCAC scheme ($\lambda_2=2\lambda_1$).

Fig. 5-17 shows the effect of varying the new call rejection threshold, T_0 on Pb and Pd for the OJCAC scheme when $\lambda_2=2\lambda_1$. As shown in Fig. 5-17, at low threshold values, the NCBP, Pb1 and Pb2 for the OJCAC schemes are high whereas the HCDP, Pd1 and Pd2 are low. As the threshold value, T_0 increases, Pb1 and Pb2 decrease because new calls are given more access to the available bandwidth. On the other hand, Pd1 and Pd2 increase as a result of higher degree of sharing between new and handoff calls. However, Pb1 is always less than the corresponding Pb2 because class-1 calls require less bbu than class-2 calls. For the same reason, Pd1 is less than the corresponding Pd2.

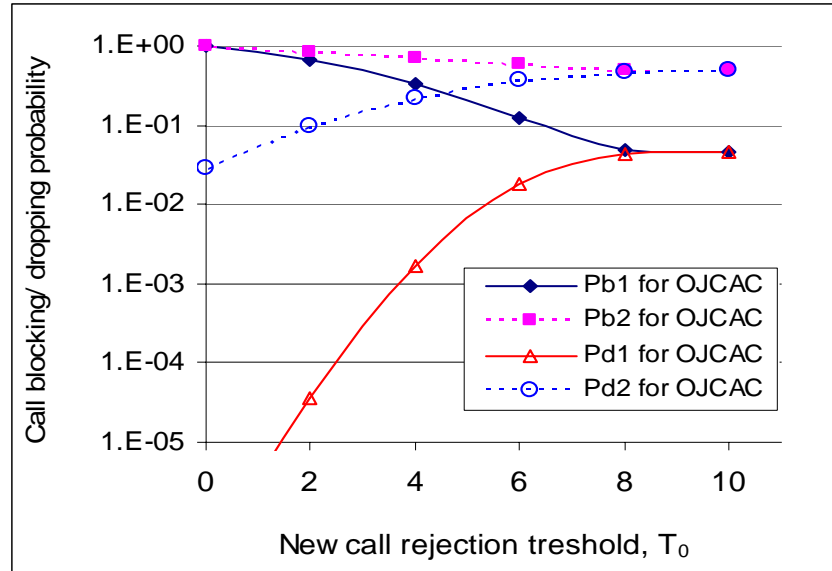


Figure 5-17. Effect of varying the new call rejection threshold, T_0 on the new call blocking probability and handoff call dropping probability of class-1 calls, ($\lambda_2=2\lambda_1$).

5.6 Summary

This chapter proposes an optimal JCAC scheme to reduce call blocking probability in heterogeneous wireless networks. Using Markov decision process, an analytical model is developed for the OJCAC scheme. New call blocking probability and handoff call dropping probability are then derived. Optimal RAT selection policy is obtained by using linear programming technique. Performance of the OJCAC scheme is compared with that of EJCAC scheme in the same heterogeneous wireless network. Results show that the OJCAC scheme improves connection-level QoS in the heterogeneous wireless network by reducing overall new call blocking probability and overall handoff call dropping probability. Moreover the OJCAC scheme prioritizes handoff calls over new calls by using a lower call rejection thresholds for new calls.

6 Dynamic Pricing for Load balancing in a Heterogeneous Wireless Network Using a Multiple-Criteria JCAC Scheme

6.1 Introduction

This chapter addresses the problem of highly unbalanced traffic load in a heterogeneous wireless network using a multiple-criteria user-centric JCAC scheme. Unbalanced traffic load is caused by independent users' preferences for specific RATs in the heterogeneous wireless network. Highly unbalanced load in heterogeneous wireless networks results in high call blocking/dropping probability and poor radio resource utilization. In this chapter, dynamic pricing is proposed to even out the unbalanced traffic load among available RATs in heterogeneous wireless networks where users' preferences are considered in making RAT selection decisions. By dynamically adjusting the service price in each of the available RATs, the proposed JCAC scheme uniformly distributes traffic load, as much as possible, among available RATs. Balancing of traffic load among the available RATs in heterogeneous wireless networks reduces overall call blocking/dropping probability and improves radio resource utilization.

6.2 Users' Preferences, Load Balancing, and Pricing

This section discusses user' preferences, load balancing, and pricing, as well as how they are interrelated in heterogeneous wireless networks.

6.2.1 Users' Preferences

Next generation wireless network will be user-centric [80]. Therefore users' preferences for a particular RAT will be considered in making call admission decisions. Users can set their preferences for a particular RAT on their mobile devices, and can even dynamically change their preferences with time. Some examples of factors that determine users'

preference for a particular RAT are service price, data rate, battery power consumption, security level, etc. These factors are briefly described as follows.

(1) Least service cost: Service cost varies from one access network to another. This variation in service cost can be attributed to differences in equipment cost and the cost of procuring spectrum license for different RATs. As a result, a user may prefer to be connected through the cheapest available RAT so that the overall service cost will be reduced.

(2) Maximum data rate: Different RATs offer different data rates for different types of calls. Most multimedia applications are adaptive. For example voice can be encoded at 16 kbps, 32 kbps, 64 kbps, and 128 kbps by choosing appropriate encoding mechanisms. Similarly, video applications can be made rate adaptive by using, for instance, a layered coding method. In layer coding method, the lowest layer (i.e., the base layer) contains the critical information for decoding the image sequence at its minimum visual quality. Additional layers provide increasing quality. As an illustration, if one would watch a 30-minute video-clip encoded at 256 kbps and 64 kbps respectively. At 256 kbps, one would see better pictures with better resolution than at 64 kbps. Therefore, a user of real-time service may prefer to be connected through the RAT with the highest data rate in order to enhance service quality. A user of non-real time service may prefer to be connected through the RAT with the highest data rate in order to reduce service-delivery time.

(3) Least battery power consumption: One of the key challenges in wireless communication is efficient use of energy stored by the batteries of mobile terminals. Efficient power utilization in mobile terminals will avoid the need for frequent batteries recharge. Different RATs, for example, use different modulation and coding techniques, which impact on mobile terminals' power consumption. Therefore, a user may prefer to be connected through a RAT that will minimize its energy consumption. Battery power consumption is usually specified using linguistic values such as very low, low, average, high, and very high.

(4) Highest network security: Different access technologies offer different levels of security. A user may prefer to be connected through the RAT with the highest security level. Security level is usually specified using linguistic values such as very low, low, average, high, and very high.

From the foregoing discussion, it can be seen that users' preferences for a particular RAT are determined by a number of factors (criteria). A major challenge in the design of multiple-criteria JCAC algorithms is how to combine many RAT selection criteria in making RAT selection decision for each arriving call. Current approaches have incorporated fuzzy logic and MADM (Multiple Attribute Decision Making) technique [19, 20, 21].

Chan *et al* [19] presented a segment selection algorithm based on the concept of fuzzy multiple objective decision making (MODM). Seven example criteria are used in the algorithm namely, signal strength, bandwidth, charging model, reliability, latency, battery status and the user's preferred segment (priority). The purpose of the segment selection algorithm is to select the most suitable RAT for a particular service class based of the criteria mentioned above.

Zhang [20] proposed an approach, which uses fuzzy MADM (multiple attribute decision making) technique. The fuzzy MADM method operates in two steps. The first step is to convert the imprecise fuzzy variable to crisp numbers. The second step is to use classical MADM technique to determine the ranking order of the candidate networks. The highest-ranking RAT is then selected for the call.

Wilson *et al* [21] proposed a decision strategy for making the optimal choice of wireless access networks. Fuzzy logic is used as the inference mechanism and a prototype is developed. The prototype uses two metrics from a candidate network, a metric from application requirements, and user defined criteria as input.

The major problem with the above algorithms is that they can lead to highly-unbalanced load among different RATs because users act independently. Thus most of the users may prefer to be connected through a particular RAT. Moreover, no analytical model has been

developed to study connection-level QoS (new call blocking probability and handoff call blocking probability) in the user-centric JCAC algorithms discussed.

6.2.2 Load Balancing

Each RAT in the heterogeneous wireless network has a maximum load capacity. If some RATs are overloaded while some other RATs are underutilized, it results in poor utilization of radio resources. Balancing of traffic load among multiple RATs in heterogeneous wireless networks allows for a better utilization of the radio resources [13].

Pillekeit *et al* [13] proposed a forced-based load balancing JCAC algorithm for co-located UMTS/GSM networks. The algorithm is triggered only when the differential load between the two networks is above a certain threshold. A scenario in which all UMTS and GSM cells are co-located and all mobile phones use circuit switched voice service is simulated. In GSM mode, the mobile terminals use a full rate codec with a data rate of 13 Kbps whereas in UMTS mode, they use a data rate of 12.2 Kbps (spreading factor of 128 is assumed). For comparison purposes, a scenario where there is no joint call admission control between the two RATs (i.e. there is independent CAC) is also simulated. Results show that with the load balancing JCAC algorithm in place, a gain of 8.4% of the overall traffic capacity can be achieved compared to the sum of the traffic capacity of the two networks with independent JCAC algorithms.

Gelabert *et al* [12] evaluated the performance of a load-balancing RAT selection algorithm for new calls in a heterogeneous wireless network which consists of a UTRAN and a GERAN. Seven collocated omnidirectional cells are considered for GERAN and UTRAN. For comparison purposes, a service-based JCAC algorithm is also simulated. The service-based JCAC algorithm allocates users according to the demanded service-type. Results also show that the load-based JCAC algorithm has a higher total aggregated throughput than that of service-based JCAC algorithm.

In the two load-balancing JCAC algorithms discussed above, users' preferences are not considered in making RAT selection decisions. In NGWN, it is very necessary to consider users' preferences in making RAT selection decision because NGWN will be user-centric.

Moreover, no analytical model is presented to investigate connection-level QoS in the previous JCAC algorithms.

6.2.3 Pricing

Pricing in communication networks has received a lot of attention in the literature. The evolution of pricing schemes for communication networks is driven by the deployment of new technologies and services [81]. In homogeneous wireless networks, it has been noted that network users act independently and sometimes “selfishly”, regardless of the prevailing network traffic conditions. Therefore, congestion has always been a problem in the network, especially during the peak period [82, 83].

To avoid (or reduce) congestion, pricing has been used as a mechanism that give users incentives to behave in ways that improve the overall utilization and performance of the network. Pricing mechanisms for wireless networks can be classified into static, dynamic, or auction-based schemes [84]. Static pricing is the simplest pricing policy, in which prices are fixed and independent of the state of the system [84]. In dynamic pricing, the network adapts prices as the traffic load changes [85]. Prices rise in accordance with demand, thereby deterring additional users from accessing the network or holding network resources for long periods, during congestion time. In auction-based pricing schemes, users attach a bid to each packet indicating the willingness to pay for the delivery of the packet, and the network serves packets in descending order of their bids [86].

Among these three classifications of pricing schemes, dynamic pricing is the most powerful and flexible mechanism [84]. In homogeneous wireless networks, dynamic pricing has been used to achieve a socially optimal bandwidth allocation [87], maximize revenue [88], obtain an incentive-compatible class allocation [89], and achieve efficient power control [90].

Hou *et al.* [91] investigated the integration of dynamic pricing scheme with call admission control in order to efficiently and effectively control the use of radio resources in homogeneous cellular networks. Their results show an improvement in resource utilization. Sarayadar *et al.* [92] incorporated a pricing mechanism into the power control

of homogeneous wireless data networks in order to improve users' utility. Vietrbo *et al* [93] proposed dynamic pricing strategies for connection-oriented services in homogeneous wireless systems. Results show 25% improvement in the network revenue with respect to the static (flat-rate) policy, and the blocking probability is halved. However, all the pricing schemes mentioned above are designed for congestion control in homogeneous wireless networks.

This work proposes the use of dynamic pricing to balance traffic load among available RATs in heterogeneous wireless networks where a user-centric JCAC scheme is employed. For any call arrival rate, the proposed JCAC scheme evenly distribute traffic load, as much as possible, among the available RATs in heterogeneous wireless networks. The objectives of the proposed scheme are to enhance connection-level QoS and improve overall radio resource utilization.

6.3 System Model

This study considers a heterogeneous wireless network, which consists of J number of RATs with co-located cells, similar to [12, 13, 16]. Wireless networks such as GSM, GPRS, UMTS, EV-DO, etc., can have the same and fully overlapped coverage, which is technically feasible, and may also save installation cost [66]. Fig. 6-1 illustrates a two-RAT heterogeneous wireless network.

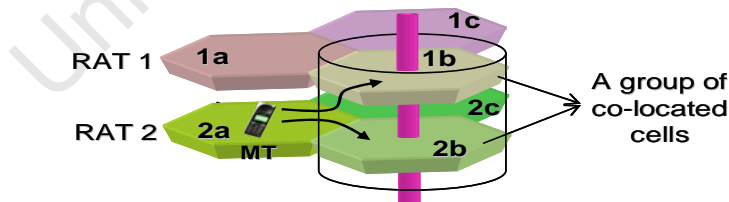


Figure 6-1. A two-RAT heterogeneous wireless network.

Radio resources are jointly managed in the heterogeneous network and each cell in RAT- j ($j=1, \dots, J$) has a total of B_j basic bandwidth units (bbu). The physical meaning of a unit of radio resources (such as time slots, code sequence, etc.) is dependent on the specific technological implementation of the radio interface [67]. However, no matter which

multiple access technology (FDMA, TDMA, CDMA, or OFDM) is used, system capacity can be interpreted in terms of effective or equivalent bandwidth [68-70]. Therefore, the bandwidth of a call represents the number of bbu that is adequate for guaranteeing the desired QoS for the call, which is similar to the approach used for homogeneous networks in [70-72]. Generally, it is assumed that packet-level QoS is stochastically assured by allocating at least the minimum effective bandwidth required to guarantee a given maximum probability on packet drop, delay, and jitter [73].

The approach used in this research is based on decomposing a heterogeneous wireless network into groups of co-located cells. As shown in Fig. 6-1, cell 1a and cell 2a form a group of co-located cells. Similarly, cell 1b and cell 2b form another group of co-located cells, and so on. Based on the following assumption commonly made in homogeneous cellular networks, it is assumed that the types and amount of traffic are statistically the same in all cells of each RATs [70, 71, 74, 75, 76, 77]. Therefore, the types and amount of traffic are statistically the same in all groups of co-located cells.

A newly arriving call will be admitted into one of the cells in the group of co-located cells where the call is located. When a mobile subscriber using a multimode terminal and having an ongoing call is moving from one group of co-located cells to another group of co-located cells, the ongoing call must be handed over to one of the cells in the new group of co-located cells. For example (Fig. 6-1), an ongoing call can be handed over from cell 2a to cell 2b or from cell 2a to cell 1b. Note that the handover consists of both horizontal and vertical handovers.

The correlation between the groups of co-located cells results from handoff connections between the cells of corresponding groups. Under this formulation, each group of co-located cells can be modelled and analyzed individually. Therefore, the research focuses on a single group of co-located cells.

The heterogeneous network supports K classes of calls. Each class is characterized by minimum and maximum bandwidth requirements, arrival distribution, and channel holding time. Each class- i call requires a discrete bandwidth value, $b_{i, w}$, where $b_{i, w}$ belongs to the

set $L_i = \{b_{i,w}\}$ for $i=1, 2, \dots, K$ and $w=1, 2, \dots, W_i$. W_i is the number of different bandwidth values that a class- i call can be allocated. $b_{i,1}$ (also denoted as $b_{i,min}$) and b_{i,W_i} (also denoted as $b_{i,max}$) are respectively, the minimum and maximum bandwidth that can be allocated to a class- i call. Note that $b_{i,w} < b_{i,(w+1)}$ for $i=1, 2, \dots, K$ and $w=1, 2, \dots, (W_i-1)$.

The proposed JCAC scheme always allocates the highest data rate available for new or handoff class- i call in the RAT in which the call is admitted. For example, in a two-RAT heterogeneous wireless network, if class-1 calls can be admitted at the following two different bandwidth levels: 32 Kbps and 64 Kbps. Assuming that the highest data rate for class-1 calls supported by RAT-1 is 32 Kbps whereas the highest data rates for class-1 calls supported by RAT-2 is 64 Kbps. If a class-1 call is admitted into RAT-1, it will be allocated 32 Kbps whereas if the call is admitted into RAT-2, it will be allocated 64 Kbps.

Following the general assumption in cellular networks, new and handoff class- i calls arrive in the group of co-located cells according to Poisson process with rate λ_i^n and λ_i^h respectively. The call holding time (CHT) of a class- i call is assumed to follow an exponential distribution with mean $1/\mu_i$ [74, 75].

To characterize mobility, the cell residence time (CRT), i.e., the amount of time during which a mobile terminal stays in a cell (same as the time it stays in a group of co-located cells) during a single visit, is assumed to follow an exponential distribution with mean $1/h$, where the parameter h represents the call handoff rate. It is assumed that the CRT is independent of the service class.

The channel holding time is the minimum of the CHT and the CRT. Because minimum of two exponentially distributed random variables is also exponentially distributed [78], the channel holding time for new class- i calls, and for handoff class- i call, is assumed to be exponentially distributed with means $1/\mu_i^n$ and $1/\mu_i^h$ respectively.

Note that this set of assumptions has been widely used for homogeneous cellular networks in the literature, and is found to be generally applicable in the environment where the number of mobile users is larger than the number of channels [78].

6.4 Proposed Scheme

In this section, the proposed JCAC scheme, which consists of the following four components, is discussed. The components are joint call admission controller, call admission rate measurement unit, price-update unit, and bandwidth reservation unit. The components are connected as shown in Fig. 6-2, are described in the following.

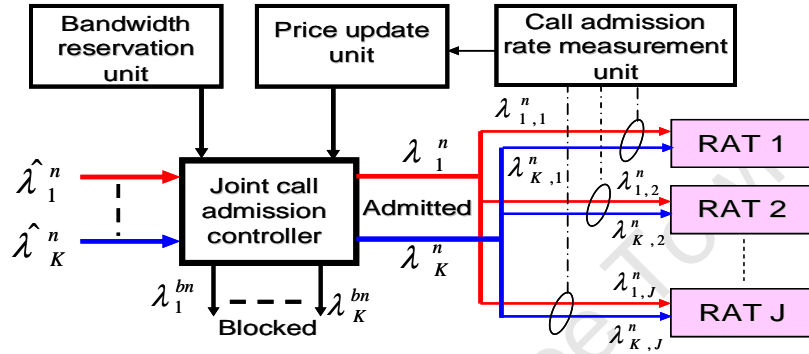


Figure 6-2. Proposed JCAC scheme.

6.4.1 Joint Call Admission Controller

The joint call admission controller implements the JCAC algorithm. The basic function of the JCAC algorithm is to make call admission and RAT selection decisions. During call setup, a multi-mode mobile terminal requesting a service sends a request to the joint call admission controller which implements the JCAC algorithm. The service request contains the call type (new or handoff), service class, minimum and maximum bandwidth required, and weights assigned by the user to each of the RAT selection criterion. The weight is used in determining the user's preference for a particular RAT. Based on the service request information, the JCAC algorithm decides which of the available RATs is most suitable for the incoming call and then notifies the mobile terminal of its decision. For an incoming new call, the response will either be "call accepted into RAT-j ($RAT-j \in H$)" or "call blocked" where H is the set of available RATs. For an incoming handoff call, the response will either be "call accepted into RAT-j ($RAT-j \in H$)" or "call dropped". Fig. 6-3 illustrates the call request/ response procedure.

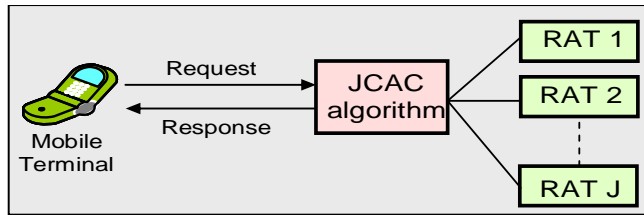


Figure 6-3. Call request/response procedure.

Fig. 6-4 shows the assignment of weight to each of the RAT selection criteria by a user. As shown in Fig. 6-4, for each class of calls (e.g., voice call, video call, etc.) each user will assign weights to the RAT selection criteria. This assignment of weights for a particular class of calls is done once and it will always be used in selecting a RAT for that class of calls for the user. However, users can make changes to the weights previously assigned to the selection criteria. The weight represents the relative importance of each RAT selection criterion to each user.

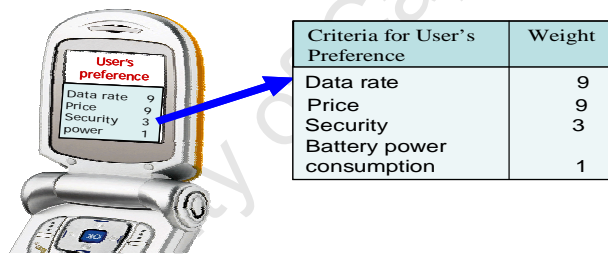


Figure 6-4. Assignment of weight to each of the RAT selection criteria.

6.4.1.1 Fuzzy MADM Technique

The JCAC algorithm uses fuzzy MADM technique to select the most appropriate RAT for each user. Generally, fuzzy MADM technique consists of two stages [20, 94]. In the first stage, fuzzy data are converted into real data. In the second stage, classical MADM is used to determine the ranking order of the available RATs.

Some of the selection criteria used in determining a user's preference for a particular RAT (e.g., data rate) are specified using real numbers whereas other criteria (e.g., security level) are specified using linguistic values such as very high, high, medium, low, very low. The linguistic terms are first converted to fuzzy numbers using a conversion scale. Then the

result fuzzy numbers are converted to crisp numbers. For instance, if five linguistic terms are used to represent the possible user preference: very low, low, medium, high and very high, these linguistic terms are first converted to fuzzy numbers using the conversion scale shown in Fig. 6-5, where both the performance score x and membership function are in the range from 0 to 1. A fuzzy scoring method is used to convert each fuzzy number to a corresponding crisp value. For example, the five fuzzy numbers shown in Fig. 6-5 are converted to 0.091, 0.283, 0.5, 0.717, and 0.909 respectively [20, 94]. Chen *et al* have proposed eight different conversion scales with different number of linguistic terms [94].

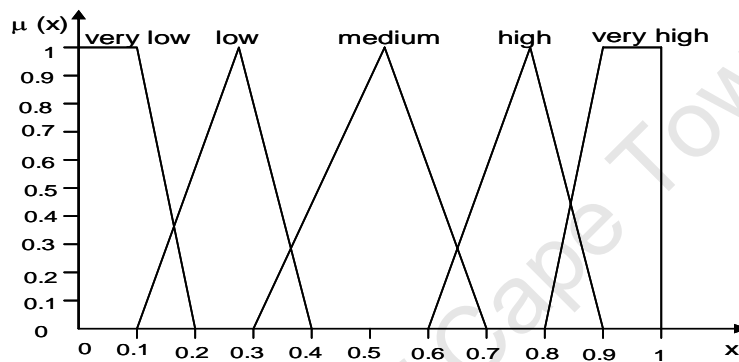


Figure 6-5. Linguistic terms to fuzzy number conversion scale.

In MADM problems, decision makers often need to select or rank alternatives that are associated with non-commensurate and conflicting attributes. The decision makers' preference information is often used to rank alternatives or select the most desirable one. There are many classical MADM methods such as SAW (Simple Additive Weighting), TOPSIS (Technique for Order Preference by Similarity to Ideal Solution), AHP (Analytical Hierarchical Process), etc. The SAW method is the most widely used method [95]. Therefore, the SAW method is used in this research.

The JCAC MADM problem involves a set of J alternative RAT- j ($j=1,2,\dots,J$). These alternative RATs are to be evaluated for each arriving call with respect to a set of N -criteria (or attributes), which are independent of each other. A decision matrix, D for J alternative RATs and N criteria is given as:

$$D = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,N} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ c_{J,1} & c_{J,2} & \cdots & c_{J,N} \end{bmatrix} \quad (6.1)$$

where $c_{j,n}$ represents the performance rating of RAT- j ($j=1,2,\dots,J$) on criterion- n ($n=1,2,\dots,N$).

The SAW method requires the normalization of the decision matrix D . Usually, there are benefit criteria and cost criteria in MADM problems [96, 95], and the ‘dimension’ of the criteria may be different. In order to measure all criteria in dimensionless units and to facilitate their comparison, normalization is necessary. There are different ways of normalizing the values of the criteria. In this work, each normalized value $b_{j,n}$ of the normalized decision matrix (\bar{D}) is calculated as follows [95].

$$b_{j,n} = \begin{cases} \frac{c_{j,n}}{\max \{c_{j,n} \mid j=1, 2, \dots, J\}}, & \text{for benefit criterion } C_n \\ \frac{\min \{c_{ij} \mid j=1, 2, \dots, J\}}{c_{j,n}}, & \text{for cost criterion } C_n \end{cases} \quad (6.2)$$

where $b_{j,n}$ is the normalized performance rating of RAT- j on criterion- n , $j = 1, 2, \dots, J$; $n = 1, 2, \dots, N$.

Each user would give his/her preference information for a particular RAT for each class of calls. This preference information is specified in terms of weight attached to each network criterion by the user. The weighing vector, W_c^i represents the relative importance of the criteria to user- c of class- i call, and it is given as:

$$W_c^i = (w_1 \quad w_2 \quad \cdots \quad w_N) \quad (6-3)$$

The normalized weighing vector, \bar{W}_c^i is obtained as follows:

$$\bar{W}_c^i = (\bar{w}_1 \quad \bar{w}_2 \quad \dots \quad \bar{w}_N)$$

$$\bar{w}_i = \frac{w_i}{\sum_{i=1}^N w_i} \quad i = 1, 2, \dots, N \quad (6.4)$$

The weight shows the relative importance of each criterion to users. This research uses a 10-point scale $\{0,1,2,3,4,5,6,7,8,9\}$ for weight assignments. For example, if a criterion C_1 has weight 1 and criterion C_2 has weight 5, then C_2 is considered to be five times more important to the user than C_1 . If a criterion is assigned weight 0 by a user, the criterion is not important to the user, and therefore will have no effect in making RAT selection decision for the user.

The normalized decision matrix (\bar{D}) and the normalized weighing vector (\bar{W}_c^i) are used to rank the alternatives RATs for each arriving class- i call. Using SAW, the preference rating ($v_{i,j}^c$) of user $-c$ of class- i calls for RAT- j is obtained by:

$$V_{i,j}^c = \sum_{n=1}^N \bar{w}_n b_{j,n}, \quad \text{for } j = 1, 2, \dots, J \quad (6.5)$$

Selection of the most appropriate RAT for each arriving call can be based on strict preference or flexible preference. In strict preference, the JCAC algorithm selects the highest ranking RAT for the incoming class- i call. If the highest ranking RAT cannot accommodate the call due to unavailability of radio resources, the call is blocked.

In flexible preference, the call is admitted into the highest ranking RAT which has enough radio resources to accommodate it. This implies that if the highest ranking RAT cannot accommodate the call due to unavailability of radio resources, the call is admitted into the next highest ranking RAT, and so on. The call is only blocked if none of the available RATs has enough bbu to accommodate the call.

6.4.1.2 Illustration of a User-Centric JCAC Algorithm

The fuzzy MADM procedure is illustrated in the following example using single-class real-time calls, a three-RAT heterogeneous wireless network, and four RAT-selection

criteria. The criteria are service price (C_1), data rate (C_2), security (C_3), power consumption (C_4). The three RATs are assumed to have equal capacity. The objective of the JCAC algorithm is to select the most appropriate RAT for each call based on individual user's preferences. Class-1 calls arrive according to Poisson process, and the weighing vector w_c^1 is randomly generated for each user- c of class-1 call. A simulation of 1000 calls is carried out. Table 6-1 shows the RAT selection criteria for incoming calls. As shown in Table 8, price and battery power consumption are cost criteria whereas data rate and security are benefit criteria. Each RAT can support the incoming calls at different data rate. Calls admitted into RAT-1, RAT-2, and RAT-3 are allocated 32 Kbps (1 bbu), 64 Kbps (2 bbu), and 128 Kbps (4 bbu) respectively. For the real-time calls, service price is given in cents per bbu per minute.

Table 6-1. RAT selection criteria for incoming calls.

RATs	Criteria			
	Initial Price	Data rate (Kbps)	security	Battery power consumption
RAT-1	0.5	32	medium	low
RAT-2	0.5	64	very high	high
RAT-3	0.5	128	high	medium

From Table 6-1, the decision matrix D is obtained as:

$$D = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT - 1 \\ RAT - 2 \\ RAT - 3 \end{matrix} & \begin{bmatrix} 0.5 & 32 & \text{medium} & \text{low} \\ 0.5 & 64 & \text{very high} & \text{high} \\ 0.5 & 128 & \text{high} & \text{medium} \end{bmatrix} \end{matrix}$$

The linguistic terms used in the decision matrix (D) are converted to crisp numbers using the conversion scale shown Fig. 6-5 [20, 94]. The following is the resulting matrix after converting the linguistic terms into crisp values.

$$D = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT-1 \\ RAT-2 \\ RAT-3 \end{matrix} & \begin{bmatrix} 0.5 & 32 & 0.500 & 0.283 \\ 0.5 & 64 & 0.910 & 0.717 \\ 0.5 & 128 & 0.717 & 0.500 \end{bmatrix} \end{matrix}$$

The following is the normalized decision matrix using (6.2).

$$\bar{D} = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT-1 \\ RAT-2 \\ RAT-3 \end{matrix} & \begin{bmatrix} 1.000 & 0.250 & 0.549 & 1.000 \\ 1.000 & 0.500 & 1.000 & 0.395 \\ 1.000 & 1.000 & 0.788 & 0.566 \end{bmatrix} \end{matrix}$$

Given that the weighing vector assigned to the first call by user-1 is:

$$W_1^1 = (5 \quad 9 \quad 3 \quad 1)$$

The following is the normalized weighing vector using (6.4):

$$\bar{W}_1^1 = (0.2778 \quad 0.5000 \quad 0.1667 \quad 0.0556)$$

Using (6.5), the ranking values, $v_{1,j}^1$ of RAT-1, RAT-2, and RAT-3 are obtained respectively as 0.5499, 0.7164, and 0.9405. Applying strict user's preference for RAT selection, RAT-3 is selected for the incoming class-1 call because $v_{1,3}^1 > v_{1,2}^1 > v_{1,1}^1$.

For 1000 class-1 calls with randomly generated weighing vectors, Fig. 6-6 shows the proportion of calls that are admitted into RAT-1, RAT-2 and RAT-3. This proportion also indicates the mean admission rate of the calls into each of the available RATs.

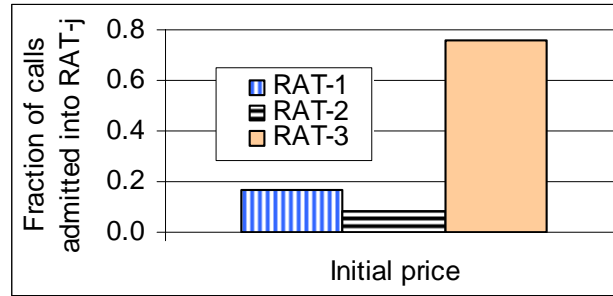


Figure 6-6. Proportion of calls admitted into each RAT.

Due to uneven distribution of the traffic load among the available RATs (Fig. 6-6), if a strict preference is used for RAT selection, the overall call blocking/dropping probability in the heterogeneous network will be high resulting in poor connection-level QoS. The reason is because RAT-3 will soon be fully loaded, and many calls for which RAT-3 has been selected will subsequently be blocked. On the other hand, if flexible preference is used, it will result in low users' satisfaction because many calls will not be admitted into the RAT which has the highest ranking value ($v_{i,j}^c$) with respect to their individual weighing vectors. It means then that users are forced to have their calls admitted into a RAT other than the one they mostly prefer, which contradicts the primary objective of user-centricity.

These are the major drawbacks of user-centric JCAC algorithms. These drawbacks can be overcome by giving users incentives to act in such a way that the traffic load in the heterogeneous wireless network is evenly distributed among the available RATs. The initial service price in each RAT is replaced with new prices shown in Table 6-2. Fig. 6-7 shows the proportion of calls admitted into each RAT using the initial and new prices.

Table 6-2. Initial and new prices used in the decision matrix.

RATs	Initial price	New price 1	New price 2	New price 3
RAT-1	0.5	0.4	0.4	0.4
RAT-2	0.5	0.5	0.4	0.4
RAT-3	0.5	0.6	0.7	0.8

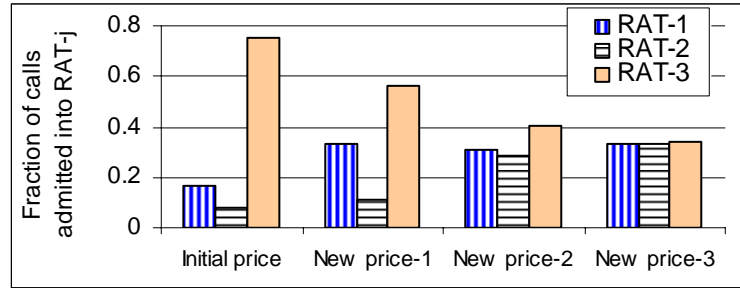


Figure 6-7. Proportion of calls admitted into each RAT.

From Fig. 6-7, it is observed that as the service price in each RAT varies, the proportion of calls admitted into each RAT varies. Fig 6-7 illustrates how pricing policy in a heterogeneous wireless network can influence user behavior, with respect to the ratio of calls admitted into each RAT. This variation of service price with respect to call admission rate in each RAT is the basic principle of the JCAC scheme proposed in this chapter. Fig 6-8 shows the proposed RAT selection method based on fuzzy MADM technique, where service price is necessarily one of the RAT selection criteria, and $C_{j,n}$ represents the rating of criterion- n (C_n) in RAT- j .

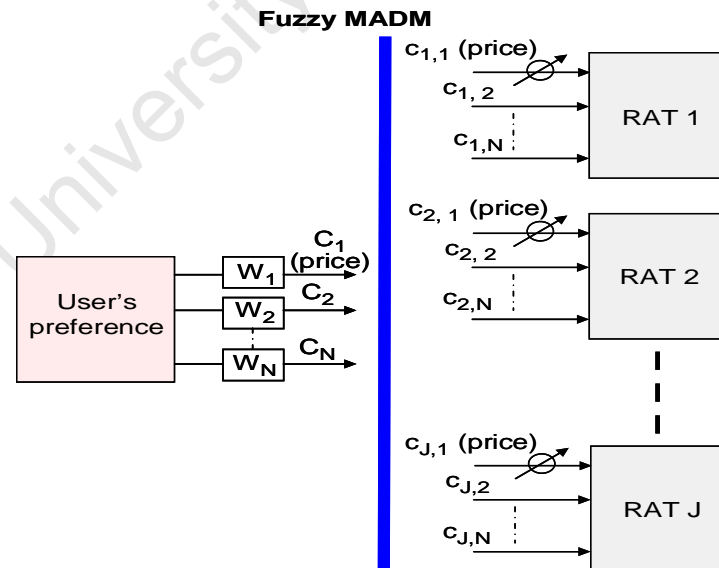


Figure 6-8. Proposed Fuzzy MADM-based RAT selection method.

As shown in Fig. 6-8, the service price is dynamically adjusted with respect to call admission rate in each RAT in order to balance the traffic load among available RATs in the heterogeneous wireless network.

6.4.2 Call Admission Rate Measurement Unit

The call admission rate measurement unit (CARMU) measures the arrival rates of different classes of calls in each RAT as shown in Fig. 6-2. This measurement is done periodically. The CARMU also measures the residual capacity (bbu) available for each class of calls in each RAT. The measured values of call arrival rates and residual bbu are used by the price update unit to calculate the new price that will be used during the next period of time.

6.4.3 Price Update Unit

The price update unit (PUU) periodically adjusts the service price in each RAT in the heterogeneous network. From the CARMU, the PUU obtains information about the residual capacity available for each class of call, and the mean call admission rate in each of the RATs. Based on the residual capacity available for class- i calls in RAT- j , the PUU determines the ideal admission rate into each RAT that is required for load balancing. It then finds the difference between the ideal admission rate and the measured admission rate in each RAT. The difference between the ideal admission rate and the measured admission rate is used to determine the new price for each class- i calls in RAT- j during the next period, T . Note that once a call is admitted into a RAT at a particular service price, the price will be used during the entire call residence time.

Fig. 6-9 shows the price update process. The time between two consecutive price updates is referred to as the Fixed Price Period (FPP). All calls admitted into a particular RAT- j during the FPP will be charged using the same service price. The FPP is chosen to be relatively long compared to a typical call holding time. The price update takes place at the end of each FPP as shown in Fig. 6-9.

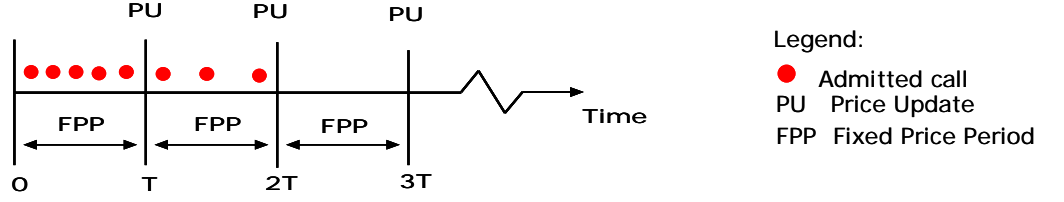


Figure 6-9. Price update process.

For new class- i calls, let $x_{i,j}^n$ denote the residual bbu available in RAT- j , $\alpha_{i,j}$ denotes the fraction of the residual bbu available in RAT- j over the summation of the residual bbu available in all RATs, $r_{i,j}$ denotes the price per unit bbu per unit time in RAT- j , $\lambda_{i,j}^n$ denotes the measured call admission rate in RAT- j , and $\bar{\lambda}_{i,j}^n$ denotes the calculated ideal admission rate in RAT- j . For handoff class- i calls, the corresponding values are $y_{i,j}$, $\beta_{i,j}$, $r_{i,j}$, $\lambda_{i,j}^h$, and $\bar{\lambda}_{i,j}^h$. Then:

$$\alpha_{i,j} = \frac{x_{i,j}^n}{\sum_{j=1}^J x_{i,j}^n} \quad \forall i, j \quad (6.6)$$

$$\sum_{j=1}^J \alpha_{i,j} = 1 \quad \forall i \quad (6.7)$$

Similarly,
$$\beta_{i,j} = \frac{y_{i,j}^h}{\sum_{j=1}^J y_{i,j}^h} \quad \forall i, j \quad (6.8)$$

$$\sum_{j=1}^J \beta_{i,j} = 1 \quad \forall i \quad (6.9)$$

When a new or handoff call arrives into a group of co-located cells, the JCAC algorithm selects the most suited RAT for the call. The action of selecting a RAT for each arriving new or handoff call in the group of co-located cells leads to splitting of the arrival process. Fig. 6-10 illustrates the splitting of the arrival among J number of RATs in the group of co-located cells. Note that the arrival rates of a split poisson process are also Poisson [79].

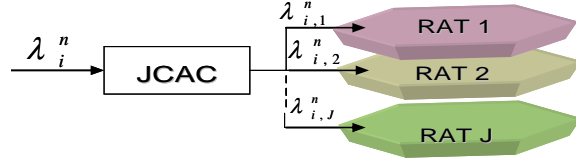


Figure 6-10. Splitting of the arrival process.

For uniform load distribution during any period, T , the ideal arrival rates for new and handoff class- i calls in RAT- j are obtained as follows:

$$\bar{\lambda}_{i,j}^n = \frac{\alpha_{i,j} \cdot \lambda_i^n}{b_{i,j}} \quad \forall i, j \quad (6.10)$$

$$\bar{\lambda}_{i,j}^h = \frac{\beta_{i,j} \cdot \lambda_{i,j}^h}{b_{i,j}} \quad \forall i, j \quad (6.11)$$

where $b_{i,j}$ is the bbu allocated class- i calls in RAT- j and λ_i is the arrival rate of class- i calls into the group of co-located cells. Note that the JCAC algorithm always allocate for each admitted class- i call, the maximum bbu available for class- i calls in the RAT where it is admitted. The difference between the ideal mean admission rate and the measured mean admission rate is obtained as follows:

$$\partial \lambda_{i,j} = (\lambda_{i,j}) - (\bar{\lambda}_{i,j}) \quad (6.12)$$

where $\lambda_{i,j} = (\lambda_{i,j}^n + \lambda_{i,j}^h)$ and $\bar{\lambda}_{i,j} = (\bar{\lambda}_{i,j}^n + \bar{\lambda}_{i,j}^h)$

Two price-update functions are considered for the proposed JCAC scheme. They are linear and exponential price-update functions. In the linear price update function, new price is calculated from the old price as follows:

$$\varphi_{j,new}^i = \varphi_{j,old}^i + \delta \varphi_j^i \quad (6.13)$$

In exponential price-update function, new price is calculated from the old price as follows:

$$\begin{aligned} \varphi_{j,new}^i &= \varphi_{j,old}^i e^{\delta \varphi_j^i} \\ \varphi_{j,min}^i &\leq \varphi_{j,new}^i, \varphi_{j,old}^i \leq \delta \varphi_{j,max}^i \end{aligned} \quad (6.14)$$

where $\varphi_{j,new}^i$ is the new price per bbu per unit time for class- i calls in RAT- j , $\varphi_{j,old}^i$ is the current price per bbu per unit time for class- i calls in RAT- j , and $\delta\varphi_j^i$ is the change in price per bbu per unit time in RAT- j . $\varphi_{j,min}^i$ and $\varphi_{j,max}^i$ are respectively, the minimum and maximum price per bbu per unit time in RAT- j . The change in price ($\delta\varphi_j^i$) is calculated using (10) as follows:

$$\delta\varphi_j^i = z \left(\frac{\lambda_{i,j} - \bar{\lambda}_{i,j}}{\bar{\lambda}_{i,j}} \right) \quad (6.15)$$

where z is a positive scalar step size. There are several different types of step size rules that can be used. Examples are as constant step size, diminishing step size, increasing step size. Step size can also be dynamically chosen based on the value of $\lambda_{i,j}^n - \bar{\lambda}_{i,j}^n$. For simplicity, a constant step size, which has been used for homogeneous wireless network in [97], is used in this research. Note that the value of $\delta\varphi_j^i$ is positive if $\lambda_{i,j} > \bar{\lambda}_{i,j}$ and negative if $\lambda_{i,j} < \bar{\lambda}_{i,j}$.

In equation 6.14, the operator(s) determine the minimum and maximum service prices in each RAT. This flexibility ensures that the profit of the operator(s) does not fall below the expected amount.

6.4.4 Bandwidth Reservation Unit

In order to maintain lower handoff dropping probability, the proposed scheme reserves certain bandwidth exclusively for handoff calls in all the cells of each group of co-located cells. Fig. 6-11 shows the bandwidth reservation policy for the heterogeneous network.

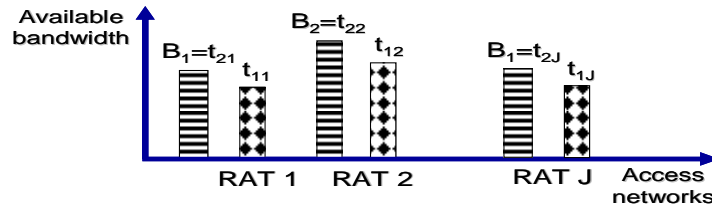


Figure 6-11. Bandwidth reservation policy.

The policy reserves bandwidth for aggregate handoff calls, thus giving them priority over new calls such that:

$$t_{1,j} \leq t_{2,j} = B_j \quad \forall j \quad (6.16)$$

where $t_{1,j}$ denotes the threshold for rejecting new calls in RAT- j , and $t_{2,j}$ denotes the threshold for rejecting handoff calls in RAT- j . B_j denotes the total number of bbu available in RAT- j .

6.5 Markov Model

The JCAC policy described in section IV can be modeled as a multi-dimensional Markov chain. The state space of the group of co-located cells can be represented by a $(2*K*J)$ -dimensional vector given as:

$$\Omega = (m_{i,j}, n_{i,j} : i=1, \dots, k, \quad j=1, \dots, J) \quad (6.17)$$

The non-negative integer $m_{i,j}$ denotes the number of ongoing new class- i calls in RAT- j , and the non-negative integer $n_{i,j}$ denotes the number of ongoing handoff class- i calls in RAT- j . Let S denote the state space of all admissible states of the group of co-located cells as it evolves over time. An admissible state s is a combination of the numbers of users in each class that can be supported simultaneously in the group of co-located cells while maintaining adequate QoS and meeting resource constraints.

The state S of all admissible states in the group of co-located cells is given as:

$$S = \{ \Omega = (m_{i,j}, n_{i,j} : i = 1, \dots, k, \quad j = 1, \dots, J) : \sum_{i=1}^K m_{i,j} \cdot b_{i,j} \leq t_{1,j}^n \quad \forall j \wedge \sum_{i=1}^k m_{i,j} \cdot b_{i,j} + \sum_{i=1}^k n_{i,j} b_{i,j} \leq B_j \quad \forall j \} \quad (6.18)$$

where $b_{i,j}$ is the bandwidth allocated to class- i calls in RAT- j . The constraints simply state that during any period, T , the sum of the bandwidth units of all admitted class- i calls in each RAT- j cannot be more than the total bandwidth units available for that class of

calls. Given that the system is in the current state, s , for the JCAC scheme, the state transition could be triggered by any of the following events:

1) Admission of a new class- i call into RAT- j with the successor state s_1^{+1} and transition rate $q(s, s_1^{+1})$. It follows that

$$q(s, s_1^{+1}) = \lambda_{i,j}^n, \quad (s, s_1^{+1} \in S) \quad (6.19)$$

2) Admission of a handoff class- i call into RAT- j with the successor state s_2^{+1} and transition rate $q(s, s_2^{+1})$. It follows that

$$q(s, s_2^{+1}) = \lambda_{i,j}^h, \quad (s, s_2^{+1} \in S) \quad (6.20)$$

3) Departure of a new class- i call from RAT- j with the successor state s_1^{-1} and transition rate $q(s, s_1^{-1})$. It follows that

$$q(s, s_1^{-1}) = m_{i,j} \mu_i^n, \quad (s, s_1^{-1} \in S) \quad (6.21)$$

4) Departure of a handoff class- i call from RAT- j with the successor state s_2^{-1} and transition rate $q(s, s_2^{-1})$. It follows that

$$q(s, s_2^{-1}) = n_{i,j} \mu_i^h, \quad (s, s_2^{-1} \in S) \quad (6.22)$$

where $s, s_1^{+1}, s_1^{-1}, s_2^{+1}$, and s_2^{-1} are the following matrices.

$$s = \begin{bmatrix} m_{11} \cdots & m_{1j} \cdots & m_{1J} \\ \vdots & \vdots & \vdots \\ m_{i1} \cdots & m_{ij} \cdots & m_{iJ} \\ \vdots & \vdots & \vdots \\ m_{K1} \cdots & m_{Kj} \cdots & m_{KJ} \\ n_{11} \cdots & n_{1j} \cdots & n_{1J} \\ \vdots & \vdots & \vdots \\ n_{i1} \cdots & n_{ij} \cdots & n_{iJ} \\ \vdots & \vdots & \vdots \\ n_{K1} \cdots & n_{Kj} \cdots & n_{KJ} \end{bmatrix}, \quad s_1^{\pm 1} = \begin{bmatrix} m_{11} \cdots & m_{1j} \cdots & m_{1J} \\ \vdots & \vdots & \vdots \\ m_{i1} \cdots & m_{ij} \pm 1 \cdots & m_{iJ} \\ \vdots & \vdots & \vdots \\ m_{K1} \cdots & m_{Kj} \cdots & m_{KJ} \\ n_{11} \cdots & n_{1j} \cdots & n_{1J} \\ \vdots & \vdots & \vdots \\ n_{i1} \cdots & n_{ij} \cdots & n_{iJ} \\ \vdots & \vdots & \vdots \\ n_{K1} \cdots & n_{Kj} \cdots & n_{KJ} \end{bmatrix}, \quad s_2^{\pm 1} = \begin{bmatrix} m_{11} \cdots & m_{1j} \cdots & m_{1J} \\ \vdots & \vdots & \vdots \\ m_{i1} \cdots & m_{ij} \cdots & m_{iJ} \\ \vdots & \vdots & \vdots \\ m_{K1} \cdots & m_{Kj} \cdots & m_{KJ} \\ n_{11} \cdots & n_{1j} \cdots & n_{1J} \\ \vdots & \vdots & \vdots \\ n_{i1} \cdots & n_{ij} \pm 1 \cdots & n_{iJ} \\ \vdots & \vdots & \vdots \\ n_{K1} \cdots & n_{Kj} \cdots & n_{KJ} \end{bmatrix} \quad (6.23)$$

Joint call admission decisions are taken in the arrival epoch. Every time a new or handoff class- i call arrives in the group of co-located cells, the JCAC algorithm decides whether or not to admit the call, and in which RAT to admit it. Note that call admission decision is made only at the arrival of a call, and no call admission decision is made in the group of co-located cells when a call departs. When the system is in state s , an accept/reject

decision must be made for each type of possible arrival, i.e., an arrival of a new class- i call, or the arrival of a handoff class- i call in the group of co-located cells. The following are the possible JCAC decisions in the arrival epoch.

1) Reject the class- i call (new or handoff) in the group of collocated cells, in which case the state s does not evolve.

2) Admit the class- i call into RAT- j in which case the state s evolves.

Thus, the call admission action space A can be expressed as follows:

$$A = \{a = (a_1^n, \dots, a_k^n, a_1^h, \dots, a_k^h) : \quad (6.24)$$

$$a_i^n, a_i^h \in (0, 1, \dots, j, (j+1), \dots, J), i = 1, \dots, k\}$$

where a_i^n denotes the action taken on arrival of a new class- i call within the group of co-located cells, and a_i^h denotes the action taken on arrival of a handoff class- i call from an adjacent group of co-located cells. a_i^n (or a_i^h) = 0 means reject the new class- i (or handoff class- i) call. a_i^n (or a_i^h) = 1 means accept the new class- i (or handoff class- i) call into RAT-1. a_i^n (or a_i^h) = j means accept the new class- i (or handoff class- i) call into RAT- j .

Based on its Markovian property, the proposed JCAC scheme can be model as a $(2 \cdot K \cdot J)$ -dimensional Markov chain. During period T , Let $\rho_{new\ i,j}$ and $\rho_{han\ i,j}$ denote the load generated by new class- i calls and handoff class- i calls, respectively, in RAT- j . Then,

$$\rho_{new\ i,j} = \frac{\lambda_{i,j}^n}{\mu_i^n} \quad \forall i, j, \quad (6.25)$$

$$\rho_{han\ i,j} = \frac{\lambda_{i,j}^h}{\mu_i^h} \quad \forall i, j \quad (6.26)$$

From the steady state solution of the Markov model, performance measures of interest can be determined by summing up appropriate state probabilities. Let $P(s)$ denote the steady state probability that the group of collocated cells is in state s ($s \in S$). From the detailed balance equation, $P(s)$ is obtained as:

$$P(s) = \frac{1}{G} \prod_{j=1}^J \prod_{i=1}^k \frac{(\rho_{new\ i,j})^{m_{i,j}} (\rho_{han\ i,j})^{n_{i,j}}}{m_{i,j}! n_{i,j}!} \quad \forall s \in S \quad (6.27)$$

where G is a normalization constant given by:

$$G = \sum_{s \in S} \prod_{j=1}^J \prod_{i=1}^k \frac{(\rho_{new\ i,j})^{m_{i,j}}}{m_{i,j}!} \frac{(\rho_{han\ i,j})^{n_{i,j}}}{n_{i,j}!} \quad (6.28)$$

6.5.1 New Call Blocking Probability

Using strict preference for RAT selection, a new class- i call is blocked in the group of co-located cells if the highest ranked (selected) RAT- j does not have enough bbu to accommodate the new call. Let $S_{bi}^j \subset S$ denote the set of states in which a new class- i call is blocked in RAT- j . It follows that:

$$S_{bi}^j = \{s \in S : ((b_{i,j} + \sum_i^k m_{i,j} b_{i,j} > t_{i,j}^n \vee b_{i,j} + \sum_{i=1}^k m_{i,j} b_{i,j} + \sum_{i=1}^k n_{i,j} b_{i,j} > B_j))\} \quad (6.29)$$

Thus the new call blocking probability (NCBP), P_{bi}^j , for a class- i call in RAT- j is given by:

$$P_{bi}^j = \sum_{s \in S_{bi}^j} P(s) \quad (6.30)$$

During period, T , the overall new class- i call blocking probability in the group of co-located cells is given as:

$$P_{bi} = \sum_{j=1}^J \frac{\lambda_{i,j}}{\sum_{j=1}^J \lambda_{i,j}} \cdot P_{bi}^j \quad (6.31)$$

6.5.2 Handoff Call Dropping Probability

A handoff class- i call is dropped in the group of co-located cells if the highest-ranked (selected) RAT- j does not have enough bbu to accommodate the handoff call. Let $S_{di}^j \subset S$ denote the set of states in which a handoff class- i call is dropped in RAT- j in the group of co-located cells. It follows that:

$$S_{di}^j = \{s \in S : b_{i,j} + \sum_{i=1}^k (m_{i,j} + n_{i,j}) b_{i,j} > B_j\} \quad (6.32)$$

Thus the handoff call dropping probability (HCDP) for a class- i call, P_{di} , in the group of

co-located cells is given by:

$$P_{d_i}^j = \sum_{s \in S_{d_i}^j} P(s) \quad (6.33)$$

During period T , the overall handoff class- i call blocking probability is given as:

$$P_{d_i} = \sum_{j=1}^J \frac{\lambda_{i,j}}{\sum_{j=1}^J \lambda_{i,j}} \cdot P_{d_i}^j \quad (6.34)$$

6.5.3 Percentage Utilization of Each RAT

During period T , the average utilization of each RAT- j in the group of co-located cells can be obtained by summing up for all the admissible state s ($s \in S$) in the group of co-located cells, the product of the system utilization in RAT- j in a particular state s ($s \in S$) and the probability $P(s)$ of the system being in that state. The average utilization of RAT- j , U^j in the group of co-located cells can be derived as follows:

$$U^j = \sum_{s \in S} P(s) \sum_{i=1}^k b_{i,j}(m_{i,j} + n_{i,j}) \quad \forall j \quad (6.35)$$

Percentage utilization of each RAT in the heterogeneous wireless network is obtained as:

$$\bar{U}^j = \frac{U^j}{\sum_{j=1}^J U^j} \times 100\% \quad \forall j \quad (6.36)$$

6.6 Performance Evaluation

In this section, the performance of the proposed user-centric JCAC scheme is evaluated through simulation, and compared with the performance of a JCAC scheme that does not incorporate dynamic pricing such as previously proposed in [20]. The performance of the proposed scheme is evaluated using a three-RAT heterogeneous wireless network with subscribers having a single class real-time calls that can be admitted with three different bandwidth levels. The calls are allocated 1bu (16 Kbps), 2 bu (32 Kbps), and 4 bu (64 Kbps) when admitted into RAT-1, RAT-2, and RAT-3 respectively. Note that for adaptive real-time services, higher data rate enhances the resolution and the quality of the received service. Four RAT selection criteria are considered as shown in Table 6-3. The price update module sets the initial service price in each RAT to 0.5 cent per bbu per minute.

The prices are periodically updated by the JCAC scheme. For each arriving call, weights are randomly assigned to the RAT selection criteria. Note that the weights indicate the user's preference for a RAT. The arrival rate of handoff class- i calls in the group of co-located cells is assumed to be proportional to the arrival rate of new class- i calls by: $\lambda_i^h = (h/\mu_i) \lambda_i^n$ where h is the handoff rate.

Table 6-3. RAT selection criteria.

Criteria	Initial Price (C_1)	Data rate (C_2)	Security (C_3)	Battery power consumption (C_4)
RAT-1	0.5	1(i.e. 16 Kbps)	medium	low
RAT-2	0.5	2(i.e. 32 Kbps)	very high	high
RAT-3	0.5	4(i.e. 64 Kbps)	high	medium

Other parameters used in the simulations are, $z = \{0.1, 0.2, 0.3\}$, $\lambda^n = \{1, 3\}$, $t_{1,1} = t_{1,2} = t_{1,3} = 10$. $B_1 = B_2 = B_3 = 20$, $\mu_1 = 0.5$, $h = 0.5$. From Table 6-3, the decision matrix D is obtained as:

$$D = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT - 1 \\ RAT - 2 \\ RAT - 3 \end{matrix} & \begin{bmatrix} 0.5 & 1 & \text{medium} & \text{low} \\ 0.5 & 2 & \text{very high} & \text{high} \\ 0.5 & 4 & \text{high} & \text{medium} \end{bmatrix} \end{matrix}$$

The linguistic terms used in the decision matrix, D , are converted to crisp numbers using the conversion scale shown Fig. 6-5 [20, 94]. The following is the resulting matrix after converting the linguistic terms into crisp values.

$$D = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT - 1 \\ RAT - 2 \\ RAT - 3 \end{matrix} & \begin{bmatrix} 0.5 & 1 & 0.500 & 0.283 \\ 0.5 & 2 & 0.910 & 0.717 \\ 0.5 & 4 & 0.717 & 0.500 \end{bmatrix} \end{matrix}$$

The following matrix \bar{D} is the normalized decision matrix using (6.2).

$$\bar{D} = \begin{matrix} & C_1 & C_2 & C_3 & C_4 \\ \begin{matrix} RAT - 1 \\ RAT - 2 \\ RAT - 3 \end{matrix} & \begin{bmatrix} 1.000 & 0.250 & 0.549 & 1.000 \\ 1.000 & 0.500 & 1.000 & 0.395 \\ 1.000 & 1.000 & 0.788 & 0.566 \end{bmatrix} \end{matrix}$$

Based on the normalized weighing vector (\bar{w}_c^1) for each incoming call and the normalized decision matrix \bar{D} , the JCAC scheme determines the most suitable RAT for each incoming call using (6.5).

The research considers three user scenarios namely: (1) general user scenario, (2) high-price-sensitive user scenario, and (3) low-price-sensitive user scenario.

6.6.1 General User Scenario

In general-user scenario, weights assigned by users to each of the RAT selection criterion are randomly chosen from 0 to 9. Fig. 6-12 shows the performance of the proposed JCAC scheme with respect to call blocking/dropping probability.

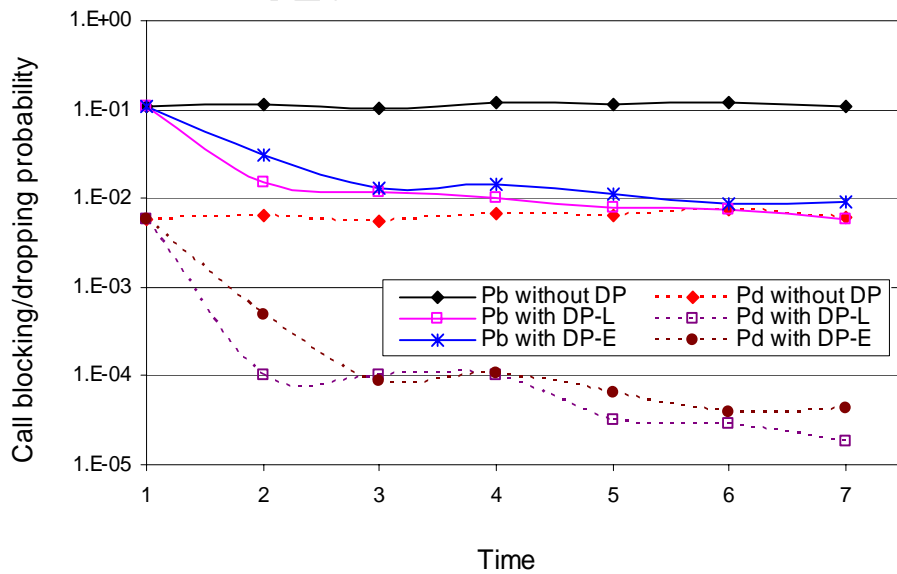


Figure 6-12. Call blocking/dropping probability versus price update time ($\lambda^n = 1, z = 0.1$).

As shown in Fig. 6-12, *Pb without DP* denotes the new call blocking probability of the JCAC scheme without dynamic pricing. *Pd without DP* denotes the handoff call dropping probability of the JCAC scheme without dynamic pricing. *Pb with DP-L* denotes the new call blocking probability of the JCAC scheme with dynamic pricing using linear price-update function. *Pb with DP-E* denotes the new call blocking probability of the JCAC scheme with dynamic pricing using exponential price-update function. *Pd with DP-L* denotes the handoff call dropping probability of the JCAC scheme with dynamic pricing using linear-price update function. *Pd with DP-E* denotes the handoff call dropping probability of the JCAC scheme with dynamic pricing using exponential price-update function.

It can be seen that for each of the JCAC schemes, Pd is always lower than the corresponding Pb. This shows that handoff calls are prioritized over new calls by using different call rejection thresholds (Fig. 6-11) for new and handoff calls. Note that lower Pd implies better connection-level QoS. It can also be seen that both *Pb with DP-L* and *Pb with DP-E* are lower than *Pb without DP*. The reason is that *JCAC scheme with dynamic pricing (DP)* is able to adjust the service price in each RAT with time in order to even out the unbalance traffic among the available RAT. As a result, JCAC scheme with DP (using either linear or exponential price function) reduces the new blocking probability in the heterogeneous wireless network. For the same reason, both *Pd with DP-L* and *Pd with DP-E* are also lower than *Pd without DP*. It is also observed that for the proposed JCAC scheme with DP, using either linear or exponential price update function achieves similar results with respect to call blocking/dropping probability.

Fig. 6-13 shows the performance of the proposed JCAC scheme with respect to call blocking/dropping probability. All the parameters used are the same as those used in Fig. 6-12 except for the new call arrival rate, which is increased from 1 to 3. In Fig 6-13, the same trend is observed for all the Pb and Pd as in Fig. 6-12. However, the values of Pb are higher than the corresponding values of Pb in Fig. 6-12. This increment in values of Pb is expected because call blocking probability generally increases with increase in call arrival rate. Similarly, values of Pd are higher than the corresponding values of Pd in Fig. 6-12.

The increment in values of Pd is also expected because call dropping probability generally increases with increase in call arrival rate.

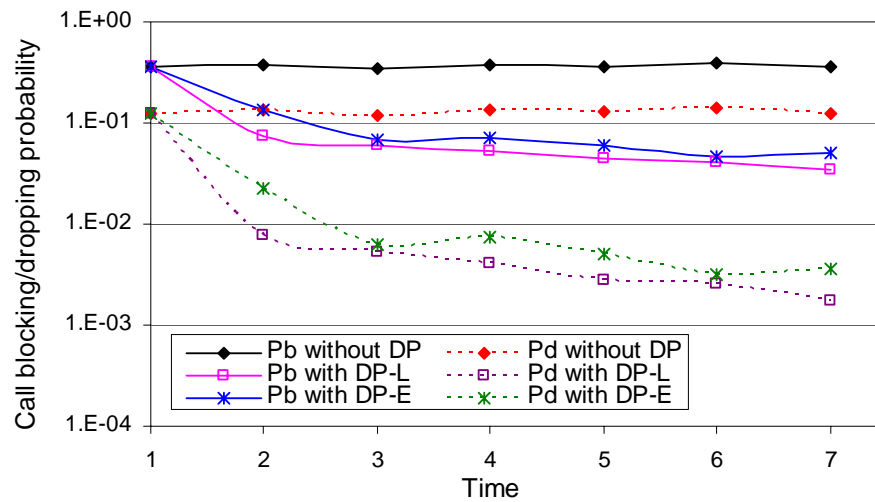


Figure 6-13. Call blocking/dropping probability versus price update time ($\lambda^n = 3, z=0.1$).

Fig. 6-14 and Fig 6-15 show the effect of varying the step size, z on call blocking/dropping probability of the JCAC scheme with dynamic pricing. The parameters used are the same as those use in Fig 6-13 except for the value of z which is increased from 0.1 to 0.2 (Fig. 6-14) and from 0.1 to 0.3 (Fig. 6-15). As show in Fig. 6-13, Fig. 6-14 and Fig. 6-15, Pb and Pd of the JCAC scheme with DP (using either linear or exponential price update function) improve (reduce) as the value of z increases. Higher values of z enable the proposed JCAC scheme to be more reactive to the unbalanced traffic load among the available RATs.

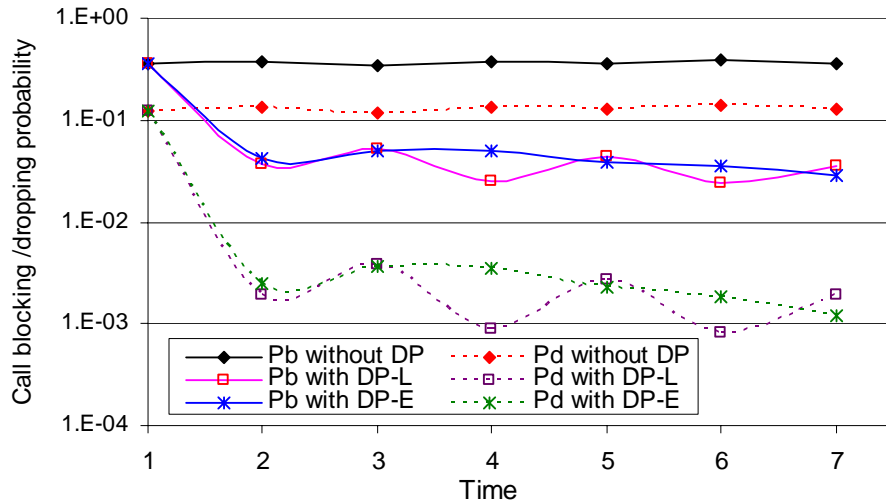


Figure 6-14. Call blocking/dropping probability versus price update time ($\lambda^n=3$, $z=0.2$).

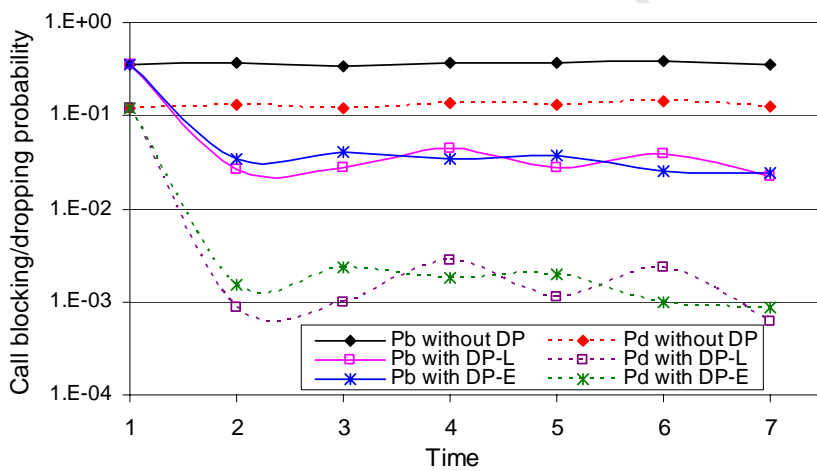


Figure 6-15. Call blocking/dropping probability versus price update time ($\lambda^n=3$, $z=0.3$).

Fig. 6-16 shows the percentage of load in each RAT for the JCAC scheme without dynamic pricing. As shown in Fig. 6-16, for a given call arrival rate, it is observed that the traffic load is highly unbalanced, and there is only a slight variation with time in the percentage of load among available RATs.

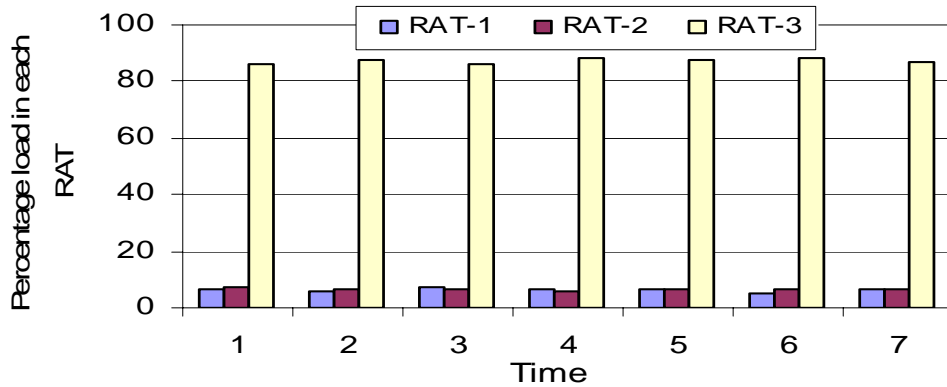


Figure 6-16. Percentage of load in each RAT versus time for JCAC without dynamic pricing ($\lambda^n = 3$).

Fig. 6-17 and Fig. 6-18 show the percentage of load in each RAT for the JCAC scheme with dynamic pricing using linear and exponential price update function respectively. As shown in Fig. 6-17 and Fig. 6-18, for a given call arrival rate, the proposed JCAC scheme with dynamic pricing tries to even out the highly-unbalanced traffic load among available RATs with time.

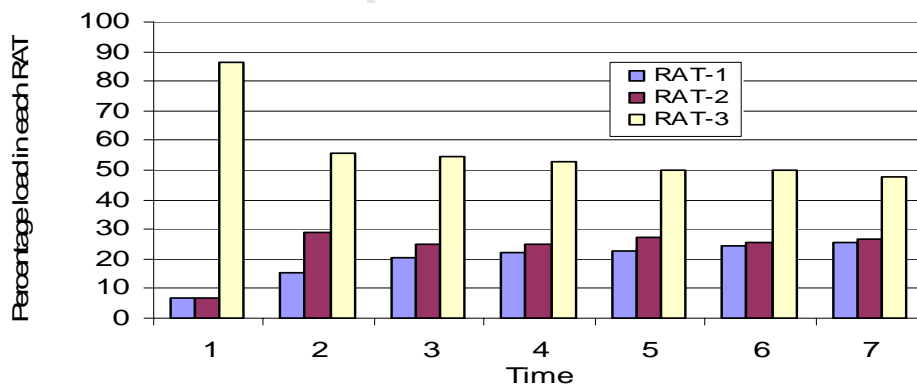


Figure 6-17. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3, z=0.1$).

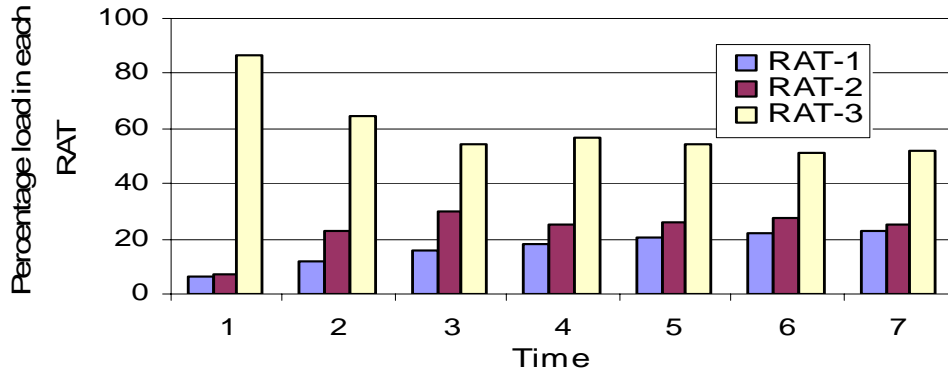


Figure 6-18. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n = 3, z=0.1$).

Fig. 6-19 and Fig. 6-20 show the percentage of load in each RAT for the JCAC scheme with dynamic pricing using linear and exponential price update function respectively. Parameter used in Fig. 6-19 and Fig. 6-20 are the same as those used in Fig. 6-17 and Fig. 6-18 except for the value of z that is increased from 0.1 to 0.2. Comparing Fig. 6-17, Fig. 6-18, Fig. 6-19, and Fig. 6-20 with Fig. 6-16, it can be seen that by varying the service price in each RAT, the proposed JCAC scheme improves the distribution of traffic load among available RATs in the heterogeneous wireless network.

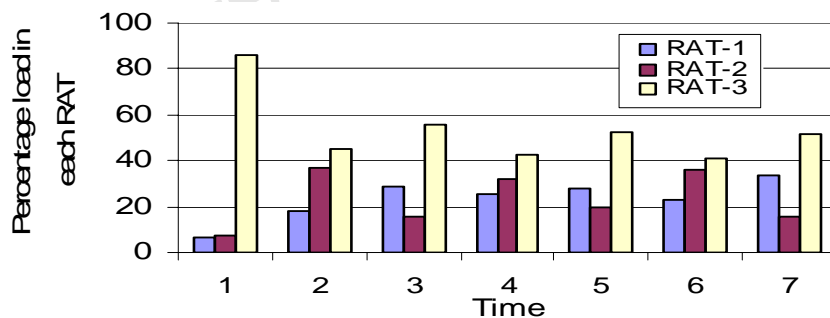


Figure 6-19. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3, z=0.2$).

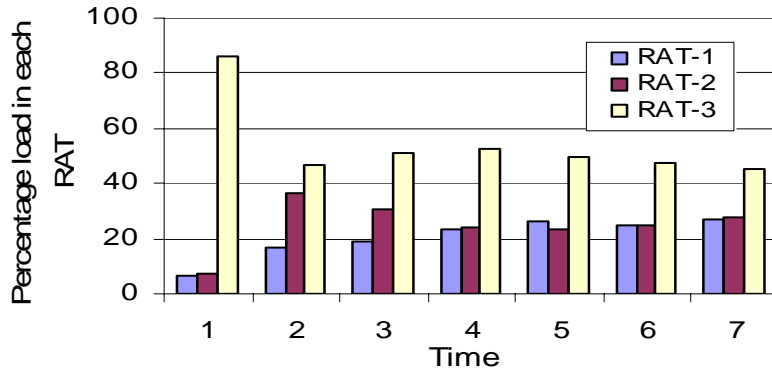


Figure 6-20. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n=3, z=0.2$).

Fig. 6-21 and Fig. 6-22 show the variation in service price with time for the JCAC scheme with dynamic pricing using linear and exponential price update function respectively. In Fig. 6-21 and Fig. 6-22, it is observed that though the service prices in RAT-1 and RAT-2 are almost constant with time, the price update unit keeps on increasing the service price in RAT-3 in order to achieve load balancing. However, after the fourth period of adjusting the service price in Fig. 6-21 and Fig 6-22, there is no longer any significant change in the load distribution among available RATs as shown in Fig. 6-17, Fig. 6-18, Fig. 6-19, and Fig. 6-20.

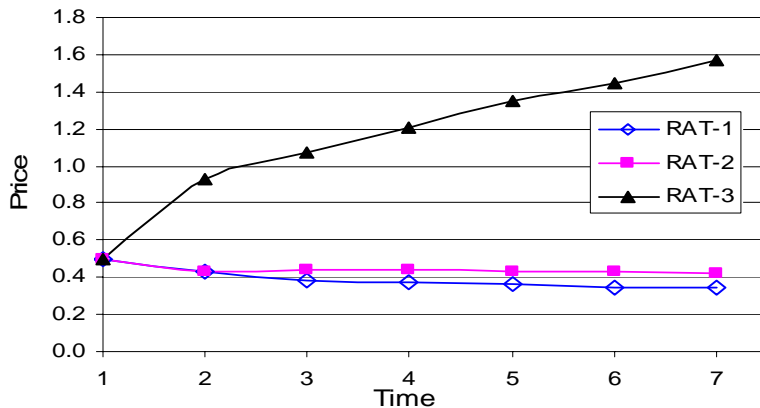


Figure 6-21. Variation of service price with update time using linear price-update function ($\lambda^n=3, z=0.1$).

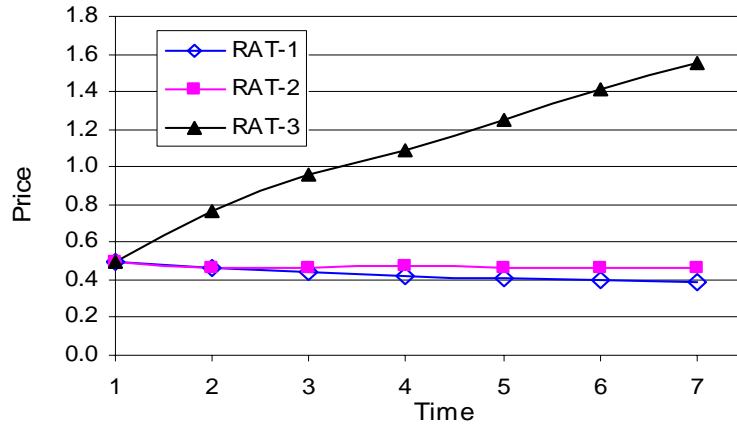


Figure 6-22. Variation of service price with update time using exponential price-update function ($\lambda^n = 3, z=0.1$).

6.6.2 High-Price-Sensitive User Scenario

In high-price sensitive user scenario, users attached high weights (great importance) to price. For illustration, weights assigned to price by users are randomly chosen from 6 to 9 whereas weights assigned to other criteria are randomly chosen from 0 to 9. Fig. 6-23 shows the performance of the proposed JCAC scheme with respect to call blocking/dropping probability. For each of the JCAC schemes, it is observed that P_d is always lower than the corresponding P_b . It can also be seen that both P_b with $DP-L$ and P_b with $DP-E$ are lower than P_b without DP . Similarly, both P_d with $DP-L$ and P_d with $DP-E$ are lower than P_d without DP .

Moreover, it is observed in Fig. 6-23 that both P_b with $DP-L$ and P_b with $DP-E$ are lower than the corresponding values for the general user scenario (Fig. 6-15). Similarly, both P_d with $DP-L$ and P_d with $DP-E$ in Fig. 6-23 are lower than the corresponding values for the general user scenario (Fig. 6-15). This shows that in a high-price sensitive user scenario, the proposed JCAC scheme with dynamic pricing achieves lower call blocking/dropping probability than in general user scenario. In other words, the performance of the proposed JCAC scheme increases as users' price sensitivity increases.

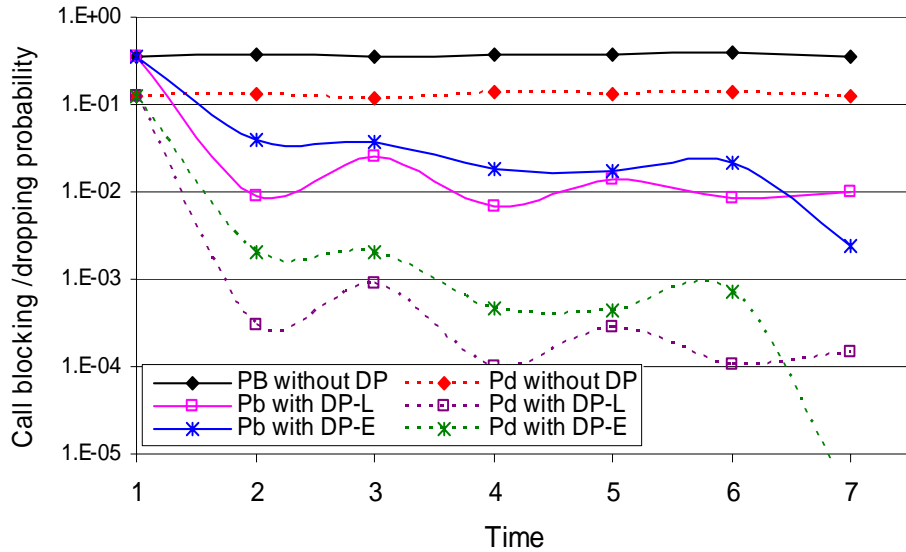


Figure 6-23. Call blocking/dropping probability versus price update time ($\lambda^n=3, z=0.1$).

Fig. 6-24 and Fig. 6-25 show the percentage of load in each RAT for the JCAC scheme with dynamic pricing using linear and exponential price update functions respectively. It is observe that in a high-price-sensitive user scenario, the proposed JCAC scheme with dynamic pricing, balances the traffic load among available RAT in the heterogeneous wireless network.

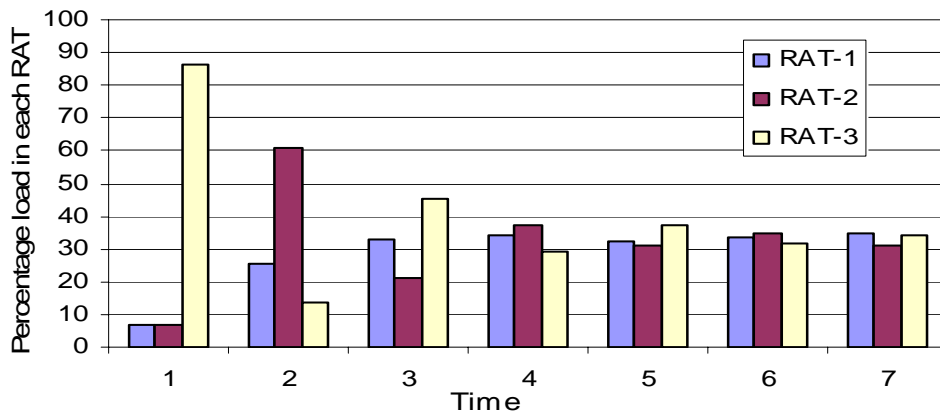


Figure 6-24. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n=3, z=0.1$).

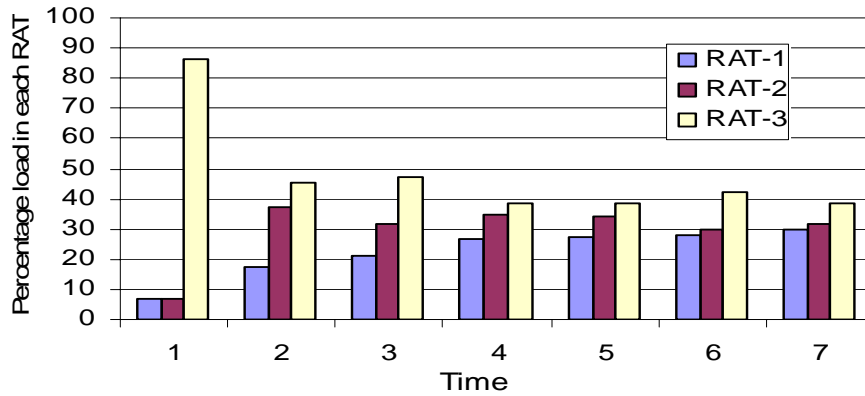


Figure 6-25. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential function ($\lambda^n=3, z=0.1$).

Fig. 6-26 and Fig. 6-27 show the variation in service price with time for the JCAC scheme with dynamic pricing using linear and exponential price update functions respectively. Comparing Fig. 6-24 with Fig. 6-26 and Fig. 6-25 with Fig. 6-27, it is observed that in a high-price-sensitive user scenario, variation in service price achieves significant load balance among available RATs in the heterogeneous wireless network. In Fig 6-26 and Fig 6-27, it can also be seen that after the fourth price-update period, the service price in each RAT remains almost constant with time.

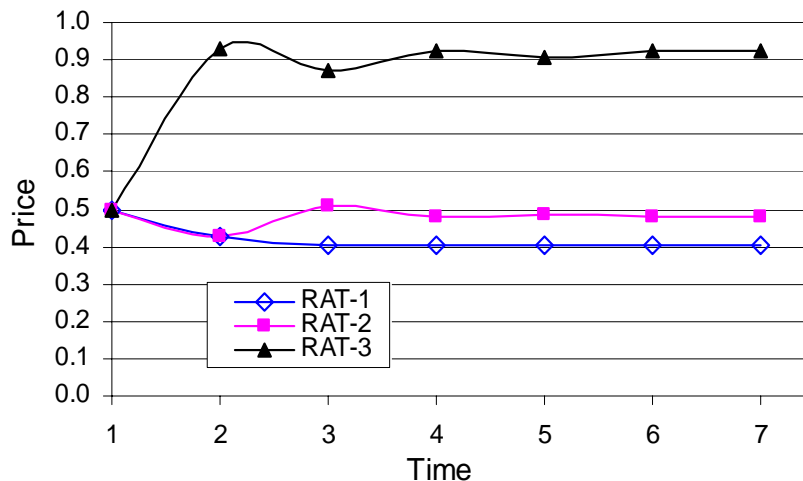


Figure 6-26. Variation of service price with update time using linear price-update function ($\lambda^n=3, z=0.1$).

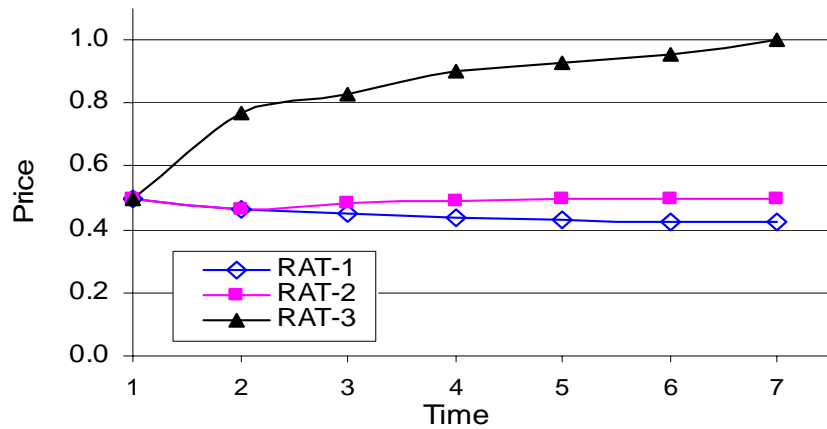


Figure 6-27. Variation of service price with update time using exponential price-update function ($\lambda^n = 3, z=0.1$).

6.6.3 Low-Price-Sensitive User Scenario

In low-price-sensitive user scenario, users attach low weight (less importance) to service price. For illustration, weights assigned to price by users are randomly chosen from 0 to 3 whereas weights assigned to other criteria are randomly chosen from 0 to 9.

Fig. 6-28 shows the performance of the proposed JCAC scheme with respect to call blocking/dropping probability. For each of the JCAC schemes, it is observed that P_d is always lower than the corresponding P_b . It can also be seen that both P_b with $DP-L$ and P_b with $DP-E$ are lower than P_b without DP . Similarly, P_d with $DP-L$ and P_d with $DP-E$ are lower than P_d without DP . Moreover, it is observed in Fig. 6-28 that for the JCAC scheme with dynamic pricing, both P_b with $DP-L$ and P_b with $DP-E$ are higher than the corresponding values for the general user scenario (Fig. 6-15). Similarly, both P_d with $DP-L$ and P_d with $DP-E$ are higher than the corresponding values for the general user scenario (Fig. 6-15). The above comparison implies that the performance of the proposed JCAC scheme decreases as users' price sensitivity decreases. However, Fig. 6-28 shows that even in a low-price-sensitive user scenario, the proposed JCAC scheme with dynamic pricing still reduces call blocking/dropping probability in the heterogeneous wireless network.

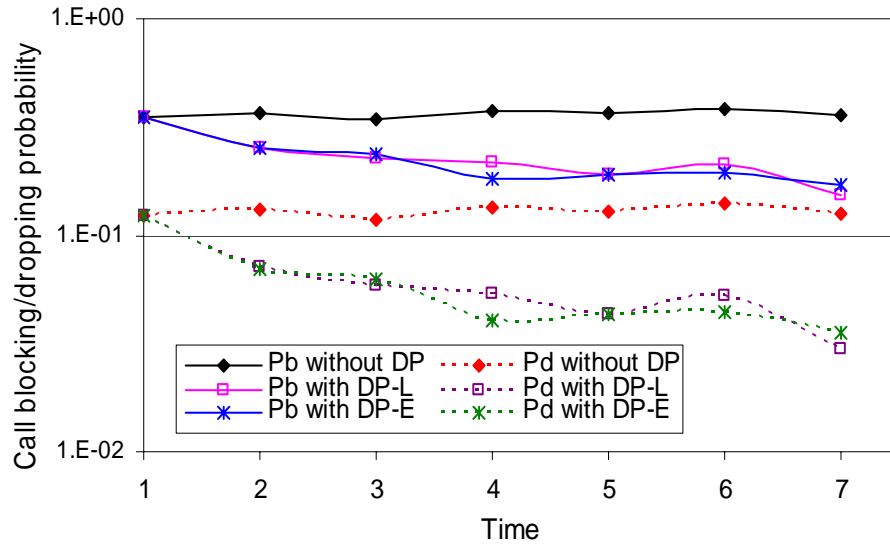


Figure 6-28. Call blocking/dropping probability versus price update time ($\lambda^n = 3, z=0.1$).

Fig. 6-29 and Fig 6-30 show the percentage of load in each RAT for the JCAC scheme with dynamic pricing using linear and exponential price update function respectively. It is observed that in a low-price-sensitive user scenario, there is a limit to load balancing that can be achieved among available RATs by the proposed JCAC scheme.

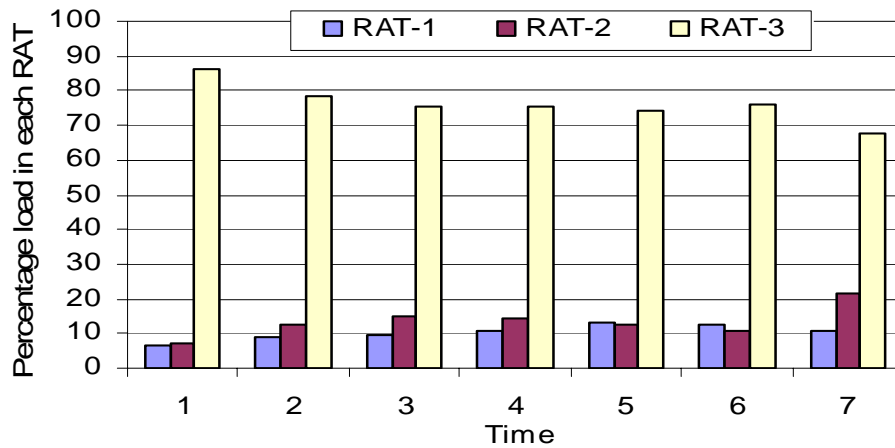


Figure 6-29. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using linear price-update function ($\lambda^n = 3, z=0.1$).

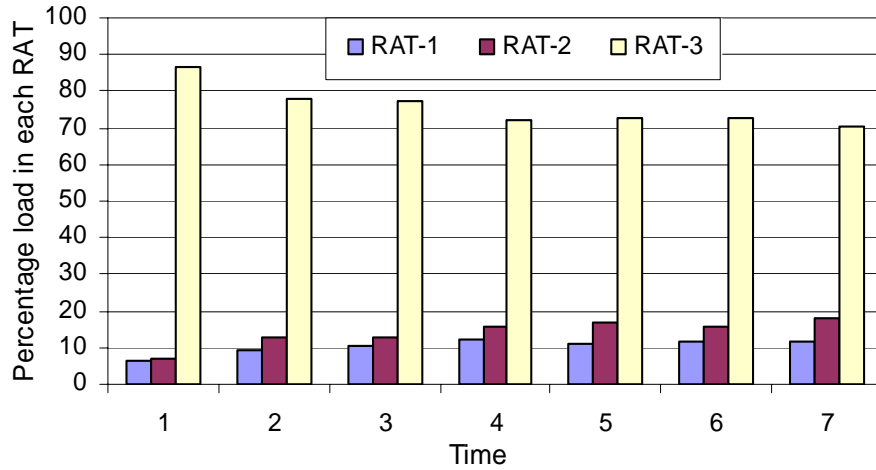


Figure 6-30. Percentage of load in each RAT versus price update time for JCAC with dynamic pricing using exponential price-update function ($\lambda^n=3, z=0.1$).

Fig. 6-31 and Fig. 6-32 show the variation in service price with time for the JCAC schemes with dynamic pricing using linear and exponential price update function respectively. Comparing Fig. 6-29 with Fig. 6-31 and Fig. 6-30 with and Fig. 6-32, it is observed that in a low-price-sensitive user scenario, a high variation in service price is necessary to achieve any reasonable load balance among available RATs in a heterogeneous wireless network. It can also be seen from Fig. 6-29, Fig. 6-30, Fig. 6-31, and Fig. 6-32 that, even with high price variation in a low-price-sensitive scenario, there is a limit to load balancing that can be achieved among available RATs.

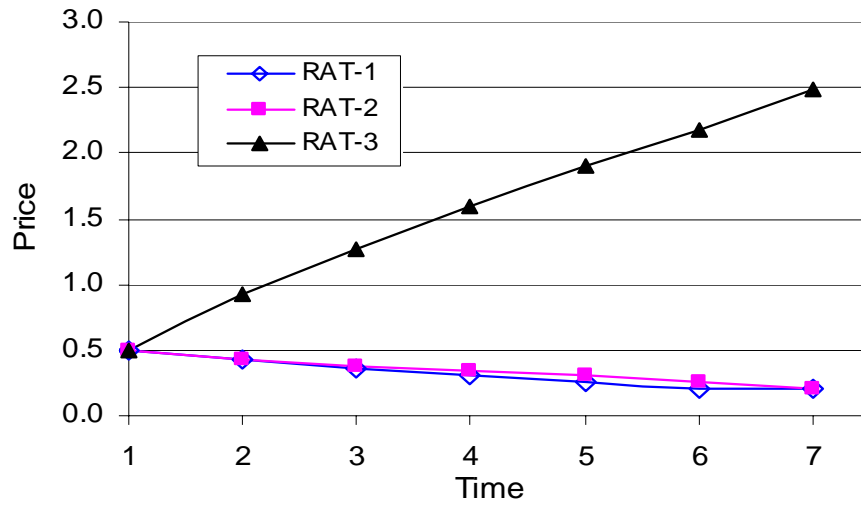


Figure 6-31. Variation of service price with update time using linear function ($\lambda^n = 3$, $z=0.1$).

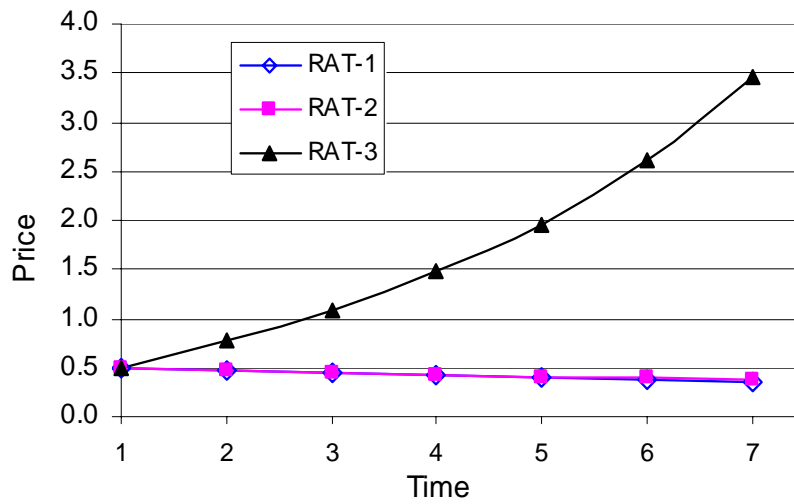


Figure 6-32. Variation of service price with update time using exponential function ($\lambda^n = 3$, $z=0.2$).

6.7 Summary

In this chapter, dynamic pricing has been proposed to even out the highly unbalanced traffic load caused by independent users' preferences in heterogeneous wireless networks where a user-centric JCAC scheme is employed. By dynamically adjusting the service price in each of the available RATs, the proposed scheme improves the distribution of traffic load in the heterogeneous wireless network. The JCAC scheme uses fuzzy MADM to select the most suited RAT for each incoming call thereby enhances users' satisfaction. Using a Markov model, new call blocking probability, handoff call probability, and percentage of load in each RAT are derived for the heterogeneous wireless network. The performance of the proposed JCAC scheme is illustrated using a three-RAT heterogeneous wireless network. Three different scenarios are considered namely general user scenario, high price-sensitive scenario, and low price-sensitive scenario. Results show that the proposed JCAC scheme with dynamic pricing achieves lower call blocking/dropping probability than the JCAC scheme that does not incorporate dynamic pricing. Results also show that the proposed scheme improves traffic load distribution among available RATs in a heterogeneous wireless network. Performance of the proposed scheme improves with increase in user-price sensitivity.

7 Conclusion and Future Work

Efficient radio resource utilization and QoS provisioning are major issues in heterogeneous wireless networks. This thesis focuses on joint call admission control and bandwidth management to enhance radio resource utilization and QoS provisioning in heterogeneous wireless networks. The summary of the contributions of the thesis and directions for future research are presented in the following subsections.

7.1 Summary of Contributions

Three JCAC schemes have been developed to enhance radio resource utilization and QoS provisioning in heterogeneous wireless networks. The main contributions of the thesis are summarized under the three schemes as follows.

7.1.1 Adaptive Bandwidth Management and Joint Call Admission Control to Enhance System Utilization and QoS in Heterogeneous Wireless Networks

In chapter 4, an adaptive bandwidth management and joint call admission control scheme is developed for heterogeneous wireless networks. The objectives of the adaptive JCAC scheme are to enhance average system utilization, uniformly distribute traffic load among available RATs, guarantee QoS requirements of all accepted calls, and reduce new call blocking probability and handoff call dropping probability in heterogeneous wireless networks. Using Markov decision process, an analytical model is developed for the adaptive JCAC scheme. Performance of the proposed JCAC scheme is compared with the performance of a JCAC scheme that does not incorporate adaptive bandwidth management. Numerical results show an improvement in average system utilization of up to 20%. Results also show an improvement in connection-level QoS in the heterogeneous wireless network. Moreover, the adaptive JCAC scheme prioritizes handoff calls over new calls by using a lower call rejection threshold for new calls.

7.1.2 Reduction of Call Blocking Probability through Optimal Call Allocation Policy

In chapter 5, an optimal RAT selection JCAC scheme is developed to reduce call blocking probability in heterogeneous wireless networks. The algorithm makes call admission decisions such that overall call-blocking probability is minimized in the heterogeneous wireless network. Optimal splitting of arrival calls is determined using linear programming optimization technique. Numerical results show that the algorithm reduces call-blocking probability in the heterogeneous wireless network. In the extreme case, the optimal JCAC scheme reduces the overall new call blocking probability to about 17% of the new call blocking probability of an equal-sharing JCAC scheme. Moreover, the optimal JCAC scheme reduces the overall handoff call dropping probability to about 37% of the handoff call dropping probability of an equal-sharing JCAC scheme.

7.1.3 Dynamic Pricing for Balancing Traffic Load in Heterogeneous Wireless Network Using Multiple-Criteria JCAC

In chapter 6, dynamic pricing is proposed to balance traffic load among available RATs in heterogeneous wireless networks where users' preferences are considered in making RAT selection decisions. Independent users' preferences in heterogeneous wireless networks often lead to highly unbalanced network load among available RATs, which in-turn increases overall call blocking/ dropping probability, and reduces radio resource utilization. By dynamically adjusting the service price in each of the available RATs, the proposed JCAC scheme evens out, as much as possible, the unbalanced traffic load caused by independent users' preferences. The performance of the proposed JCAC scheme is compared with the performance of a scheme that does not incorporate dynamic pricing. Numerical results show that the proposed JCAC scheme reduces new call blocking probability and handoff call dropping probability, and improves radio resource utilization in the heterogeneous wireless network. Results also show that the performance of the proposed JCAC scheme increases as users' price sensitivity increases.

The three JCAC schemes proposed in this thesis are applicable to a single operator with multiple access networks. The JCAC schemes are also applicable to different operators of access networks provided there is cooperation among the different operators.

7.2 Future Work

In this subsection, directions for future research are discussed.

7.2.1 Extension of JCAC Scheme to Cover a Non Fully-Overlapping Heterogeneous Wireless Network

The schemes developed in this thesis assume a heterogeneous wireless networks with fully overlapping coverage such as cellular networks with co-located cells. The JCAC and bandwidth management schemes proposed can be extended to cover a heterogeneous wireless network environment with non fully-overlapping RATs such as WiFi/WiMAX/cellular. In a heterogeneous wireless network with non fully-overlapping RATs, the current location of a mobile terminal will determine the number of RATs that are accessible to the mobile terminal. Therefore, it will be necessary to first develop a sophisticated mobility model. A JCAC scheme can then be developed based on the mobility model, coverage of individual RATs, current load in each RAT, and other factors of interest in the heterogeneous wireless network.

7.2.2 Consideration of Packet-Level QoS Metrics

This research focuses on connection-level QoS metrics. Evaluation of packet level-QoS metrics such as delay, jitter, and packet loss is outside the scope of the thesis. Consequently, it is assumed that packet-level QoS is stochastically assured by allocating at least the minimum effective bandwidth required to guarantee a given maximum probability on packet drop, delay, and jitter. Future work on the proposed JCAC and bandwidth management schemes could be carried of to investigate packet-level QoS metrics.

7.2.3 Consideration of Other Step Size Rules

In chapter 6, constant step size rule is considered for the proposed scheme in order to simplify the design. Future work could consider other step size rules such as diminishing step size, increasing step size, and dynamically chosen step size.

7.2.4 Combination of JCAC with JSS

Another possibility of future work is to combine JCAC with JSS for a greater flexibility. For a heterogeneous wireless network that consist of J number of RATs in set H , the combined JCAC and JSS scheme will operate in two stages. In the first stage, the JCAC algorithm will select a set N of RATs among the available RATs for the incoming call ($N \subseteq H$). In the second stage, the JSS algorithm will split the traffic of an incoming call among the selected RATs contained in set N provided the number of elements in set N is greater than 1.

7.2.5 Investigation of the Effect of Variation in Service Price on Operator' Revenue

Existing CAC algorithms use dynamic pricing to vary service price with time in homogeneous wireless networks. The variation in service price is based on users' demand for services at different times of the day. In this thesis, the User-centric JCAC scheme proposed uses dynamic pricing to vary service price across different RATs based on users' demand for each RAT.

Another possibility of future work is to employ dynamic pricing to vary the service price in each RAT based on users' demand for services during different times of the day and based on users' demand for each of the available RATs. Variation in service price in each RAT based on users' demand for services during different times of the day will address the problem of congestion in the entire heterogeneous wireless network. This congestion usually occurs during certain hours of the day (peak time). On the other hand, variation in

service price based on users' demand for specific RATs will address the problem of congestion that occurs in certain RATs as a result of users' independent preferences. Performance metric such as *operator's revenue against price* can be used to investigate the effect of variation in service price on the operator' revenue at different times of the day.

References

- [1] K. Murray, R. Mathur, and D. Pesch, "Network Access and Handover Control in Heterogeneous Wireless Networks for Smart Space Environments," First International Workshop on Management of Ubiquitous Communications and Services (MUCS '03), Waterford, Ireland, December 11, 2003.
- [2] J. P´erez-Romero, O. Sallent, R. Agust´ı, and M. A. D´ıaz-Guerra, "Radio Resource Management Strategies in UMTS," John Wiley & Sons, New York, NY, USA, 2005.
- [3] R. Agust´ı, O. Sallent, J. P´erez-Romero, and L. Giupponi, "A Fuzzy-Neural Based Approach for Joint Radio Resource Management in a Beyond 3G Framework," in Proceedings of the 1st International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE '04), Dallas, Texas, USA, October 2004, pp. 216–224.
- [4] N. R. Figueira and J. Pasquale, "Providing Quality of Service for Wireless Links: Wireless/Wired Networks," IEEE Personal Communications, vol. 6, no. 5, October 1998, pp. 42–51.
- [5] N. Nasser and H. Hassanein, "An Optimal and Fair Call Admission Control Policy for Seamless Handoff in Multimedia Wireless Networks with QoS Guarantees," IEEE Global Telecommunications Conference (Globecom), Dallas, Texas, November 2004, pp. 3926 – 3930.
- [6] S. S. Rappaport, "The Multiple-Call Hand-off Problem in High-Capacity Cellular Communications Systems," IEEE Transactions on Vehicular Technology, vol. 40, no. 3, August 1991, pp. 546-557.
- [7] N.Tang, S. Tsui, and L. Wang, "A Survey of Admission Control Algorithms," http://www.cs.ucla.edu/~tang/papers/admission_control_paper.ps, UCLA class paper, March 1998.
- [8] R. M. Rao, Cristina, and H. V. Poor, "Call Admission Control in Wireless Multimedia Networks: An Overview of CAC Principles in DS-CDMA Networks," IEEE Signal Processing Magazine, vol. 21, Issue 5, 2004, pp. 51-58.
- [9] M. H. Ahmed, "Call Admission Control in Wireless Networks: A Comprehensive Survey," IEEE Communications Surveys and Tutorials, vol. 7, no. 1, 1st Quarter 2005, pp. 2-21.
- [10] D. Karabudak, C. Hung, and B. Bing, "A Call Admission Control Scheme Using Genetic Algorithms," Symposium on Applied Computing (SAC '04), Nicosia, Cyprus, March 14 -17, 2004.
- [11] V. A. de Sousa Jr., R. A. de O. Neto, F. de S. Chaves, L. S. Cardoso, J. F. Pimentel, and F. R. P. Cavalcanti, "Performance of Access Selection Strategies in Cooperative Wireless Networks using Genetic Algorithms," Proceedings of the 15th World

- Wireless Research Forum Meeting (WRRF '05), December 8-9, Paris, France, 2005.
- [12] X. Gelabert, J. Perez-Romero, O Sallent, and R. Agusti, "On the Stability of Load Balancing Principles in Heterogeneous Wireless Access Networks," 1st International Wireless Summit, Aalborg Denmark, 17-22 September, 2005.
 - [13] A. Pillekeit, F. Derakhshan, E. Jugl, and A. Mitschele-Thiel, "Force-Based Load Balancing in Co-located UMTS/GSM Networks," Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04), vol. 6, Los Angeles, Calif, USA, September 2004, pp. 4402–4406.
 - [14] W. Zhang, "Performance of Real-time and Data Traffic in Heterogeneous Overlay Wireless Networks," Proceedings of the 19th International Teletraffic Congress (ITC '05), Beijing, China, 2005.
 - [15] J. Pérez-Romero, O. Sallent, and R. Agustí, "Policy-based Initial RAT Selection Algorithms in Heterogeneous Networks," 7th IFIP International Conference on Mobile and Wireless Communications Networks, Marrakech, Morocco, September 19-21, 2005.
 - [16] J. Pérez-Romero, O. Sallent, R. Agustí, L. Wang, and H. Aghavmi, "Network-Controlled Cell-Breathing for Capacity Improvement in Heterogeneous CDMA/TDMA Scenarios," IEEE Wireless Communications and Networking Conference (WCNC '06), Las Vegas, NV, USA, 3-6 April, 2006.
 - [17] O.E. Falowo and H.A. Chan, "Joint Call Admission Control for Integrated UMTS-WLAN Network," Proceedings of South African Telecommunication Networks Conference, Zulu Natal, South Africa, 11-14 September, 2005.
 - [18] R. B. Ali and S. Pierre, "An efficient Predictive Admission Control Policy for Heterogeneous Wireless Bandwidth Allocation in Next Generation Mobile Networks," International Conference on Communications and Mobile Computing (IWCMC'06), Vancouver, Canada, July 3-6, 2006.
 - [19] 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, vol. 39, no.12, December 2001, pp. 42-51.
 - [20] W. Zhang, "Handover Decision Using Fuzzy MADM in Heterogeneous Networks," Proceeding of IEEE Wireless Communications and Networking Conference (WCNC'04), Atlanta, GA, March 2004.
 - [21] A. L. Wilson, A. Lenaghan, and R. Malyan, "Optimizing Wireless Network Selection to Maintain QoS in Heterogeneous Wireless Environments," Proceedings of World Communication Forum, Demark, September, 2005.
 - [22] J. Chen, K. Yu, Y. Ji, and P. Zhang, "Non-Cooperative Distributed Network Resource Allocation in Heterogeneous Wireless Data Networks," Proceedings of IST Mobile Summit, Mykonos, Greece, 5-7.June, 2006.

- [23] G. Wu, M. Mizuno, and P.J.M. Havinga, "MIRAI Architecture for Heterogeneous Network," *IEEE Communication Magazine*, vol. 40, no. 2, Feb. 2002.
- [24] E. Vanem, S. Svaet, and F. Paint, "Effects of Multiple Access Alternatives in Heterogeneous Wireless Networks," *IEEE Wireless and Networking*, vol. 3, 2003, pp. 1696–1700.
- [25] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. Mäkelä, R. Pichna, J. Vallström, "Handoff in Hybrid Mobile Data Networks," *IEEE Personal Communications Magazine*, vol.7, no.2, April 2000, pp. 34-47.
- [26] J. McNair and F. Zhu, "Vertical Handoffs in Fourth-Generation Multi-Network Environments," *IEEE Wireless Communications*, vol. 11, no. 3, 2004, pp. 8–15.
- [27] S. Glass, T. Hiller, S. Jacobs, and C. Perkins, "Mobile IP Authentication, Authorization, and Accounting Requirements," Internet draft (work in progress), 2000.
- [28] S. Y. Hui and K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems," *IEEE Communications Magazine*, vol. 41, no. 12, Dec. 2003, pp. 54-59.
- [29] S. J. Lincke, "Vertical Handover Policies for Common Radio Resource Management," *International Journal of Communication*, 2005: Published Online: 15 March, 2005.
- [30] M. O'Callaghan, N. Gawley, M. Barry, and S. McGrath, "Admission Control for Heterogeneous Networks," 13th IST Mobile and Wireless Communications Summit, Lyon, June 27 – 30, 2004.
- [31] A. Tölli, P. Hakalin, and H. Holma, "Performance Evaluation of Common Radio Resource Management (CRRM)," *Proceedings of the IEEE International Conference on Communications (ICC '02)*, vol. 5, New York, NY, USA, April-May 2002, pp. 3429–3433.
- [32] A. Bria, "Joint Resource Management of Cellular and Broadcasting Systems – Research Challenges," *Proceedings of RVK05 Conference*, Linköping, 2005.
- [33] A. Tölli and P. Hakalin, "Adaptive Load Balancing Between Multiple Cell Layers," *Vehicular Technology Conference (VTC 2002)* vol. 3, September 24–28, 2002, pp. 1691–1695.
- [34] K.H. Suleiman, H.A. Chan, and M.E. Dlodlo, "Load Balancing in the Call Admission Control of Heterogeneous Wireless Networks," *Proceedings of International Conference on Communications and Mobile Computing (IWCMC'06)*, Vancouver, Canada, July 3–6, 2006, pp. 245–250.
- [35] W. Song, H. Jiang, W. Zhuang, and X. Shen, "Resource Management for QoS Support in WLAN/Cellular Interworking," *IEEE Network Magazine*, vol. 9, no. 5, 2005, pp. 12–18.

- [36] I. Koo, A. Furuskar, J. Zander, and K. Kim, "Erlang Capacity of Multi-access Systems with Service-Based Access Selection," *IEEE Communications Letters*, vol. 8, no. 11, 2004, pp. 662–664, ISSN: 1089-7798.
- [37] O.E. Falowo and H.A. Chan, "Joint Call Admission Control for Next Generation Wireless Networks," *Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering*, Ottawa, Canada, May 7–10, 2006.
- [38] O. Ormond, J. Murphy, and G. Muntean, "Utility-Based Intelligent Network Selection in Beyond 3G Systems," *IEEE International Conference on Communications (ICC 2006)*, Istanbul, Turkey, June 11–15.
- [39] M.L. Puterman, "Markov Decision Processes: Discrete Stochastic Dynamic Programming," John Wiley & Sons, New York, 1994, ISBN 0-471-61977-9.
- [40] O.E. Falowo and H.A. Anthony, "Fuzzy Logic Based Call Admission Control for Next Generation Wireless Networks," *Proceedings of 3rd International Symposium on Wireless Communication Systems*, Valencia, Spain, September 5–8, 2006.
- [41] S. Chen, C. Hwang, and F.P. Hwang, "Fuzzy Multiple Attribute Decision Making," Springer-Verlag, 1992.
- [42] 3GPP TR 25.881 v5.0.0 Improvement of RRM across RNS and RNS/BSS.
- [43] 3GPP TR 25.891 v0.3.0 Improvement of RRM across RNS and RNS/BSS (Post Rel-5) (Release 6).
- [44] J. Pe´rez-Romero, O. Sallent, and R. Agustí, "On Evaluating Beyond 3G Radio Access Networks: Architectures, Approaches and Tools," *IEEE Vehicular Technology Conference*, Stockholm, Sweden, May 30–June 1, 2005.
- [45] A. Furuskar and J. Zander, "Multi-service Allocation for Multi-Access Wireless Systems," *IEEE Transactions on Wireless Communications* vol. 4, January, 2005, pp.174–184.
- [46] P. Demestichas and V. Stavroulaki, "Issues in Introducing Resource Brokerage Functionality in B3G, Composite Radio Environments," *IEEE Wireless Communications Magazine* vol. 11, 2004, pp. 32–40.
- [47] F. Yu and V. Krishnamurthy, "Optimal Joint Session Admission Control in Integrated CDMA/WLAN Networks with Vertical Handoff," *IEEE Transactions on Mobile Computing*, vol. 6, no. 1, 2007.
- [48] N. Nasser, "A Theoretical Approach for Service Provider Decision in Heterogeneous Wireless Networks," *IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, Nevada, USA, January 2006, pp. 1249–1252.
- [49] V.P. Kafle, E. Kamioka, and S. Yamada, "Maximizing User Satisfaction Based on Mobility in Heterogeneous Mobile Communication Networks," *IEICE Transactions on Communications E88-B*, vol. 7, 2005, pp. 2709–2717.

- [50] C. Ho and C. Lea, "Improving Call Admission Policies in Wireless Networks," *Wireless Networks*, vol. 5, issue 4, 1999, pp. 257-265.
- [51] Y. C. Lai and S. F. Tsai, "A Fair Admission Control for Large Bandwidth Multimedia Applications," *Proceedings of GLOBECOM' 2002*, San Antonio, Texas, November 2002, pp. 317-322.
- [52] D.Z. Deniz and N.O. Mohamed, "Performance Analysis of a Call Admission Control Strategy in a Broadband Node with Narrowband and Wideband Traffic," *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS'02)*, San Diego, CA, USA, July 14-18, 2002, pp.749- 755.
- [53] T. Kanter, "Going Wireless: Enabling an Adaptive and Extensible Environment", *ACM Journal on Mobile Networks and Applications*, vol. 8, no. 1, February 2003, pp. 37-50.
- [54] D. Hong and S. S. Rappaport, "Traffic Model and Performance Analysis of Cellular Radio Telephony Systems with Prioritized and Non-prioritized Handoff Procedures," *IEEE Transaction of Vehicular. Technology*, vol. 35, August. 1986.
- [55] C. Chang, C. J. Chang, and K. R. Lo, "Analysis of a Hierarchical Cellular System with Reneging and Dropping for Waiting New Calls and Handoff Calls," *IEEE Transaction on Vehicular Technology*, vol. 48, no. 4, 1999, pp. 1080–1091.
- [56] S. S. Rappaport and C. Purzynski, "Prioritized Resource Assignment for Mobile Cellular Communication Systems with Mixed Services and Platform Types," *IEEE Transaction on Vehicular Technology*, vol. 45, no. 3, August. 1996.
- [57] H. Chen, S. Kumar, and C.C. Jay Kuo, "Differentiation QoS Aware Priority Handoff in Cell-Based Multimedia Wireless Network," *Electronic Imaging IS&T/SPIE 12th International Symposium*, San Jose, CA, January 2000.
- [58] H. Chen, S. Kumar, and C.-C. Jay Kuo, "Dynamic Call Admission Control Scheme for QoS Priority Handoff in Multimedia Cellular Systems," *Proceedings of IEEE Wireless Communications and Networking Conference*, Orlando, vol. 1, March, 2002, pp. 114-118.
- [59] S. Mahmood¹, H. Sirisena¹, and K. Pawlikowski, "Call Admission Control for Mobile Multimedia Wireless Networks: DS-CDMA Cellular Systems," *IEEE Journal in Selected Areas in Communication*, vol. 12, May 1994, pp. 638-644.
- [60] S. Tekinay and B. Jabbari, "A Measurement-Based Prioritization Scheme for Handovers in Mobile Cellular Networks," *IEEE Journal in Selected Areas in Communication*, vol. 10, no. 8, May 1992.
- [61] C. Lindemann, M. Lohmann, and A. Thummler, "Adaptive Call Admission Control for QoS/Revenue Optimization in CDMA Cellular Networks," *ACM Journal on Wireless Networks (WINET)*, vol. 10, 2004, pp. 457-472.

- [62] M. M. Islam and M. Murshed, "Min-Max Fairness Scheme for Resource Allocation in Cellular Multimedia Networks," International Conference on Information Technology: Coding and Computing (ITCC'05), volume II, 2005, pp. 265-270.
- [63] S.Sen, J. Jawanda, K.Basau, and S. K. Das, "Quality of Service Degradation Strategies for Multimedia Wireless Networks," Proceedings of Vehicular Technology Conference, Ottawa Canada, May 1998, pp. 1884-1888.
- [64] N. Bartolini and I. Chamtac, "Call Admission Control in Wireless Multimedia Networks," Proceedings of the 13th IEEE International Symposium on Personal Indoor Mobile Radio Communication (PIMRC2002), vol. 1, September 2002, pp. 285- 289.
- [65] M. Aboelaze, "A Call Admission Protocol for Cellular Networks that Support Differentiated Fairness," Proceeding of the IEEE 59th Vehicular Technology Conference (VTC2004), May 2004.
- [66] H. Holma and A. Toskala, "WCDMA for UMTS," John Wiley & Sons, New York, NY, USA, 2nd edition, 2001.
- [67] V. Pla, J. M. Giménez-Guzmán, J. Martínez, and V. Casares-Giner, "Optimal Admission Control using Handover Prediction in Mobile Cellular Networks," Proceedings of the 2nd International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs '04), Ilkley, West Yorkshire, UK, July 2004.
- [68] G. Kesidis, J. Walrand, and C.-S. Chang, "Effective Bandwidths for Multi-class Markov Fluids and other ATM Sources," IEEE/ACM Transactions on Networking, vol. 1, no. 4, 1993, pp. 424–428.
- [69] M. Gałbowski, M. Stasiak, A. Wiśniewski, and P. Zwierzykowski, "Uplink Blocking Probability Calculation for Cellular Systems with WCDMA Radio Interface, Finite Source Population and Differently Loaded Neighbouring Cells," Proceedings of the 11th Asia-Pacific Conference on Communications (APCC '05), Perth, Western Australia, October 2005, pp. 138–142.
- [70] N. Nasser and H. Hassanein, "Dynamic Threshold-Based Call Admission Framework for Prioritized Multimedia Traffic in Wireless Cellular Networks," Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '04), vol. 2, Dallas, Texas, USA, November-December, 2004, pp. 644–649.
- [71] Y. Guo and H. Chaskar, "Class-Based Quality of Service over Air Interfaces in 4G Mobile Networks," IEEE Communications Magazine, vol. 40, no. 3, 2002, pp. 132–137.
- [72] S.-P. Chung and J.-C. Lee, "Performance Analysis and Overflowed Traffic Characterization in Multi-Service Hierarchical Wireless Networks," IEEE Transactions on Wireless Communications, vol. 4, no. 3, 2005, pp. 904–918.

- [73] F. Houeto and S. Pierre, "Characterization of Jitter and Admission Control in Multi-Service Networks," *IEEE Communications Letters*, vol. 8, no. 2, 2004, pp. 125–127.
- [74] J. Wang, Q.-A. Zeng, and D. P. Agrawal, "Performance Analysis of a Pre-emptive and Priority Reservation Handoff Scheme," *Transactions on Mobile Computing*, vol. 2, no. 1, 2003, pp. 65–75.
- [75] X. Bo and Z. Chen, "On Call Admission and Performance Evaluation for Multi-Service CDMA Networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 8, no. 1, 2004, pp. 98–108.
- [76] H. Chen, Q-A. Zeng, and D. P. Agrawal, "A Novel Analytical Modelling for Optimal Channel Partitioning in the Next Generation Integrated Wireless and Mobile Networks," *ACM MSWiM Workshop*, September 28, 2002.
- [77] S. K. Das, S. K. Sen, K. Basu, and H. Lin, "A Framework for Bandwidth Degradation and Call Admission Control Schemes for Multi-Class Traffic in Next-Generation Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, December 2003, pp. 1790 - 1801.
- [78] P. V. Orlik and S. S. Rappaport, "On the Handoff Arrival Process in Cellular Communications," *Wireless Networks*, vol. 7, no. 2, 2001, pp. 147–157.
- [79] D. P. Bertsekas and J. N. Tsitsiklis, "Introduction to Probability," Athena Scientific, Belmont, Mass, USA, 2002.
- [80] S. Frattasi, H. Fathi, F. Fitzek, Kiho Chung, and R. Prasad, "4G: A User-Centric System," *Mobile E-Conference (Me)*, August, 2004.
- [81] V. Ozianyi and N. Ventura, "Pricing Profiles for 3G Services with QoS Guarantees," 13th International Conference on Telecommunications, Funchal, Madeira island, Portugal, 9-12 May, 2006.
- [82] J. Hou, J. Yang, and S. Papavassiliou, "Integration of Pricing with Call Admission Control to Meet QoS Requirements in Cellular Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, September 2002.
- [83] S. Yaipairoj and F. Harmantzis, "A Dynamic Pricing Model for Data Services in GPRS Networks," *Proceedings of IEEE Workshop on Adaptive Wireless Networks (AWiN 2004) (in conjunction with GLOBECOM 2004)*, pp. 453-458.
- [84] P. Marbach, "Analysis of a Static Pricing Scheme for Priority Services," *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, April 2004.
- [85] E.D. Fitkov-Norris and A. Khanifar, "Dynamic Pricing in Mobile Communication Systems," *First International Conference on 3G Mobile Communication Technologies*, 2000, pp. 416–420.
- [86] M. Bitsaki, G. Stamoulis, and C. Courcoubetis, "An Efficient Auction Mechanism for Hierarchically Structured Bandwidth Markets," *Proceedings of ICQT'04, LNCS*

3266, Barcelona, Spain, Oct 2004, Springer.

- [87] K. Park, M. Sitharam, and S. Chen, "Quality of Service Provision in Noncooperative Networks: Heterogenous Preferences, Multi-Dimensional QoS Vectors, and Burstiness," Proceedings of International Conference on Information and Computation Economies, 1998, pp. 111–127.
- [88] I. Chen, O. Yilmaz, and I. Yen, "Admission Control Algorithms for Revenue Optimization with QoS Guarantees in Mobile Wireless Networks," Wireless Personal Communications, vol. 38, Springer 2006, pp. 357–376.
- [89] H. Mendelson and S. Whang, "Priority Pricing for the M/M/I Queue," Operational Research, vol. 38, no. 5, 1990, pp. 870–883.
- [90] U. Cem, N. Saraydar, and B. Mandayamy, "Efficient Power Control via Pricing in Wireless Data Networks," IEEE Transactions on Communications, vol. 50, no. 2, February 2002.
- [91] J. Hou, J. Yang, and S. Papavassiliou, "Integration of Pricing with Call Admission Control for Wireless Networks," Vehicular Technology Conference, 2001, vol. 3, 2001, pp. 1344 -1348.
- [92] C. Sarayadar, N. Mandayam, and D. Goodman, "Pareto Efficiency of Pricing-Based Power Control in Wireless Data Networks", IEEE GLOBECOM 1999, pp. 231-235.
- [93] E. Viterbo and C. F. Chiasserini, "Dynamic Pricing For Connection-Oriented Services in Wireless Networks," 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, San Diego, CA, USA, 30 September – 3 October, 2001.
- [94] S. Chen, C. Hwang and F. P. Hwang, "Fuzzy Multiple Attribute Decision Making," Springer-Verlag, 1992.
- [95] C. L. Hwang and K. Yoon, "Multiple Attribute Decision Making-Methods and Application: A State of the Art Survey," Springer-Verlag, Berlin, 1981.
- [96] T. L. Saaty, "Analytic Hierarchy Process," McGraw-Hill, 1980.
- [97] N. Jin and G. Venkitachalam, "Dynamic Congestion-Based Pricing of Bandwidth and Buffer," IEEE/ACM Transactions on Networking, vol. 13, no. 6, December 2005.