

Wireless Optimisation Based on Economic Criteria

by

Siew Lee Hew

B.E. (IT & T, HONS. 1)

Thesis submitted for the degree of

Doctor of Philosophy

in

School of Electrical and Electronic Engineering,
Faculty of Engineering, Computer and Mathematical Sciences

The University of Adelaide, Australia

2006

© Copyright 2006
Siew Lee Hew
All Rights Reserved



Typeset in L^AT_EX 2_ε
Siew Lee Hew

Statement of Originality

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of the thesis, when deposited in the University Library, being available for loan and photocopying.

Signed

Date

Acknowledgements

I am highly indebted to my principal supervisor Professor Langford B. White for his guidance, support, motivation and encouragement throughout my candidature. Working with such an inspiring and excellent mentor has been an insightful journey for me. I still have a lot to learn from him, who never ceases to explore new knowledge.

I have profited immensely from various discussions with my two other excellent supervisors, Dr. Paul Chapman and Dr. Matthew Sorell. Particularly, my simulation work on Smart Pricing with them at the beginning of this research has sparked the idea for our work in Chapter 3. I am also grateful to them for patiently helping me to improve my writing skills.

This research would not have been possible without the generous financial help in the form of the Australian Postgraduate Award (APA) and the CRC for Smart Internet Technology top-up scholarships. I would also like to thank the CRC for Smart Internet, University of Adelaide, and ACORN for sponsoring my workshop and conference travels.

My sincere gratitude also goes to my wonderful colleagues at the Centre for Internet Research (CIR) who have given me a friendly second home for the past three years. Extra kudos to those (especially Ange) who gave me chocolates and lollies from time to time! I am also thankful to the friendly people working at the main office of the School of Electrical and Electronic Engineering for their help regarding various matters.

Last but not least, I would like to thank my partner and family. Without your encouragement, emotional support and understanding, I will never be able to complete this remarkable but often challenging journey.

Siew-Lee Hew

April 2006

Abstract

The rapid growth in demand due to the emergence of mobile communication services with variable rates, coupled with the resource scarcity of mobile air interface, has encouraged researchers to find technological solutions to increase spectral efficiency in order to support different levels of Quality of Service (QoS). Radio resource management (RRM) plays a major role in QoS provisioning and congestion control for wireless networks. The main problem with the congestion control mechanisms provided by current RRM schemes is that they are mostly reactive, triggered only when congestion occurs. The common, traditional solution to congestion has been for system planners to *over-engineer* a network by assigning more resources than are necessary. This approach is very costly because busy periods are usually brief, causing the network to be often under-utilised outside of these periods. Current static, usage-based pricing models also fail to assist in traffic shaping to even out loads.

Economic modelling offers a new perspective into current RRM schemes and enables efficient utilisation of scarce resources and congestion prevention based on concepts such as utility, price, Pareto optimality and game theory. Dynamic pricing has been proposed as a mechanism to encourage users to adapt their resource consumption level according to network conditions. A good pricing model can provide the necessary positive incentives to increase users' arrival rate when the network load is relatively low and negative incentives for users to defer their usage when the load is relatively high. In this dissertation, we propose an economic framework for pricing and RRM for 3G and beyond systems. Our aim is two-fold: to calculate an optimal integrated dynamic pricing and RRM policy; and to allocate scarce network resources in a fair and Pareto-optimal manner.

The optimal integrated dynamic pricing and RRM policy is computed based on the stochastic distribution of users' budget, arrivals, handoffs and departures. Our results

show that the integrated policy is superior in terms of average reward improvement and congestion prevention to current schemes that use static pricing models. In interference-based networks such as WCDMA, we suggest users be charged according to their noise rise factor, i.e. an estimate of the amount of interference generated by the call. This interference-based pricing model improves on the conventional load-based model in by delivering higher revenue and lower call blocking and handoff probabilities.

Using the axiomatic bargaining concepts from cooperative game theory, we derive a class of fair and Pareto-optimal bargaining solutions that allocate wireless resources based on users' minimum and maximum rate requirements. We propose two models: symmetric and asymmetric. In the latter, resource is allocated according to the price paid by the users. An important significance of the asymmetric bargaining model is that this solution is still Pareto-optimal and fair according to the users' bargaining power. Our approach is also a departure from current works using noncooperative game theory that can only achieve an inefficient outcome, i.e. the Nash equilibrium; or cooperative game theory that focus on only one solution on the Pareto-optimal boundary. By analysing a range of bargaining solutions instead of specific ones, operators can proceed to select the best outcome out of these Pareto-optimal solutions based on criteria like revenue.

Publications

1. S-L. Hew and L. B. White, “Interference-based dynamic pricing and radio resource management for WCDMA networks”, in *Proc. of 61st IEEE Vehicular Technology Conference (VTC2005-Spring)*, May 30 - June 1 2005, Stockholm, Sweden.
2. S-L. Hew and L. B. White, “Optimal integrated call admission control and dynamic pricing with handoffs and price-affected arrivals”, in *Proc. of Asia-Pacific Conference on Communications (APCC’05)*, Oct. 3-5 2005, Perth, Australia.
3. S-L. Hew and L. B. White, “Fair resource bargaining solutions for cooperative multi-operator networks”, in *Proc. of IEEE International Zurich Seminar on Communications (IZS’06)*, Feb. 22-24 2006, Zurich, Switzerland.
4. S-L. Hew and L. B. White, “A game theoretical approach for resource bargaining in shared WCDMA networks: symmetric and asymmetric models”, in *Proc. of 1st IEEE International Conference on Wireless Broadband and Ultra Wideband Communications (AusWireless’06)*, Mar. 13-16 2006, Sydney, Australia.
5. S-L. Hew and L. B. White, “Integrated call admission control and dynamic pricing in multiservice networks with price-affected arrivals, retrials and substitutions”, submitted for publication.
6. S-L. Hew and L. B. White, “Interference-based dynamic pricing for WCDMA networks using neuro-dynamic programming”, accepted for publication in *IEEE Trans. on Automatic Control*.
7. S-L. Hew and L. B. White, “Cooperative resource allocation games in shared networks”, submitted to *IEEE Trans. on Wireless Communications*, in Feb. 2006.

Dedication

This dissertation is dedicated to my parents, *Hew Peng Fah* and *Tan Hong Gaik*, for their sacrifice, valuable teachings and unconditional support all these years.

Contents

Statement of Originality	iii
Acknowledgements	v
Abstract	vii
Publications	ix
Dedication	xi
Contents	xiii
List of Figures	xix
List of Tables	xxiii
Chapter 1. Introduction	1
1.1 Background and Motivation	3
1.1.1 Radio Resource Management	3
1.1.2 Economic Modelling	7
1.1.3 Pricing and Charging Models in Communication Networks	9
1.2 An Economic Framework for Pricing and RRM	12

Contents

1.3	Organisation and Original Contributions	14
1.3.1	Chapter 2: Background	14
1.3.2	Chapter 3: Integrated Call Admission Control and Dynamic Pricing Cellular Networks	14
1.3.3	Chapter 4: Interference-based Radio Resource Management and Dynamic Pricing for WCDMA Networks	15
1.3.4	Chapter 5: Cooperative Resource Bargaining Games for Shared Networks	17
1.3.5	Chapter 6: Conclusion	18
Chapter 2. Background		19
2.1	WCDMA System	19
2.1.1	Wireless Resource Parameters	20
2.2	Economic Modelling Concepts	23
2.2.1	Game Theory	23
2.2.2	Pricing Schemes for Wired Networks	29
2.2.3	Pricing Schemes for Wireless Networks	34
2.3	Optimal Control Theory	39
2.3.1	Dynamic Programming	39
2.3.2	Neuro-Dynamic Programming	41
Chapter 3. Integrated Dynamic Pricing and Call Admission Control		43
3.1	Introduction	43
3.2	System Model	46
3.3	Markov Decision Problem Formulation	49
3.3.1	State Space and Transition Rates	49

3.3.2	Event and Control Space	51
3.3.3	Revenue Maximisation Problem	52
3.4	Calculation of Stationary Probabilities	57
3.4.1	Level-Dependent Quasi-Birth-Death Process	58
3.4.2	Matrix Analytic Methods	58
3.4.3	System Characteristics	59
3.5	Numerical Results	61
3.5.1	Policy Comparison	63
3.5.2	Congestion Control	67
3.5.3	Price Discrimination	70
3.5.4	Exponential WTP Distribution	71
3.6	Conclusions	72
 Chapter 4. Interference-based Dynamic Pricing and RRM		 75
4.1	Introduction	75
4.2	System Model	77
4.2.1	Load-based Pricing	78
4.2.2	Interference-based Pricing	79
4.3	Dynamic Programming Formulation	82
4.4	Neuro-Dynamic Programming Formulation	84
4.4.1	Approximation Architecture	85
4.4.2	TD(0) Learning Algorithm	87
4.4.3	Exploitation and Exploration	88
4.5	Numerical Results	89
4.5.1	Normal Traffic Load	92

Contents

4.5.2	Heavy Traffic Load	94
4.5.3	Effects of Exploration	96
4.5.4	Effects of a Price Sliding Window	97
4.6	Conclusions	98

Chapter 5. Cooperative Resource Allocation Games in Shared Networks 101

5.1	Introduction	101
5.2	Network Sharing Models	105
5.3	System Model	106
5.3.1	Uplink Load Factor	107
5.3.2	Downlink Load Factor	108
5.4	Cooperative Game Theory Framework	109
5.5	Resource Bargaining in WCDMA	112
5.5.1	Symmetric Bargaining	112
5.5.2	Asymmetric Bargaining	114
5.5.3	Revenue Optimisation	117
5.6	Resource Sharing Among Competing Operators	118
5.7	Numerical Analysis and Discussions	120
5.8	Conclusion	123

Chapter 6. Conclusion 125

6.1	Summary	125
6.2	Potential Directions for Further Research	128
6.2.1	Online Dynamic Pricing Schemes	128
6.2.2	Stochastic Pricing and Resource Allocation Games	129

6.2.3 Noncooperative Implementation of the Cooperative Bargaining Solutions in Self-organising Networks	130
Appendix A. Appendix	133
A.1 Infinitesimal Generator	133
Bibliography	135

List of Figures

1.1	A typical hourly arrival pattern in a mobile network.	2
1.2	Radio Resource Management model consists of call admission control, rate control and power control.	4
1.3	An Economic Framework for Pricing and Radio Resource Management. . .	13
2.1	The Nash, Raiffa-Kalai-Smorodinsky and utilitarian bargaining solutions are special instances on the Pareto-optimal boundary. They can be obtained by varying the value of β in the preference function.	27
3.1	Integrated call admission control and dynamic pricing model.	44
3.2	The probability distribution function of WTP using variable shapes β . . .	46
3.3	Transition Diagram of state $s = (\mathbf{x}, \mathbf{n})$	50
3.4	WTP pdfs and access probabilities vs. price with $\beta = 3.5$ and 1.0	62
3.5	Optimal revenue per stage J^* for WTP shape $\beta = 3.5$	63
3.6	Optimal call admission policy of OCADP. The x and y axes indicate the state of service 1, $s_1 = (n_1, x_1)$, and 2, $s_2 = (n_2, x_2)$, respectively. The combination of s_1 and s_2 forms a state $s = (s_1, s_2)$ in the system. The CAC policy of each state is indicated by the following symbols: 'o' represents $\mathbf{u}_c = (0, 1)$ and ' \otimes ' represents $\mathbf{u}_c = (0, 0)$. States with ' \square ' are states where the bandwidth of the system is fully utilised. In other states, $\mathbf{u}_c = (1, 1)$. . .	64
3.7	Optimal differential reward rate, $h^*(s)$	65

List of Figures

3.8	Optimal prices per bandwidth time for OCADP with $B = 20$, $\psi = (1, 1)$ per bandwidth time and $\mathbf{b} = (2, 2)$. The x and y axes indicate the state of service 1, i.e. $s_1 = (n_1, x_1)$, and 2, i.e. $s_2 = (n_2, x_2)$ respectively. The combination of s_1 and s_2 forms a state.	66
3.9	Stationary probabilities of a system with $B = 20$ and $\mathbf{b} = (2, 2)$. The x and y axes indicate that the state of service 1, $s_1 = (n_1, x_1)$, and 2, $s_2 = (n_2, x_2)$, respectively. The intersection of s_1 and s_2 forms a state $s \in \mathcal{S}$. OCADP controls network congestion by shifting stationary probabilities from states with high number of busy connections to less busy states.	68
3.10	Stationary probabilities of a system with $B = 20$ and $\mathbf{b} = (2, 2)$. OCADP controls network congestion by shifting stationary probabilities from states with high number of busy connections to less busy states.	69
3.11	Optimal reward per stage J^* under OCADP (with and without price discrimination).	71
3.12	Optimal reward per stage J^* for WTP shape $\beta = 1.0$	72
4.1	The interference level in the network, indicated by the system noise rise, rises exponentially as the system load factor increases.	80
4.2	The noise rise factor, i.e. the interference generated by a call, increases exponentially as the system load factor increases although the individual load factor remains the same throughout.	80
4.3	(a) A general feature-based approximation architecture. (b) An approximation architecture with feature vector $\boldsymbol{\theta}_k$ and future reward rate.	85
4.4	Average reward per time, \tilde{J}_k , and parameter vector, $\boldsymbol{\theta}_k$, averaged for every 250 steps, under O-NRF.	91
4.5	Average reward and proportion of reward obtained under normal load.	92
4.6	The relationship between ILF/NRF and System Load Factor	93
4.7	Blocking, dropping and access probabilities under normal load.	94
4.8	Average load factor and noise rise under normal load.	95

4.9	Average reward and proportion of reward obtained under heavy traffic load.	95
4.10	Average load factor and noise rise under heavy traffic load.	96
4.11	Average reward and proportion of reward obtained under policies O-ILF and O-NRF with 20% to 80% state exploration.	97
4.12	Results under heavy traffic load using a price sliding window.	99
5.1	Resource allocation games in a shared network.	103
5.2	Models of network sharing.	105
5.3	Geometrical interpretation of the symmetric and asymmetric Nash and Raiffa solutions.	121
5.4	Symmetric and asymmetric bargaining solutions with $\hat{\tau} = (0.5, 0.5)$ and $\hat{\tau} = (0.70, 0.30)$ respectively.	122
5.5	Operator's revenue using $\beta = 0, 1$ for varying bargaining powers.	122
6.1	Summary of the contributions of this dissertation and key distinctions with existing works.	126
6.2	Modified feature-based approximation architecture with parameter estimation.	129

List of Tables

2.1	Summary of the use of pricing in wireless networks.	36
3.1	System parameters derived using stationary distribution of the system. . .	60
3.2	Simulation parameters for AASP, OCASP, AADP and OCADP policies. . .	62
4.1	Simulation parameters for S-ILF, S-NRF, O-ILF and O-NRF policies. . . .	90

Chapter 1

Introduction

The high demand for information exchange, evident since the introduction of first generation analog cellular networks in the early 1980s, has accelerated the development of wireless communication systems. These analog systems were based on Frequency Division Multiple Access (FDMA). In the early 1990s, second generation (2G) cellular systems based on digital technology were introduced. The deployment of 2G was a huge success story because of the revolutionary technology, high-quality speech services and global mobility that it provided. The Global Standard for Mobile Communications (GSM), based on Time Division Multiple Access (TDMA), was deployed in Europe in 1992. In the United States, there are two digital standards, i.e. TDMA-based IS-54 and narrowband Code Division Multiple Access (CDMA)-based IS-95. Personal Digital Cellular (PDC) was also introduced in the 1990s and mainly used in Japan. However, 2G is primarily designed for voice communication and as a secondary feature, it provides circuit-switched data services but only at low data rate. There was an obvious need for a greater data capacity.

The transition of the 2G to the much-hyped third generation (3G) has begun and taken off around the world in the past few years. 3G systems such as Wideband Code Division Multiple Access (WCDMA) and cdma2000 can provide high data rates up to 2 Mbit/s and support a broad range of multimedia services including voice, data and video to mobile users. Although there are as yet no solid specifications on the beyond 3G (B3G) or fourth generation (4G) technology, it is clear that 4G will support higher data rates

than 3G and will efficiently integrate different modes of wireless communications. Data rates in 4G systems are expected to be as high as 20 Mbit/s [11, 51].

The emergence of applications with very different throughput, loss rate and delay underscores the need for a network capable of supporting different levels of quality of service (QoS). Since radio spectrum is a scarce resource, efficient Radio Resource Management (RRM) is one of the most important and challenging engineering issues in 3G and 4G mobile communication systems. Researchers have been focusing on jointly or singularly optimising parameters such as capacity, modulation scheme, coding scheme, transmission power and bandwidth. The suggested solutions invariably involve installation of new infrastructure. For example, due to the limitations of the radio spectrum, micro- or pico-cellular architectures are used to provide a higher capacity [130]. However, handoffs occur more frequently in these micro- and pico-cells due to their small coverage areas.

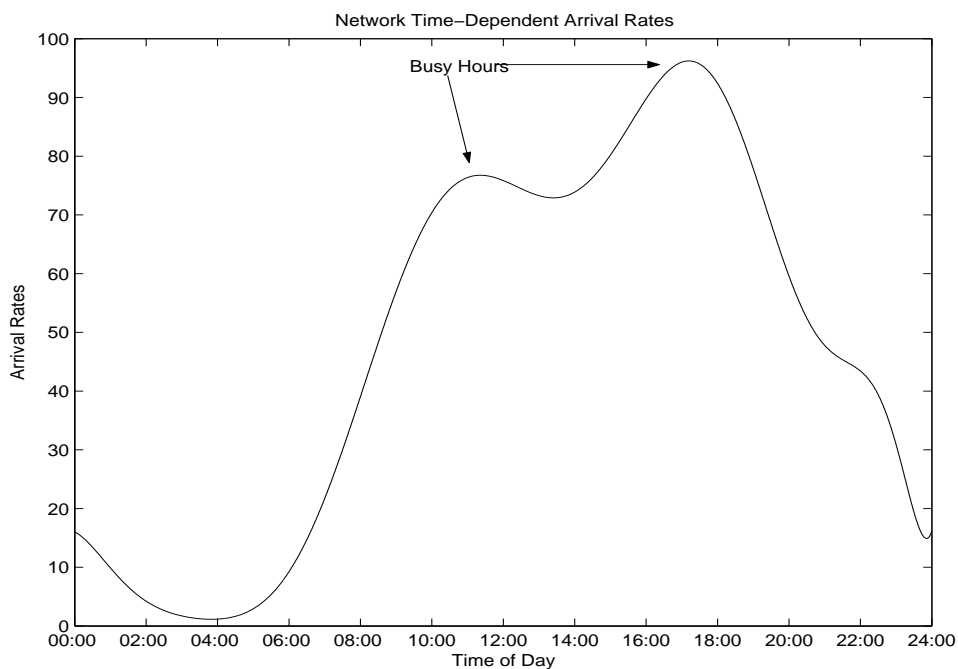


Figure 1.1. A typical hourly arrival pattern in a mobile network.

In order to satisfy high resource demand, it is a common strategy for network designer to *over-engineer* a network or cell so that the probability of users being blocked is conservatively minimised during the busiest hours. This means that the ever-increasing peak traffic requirements during these busiest hours (as shown in Fig. 1.1) will trigger a network upgrade to boost capacity from time to time. It is clear that provisioning a

network always to meet peak demand, which is only typically several hours in a day, is very costly for network operators. The vast differences between peak and off-peak demand patterns mean that network resources will be under-utilised outside of these peak hours and that idle capacity is expensive to maintain. A strategy is needed to both limit peak demand and increase utilisation during off-peak hours. This has in turn encouraged researchers to investigate alternative methods for optimal use of the available resources and to contend with the considerable costs involved in obtaining spectrum and site licences as well as upgrading infrastructure to boost network capacity.

In this dissertation, we offer a new economic perspective into the important issues of efficient management of scarce network resources and congestion control in wireless networks. Our research is multidisciplinary and combines concepts from telecommunications, economics, decision theory, queueing theory, artificial intelligence and game theory to address these issues. The rest of this chapter is organised as follows. In Section 1.1, we discuss the background of this research. In particular, we will analyse existing RRM schemes and pricing models for wired and wireless networks and highlight the need for improvements. In Section 1.2, we introduce our proposal for an economic framework for pricing and RRM. Finally in Section 1.3, we give an overview on the organisation of this dissertation and our original contributions.

1.1 Background and Motivation

1.1.1 Radio Resource Management

Congestion control measures can be either *preventive* and *reactive*. Preventive congestion control procedures strive to deflect congestion before it occurs by avoiding traffic patterns that can lead to congestion. Reactive methods will only be triggered to control congestion when it occurs. A good congestion control strategy should consist of both methods. RRM plays a major role in congestion control and QoS provisioning for wireless communication systems. The family of RRM algorithms can be classified into three categories: Call Admission Control (CAC), which manages admission of new and handoff arrivals; Rate Control, which determines users' transmission rate; and Power Control, which controls

1.1 Background and Motivation

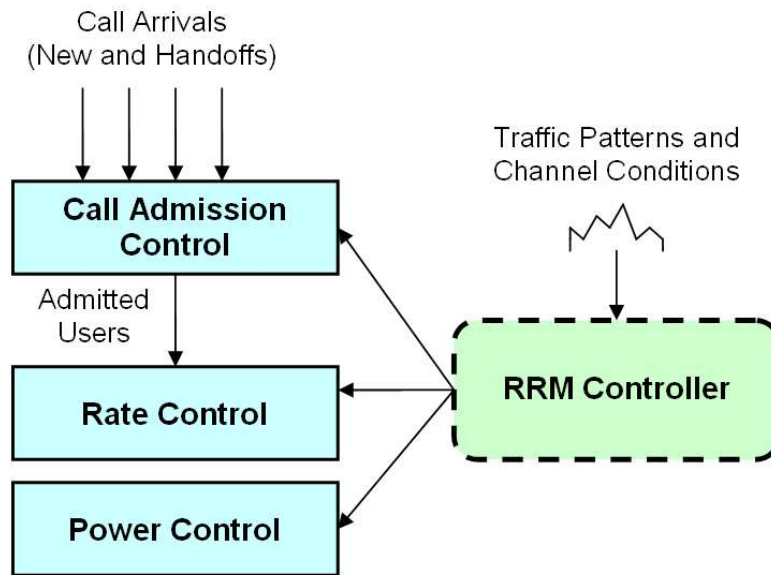


Figure 1.2. Radio Resource Management model consists of call admission control, rate control and power control.

the transmission power of the mobile stations (MSs) and the base station (BS). We depict these functions and the general RRM model in Fig. 1.2.

CAC determines whether to admit or reject a call upon its arrival and is one of the most important aspects of RRM. The main objective is to admit as many calls as possible to achieve high utilisation while maintaining the QoS guarantees of ongoing calls. QoS is guaranteed in terms of signal quality and call dropping probability. In interference-limited wireless networks such as WCDMA, CAC admits new users only if the minimum signal quality and transmission rates can be maintained for existing users. CAC schemes can provide preventive and/or reactive congestion control [1]. In preventive CAC, admission or rejection is based on some assessment of the QoS constraints. In reactive CAC, all users are admitted and transmission will only begin after some QoS measurements at the beginning of the call.

There is a considerable literature on CAC schemes, most of which can be classified into *number-based*, *interference-based* and *throughput-based*. In number-based CAC, any call will be blocked if N users are connected. Setting an effective CAC threshold for N can be based on the QoS requirement for the upper bound on packet error probability [111] or tolerable interference level [53]. Since the capacity in a WCDMA system varies its interference level, this scheme is generally inaccurate and nonadaptive [130].

In interference-based CAC, a new call is blocked if the interference level exceeds a pre-determined level. This scheme is applied on a call-by-call basis. The admission of a new user can gracefully degrade the performance of other users in the system. Therefore, this scheme requires more overheads in terms of hardware and computation. It also relies heavily on the reliability of real-time interference measurements. Examples of this scheme can be found in [7, 8, 54]. In throughput-based CAC, a call is accepted or blocked if the additional load introduced by the new user does not result in the system load exceeding some pre-determined threshold. In WCDMA, cell loading is represented by the system load factor on the uplink and downlink (see definition in Section 2.1) [49].

A good CAC scheme needs to balance the call blocking of new users and call dropping of handoff users in order to provide the desired QoS requirements. Specifically, the denial of service to new calls is better than the unreliability of service to existing calls. Therefore, handoff calls receive less stringent admission criterion than new calls. Various handoff-based CAC schemes have been proposed and their strategy for minimising call dropping can be classified into [33, 75]: guard channel and queueing priority schemes. In guard channel schemes, some channels are reserved for handoff calls. A new call is rejected whenever the number of free channels is less than the guard level. Although this scheme reduces the probability of dropping a handoff call, it may also result in under-utilisation of system resources. In queueing priority scheme, all calls are accepted whenever there are free channels. When all channels are busy, either the new calls, handoff calls or all calls are queued.

Rate control and power control in wireless networks are two inter-related problems. The achievable data throughput rate depends on the transmission power and propagation loss of the channel. In FDMA/TDMA-based systems, power control is motivated by the need to manage co-channel interference, which is caused by frequency reuse due to limited available frequency. In CDMA-based systems like WCDMA, power control is needed to remove near-far effect on the uplink, mitigate fading and compensate changes in propagation conditions. On the uplink, power control should make signal powers from different users nearly equal at the base station in order to maximise the total capacity in a cell. Without power control, all users transmit signals to the base station with the same transmission power, without taking into account the propagation loss due to multipath fading and shadowing. The signals of the users that are closer to the base station will cause significant interference to the signals of other mobiles located further away. On the

1.1 Background and Motivation

downlink, the total transmission power of the base station should be kept at the minimum required level in order to decrease the interference to users in other cells.

Power control schemes for CDMA-based systems have been extensively surveyed in [82]. In general, power control schemes can be categorised, among other criteria, according to their quality measure, power update frequency and strategy, whether they are centralised or decentralised and closed- or open-loop. Signal quality is very subjective and can be measured in terms of the signal strength, signal-to-interference ratio (SIR) and bit error rate (BER). In strength-based systems, the base station will give commands to increase or decrease transmission power to the users based on the signal strength received. This scheme is easy to implement using signal quality measurements. However, the problem of *positive feedback* might occur in SIR-based systems. Positive feedback occurs when a user is under the instruction to increase its transmission power in order for the base station to receive a desirable SIR. This results in extra interference to other users, who in turn also increase their transmission power and cause SIR deterioration for all users. BER is a better measure of signal quality than SIR because of the time-varying nature of the latter in real systems.

Even with perfect admission control, congestion might happen due to a variety of factors such as the deterioration of the wireless environment, the mobility of users, users' activity and power control imperfection [130]. As the number of users increases in a cell, excessive interference causes traffic loss and deterioration in signal quality. When congestion occurs, the RRM controller can either drop some calls; decrease the transmission rate of all or some users; or reduce the number of simultaneous transmissions. The main problem is that these congestion control mechanisms are all reactive and will only be instantiated when congestion occurs. It is clear that current admission, rate and power control schemes do not provide sufficient preventive congestion control measures. When the network load is extremely high during the busiest hours of the day, no matter how RRM parameters are adjusted, these schemes cannot guarantee QoS to users because no incentives are provided for users to regulate and adapt their consumption according to the network conditions [50]. Instead, more focus needs to be placed on providing *negative incentives* for users to defer their usage to avoid or relieve congestion when network load is high; and *positive incentives* to encourage arrivals and resource consumption when the network is lightly loaded. These incentives can be provided in the form of pricing.

1.1.2 Economic Modelling

Economic modelling offers a new perspective into RRM and enables efficient utilisation of scarce resources and congestion prevention. It is based on the notions of utility functions, congestion pricing, Pareto optimality and game theory. Formally, economics is concerned with the production, sale and purchase of commodities, and with how consumers and producers interact in a market [116]. In general, economics is a social science which studies human activity involved in meeting needs and wants. The Father of Economics, Adam Smith, once defined economics as “the science of wealth” and “the science related to the laws of production, distribution and exchange”. Economics can be divided into two main branches: *microeconomics* and *macroeconomics*. Microeconomics, which is more relevant to our research, examines the behaviour of individual entities such as businesses and individuals in order to understand the decision making in the face of scarcity and the allocative consequences on these decisions. Macroeconomics offers a higher-level perspective and examines economy at an aggregated state, national and international level.

Communication resources such as bandwidth, code and power can be viewed as traditional economic goods and well-established economic ideas can be applied. These resources are non-storable goods and if not utilised, will be wasted and are non-recoverable. When one additional user is admitted into the network, he/she causes QoS degradation to other existing users. In economic terms, this phenomenon is called a *congestion externality*, i.e. the congestion, delay and cost of exclusion a user imposes on others [68]. Wireless resources are limited and scarce. In fact, scarcity is central to the economic problem. Scarcity is a situation where people’s wants exceed their resources and implies that people must make a choice – to forgo one thing in favour of another. Price is a measure of relative scarcity, i.e. of supply relative to demand. When the resource becomes relatively scarce, because supply constricts or demand expands, the price increases. This restricts the supply, making it available to those willing to pay the higher price. Therefore, price is an effective arbitration mechanism to affect user behaviour and control congestion, especially in the face of scarcity.

With different prices associated with different levels of scarcity, users have to analyse the trade-off between convenience and money. For example, when the price is relatively high, users can delay their calls and pay less when network resource becomes less scarce. Otherwise, they can pay more to access the network immediately out of convenience.

1.1 Background and Motivation

As Adam Smith once stated, these users are simply “paying the difference out of their convenience”. The model economists use to explain how prices are determined in a market economy is called the *supply and demand model*. For example, exchange rates in the currency market are determined by the relative supply and demand of different currencies. Supply is a positive relationship between the price of and the quantity supplied of the good by producers, higher prices give producers more incentive to produce. Demand is a negative relationship between price and the quantity people will buy at each price. Demand is also defined as the utility maximising choice of a consumer when constrained by price. The price and quantity demanded are negatively related, i.e. consumers will demand less of a good if its price increases. When producers and consumers interact in a market, the equilibrium price is at the intersection of the supply and demand curves. At this point, the quantity producers are willing to supply at the current price equals the quantity consumers demand at that price, resulting in no shortage or surplus in resources.

Network user preferences, which determine demand at various prices, may be modelled using utility functions, which describe how sensitive users are to changes in network performance. A utility is defined as the benefit or satisfaction that a person gets from the consumption of goods and services. Utility function is usually a non-increasing function of the amount of consumption. Higher values of utility indicate increased satisfaction and the consumer’s objective of maximum satisfaction. Utility can be used to indicate users’ level of satisfaction numerically and may be modelled using users’ willingness to pay, amount of allocated resources, the perceived blocking probability or a combination of QoS parameters in general [23], [50]. The concept of utility function is also used in Pareto maximisation. Pareto efficiency, or Pareto optimality, is a central theory in economics with various applications in game theory, engineering and social science. An allocation is Pareto optimal if there is no wasted utility, i.e. it is impossible to make any one party better off without making any other worse off. A formal overview on game theory will be provided in Section 2.2.1 of Chapter 2.

Consumers always try to minimise their expenditure for a given level of utility or maximise their utility for a given budget. These problems are as the *expenditure minimisation problem* and *utility maximization problem* respectively. A producer, like a network operator, is a supplier of different services. Profit is the difference between the prices at which these services can be sold and the cost of production. On the other hand, total *revenue* is the total number of dollars the producer receives from customers who purchase

its products. Producers aim to maximise their revenue, a problem known as the *revenue maximisation problem*. Communications services are usually costly to produce but cheap to reproduce [26]. When the producers aim to maximise the aggregate utilities of all consumers, the problem is known as the *social welfare maximisation problem*.

1.1.3 Pricing and Charging Models in Communication Networks

Billing, charging and accounting are the most critical operational support activities conducted by network or service providers. Based on the 3rd Generation Partnership Project (3GPP) standards, charging is the process of collecting information about chargeable events; accounting is the process of determining revenue sharing among operators in cases of roaming and other services; and billing is the process of employing specific pricing policies and issuing bills for the users [30]. In this research, our main focus is on charging models. The role of pricing can be analysed from two perspectives. From the economic perspective, it enables the network operator to recover its investments and ongoing costs; and make profits to finance future capacity expansion. It is also an important marketing tool for the provider to attract new subscribers. Secondly, and perhaps more interestingly for engineers, pricing is a mechanism for traffic shaping and congestion control that can efficiently allocate and influence the way users utilise scarce network resources. Pricing policies can be *static*, in which prices are independent of the current network utilisation or *dynamic*, in which prices fluctuate as a result of some network conditions. Both assume that the network operator has mechanisms to set prices, and to perform accounting and billing for network usage on a per-user basis.

The most dominant charging model used in the mobile telecommunications is the static, flat-rate model. Users are mainly billed based on their type of subscription and other parameters such as call duration, type of communication and location of destination. Such widespread use of flat-rate pricing is attributed to the fact that most mobile service providers emerged from the traditional fixed-line, telephony world where distance- and time-based billing are common [30]. However, with almost all users paying a flat-rate charge, they have little incentive to manage their consumption. Other static charging models proposed for wired networks include time-of-day pricing, Paris-Metro pricing [83], priority pricing [22, 23], reservation-based pricing [87, 88], edge pricing [104] and usage-based pricing based on effective bandwidth [25, 26]. We will elaborate more on these

1.1 Background and Motivation

schemes in Chapter 2. We also refer our readers to [27] for a comprehensive comparisons of pricing schemes for broadband IP network.

Under such static models, users act independently and selfishly without considering the current network conditions. If price is set too low, self-serving users will tend to over-use and cause congestion. On the other hand, too high a price will discourage users from accessing the network. In addition, flat-rate pricing models are inadequate in providing QoS choices to users, especially in the event of congestion. For example, some users (e.g. business users) are willing to pay more to always have access to the network and also to receive high signal quality at all times. However, these users might be blocked by admission control during congestion because current static pricing models are not adaptive to user preferences and network conditions. Such lack of choice translates into a loss of opportunity for operators to generate additional revenue and it restricts the utility derived from a given network capacity.

Dynamic pricing models adjust prices according to the demand pattern and congestion level in a network. Monetary incentive can influence the way users utilise resources, so that important resources are not wasted by users who do not value them enough. They offer flexibility to react to fluctuations in the incoming traffic in order to achieve some system performance objectives. This enables the tracking of the optimal prices to be charged by a network at a given time in order to achieve some system objectives such as revenue and social welfare. However, these schemes are in general more complex than static pricing schemes and expensive to implement. Perhaps the absence of dynamic pricing models might be attributed to the lack of sophisticated billing and accounting mechanisms to support such models. Dynamic pricing models for wired networks, such as the Internet, can be categorised into three main approaches: auction-based [63,68,117], shadow pricing [38,61] and stochastic control [89,90]. These pricing schemes will be outlined in detail in Section 2.2.2 in Chapter 2.

In the auction-based approach, the network or service provider allocates resources according to users' bids. For example, in smart-market pricing [68], each packet contains a bid in its header and will only be serviced by routers if the bid exceeds a threshold. The shadow pricing approach, first proposed in [61], involves solving a social welfare maximisation problem and the Lagrangian multiplier of the optimisation problem can be interpreted as the *shadow price*. When resource allocation is decentralised, the system is

optimal when the users' demand, based on the shadow price, coincides with the network's optimal choice of allocation. Both the auction-based and shadow pricing approaches are designed to optimise resource allocation for a fixed number of users in the system and only provide reactive congestion control. The final approach is different. It makes pricing decisions based on the stochastic attributes of the network traffic and is both preventive and reactive in terms of congestion control. The problem is modelled as a Markov decision problem and the outcome, i.e. a congestion-dependent pricing policy, is obtained using dynamic programming techniques.

In the case of wireless networks, the bulk of the pricing literature is motivated by power control. There are two main approaches: shadow pricing [70,108,109] and noncooperative power control game with pricing [32,66,67,71,98,99,126,131]. Similarly to [38,61], the first approach computes a shadow price to achieve an optimal allocation that maximises the social welfare of all users. In recent years, there has been an increasing trend to use game theory to examine resource allocation problems. In the noncooperative game-theoretic approach with pricing, resource allocation is decentralised and users determine their own transmission power. The common approach is to first define a suitable user's utility function for the problem. The utility functions proposed are the throughput per terminal life [32, 71, 98, 99], the sigmoid function of SIR [66, 126] and the step function of Signal-to-Interference-plus-Noise-Ratio (SINR) [67, 131]. Users then enter into a decentralised, noncooperative game to select the transmission power that maximises their utility. The outcome of the game is known as the Nash Equilibrium, which is not optimal (or inefficient) due to the lack of cooperation among the self-serving users. In order to achieve an outcome that gives Pareto improvements, a usage-based price per transmission power is incorporated into the models.

We summarise the drawbacks of these schemes as follows. Firstly, price only acts as an internal mechanism and does not reflect the actual charges that users pay. It is a reactive congestion control measure and only provides negative incentive to discourage selfish usage of network resources such as power. If users are actually charged according to a different pricing model, it is not clear whether they will react to the price per transmission power announced by the base station. Furthermore, the price in the users' utility functions is only a static parameter that is optimised for a fixed number of users in the network. In reality, the network providers can make use of their knowledge and historical data of the stochastic nature of users' traffic to develop an effective scheme. Nothing is mentioned

1.2 An Economic Framework for Pricing and RRM

about how to determine the optimal price as more users arrive to the network. In these schemes, the important role of pricing to provide positive incentive to encourage usage when the network is lightly loaded has been ignored. Moreover, the outcome of the noncooperative games, i.e. the Nash equilibrium is well-known to be inefficient. Even with the introduction of prices, which provides some Pareto improvements, a socially optimal outcome cannot be achieved. The resulting degree of efficiency loss is known as the price of anarchy. For example, selfish behaviour of users leads to an efficiency loss of up to 25% in a network with inelastic supply [57]. These noncooperative solutions also burden users' mobiles, which have limited battery power, with additional complex computations. The crucial issue of fairness in resource allocation has also been ignored in these works.

1.2 An Economic Framework for Pricing and RRM

Up to this point, we have outlined current RRM models and discussed the role of pricing in congestion control and helping network operators to improve revenue. Existing RRM schemes focus on relieving congestion when it occurs and fail to provide the necessary incentives to discourage resource consumption before the network becomes relatively congested or encourage consumption otherwise. In this section, we will introduce our proposal for an economic framework for pricing and RRM. This integrated model is illustrated in Fig. 1.3. We propose a two-tier model to

- calculate an optimal integrated dynamic pricing and RRM policy in order to influence the demand pattern of incoming users, minimise handoff call dropping and maximise operator's revenue; and
- allocate scarce resources to admitted users in a fair and Pareto-optimal manner using bargaining concepts from cooperative game theory.

We compute the optimal integrated policy based on the stochastic distribution of users' willingness to pay, arrival and departure using dynamic programming and neurodynamic programming (NDP). The former technique is suitable for smaller networks such as the fixed-capacity, cellular network that we will consider in Chapter 3. We then extend this concept to an interference-limited WCDMA cell in Chapter 4. This problem has

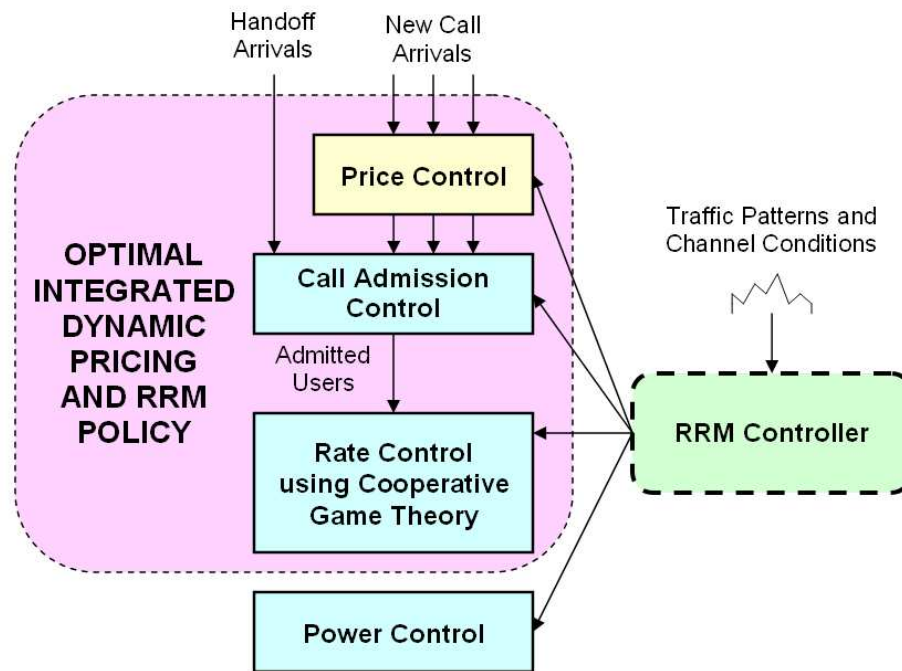


Figure 1.3. An Economic Framework for Pricing and Radio Resource Management.

a large state space and simulation-based methods such as NDP is more suited to such problem. Unlike the pricing proposals for wireless networks discussed in the previous section, the calculated state-dependent dynamic prices will reflect the actual admission price that users will pay. Based on the prices advertised by the base station when users make a call, they can choose the service class that best suits their expectation in terms of QoS and cost of access. In our models, to minimise accounting and billing overheads, we assume that once users are granted access, their admission price will be honoured by system throughout their call. The final call charge still depends on the amount of resources utilised by the users. Of course the use of dynamic pricing models does not always mean the users will end up paying more. Price-sensitive users can defer their calls until the network load is relatively low and benefit from lower prices.

Efficient allocation of scarce resources after users are admitted into system is an equally important problem that will be dealt with in Chapter 5. There is an increasing trend to apply game theory in various power control and resource allocation problems (a survey is available in [4]). The noncooperative game model, which is the most common in existing literature, leads to a solution that is not Pareto-optimal and in some cases unfair. In order to achieve an optimal operating point, arbitration is necessary [19]. In this research, we shift the focus to cooperative game theory. Cooperative game theory provides

1.3 Organisation and Original Contributions

an excellent framework for this resource allocation problem, in that it considers fairness by means of the axiomatic properties a solution must possess and efficiency through the notion of Pareto optimality. The notion of axiomatic bargaining in cooperative game theory provides a good analytical framework to derive a desirable operative point that is fair and Pareto-optimal. An allocation is Pareto-optimal if there is no wasted utility, i.e. it is impossible to make any one party better off without making any other worse off. Such an outcome is said to be efficient.

1.3 Organisation and Original Contributions

1.3.1 Chapter 2: Background

In Chapter 2, we discuss the technical background of this research in WCDMA system and radio resource parameters; economic modelling concepts of game theory; and stochastic decision theory concepts of dynamic programming and NDP. We will also review the vast number of existing static and dynamic pricing proposals for wired, broadband networks such as the Internet. Then, we will analyse the existing power-control-motivated pricing proposals for wireless networks. Dynamic programming is used in Chapter 3, NDP in Chapter 4 and game theory in Chapter 5.

1.3.2 Chapter 3: Integrated Call Admission Control and Dynamic Pricing Cellular Networks

In Chapter 3, we present our model of optimal integrated call admission control and dynamic pricing for fixed-capacity networks. Our model captures the price-affected behaviour of users; considering the effects of price on users' arrivals, retrials, substitutions and departures. We show, via the computation of optimal policy and average reward using dynamic programming, that the performance of the integrated policy is superior to other conventional policies that consider call admission control and dynamic pricing as separate problems. Due to the limitations of the dynamic programming method, this model is designed for small, fixed-capacity cellular systems. We summarise the original contributions of this chapter as follows:

- **Bandwidth reservation for handoff calls based on satisfaction revenue**

It has been widely accepted that handoff call requests have higher priority than new calls since premature termination of a call is less desirable than a rejection in the first place. In our model, we capture the importance of handoff calls by associating a pseudo revenue called satisfaction revenue. Satisfaction revenue should be set higher than the actual revenue from admitting a new call so that handoff calls have higher admission priority than new calls.

- **Price-affected arrival model with retrial, substitution and abandonment**

We propose an advanced arrival model that incorporates retrials, abandonments and substitution effects among services and through time. When users are blocked or have insufficient budget to begin a call, they will remain in the *orbit* of the system until they receive service, use another service as a substitute or abandon their intention completely. The queueing model is found to have a level-dependent quasi-birth-death [16] structure and we derive various system parameters through its stationary distribution.

- **Reduction of computational cost using non-discriminatory pricing scheme**

Price discrimination refers to the practice of varying the price of a product (i.e. bandwidth in our case) between users or applications to improve revenue. For example, the price per kbit of a voice call differs from that of an SMS. However, when price discrimination is avoided, the computational cost of the optimal policy is reduced. This is because the optimal price control only has to compute one price, instead of J prices for J services. We will show that an integrated policy with non-discriminatory pricing closely approximates that of one which enforces price discrimination.

1.3.3 Chapter 4: Interference-based Radio Resource Management and Dynamic Pricing for WCDMA Networks

In Chapter 4, we extend our proposal to soft-capacity, CDMA-based systems. We will present our model of optimal interference-based radio resource management and dynamic pricing for soft-capacity. We formulate the integrated model as an NDP problem and solve

1.3 Organisation and Original Contributions

for the optimal policy using a simulation-based temporal difference algorithm. The dynamic programming method that we have used in Chapter 3 is unsuitable for this problem due to its large state space. We will prove that this *interference-based* scheme provides improvement in terms of congestion control and efficient utilisation in an interference-based system over schemes that charge users based on their bandwidth usage. We will also study the effects of heavy traffic load, action exploration and the introduction of a price sliding window to reduce the size of the price control space on our results. We summarise the original contributions of this chapter as follows:

- **Interference-based dynamic pricing and noise rise factor**

Existing pricing schemes for wireless networks charge users based on the amount of bandwidth used in terms of the transmission rate used. However, in interference-based networks such as WCDMA, the relationship between one's transmission rate and the interference imposed on others as the system loading reaches its threshold is *not* linear. We introduce a parameter called *noise rise factor* as a basis for setting the price. This parameter quantifies the expected amount of interference generated by a call based on users' expected transmission rate, service type and the current system load. As we will prove in Chapter 4, interference-based pricing scheme is better than load-based pricing in terms of congestion control and reward improvement.

- **Optimal integrated dynamic pricing and RRM with QoS guarantee**

Our work is the first to analyse conventionally separate but inter-related problems of pricing and RRM in terms of call admission control and rate allocation as a joint problem. Our model incorporates QoS guarantee by requiring users to select their minimum and maximum transmission rates during admission and allow the network operator to vary their rates within these upper and lower bounds throughout the call.

- **Neuro-dynamic programming (NDP) method with action-based approximation architecture**

Using NDP, the design of a suitable approximation architecture is highly-dependent on the nature of one's problem. The essence of problem is captured using *features* that closely approximate the actual reward function. In this problem, we propose to have the features of the problem dependent on the price and RRM actions. The

motivation for this proposal stems from our experience of using dynamic programming in Chapter 3, where state transitions depend on the actions that are chosen by the optimal policy.

1.3.4 Chapter 5: Cooperative Resource Bargaining Games for Shared Networks

Chapters 3 and 4 focus on the computation of an optimal integrated policy to manage incoming arrivals such that the long-term average reward of the operator is maximised. In Chapter 5, we apply the bargaining theory from cooperative game theory to deal with the issue of radio resource allocation for connected users. The aim of this work is to devise a strategy for network operator to allocate resources in a fair and Pareto-optimal manner to its users and to smaller providers known as Mobile Virtual Network Operators (MVNOs). The MVNOs purchase resources from the network operator in order to resell them to their own subscribers. Our work is the first comprehensive treatment of resource allocation in shared networks. We summarise the original contributions of this chapter as follows:

- **Symmetric bargaining: fair and Pareto-optimal resource allocation**

Conventional works have only considered specific bargaining solutions (see Section 2.2.1) such as the Nash bargaining solution [76]. Other bargaining solutions on the Pareto-optimal boundary such as the Raiffa-Kalai-Smorodinsky solution have been relatively ignored. Based on the preference function concept proposed by Cao [18], we derive the explicit formulas of a class of bargaining solutions between Nash and Raiffa-Kalai-Smorodinsky. The benefit of having such closed-form solutions is immense, enabling easy implementation and avoiding the use of complex derivation algorithm. The solutions derived are fair and Pareto optimal according to the users' minimum and maximum resource requirements.

- **Asymmetric bargaining: resource allocation based on bargaining powers**

The symmetric bargaining model enforces fairness according to users' minimum and maximum resource requirements. In this model, we allow the amount of resource allocated to depend on the price paid by the users. For example, the price can be the admission price derived in Chapters 3 and 4. However, if the price is allowed to

1.3 Organisation and Original Contributions

vary during a call, users can submit bids to the operator to influence the bargaining outcome. An important significance of this asymmetric bargaining model is that, unlike conventional auction models, the outcome is still Pareto-optimal and fair according to the users' bargaining power. This asymmetric model also provides opportunity for the operator to select a bargaining solution out of the class of solutions derived to maximise its revenue.

- **Cooperative resource sharing in multi-operator networks**

In networks shared by more than one licensed 3G operator, operators can benefit from temporary resource exchange due to dissimilar usage patterns and non-coincident peak demand. In our model, we classify operators according to their resource surplus or deficit. The model keeps a history of the amount of resources they have obtained in the past when there has been a deficit and contributed when there has been a surplus. Operators who have contributed more in the past will therefore be allocated a bigger share of the common pool of resources when deficit occurs.

1.3.5 Chapter 6: Conclusion

Finally, in Chapter 6, we present the conclusions and re-highlight the significance of our work. We will also suggest areas for some possible extension of our work.

Chapter 2

Background

In this chapter, we will provide some background on the concepts used throughout this dissertation. In Section 2.1, we will give an overview of WCDMA and define some resource usage parameters. Then in Section 2.2, we elaborate on the concepts of game theory and discuss in detail existing pricing schemes for wired and wireless networks. Finally in Section 2.3, we introduce the dynamic programming and neuro-dynamic programming used in Chapters 3 and 4 respectively.

2.1 WCDMA System

Wideband Code Division Multiple Access (WCDMA) has emerged as the most widely adopted third generation (3G) air interface technology [49]. WCDMA is a wideband Direct-Sequence CDMA (DS-SS-CDMA) system, in which user information bits are spread over a wide bandwidth by multiplying the user data with quasi-random bits (called *chips*) derived from CDMA spreading codes. An important advantage of WCDMA is its ability to support variable and very high bit rates of up to 2 Mbps through the use of variable spreading factors and multiple codes. The chip rate of 3.84 Mcps leads to a carrier bandwidth of approximately 5 MHz. DS-SS-CDMA systems with a bandwidth of about 1 MHz, such as IS-95, are commonly known as *narrowband* CDMA systems. Each user is allocated frames of 10 ms duration, during which the user's data rate is kept constant. However, the data capacity among the users can change from frame to frame.

2.1 WCDMA System

WCDMA supports four traffic classes: conversational (voice, video telephony); streaming (streaming multimedia); interactive (web browsing, network games); and background traffic (background download of emails, SMS, MMS). The speech codec in UMTS employs the Adaptive Multi-Rate (AMR) technique. The multi-rate speech coder in a single integrated speech codec with eight source rates: 12.2 (GSM-EFR), 10.2, 7.95, 7.40 (IS-641), 6.70 (PDC-EFR), 5.90, 5.15 and 4.75 kbps. The AMR bit rates can be controlled by the radio access network. Some of the modes are the same as in existing cellular networks to facilitate interoperability. The AMR speech codec is capable of switching its bit rate every 20 ms speech frame, which corresponds to 160 samples at the sampling frequency of 8000 samples per second, upon command.

Traditional definitions of capacity of networks are either related to the number of calls they can handle (pole capacity) or to the arrival rate that guarantees that the rejection rate (or outage) is below a given fraction (Erlang capacity). One of the main features of CDMA-based system is *soft-capacity*, where the cell capacity is not determined by available resources as in the case of TDMA. The capacity of a particular cell is not known exactly as it depends on the user density distribution, their movements and the propagation conditions in the cell and its neighbouring cells. As the situation in the cell changes dynamically, the instantaneous capacity varies, leading to QoS fluctuations. Capacity analysis has been considered in [122] for a voice-based CDMA system, in [34] for CDMA system with multirate sources, [48] for downlink WCDMA, [42] for multiservice WCDMA networks with variable GoS and [110] for WCDMA system with multimedia packet transmission.

2.1.1 Wireless Resource Parameters

Uplink

In order for a signal to be received on CDMA-based systems, the ratio of its received power to the sum of the background noise and interference must be greater than a given target [49, 109, 122]. On the uplink, the constraint is given by

$$\frac{W}{\nu_i x_i} \frac{g_i p_i}{\sum_{j \neq i}^N g_j p_j + I_{\text{other}} + \sigma^2} \geq \left(\frac{E_b}{N_0} \right)_i, \quad (2.1)$$

where W = WCDMA chip rate,
 ν_i = activity factor,
 N = number of users,
 x_i = allocated transmission rate,
 g_i = path gain between the base station and user i
 p_i = transmission power of i th user,
 σ^2 = background thermal noise power,
 I_{other} = other-cell interference, and
 $(E_b/N_0)_i$ = bit-energy-to-noise density ratio to meet a target BER level.

$g_i p_i$ is the received power at the base station from user i . The total interference received at the base station is denoted as I_{total} , where

$$I_{\text{total}} = \sum_{j \neq i}^N g_j p_j + I_{\text{other}} + \sigma^2. \quad (2.2)$$

The other-cell interference, I_{other} , can be taken into account by some constant f [123], i.e. $I_{\text{other}} = f \sum_{j=1}^N g_j p_j$. Interference coefficient f typically has values between 0.1 and 0.6 [49]. Therefore, the capacity of the cell depends on the load of the neighbouring cells. The cell has less capacity if the neighbouring cells have a high load.

Definition 2.1. *The individual uplink load factor of user i is defined as the ratio of one's received power at the base station with respect to the total uplink interference, i.e.*

$$\eta_i^{\text{UL}} = \frac{g_i p_i}{I_{\text{total}}} = \frac{1}{\frac{W}{\left(\frac{E_b}{N_0}\right)_i \nu_i x_i} + 1}. \quad (2.3)$$

When other-cell interference is considered, the individual uplink load factor becomes

$$\eta_i^{\text{UL}} = \frac{1 + f}{\frac{W}{\left(\frac{E_b}{N_0}\right)_i \nu_i x_i} + 1}. \quad (2.4)$$

Definition 2.2. *The uplink system load factor is defined as the sum of uplink individual load factors of all users in the cell, i.e.*

$$\eta_{\text{sys}}^{\text{UL}} = \sum_{j=1}^N \eta_j^{\text{UL}} < \eta_{\text{max}}^{\text{UL}} < 1 \quad (2.5)$$

The constraint on the uplink system load factor asserts that the uplink is interference-limited. Even when there are no power constraints on users' transmit power, they cannot

2.1 WCDMA System

increase their power without bound due to the interference they will cause others. In actual systems, the load threshold η_{\max}^{UL} must be significantly less than one, usually between 0.5 and 0.7, due to the limited transmission power of users, imperfect power control, path loss, shadowing and inter-cell interference. In radio planning, the cell loading can be translated into another parameter called noise rise or interference margin.

Definition 2.3. *The uplink noise rise is defined as the total received power to the noise power:*

$$\alpha_{\text{sys}}^{\text{UL}} = \frac{I_{\text{total}}}{\sigma^2} = \frac{1}{1 - \sum_{j=1}^N \eta_j^{\text{UL}}} = \frac{1}{1 - \eta_{\text{sys}}^{\text{UL}}}. \quad (2.6)$$

Downlink

On the downlink, WCDMA employs orthogonal codes to separate users. However, users will still see part of the signal as multi-access interference (MAI) due to multipath propagations and neighbouring cell transmissions. When the orthogonality factor of the i th user $\theta_i = 1$, signals are perfectly orthogonal. Typically, θ_i is between 0.4 and 0.9 [49]. On the uplink, transmission is asynchronous and therefore the signals are not orthogonal. The downlink load factor depends on the orthogonality factor and the maximum transmission power available at the base station. Assuming perfect power control [49, 108],

$$\frac{W}{\nu_i x_i \theta_i g_i} \frac{g_i p_i}{\sum_{j \neq i}^N p_j + \sigma^2} \geq \left(\frac{E_b}{N_0} \right)_i, \quad (2.7)$$

where θ_i is the orthogonality factor of the codes used in the downlink. The total transmission power to the users on the downlink is limited by the maximum power p^{\max} the base station can transmit, i.e.

$$\sum_{j=1}^N p_j \leq p^{\max}. \quad (2.8)$$

Hence, the downlink is power-limited.

Definition 2.4. *Assuming perfect power control, the downlink individual load factor is defined as*

$$\eta_i^{\text{DL}} = \frac{1 + \frac{\sigma^2}{g_i \theta_i p^{\max}}}{1 + \frac{W}{\theta_i \left(\frac{E_b}{N_0} \right)_i \nu_i x_i}}. \quad (2.9)$$

Definition 2.5. *The downlink system load factor is defined as the sum of downlink individual load factors of all users in the cell, i.e.*

$$\eta_{\text{sys}}^{\text{DL}} = \sum_{j=1}^N \eta_j^{\text{DL}} \leq 1. \quad (2.10)$$

In the case of multiple cells, the intercell coefficient f_i depends on the position of the user within the cell and is therefore different for each user. If only the average value is considered, θ_i and f_i can be replaced by with $\bar{\theta}$ and \bar{f} .

2.2 Economic Modelling Concepts

2.2.1 Game Theory

Game theory [37, 72, 85] is a tool for analysing the interaction of decision makers with conflicting objectives. Economists have long used it as a tool for examining the actions of economic agents such as firms in a market. Game theory must be understood as situations of *conflict* and *cooperation* between intelligent and rational individuals (or groups of people) for which the objectives are generally more complex than just beating their opponents. Game theory deals with all real-life situations where *rational* people interact with each other, that is when one individual's actions depend essentially on what other individuals might do. The rational choice is the action chosen by a decision-maker that is at least as good as every available action according to his or her preferences.

There are two main branches of game theory: *noncooperative* and *cooperative* games. These two branches differ in how the interdependence and interactions among the players are formalised. In noncooperative game theory, a game is a detailed model of all available decisions available for the players. By contrast, cooperative game theory ignores these details and only describes the optimal outcome when the players come together.

Noncooperative Games

In noncooperative games, decisions are based on players' perceived self-interest and they are unable to make binding agreements except for those explicitly allowed by the rules

2.2 Economic Modelling Concepts

of the games. Noncooperative games are not defined as games in which players do not cooperate, but as games in which any cooperation must be self-enforcing. Games in which cooperation is exogenous are termed cooperative games. For non-cooperative games, there are two types of standard representations: extensive form and strategic form models. The extensive form provides an exact description of players' successive moves and the payoff obtained in the form of a game tree. The extensive representation is only suitable for simple games as the number of nodes can approach infinity for large games with many players.

A strategic game is a model of interacting decision makers. An n -player game consists of: a set of *players*, labelled $i = 1, \dots, n$; the set of *actions*, X_1, \dots, X_n , available to the players; and the *utility* functions, $u_i(x_1, \dots, x_n)$ $x_i \in X_i$ for every i , that describe the preferences over the set of actions. A strategy is a complete specification of how a player intends to play in every contingency that might arise. The game is viewed from the point view of a single player. Each player takes a decision without the knowledge of the decisions taken by their opponents. The *Nash equilibrium* of a strategic game with ordinal preferences is an action profile from which no user may gain by unilaterally deviating.

Definition 2.6. *An outcome $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ of an n -player game (X_1, \dots, X_n) is a Nash equilibrium if the following is satisfied:*

$$u_i(x_i^*, x_{-i}^*) \geq u_i(x_i', x_{-i}^*), \quad \text{for all } x_i \in X_i \text{ and } 1 \leq i \leq n. \quad (2.11)$$

Note that $\mathbf{x}_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

The definition implies neither that a strategic game necessarily has a Nash equilibrium, nor that it has at most one. Note that Nash equilibria do not always entail the same payoff (or utility). The definition of a Nash equilibrium is designed to model a *steady state* among experienced players where no player has any incentive to deviate from his or her strategy given that the other players do not deviate. Nash, in [77], showed that if each player in an n -player game has a finite number of pure strategies, then the game has a Nash equilibrium in pure (one action or the other) and mixed (probabilistic) strategies. In a two-player game, the outcome $x^* = (x_1^*, x_2^*)$ is a Nash equilibrium if and only if x_1^* yields the maximum payoff given that player 2 chooses x_2^* and vice versa. Such a strategy is called a *best response*.

Pareto optimality, which is also known as Pareto efficiency, is an important concept to compare different outcomes of the game. A Pareto optimal solution is defined being efficient. Not every Nash equilibrium is Pareto efficient.

Definition 2.7. *An action vector \mathbf{x} is said to be Pareto optimal if for every other vector \mathbf{y} in the feasible region that*

$$u_i(y_i) \leq u_i(x_i), \text{ for all } i, \quad (2.12)$$

with at least one strict inequality for one i . In other words, an action vector is said to be Pareto optimal if it is impossible to improve the utility of any player without harming another player.

Cooperative Games

Cooperative game theory abstracts from the procedures and details of reaching an outcome and concentrates on the possibility for an agreement. It studies the frictionless negotiations among rational players who can make binding agreements about the rule of the game. Commitments are fully binding and enforceable, i.e. cooperation is exogenous.

Bargaining Games

The formal theory of bargaining was founded by Nash in his two seminal papers [76, 78]. The final outcome is of the main interest here and it is often convenient to analyse the domain of all possible outcomes in order to find an efficient outcome. The desirable solution can be expressed in terms of axioms, which ideally should incorporate some fairness and efficiency features of the solution. A bargaining problem is represented as a pair $(\mathcal{S}, \mathbf{d})$ in the utility space. $\mathcal{S} \subseteq \mathfrak{R}^N$ is the compact and convex set of all utility vectors and $\mathbf{d} \in \mathfrak{R}^N$ is the disagreement or threat point. The disagreement point represents the minimum utility level that the bargainers will obtain if negotiations fail. The set \mathcal{S} must include points that dominate the disagreement point, i.e. there is a positive surplus to be divided among the players once their minimum requirements are reached. The main question then is “*how should this surplus be divided?*”. Approaches to bargaining fall into two divisions: *strategic bargaining* and *axiomatic bargaining*. The former describes what the outcome will be and the latter emphasises how negotiations can evolve to reach an outcome.

2.2 Economic Modelling Concepts

Strategic bargaining [74] studies the exact specification of the negotiation procedure (such as the timing, communication devices and threats) and identifies the rational behaviour in them. An example of a strategic bargaining procedure is Rubinstein's model of alternating offers. The negotiation is modelled explicitly as a real-time game. Suppose that there are two players bargaining over the division of a surplus of 1. In period 1, Player 1 will make a proposal of the division. Player 2 can either reject or accept this proposal. If it is accepted, the negotiation ends. If it is rejected, Player 2 will make a proposal in period 2. Then, Player 1 must respond. The negotiation continues until an outcome is reached. In the T -period horizon game, the disagreement will be imposed after T proposals have been rejected.

Axiomatic bargaining [95, 96] assumes some desirable properties about the outcome of the bargaining process and then identifies process rules that guarantee this outcome. It ignores the negotiation process completely. Nash proposed the Nash bargaining solution while he was still an undergraduate in his paper [76]. Nash specifies four axioms, which impose properties that a bargaining solution should satisfy:

- **Symmetry:** Two players with symmetric utilities get the same payoff. It ensures the solution yields a fair outcome.
- **Pareto Optimality:** The solution is on the Pareto boundary. This axiom reflects the rationality of the players. If they work together, they would not accept the disagreement point as the outcome when they can do better than that.
- **Invariance with respect to affine transformation:** If the utility functions are rescaled, the solution should be rescaled in the same fashion. For example, a change of currency of the players' utility implies a change in the currency of the outcome.
- **Independence of Irrelevant Alternatives (IIA):** Suppose the solution for bargaining problem $(\mathcal{S}, \mathbf{d})$ is \mathbf{s}^* . Now consider a new bargaining $(\mathcal{S}', \mathbf{d})$, where $\mathcal{S}' \subseteq \mathcal{S}$. Then the solution for the new bargaining problem is also \mathbf{s}^* . That is, the solution is independent of "alternatives" that are deemed irrelevant because they were not chosen in the larger \mathcal{S} , so their absence should not alter the outcome.

Definition 2.8. *The Nash bargaining solution satisfies the axioms of symmetry, Pareto optimality, invariance with respect to affine transformation and IIA and is defined as*

$$f(\mathcal{S}, \mathbf{d}) = \arg \max_{\mathbf{s} \in \mathcal{S}} \prod_{i=1}^N (s_i - d_i), \quad (2.13)$$

subjected to constraints $s_i \geq d_i$, $i = 1, \dots, N$. The product of $(s_i - d_i)$, $\forall i$, is called the Nash product.

Axiom IIA received a number of criticisms from a number of researchers. Kalai and Smorodinsky [59] (and Raiffa [93] in an earlier work) proposed to drop IIA and replace it by the axiom of individual monotonicity. That axiom states that players' assigned utility should be proportional to their aspiration level or maximum gain.

Definition 2.9. *The Raiffa-Kalai-Smorodinsky bargaining solution satisfies the axioms of symmetry, Pareto optimality, invariance with respect to affine transformation and monotonicity and it is the intersection point of a line connecting the disagreement point and the utopia point of the Pareto optimal boundary.*

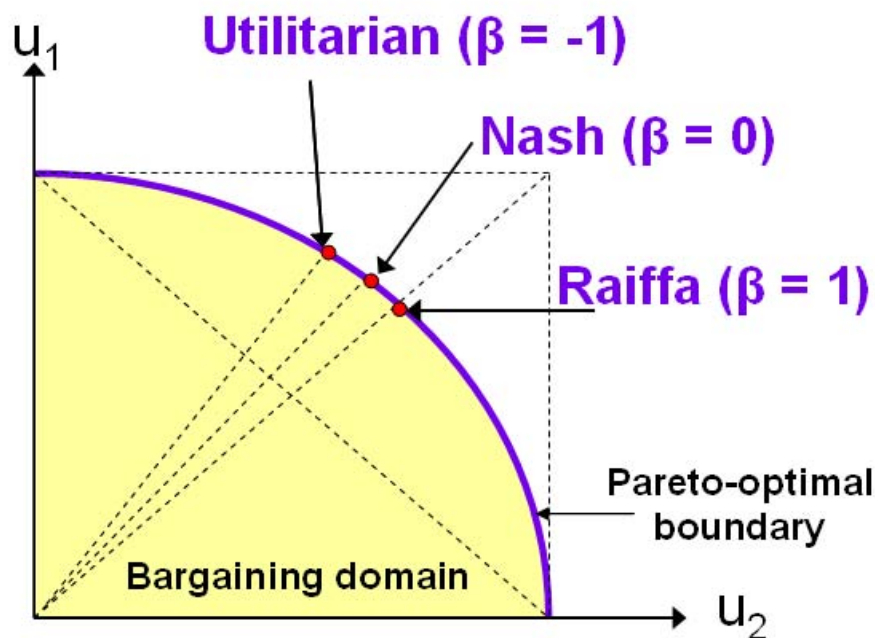


Figure 2.1. The Nash, Raiffa-Kalai-Smorodinsky and utilitarian bargaining solutions are special instances on the Pareto-optimal boundary. They can be obtained by varying the value of β in the preference function.

2.2 Economic Modelling Concepts

Besides these solutions, there are other axiomatic bargaining solutions such as utilitarian and egalitarian (see [118]). Cao in [18] showed that these solutions can be expressed using preference functions. By varying a parameter β , which represents the trade off between one's gain and the losses of others, a class of bargaining solutions with Nash, Raiffa-Kalai-Smorodinsky and modified Thomson as special cases can be obtained (see Fig. 2.1). We will apply this concept to efficient radio resource management in a WCDMA network in our work in Chapter 5. For a more comprehensive introduction to the theory of bargaining, we refer our readers to [101].

Coalitional Games

The bargaining games introduced previously assume that the final outcome is achieved with the cooperation of all players. By contrast, coalitional games allow the formation of coalitions that influence the final outcome. The solution concepts can be divided into two families [41]:

- **Valuation approach:** This solution concept was introduced by Shapley [103] and is a powerful tool for evaluating the power structure in a coalitional game. Shapley suggested to summarise the complex possibilities facing each player in a game in coalitional form by a single number expressing the value of playing the game, i.e. the Shapley value. The Shapley value satisfies the axioms of group rationality, symmetry, dummy player condition and additivity. It can be interpreted as the expected marginal contribution of each player when he/she enters a coalition.
- **Domination approach:** This concept uses domination or objection to derive results concerning stability and coalition formation. The aim is to look at deviation possibilities of coalitions for delimiting the set of outcomes which are “unobjectionable”. The concepts in this approach are the core, stable set, bargaining set, kernel and nucleolus.

2.2.2 Pricing Schemes for Wired Networks

Static Pricing

The simplest form of static pricing is fixed pricing, schemes that charge a fixed price independently of either call duration or bandwidth usage. Although this scheme is simple, easy to budget for and has little overhead for billing and accounting, it is unresponsive to the fact that resource demand varies throughout the day, the month and the year. It lacks flexibility and fails to provide incentives to shape their demand in response to network congestion.

- **Flat Pricing**

Under flat pricing [6], users are charged a fixed amount per time unit such as per month, regardless of the actual usage. This charging method is akin to that of *all-you-can-eat* restaurants. At such restaurants, a customer is charged not for his own food consumption, but rather for the average amount that similar customers have eaten in the past. The advantage of flat pricing is that it is simple to understand from the users' perspective and easy to deploy since no measurements are required to record users' usage for the purpose of billing and accounting. However, it is obvious that users have no incentive to alter their usage behaviour in response to the network providers' need for congestion control and traffic management. The lack of penalty for over-use encourages heavy users for using more than the average amount of resources.

- **Time-of-Day Pricing**

The definition of static pricing also encompasses time-of-day pricing, schemes that try to accommodate peak and off-peak network periods. This policy attempts to take advantage of demand elasticity by utilizing historical information about expected peak load and congestion periods [87]. The peak period, which is during business hours, has a much higher call arrival rate. Off-peak periods such as late evening or early morning are periods where network utilization is predicted to be low. Rates are lower during this off-peak time in hopes of stimulating user demand for resources. Although this scheme is marginally better than fixed pricing, being more responsive to users' arrival patterns, it only has two variations in a day based on what network operators anticipate the load would be

2.2 Economic Modelling Concepts

without variation. It does not vary price with respect to the actual load and so will not be able to react to unexpected increase in network traffic. One wad in which this can be a problem is that potential users during expected peak hours might be discouraged to use the network although it is not in a congested state.

- **Paris-Metro Pricing**

This pricing scheme is based on the concept of travel class, as used in public transport systems [83]. The network is partitioned into several logical subnetworks that operate on a best-effort basis. Each subnetwork is priced differently, for example based on customer surveys. Depending on the expected network congestion, users select one of these logical networks (and therefore price) that is within their budget constraint. As a result of differential pricing, differential quality can be provided for users in best-effort networks. As with flat pricing, this scheme is simple to use and easy to deploy. However, Paris-Metro pricing does not provide any individual QoS guarantee. During periods of congestion, price-insensitive users might choose a higher-priced network in expectation of a better service. However, this may lead to congestion in the higher-priced network and thus cause instability.

- **Priority- and Reservation-based Pricing**

Priority pricing is introduced in [22, 23] and uses the concept of *price discrimination* to price users differently based on their price elasticity demand. Under this scheme, users are forced to indicate the level of their resource use by selecting a priority level set by the provider. High priority services are charged accordingly. During congestion, traffic transmitted by priority level and low-priority traffic may be delayed or dropped. Under the assumption that users are sensitive to the network's performance and price signals, prices can be set carefully such that the overall user satisfaction is higher under priority pricing than under flat and Paris-Metro pricing. However, this scheme assumes knowledge of users' utility function over all time frames throughout the transmission. This model is further extended by [87, 88] to allow for bandwidth reservation. In this reservation-based scheme, users choose a class of service and pay a fee per packet based on the class of service chosen. At the time of call setup, users not only determine the class of service they desire, but also how long they need the network resources. Based on this information,

the network will simply reserve the necessary resources to accommodate the call or reject it. Unfortunately, this policy requires enormous a priori knowledge, which is difficult and not always realistic to obtain.

- **Effective-Bandwidth Pricing**

A usage-based charge that is based on effective bandwidth [62] is constructed in [60] and [24] where charge is a function of both *static* parameters which are part of the traffic contract (such as peak rate and average rate parameters) and *dynamic* parameters (corresponds to the actual traffic of the connection, such as volume and duration). Charges depend upon both *a priori*, i.e. traffic contract parameters, and *a posteriori*, i.e. actual usage, measurement of network resources. Static parameters (a priori) are policed while dynamic parameters (a posteriori) are measured and effective bandwidth is bounded by a linear function of the measured parameters. The charge proposed is incentive compatible and hence a rational user will seek to minimise his/her charge by minimise resource use, i.e. minimise its effective bandwidth subjected to peak and average usage. This is just one example of constraints that can arise as part of the traffic contract between the network and the user. On the other hand, the network will maximise the number of users subject to the limits of capacity and buffer size of the network. The authors argued that *fair* usage-based charging schemes, i.e. schemes that capture the relative amount of resources used by connections, are required to achieve economic efficiency.

- **Edge Pricing**

Edge pricing [104] provides a conceptual shift of focus to locally computed charges based on simple expected values of congestion and route. Shenker offers the criticism that those usage-based pricing schemes which optimise overall user benefit ignore architectural and technical issues. Instead, this model charges for usage at the *edge* of the network scope for the subscriber, rather than along the expected path of the source and destination of the calling session. The networks in turn cross-charge each other at the network edges. This approach has the advantage of capturing billing data locally without having to exchange data of all sessions with other networks and partners for subscriber billing, as for current roaming arrangements between mobile markets. However, this model lacks visibility of the routing via external networks and the costs of the traffic to both networks.

Dynamic Pricing

Dynamic pricing could provide an additional strategy for stimulating demand and encouraging more efficient use of available resources. In these schemes, the price of calls changes as demand fluctuates. Various dynamic pricing schemes have been proposed for packet-based networks. In general, there are three major approaches for in the design of seller-centric dynamic pricing schemes for wired networks:

- Auction-based approach: The seller, i.e. the service provider, engages the users to bid for the network resources. Resources are allocated according to users' bids.
- Shadow pricing approach: The seller derives a *shadow price* that maximises the aggregate utility of all users. Resources are allocated according to the ratio of users' payment with respect to the shadow price. This scheme is designed for a fixed number of users.
- Stochastic control approach: The seller computes an optimal dynamic pricing policy based on its knowledge of the stochastic nature of users' arrival and departure.

The first two approaches are designed for a fixed number of users in the system. Congestion control measures provided by them are *reactive* because prices are optimised for these users and will only increase when the aggregate resource demand increases. By contrast, the stochastic control has the ability to set low prices to encourage arrivals when the system is lightly loaded, in addition to setting high prices to alleviate congestion when the system approaches a busy state. Therefore, it is both *preventive* and reactive.

- **Auction-based Approach**

Auction-based **smart-market pricing** has users inform the network of their willingness to pay for a transmission of a packet prior to transmission. It has been proposed in [68], where it has been shown to take into account the issues of capacity expansion and the social cost imposed on other users. To register user information, each packet carries a *bid* in the packet header and packets are serviced at each router if their bids exceed some threshold, which is also the actual charge. This threshold is chosen to be a market clearing price, ensuring that the network is fully utilized. This scheme addresses the important problem of *congestion externalities*, the congestion that one user imposes on others.

A similar approach has been considered in [117] for service provisioning in a connection-oriented network. A nonlinear mathematical programming model that maximises the expected revenue is proposed in [63]. The problem is solved using an auction algorithm to search for the optimal prices, subject to the constraints on the resource allocation implied by finite resources and QoS guarantees. The implementation of auction-based pricing involves major structural changes not only to network management but also to users' applications, which must be able to submit bids to the network. In addition, accurate bids cannot be submitted without precisely knowing the delay associated with each bid [104].

- **Shadow Pricing Approach**

The concept of proportional fairness pricing, i.e. that a resource allocation is fair if it is in proportion to users' willingness to pay, was proposed in [61] and further developed in [38] for fixed capacity networks. This technique applies optimisation techniques to maximise operator's revenue and the aggregate utilities of all users. In the latter, the solution is called the *social optimum* outcome. Suppose that R users share a communication link with capacity $C > 0$. Denote d_r as the rate allocated to user r and $U_r(d_r)$ as the utility received by user r when d_r is received. The system optimal rates can be obtained from the following problem:

$$\begin{aligned} \max_{\mathbf{d}=(d_1,\dots,d_R)} \quad & \sum_{r=1}^R U_r(x_r) \\ \text{s.t.} \quad & \sum_{r=1}^R d_r \leq C \\ & d_r \geq 0, r = 1, \dots, R. \end{aligned} \tag{2.14}$$

Depending on the characteristics of the objective and constraint functions, this problem can be solved using linear programming, nonlinear programming and integer programming. When the objective function is continuous and the feasible region is compact, an optimal solution exists. If the feasible region is convex and the functions $U_r(d_r)$ are strictly concave, the optimal solution is unique. The Lagrangian function of problem (2.14) can be written as:

$$\mathcal{L}(\lambda, \boldsymbol{\mu}) = \sum_{r=1}^R U_r(x_r) - \lambda \left(\sum_{r=1}^R d_r - C \right) + \sum_{r=1}^R \mu_r d_r. \tag{2.15}$$

The Lagrange multiplier, λ , is interpreted as the *shadow price* of using a unit of resource.

2.2 Economic Modelling Concepts

Users' utility function are usually unavailable to the resource manager. If they are charged a shadow price μ , the utility maximisation problem of the user is given by

$$\begin{aligned} \max_{d_r} \quad & U_r(d_r) - \lambda d_r \\ \text{s.t.} \quad & d_r \geq 0. \end{aligned} \tag{2.16}$$

Similarly, users can also submit a resource bid w_r to the resource manager based on shadow price λ . The decentralised utility maximisation problem becomes

$$\begin{aligned} \max_{w_r} \quad & U_r\left(\frac{w_r}{\lambda}\right) - w_r \\ \text{s.t.} \quad & d_r \geq 0. \end{aligned} \tag{2.17}$$

A system optimum is achieved when users' choices of resource usage in (2.16) or bid in (2.17) coincide with the network's choice of allocated rates in equilibrium in (2.14). Such an allocation of resources guarantees economic efficiency, since the sum of users' utilities is optimised. A particular application of with this approach is in an ATM network offering available bit rate service. Assuming the users are price-anticipating, Johari and Tsitsiklis extended this resource allocation method to a congestion game [56].

- **Stochastic Control Approach**

The dynamic pricing mechanisms we have introduced thus far only deal with a fixed number of users and do not take into account any stochastic knowledge of the system. In [90], a stochastic control model for **congestion-dependent pricing** based on dynamic user arrivals is proposed. The optimal pricing policy for the single link case is computed using dynamic programming (see Section 2.3.1), with the objective of maximising revenue or social welfare. The authors conclude that static pricing is asymptotically optimal under extreme conditions, i.e. very large or many relatively small users. Later in [89], Paschalidis and Liu extended the model to a network setting with more than one node. The optimal prices for the network case are computed using a simulation-based approach. This approach has been further developed in [124] and [91].

2.2.3 Pricing Schemes for Wireless Networks

While the previous section has shown that there is considerable literature which applies pricing schemes in telecommunication networks, research into the design of pricing

schemes for wireless networks has been relatively new and few. The use of pricing in wireless networks has been tightly coupled with power control and rate allocation. We summarise the resource control via pricing approach in Table 2.1. The common approach is to associate users' satisfaction or requirement with a utility function and achieve a socially optimal outcome using pricing. In these works, prices are only used as an internal control mechanism and do not reflect the actual charges that end-user pay. We classify the approaches used in existing proposals into two categories:

- **Shadow pricing approach:** Similar to [38, 61], the seller computes a shadow price that maximises the social welfare of all users. Facing such a shadow price, the users then make resource request such that their net utility, i.e. utility minus cost, is maximised. Hence, a social welfare optimum can be achieved in a decentralised manner if the resource requests of users coincide with the optimal allocation of the network.
- **Noncooperative game-theoretic approach:** The main focus of the works that fall into this category is decentralised, noncooperative power control. Noncooperative game theory was first used as a framework for the power control problem in [39]. Users adjust their transmit power so that their utility is maximised. The stable point at which no user can unilaterally improve their utility is called the Nash equilibrium of the system. Due to the competitive nature of the users' interaction, a Nash equilibrium does not always exist, and even if it does, it is likely to be inefficient. A linear, static pricing scheme is used to discourage selfish behaviour such that a solution that provides Pareto improvements (see Definition 2.7) over the Nash equilibrium can be achieved.

- **Shadow Pricing Approach**

Marbach and Berry [70] proposed a model for resource allocation and pricing for the downlink of time-slotted CDMA systems, where the base station can only transmit to a single user at any given time. Users' utility is a function of their throughput, which is in terms of the number of packets transmitted. The pricing framework analysed is receiver driven, i.e. the receiver pays for service. Two pricing models are considered: a pre-determined pricing scheme and a resource auction mechanism. In the first model, users agree to pay a pre-determined price and the base station allocate resources such that the

2.2 Economic Modelling Concepts

Table 2.1. Summary of the use of pricing in wireless networks.

Paper	Summary
[70]	<p>Problem: Downlink rate allocation with shadow pricing</p> <p>Utility function: Concave function of throughput</p> <p>Pricing function: Linear function of throughput</p> <p>Objective: Resource manager maximises throughput, revenue and social welfare.</p>
[108,109]	<p>Problem: Rate and signal quality assignment with shadow pricing</p> <p>Utility function: Product of transmission rate and signal quality</p> <p>Pricing function: Linear function of utility</p> <p>Objective: Resource manager maximises social welfare. Users maximises their net utility.</p>
[32,71,98,99]	<p>Problem: Uplink power control using noncooperative game theory</p> <p>Utility function: Throughput per terminal life</p> <p>Pricing function: Linear function of transmit power</p> <p>Objective: Users maximises their net utility.</p>
[126]	<p>Problem: Downlink power control using noncooperative game theory</p> <p>Utility function: Sigmoid function of SIR</p> <p>Pricing function: Linear function of transmit power</p> <p>Objective: Users maximises their net utility.</p>
[66]	<p>Problem: Downlink power control using noncooperative game theory</p> <p>Utility function: Sigmoid, concave or convex function of SIR</p> <p>Pricing function: Linear function of transmit power</p> <p>Objective: Users maximises their net utility.</p>
[67,131]	<p>Problem: Downlink power control with partial cooperation</p> <p>Utility function: Step function of SINR</p> <p>Pricing function: Linear function of transmit power and code</p> <p>Objective: Users maximises their net utility.</p>

revenue or social welfare is maximised. This pricing scheme is similar to the static pricing scheme with service differentiation. The second model is similar to problem (2.17). Based on the knowledge of the utility functions of all users, the base station announces a shadow

price that maximises social welfare. Users respond by submitting bids that indicate the amount of resources they need. In each frame, users pay a price that is equal to their bid, which is independent of the amount of data they receive.

In [108, 109], joint optimisation of the transmission rate and signal quality, which is defined in terms of target energy-bit-to-noise-density ratio, is considered. Users' utility is defined in terms of the product of their rate and target energy-bit-to-noise-density ratio. However, in the case of elastic traffic, this joint problem can be decoupled into two problems. In particular, the base station first selects an optimal target bit-energy-to-noise-density ratio for each user and announces a shadow price per unit resource λ . Similar to problem (2.16), the users react to the price by selecting a transmission rate that maximises their net utility. Users are charged linearly with respect to the utility obtained.

- **Noncooperative Game-Theoretic Approach**

In [32, 98, 99], utility is a function of the throughput (bits transmitted) per terminal battery lifetime, which is measured in bits/joule. The focus of these works is the transmission of wireless data on the uplink. Users enter into a non-cooperative game to maximise their individual utilities by adjusting their transmitter powers. The authors show that the resulting noncooperative game has a Nash equilibrium that is inefficient. Pricing is used to achieve a more Pareto efficient equilibrium solution. The base station informs each user of a fixed price per unit transmit power and each self-optimising user adjusts its power level to maximise their net utility (utility minus cost of power allocation). They then show that there exists equilibria in the noncooperative power game with pricing and that these equilibria are Pareto-superior compared to the equilibrium of the game without pricing. However, a socially optimum power solution cannot be obtained even with pricing. The single-cell problem presented in [99] is extended to a multi-cell environment in [98], where each user experiences interference from other terminals outside its cell in addition to the ones within the same cell. A similar utility function is considered in [71].

A framework for downlink utility-based power control (UBPC) for systems with softened SIR requirements is presented in [126]. Similar to [32, 98, 99], the problem is modelled as a noncooperative game, in which the BS informs each user of a fixed price per unit power and each user requests a power level that maximises its net utility (utility minus cost). There is no constraint on the total power transmitted. A sigmoid-like function

2.2 Economic Modelling Concepts

is used to model users' utility as a function of SIR requirements. By adjusting the parameters of the sigmoid function, the utility functions of voice and data users can be treated in a unified way. By softening the SIR requirement, the algorithm suggested is exempt from the divergence problem of models with hard SIR requirements. When the transmission environment becomes very hostile, transmissions that do not have positive net utility will be totally shut off by UPBC and the system will still be feasible. The cost function used is a linear function of the transmit power. However, the authors did not provide an algorithm for how to obtain the optimal price.

A power and code allocation policy using pricing for a single-cell CDMA network is considered in [67]. The focus is on voice users and users' utility is a step function of the signal-to-interference plus noise ratio (SINR). The BS announces a price per unit transmitted power α_p and a price per code α_c . Each user responds by requesting the amount of resources that maximise his/her individual surplus (utility minus cost). The total charge for the user with power P_k is therefore $\alpha_c + \alpha_p P_k$. Their motivation is power and code allocation and the prices set serve only as internal network parameters to guide the resource allocation. Therefore, users may not actually pay the prices set by the network. This work is extended to the two-cell setting by [131].

A distributed algorithm similar to [126] is proposed in [66] to obtain an approximation to the social optimal downlink power allocation for multi-class CDMA networks. Unlike in [126], a total power constraint is exercised. The algorithm consists of two stages: mobile selection, i.e. the base station selects mobiles to which power is allocated, and power allocation, i.e. the base station allocates socially optimal power to the mobiles selected. Since the base station needs some cooperation from unselected mobiles to not to participate in the power allocation game in the second stage, the problem is expressed as a partial-cooperative power allocation game. The algorithm also incorporates a dynamic pricing component. Instead of using a fixed price per unit power, as in [32,98,99,126], the base station dynamically adjusts the price to obtain a good approximation to the socially optimal power allocation problem, i.e. when the total power allocated is closest to the total power constraint.

2.3 Optimal Control Theory

2.3.1 Dynamic Programming

The concept of dynamic programming was introduced by Bellman in 1957 [10]. In this section, we will provide an overview on Markov decision process (MDP) and dynamic programming. Treatments of dynamic programming can be found in a number of textbooks and particularly, we recommend [12] and [92] for further reference. Markov decision theory provides powerful tools for the analysis of probabilistic sequential decision processes and has been widely applied in inventory control, maintenance, computer science and resource allocation. An MDP is characterised by the following:

- **States:** A state i represents some properties of the system. For example, the number of users, n , in a system. The state space of the system is denoted as \mathcal{S} .
- **State transitions:** A transition describes a change in the system due to its stochastic nature or a decision by the decision maker. For example, an arrival or departure that respectively increases or decreases the number of users in a system.
- **Transition rates and probabilities:** The transition rate q_{ij} describes how often state transitions occur. For example, the transition rate from state $s = i$ to j depends on the arrival rate of users. A continuous-time MDP can be converted to its discrete-time equivalent using uniformisation techniques. In the equivalent process, state transitions are described by transition probability p_{ij} , where $\sum_{j \in \mathcal{S}} p_{ij} = 1$, for every $i \in \mathcal{S}$.
- **Control actions:** State transitions can be controlled by the users or the system owner. Control actions u refer to the decisions that the owner can make in order to achieve some system objectives. For example, the owner might block an arrival so that the system state remains the same. The control space is denoted as \mathcal{U} .
- **Reward function:** Reward function $g(i, u, j)$ gives the amount of reward obtained once a decision is executed in state $i \in \mathcal{S}$ and resulting a transition to state $j \in \mathcal{S}$. The reward function is set up such that a long-term objective is met.

2.3 Optimal Control Theory

- **Policy:** The control u depends on the state i and the rule by which the controls are selected is called a policy. Once an optimal policy is computed off-line, the decision-maker can refer to the policy every time an action has to be taken.

In computing the optimal control u in state i , it is not enough to compare the magnitude of the reward $g(i, u, j)$. The desirability of going to the next state j is equally important. The next state is ranked using the optimal reward-to-go of state j , denoted as $J^*(j)$.

Definition 2.10. *The Bellman equation for a system with state space \mathcal{S} , control space \mathcal{U} , reward function $g(i, u, j)$ and transition probabilities $p_{ij}(u)$ for $i, j \in \mathcal{S}$ is given as*

$$J^*(i) = \max_{\mathbf{u} \in \mathcal{U}(i)} E [g(i, u, j) + J^*(j)|i, u], \quad \forall i, \quad (2.18)$$

where j is the state subsequent to i , and $E[\cdot|i, u]$ denotes the expected value with respect to j , given i and u . J^*

For our work in Chapter 3 and 4, we are only interested in the average reward problem because immediate actions are as important as future actions. In that case, the Bellman equation can be rewritten as:

$$J^* + h(i) = \max_{\mathbf{u} \in \mathcal{U}(i)} \left\{ \sum_{j \in \mathcal{S}} p_{ij}(u) [g(i, u, j) + h(j)] \right\}, \quad (2.19)$$

where J^* is now defined as the optimal expected revenue per state transition and $h(i)$ is the relative or differential value of state i . The objective is to compute an optimal policy such that the long-term, expected reward is maximised.

The Bellman equation can be solved using dynamic programming algorithms such as value iteration, policy iteration or linear programming. Dynamic programming methods are closely related to heuristic search. Like a heuristic search algorithm, dynamic programming is an off-line procedure for designing an optimal control policy. However, unlike heuristic search algorithms, dynamic programming produces an optimal closed-loop policy instead of an open-loop policy for a given initial state. The main drawback of this method is that the computational complexities of classical dynamic programming algorithms increases exponentially with the size of the state space. For this reason, an exact solution using dynamic programming is feasible only when the number of states is quite small.

2.3.2 Neuro-Dynamic Programming

Dynamic programming offers a class of algorithms for computing optimal control policies. For many problems where the number of states and controls are very large, dynamic programming suffers from the well-known Bellman *curse of dimensionality* due to overwhelming computational requirements. In such situations, a suboptimal solution is required. Neuro-dynamic programming (NDP) [13] refers to approximate methods that centre around the evaluation and approximation of the optimal reward-to-go function, possibly through simulation and/or the use of neural networks.

Instead of computing the differential reward function $h(i)$ for every state $i \in \mathcal{S}$, NDP uses a compact representation $\tilde{h}(\cdot, \boldsymbol{\theta})$ to approximate $h^*(\cdot)$, using parameter vector $\boldsymbol{\theta}$. Naturally, we want to define the general structure of $\tilde{h}(\cdot, \boldsymbol{\theta})$ and calculate parameter vector $\boldsymbol{\theta}$ so as to minimise the error between the functions $h^*(\cdot)$ and $\tilde{h}(\cdot, \boldsymbol{\theta})$. The process of tuning parameter vector $\boldsymbol{\theta}$ is often referred to as training or learning. The average reward per time J^* is approximated by tunable scalar \tilde{J} . If $\tilde{h}(\cdot, \boldsymbol{\theta})$ and \tilde{J} are close enough to the $h^*(s)$ and J^* , then the greedy control policy induced is, in some sense, close to an optimal policy. Hereafter, we denote the k^{th} step estimate of $\tilde{h}(\cdot, \boldsymbol{\theta})$ and \tilde{J} as $\tilde{h}(\cdot, \boldsymbol{\theta}_k)$ and \tilde{J}_k respectively. There are two major parts in NDP:

- **Approximation architecture:** The approximation architecture of the problem refers to the development of an approximate representation of the differential reward function $\tilde{h}(\cdot, \boldsymbol{\theta})$ so that the problem does not suffer from the curse of dimensionality. This should provide an acceptably close approximation to the function itself. The design of an appropriate approximation architecture is usually problem-dependent. Approximation architectures can be classified into two main categories: linear and nonlinear.
- **Simulation and learning:** Using simulation (or learning by experience) to compute the approximate reward-to-go is a key distinguishing aspect of NDP. The significant advantage of this method is that a detailed model of the system is not necessary. During training or learning, the parameter vector $\boldsymbol{\theta}$ is systematically tuned. Examples of learning algorithms are temporal difference algorithm TD(λ), Q-learning and λ -policy iteration. These algorithms are derived based on dynamic programming ideas of value and policy iterations.

Chapter 3

Integrated Dynamic Pricing and Call Admission Control

3.1 Introduction

Call admission control (CAC) and dynamic pricing have been proposed as arbitration mechanisms to regulate traffic and reduce congestion in a network. CAC is a provisioning strategy that limits the number of call connections. It means that users are not automatically admitted, even when there are resources available. Dynamic pricing makes adjustments often, according to the demand pattern and congestion level in order to influence the way users utilise network resources, especially to allocate limited resources to those who value them most. Dynamic pricing also enhances network operators' ability to recover costs and make profits to finance capacity expansions.

In this chapter, we provide a generic framework to formulate an integrated call admission and dynamic pricing problem for a multiservice, single-cell, cellular network with fixed capacity as a Markov Decision Problem (MDP) [92] and solve it using Dynamic Programming methods [12], with the objective of maximising the long-term expected revenue. Given a particular configuration of network users, the objective is to determine both whether or not to accept a new connection and the optimal price per bandwidth time to shape demand. Since premature termination of ongoing calls is more undesirable than rejection of new call requests, it has been widely accepted that a system should allocate a higher priority to handoff call requests. Our model allows for bandwidth reservation

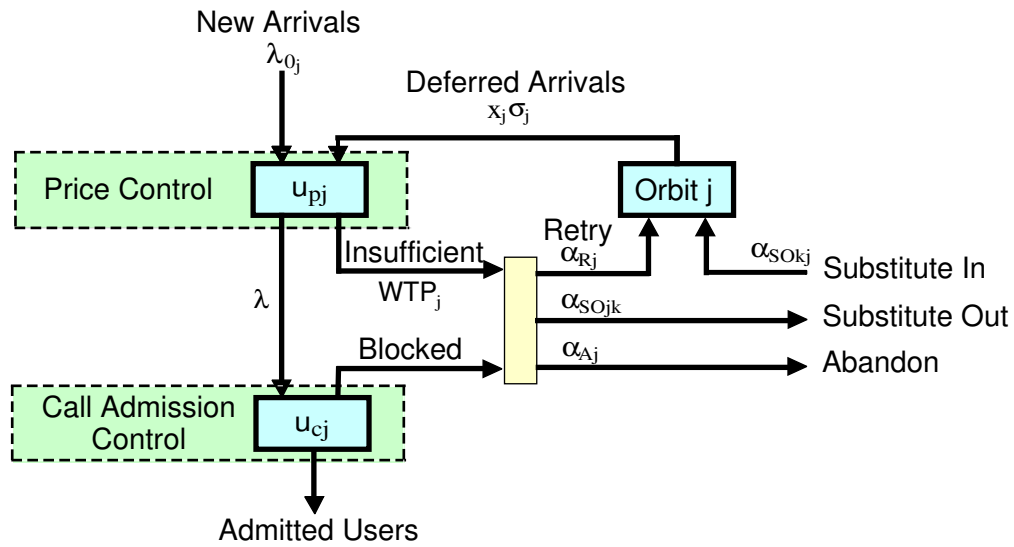


Figure 3.1. Integrated call admission control and dynamic pricing model.

for handoff calls by associating a *satisfaction revenue* (SR) with the admission of handoff calls. SR is not real income to the service operator but rather an incentive which favours handoff calls.

Our approach is similar to the stochastic control approach used in [90] and [89]. Their proposal, which does not consider handoffs, is a special case of ours since our model jointly optimises dynamic pricing and CAC policy to handle new and handoff calls. Also, we allow an incoming call request to be rejected even though the user has sufficient budget and enough bandwidth is available, in order to give way to future connection requests that generate higher expected long-term revenue. Our work is also different from a number of papers on CAC such as [5], [33] and [20], which assume that prices are fixed and are concerned with admission decisions only. The call admission problem that we analyse is similar to the *stochastic knapsack* problem in [5], where an optimal call admission policy is derived from a fluid model.

A similar problem, but with different approach, has been considered in [50]. In that work, the pricing component of the integrated system is only active when the arrival rate of the system exceeds a pre-determined optimal level. Their approach does not provide opportunity to the network operator to generate additional revenue when the traffic level is below average. Therefore, in addition to congestion control during high-traffic periods, we use price as an incentive to encourage arrivals during low-traffic periods. Our CAC

approach is also better because the number of guard channels can be optimally calculated to reserve resources for future-arriving users so that the long-term revenue is maximised. Apart from the differences highlighted above, our problem is different because we consider the integrated problem in a decision-theoretic framework under an explicit model of users' reaction to prices. The integrated control problem is depicted in Fig. 3.1 and will be explained in the following sections.

A major assumption in all of the papers mentioned is that users arrive independently according to a Poisson process and will be automatically lost if blocked. This queueing model creates an underestimation of the actual arrival rate in the system, especially during congestion. We consider an advanced arrival model that incorporates retries and substitution effects among services and through time. We assume that some users remain in the vicinity of the system when they are blocked. They have a choice to defer their call requests or to use another service as a substitute. These deferred users, together with new arrivals, provide a more realistic estimate of the actual arrival rate to the system than do conventional arrival models. Unlike the immediate substitution model proposed in [89], our model allows both immediate and deferred substitutions and considers the rate of substitution to be a random variable.

The system model is found to have a level-dependent quasi birth death (LDQBD) structure. Using matrix-analytic methods (see [65, 79]), we can derive various system characteristics through the stationary distribution of the Markov chain. Finally, we consider a simplified version of the control problem that sets the price per bandwidth time for all services to an optimal value, instead of one for each service, that maximises the total expected revenue. The original optimal policy, which allows for price variation among the services, is an example of *price discrimination*. The near-optimal case (no price discrimination) provides lower expected revenue but has the benefit of reduced complexity because it considers a smaller price control space.

The remainder of the chapter is organised as follows. In Section 4.2, we present our system model. We then formulate our integrated problem as an MDP in section 4.3. The stationary distribution of the model is then derived in Section 3.4 to obtain system characteristics. We then present computational results and discuss price discrimination in section 4.5. Finally, we present our main conclusions and suggestions for possible extensions in section 4.6. This work has been partially presented in [45].

3.2 System Model

We consider a network with a total capacity of B units of bandwidth and J classes of service. The price per unit time for using one unit of bandwidth of service class j is denoted by u_{p_j} , $j = 1, 2, \dots, J$. Each service class j , which uses b_j units of bandwidth, is characterised by its Poisson-distributed new and handoff call arrival rates λ_j^n and λ_j^h ; and exponentially-distributed call holding time $1/\mu_j$. The satisfaction that users gained from a call is quantified by their *willingness to pay* (WTP) Ψ_j , which is a Weibull-distributed parameter with mean ψ_j and shape β_j . WTP is similar to the concept of *utility* and measures how much a user values the call.

The Weibull distribution is one of the most widely used lifetime distribution in reliability engineering. We use the Weibull distribution to model users' WTP because it is versatile and can take up the characteristics of other types of distributions, based on the value of the shape parameter [86]. Within the telecommunications framework, it has been used to model the traffic characteristics of packet audio streams in [21], to simulate data traffic in wireless networks in [106] and to characterise fading channels in [129]. As shown in Fig. 3.2, different values of the shape parameter can have different effects on the

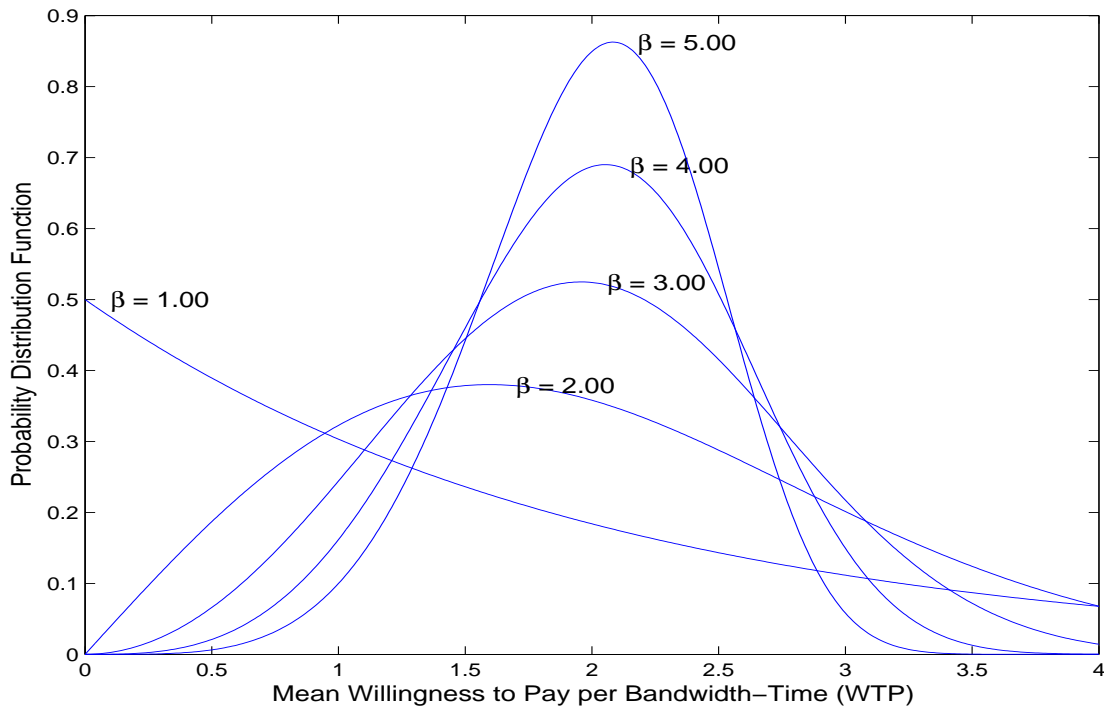


Figure 3.2. The probability distribution function of WTP using variable shapes β .

behaviour of the distribution. For example, when $\beta = 1$, the probability density function of the Weibull reduces to that of an exponential distribution. Exponential demand functions, as used in [50] and [36], are special instances of the WTP distribution considered in this work. As mentioned in [50] and [36], the parameters that define the utility function must be identified by adequate market research. CAC is triggered by call connection requests. At the time that a new or handoff call is accepted, the system must have at least b_j units of bandwidth available.

Definition 3.1. *A CAC policy determines the state-dependent admission policy:*

$$\mathbf{u}_c(s) = (\mathbf{u}_c^h(s), \mathbf{u}_c^n(s)) = (u_{c_1}^h(s), \dots, u_{c_J}^h(s), u_{c_1}^n(s), \dots, u_{c_J}^n(s)) \quad (3.1)$$

for all states $s \in \mathcal{S}$, where $u_{c_j}^n, u_{c_j}^h \in \{0, 1\}$. A handoff (new) connection can either be accepted with $u_{c_j}^h = 1$ ($u_{c_j}^n = 1$) or rejected with $u_{c_j}^h = 0$ ($u_{c_j}^n = 0$).

Definition 3.2. *Price control determines the state-dependent pricing policy:*

$$\mathbf{u}_p(s) = (u_{p_1}(s), \dots, u_{p_J}(s)), \quad (3.2)$$

with $u_{p_j}(s) \in \mathcal{U}_{p_j}$, to regulate new call arrivals. \mathcal{U}_{p_j} is the set of possible values of u_{p_j} .

We will use $u_{p_j}(s)$ and u_{p_j} , and $u_{c_j}(s)$ and u_{c_j} interchangeably in this chapter. A new user will decide to either make a connection request if his or her budget is sufficient to cover the expected call cost, i.e. $\Psi_j \geq u_{p_j} b_j / \mu_j$ or defer the request otherwise. When blocked, handoff users will leave the system immediately. However, new users of service j who are rejected by control or have insufficient WTP will either retry later with probability α_{R_j} , substitute out to another service $k \neq j$, with probability $\alpha_{SO_{j,k}}$ or abandon the system with probability α_{A_j} . New users of service j who retry later and users of service $k \neq j$ who substitute into service j are said to be in *orbit* and will independently generate requests for service at exponentially-distributed time intervals with mean $1/\sigma_j$ until they obtain service, abandon or substitute out. In other words, our model incorporates substitution effects among classes and through time, i.e. users can use another service as a substitute or defer their call until the price falls below their WTP.

Definition 3.3. *The probability of having the sufficient WTP to make a call is defined as the access probability. The access probability of service j is given as*

$$\alpha_{P_j} = 1 - F_W(u_{p_j} b_j / \mu_j), \quad (3.3)$$

3.2 System Model

where $F_W(y) = P[\Psi_j \leq y | \psi_j, \beta_j]$ is the cumulative distribution function of the Weibull-distributed Ψ_j .

New call arrival rate λ_{0j} is the maximum arrival rate, limited by the access probability α_{Pj} . Thus, the total price-affected new call arrival rate to a service is

$$\lambda_j^T = \alpha_{Pj}(\lambda_j^n + x_j\sigma_j), \quad (3.4)$$

where x_j is the number of users in orbit and $x_j\sigma_j$ is the retrial rate of these users. The access probability can be seen as an *arrival gate* that controls the flow of price-affected arrivals of new users to the system. By varying the admission price per bandwidth time u_{Pj} for every state in the system, the new arrival rate of a particular service can be encouraged or discouraged. Note that handoff arrivals are not price-affected.

Our pricing policy relies on the operator's ability to estimate the statistical distribution of users' WTP. In reality, this information can be obtained in a number of ways. For example, information on users' call budget can be extracted from a network operator's historical data on users' spending patterns. With suitable incentives offered by the network operator, users can also willingly share their WTP with their operator. Although it is expected that most users would like to spend as little as possible and some would only indicate their minimum WTP, higher-end users would place a higher value on a call during congestion when their initial WTP is not sufficient. This procedure can be automated by including a simple call-budgeting program into users' mobile device.

The objective is to exercise call admission and pricing to maximise the long-term expected revenue. Intuitively, an optimal integrated admission and price control policy should regulate demand by providing incentives for users to access the network when bandwidth utilisation is low and defer low-valued calls when bandwidth utilisation is high so that resources can be allocated to users who value them most. We assume that a fixed block of bandwidth is allocated to each call for the entire duration of that call, meaning that the quality of service (QoS) of all calls is guaranteed by the network. Hence, our approach fits well with a conventional, circuit-switched network but also applies to network in which a virtual circuit is provided to each call.

3.3 Markov Decision Problem Formulation

In this section, we formulate our problem as an infinite-horizon MDP with a finite set of states. MDP has been a popular paradigm for sequential decision making problems under uncertainty and dynamic programming provides a framework for studying such problems. The state evolves through time according to given transition probabilities, a function of the system parameters and state-dependent CAC and pricing decisions \mathbf{u}_c and \mathbf{u}_p described. The interval between two transitions is referred to as a *stage*.

3.3.1 State Space and Transition Rates

We first define the set of states where the system is full and no new or handoff calls can be admitted as follows:

$$\mathcal{S}_{\text{full}} = \{(\mathbf{x}, \mathbf{n}) \mid \sum_{j=1}^J n_j b_j = B\}. \quad (3.5)$$

The system can be described by a Markov Chain on the state space:

$$\mathcal{S} = \{(\mathbf{x}, \mathbf{n}) : 0 \leq x_j \leq X_j, (\mathbf{x}, \mathbf{n}) \notin \mathcal{S}_{\text{full}}, \forall j = 1, \dots, J\}.$$

where vectors $\mathbf{n} = (n_1, \dots, n_J)^T$, $\mathbf{x} = (x_1, \dots, x_J)^T$ and $\mathbf{b} = (b_1, \dots, b_J)^T$ represent the number of active connections, the number of users in orbit and the amount of bandwidth required for all services respectively. To ensure that the state space remains finite, we limit the number of users in orbit to \mathbf{X} . When the number of users of service j in orbit reaches X_j , the blocked calls will be lost and no users from other services can substitute in and they have no further influence on the system. This truncation method is often used in the analysis of retrial systems to reduce the complexity involved [31]. Methods for choosing appropriate level of truncation are discussed in [80] and [15].

We will now derive the transition rates from a state $s = (\mathbf{x}, \mathbf{n})$. The number of active connections in the system increases due to arriving handoff users at a rate of

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x}, \mathbf{n} + \mathbf{e}_j)) = \lambda_j^h u_{c_j}^h = \lambda_{0j}(u_{c_j}^h). \quad (3.6)$$

The value of an indicator function $I(a)$ is 1 if condition a is true and 0 otherwise. New and retrying users in orbit increase the number of connections at rates

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x}, \mathbf{n} + \mathbf{e}_j)) = \alpha_{P_j}(u_{p_j}) \lambda_j^n u_{c_j}^n = \lambda_{1j}(u_{p_j}, u_{c_j}^n) \quad (3.7)$$

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j)) = \alpha_{P_j}(u_{p_j}) x_j \sigma_j u_{c_j}^n = \lambda_{2j}(u_{p_j}, u_{c_j}^n) \quad (3.8)$$

3.3 Markov Decision Problem Formulation

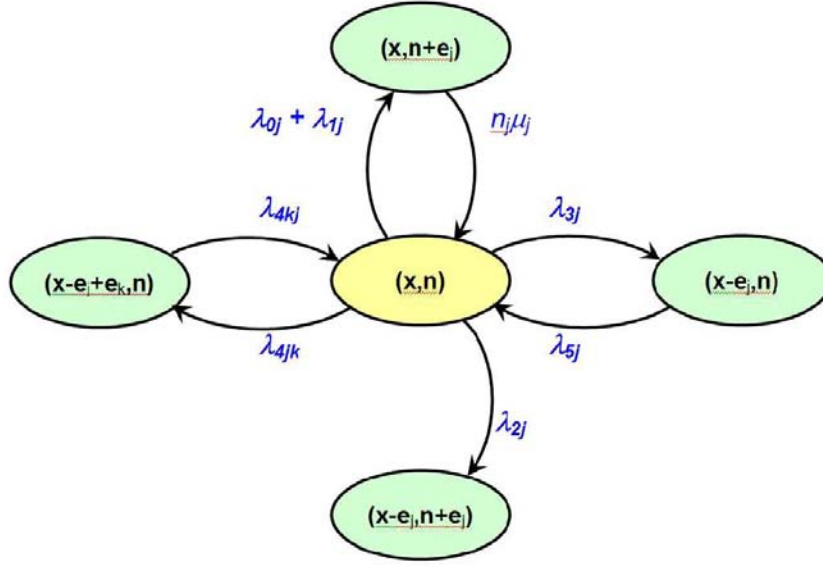


Figure 3.3. Transition Diagram of state $s = (\mathbf{x}, \mathbf{n})$.

respectively, where \mathbf{e}_j is a unit vector with 1 in its j^{th} position. The number of connections will decrease at a rate of:

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x}, \mathbf{n} - \mathbf{e}_j)) = n_j \mu_j. \quad (3.9)$$

The number of users in orbit j will only decrease if users abandon service at a rate of

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mathbf{e}_j, \mathbf{n})) = (1 - \alpha_{P_j}(u_{P_j})u_{C_j}^n) \alpha_{A_j} x_j \sigma_j = \lambda_{3j}(u_{P_j}, u_{C_j}^n) \quad (3.10)$$

or substitute out to another service k , $k \neq j$, at

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n})) = (1 - \alpha_{P_j}(u_{P_j})u_{C_j}^n) \alpha_{SO_{jk}} x_j \sigma_j = \lambda_{4jk}(u_{P_j}, u_{C_j}^n). \quad (3.11)$$

The number of users in orbit j will increase if new users retry at a rate of

$$\begin{aligned} q((\mathbf{x}, \mathbf{n}), (\mathbf{x} + \mathbf{e}_j, \mathbf{n})) &= (1 - \alpha_{P_j}(u_{P_j})u_{C_j}^n) \alpha_{R_j} \lambda_k^T + \sum_{k \neq j}^J (1 - \alpha_{P_k}(u_{P_k})u_{C_k}^n) \alpha_{SO_{kj}} \lambda_k^T \\ &= \lambda_{5j}(\mathbf{u}_P, \mathbf{u}_C^n) \end{aligned} \quad (3.12)$$

or users of another service, say k , where $k \neq j$, substitute in at a rate of

$$q((\mathbf{x}, \mathbf{n}), (\mathbf{x} + \mathbf{e}_j - \mathbf{e}_k, \mathbf{n})) = (1 - \alpha_{P_k}(u_{P_k})u_{C_k}^n) \alpha_{SO_{kj}} x_k \sigma_k = \lambda_{4kj}(u_{P_k}, u_{C_k}^n). \quad (3.13)$$

Note that (3.11) and (3.13) are actually the same. In order to avoid calculating the same rates twice, we only need to consider the rates associated with a user substituting

out for each service. We also note that transitions $q((\mathbf{x}, \mathbf{n}), (\mathbf{x} - \mathbf{e}_k, \mathbf{n} + \mathbf{e}_j))$ are disallowed and set to zero. This means that users substituting in from service k need to enter the orbit of service j before reattempting. Users who reattempt unsuccessfully and decide to reattempt again and new users who abandon do not change the state of the system, and therefore need not be considered. With reference to Fig. 3.3, the transition rates from state $s = (\mathbf{x}, \mathbf{n})$ are summarised as follows:

$$q(s, s') = \begin{cases} -\nu(s) & \text{if } s' = (\mathbf{x}, \mathbf{n}) \\ \lambda_{0j}(u_{c_j}^h) + \lambda_{1j}(u_{p_j}, u_{c_j}^n) & \text{if } s' = (\mathbf{x}, \mathbf{n} + \mathbf{e}_j) \\ \lambda_{2j}(u_{p_j}, u_{c_j}^n) & \text{if } s' = (\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j) \\ \lambda_{3j}(u_{p_j}, u_{c_j}^n) & \text{if } s' = (\mathbf{x} - \mathbf{e}_j, \mathbf{n}) \\ \lambda_{4jk}(u_{p_j}, u_{c_j}^n) & \text{if } s' = (\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) \\ \lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c^n) & \text{if } s' = (\mathbf{x} + \mathbf{e}_j, \mathbf{n}) \\ n_j \mu_j I(n_j > 0) & \text{if } s' = (\mathbf{x}, \mathbf{n} - \mathbf{e}_j) \\ 0 & \text{otherwise.} \end{cases} \quad (3.14)$$

where the total transition rate out of state s is given by

$$\begin{aligned} \nu(s) = & \sum_{j=1, k \neq j}^J [\lambda_{0j}(u_{c_j}^h) + \lambda_{1j}(u_{p_j}, u_{c_j}^n) + \lambda_{2j}(u_{p_j}, u_{c_j}^n) + \lambda_{3j}(u_{p_j}, u_{c_j}^n) + \lambda_{4jk}(u_{p_j}, u_{c_j}^n) \\ & + \lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c^n) + n_j \mu_j]. \end{aligned} \quad (3.15)$$

3.3.2 Event and Control Space

Even though the process evolves in continuous time, we only have to consider the state of the network when certain events take place. We say that an event happens at a certain time if any of the transitions derived in the previous section occurs. Let Ω denote the set of possible events, i.e. $\Omega = \{\omega \mid \omega \in \{0, 1, 2, 3, 4, 5, 6, 7\}^J\}$. The list of possible events corresponds to the possible transitions outlined previously, i.e. event 0 indicates a handoff arrival, 1 indicates a new arrival, 2 indicates an arrival from orbit and so on. Event 7 indicates no event occurring. For each state $s \in \mathcal{S}$ and event $\omega \in \Omega$, $U(s, \omega)$ is the set of available decisions:

$$U(s, \omega) = \begin{cases} \{\mathcal{U}_c \times \mathcal{U}_p\} & \text{if } \omega \in \Omega_a \\ \{\mathcal{U}_p\} & \text{if } \omega \notin \Omega_a, \end{cases} \quad (3.16)$$

3.3 Markov Decision Problem Formulation

where Ω_a is the set of all events consisting of arrival events 1 or 2. \mathcal{U}_c and \mathcal{U}_p denote the set of all possible call admission and price control decisions and are defined as $\mathcal{U}_c = \{\mathbf{u}_c \mid u_{c_j}^n \in \{0, 1\}, \forall j\}$ and $\mathcal{U}_p = \{\mathbf{u}_p \mid u_{p_j} \in \mathcal{U}_{p_j}, \forall j\}$ respectively.

Given that the system is in state $s \in \mathcal{S}$ with control actions $u \in \mathcal{U}$ available and an event $\omega \in \Omega$ occurred, the next state, $s' \in \mathcal{S}$, is given by a function $f : \mathcal{S} \times \Omega \times \mathcal{U}$ such that

$$f(s, \omega, \mathbf{u}) = \begin{cases} (\mathbf{x}, \mathbf{n} + \mathbf{e}_j) & \text{if } \omega_j = 0, u_{c_j}^h = 1 \\ (\mathbf{x}, \mathbf{n} + \mathbf{e}_j) & \text{if } \omega_j = 1, u_{c_j}^n = 1 \\ (\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j) & \text{if } \omega_j = 2, u_{c_j}^n = 1 \\ (\mathbf{x} - \mathbf{e}_j, \mathbf{n}) & \text{if } \omega_j = 3 \\ (\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) & \text{if } \omega_j = 4 \\ (\mathbf{x} + \mathbf{e}_j, \mathbf{n}) & \text{if } \omega_j = 5 \\ (\mathbf{x}, \mathbf{n} - \mathbf{e}_j) & \text{if } \omega_j = 6 \\ (\mathbf{x}, \mathbf{n}) & \text{otherwise.} \end{cases} \quad (3.17)$$

3.3.3 Revenue Maximisation Problem

Using uniformisation [12], the continuous-time MDP can be transformed into its discrete-time equivalence with the so-called uniform transition rate, where the total transition rate out of any state is bounded by ν . The transition probabilities $p(s, \omega, \mathbf{u})$ for state s are then given by

$$p(s, \omega, \mathbf{u}) = \begin{cases} \frac{\lambda_{0j}(u_{c_j}^h)}{\nu} & \text{if } \omega_j = 0 \\ \frac{\lambda_{1j}(u_{p_j}, u_{c_j}^n)}{\nu} & \text{if } \omega_j = 1 \\ \frac{\lambda_{2j}(u_{p_j}, u_{c_j}^n)}{\nu} & \text{if } \omega_j = 2 \\ \frac{\lambda_{3j}(u_{p_j}, u_{c_j}^n)}{\nu} & \text{if } \omega_j = 3 \\ \frac{\lambda_{4jk}(u_{p_j}, u_{c_j}^n)}{\nu} & \text{if } \omega_j = 4 \\ \frac{\lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c)}{\nu} & \text{if } \omega_j = 5 \\ \frac{n_j \mu_j}{\nu} & \text{if } \omega_j = 6 \\ 1 - \frac{\nu(s)}{\nu} & \text{and if } \omega_j = 7, \end{cases} \quad (3.18)$$

where the total transition rate out of state $s \in \mathcal{S}$ is given by (3.15).

The revenue rate collected by the system at a particular state $s = (\mathbf{x}, \mathbf{n})$ is given by the reward function

$$g(s, \omega, \mathbf{u}) = \begin{cases} SR_j & \text{if } \omega_j = 0 \text{ and } u_{c_j}^h = 1 \\ r_j(u_{p_j}(s)) = u_{p_j}(s)b_j/\mu_j & \text{if } \omega_j = 1, 2 \text{ and } u_{c_j}^n = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (3.19)$$

Reward $r_j(u_{p_j}(s))$ is the revenue collected when a user of class j is admitted. In order to reflect the higher importance of accepting handoff calls, SR_j should be greater than the actual revenue provided by the admission of new call requests of class j . A user is admitted based on a single admission price that will not change for the entire duration of the call.

Proposition 3.1. *All policies $\Lambda = \{\mathbf{u} = (\mathbf{u}_c(s), \mathbf{u}_p(s)) \mid s \in \mathcal{S}, \mathbf{u} \in \mathcal{U}\}$ are unichain if*

$$\alpha_{P_j}(u_{p_j}(s)) < 1 \quad (3.20)$$

for all states $s \in \mathcal{S}$.

Proof. A policy Λ is said to be unichain if the corresponding Markov chain has no two disjoint closed sets of states [92, 107]. When a state s_K is accessible from s_0 , i.e. $s_0 \rightarrow s_K$, there exists a sequence of states $s_1, s_2, \dots, s_{K-1} \in \mathcal{S}$ such that

$$p(s_1|s_0, \mathbf{u}(s_0))p(s_2|s_1, \mathbf{u}(s_1)), \dots, p(s_K|s_{K-1}, \mathbf{u}(s_{K-1})) > 0. \quad (3.21)$$

To prove that there is no disjoint closed set of states, we show that all states $(\mathbf{x}, \mathbf{n}) \in \mathcal{S}$ are accessible from $(\mathbf{0}, \mathbf{0})$ when condition (3.20) is satisfied. We show our proof in two parts. First, we show that $(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{x}, \mathbf{0})$ for all $(\mathbf{x}, \mathbf{n}) \in \mathcal{S}$. Next, we prove that $(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{0}, \mathbf{n})$.

Part 1: For every states $(\mathbf{x}, \mathbf{n}) \in \mathcal{S}$ and service j , (\mathbf{x}, \mathbf{n}) is accessible from $(\mathbf{x}, \mathbf{0})$, i.e. $(\mathbf{x}, (n_1, \dots, n_j, \dots, n_J)) \rightarrow (\mathbf{x}, (n_1, \dots, 0, \dots, n_J))$, $n_j > 0$,

$$\begin{aligned} & p((\mathbf{x}, \mathbf{n} - \mathbf{e}_j)|(\mathbf{x}, \mathbf{n}))p((\mathbf{x}, \mathbf{n} - 2\mathbf{e}_j)|(\mathbf{x}, \mathbf{n} - \mathbf{e}_j)) \dots p((\mathbf{x}, \mathbf{n} - n_j\mathbf{e}_j)|(\mathbf{x}, \mathbf{n} - (n_j - 1)\mathbf{e}_j)) \\ &= \frac{n_j\mu_j}{\nu} \frac{(n_j - 1)\mu_j}{\nu} \dots \frac{\mu_j}{\nu} = n_j! \left(\frac{\mu_j}{\nu}\right)^{n_j} > 0. \end{aligned} \quad (3.22)$$

The probability of a user of service j departing from service is independent of the optimal control \mathbf{u} . Since $(\mathbf{x}, (n_1, \dots, n_j, \dots, n_J)) \rightarrow (\mathbf{x}, (n_1, \dots, 0, \dots, n_J))$ for all $j \in \{1, \dots, J\}$, it is easy to see that

$$(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{x}, \mathbf{0}). \quad (3.23)$$

3.3 Markov Decision Problem Formulation

Part 2: Denote $s_k = (\mathbf{x} - k\mathbf{e}_j, \mathbf{n})$, $k = 0, \dots, x_j$, the probability of accessing state $((x_1, \dots, 0, \dots, x_j), \mathbf{n})$ from state $((x_1, \dots, x_j, \dots, x_j), \mathbf{n})$ is given by

$$\begin{aligned}
& p(s_1|s_0, \mathbf{u}(s_0))p(s_2|s_1, \mathbf{u}(s_1)) \dots p(s_{x_j}|s_{x_j-1}, \mathbf{u}(s_{x_j-1})) \\
&= \frac{(1 - \alpha_{P_j}(u_{P_j}(s_0))u_{C_j}^n(s_0))\alpha_{A_j}x_j\sigma_j}{\nu} \dots \frac{(1 - \alpha_{P_j}(u_{P_j}(s_{x_j-1}))u_{C_j}^n(s_{x_j-1}))\alpha_{A_j}\sigma_j}{\nu} \\
&= x_j! \left(\frac{\alpha_{A_j}\sigma_j}{\nu}\right)^{x_j} \prod_{k=0}^{x_j-1} [1 - \alpha_{P_j}(u_{P_j}(s_k))u_{C_j}^n(s_k)]. \tag{3.24}
\end{aligned}$$

It is obvious that $(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{x} - x_j\mathbf{e}_j, \mathbf{n})$ when $\alpha_{P_j}(u_{P_j}(s_k)) < 1$ for all $k = 1, \dots, x_j - 1$ and $u_{C_j}^n(s_k) \in \{0, 1\}$. In that case, $(\mathbf{0}, \mathbf{n})$ is accessible from all states $(\mathbf{x}, \mathbf{n}) \in \mathcal{S}$, i.e.

$$(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{0}, \mathbf{n}). \tag{3.25}$$

Based on (3.23) and (3.25), we can conclude that all states $(\mathbf{x}, \mathbf{n}) \in \mathcal{S}$ are accessible from $(\mathbf{0}, \mathbf{0})$ and vice versa, i.e.

$$(\mathbf{x}, \mathbf{n}) \rightarrow (\mathbf{0}, \mathbf{0}) \tag{3.26}$$

when $\alpha_{P_j}(u_{P_j}(s_k)) < 1$. Therefore, no disjoint closed set of states exists for all policies when condition (3.20) is satisfied. \square

For unichain models, all stationary policies have constant gain. It is shown in [92] that whenever a stationary policy has nonconstant gain, a stationary policy can be constructed to dominate the nonconstant gain policy. For simplicity, we assume that condition (3.20) is always satisfied by only allowing prices $u_{P_j}(s_k)$ that result in $\alpha_{P_j}(u_{P_j}(s_k)) < 1$. The average reward-to-go function is given by the Bellman equation:

$$J^* + h(s) = \max_{\mathbf{u} \in \mathcal{U}(s, \omega)} \left[\sum_{\omega \in \Omega} p(s, \omega, \mathbf{u}) [g(s, \omega, \mathbf{u}) + h(f(s, \omega, \mathbf{u}))] \right] \tag{3.27}$$

where J^* and $h(s)$ denote the optimal expected reward per stage and the relative or differential reward rate of state $s \in \mathcal{S}$, respectively. A stage here means a transition in the uniformised chain.

For example, when $s = (\mathbf{x}, \mathbf{n})$, the Bellman equation becomes

$$\begin{aligned}
J^* + h(\mathbf{x}, \mathbf{n}) &= \max_{\mathbf{u}_p, \mathbf{u}_c} \left[\sum_{j=1}^J \left(\frac{\lambda_{0j}(u_{c_j}^h)}{\nu} [SR_j + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j)] + \frac{\lambda_{1j}(u_{p_j}, u_{c_j}^n)}{\nu} [r_j(u_{p_j}(s)) + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j)] \right) \right. \\
&+ \sum_{j=1}^J \left(\frac{\lambda_{2j}(u_{p_j}, u_{c_j}^n)}{\nu} [r_j(u_{p_j}(s)) + h(\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j)] + \frac{\lambda_{3j}(u_{p_j}, u_{c_j}^n)}{\nu} h(\mathbf{x} - \mathbf{e}_j, \mathbf{n}) \right) \\
&+ \sum_{j=1}^J \sum_{k \neq j}^J \frac{\lambda_{4jk}(u_{p_j}, u_{c_j}^n)}{\nu} h(\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) + \sum_{j=1}^J \frac{\lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c^n)}{\nu} h(\mathbf{x} + \mathbf{e}_j, \mathbf{n}) \\
&\left. + \sum_{j=1}^J \frac{n_j \mu_j}{\nu} h(\mathbf{x}, \mathbf{n} - \mathbf{e}_j) + \left(1 - \frac{\nu(\mathbf{x}, \mathbf{n})}{\nu} \right) h(\mathbf{x}, \mathbf{n}) \right],
\end{aligned}$$

where $\nu(\mathbf{x}, \mathbf{n})$ is defined in (3.15). Rearranging the previous equation, we have

$$\begin{aligned}
J^* + h(\mathbf{x}, \mathbf{n}) &= \sum_{j=1}^J \frac{n_j \mu_j}{\nu} [h(\mathbf{x}, \mathbf{n} - \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] + h(\mathbf{x}, \mathbf{n}) \\
&+ \max_{\mathbf{u}_p, \mathbf{u}_c} \left[\sum_{j=1}^J \frac{\lambda_{0j}(u_{c_j}^h)}{\nu} [SR_j + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] \right. \\
&+ \sum_{j=1}^J \frac{\lambda_{1j}(u_{p_j}, u_{c_j}^n)}{\nu} [r_j(u_{p_j}(s)) + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] \\
&+ \sum_{j=1}^J \frac{\lambda_{2j}(u_{p_j}, u_{c_j}^n)}{\nu} [r_j(u_{p_j}(s)) + h(\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] \\
&+ \sum_{j=1}^J \frac{\lambda_{3j}(u_{p_j}, u_{c_j}^n)}{\nu} [h(\mathbf{x} - \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \\
&+ \sum_{j=1}^J \sum_{k \neq j}^J \frac{\lambda_{4jk}(u_{p_j}, u_{c_j}^n)}{\nu} [h(\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \\
&\left. + \sum_{j=1}^J \frac{\lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c^n)}{\nu} [h(\mathbf{x} + \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \right].
\end{aligned}$$

The Bellman equation can then be expressed in terms of the operators for handoff and new users, $H_j^h[h(\mathbf{x}, \mathbf{n})]$ and $H_j^n[h(\mathbf{x}, \mathbf{n})]$, as follows:

$$\begin{aligned}
J^* + h(\mathbf{x}, \mathbf{n}) &= h(\mathbf{x}, \mathbf{n}) + \\
&\frac{1}{\nu} \sum_{j=1}^J [\lambda_j^h H_j^h[h(\mathbf{x}, \mathbf{n})] + H_j^n[h(\mathbf{x}, \mathbf{n})] + n_j \mu_j [h(\mathbf{x}, \mathbf{n} - \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})]].
\end{aligned}$$

3.3 Markov Decision Problem Formulation

The operator $H_j^h[h(\mathbf{x}, \mathbf{n})]$ represents the choice between accepting and rejecting a handoff users, taking into account the immediate and future expected reward. Given that $\sum_{j=1}^J n_j b_j \leq B - b_j$, $H_j^h[h(\mathbf{x}, \mathbf{n})]$ is given by

$$H_j^h[h(\mathbf{x}, \mathbf{n})] = \max_{u_{c_j}^h \in \{0,1\}} [u_{c_j}^h (SR_j + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n}))]. \quad (3.28)$$

Therefore, a handoff user will only be admitted only if the satisfaction revenue it provides exceeds the loss of expected revenue, or *opportunity cost*, due to an additional user. The optimal handoff CAC action is given by

$$u_{c_j}^h = \begin{cases} 1 & \text{if } SR_j > h(\mathbf{x}, \mathbf{n}) - h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) \\ 0 & \text{otherwise.} \end{cases} \quad (3.29)$$

$H_{jk}^n[h(\mathbf{x}, \mathbf{n})]$, represents the integrated CAC and pricing control operator for new users, is defined as follows:

$$\begin{aligned} H_j^n[h(\mathbf{x}, \mathbf{n})] = \max_{\mathbf{u}_p, \mathbf{u}_c} & \left[\lambda_{1j}(u_{p_j}, u_{c_j}^n) [r_j(u_{p_j}(s)) + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] + \lambda_{2j}(u_{p_j}, u_{c_j}^n) r_j(u_{p_j}(s)) \right. \\ & + \lambda_{2j}(u_{p_j}, u_{c_j}^n) [h(\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] + \lambda_{3j}(u_{p_j}, u_{c_j}^n) [h(\mathbf{x} - \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \\ & \left. + \sum_{k \neq j}^J \lambda_{4jk}(u_{p_j}, u_{c_j}^n) [h(\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] + \lambda_{5j}(\mathbf{u}_p, \mathbf{u}_c^n) [h(\mathbf{x} + \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \right]. \end{aligned}$$

If the system is allowed to set a very high price u_{p_j} such that the corresponding access probability $\alpha_{p_j}(u_{p_j}) = 0$, price can be used as the sole admission control parameter for new users. In that case,

$$u_{c_j}^n(s) = I(\alpha_{p_j}(u_{p_j}(s)) = 0) I(s \notin \mathcal{S}_{\text{full}}) \quad (3.30)$$

and the operator $H_{jk}^n[h(\mathbf{x}, \mathbf{n})]$ reduces to

$$\begin{aligned} H_j^n[h(\mathbf{x}, \mathbf{n})] = \max_{\mathbf{u}_p} & \left[\lambda_{1j}(u_{p_j}) [r_j(u_{p_j}(s)) + h(\mathbf{x}, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] \right. \\ & + \lambda_{2j}(u_{p_j}) [r_j(u_{p_j}(s)) + h(\mathbf{x} - \mathbf{e}_j, \mathbf{n} + \mathbf{e}_j) - h(\mathbf{x}, \mathbf{n})] + \lambda_{3j}(u_{p_j}) [h(\mathbf{x} - \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \\ & \left. + \sum_{k \neq j}^J \lambda_{4jk}(u_{p_j}) [h(\mathbf{x} - \mathbf{e}_j + \mathbf{e}_k, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] + \lambda_{5j}(\mathbf{u}_p) [h(\mathbf{x} + \mathbf{e}_j, \mathbf{n}) - h(\mathbf{x}, \mathbf{n})] \right]. \quad (3.31) \end{aligned}$$

For any stationary policy for unichain Markov models, the optimal expected revenue per stage is independent of the initial state. It has been argued that the standard infinite-horizon average reward dynamic programming theory applies and there exists a stationary

policy which is optimal [12]. There are a number of methods for obtaining the values of J^* and $h(s)$. Most of them are value iteration and policy iteration algorithms, which can be computed off-line and before the system starts to operate. A policy Λ^* is said to be optimal if $J_{\Lambda^*}^*(s) \geq J_{\Lambda}^*(s)$ for every other policy Λ . If policy iteration is used, we can simplify (4.19) by exploiting the repetitive Quasi-Birth-Death (QBD) structure of the chain (see Section 3.4) [125]. (4.19) can be rewritten using dynamic programming operator T , which maps the set of functions on the state space to itself:

$$(Th)(s) = J^* + h(s). \quad (3.32)$$

Then, $(Th^k)(s)$ is just the total optimal expected average reward in a k -stage problem, starting from state s . The natural version of the value iteration method for the average reward problem is simply to generate successively the finite horizon optimal costs $T^k h_o$, $k = 1, 2, \dots$, starting with the zero function h_o . It is then intuitive that the k -stage average reward $\lim_{k \rightarrow \infty} \frac{T^k h_o}{k}$ converges to the optimal average reward vector (as proven in [12]). However, some components of $T^k h_o$ typically diverge to ∞ or $-\infty$ and therefore direct computation of $\lim_{k \rightarrow \infty} \frac{T^k h_o}{k}$ is numerically impractical. To solve this problem, we introduce the minimum and maximum error bounds:

$$\underline{c}_k = \min_s [(Th^k)(s) - h^k(s)], \quad (3.33)$$

$$\bar{c}_k = \max_s [(Th^k)(s) - h^k(s)] \quad (3.34)$$

and stop the iteration when $\bar{c}_k - \underline{c}_k \leq \epsilon$.

3.4 Calculation of Stationary Probabilities

In this section, we will present the method for deriving performance measures of the system from its stationary distribution, which will be used in the results analysis in Section 4.5. Firstly, we analyse the Level-Dependent Quasi-Birth-Death (LDQBD) structure of the system model in Section 3.4.1. Secondly, we present the matrix analytic methods used in 3.4.2. Finally, we provide the derivation of important system characteristics from the stationary distribution in 3.4.3.

3.4 Calculation of Stationary Probabilities

3.4.1 Level-Dependent Quasi-Birth-Death Process

The system model is found to have a general LDQBD structure (see [65] and [15]). An LDQBD differs from a QBD process in that the transition rates at each *level* can be dependent upon the level the process is in. The level of a QBD process is usually indicated by its first entry in the system state. In our case, a level is best indicated by the number of users in orbit of Service 1. The LDQBD process has an infinitesimal generator \mathbf{Q} of the block partitioned form:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1^{(0)} & \mathbf{Q}_0^{(0)} & \mathbf{0} & & \\ \mathbf{Q}_2^{(1)} & \mathbf{Q}_1^{(1)} & \mathbf{Q}_0^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{Q}_2^{(X_1-1)} & \mathbf{Q}_1^{(X_1-1)} & \mathbf{Q}_0^{(X_1-1)} \\ & & \mathbf{0} & \mathbf{Q}_2^{(X_1)} & \mathbf{Q}_1^{(X_1)} \end{pmatrix} \quad (3.35)$$

where $\mathbf{Q}_0^{(x_1)}$ and $\mathbf{Q}_2^{(x_1)}$ are non-negative matrices. $\mathbf{Q}_1^{(x_1)}$ has non-negative off-diagonal entries. They give the rates of going up one level (forward transitions), staying in the same level (local transitions) or going down one level (backward transitions) respectively. We say that the process is skip-free in the levels. Forward and backward transitions are events that are associated with the arrival and departure of a user in the orbit of service 1 respectively while local transitions include all other events. The transition rates have been described in (3.14). The states can be partitioned into levels:

$$l(k) = \{s = (k, x_2, \dots, x_J, n_1, \dots, n_J) : s \in \mathcal{S}\}. \quad (3.36)$$

Position $(x_2, \dots, x_J, n_1, \dots, n_J)$ within the level is termed its *phase*. Matrices $\mathbf{Q}_0^{(x_1)}$, $\mathbf{Q}_1^{(x_1)}$ and $\mathbf{Q}_2^{(x_1)}$ are of size $L \times L$, where $L = \prod_{j=0}^J (X_j + 1)M$ and M is the number of possible values of vector \mathbf{n} given that the system has a capacity of B . The detailed entries of $\mathbf{Q}_0^{(x_1)}$, $\mathbf{Q}_2^{(x_1)}$ and $\mathbf{Q}_1^{(x_1)}$ for a two-service network are detailed in the Appendix.

3.4.2 Matrix Analytic Methods

In order to obtain stationary probabilities vector π , we need to solve the simultaneous equations $\pi\mathbf{Q} = 0$ and $\pi\mathbf{e} = 1$. The key to a general solution for generator \mathbf{Q} is the assumption that a geometric relation holds among the stationary probability vectors $\pi^{(x_1)}$, expressed as follows:

$$\pi^{(x_1)} = \pi^{(x_1-1)}\mathbf{R}_{x_1-1}, \quad x_1 \geq 1. \quad (3.37)$$

The family of $\{\mathbf{R}_{x_1}, x_1 \geq 0\}$ matrices is defined as the expected sojourn time in k per unit of sojourn time in $k - 1$ and satisfy the following non-linear matrix equation:

$$\mathbf{Q}_0^{(x_1)} + \mathbf{R}_{x_1} \mathbf{Q}_1^{(x_1)} + \mathbf{R}_{x_1} \mathbf{R}_{x_1+1} \mathbf{Q}_2^{(x_1+1)} = 0, x_1 \geq 0. \quad (3.38)$$

\mathbf{R}_{x_1} is the minimal non-negative solution of the above equation, whose solutions are not necessarily unique. The matrices $\{\mathbf{U}_{x_1}, x_1 \geq 1\}$ are defined by

$$\mathbf{U}_{x_1} = \mathbf{Q}_1^{(x_1)} + \mathbf{R}_{x_1} \mathbf{Q}_2^{(x_1+1)}, x_1 \geq 1. \quad (3.39)$$

Suppose that the LDQBD starts in level x_1 and we observe the process only at time points when it is in level x_1 , before it visits level $x_1 - 1$ for the first time. If we call this partial process B_k , then \mathbf{U}_{x_1} is the infinitesimal generator for B_{x_1} . Assuming that we know truncation level \mathbf{X} , the stationary vector π can be calculated iteratively using the equations defined earlier using the following algorithm [65]:

Step 1 Initialise $\mathbf{U}_{X_1} = \mathbf{Q}_1^{(X_1)}$.

Step 2 Recursively calculate \mathbf{R}_k for $k = X_1 - 1$ to 0 using $\mathbf{R}_k = \mathbf{Q}_0^{(k-1)}(-\mathbf{U}_{k+1})^{-1}$ and $\mathbf{U}_k = \mathbf{Q}_1^{(k)} + \mathbf{R}_k \mathbf{Q}_2^{(k+1)}$.

Step 3 Solve $\pi^{(0)}(\mathbf{Q}_1^{(0)} + \mathbf{R}_0 \mathbf{Q}_2^{(1)}) = \mathbf{0}$ subject to normalizing equation $\pi \mathbf{e} = \sum_{k=0}^{X_1} \pi^{(k)} = \pi^{(0)} \sum_{k=0}^{X_1} \left[\prod_{m=0}^{k-1} \mathbf{R}_m \mathbf{1} \right] = 1$.

Step 4 For $k = 1$ to X_1 , compute $\pi^{(k)}$ using $\pi^{(k)} = \pi^{(k-1)} \mathbf{R}_{k-1}$.

3.4.3 System Characteristics

The new and handoff call blocking probabilities can be calculated using the stationary distribution under a given policy Λ .

Definition 3.4. *The handoff call blocking probability of service class j is defined as*

$$P_{H_j} = \sum_{s \in \mathcal{S}} (1 - u_{c_j}^h(s)) \pi(s) \quad (3.40)$$

where $\pi(s)$ is the stationary probability of state $s \in \mathcal{S}$.

3.4 Calculation of Stationary Probabilities

Definition 3.5. The new call blocking probability of service class j under admission control $u_{c_j}^n(s)$, $s \in \mathcal{S}$, is given as

$$P_{B_j} = \sum_{s \in \mathcal{S}} (1 - u_{c_j}^n(s)) \pi(s), \quad (3.41)$$

where $\pi(s)$ is the stationary probability of state $s \in \mathcal{S}$.

Table 3.1. System parameters derived using stationary distribution of the system.

System Parameter	Definition
Expected price	$\bar{P}_j = \sum_{s \in \mathcal{S}} u_{p_j}(s) \pi(s)$
Channel utilisation	$U = \sum_{c=1}^B \sum_{s \in \mathcal{S}_{\text{full}}} \frac{c}{B} \pi(s)$
Expected number of users in orbit	$\bar{X}_j = \sum_{i=0}^{X_j} \sum_{s \in \mathcal{S}} i \pi(s x_j = i)$
Expected number of users in service	$\bar{N}_j = \sum_{i=0}^{\lfloor B/b_j \rfloor} \sum_{s \in \mathcal{S}} i \pi(s n_j = i)$
Expected number of users in system	$N_T = \sum_{j=0}^J \bar{X}_j + \bar{N}_j$
Mean time spent by users in orbit	$W_{oj} = \frac{\bar{X}_j}{\bar{\alpha}_{P_j} (\lambda_{1j} + \bar{X}_j \sigma_j)}$
Mean time spent by users in system	$W_{sj} = \frac{\bar{X}_j + \bar{N}_j}{\bar{\alpha}_{P_j} (\lambda_{1j} + \bar{X}_j \sigma_j)}$
Mean number of retrials	$r_j = \sigma_j W_{oj}$

Other important system parameters are presented in Table 3.1. Note that the mean time spent by an arbitrary network user in orbit, W_{oj} and in the system, W_{sj} , are derived using Little's Theorem [94]. The theorem states that the average number of users can be determined from the product of the average user arrival rate and time spent in the service.

3.5 Numerical Results

In this section, we will firstly discuss the choice of parameters in our numerical experiments. Then, we analyse the results in terms of reward maximisation in Section 3.5.1, congestion control in Section 3.5.2, exponential WTP distribution in Section 3.5.4 and price discrimination in Section 3.5.3.

We consider a simple problem with two services because the size of the state space grows exponentially with the number of services in the system. We compare our *Optimal Call Admission and Dynamic Pricing* (OCADP) policy numerically in the revenue maximisation problem with three other policies:

1. Always Accept and Static Pricing (AASP)
2. Optimal Call Admission and Static Pricing (OCASP)
3. Always Accept and Dynamic Pricing (AADP)

Definition 3.6. *An Always Accept admission policy always admit an additional user if sufficient bandwidth is available, i.e.*

$$u_{c_j}(s) = I(s \notin \mathcal{S}_{\text{full}}). \quad (3.42)$$

This strategy is equivalent to having no call admission control at all. On the other hand, the Optimal Call Admission policy, as defined in Definition 3.1, is a state-dependent policy that may, depending on the network parameters, reject a connection request even though sufficient bandwidth is available. For both AASP and OCASP, we use a standard static pricing mechanism that sets the price per bandwidth time to the average price. For the dynamic pricing setting (see Definition 3.2) in AADP and OCADP, we allow price u_{p_j} and $\alpha_{p_j}(u_{p_j})$ to vary in order to maximise the long-term expected revenue.

To illustrate the benefits of our integrated policy, we also purposely set up service 2 to generate higher expected revenue than service 1. In this simulation, the SR of each service is set to 3 times of the expected reward, r_j , to reflect the higher priority of handoff calls. However, as SR increases, so does the number of guard channels and, consequently, more new calls will be blocked. All other parameters are listed in Table 3.2. We will analyse the results for two WTP distributions (see Fig. 3.4), i.e. with $\beta = 3.5$ and 1.0. Both have the same mean WTP $\psi = (1, 1)$ per bandwidth time but different distribution.

3.5 Numerical Results

Table 3.2. Simulation parameters for AASP, OCASP, AADP and OCADP policies.

Simulation Parameter	Value
Total bandwidth	$B = 20$
Maximum number of users in orbit	$\mathbf{X} = (3, 4)$
New arrival rate	$\boldsymbol{\lambda}^n = (1.0, 2.0)$
Handoff arrival rate	$\boldsymbol{\lambda}^h = 0.25\boldsymbol{\lambda}^n$
Call hold time	$1/\boldsymbol{\mu} = (1, 2)$
Reattempt rate	$\boldsymbol{\sigma}_j = (0.5, 0.5)$
Bandwidth usage	$\mathbf{b} = (2, 2)$
Probability of reattempt	$\boldsymbol{\alpha}_R = (0.6, 0.6)$
Probability of substitution out	$\boldsymbol{\alpha}_{SO} = (0.2, 0.2)$
Probability of abandonment	$\boldsymbol{\alpha}_A = (0.2, 0.2)$
Satisfaction revenue	$SR_j = 3r_j, j = 1, 2$

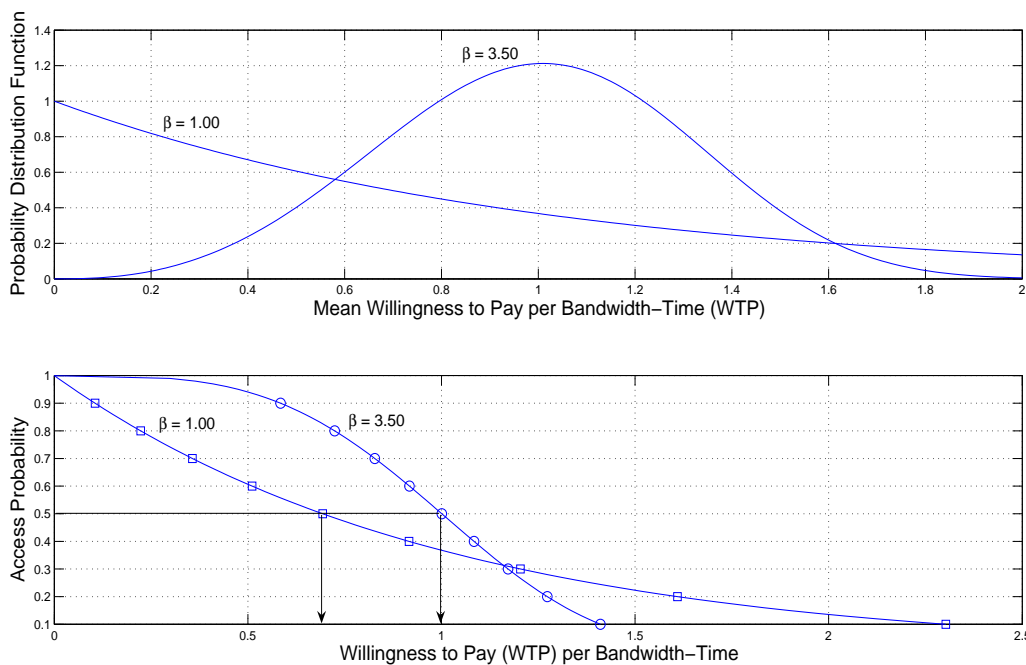


Figure 3.4. WTP pdfs and access probabilities vs. price with $\beta = 3.5$ and 1.0.

3.5.1 Policy Comparison

The advantage of integrating CAC with dynamic pricing is evident. Integrated policy OCADP, as shown in Fig. 3.5, generates the highest optimal expected reward per stage, J^* , followed by AADP, OCASP and AASP. Under policy AASP, the average reward increases at first due to the higher total new call arrival rate. However, as the new call arrival rate increases from $\lambda = (1.0, 2.0)$ to $(1.0, 8.0)$, AASP is unable to handle the additional traffic load and results in a decrease in the average reward. Policy OCASP improves over AASP by exercising optimal CAC to give priority to handoff calls and higher revenue-generating new calls of service 2. The effect of optimal CAC becomes more significant as the total new call arrival rate increases and resources become increasingly scarce.

The dynamic pricing component in AADP and OCADP further provided reward improvements over AASP and OCASP by setting higher admission prices for new calls when bandwidth becomes scarcer. Extra reward is obtained by charging lower admission prices when traffic load is light. However, as the arrival rate to the system increases, higher prices alone will no longer be sufficient to deter high WTP users. By integrating the CAC and dynamic pricing components, the optimal policy OCADP maximises revenue

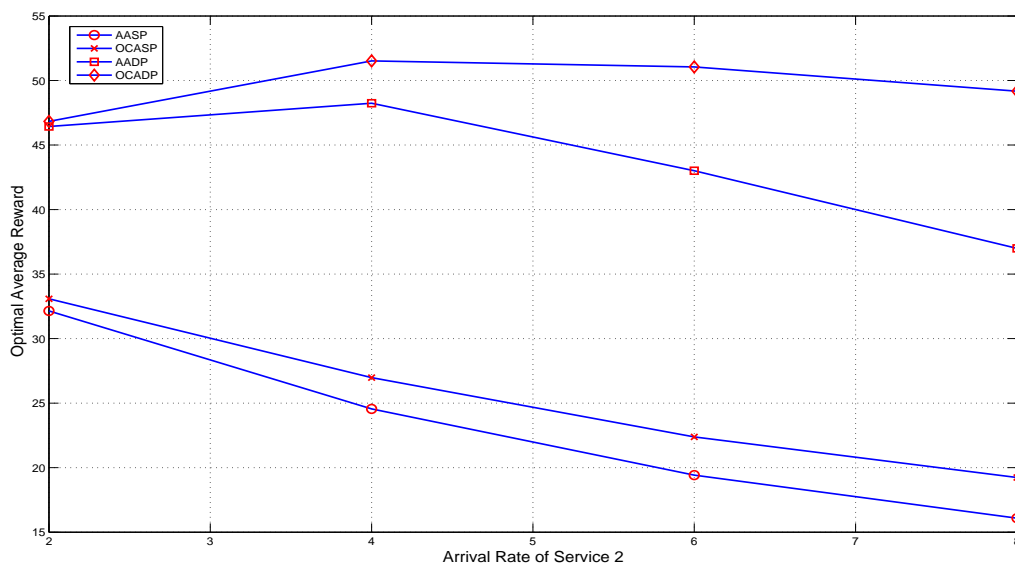


Figure 3.5. Optimal revenue per stage J^* for WTP shape $\beta = 3.5$.

3.5 Numerical Results

by exercising optimal control to give higher priority to high reward-generating new and handoff calls of service 2. The optimal policy also adjusts prices according to the system state, i.e. increasing prices as bandwidth becomes scarce to discourage arrivals of all services and decreasing prices when bandwidth consumption is low.

Assuming that the operator is allowed to charge very high prices such that users will be blocked when $\alpha_P = 0$, we focus our discussion on the CAC policy for handoff users. We illustrate the optimal CAC policy for the case of $\lambda^h = (1, 6)$ in Fig. 3.6. The policy is defined as $\mathbf{u}_c^h(s) = (u_{c_1}^h(s), u_{c_2}^h(s))$, with $u_{c_1}^h, u_{c_2}^h \in \{0, 1\}$, and can either reject users of both services with $\mathbf{u}_c^h(s) = (0, 0)$; only reject service 1 users with $\mathbf{u}_c^h(s) = (0, 1)$; only reject service 2 users with $\mathbf{u}_c^h(s) = (1, 0)$; or admit both with $\mathbf{u}_c^h(s) = (1, 1)$. In our simulation, service 2 provides higher expected reward and therefore has higher a priority. Since the SR of both services is set to three times the revenue generated from the admission of

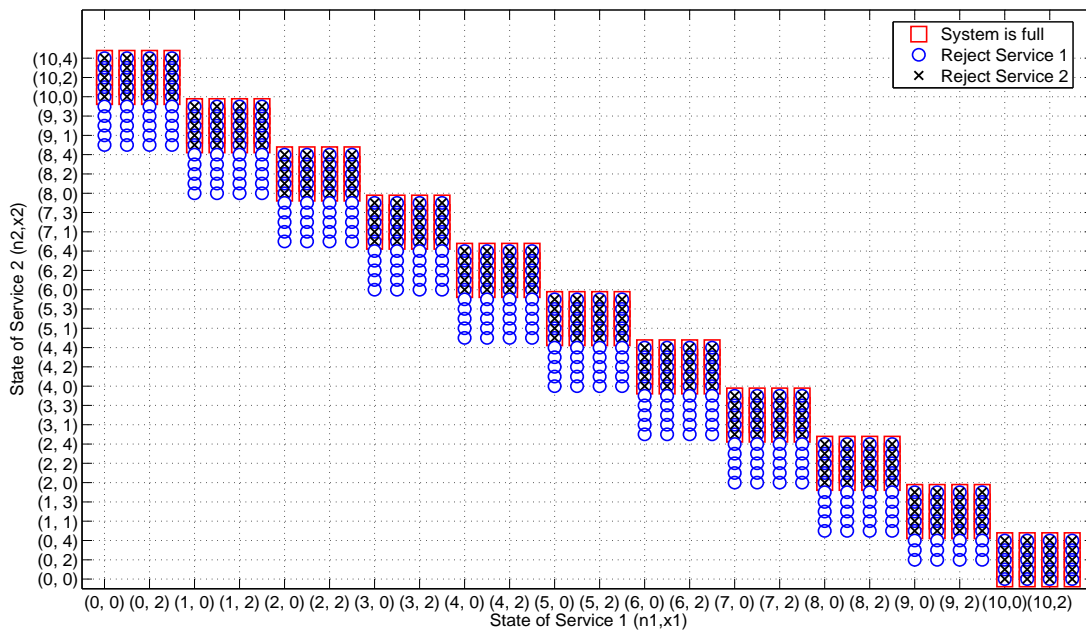


Figure 3.6. Optimal call admission policy of OCADP. The x and y axes indicate the state of service 1, $s_1 = (n_1, x_1)$, and 2, $s_2 = (n_2, x_2)$, respectively. The combination of s_1 and s_2 forms a state $s = (s_1, s_2)$ in the system. The CAC policy of each state is indicated by the following symbols: 'o' represents $\mathbf{u}_c = (0, 1)$ and 'x' represents $\mathbf{u}_c = (0, 0)$. States with '□' are states where the bandwidth of the system is fully utilised. In other states, $\mathbf{u}_c = (1, 1)$.

new users, handoff users have higher priority than the new users within the same service. However, as indicated by (3.29), the handoff users of service 1 are blocked in some states because the SR provided by them in these states is less than the expected reward gained from a future arrival of a handoff or new user of service 2. Therefore, by appropriately setting the SR of each service, the system has the flexibility to block some low priority handoff calls when the system becomes increasingly congested.

The higher priority of service 2 users is also reflected in the structure of the optimal differential reward function, $h(s)$. Referring to Fig. 3.7, we observe that $h(\mathbf{x}, \mathbf{n})$ first increases as (n_1, x_1) and (n_2, x_2) increase but then decreases as the bandwidth consumption in $s = (\mathbf{x}, \mathbf{n})$ approaches the system limit. The rate at which $h(\mathbf{x}, \mathbf{n})$ increases according to (n_2, x_2) is higher than that of (n_1, x_1) , i.e.

$$\frac{\delta h(x_1, x_2, n_1, n_2)}{(n_2, x_2)} > \frac{\delta h(x_1, x_2, n_1, n_2)}{(n_1, x_1)}. \quad (3.43)$$

Again, this shows the preference of the system to admit service 2 users in order to maximise the expected reward. As the system becomes congested, it is more profitable to deter service 1 users from receiving service by imposing higher prices and blocking them if necessary.

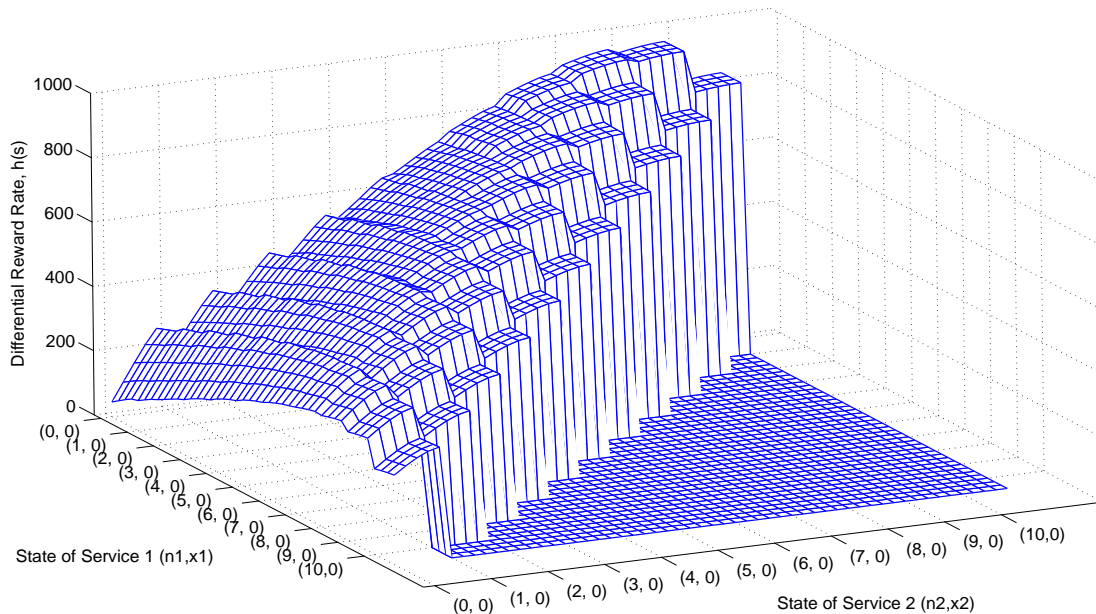
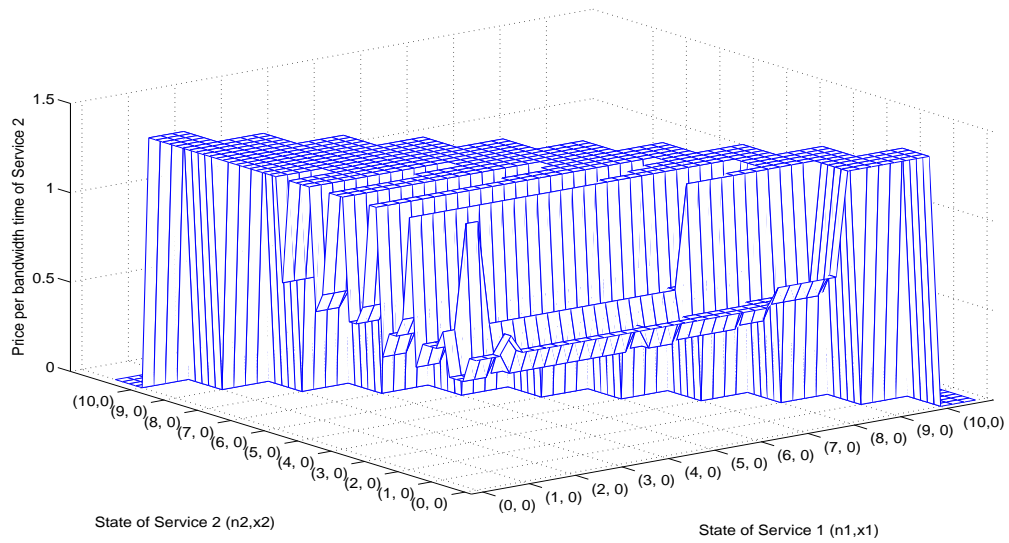
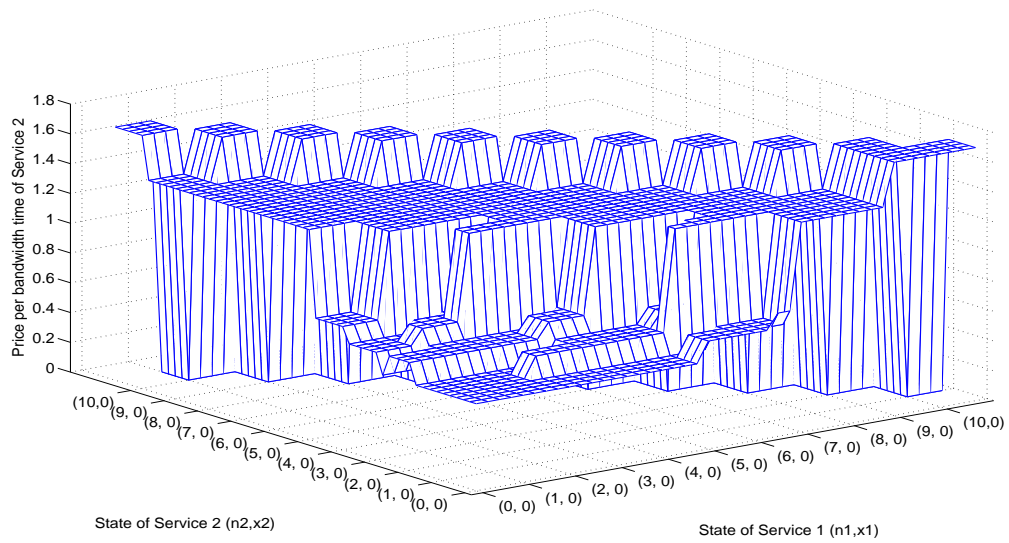


Figure 3.7. Optimal differential reward rate, $h^*(s)$.

3.5 Numerical Results



(a) Service 1



(b) Service 2

Figure 3.8. Optimal prices per bandwidth time for OCADP with $B = 20$, $\psi = (1, 1)$ per bandwidth time and $\mathbf{b} = (2, 2)$. The x and y axes indicate the state of service 1, i.e. $s_1 = (n_1, x_1)$, and 2, i.e. $s_2 = (n_2, x_2)$ respectively. The combination of s_1 and s_2 forms a state.

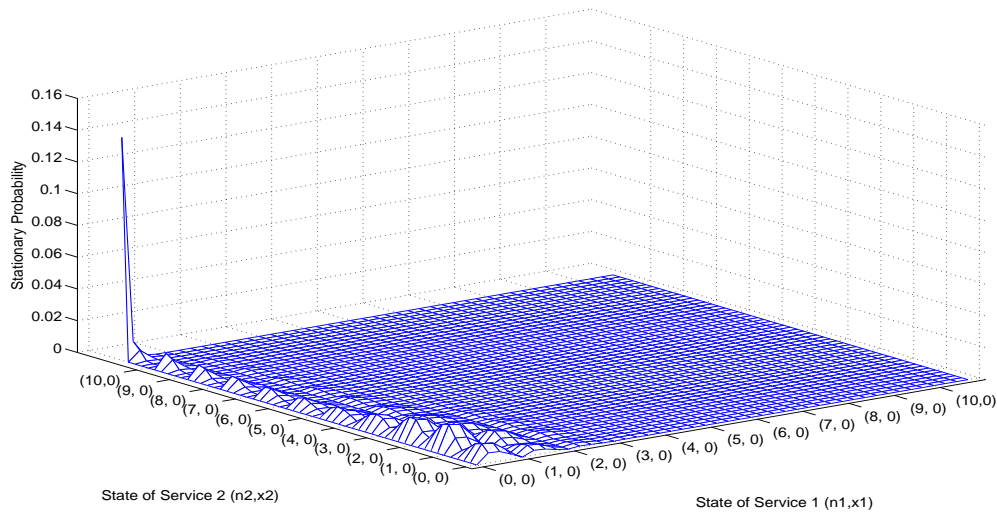
As the network becomes congested, the optimal pricing policy deters low-WTP arrivals by charging higher price per bandwidth time. The state- and service-dependent dynamic pricing policy is illustrated in Fig. 3.8. Additional revenue is generated by offering lower prices per bandwidth time to users when the network is under-utilised. This result is in line with the laws of *demand and supply* from economics which require that prices rise when demand is large relative to available supply and fall in the contrary case [116]. A static pricing scheme that offers average, state-independent price generates lower expected revenue as it fails to offer the same incentives.

3.5.2 Congestion Control

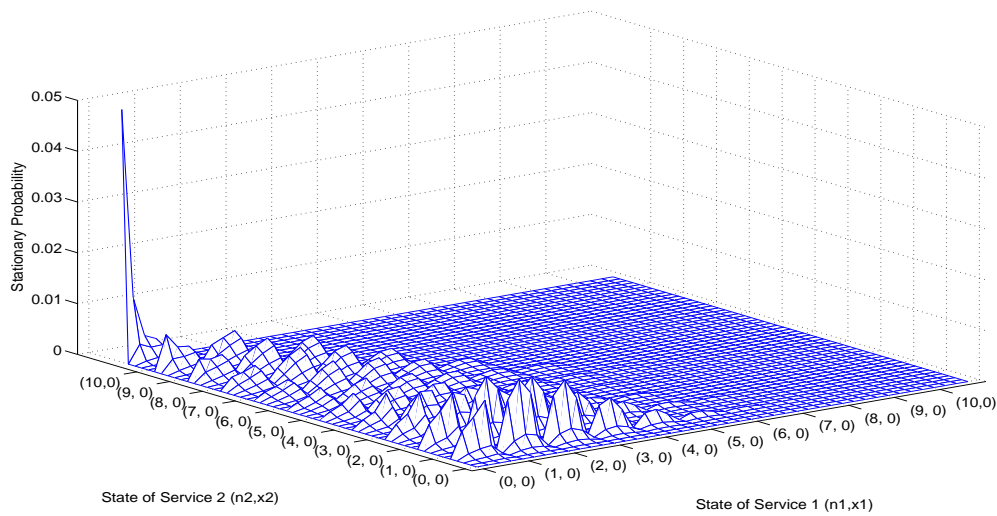
So far in this chapter, we assume that arrival rates are stationary. This leads to a stationary integrated admission and pricing policy. In practice, there are vast differences between peak and off-peak demand patterns and the new call arrival rate λ_j^n typically varies with the time of the day. For example, the peak period for a typical cell in a central business district is between 9 a.m. to 6 p.m. while the busy period in popular night spots occurs well after 6 p.m. For this exercise, we analyse the effectiveness of OCADP in congestion control for a network with $\lambda^n = (1, 2)$ and $\lambda^n = (5, 10)$ to simulate low and heavy traffic. The stationary probabilities of both cases are illustrated in Figs. 3.9 and 3.9 respectively.

Since we design service 2 to generate higher reward than service 1, the steady state probability that the system has more than 4 service 1 users is very low. The spikes in both graphs are due to the variation of the number of users in orbit. In both cases, there is an evident displacement of probabilities from congested states, i.e states with high bandwidth consumption, to less congested states when OCADP is used. In particular, the probabilities of being in congested states are shifted to other states when the arrival rates are increased to $\lambda^n = (5, 10)$. The explanation for this phenomenon is simple. OCADP blocks call requests that generate lower revenue and sets high prices to subsequent call admissions when the network is congested. Persistent users who are blocked or have insufficient WTP will be deferred to less congested periods, thereby easing the burden on the network of needing to always cater for the peak demand.

3.5 Numerical Results



(a) Under AASP with $\lambda = (1, 2)$



(b) Under OCADP with $\lambda = (1, 2)$

Figure 3.9. Stationary probabilities of a system with $B = 20$ and $\mathbf{b} = (2, 2)$. The x and y axes indicate that the state of service 1, $s_1 = (n_1, x_1)$, and 2, $s_2 = (n_2, x_2)$, respectively. The intersection of s_1 and s_2 forms a state $s \in \mathcal{S}$. OCADP controls network congestion by shifting stationary probabilities from states with high number of busy connections to less busy states.

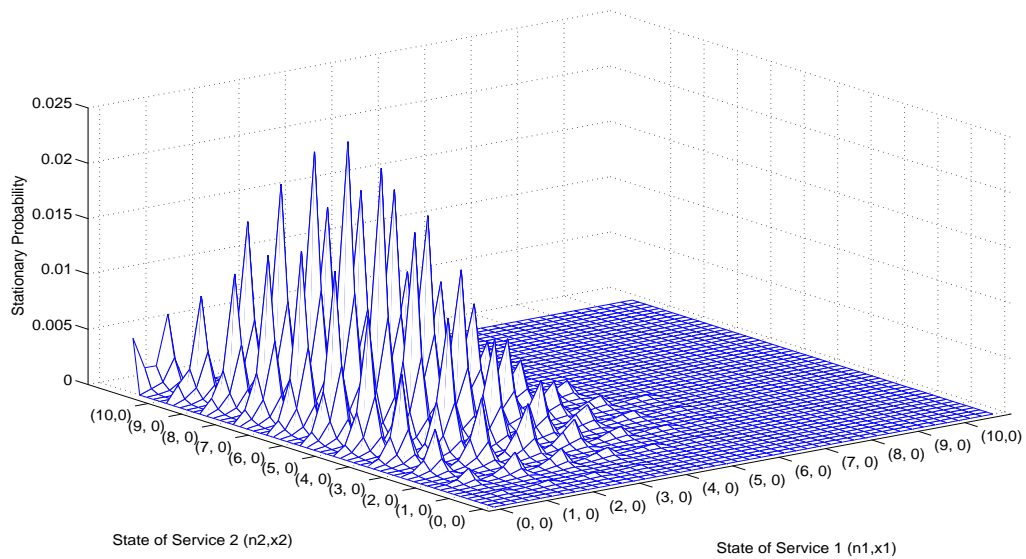
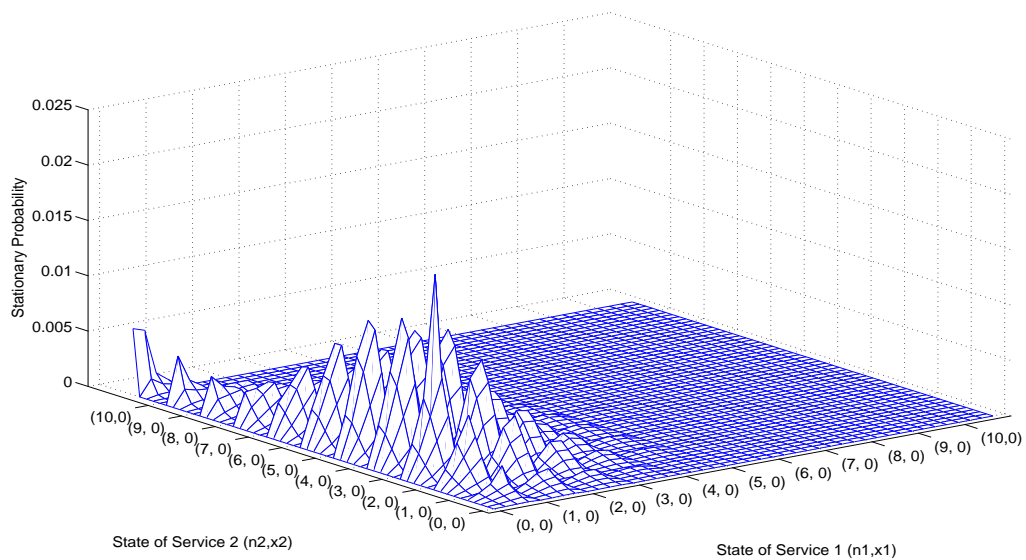
(a) Under AASP with $\lambda = (5, 10)$ (b) Under OCADP with $\lambda = (5, 10)$

Figure 3.10. Stationary probabilities of a system with $B = 20$ and $\mathbf{b} = (2, 2)$. OCADP controls network congestion by shifting stationary probabilities from states with high number of busy connections to less busy states.

3.5.3 Price Discrimination

Up to this point, we have optimised admission and price based on different expected revenue and the same WTP per bandwidth time of each service. The conclusion drawn from the previous section is that the optimal dynamic pricing scheme in AADP and OCADP is discriminatory, i.e. the optimal price per bandwidth time of each service is different in the same state, even though users have the same WTP per bandwidth time. Such price variation to optimise revenue is an example of price discrimination. We will now investigate how setting the same state-dependent price per bandwidth time, i.e. without price discrimination, for all services is a good approximation of the optimal policy. This setting is not to be confused with static pricing, where price is the same in *all* states.

In general, price discrimination refers to the practice of varying the price of a product (in this case, bandwidth) through time or between customers or applications so as to improve revenue. It is widely practised in various forms, including hotels and airlines, where it is known as yield management. All applications rely on the proposition that revenue can be increased by raising the prices of goods with inelastic demand and reducing the price of goods with elastic demand. The idea behind price discrimination is that, if demand responds more than proportionately to price (elastic demand), revenue will rise with declines in price while the opposite is true if demand responds less proportionately with price.

In this work, we do not make estimates of demand elasticity. Instead, the modelling of demand responses to price changes relies on the distribution of WTP among users, i.e. raising the price means that fewer arrivals are admitted and might mean a reduction on demand, depending on the WTP of arriving users. This amounts to saying that those users who are willing to pay more than the current price are impervious to price changes up to that point, implying they have infinitely inelastic demand up to the price which equals their WTP. The impact on revenue therefore depends on the shape and scale of the WTP among arriving users. These factors determine the average elasticity of demand for arriving users as a whole.

We analyse the two price settings, with and without price discrimination, by comparing them over a range of arrival rates for service 2. The optimal rewards per stage J^* for both cases are depicted in Fig. 3.11. It is evident that the non-price-discrimination

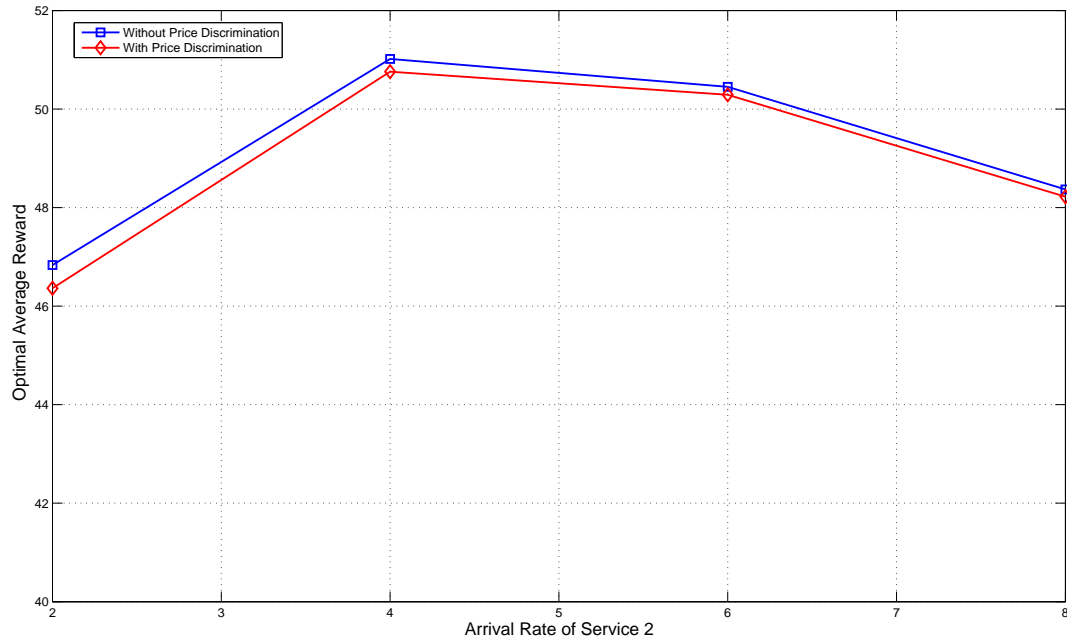


Figure 3.11. Optimal reward per stage J^* under OCADP (with and without price discrimination).

scheme closely approximates the optimal reward per stage J^* as the arrival rate to service 2 increases. Although we have established in the previous section that reward is maximised when price discrimination is practised, there are various reasons for not implementing it. Firstly, the computational cost of the Bellman equation (4.19) is reduced if price discrimination is avoided. This is because the size of the control space is now smaller and the system will only need to compute the single optimal price that maximises revenue in that state. Instead of solving for $2J$ control variables (price and admission each), the non-price-discrimination strategy has only $J + 1$ variables for every state. Secondly, the practical implementation of such scheme is simplified because only one price has to be announced to all network users. The burden of choice on users is also reduced by providing them with a single price that applies to all services.

3.5.4 Exponential WTP Distribution

When the probability distribution function of the WTP reduces to an exponential distribution by setting $\beta = 1.0$ (see Fig. 3.5), we obtained similar results as when $\beta = 3.5$ was used in Fig. 3.5. However, OCASP and AASP provide lower revenue but OCADP and

3.6 Conclusions

AADP achieve higher revenue, compared to when $\beta = 3.5$ is used. These results can be explained using Fig. 3.4. The first result is due to the higher WTP placed by users with $\beta = 1$ for the same access probability $\alpha_P = 0.5$. In the dynamic pricing case, the average optimal access probabilities for OCADP when $\beta = 1$ and 3.5 are $\alpha_P^* = (0.11, 0.25)$ and $(0.13, 0.45)$ respectively. The higher revenue is due to the WTP of users being higher at $\alpha_P = 0.25$ when $\beta = 1$ compared to $\alpha_P = 0.45$ when $\beta = 3.5$.

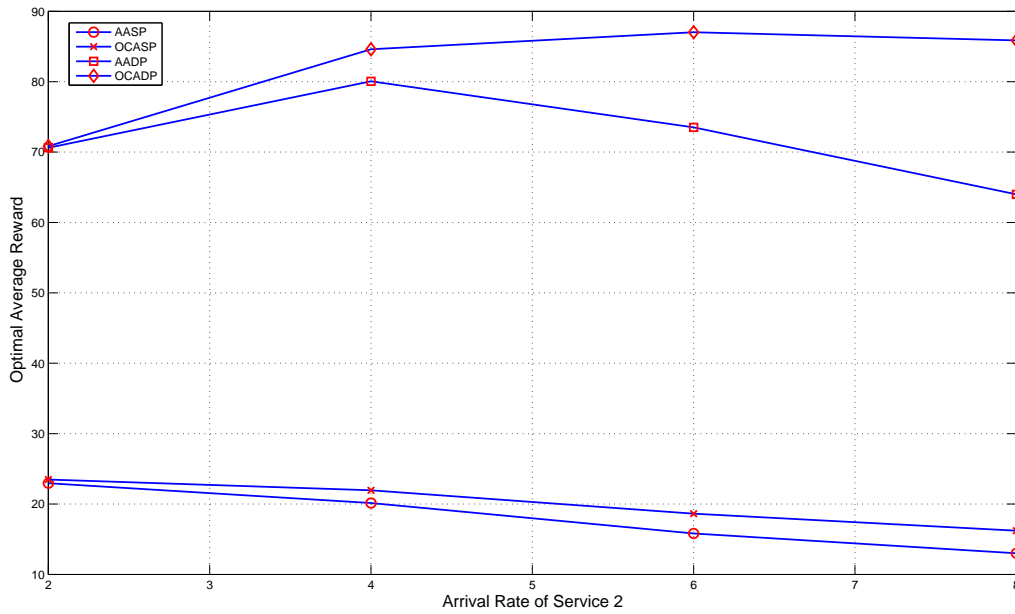


Figure 3.12. Optimal reward per stage J^* for WTP shape $\beta = 1.0$.

3.6 Conclusions

We have introduced and analysed a model for optimal integrated call admission and dynamic pricing of services on a resource-sharing, multiservice network. We have also shown how matrix analytic methods can be applied to solve for the stationary distribution and system characteristics of the model. We summarise our main results as follows:

- Integrated policy OCADP outperforms other conventional policies which consider call admission and price as separate problems. It has the flexibility to reject a new connection request when this is advantageous to the network even though there is

sufficient bandwidth to accommodate the call. This strategy provides monetary incentives to low-WTP users to access the network when the load is relatively light and allocating resources to high-WTP users when the network is relatively congested.

- OCADP is effective in congestion control because it displaces traffic from congested to less congested states. The blocking probabilities of the services decrease because occasions of high price-per-bandwidth-time shift and even out the load, resulting in lower stationary probabilities of the network being in congested states.
- The revenue-maximising optimal pricing policy is discriminatory even when users of all services have the same WTP per bandwidth time. However, a non-discriminatory pricing scheme can closely approximate the optimal results. This scheme has the benefit of reduced computation effort due to its smaller price control space.

Chapter 4

Interference-based Dynamic Pricing and RRM

4.1 Introduction

In this chapter, we extend our proposal from Chapter 3, designed for fixed-capacity, cellular networks, to soft-capacity CDMA-based systems such as WCDMA. The effective capacity of such systems is not determined by the available resources as in TDMA. Each user experiences interference from users outside its cell, in addition to the ones within the same cell. The ability to adapt to such interference, called graceful degradation, is one of the most essential features of CDMA systems. Good interference handling via radio resource management (RRM) plays an important role in increasing system capacity and providing Quality of Service (QoS) guarantee. Dynamic pricing can assist the RRM; it can act as an arbitration mechanism for efficiently allocating network resources by influencing the way users utilise scarce network resources. Dynamic pricing can also enhance operators' ability to recover costs and make profits to finance capacity expansions.

With those motivations in mind, we will study the problem of optimally integrating dynamic pricing and RRM, in terms of CAC and resource allocation, in a multiservice WCDMA network. A major difference from much of the literature on pricing outlined in Sections 2.2.2 and 2.2.3 in Chapter 2 is that we consider charging users based on their *Noise Rise Factor*, i.e. a new parameter that measures one's contribution to the total interference. We call pricing strategies of this nature *interference-based* pricing. We

4.1 Introduction

propose that the rate at which price rises as the network reaches its interference limit should depend on the nonlinear relationship between one's transmission rate and the interference imposed on others. Conventional bandwidth- or load-based pricing schemes fail to capture such relationship. This strategy is non-discriminatory in the sense that it charges the same price per unit interference to all users, regardless of their service. Unlike [43,98,126], which have been developed to deal with static scenarios and optimised for a fixed number of users on the network, our work will also use pricing to influence the rate of stochastic incoming traffic. By considering dynamic user arrivals, handoffs and departures, the objective is to develop an optimal policy that maximises the long-term, expected reward. Handoff call dropping is minimised via CAC.

The approach used here exploits another characteristic of WCDMA services, *viz* that it can operate within a range of transmission rates. For example, the UMTS Adaptive Multi-Rates (AMR) voice codec offers transmission rates that vary between 4.75 and 12.2 kbit/s for conversational voice service [49]. As the level of interference increases, the network recalculates the optimal transmission rate of all services such that existing connections can be maintained with the addition of new users. A service can be further classified according to the range of acceptable transmission rates, which reflects users' perception of QoS. Unlike typical congestion-dependent pricing schemes that increase price as the number of users increase, our model allows the possibility of maintaining the price if existing users are tolerant towards the degradation of QoS during their call and the long-term, expected reward of the operator is still maximised.

This problem is naturally formulated as a Dynamic Programming (DP) problem, but the evaluation function is too complex for an exact solution. Offline DP methods are of limited utility for problems with large state spaces because they require full expansion of all possible states and storing the reward for each state. This often leads to space complexity exponential in the number of state variables, the situation infamously known as the "curse of dimensionality". We will use Neuro-Dynamic Programming (NDP) [13], a simulation-based learning method, to solve the problem. This method has been successfully applied in a CAC problem [69] and a retailer inventory management problem in [121].

The rest of the chapter is organised as follows. In Section 4.2, we describe our network model. We then formulate our problem as DP and NDP problems in Section

4.3 and Section 4.4 respectively. Experimental results are presented in Section 4.5 and conclusions are summarised in Section 4.6. This work has been partially presented in [44].

4.2 System Model

We consider the uplink of a multi-service WCDMA system with J classes of service. New and handoff calls of class j arrive at the cell according to Poisson process with rates λ_j^n and λ_j^h respectively. The call holding time and cell residence time of a class j call are both exponentially distributed with mean $1/\mu_j$ and $1/\gamma_j$ respectively. During a connection, it is assumed that a call alternates between ON and OFF states at rate α_j and β_j . We denote the probability that a connection is active as the activity factor $\nu_j = \frac{\beta_j}{\alpha_j + \beta_j}$. Although the system cannot distinguish between idle and active periods and is unable to use the idle periods to transmit other calls, idle periods do not contribute any interference and network users will benefit from the interference reduction [128]. Users arrive with a mean budget or willingness to pay (WTP) of Ψ_j that quantifies the satisfaction gained from a call.

The system state can be represented by a column vector, $\mathbf{n} = (n_1, \dots, n_J)^T$, where n_j is the number of admitted users of service j . The state space of the system depends on the interference generated by users within and outside of the cell. In WCDMA, services can operate within a range of transmission rates. The controller jointly controls the resource allocation (i.e. transmission rate), admission price and call admission of the system. The resource allocation policy determines the state-dependent optimal transmission rate vector:

$$\mathbf{u}_r(\mathbf{n}) = (u_{r_1}, \dots, u_{r_J}), \quad u_{r_j} \in \mathcal{R}_j. \quad (4.1)$$

The transmission rate of service j users belongs to a finite set $\mathcal{R}_j = \{R_{j1}, \dots, R_{jM_j}\}$, where M_j is the number of discrete transmission rates supported by the system. The discretisation of transmission rates is due to the allocation of Orthogonal Variable Spreading Factor (OVSF) codes. The OVSF code is of length between 1024 to 4 chips, allowing transmission rates between 15 kbit/s and 1.92 Mbit/s [49]. Users' data rate is between 3 kbit/s and 768 kbit/s after data correction.

In order for a signal to be received, the ratio of its received power to the sum of the background noise and interference must be greater than a given target. When there

4.2 System Model

are \mathbf{n} users transmitting simultaneously in a given cell, the target quality is translated to the following inequality that must be satisfied for each user $i = 1, \dots, n_j$ of service $j = 1, \dots, J$ [49] [42]:

$$\frac{W}{\nu_j u_{rj}(\mathbf{n})} \times \frac{P_j}{\sigma^2 + I_{\text{own}} + I_{\text{other}} - P_j} \geq \left(\frac{E_b}{N_0} \right)_j, \quad (4.2)$$

where W is the WCDMA chip rate, ν_j is the activity factor, $u_{rj}(\mathbf{n})$ is the allocated transmission rate, P_j is the received signal power from the i th user, σ^2 is the background thermal noise power, I_{other} and I_{own} are the other-cell and own-cell interference and $(E_b/N_0)_j$ is the ratio of energy per bit to noise density required to meet predefined bit error rate (BER). The total received interference at the base station is defined as $I_{\text{total}} = \sigma^2 + I_{\text{own}} + I_{\text{other}}$. For simplicity, the other-cell interference can be taken into account by some constant f [123], i.e. $I_{\text{other}} = fI_{\text{own}}$.

Definition 4.1. *Dynamic pricing determines the state-dependent admission price policy,*

$$\mathbf{u}_p(\mathbf{n}) = u_p, \quad u_p \in \mathcal{U}_p, \quad (4.3)$$

where \mathcal{U}_p is the set of possible values of u_p . Price u_p is defined as the price per load-time in load-based pricing and price per interference-time in interference-based pricing.

4.2.1 Load-based Pricing

In load-based pricing, users are charged according to their Individual Load Factor (ILF) η_j , which is the ratio of their individual load with respect to the system loading. Assuming that the transmit power of each mobile station (MS) is perfectly controlled based on the receiving level at the base station (BS), the minimum power that the i th user of service j must transmit in order to achieve (5.1) is given by $P_j = \eta_j(u_{rj}(\mathbf{n}))I_{\text{total}}$, where $\eta_j(u_{rj}(\mathbf{n}))$ is defined as:

$$\eta_j(u_{rj}(\mathbf{n})) = (1 + f) \left(1 + \frac{W}{\left(\frac{E_b}{N_0} \right)_j u_{rj}(\mathbf{n}) \nu_j} \right)^{-1}. \quad (4.4)$$

The system load factor is defined as the sum of all individual load factors. With \mathbf{n} users, the system load factor is given by:

$$\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = \sum_{j=1}^J \eta_j(u_{rj}(\mathbf{n})) n_j. \quad (4.5)$$

When a user requests a call connection of service j with price per ILF-time u_p , they will decide to either make a connection request if their budget is sufficient to cover the expected call cost of length $1/\mu_j$ or defer the request otherwise. We denote the probability of having the sufficient WTP as the *access probability*:

$$\alpha_{p_j}(\mathbf{n}, \mathbf{u}) = \Pr\left(\Psi_j \geq \frac{u_p \eta_j(u_{rj}(\mathbf{n}))}{\mu_j}\right). \quad (4.6)$$

Access probability can be seen as an *arrival gate* that controls the flow of price-affected arrivals to the system. λ_j^n is the maximum *new* arrival rate, limited only by α_{p_j} . Since handoff calls are pre-admitted at another price, their arrival rate will be independent of the current admission price and should never be dropped on the basis of insufficient budget.

4.2.2 Interference-based Pricing

In interference-based pricing, users are charged according to the interference generated by their call. The total interference on the uplink can be estimated using the system load factor defined in (4.5). The system noise rise can be expressed as

$$\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = 10 \log_{10} \left(\frac{I_{\text{total}}}{P_N} \right) = -10 \log_{10}(1 - \eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))). \quad (4.7)$$

using $I_{\text{total}} = \sigma^2 + I_{\text{own}} + I_{\text{other}} = \sigma^2 + \sum_{j=1}^J P_j = \sigma^2 + \eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) I_{\text{total}}$. When the system is empty, the system load factor and noise rise are $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = 0$ and $\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = 0$ dB respectively. The system noise rise, defined as the ratio of the total received wideband power to the background thermal noise, is a metric for measuring the total interference in the cell. From (4.7), the system noise rise increases logarithmically with the system load factor, which depends on the individual load factor of all users. The relationship between the system noise rise and system load factor is illustrated in Fig. 4.1. We now propose a metric to measure the amount of interference generated by a call.

Definition 4.2. *The Noise Rise Factor (NRF), $\varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$, of a call with load factor $\eta_j(u_{rj}(\mathbf{n}))$ is defined as*

$$\varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = \frac{\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))}{\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))} \eta_j(u_{rj}(\mathbf{n})), \quad (4.8)$$

where $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ and $\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ are the system load factor and noise rise respectively. Note that the first component gives the noise rise per load factor. Multiplying this component with the load factor of a call gives its individual noise rise.

4.2 System Model

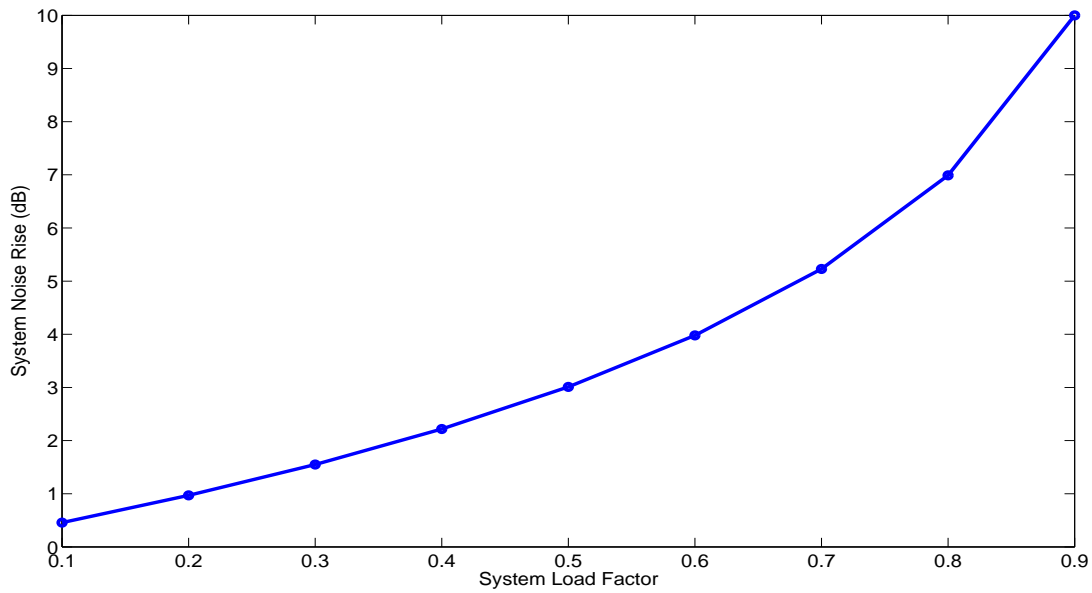


Figure 4.1. The interference level in the network, indicated by the system noise rise, rises exponentially as the system load factor increases.

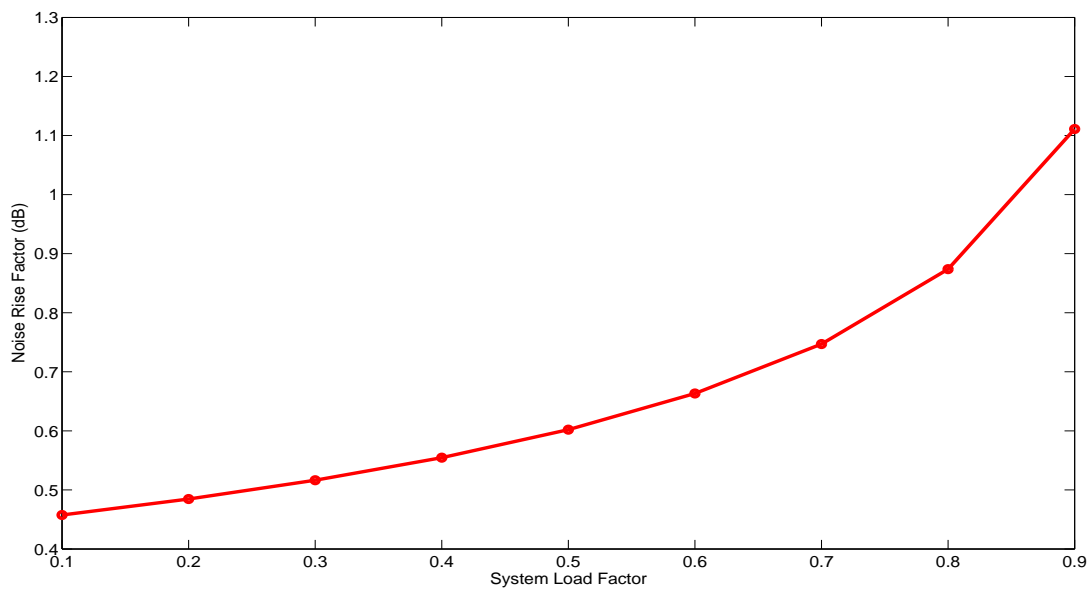


Figure 4.2. The noise rise factor, i.e. the interference generated by a call, increases exponentially as the system load factor increases although the individual load factor remains the same throughout.

For example, assume that $J = 1$ and all users have the same ILF, i.e. $\eta = 0.1$. When there is one user, say A, in the system with $\eta^A = 0.1$, the equivalent system load factor, system noise rise and system noise rise are $\eta_{\text{sys}} = 0.1$ and $\varphi_{\text{sys}} = 0.46$ dB respectively. The NRF of user A is also the same as the system noise rise, i.e. $\varphi^A = \frac{0.46(0.1)}{0.1} = 0.46$ dB. Now consider the case where there are two additional users, say B and C, in the system. User B and C also have the same ILF as A, i.e. $\eta^A = \eta^B = \eta^C = 0.1$. The addition of two new users results in an increase in the system load factor and system noise rise, each of which becomes $\eta_{\text{sys}} = 0.3$ and $\varphi_{\text{sys}} = 1.55$ dB. Given that all users have the same ILF, their NRFs are also the same. However, the NRF has now increased to $\varphi^A = \varphi^B = \varphi^C = \frac{1.55(0.1)}{0.3} = 0.52$ dB. Therefore, the amount of interference generated per load factor increases at a nonlinear rate as the number of users increases. Note that the NRF defined in (4.2) is in terms of the system noise rise in dB because of the additive nature of the noise rise. The system noise rise $\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$, also in dB, is therefore defined as the total NRF of all users:

$$\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) = \sum_{j=1}^J n_j \varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) \text{ dB}. \quad (4.9)$$

The nonlinear relationship between the NRF and the system load factor in the previous example is illustrated using Fig. 4.2 and will be derived in the following proposition.

Proposition 4.1. *There exists a nonlinear relationship between the system load factor $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ and noise rise factor $\varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$. As $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ increases, the individual noise rise factor will increase nonlinearly by*

$$\frac{10\eta_j(u_{rj}(\mathbf{n}))}{\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))} \left[\frac{1}{(\log 10)(1 - \eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})))} + \frac{\log_{10}(1 - \eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n})))}{\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))} \right], \quad (4.10)$$

where $\eta_j(u_{rj}(\mathbf{n}))$ is the individual load factor of a call of class j .

Proof. Let φ_j , η_j and η_{sys} represent $\varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$, $\eta_j(u_{rj}(\mathbf{n}))$ and $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ respectively. Equation (4.2) can be rewritten using (4.7) as

$$\varphi_j = \frac{\eta_j}{\eta_{\text{sys}}} \times 10 \log_{10} \left(\frac{1}{1 - \eta_{\text{sys}}} \right). \quad (4.11)$$

Taking the derivative of this equation with respect to η_{sys} , we have the following:

$$\begin{aligned} \frac{d\varphi_j}{d\eta_{\text{sys}}} &= 10\eta_j \left[\frac{1}{\eta_{\text{sys}}} \frac{d}{d\eta_{\text{sys}}} \log_{10} \left(\frac{1}{1 - \eta_{\text{sys}}} \right) + \log_{10} \left(\frac{1}{1 - \eta_{\text{sys}}} \right) \frac{d}{d\eta_{\text{sys}}} \frac{1}{\eta_{\text{sys}}} \right] \\ &= 10\eta_j \left[\frac{1}{\eta_{\text{sys}}} \frac{1}{\log 10} \left(\frac{1}{1 - \eta_{\text{sys}}} \right) - \frac{1}{\eta_{\text{sys}}^2} \log_{10} \left(\frac{1}{1 - \eta_{\text{sys}}} \right) \right] \end{aligned}$$

4.3 Dynamic Programming Formulation

We can obtain the relationship in (4.10) by rearranging this equation. \square

When a user requests a call connection of service j with price per NRF-time u_p , they will decide to either make a connection request if their budget is sufficient to cover the expected call cost of length $1/\mu_j$. or defer the request otherwise. The access probability of users is given by:

$$\alpha_{p_j}(\mathbf{n}, \mathbf{u}) = \Pr \left(\Psi_j \geq \frac{u_p \varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))}{\mu_j} \right). \quad (4.12)$$

CAC is triggered whenever new or handoff users request a call connection.

Definition 4.3. A CAC policy determines the state-dependent call admission vector:

$$\mathbf{u}_c(\mathbf{n}) = (\mathbf{u}_c^n(\mathbf{n}), \mathbf{u}_c^h(\mathbf{n})), \quad u_{c_j}^n, u_{c_j}^h \in [0, 1], \quad (4.13)$$

where $\mathbf{u}_c^n(\mathbf{n}) = (u_{c_1}^n, \dots, u_{c_J}^n)$ and $\mathbf{u}_c^h(\mathbf{n}) = (u_{c_1}^h, \dots, u_{c_J}^h)$ represent the admission probability of a new and handoff call request when \mathbf{n} users already exist respectively.

When $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ approaches 1, the system reaches its capacity threshold, η_{max} and $\varphi_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ approaches infinity. The admission condition of a new user of service j is:

$$\eta_{\text{sys}}(\mathbf{n} + \mathbf{e}_j, \mathbf{u}_r(\mathbf{n} + \mathbf{e}_j)) \leq \eta_{\text{max}} < 1. \quad (4.14)$$

Note that there is an important distinction between our definition of (4.14) and in [49]. Since the optimal transmission rate of service $k \neq j$, $\mathbf{u}_{r_k}(\mathbf{n})$, might be adjusted by the system to accommodate a new user of service j , the new system load factor, $\eta_{\text{sys}}(\mathbf{n} + \mathbf{e}_j, \mathbf{u}_r(\mathbf{n} + \mathbf{e}_j))$ does not necessarily equal the sum of $\eta_{\text{sys}}(\mathbf{n}, \mathbf{u}_r(\mathbf{n}))$ and $\eta_j(\mathbf{u}_r(\mathbf{n} + \mathbf{e}_j))$. A new call will only be admitted if the interference threshold constraint (4.14) is still satisfied and the user has sufficient WTP to cover the cost of the call.

4.3 Dynamic Programming Formulation

Any state transition is caused by one of the following events: an arrival of a new call; an arrival of a handoff call; and departure or handoff of an ongoing call. Since we do not keep track of the number of users in other cells, departure and handoff of ongoing calls can be treated as the same event. Let Ω denote the set of possible events, $\Omega = \{\omega \mid \omega \in \{\mathbf{0}, \omega_j^n, \omega_j^h, \omega_j^d\}, j \in [1, J]\}$, where $\mathbf{0}$, ω_j^n , ω_j^h and ω_j^d represent no state transition,

a new or handoff call arrival and a departure respectively. Pricing, resource allocation and CAC are triggered when there is a new or handoff call request. Let $\mathcal{U}(\mathbf{n}, \omega)$ be the set of available actions in state \mathbf{n} when event ω occurs:

$$U(\mathbf{n}, \omega) = \{\mathcal{U}_c \times \mathcal{U}_r \times \mathcal{U}_p\} \quad \text{if } \omega \in \omega_j^n, \omega_j^h, \quad (4.15)$$

where \mathcal{U}_c , \mathcal{U}_r and \mathcal{U}_p are the set of possible call admission, resource allocation and dynamic pricing actions.

Using uniformisation [12, 92], the continuous-time Markov Decision Problem (MDP) can be transformed into its discrete-time equivalence with the so-called uniform transition rate, where the total transition rate out of any state is bounded by τ . The transition probabilities are then given by:

$$p(\mathbf{n}, \omega, \mathbf{u}) = \begin{cases} u_{c_j}^n \alpha_{p_j}(\mathbf{n}, \mathbf{u}) \lambda_j^n / \tau & \text{if } \omega = \omega_j^n \\ u_{c_j}^h \lambda_j^h / \tau & \text{if } \omega = \omega_j^h \\ n_j (\mu_j + \nu_j) / \tau & \text{if } \omega = \omega_j^d \\ 1 - \tau(\mathbf{n}) / \tau & \text{otherwise,} \end{cases} \quad (4.16)$$

where $\tau(\mathbf{n}) = \sum_j u_{c_j}^n (\alpha_{p_j}(\mathbf{n}, \mathbf{u}) \lambda_j^n + \lambda_j^h) + n_j (\mu_j + \nu_j) \leq \tau$ is the total transition rate out of a state \mathbf{n} . Given that the system is in state $s = \mathbf{n}$ with an event $\omega \in \Omega$ and control actions $u \in \mathcal{U}(\mathbf{n}, \omega)$ available, the next state, s' , is given by a function y such that:

$$s' = y(\mathbf{n}, \omega, \mathbf{u}) = \begin{cases} \mathbf{n} + \mathbf{e}_j & \text{if } \omega = \omega_j^n, u_{c_j}^n = 1 \\ \mathbf{n} + \mathbf{e}_j & \text{if } \omega = \omega_j^h, u_{c_j}^h = 1 \\ \mathbf{n} - \mathbf{e}_j & \text{if } \omega = \omega_j^d, \\ \mathbf{n} & \text{otherwise.} \end{cases} \quad (4.17)$$

As previously introduced in Chapter 3, we will again use the term Satisfaction Revenue (SR) to denote the monetary measure of users' satisfaction with the continuation of a call when a handoff is successful. The immediate revenue collected by the system when a user is admitted is:

$$g(\mathbf{n}, \omega, \mathbf{u}) = \begin{cases} u_{c_j}^n u_p \frac{1}{\mu_j} \eta_j(u_{r_j}(\mathbf{n})) & \text{if } \omega = \omega_j^n \text{ for ILF-based pricing} \\ u_{c_j}^n u_p \frac{1}{\mu_j} \varphi_j(\mathbf{n}, \mathbf{u}_r(\mathbf{n})) & \text{if } \omega = \omega_j^n \text{ for NRF-based pricing} \\ u_{c_j}^h SR & \text{if } \omega = \omega_j^h. \end{cases} \quad (4.18)$$

4.4 Neuro-Dynamic Programming Formulation

In order to reflect the higher importance of accepting a handoff call, SR should be greater than the actual revenue collected when a new call request is accepted. The average reward-to-go function, known as the Bellman equation, is given by:

$$J^* + h(s) = \max_{\mathbf{u} \in \mathcal{U}(s, \omega)} \left[\sum_{\omega \in \Omega} p(s, \omega, \mathbf{u}) [g(s, \omega, \mathbf{u}) + h(y(s, \omega, \mathbf{u}))] \right]. \quad (4.19)$$

J^* and $h(s)$ denote the optimal average reward and the differential reward rate of state s respectively. A stage here means a transition in the uniformised chain. The optimal expected reward per stage is independent of the initial state. Standard average-reward DP theory applies and there exists a stationary policy which is optimal [12].

4.4 Neuro-Dynamic Programming Formulation

NDP refers to approximate methods that centre around the evaluation and approximation of the optimal cost-to-go function (4.19), possibly through simulation and/or the use of neural networks. In the artificial intelligence community, where the methods originated, they are also known as *reinforcement learning* [115]. Instead of computing the differential reward function $h(s)$ for every state $s \in \mathcal{S}$, NDP uses a compact representation $\tilde{h}(\cdot, \boldsymbol{\theta})$ to approximate $h^*(\cdot)$, using parameter vector $\boldsymbol{\theta}$. Naturally, we want to define the general structure of $\tilde{h}(\cdot, \boldsymbol{\theta})$ and calculate parameter vector $\boldsymbol{\theta}$ so as to minimise the error between the functions $h^*(\cdot)$ and $\tilde{h}(\cdot, \boldsymbol{\theta})$. The process of tuning parameters $\boldsymbol{\theta}$ is often referred as training or learning. The average reward per time J^* is approximated by tunable scalar \tilde{J} . If $\tilde{h}(\cdot, \boldsymbol{\theta})$ and \tilde{J} are close enough to the $h^*(s)$ and J^* , then the greedy control policy induced is, in some sense, close to an optimal policy. Hereafter, we denote the k^{th} step estimate of $\tilde{h}(\cdot, \boldsymbol{\theta})$ and \tilde{J} as $\tilde{h}(\cdot, \boldsymbol{\theta}_k)$ and \tilde{J}_k respectively.

There are two major parts in NDP: an approximation architecture to define the structure of $\tilde{h}(\cdot, \boldsymbol{\theta}_k)$ and a learning method for tuning $\tilde{h}(\cdot, \boldsymbol{\theta}_k)$ and \tilde{J} . As illustrated in Fig. 4.3(a), a general, feature-based approximation architecture involves:

- *feature extraction*, i.e. the design of a feature vector $f(\cdot)$ that enhances the approximation, and
- *function approximation*, i.e. the definition of a function to approximate $\tilde{h}(f(\cdot), \boldsymbol{\theta}_k)$.

There are several simulation-based algorithms that can be used to tune the vector θ_k and scalar \tilde{J}_k . Like [69], we will use the TD(0) algorithm for average reward problems. This algorithm preserves the same convergence properties and error guarantees as its discounted version. The TD(0) algorithm belongs to the class of Temporal Difference learning algorithms, often known as the TD(λ). In the next subsections, we will propose a modified version of the general feature-based approximation architecture and TD(0) learning algorithm that includes the decision/action \mathbf{u}_k in the feature vector $f(\cdot)$. This modified version is illustrated in Fig. 4.3(b). Finally, we discuss the tradeoffs between exploration and exploitation in the learning algorithm.

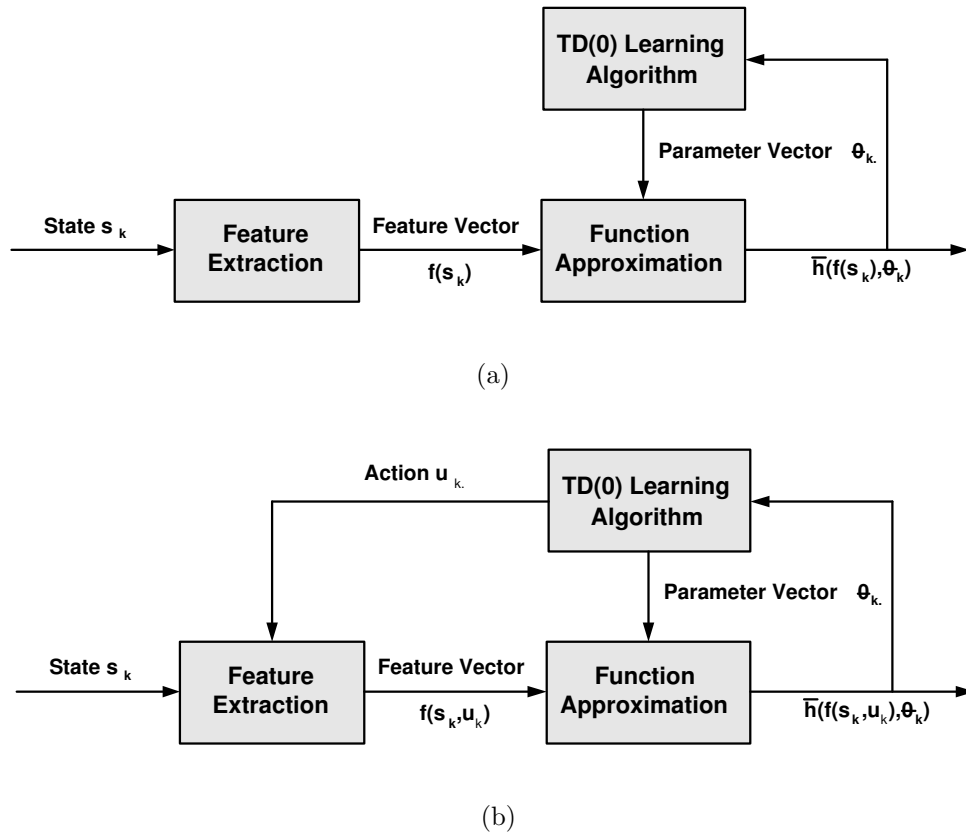


Figure 4.3. (a) A general feature-based approximation architecture. (b) An approximation architecture with feature vector θ_k and future reward rate.

4.4.1 Approximation Architecture

Feature Extraction: Feature vectors summarise what considered to be important characteristics of a state. It is often the case that the complexity of function $\tilde{h}(s_k, \theta_k)$ can be

4.4 Neuro-Dynamic Programming Formulation

reduced by feeding a set of *features* of the state into an approximation architecture. A set of features can be defined as a mapping of $f_l : \mathcal{S} \rightarrow R$. Given a collection of

$$f(s_k) = (f_1(s_k), \dots, f_L(s_k)), \quad (4.20)$$

we approximate for $\tilde{h}(f(s_k), \boldsymbol{\theta}_k)$ instead of $\tilde{h}(s_k, \boldsymbol{\theta}_k)$. The task of selecting the appropriate set of features is usually problem-dependent. Feature vectors are effective in simplifying the approximation of the optimal cost-to-go function when they can capture the nonlinearities that present in the function.

We are interested in an optimal, decision-making policy that maximises the long-term, expected reward. Based on various experiments, we discovered that the approximation architecture is most effective when the “effects” of action vector \mathbf{u}_k are captured in the feature vector $f(\cdot)$. Therefore, we redefine $f(s_k)$ in (4.20) as:

$$f(s_k, \mathbf{u}_k) = (f_1(s_k, \mathbf{u}_k), \dots, f_L(s_k, \mathbf{u}_k)). \quad (4.21)$$

From (4.19), each decision not only results in some immediate reward but also affects the reward obtained in future stages. In our case, the “effects” of \mathbf{u}_k are best described by the future reward rates, which are defined as the product of the future arrival rate due to the integrated pricing, resource allocation and CAC action and the expected call cost.

There are two sources of reward in this problem, i.e. the reward obtained from new and handoff users. The first feature $f_1(s_k, \mathbf{u}_k)$ is usually set as a scalar offset $f_1(s_k, \mathbf{u}_k) = 1$. The next J features are the future reward rates due to new users and are defined as:

$$f_l(s_k, \mathbf{u}_k) = \begin{cases} u_{c_j}^n \alpha_{p_j}(s_k, \mathbf{u}_k) \lambda_j^n u_{p_j} \eta_j(\mathbf{u}_{rk}) & \text{OR} \\ u_{c_j}^n \alpha_{p_j}(s_k, \mathbf{u}_k) \lambda_j^n u_{p_j} \varphi(s_k, \mathbf{u}_{rk}). \end{cases} \quad (4.22)$$

for load-based and interference-based pricing respectively. In the equations, $l = 2, \dots, J+1$ and $j = l - 1$. The last feature is the sum of future reward rate due to handoff users:

$$f_{L=J+2}(s_k, \mathbf{u}_k) = \sum_{j=1}^J u_{c_j}^h \lambda_j^h SR. \quad (4.23)$$

Function Approximation: Choosing an appropriate function approximator for $\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k)$ is the second stage in designing an approximation architecture. A non-linear function approximation like a multilayer perceptron (neural network) often has the

disadvantage of time-consuming tuning of $\boldsymbol{\theta}_k$. A linear approximation architecture is generally more reliable because of the linear dependence of: $\tilde{h}(f(\cdot, \mathbf{u}_k), \boldsymbol{\theta}_k)$ on $\boldsymbol{\theta}_k$. It has the form of

$$\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k) = \boldsymbol{\theta}_k^T f(s_k, \mathbf{u}_k). \quad (4.24)$$

The learning process becomes a linear regression problem. The dimension of the parameter vector $\boldsymbol{\theta}$ is equal to the number of features, L . We will only consider linear approximation function in this work. The derivative of $\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k)$ with respect to $\boldsymbol{\theta}_k$ is:

$$\nabla_{\boldsymbol{\theta}}(\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k)) = f(s_k, \mathbf{u}_k). \quad (4.25)$$

4.4.2 TD(0) Learning Algorithm

Temporal difference learning was originally proposed by Sutton [114] and has been adapted to average reward problems by Tsitsiklis and Van Roy in [120] to approximate differential reward function $h(\cdot, \boldsymbol{\theta})$ and average reward J . Parameter vector $\boldsymbol{\theta}$ is iteratively adapted based on information from simulation or observation of a Markov process. Updates of $\boldsymbol{\theta}$ and \tilde{J} occur upon each state transition that requires decision making with the objective of improving the approximation as time progresses. Starting with an initial parameter vector $\boldsymbol{\theta}_0$ and scalar J_0 , TD(0) generates a sequence of $\boldsymbol{\theta}_k$ and J_k . The number of simulation steps is set at N . At simulation step $k \leq N$, $\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k)$ is used as an approximation of $h^*(s_k)$ to compute the greedy policy that approximates π^* .

Algorithm 4.1. *Suppose that $\nabla_{\boldsymbol{\theta}}(h(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k))$ exists for every $s_k \in \mathcal{S}$ and $\boldsymbol{\theta}_k \in \mathcal{R}^L$. With initial $\boldsymbol{\theta}_0 \in \mathcal{R}^L$, $J_0 \in \mathcal{R}$ and $s_0 \in \mathcal{S}$, we generate $\boldsymbol{\theta}_k$ and J_k using the following recursive procedure:*

Step 1 *Assume that we are given state s_k and parameter vector $\boldsymbol{\theta}_k$, obtain the event $\omega_{k+1} \in \Omega$ according to the transition rates outlined in the model in Section 4.2.*

Step 2 *Choose action vector $\mathbf{u}_k \in \mathcal{U}(s_k, \omega_{k+1})$ that satisfies:*

$$\mathbf{u}_k = \arg \max_{\mathbf{u}_k \in \mathcal{U}} \left[g(s_k, \omega_{k+1}, \mathbf{u}) + \tilde{h}(f(s'_{k+1}, \mathbf{u}_k), \boldsymbol{\theta}_k) \right], \quad (4.26)$$

using $s'_{k+1} = y(s_k, \omega_{k+1}, \mathbf{u}_k)$. Each potential decision \mathbf{u}_k is evaluated in the process of feature extraction.

4.4 Neuro-Dynamic Programming Formulation

Step 3 Update vector $\boldsymbol{\theta}_k$ and scalar \tilde{J}_k with:

$$\begin{aligned} s_{k+1} &= y(s_k, \omega_{k+1}, \mathbf{u}_k) \\ d_k &= g(s_k, \omega_{k+1}, \mathbf{u}_k) - (t_{k+1} - t_k)\tilde{J}_k + \tilde{h}(f(s_{k+1}, \mathbf{u}_k), \boldsymbol{\theta}_k) - \tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k) \\ \boldsymbol{\theta}_{k+1} &= \boldsymbol{\theta}_k + \gamma_k d_k \nabla_{\boldsymbol{\theta}}(\tilde{h}(f(s_k, \mathbf{u}_k), \boldsymbol{\theta}_k)) \\ \tilde{J}_{k+1} &= \tilde{J}_k + \tau_k (g(s_k, \omega_{k+1}, \mathbf{u}_k) - (t_{k+1} - t_k)\tilde{J}_k) \end{aligned}$$

Step 4 Return to step 1 if $k \leq N$.

Scalar d_k is known as the temporal difference corresponding to the transition from s_k to s_{k+1} . The terms γ_k and τ_k are small step size parameters.

Under a fixed policy and standard diminishing step size conditions, J_k and $\boldsymbol{\theta}_k$ will converge to the average reward J^* and vector $\boldsymbol{\theta}$. The algorithm presented is known as *optimistic* TD(0) because the parameter vector $\boldsymbol{\theta}_k$ is updated according to the greedy action chosen in (4.26) during each step of the simulation. This algorithm has been widely used in practice, albeit its convergence properties have never been studied thoroughly [13] [120].

4.4.3 Exploitation and Exploration

In order to maximise the reward obtained in each step, optimistic TD(0) method relies on taking the greedy action defined by (4.26). In other words, the algorithm is *exploiting* actions that it has tried before and which have proven to be effective in producing rewards. However, to discover such actions, it has to *explore* and try actions that it has not selected before. In other words, the algorithm not only has to *exploit* what it already knows, but also to *explore* in order to make better decisions in the future. This is known as the Exploitation-Exploration Dilemma [115]. The inadequacy of state space exploration means that certain profitable alternatives are never explored and remain undiscovered. In our simulation, a greedy action is chosen with probability $1 - P_\epsilon$ and a random action, $\mathbf{u}_k \in \mathcal{U}$, is used with probability P_ϵ .

4.5 Numerical Results

In this section, we present the results obtained using the modified NDP algorithm defined previously. We will compare the following policies, which charge users based on their Individual Load Factor (ILF) or Noise Rise Factor (NRF):

- S-ILF: Static ILF-based Pricing, Always Accept and Average Transmission Rates,
- S-NRF: Static NRF-based Pricing, Always Accept and Average Transmission Rates,
- O-ILF: Optimal ILF-based Pricing, CAC and Rates, and
- O-NRF: Optimal NRF-based Pricing, CAC and Rates.

The Always Accept component in S-ILF and S-NRF always admits a new user if the system load factor constraint (4.14) is not violated. We simulate a system of three services and the parameters used are listed in Table 4.1. The value of chip rate W , other-to-own cell interference ratio f , activity factor ν and target E_b/N_0 used are the same as those used throughout [49].

The first two services are AMR voice, each with a different range of operating transmission rates. Users' choice on this range reflects their tolerance towards the degradation of service quality during high interference. Service 1 only operates within the upper half of the AMR rates provided in Table 4.1. In other words, users of service 2 are more interference-tolerant when the system is heavily loaded. The third service is a data service. We summarise the transmission rate set \mathcal{R}_j as follows: $\mathcal{R}_1 = \{7.40, 7.95, 10.20, 12.20\}$, $\mathcal{R}_2 = \{4.75, 5.15, 5.90, 6.70\} \cup \mathcal{R}_1$ and $\mathcal{R}_3 = \{16, 32, 64\}$ kbit/s. We will be experimenting with the network working at extreme capacity level, i.e. $\eta_{\max} = 0.98$. In practice, the target system load factor is between 0.50 and 0.75, which results in a system noise rise of between 3dB and 6dB [49]. The maximum load factor that can be supported by a network usually increases towards the pole capacity as the WCDMA technology matures.

To ensure a fair comparison among the policies mentioned, the same set of WTP per unit time is used for all simulations. The mean WTP per unit time of each service is proportional to its average transmission rate. We assume that the WTP of users can be fitted into a Weibull distribution with parameters shape β_j and scale ζ_j using mean WTP Ψ_j . This distribution is versatile and can take up the characteristics of other types

4.5 Numerical Results

Table 4.1. Simulation parameters for S-ILF, S-NRF, O-ILF and O-NRF policies.

Simulation Parameter	Value
WCDMA Chip Rate	$W = 3.84$ Mcps
Other-to-Own Cell Ratio	$f = 0.55$
Activity Factor	$\nu = 0.67$ (voice), 1.00 (data)
Target E_b/N_0	$\frac{E_b}{N_0} = 5.0$ dB (voice), 1.5 dB (data)
Satisfaction Revenue	$SR = 50$ per handoff call
AMR Transmission Rates	12.20, 10.20, 7.95, 7.40, 6.70, 5.90, 5.15, 4.75 kbit/s
Transmission Rates for Data	16, 32, 64 kbit/s
New User Arrival Rate	$\lambda^n = (5, 10, 10)$
Handoff User Arrival Rate	$\lambda^h = (1, 2, 2)$
User Departure Rate	$\mu + \gamma = (5, 5, 3)$
Minimum Transmission Rates	$\mathcal{R}^{\min} = (7.40, 4.75, 16.0)$ kbit/s
Maximum Transmission Rates	$\mathcal{R}^{\max} = (12.2, 12.2, 64.0)$ kbit/s
Willingness to Pay per Time	$\Psi = (0.86, 0.63, 3.98)$
Number of Simulation Steps	$N = 1.5 \times 10^6$

of distributions based on the value of its shape. As discussed in Chapter 3, the Weibull distribution has been previously used in [21, 106, 129] to model traffic characteristics and fading channels within the telecommunications framework. In reality, users' WTP can be obtained in a number of ways. For example, information on users' WTP can be extracted from a network operator's historical data on users' spending patterns. Users can also willingly inform network operator of their WTP. Although it is expected that most users would like to spend as little as possible and only indicate their minimum WTP, higher-end users would place a higher value on a call during high interference when their initial WTP is not sufficient.

Using the Weibull distribution, we can deduce the range of prices per ILF-time or NRF-time using:

$$\alpha_p^{\min} \leq \alpha_{p_j}(s_k, \mathbf{u}_k) \leq \alpha_p^{\max}, \quad (4.27)$$

where α_p^{\min} and α_p^{\max} are the minimum and maximum access probabilities set by the operator. Then, u_p can be calculated from the α_{p_j} obtained from (4.6) and (4.12) using average transmission rates. The minimum price that corresponds to α_p^{\max} should be set such as to recover the cost needed to deliver the service. We set α_p^{\min} and α_p^{\max} to 0.1 and 0.8 respectively and select 35 uniformly distributed prices between them as the price decision space, \mathcal{U}_p . The chosen size, denoted as $\#\mathcal{U}_p$, is based on various experiments that indicate that further increase of the pricing space will not provide significantly better results.

The choice of step sizes τ and γ are crucial to convergence and after some trial and error, they are set to 10^{-3} and 10^{-8} respectively throughout the simulation. For all cases, we use the same random number seed and run the simulation for N steps. Although we do not need to run the TD(0) algorithm for static policies S-ILF and S-NRF, their average reward can be approximated using the update rule for J_k in step 3 of the algorithm. In the following subsections, we present the results for the following: (i) simulations under normal traffic load (ii) simulations under heavy traffic load, i.e. twice the arrival rates

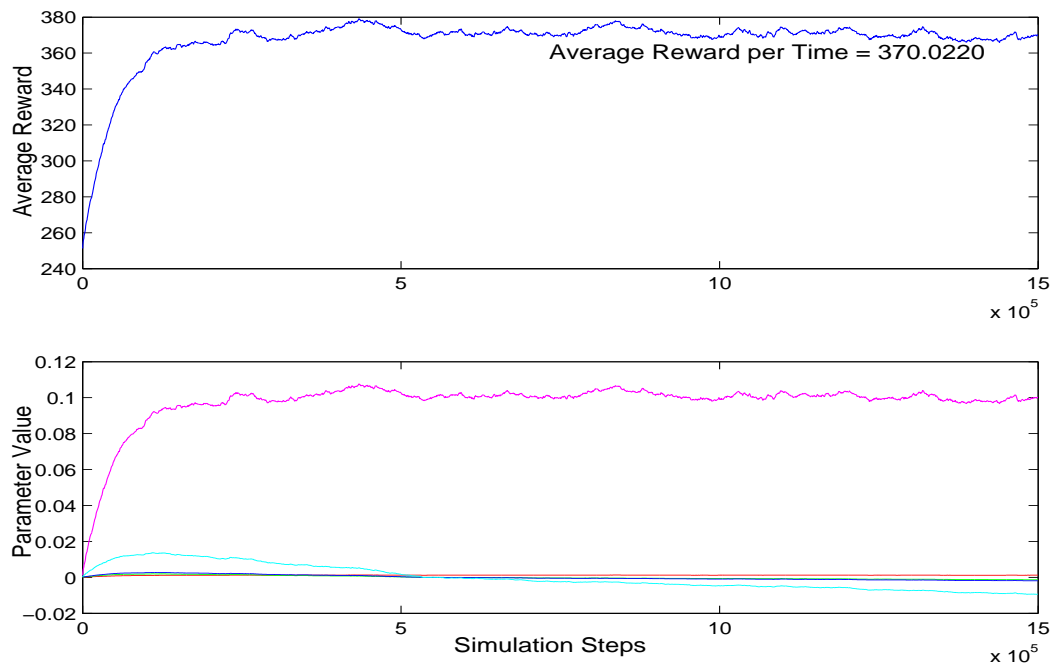


Figure 4.4. Average reward per time, \tilde{J}_k , and parameter vector, $\tilde{\theta}_k$, averaged for every 250 steps, under O-NRF.

4.5 Numerical Results

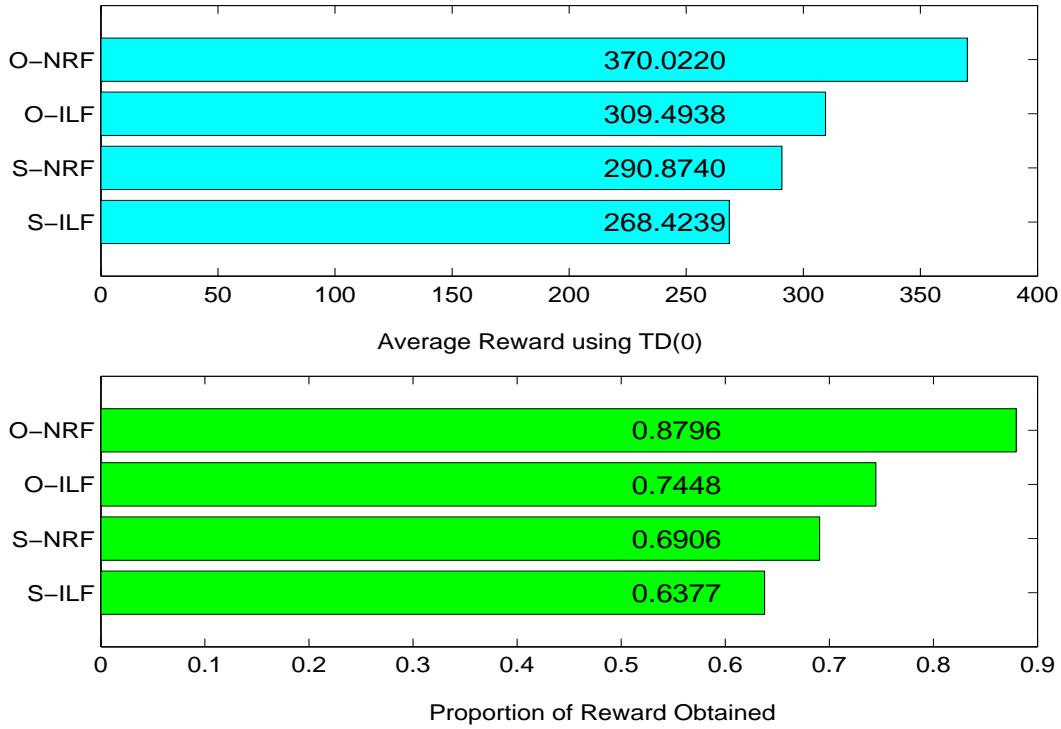


Figure 4.5. Average reward and proportion of reward obtained under normal load.

in Table 4.1, (iii) the effects of active exploration on the results and (iv) the effects of a price sliding window on the average reward.

4.5.1 Normal Traffic Load

The estimation of average reward per unit time \tilde{J}_k and parameter vector θ_k for O-NRF during training are shown in Fig. 4.4. The average reward \tilde{J}_N and proportion of the reward obtained are illustrated in Fig. 4.5. The proportion of reward obtained is the ratio between the average reward and its potential reward if no users are blocked or dropped due to insufficient budget or the violation of constraint (4.14). Optimal policy O-NRF accumulated the highest average reward and proportion of reward obtained, followed by O-ILF, S-NRF and S-ILF. The improvement of O-NRF is about 38% over S-ILF, 27% over S-NRF and 19% over O-ILF. As we have verified using Proposition 4.1, NRF increases exponentially as the system load factor approaches 1. By contrast, as indicated in Fig. 4.6, ILF remains constant regardless of the level of interference in a network. Even when static price per unit NRF is used, the price per unit time will rise with the level of system interference because the noise rise generated by a call has increased. This helps

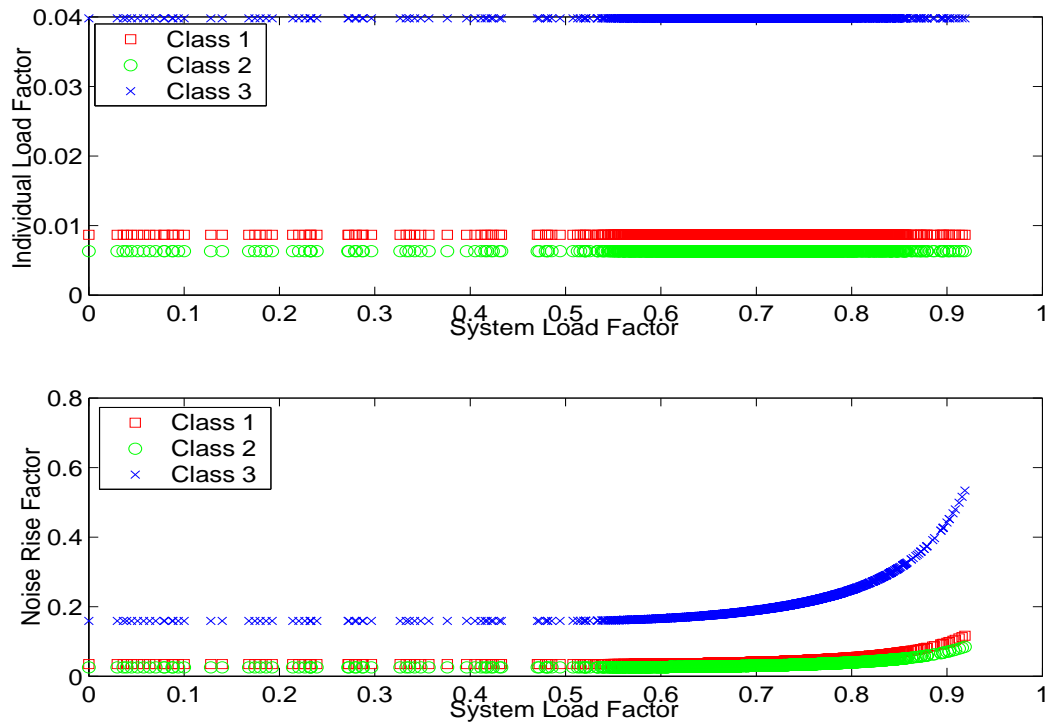


Figure 4.6. The relationship between ILF/NRF and System Load Factor

to prevent low-WTP users from entering the system when the interference level is high, thus avoiding them to further aggravate the situation.

The access, blocking and dropping probabilities of all policies are given in Fig. 4.7. The average access probability is lower with NRF-based pricing compared to their ILF counterpart. This re-emphasises the earlier point about the exponential increase of price during high interference level in the network. The higher access probability when O-NRF is used, compared to S-NRF, is due to the flexibility of the optimal policy to offer low prices to users when interference is low in the network. Even though the average access probabilities of O-NRF and S-ILF are close, the optimal policy provides a significantly better average reward due to resource allocation and CAC, in addition to interference-based pricing.

The NRF-based policies are also far more effective in controlling the blocking of new users and dropping of handoff users. The blocking and dropping probabilities decrease dramatically to negligible values when interference-based pricing is used in S-NRF and O-NRF. The results from Fig. 4.8 affirm that interference-based pricing as an effective mechanism for congestion control. The load factor of all policies is well below its constraint

4.5 Numerical Results

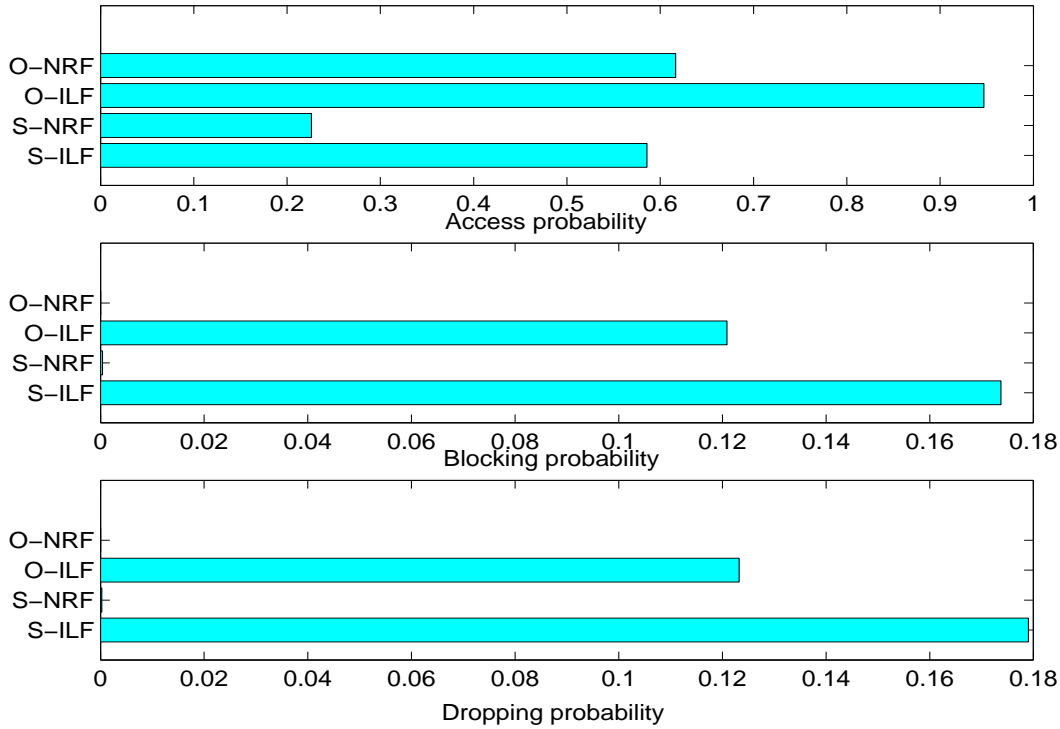


Figure 4.7. Blocking, dropping and access probabilities under normal load.

$\eta_{\max} = 0.98$, which translates to a maximum noise rise of about 17 dB, under normal traffic load. However, the system load factor and noise rise of O-NRF and S-NRF are notably lower. Compared to S-NRF, the use of O-NRF results in higher system load factor and noise rise because more users are allowed access to service. This results in more efficient usage of system resources.

4.5.2 Heavy Traffic Load

To simulate the system under heavy traffic conditions, we now double the arrival rates for both new and handoff users. The arrival rates become $\lambda^n = (10, 20, 20)$ and $\lambda^h = (2, 4, 4)$. All other simulation parameters from Table 4.1 remain the same. The average reward and proportion of reward obtained are shown in Fig. 4.9. Again, O-NRF collected significantly better average reward and proportion of reward obtained compared to other policies. The average reward collected by O-NRF after N simulation steps is 76%, 23% and 51% more than of S-ILF, S-NRF and O-ILF respectively.

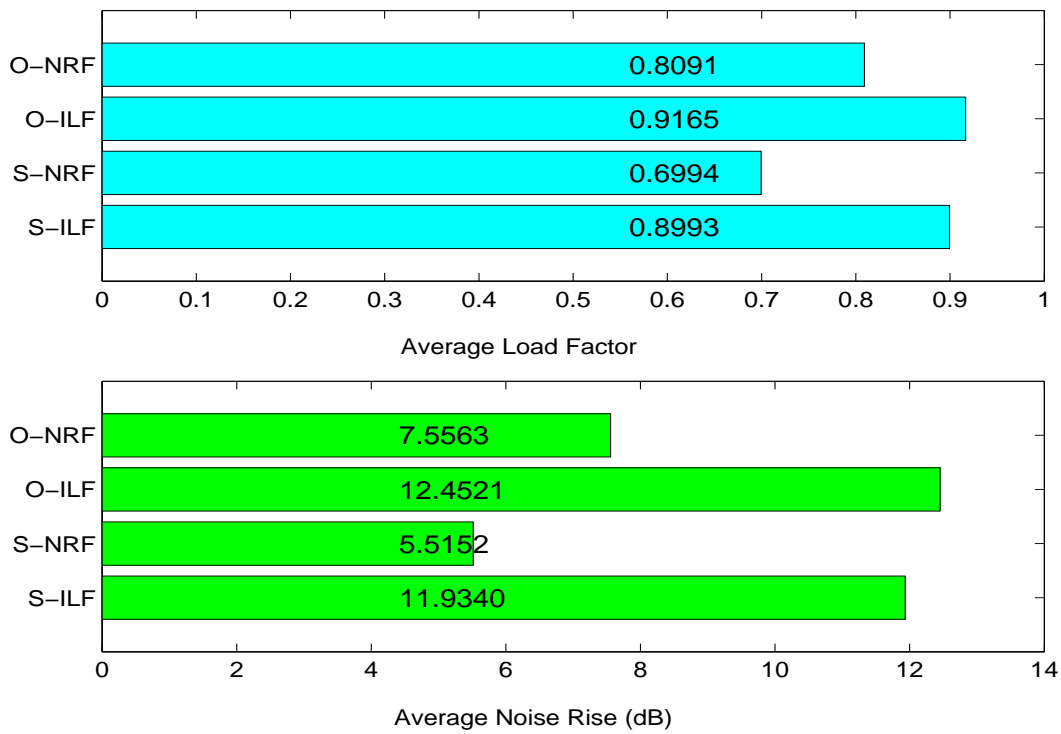


Figure 4.8. Average load factor and noise rise under normal load.

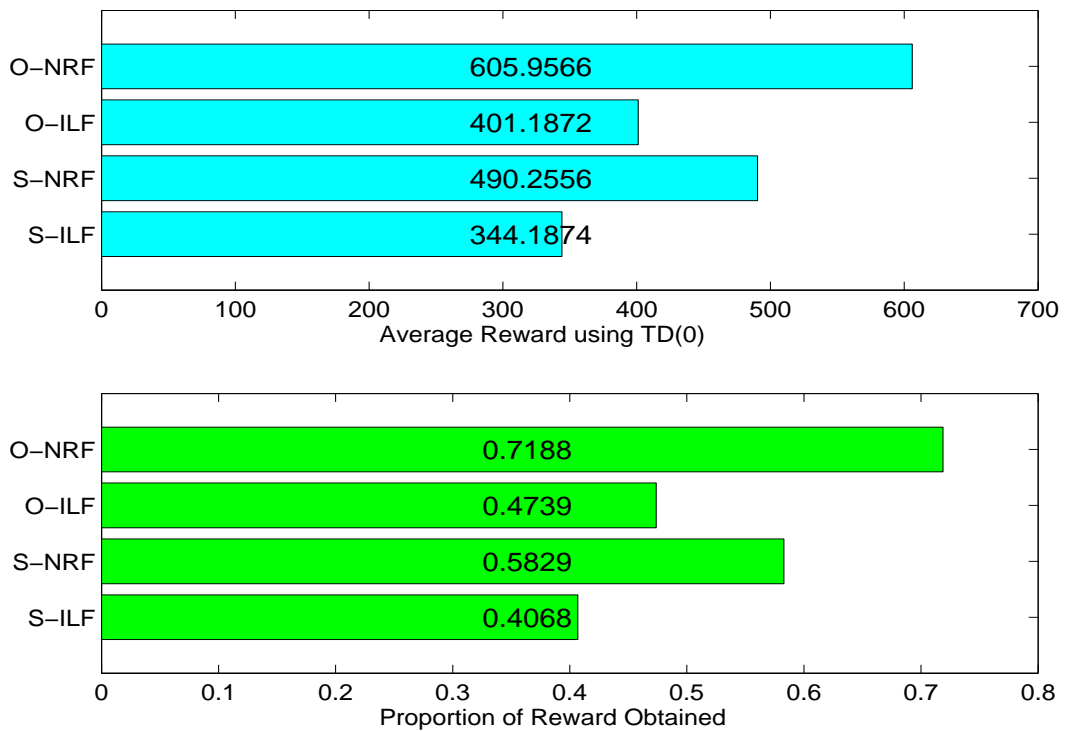


Figure 4.9. Average reward and proportion of reward obtained under heavy traffic load.

4.5 Numerical Results

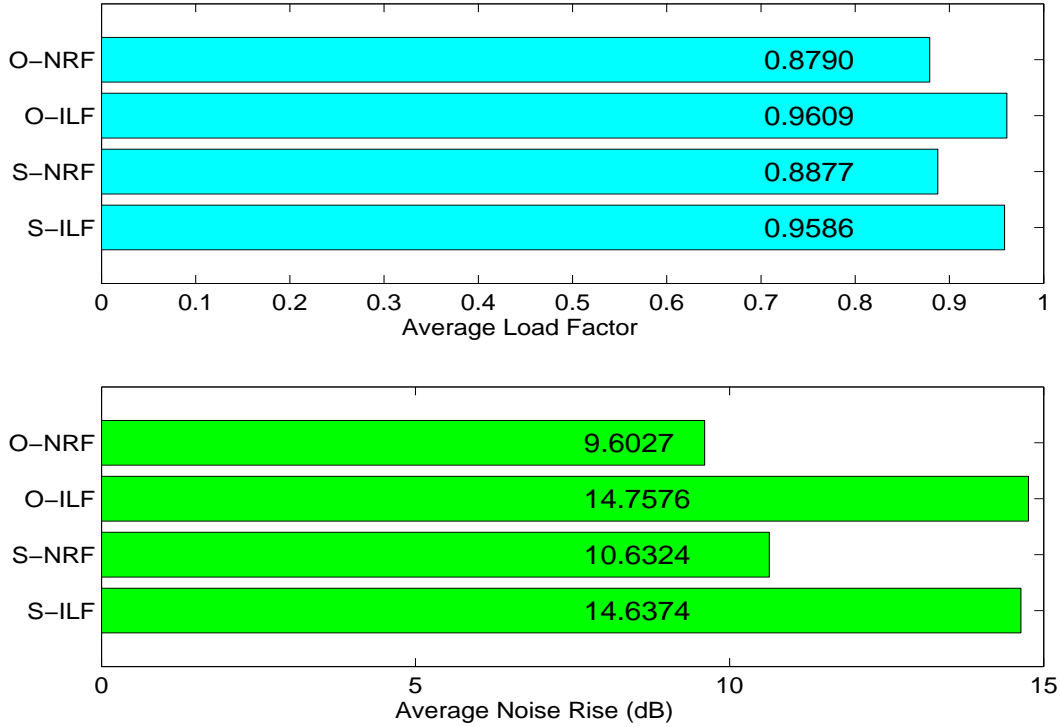


Figure 4.10. Average load factor and noise rise under heavy traffic load.

Due to the higher traffic load, the proportion of reward obtained by all policies are less than under normal traffic load because of the interference constraint of the system. The system load factor and noise rise (see Fig. 4.10) are correspondingly higher as well. However, O-NRF and S-NRF are more effective in controlling the system load factor and noise rise below $\eta_{\max} = 0.98$ and $\varphi_{\max} = 17$ dB respectively. Contrary to the initial expectation, the performance of O-ILF is worse than S-NRF. O-ILF only collected 47% of the available reward per time, compared to 58% by the S-NRF. A closer look at the simulation statistics indicates that the O-ILF is unable to find the higher prices during high interference. We will discuss in the next section how exploration can improve the performance of O-ILF.

4.5.3 Effects of Exploration

The results of O-NRF and O-ILF in the previous two subsections are solely based on the exploitation of the system current knowledge on actions that are effective in generating reward in the past. The lack of action exploration results in the load-based pricing optimal policy O-ILF performing worse than S-NRF in the previous section. Therefore,

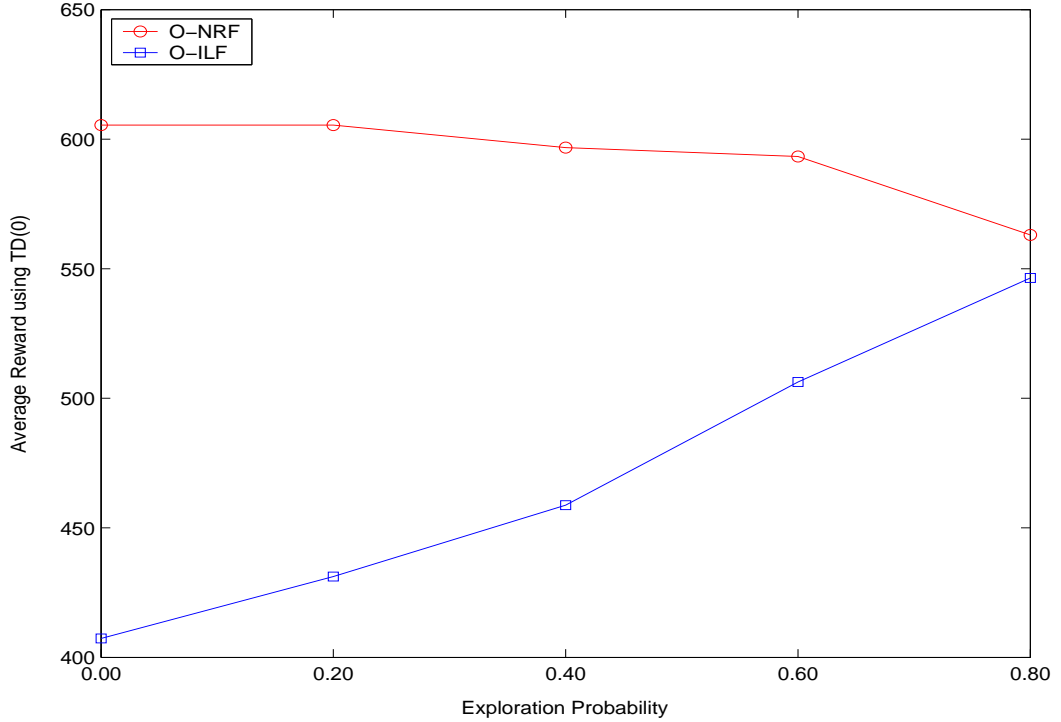


Figure 4.11. Average reward and proportion of reward obtained under policies O-ILF and O-NRF with 20% to 80% state exploration.

we repeat the simulation under heavy traffic using exploration probabilities P_ϵ of 0.20, 0.40, 0.60 and 0.80. For example when $P_\epsilon = 0.20$, a random action $\mathbf{u}_k \in \mathcal{U}$ is chosen in 20% of the decision-making steps, instead of the greedy action given by (4.26). Fig. 4.11 shows that the average reward of O-ILF improve with the increase of exploration. Its \tilde{J}_N even exceeds the average reward of S-NRF in the previous subsection when $P_\epsilon \geq 0.60$. Correspondingly, the proportion of reward obtained by O-ILF also increases with P_ϵ . However, exploration does not improve the performance of O-NRF at all. It produces worse results. The average reward collected decreases from 605.9566 when $P_\epsilon = 0$ to 571.5817 when $P_\epsilon = 0.8$. Simulation results of O-NRF indicate that random selection of lower prices reduced the average reward collected.

4.5.4 Effects of a Price Sliding Window

Up to this point, we have optimised O-ILF and O-NRF using $\#\mathcal{U}_p = 35$ uniformly distributed price levels given by (4.27). To accelerate the computation of \tilde{J}_k and θ_k using TD(0) learning algorithm, we restrict the size of the price decision space \mathcal{U}_p using a price

4.6 Conclusions

sliding window. We denote the size of the sliding window as SW . Given that the current price decision is \mathbf{u}_{pk} and $\mathcal{U}_p(i) = \mathbf{u}_{pk}$, where i is the index of \mathbf{u}_{pk} in \mathcal{U}_p , the price decision space for the next event is limited to:

$$\mathcal{U}_p^{k+1} = \mathcal{U}_p(a, \dots, i, \dots, b), \quad (4.28)$$

where $a = \max(1, i - \lfloor SW/2 \rfloor)$ and $b = \min(a + SW - 1, \#\mathcal{U}_p)$. The size of the price decision space is $\#\mathcal{U}_p^{k+1} = SW$. For example, if the current price index is $i = 9$ and $SW = 9$, $a = 5$ and $b = 13$, and therefore $\mathcal{U}_p^{k+1} = \mathcal{U}_p(5, \dots, 13)$.

The motivation for this strategy is based on the observation that a huge leap of price is unlikely during the simulation. A change from the minimum to the maximum price is unusual when a state transition of \mathbf{n} to $\mathbf{n} + \mathbf{e}_j$ occurs. The restriction on how much prices should change from one step to another also prevents sharp changes in price.

We repeat the simulation under heavy traffic load using no exploration and varying sizes of price sliding window of $SW = 5, 10, 15, 20, 25$ and 35 , the last of which is equivalent to having no sliding window. The average reward and proportion of reward obtained in all cases are displayed in Fig. 4.12. The average reward decreases from 605.96 when no sliding window is used to 588.94 when $SW = 5$, a total reduction of only 2.89%. The number of seconds needed for the execution of the simulation also decreases from 4817s when there is no sliding window to 4641s when $SW = 5$. This saving in computation time is expected to be more significant when more services are simulated.

4.6 Conclusions

In this chapter, we formulated an integrated dynamic pricing and radio resource management problem for an interference-limited network as a Neuro-Dynamic Programming problem. We also modified the conventional feature-based approximation architecture to have the effects of each decision on future reward rate be evaluated before the decision is made. A new parameter, Noise Rise Factor, is suggested as a basis for setting price in an interference-limited network. This parameter can effectively capture the positive, non-linear relationship between resource usage and the interference generated as the system loading increases. The average-reward TD(0) has been successfully applied and adapted to a pricing problem. We summarise the key results of this chapter as follows:

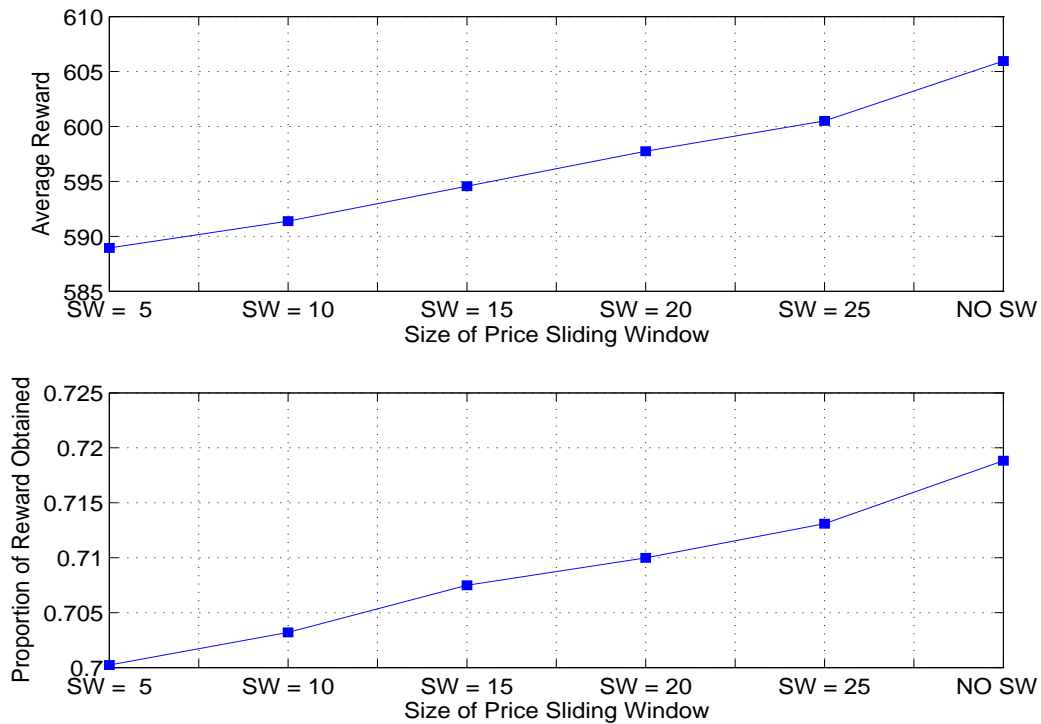


Figure 4.12. Results under heavy traffic load using a price sliding window.

- Optimal interference-based dynamic pricing and RRM policy increases average reward and reduces blocking and dropping probabilities compared to load-based schemes.
- Although the average reward obtained under optimal integrated policy increases, the average system load factor and noise rise actually decrease. This means that the optimal policy can achieve more with less! Higher reward in terms of actual and satisfaction revenue can be generated without overloading the system and generating high interference. The gain in average reward is also due to the admission and allocation of resources to users with high WTP.
- Action exploration during TD(0) learning can improve the integrated load-based and RRM policy. However, active exploration does not always lead to the improvement of the reward obtained in the case of the interference-based optimal policy.
- By implementing a price sliding window that restricts the size of the price decision space, the sub-optimal policy closely approximates that of the optimal policy with full choice of prices. This idea is based on the observation that price is unlikely to increase or decrease a lot during simulation.

Chapter 5

Cooperative Resource Allocation Games in Shared Networks

5.1 Introduction

Previously, in Chapters 3 and 4, our focus has been on the computation of an integrated admission pricing and resource management policy that maximises the operator's long-term expected reward in terms of actual and satisfaction revenue. In Chapter 4, we suggested users to preselect their minimum and maximum resource requirements at admission and allow the network operator to vary their transmission rates throughout their call. QoS was guaranteed in that at least their minimum transmission rate will be delivered. In this chapter, we will shift the focus to providing *fair* and *Pareto-optimal* resource allocation for connected users. As defined in Definition 2.7, an allocation is Pareto-optimal if there is no wasted utility, i.e. it is impossible to make any one party better off without making any other worse off. Such an outcome is also said to be efficient. The main question we are answering is:

“Given that all the users have received their minimum resource requirement and there is a resource surplus, how should the surplus be divided among them such that the allocation is fair and there are no wasted resources?”.

The notion of axiomatic bargaining (see Section 2.2.1) in cooperative game theory provides a good analytical framework to derive a desirable operating point that is both fair

5.1 Introduction

and Pareto-optimal. The issue of fairness have been mostly ignored in problems where the objective is to maximise the total utility or minimise the transmission power under some constraints such as [38,61,69,84,108,109]. Fairness is also not considered in power-control-motivated pricing literature using noncooperative game theory in [32,66,67,71,98,99,126,131]. It is also well known that the Nash equilibria from these noncooperative power control games are Pareto inefficient and the pricing mechanisms proposed only provide some Pareto improvements. The resulting degree of efficiency loss is known as the price of anarchy [57]. One measure of fairness that existed early in network literature is the notion of *max-min fairness*, which maximises the allocation for the most poorly treated user. However, max-min fairness gives priority to the worst performer, which in turn will reduce system performance. A notion of *proportional fairness*, where an allocation is made such that the sum proportional gains cannot be reduced, has been introduced by Kelly in [61]. It has been shown in [127] that proportional fairness is indeed a Nash bargaining solution.

Previous work on other resource allocation problems using cooperative game theory emphasise only *one* bargaining solution such as the Nash bargaining solution [18,40,127] and Raiffa-Kalai-Smorodinsky (Raiffa, hereafter) bargaining solution [35]. Our work is a major improvement over these works because it presents and analyses a class of fair and Pareto-optimal bargaining solutions based on the concept of preference functions developed by Cao for a two-user problem in [18]. To the best of our knowledge, this work is the first that provides explicit formulas for the Raiffa solution and all solutions between the Nash and Raiffa solutions on the Pareto-optimal boundary for both the symmetric and asymmetric cases. Our approach is better because it enables us to find a range of solutions on the Pareto-optimal boundary with Nash and Raiffa as special cases and choose the best solution within these optimal solutions that maximises some other criteria such as revenue. Besides, network operators can implement the solutions without any complex derivation algorithms such as in [40].

Unlike our work in Chapters 3 and 4, we will also consider the presence of other providers who share network resources with the main operator in the network. The high cost associated with the rollout of 3G services encourages operators to share network infrastructure. Network sharing among competing operators opens up a whole new range of research opportunities, especially in devising Radio Resource Management (RRM) strategies in a shared network. This work is the first comprehensive treatment of resource

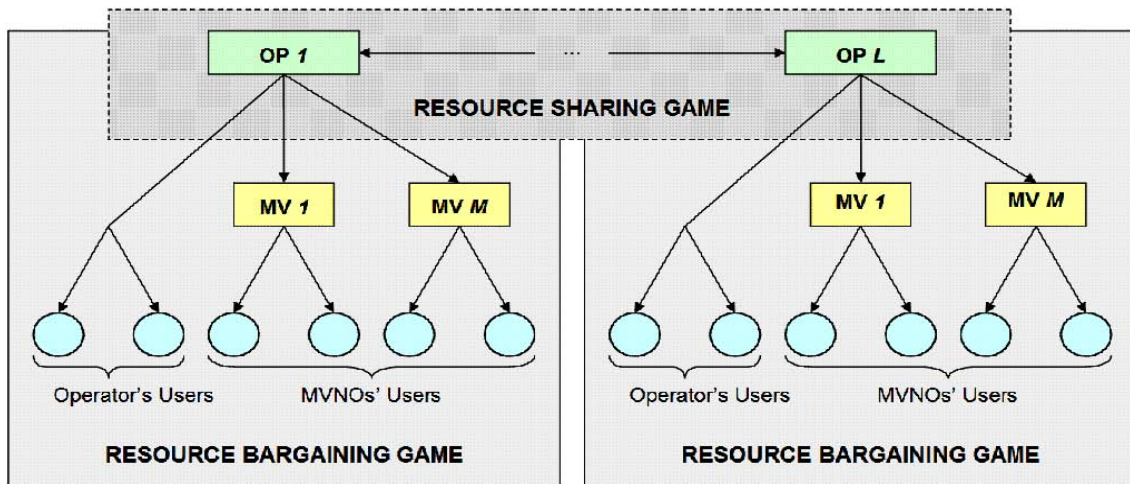


Figure 5.1. Resource allocation games in a shared network.

allocation problems in shared networks. Our focus is on networks that are completely-shared, where the operators share the core network, gateway core, RAN and sites, as in the case of Mobile Virtual Network Operator (MVNO), as well as geographical network sharing. Referring to Fig. 5.1, the resource allocation problem in a shared network can be divided into two sub-problems:

- *resource sharing* among the operators, i.e. the co-owners of the network; and
- *resource bargaining* among the users and MVNOs of each operator.

MVNOs, which are smaller service providers who do not own a 3G license, can be treated similarly to other network users with minimum and maximum resource requirements. These virtual operators purchase resources from the main operator with the intention of redistributing them to their subscribers. There are few existing published works that explore resource allocation strategies in a shared network environment. A simple admission control strategy with non-preemptive priority queueing has been proposed in [55], which sets the call admission priority of an operator according the ratio of its pre-agreed guaranteed load and its current load. In [52], the authors discuss a framework to manage radio resources using Service Level Agreements (SLA) among a network operator and its MVNOs. The downside of this proposal is that the SLA needs to be repeatedly renegotiated when the users traffic characteristics evolve.

5.1 Introduction

In our asymmetric bargaining model, players are allowed to influence the bargaining outcome using their bargaining powers. This is achieved by relaxing the axiom of symmetry used in [18]. Bargaining power is defined as the ratio of one's to the others' price or bid. If the price per unit resource is fixed throughout a call, as in Chapters 3 and 4, the bargaining power of users' is proportionate to their admission price compared to the total revenue. However, if prices are allowed to vary during a call, players can submit bids to the arbitrator to influence the bargaining outcome. The bargaining power is then interpreted as the ratio of one's bid with respect to others'. Asymmetric bargaining provides an opportunity to the operators to maximise their revenue and allocate scarce resources to users who need them most via the optimisation of parameter β . Unlike other resource auction approaches such as [57], the solution that we propose is still Pareto-optimal, so long as that appropriate value of β is selected. Pareto optimality is crucial in wireless networks because no scarce resources will be wasted.

We then extend our axiomatic bargaining approach to cooperative resource sharing among network operators. Although the amount of resources assigned to each operator is usually well-specified when the partnership is formed, operators who experience dissimilar demand patterns and non-coincident peak usage can benefit from exchanging their resources temporarily. We tackle the resource sharing game by categorising the operators as buyers, sellers or dummy players. To avoid the situation where the operators take additional resources but refuse to contribute, the allocation is history-dependent and will be based on the amount of resources the buyers contributed and/or obtained in the past. This problem is similar to the problem of electricity exchange between independent power producers studied in [97]. The operators can then redistribute their resources to their users and MVNOs.

The rest of the chapter is organised as follows. In Section 5.2, we will give an overview of network sharing in 3G networks. In Section 5.3, we present our system model. Then, we introduce our model of resource bargaining and derive the solutions in Sections 5.4 and 5.5 respectively. The resource sharing model among the operators is proposed in Section 5.6. Result analysis and conclusion are presented in Sections 5.7 and 5.8. This work has been partially presented in [46] and [47].

5.2 Network Sharing Models

Network infrastructure sharing has become a popular strategy among operators in the roll-out of 3G services, especially in the wake of substantial investments in licensing and slow 3G user growth. Operators are attracted to network sharing because of the lower capital expenditure (CAPEX) in infrastructure establishment and reduced operation expenditure (OPEX) in the long run. For example, a greenfield operator can save considerable costs by sharing its infrastructure with an incumbent operator. The acceleration of roll-out of 3G services, enabled by substantial cost savings, facilitates an earlier user acceptance of WCDMA and its related services.

In addition, operators can increase coverage by sharing or having complementary, geographically separated sites, especially in low-density suburban and rural areas where it is more cost-effective to share. The reduction in the number of new sites being built for base stations will also have positive environmental effects. Beyond the CAPEX and OPEX reductions, network sharing is argued to contribute far more to the 3G business case than cost reductions by allowing the value chain to be disaggregated into operators and other entities in the network [9]. Referring to Fig. 5.2, there are several sharing models available [29, 81, 105]:

- *Site sharing*: Operators share the site for the base station, the transmission to the Radio Network Controller (RNC) and equipment such as antenna towers, antenna environmental facilities, physical access and power supplies.

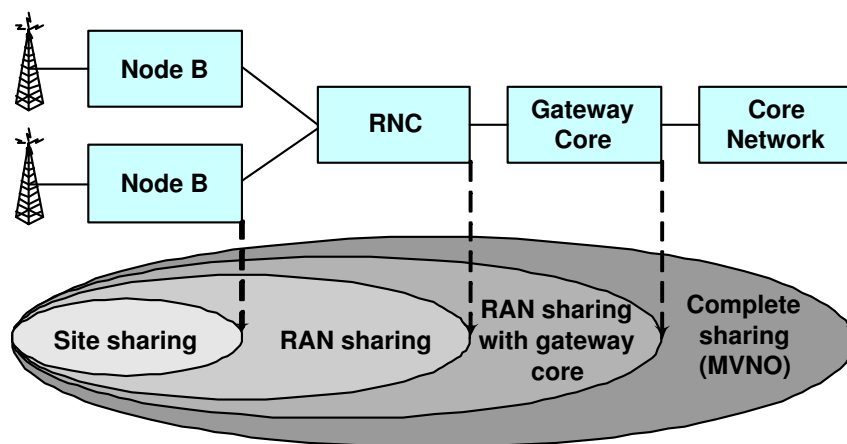


Figure 5.2. Models of network sharing.

5.3 System Model

- *Radio Access Network (RAN) sharing*: Operators share Node Bs and RNC but maintain their own gateway core and Core network. Operators have their own transceivers within the Node-B and use their own licensed frequencies. Operators still have a high level of independence.
- *RAN sharing with gateway core*: Operators have their own Core network and share a common RAN and gateway core. Operators use licensed frequencies from one operator only.
- *Mobile Virtual Network Operators (MVNOs)*: This approach is primarily used by MVNOs who do not have a 3G licence and have little infrastructure of their own except for their own home location register (HLR) and billing systems. At least one operator, who owns the network and has a 3G license, and radio-less MVNOs share a full-scale 3G infrastructure (RAN, Core network and backbone).
- *Geographical network sharing*: Operators have complementary 3G infrastructure in different areas of a country and share them via national roaming to extend coverage.

The last two models involve complete sharing of network infrastructure. A number of white papers, i.e. [29,81,105], have quantified the benefits of network sharing. Site sharing typically brings an overall saving of 20-30% while the savings increase to up to 40% when RAN sharing is also used.

5.3 System Model

We first consider a shared network with one operator and M MVNOs, denoted by $m \in \{1, \dots, M\}$. The cooperative resource sharing model for networks shared by more than one operator will be proposed in Section 5.6. Apart from serving its users, the operator sells unused resources to its MVNOs. These MVNOs do not own any resources and only have the ability to purchase them from the network operator and then resell them to their users. It is reasonable to assume that there is a pre-existing SLA between the operator and each MVNO to guarantee it at least R_m^{\min} units of resources. We denote the number of users associated with the operator or any of the m th MVNO as N_0 and N_m respectively.

Assume that the services provided by the operator are elastic and defined by a range of transmission rates bounded by minimum and maximum R^{\min} and R^{\max} . For example, the

UMTS Adaptive Multi-Rate (AMR) codec offers transmission rates that vary between 4.75 and 12.2 kbps for conversational voice service [49]. The transmission rate can be dynamically adjusted every 20 ms. We assume that users can select their acceptable QoS level by setting their range of transmission rates. In order to allocate *resources* in a fair and Pareto-optimal way, we first need to derive the meaning of a unit of resource. In a WCDMA network, resources can be expressed in terms of the load factors of the users. The load factors are commonly used to make a semi-analytical prediction of the capacity of a WCDMA cell without performing system-level simulations [49].

5.3.1 Uplink Load Factor

Consider a single WCDMA cell. In order for a signal to be received, the ratio of its received power to the sum of the background noise and interference must be greater than a given target. The target quality is translated to the following inequality that must be satisfied for each user $i = \{1, \dots, N\}$ [122], [109]:

$$\frac{W}{\nu_i x_i} \frac{g_i p_i}{\sum_{j \neq i}^N g_j p_j + \sigma^2} \geq \left(\frac{E_b}{N_0} \right)_i, \quad (5.1)$$

where W is the WCDMA chip rate, ν_i is the activity factor, x_i is the allocated transmission rate, g_i is the path gain between the base station and user i , p_i is transmission power, σ^2 is the background thermal noise power, $I_i = \sum_{j \neq i}^N g_j p_j$ is interference received by the base station from all the other users within the same cell and $\left(\frac{E_b}{N_0} \right)_i$ is the target bit-energy-to-noise-density required to meet predefined bit error rate (BER). In the case of multiple cells, the interference from other cells can be taken into account by using a coefficient f , i.e. $I_i = (1 + f) \sum_{j \neq i}^N g_j p_j$. Interference coefficient f typically has values between 0.1 and 0.6 [49].

Assuming perfect power control and solving the set of equations in (5.1), we obtain the following:

$$g_i p_i = \eta_i^{\text{UL}} \left(\sum_{j=1}^N g_j p_j + \sigma^2 \right) = \frac{\sigma^2 \eta_i^{\text{UL}}}{1 - \sum_{j=1}^N \eta_j^{\text{UL}}}, \quad (5.2)$$

5.3 System Model

where the load factor η_i^{UL} and total interference I are respectively given by

$$\eta_i^{\text{UL}} = \frac{1}{\frac{W}{\left(\frac{E_b}{N_0}\right)_i \nu_i x_i} + 1} \quad (5.3)$$

$$I = \sum_{j=1}^N g_j p_j = \frac{\sigma^2 \sum_{j=1}^N \eta_j^{\text{UL}}}{1 - \sum_{j=1}^N \eta_j^{\text{UL}}}. \quad (5.4)$$

The number of users that can be supported by the network is limited by the maximum uplink system load factor allowed, i.e.

$$\sum_{j=1}^N \eta_j^{\text{UL}} \leq \bar{\eta}^{\text{UL}} < 1. \quad (5.5)$$

When the total load factor approaches unity, the system reaches its pole capacity and the total interference increases to infinity in (5.4). If this constraint is violated, the target $\left(\frac{E_b}{N_0}\right)_i$ for all users will not be satisfied. We say that the uplink is interference-limited, i.e. users cannot increase their power without bound because of the increased interference they caused to other users. The corresponding transmission rate allocated is

$$x_i = \frac{W}{\left(\frac{E_b}{N_0}\right)_i \nu_i \left(\frac{1}{\eta_i^{\text{UL}}} - 1\right)}, \quad (5.6)$$

which increases according to the load factor allocated.

5.3.2 Downlink Load Factor

In the downlink, the target signal quality of user i is [109]

$$\frac{W}{\nu_i x_i} \frac{g_i p_i}{\theta_i g_i \sum_{j \neq i}^N p_j + \sigma^2} \geq \left(\frac{E_b}{N_0}\right)_i, \quad (5.7)$$

where $I_i = \theta_i g_i \sum_{j \neq i}^N p_j$ and θ_i is the orthogonality factor of the codes used in the downlink. Although WCDMA employs orthogonal codes, users will receive part of the base station signal due to multipath propagation. In the uplink, transmission is asynchronous and therefore the signals are not orthogonal. Typically, the orthogonality factor θ_i can be approximated by an average value $\bar{\theta}$, which is usually between 0.5 (ITU Vehicular A channel) and 0.9 (ITU Pedestrian B channel) [49]. The total transmission power in the downlink is limited by the maximum power that the base station can transmit, i.e.

$$\sum_{j=1}^N p_j \leq p^{\text{max}}. \quad (5.8)$$

Assuming perfect power control, the transmission power to the i th user can be expressed as

$$p_i = \frac{\sum_{j=1}^N p_j + \frac{\sigma^2}{g_i \theta_i}}{1 + \frac{W}{\theta_i \left(\frac{E_b}{N_0}\right)_i \nu_i x_i}}. \quad (5.9)$$

Using (5.9) and (5.7), the downlink load factor is given as

$$\eta_i^{\text{DL}} = \frac{1 + \frac{\sigma^2}{g_i \theta_i p^{\text{max}}}}{1 + \frac{W}{\theta_i \left(\frac{E_b}{N_0}\right)_i \nu_i x_i}}. \quad (5.10)$$

Unlike in the uplink, the downlink load factor depends on the orthogonality factor and path gain between the user and the base station. Similar to the uplink, the total downlink load factors must satisfy

$$\sum_{i=1}^N \eta_i^{\text{DL}} \leq 1. \quad (5.11)$$

Using $\hat{\eta}_i^{\text{DL}} = \frac{\eta_i^{\text{DL}}}{1 + \frac{\sigma^2}{g_i \theta_i p^{\text{max}}}}$ to express (5.9) in terms of the downlink load factor, we have the following:

$$p_i = \hat{\eta}_i^{\text{DL}} \left(\frac{\sum_{j=1}^N \hat{\eta}_j^{\text{DL}} \frac{\sigma^2}{g_j \theta_j}}{1 - \sum_{j=1}^N \hat{\eta}_j^{\text{DL}}} + \frac{\sigma^2}{g_i \theta_i} \right). \quad (5.12)$$

Given η_i^{DL} , the transmission rate allocated to the i th user is then expressed as

$$x_i = \frac{\frac{W}{\theta_i \left(\frac{E_b}{N_0}\right)_i \nu_i}}{\frac{1}{\eta_i^{\text{DL}}} \left(1 + \frac{\sigma^2}{g_i \theta_i p^{\text{max}}}\right) - 1}, \quad (5.13)$$

which increases as the allocated downlink load factor η_i^{DL} and path gain g_i increase.

5.4 Cooperative Game Theory Framework

In the bargaining framework, the players or bargainers in our problem are the MVNOs and users of the operators. Therefore, there are a total number of $N = N_0 + M$ players in the network. Define \mathcal{S} as the bargaining domain or the feasible set of all possible outcomes. \mathcal{S} is assumed to be convex, closed and bounded sets of \mathbb{R}^N . Each player $i \in \{1, \dots, N\}$ competes for the use of resources and has a

5.4 Cooperative Game Theory Framework

- *utility function* $u_i = \eta_i$, which is represented by the allocated load factor. Any point $\mathbf{u} \in \mathcal{S}$ represents an outcome or solution of the game.
- *desired initial performance* u_i^{\min} , which is the minimal performance required by the user without any cooperation in order to enter the game. It is also known as the *disagreement point* or *threat point*. Players will not enter the game if u_i^{\min} is not achievable.

The bargaining problem and outcome can be defined as $(\mathcal{S}, \mathbf{u}^{\min})$ and $F(\mathcal{S}, \mathbf{u}^{\min}) \in \mathcal{S}$ respectively. Approaches to bargaining fall into one of two divisions: *strategic* and *axiomatic* bargaining. Strategic bargaining, such as the Rubinstein's model of bargaining [74], assumes that there is a bargaining process where the solution is achieved in a series of offers and counteroffers. The bargaining solution emerges as the equilibrium of a sequential game. The need for a bargaining process among the players is time-consuming and therefore unsuitable for WCDMA network with many users.

Axiomatic bargaining ignores the bargaining process and assumes some desirable properties about the outcome of the bargaining process and then identifies process rules or axioms that guarantee this outcome. The operator serves as the arbitrator in the cooperative resource bargaining game. Nash specifies four axioms, which impose properties that a bargaining solution should satisfy:

- A1 *Invariance with respect to affine transformation*: If \mathbf{u}^* is the solution to $(\mathcal{S}, \mathbf{u}^{\min})$ and y is any positive affine transformation, the solution to $(y(\mathcal{S}), y(\mathbf{u}^{\min}))$ is $y(\mathbf{u}^*)$.
- A2 *Symmetry*: If the bargaining problem is symmetric, in the sense that (e.g. $N = 2$) $u_1^{\min} = u_2^{\min}$ and $(u_1, u_2) \in \mathcal{S} \Leftrightarrow (u_2, u_1) \in \mathcal{S}$, then $F_1(\mathcal{S}, \mathbf{u}^{\min}) = F_2(\mathcal{S}, \mathbf{u}^{\min})$. This means that two players with symmetric utilities get the same payoff.
- A3 *Pareto Optimality*: The bargaining solution will be on the Pareto boundary. If $(\mathcal{S}, \mathbf{u}^{\min})$ is a bargaining problem and $\mathbf{u}, \mathbf{u}' \in \mathcal{S}$ and $u'_j > u_j, j = 1, \dots, N$, then the outcome $F(\mathcal{S}, \mathbf{u}^{\min}) \neq \mathbf{u}$.
- A4 *Independence of Irrelevant Alternatives*: If $(\mathcal{S}, \mathbf{u}^{\min})$ and $(\mathcal{S}', \mathbf{u}^{\min})$ are bargaining problems with $\mathcal{S} \subseteq \mathcal{S}'$ and $F(\mathcal{S}', \mathbf{u}^{\min}) \in \mathcal{S}$, then $F(\mathcal{S}, \mathbf{u}^{\min}) = F(\mathcal{S}', \mathbf{u}^{\min})$.

Axiom A4 has received a number of criticisms. In particular, Kalai and Smorodinsky [59] (and Raiffa in an earlier work [93]) argued that one's gain should be proportional to its

maximum gain but the Nash solution fails to satisfy this requirement. They retained A1-A3 and proposed a new axiom:

A5 *Monotonicity*: If $\mathcal{S} \subseteq \mathcal{S}'$ ($N = 2$), $u_1(\mathcal{S}') = u_1(\mathcal{S})$ and $u_2(\mathcal{S}') \geq u_2(\mathcal{S})$, $F_2(\mathcal{S}', \mathbf{u}^{\min}) \geq F_2(\mathcal{S}, \mathbf{u}^{\min})$.

Cao in [18] explained that the Nash and Raiffa solutions represent different solution points on the Pareto boundary. There is no special reason why they should be chosen and one might choose another point on the boundary if one dislikes the properties of the Nash and Raiffa solutions. Bargaining solutions can be analysed using players' preference function. In the two-user case, with disagreement points $u_1^{\min} = u_2^{\min} = 0$, the players' preference functions are defined as [18]

$$v_1 = u_1 + \beta(1 - u_2) \tag{5.14}$$

$$v_2 = u_2 + \beta(1 - u_1), \tag{5.15}$$

where $0 \leq u_1, u_2 \leq 1$ and β is a weighting factor that measures the trade-off between one's gain and another's loss. The bargaining outcome, \mathbf{u}^* is the solution to $\mathbf{u}^* = \arg \max_{\mathbf{u}} (v_1 v_2)$. The special cases of $\beta = 0, 1, -1$ correspond to the Nash, Raiffa and modified Thomson solutions. The Nash solution only considers individual gains and ignores how much other players may gain or lose. The Raiffa solution places the same weight on individual gain and other players' losses. The modified Thomson solution, also known as the relative utilitarian outcome, maximises the sum of all players' normalised utilities.

For the multi-player case, we define the i th player's preference function with minimum and maximum utility, u_i^{\min} and u_i^{\max} , as follows:

$$v_i(\beta) = u_i - u_i^{\min} + \frac{\beta}{N-1} \left(\sum_{j \neq i} u_j^{\max} - u_j \right), \tag{5.16}$$

where $\beta = 0, 1, -(N-1)$ corresponds to the Nash, Raiffa and utilitarian solutions respectively. Our definition does not require u_i to be normalised by its maximum value since u_i^{\max} is included and it is general enough to include the special case of normalised utility in [18] and [28]. The β -dependent bargaining outcome, $\mathbf{u}^*(\beta)$, is the solution to

$$\mathbf{u}^*(\beta) = \arg \max \prod_{i=1}^N v_i(\beta). \tag{5.17}$$

We call this the symmetric *parameterised* solution of the bargaining problem.

5.5 Resource Bargaining in WCDMA

In our WCDMA resource bargaining problem, η_i^{\min} and η_i^{\max} correspond to the minimum and maximum acceptable load factors based on the player's requirement of the minimum and maximum transmission rates, x_i^{\min} and x_i^{\max} , defined in (5.3) and (5.10). If the player is a user of the operator, $R^{\min} \leq x_i^{\min}, x_i^{\max} \leq R^{\max}$, $i \in \{1, \dots, N\}$. For the m th MVNO, its guaranteed minimum resource allocation is R_m^{\min} , which can be in terms of the transmission rate or load factor. Its maximum requirement will be in terms of the maximum requirements of all of its N_m users.

5.5.1 Symmetric Bargaining

We first derive the symmetric bargaining problem with $u_i = \eta_i$ where all players are assumed to have equal bargaining power. The symmetric resource bargaining problem is defined as (P1):

$$\begin{aligned} \max_{\boldsymbol{\eta}} \quad & \prod_{i=1}^N \left(\eta_i - \eta_i^{\min} + \frac{\beta}{N-1} \sum_{j \neq i} (\eta_j^{\max} - \eta_j) \right) \\ \text{s.t.} \quad & \eta_i \geq \eta_i^{\min}, \eta_i \leq \eta_i^{\max}, \sum_{i=1}^N \eta_i \leq T. \end{aligned} \quad (5.18)$$

Referring to (5.5) and (5.11), the resource constraint parameter T corresponds to $\bar{\eta}^{\text{UL}}$ and 1 for the uplink and downlink respectively. Note that a similar formulation has been considered in [127] but the authors only focus on the Nash solution, i.e. $\beta = 0$. We are interested in deriving a range of bargaining solutions, parameterised by β , on the Pareto boundary.

Proposition 5.1. *Under the assumption of $\eta_i^{\min} \leq \eta_i \leq \eta_i^{\max}$, $\sum_{i=1}^N \eta_i \leq T$ and $\sum_{i=1}^N \eta_i^{\min} < T < \sum_{i=1}^N \eta_i^{\max}$, the symmetric bargaining solution, parameterised by weighting factor β , $-(N-1) < \beta \leq 1$, of the problem (P1) is given by*

$$\eta_i(\beta) = \min\{\tilde{\eta}_i(\beta), \eta_i^{\max}\}, \quad (5.19)$$

$$\text{where } \tilde{\eta}_i(\beta) = \frac{T}{N} + \frac{(N-1)(N\eta_i^{\min} - \sum_{j=1}^N \eta_j^{\min}) + \beta(N\eta_i^{\max} - \sum_{j=1}^N \eta_j^{\max})}{N(N-1+\beta)}. \quad (5.20)$$

Proof. Let bargaining domain \mathcal{S} be a nonempty, convex and compact set. The optimisation of (P1) is equivalent to the optimisation of its logarithm [127]. The Lagrangian equation of the equivalent problem is given as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\eta}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \gamma) &= \sum_{i=1}^N \ln \left(\eta_i - \eta_i^{\min} + \frac{\beta}{N-1} \sum_{j \neq i} (\eta_j^{\max} - \eta_j) \right) - \sum_{i=1}^N \lambda_i (\eta_i^{\min} - \eta_i) \\ &\quad - \sum_{i=1}^N \mu_i (\eta_i - \eta_i^{\max}) - \gamma \left(\sum_{i=1}^N \eta_i - T \right). \end{aligned}$$

Let $f(\eta_i) = \frac{1}{\eta_i - \eta_i^{\min} + \frac{\beta}{N-1} \sum_{j \neq i} (\eta_j^{\max} - \eta_j)}$, the necessary and sufficient Karush-Kuhn-Tucker (KKT) conditions for optimality for $i \in \{1, \dots, N\}$ are

$$f(\eta_i) - \frac{\beta}{N-1} \sum_{j \neq i} f(\eta_j) = \lambda_i - \mu_i + \gamma \quad (5.21)$$

$$\gamma \left(\sum_{i=1}^N \eta_i - T \right) = 0. \quad (5.22)$$

When constraints $(\sum_{i=1}^N \eta_i - T)$ is active and $(\eta_i^{\min} - \eta_i)$ and $(\eta_i - \eta_i^{\max})$ are inactive for all $i \in \{1, \dots, N\}$, $\lambda_i = \mu_i = 0$ and $\gamma \geq 0$. Solving these equations, we have $f(\eta_i) = f(\eta_j)$ for all $j \neq i$, $i, j \in [1, N]$ or

$$\eta_i(\beta) = \eta_j + \frac{(N-1)(\eta_i^{\min} - \eta_j^{\min}) + \beta(\eta_i^{\max} - \eta_j^{\max})}{N-1+\beta}. \quad (5.23)$$

Using (5.23) and condition $\sum_{j=1}^N \eta_j = T$, the solution η_i can be derived accordingly. \square

Proposition 5.2. *When players take into account the utility loss of other players in their preference function by setting weighting factor $0 < \beta \leq 1$, the absolute gap between the new outcome and the Nash solution increases by up to*

$$\Delta_i = \frac{(N-1)(\eta_i^{\max} - \eta_i^{\min}) - (\sum_{j \neq i}^N \eta_j^{\max} - \eta_j^{\min})}{N^2} \quad (5.24)$$

when $\beta = 1$. The utility, measured in terms of the load factor, of i th player using the Raiffa solution ($\beta = 1$) is more than the Nash solution if $\eta_i^{\max} - \eta_i^{\min} > \sum_{j \neq i}^N \frac{\eta_j^{\max} - \eta_j^{\min}}{N-1}$.

Proof. From Proposition 5.1, the Nash ($\beta = 0$) and Raiffa ($\beta = 1$) solutions are respectively given by $\eta_i(\beta) = \min\{\tilde{\eta}_i(\beta), \eta_i^{\max}\}$ and

$$\tilde{\eta}_i(0) = \eta_i^{\min} + \frac{T - \sum_{j=1}^N \eta_j^{\min}}{N} \quad (5.25)$$

$$\tilde{\eta}_i(1) = \tilde{\eta}_i(0) + \Delta_i. \quad (5.26)$$

The second part of the proposition is obvious. \square

5.5 Resource Bargaining in WCDMA

The Nash solution in (5.25) only takes into account the individual gain $\eta_i - \eta_i^{\min}$ and coincides with the two-user bargaining outcome derived in [74]. The result is known as the *split-the-difference* rule. The Raiffa solution in (5.26) places the same importance on one's gain and the losses of others. This solution is fairer to players with high maximum utility. On the other hand, as β approaches $-(N - 1)$, more weight is placed on other players' gain. When $\beta = -(N - 1)$, the problem maximises the sum of utilities of all players. However, there is no trivial solution for this problem as $(N - 1 + \beta)$ approaches 0 in (5.1) when β approaches $-(N - 1)$. When $-(N - 1) < \beta < 0$, the weight on other players' utility is less than 1.

5.5.2 Asymmetric Bargaining

Imposing the axiom of symmetry A2 in (P1) assumes that all players have equal bargaining skills. In practice, the bargaining outcome may be influenced by other variables such as the tactics employed by the bargainers, the negotiation procedure and the information structure [74]. In our asymmetric resource bargaining model, we allow the final outcome to be influenced by the price paid by all players. Suppose that each player $i \in \{1, \dots, N\}$ can submit a bid $\tau_i \in \mathbb{R}$ to the network operator, which is also the arbitrator. We then define the asymmetric resource bargaining problem (P2) as follows:

$$\begin{aligned} \max_{\boldsymbol{\eta}} \quad & \prod_{i=1}^N \left(\eta_i - \eta_i^{\min} + \frac{\beta}{N-1} \sum_{j \neq i} (\eta_j^{\max} - \eta_j) \right)^{\tau_i} \\ \text{s.t.} \quad & \eta_i \geq \eta_i^{\min}, \eta_i \leq \eta_i^{\max}, \sum_{i=1}^N \eta_i \leq T. \end{aligned} \quad (5.27)$$

Proposition 5.3. *Under the assumption of $\eta_i^{\min} \leq \eta_i^{\text{AS}} \leq \eta_i^{\max}$, $\sum_{i=1}^N \eta_i \leq T$ and $\sum_{i=1}^N \eta_i^{\min} < T < \sum_{i=1}^N \eta_i^{\max}$ the asymmetric bargaining solution, parameterised by weighting factor β , $-(N - 1) < \beta \leq 1$, of the problem (P2) is given by*

$$\eta_i^{\text{AS}}(\beta) = \min\{\tilde{\eta}_i^{\text{AS}}(\beta), \eta_i^{\max}\}, \quad (5.28)$$

$$\begin{aligned} \tilde{\eta}_i^{\text{AS}}(\beta) = \hat{\tau}_i T + & \frac{(N-1)(\eta_i^{\min} - \hat{\tau}_i \sum_{j=1}^N \eta_j^{\min})}{N-1+\beta} \\ & + \frac{\beta[(1-N\hat{\tau}_i)T + \eta_i^{\max} + ((N-1)\hat{\tau}_i - 1) \sum_{j=1}^N \eta_j^{\max}]}{N-1+\beta}. \end{aligned} \quad (5.29)$$

$\hat{\tau}_i = \frac{\tau_i}{\sum_{j=1}^N \tau_j}$ can be interpreted as the bargaining power of the i th player and the sum of all bargaining powers is equal to one. When the bids submitted by all players are the same, the asymmetric solution (5.28) is the same as the symmetric solution derived in (5.19).

Proof. The derivation is similar to the one in the previous section and will therefore be omitted. It is easy to see that the symmetric solution (5.19) is a special instance of the asymmetric solution. The second part of the proof can be obtained using $\hat{\tau}_i = \tau_i = \frac{1}{N}$. \square

Proposition 5.4. *Similar to Proposition 5.2, when players take into account the utility loss of other players, i.e. $0 < \beta \leq 1$, the absolute gap between the new outcome and the Nash solution increases by up to*

$$\Delta_i^{\text{AS}} = \frac{(1 - N\hat{\tau}_i)T + \hat{\tau}_i(N - 1)\eta_i^{\text{max}} + [(N - 1)\hat{\tau}_i - 1] \sum_{j \neq i}^N \eta_j^{\text{max}} - \eta_i^{\text{min}} + \hat{\tau}_i \sum_{j=1}^N \eta_j^{\text{min}}}{N} \quad (5.30)$$

when $\beta = 1$.

Proof. Using Proposition 5.3, the Nash and Raiffa solutions are respectively given by

$$\tilde{\eta}_i^{\text{AS}}(0) = \eta_i^{\text{min}} + \hat{\tau}_i(T - \sum_{j=1}^N \eta_j^{\text{min}}) \quad (5.31)$$

$$\tilde{\eta}_i^{\text{AS}}(1) = \tilde{\eta}_i^{\text{AS}}(0) + \Delta_i^{\text{AS}}. \quad (5.32)$$

When the Raiffa solution is used, the utility of the i th player will only be greater than the utility derive from the Nash solution, i.e. $\tilde{\eta}_i^{\text{AS}}(1) > \tilde{\eta}_i^{\text{AS}}(0)$, when $\Delta_i^{\text{AS}} > 0$. \square

The asymmetric Nash and Raiffa solutions derived in (5.31) and (5.32) exhibit the same properties as the symmetric solutions in the previous section. The asymmetric Nash solution only varies according to the minimum load factor requirements and bargaining powers of all players. The asymmetric Raiffa solution also takes into account the players' maximum load factor requirements. In Proposition 5.3, the bargaining solution $\tilde{\eta}_i^{\text{AS}}(\beta)$ is limited by its maximum bound η_i^{max} . In cases where $\tilde{\eta}_i^{\text{AS}}(\beta) > \eta_i^{\text{max}}$ for some $\beta \in [0, 1]$, the excess can be redistributed to other players with $\tilde{\eta}_j^{\text{AS}}(\beta) < \eta_j^{\text{max}}$ for some $i, j = 1, \dots, N$ and $i \neq j$. One way of doing this is to use Proposition 5.3 to reallocate the total excess, $T' = \sum_{i \in \mathcal{I}_E} \tilde{\eta}_i^{\text{AS}}(\beta) - \eta_i^{\text{max}}$ to other players $j \notin \mathcal{I}_E$ who haven't reached their maximum requirement. $\mathcal{I}_E = \{i : \tilde{\eta}_i^{\text{AS}}(\beta) > \eta_i^{\text{max}}, i = 1, \dots, N\}$ denotes the set of players with surplus. Another way of redistributing the excess is to choose a feasible range of β such that $\eta_i^{\text{min}} \leq \eta_i^{\text{AS}}(\beta) \leq \eta_i^{\text{max}} \forall i$.

5.5 Resource Bargaining in WCDMA

Proposition 5.5. For a given bid vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)$, β must satisfy the following condition such that $\eta_i^{\min} \leq \tilde{\eta}_i^{\text{AS}}(\beta) \leq \eta_i^{\max}$ for all players $i = 1, \dots, N$:

$$\beta \begin{cases} \in [\beta^{\min}, \beta^{\max}] & \text{if } \beta^{\min} < \beta^{\max} \\ \notin [0, 1] & \text{otherwise.} \end{cases} \quad (5.33)$$

Minimum and maximum constraints, β^{\min} and β^{\max} , are defined as:

$$\beta^{\min} = \arg \max \{ \beta_i^1 : \beta_i^1 \geq 0, i = 1, \dots, N \} \quad (5.34)$$

$$\beta^{\max} = \arg \min \{ \beta_i^2 : 0 \leq \beta_i^2 \leq 1, i = 1, \dots, N \} \quad (5.35)$$

using the following individual constraints on β for each player $i = 1, \dots, N$:

$$\beta_i^1 = \frac{(N-1)[\eta_i^{\max} - \eta_i^{\min} - \hat{\tau}_i(T - \sum_{j=1}^N \eta_j^{\min})]}{(1 - (N-1)\hat{\tau}_i)(T - \sum_{j=1}^N \eta_j^{\max})} \quad (5.36)$$

$$\beta_i^2 = \frac{-(N-1)\hat{\tau}_i(T - \sum_{j=1}^N \eta_j^{\min})}{\eta_i^{\max} - \eta_i^{\min} + (1 - (N-1)\hat{\tau}_i)(T - \sum_{j=1}^N \eta_j^{\max})}. \quad (5.37)$$

Proof. This proposition can be obtained by first deriving the individual minimum and maximum constraints β_i^1 and β_i^2 on β of each player i using $\tilde{\eta}_i(\beta) \leq \eta_i^{\max}$ and $\tilde{\eta}_i(\beta) \geq \eta_i^{\min}$:

$$\tilde{\eta}_i(\beta) \leq \eta_i^{\max} \Rightarrow \beta \leq \beta_i^1 \quad (5.38)$$

$$\tilde{\eta}_i(\beta) \geq \eta_i^{\min} \Rightarrow \beta \leq \beta_i^2 \text{ if } \beta_i^2 < 0 \text{ or } \beta \geq \beta_i^2 \text{ if } \beta_i^2 \geq 0. \quad (5.39)$$

Since we are considering β in the range of $[0, 1]$, we can ignore $\beta \leq \beta_i^2$ when $\beta_i^2 < 0$. The resulting β^{\min} and β^{\max} are the maximum of β_i^1 and minimum of β_i^2 , respectively. \square

The second case in (5.33) corresponds to the situation where there is no feasible range of β within $[0, 1]$. This situation occurs because one or more players have very high bargaining power compared to the others, as revealed through submission of very high bids. In practical situations where there are many players in the resource bargaining game, it is unlikely to have a few players who are very dominant and have incentives to submit very high bids that return $\tilde{\eta}_i > \eta_i^{\max}$. In any case, the network operator can set minimum and maximum bounds, τ_i^{\min} and τ_i^{\max} , on the bid submitted by each player to prevent this situation from occurring. The bounds are given by:

$$\tau_i^{\min} = \frac{\beta(\eta_i^{\min} + T - \sum_{j \neq i}^N \eta_j^{\max}) \sum_{j \neq i}^N \tau_j}{(N-1 - N\beta)T + (\sum_{j=1}^N N\beta\eta_j^{\max} - (N-1)\eta_j^{\min}) - \beta(\eta_i^{\min} + \eta_i^{\max})} \quad (5.40)$$

$$\tau_i^{\max} = \frac{[(N-1)(\eta_i^{\max} - \eta_i^{\min}) + \beta(T - \sum_{j=1}^N \eta_j^{\max})] \sum_{j \neq i}^N \tau_j}{(N-1 - N\beta)T + \sum_{j=1}^N (N\beta\eta_j^{\max} - (N-1)\eta_j^{\min}) - (N-1)(\eta_i^{\max} - \eta_i^{\min})} \quad (5.41)$$

5.5.3 Revenue Optimisation

The asymmetric bargaining solutions proposed in Proposition 5.3 provide a range of solutions with varying emphasis on one's gain and others' losses. When the bargaining power of all players are the same (i.e. Proposition 5.1), the selection of β is arbitrary. However, when players submit asymmetric bids, β can be chosen such that the total revenue received by the operator is optimal. Therefore, we consider the following revenue optimisation problem (P3):

$$\max_{\beta} R(\beta) = \sum_{i=1}^N \eta_i^{\text{AS}}(\beta) \tau_i \quad (5.42)$$

where $R(\beta)$ is the sum of revenue generated.

Proposition 5.6. *Given τ_i of all players, the optimal β that maximises the total revenue obtained by the network operator is on the boundary of the feasible range:*

$$\beta = \begin{cases} \beta^{\max} & \text{if } \sum_{i=1}^N \tau_i \Delta_i^{\text{AS}} \geq 0 \\ \beta^{\min} & \text{otherwise.} \end{cases} \quad (5.43)$$

where β^{\min} and β^{\max} are derived in Proposition 5.5 and Δ_i^{AS} is the maximum gap between the asymmetric Nash and Raiffa solutions in Proposition 5.4. The selection of β is arbitrary when $\sum_{i=1}^N \tau_i \Delta_i^{\text{AS}} = 0$ in (5.43).

Proof. The first-order derivative of the revenue function is

$$\frac{dR(\beta)}{d\beta} = \sum_{i=1}^N \tau_i \frac{d}{d\beta} \eta_i^{\text{AS}}(\beta) = \frac{N(N-1)}{(N-1+\beta)^2} \sum_{i=1}^N \tau_i \Delta_i^{\text{AS}}.$$

The first-order necessary condition for optimality is $\frac{dR(\beta)}{d\beta} = 0$, which there is no trivial solution. The second derivative of $R(\beta)$ is given as

$$\frac{d^2R(\beta)}{d\beta^2} = -\frac{2N(N-1)}{(N-1+\beta)^3} \sum_{i=1}^N \tau_i \Delta_i^{\text{AS}}.$$

It is then easy to establish that $R(\beta)$ is a monotonically increasing function of β if $\sum_{i=1}^N \tau_i \Delta_i^{\text{AS}} > 0$ or a monotonically decreasing function of β if $\sum_{i=1}^N \tau_i \Delta_i^{\text{AS}} < 0$. Therefore, the optimal β is on the boundary of the feasible range of β . \square

5.6 Resource Sharing Among Competing Operators

Up until this point, we have considered the resource allocation problem in networks shared by *one* network operator, M MVNOs and N users of the operator. Unlike the MVNOs, the network operator owns the network and plays the role of arbitrator in the cooperative resource allocation game. In this section, we present a model for cooperative resource sharing among several operators who co-own the network as shown in Fig. 5.2. Assume that L operators are sharing T amount of resources with sharing factor $\boldsymbol{\alpha}^{\text{SLA}} = (\alpha_1^{\text{SLA}}, \dots, \alpha_L^{\text{SLA}})$ and $\sum_{l=1}^L \alpha_l^{\text{SLA}} = T$. The sharing factor $\boldsymbol{\alpha}^{\text{SLA}}$ is pre-specified in the SLAs among the operators. However, these operators are likely to have different users' characteristics and arriving patterns and, depending on the time-of-day, might require fewer or more resources than the specified $\boldsymbol{\alpha}^{\text{SLA}}$. Because of these differences, temporary exchanges of resources will provide benefits to the operators in terms of better overall communication quality for their users.

Denote the current maximum resource demand of the l th operator, $l = 1, \dots, L$, as α_l^{max} . Depending on α_l^{max} , the operators will either play the role of *buyer* if they have a resource deficit; *seller* if they have a resource surplus; or *dummy* if their α_l^{SLA} is sufficient or they choose not to participate. The set of buyers, \mathcal{I}_B , and sellers, \mathcal{I}_S can then be defined as:

$$\mathcal{I}_B = \{l : \alpha_l^{\text{SLA}} < \alpha_l^{\text{max}}, l = 1, \dots, L\} \quad (5.44)$$

$$\mathcal{I}_S = \{l : \alpha_l^{\text{SLA}} > \alpha_l^{\text{max}}, l = 1, \dots, L\}. \quad (5.45)$$

The total surplus of the sellers is then given as $C = \sum_{l \in \mathcal{I}_S} \alpha_l^{\text{SLA}} - \alpha_l^{\text{max}}$. When there is more than one buyer and surplus C exceeds zero but not enough to cover all requests, i.e. $C < \sum_{l \in \mathcal{I}_B} \alpha_l^{\text{max}} - \alpha_l^{\text{SLA}}$, the buyers enter into a cooperative resource sharing game. Much of the success of this game depends on the cooperation and willingness of the operators to share unused resources. To avoid the *tragedy of the commons*, where the operators become a buyer when they need additional resources but refuse to become a seller when there is a surplus, we capture the operators' history of additional resources obtained and contributed in the current allocation problem.

Let α_l^t be the amount of additional resources allocated to the $l \in \mathcal{I}_B$ operator (as a buyer) or the amount shared by the $l \in \mathcal{I}_S$ operator (as a seller) at iteration t , the average

resource allocation in the past k iterations at iteration t is

$$\hat{\alpha}_l^t = \frac{1}{k}(\alpha_l^t + \sum_{m=1}^{k-1} \alpha_l^{t-m} F). \quad (5.46)$$

Note that $\alpha_l^t > 0$ if $l \in \mathcal{I}_B$, $\alpha_l^t < 0$ if $l \in \mathcal{I}_S$ and $\alpha_l^t = 0$ otherwise. Given $\sum_{l \in \mathcal{I}_B} \alpha_l^t \leq C$, it is easy to see that $\sum_{l \in \mathcal{I}_B} \hat{\alpha}_l^t \leq \frac{1}{k}(C + \sum_{l \in \mathcal{I}_B} \sum_{m=1}^{k-1} \alpha_l^{t-m})$. Therefore, the history-dependent cooperative resource sharing game at iteration t leads to the following problem (P4):

$$\begin{aligned} \max_{\hat{\alpha}^t} \quad & \prod_{l \in \mathcal{I}_B} \left(\hat{\alpha}_l^t + \frac{\beta}{N-1} \sum_{k \neq l, l, k \in \mathcal{I}_B} (\alpha_k^{\max} - \alpha_k^{\text{SLA}} - \hat{\alpha}_k^t) \right) \\ \text{s.t.} \quad & \hat{\alpha}_l^t \geq \hat{\alpha}_l^{\min, t}, \hat{\alpha}_l^t \leq \hat{\alpha}_l^{\max, t}, \sum_{l \in \mathcal{I}_B} \hat{\alpha}_l^t \leq \frac{1}{k}(C + \sum_{l \in \mathcal{I}_B} \sum_{m=1}^{k-1} \alpha_l^{t-m}), \end{aligned} \quad (5.47)$$

where $\hat{\alpha}_l^{\min, t} = \frac{1}{k} \sum_{m=1}^{k-1} \alpha_l^{t-m}$, $\hat{\alpha}_l^{\max, t} = \frac{1}{k}(\alpha_l^{\max} - \alpha_l^{\text{SLA}} + \sum_{m=1}^{k-1} \alpha_l^{t-m})$ and $\alpha_l^{\max} - \alpha_l^{\text{SLA}}$ is the maximum amount of additional resources required.

The amount of additional resources allocated to the $l \in \mathcal{I}_B$ operator depends on the amount obtained and contributed in the past k iterations. Consider allocation $\hat{\alpha}_1^t$ and $\hat{\alpha}_2^t$ with $1, 2 \in \mathcal{I}_B$ and $\hat{\alpha}_1^t = \hat{\alpha}_2^t$. When $\sum_{m=1}^{k-1} \alpha_1^{t-m} > \sum_{m=1}^{k-1} \alpha_2^{t-m}$, i.e. operator 1 has been allocated more and/or contributed less in the past, it is easy to see from (5.46) that operator 1 will be allocated less in the current iteration or $\alpha_1^t < \alpha_2^t$. Therefore, operators who always obtain additional resources from others need to contribute often to bring down their $\hat{\alpha}_l^t$ in order to continue benefiting from the resource sharing game. Similar to the symmetric bargaining game outlined in Section 5.5, the solution to this game is the same as in Proposition 5.1 with $\eta_i = \hat{\alpha}_i^t$, $\eta_i^{\min} = \hat{\alpha}_i^{\min, t}$, $\eta_i^{\max} = \hat{\alpha}_i^{\max, t}$ and $T = \frac{1}{k}(C + \sum_{l \in \mathcal{I}_B} \sum_{m=1}^{k-1} \alpha_l^{t-m})$. Using the outcome of the game, each operator can then distribute $T' = \alpha_l^{\text{SLA}} + \alpha_l^t$ amount of resources to its users and MVNOs using Propositions 5.1-5.6.

The main challenge of this resource sharing problem is to minimise the disclosure of commercially sensitive information among these competing operators. In our model, the only information that the operators need to exchange is their maximum resource requirement α_l^{\max} . This parameter can be computed independently by each operator using criteria that are important to them. For example, α_l^{\max} of the l th operator can be a function of the aggregate maximum resource requirement of its MVNOs and users or QoS parameters such as the probability of communication loss of existing users or the blocking probability of incoming users.

5.7 Numerical Analysis and Discussions

In order to achieve the bargaining outcome in Proposition 5.1, the operator requires knowledge of each player's η_i^{\min} and η_i^{\max} . This is not difficult to achieve in real implementation because users can select their acceptable transmission rate range at the beginning of a call or even change it during the call. Moreover, for each service, the operator can set up several classes with varying guaranteed quality for its users. Using (5.3) and (5.10), the operator can calculate the minimum and maximum load factor requirements of its users for both uplink and downlink, respectively. For the MVNOs, η_i^{\min} is specified in their SLA with the operator and η_i^{\max} is a function of the total maximum load factor requirements of the users supported by them. The MVNOs can in turn redistribute the resources allocated in a similar manner using (5.19).

The Nash and Raiffa bargaining solutions that we have derived satisfy different axioms and are both on the Pareto-optimal boundary. The Nash solution maximises the Nash product, i.e. the product of the gain of all players. The Raiffa solution also considers the size of the bargaining domain of each player, i.e. how much other players give up in addition to one's gain. To illustrate this, we consider the following simple game with $N = 2$. Player 1 is a user and Player 2 is an MVNO, which has two users with the same maximum load factor 0.5. Suppose that $T = 1$ and the minimum and maximum requirements of the players are $\eta^{\min} = (0.1, 0.2)$, $\eta^{\max} = (0.70, 1.00)$ respectively. We analyse two scenarios, i.e. $\hat{\tau} = (0.5, 0.5)$ and $\hat{\tau} = (0.7, 0.3)$. The latter corresponds to the case where Player 1 increases his/her bid to the operator.

The solutions are $\boldsymbol{\eta}(0) = (0.45, 0.55)$ and $\boldsymbol{\eta}(1) = (0.40, 0.60)$ for the symmetric case; and $\boldsymbol{\eta}^{\text{AS}}(0) = (0.59, 0.41)$ and $\boldsymbol{\eta}^{\text{AS}}(1) = (0.53, 0.47)$ for the asymmetric case. The geometrical interpretations of the Nash and Raiffa solutions for both cases are illustrated in Fig. 5.3. The solid line is the Pareto-optimal boundary, i.e. the outcomes that satisfy $\sum_{i=1}^N \eta_i = T$ and $\eta_i^{\min} \leq \eta_i \leq \eta_i^{\max}$ for all i . All solutions on this boundary are Pareto-optimal or efficient but not necessarily fair. The axiomatic bargaining theory used in this work characterises the fair solutions on this boundary via axioms A1-5. The area under the boundary, with $\eta_i \geq \eta_i^{\min} \forall i$, is the bargaining domain of the problem. The symmetric Nash solution is the tangent point of the hyperbola $\prod_{i=1}^N (\eta_i - \eta_i^{\min}) = \text{constant}$ and only takes into account each player's individual gain $\eta_i - \eta_i^{\min}$. The symmetric Raiffa solution is the intersection point between a line from the disagreement point, $(\eta_1^{\min}, \eta_2^{\min})$,

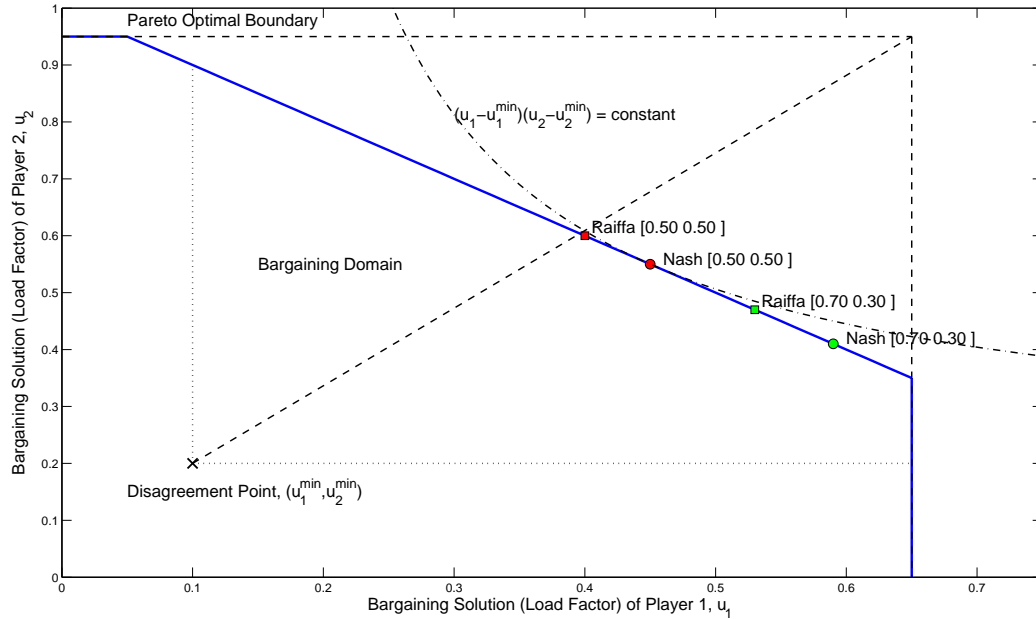


Figure 5.3. Geometrical interpretation of the symmetric and asymmetric Nash and Raiffa solutions.

to the maximum requirement point, $(\eta_1^{\max}, \eta_2^{\max})$, and the Pareto-optimal boundary. In other words, the Raiffa solution is the maximal point in the bargaining domain on that line. For the asymmetric case, Player 1 has a higher bargaining power and is therefore allocated more resources. Hence, by varying the parameters β and τ , all bargaining solutions on the Pareto-optimal boundary can be reached by the explicit solution derived in Proposition 5.3. Unlike our approach, the solutions derived using non-cooperative game theory in [3, 98, 99, 113, 126] are not Pareto-optimal.

The symmetric and asymmetric bargaining solutions are depicted in Fig. 5.4 in solid lines and dashed lines, respectively. From Propositions 5.1 and 5.3, as β approaches $-(N - 1)$, the bargaining solutions for Player 1 and 2 increases to $+\infty$ and decreases to $-\infty$, respectively. Therefore, we will only concentrate of solutions with $\beta \in [0, 1]$. As β increases from 0 to 1, the load factor allocated to Player 2, which has a higher maximum rate requirement, increases. Both solutions are Pareto-optimal. However, by the axiom of monotonicity A5, the Raiffa solution is at the point where each player’s gain is proportional to its maximum gain and therefore “fairer” to player 2. However, Player 1 is able to increase her load factor by submitting a higher bid, τ_1 , to the operator. In that case, Player 2’s bargaining power $\hat{\tau}_2 = \frac{\tau_2}{\tau_1 + \tau_2}$ decreases. The asymmetric Nash and

5.7 Numerical Analysis and Discussions

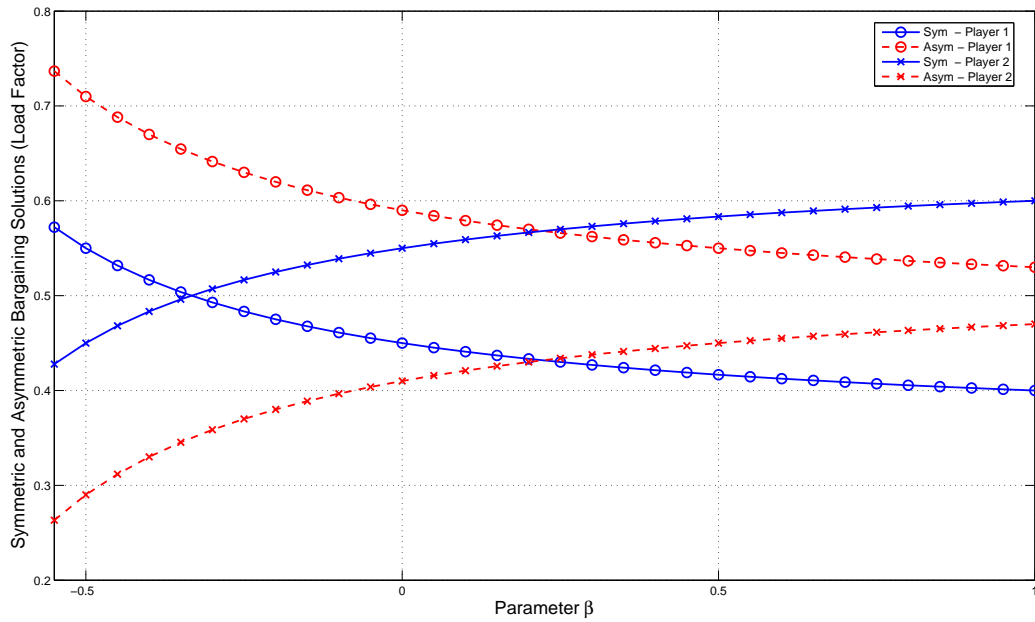


Figure 5.4. Symmetric and asymmetric bargaining solutions with $\hat{\tau} = (0.5, 0.5)$ and $\hat{\tau} = (0.70, 0.30)$ respectively.

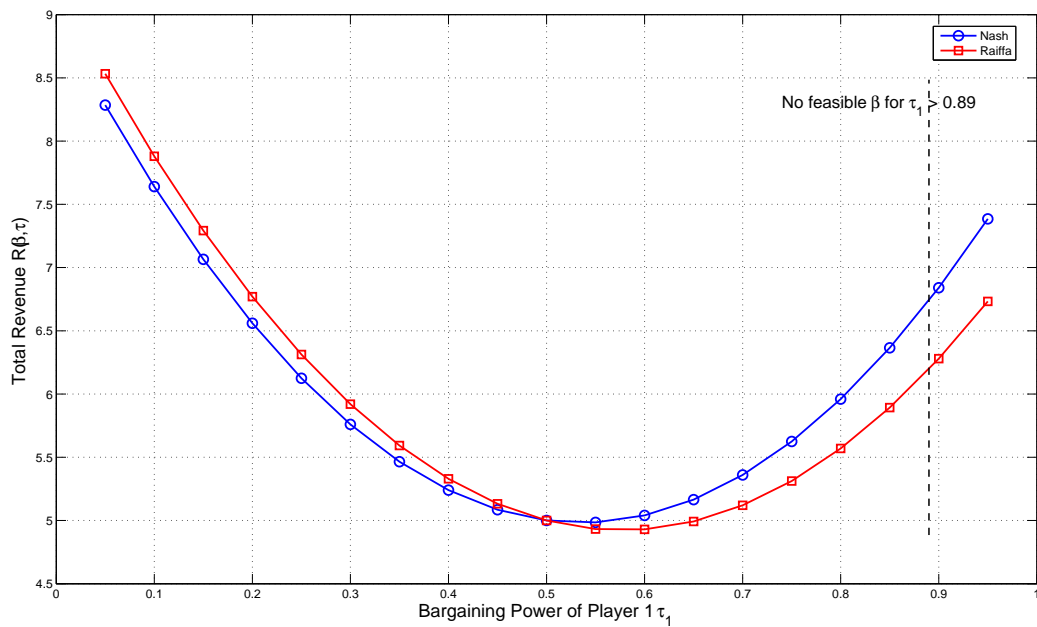


Figure 5.5. Operator's revenue using $\beta = 0, 1$ for varying bargaining powers.

Raiffa solutions are given by $\boldsymbol{\eta}^{\text{AS}}(0) = (0.59, 0.41)$ and $\boldsymbol{\eta}^{\text{AS}}(1) = (0.53, 0.47)$. Given $\boldsymbol{\eta}$, the uplink or downlink transmission rate of the players can then be determined using (5.6) and (5.13), respectively. In practice, players can submit their bids, τ_i , asynchronously or the operator can set up a number of price levels for them to select from.

Numerical analysis of Propositions 5.5 and 5.6 is depicted in Fig. 5.5. We plot the total revenue of the system, i.e. $R(\beta) = \sum_{i=1}^N \tau_i \tilde{\eta}_i(\beta, \boldsymbol{\tau})$, against various combinations of $\hat{\boldsymbol{\tau}} = (\hat{\tau}_1, \hat{\tau}_2)$ for $\beta = 0$ (Nash) and 1 (Raiffa). When the bids submitted by the players are symmetric, i.e. $\tau_1 = \tau_2$, the total revenue collected using the Nash solution coincides to that of Raiffa. In other words, the selection of $\beta \in [0, 1]$ is arbitrary since all solutions within that range are Pareto-optimal. However, when the bids are asymmetric, the total revenue is maximised by either using the Nash solution if $\sum_{i=1}^2 \tau_i \Delta_i^{\text{AS}} < 0$ or the Raiffa solution otherwise. All other solutions using $0 < \beta < 1$ are bounded by these two solutions. Also, we have established in Proposition 5.5 that a feasible range of β can be computed such that $\eta_i^{\text{min}} \leq \tilde{\eta}_i \leq \eta_i^{\text{max}}$ for $i = 1, 2$. For example, by Proposition 5.5, when $\hat{\boldsymbol{\tau}} = (0.30, 0.70)$, $\beta_{\text{min}} = 0$ and $\beta_{\text{max}} = 1.00$ and when $\hat{\boldsymbol{\tau}} = (0.85, 0.15)$, $\beta_{\text{min}} = 0.50$ and $\beta_{\text{max}} = 1.00$. In these cases, the total revenue is maximised when $\beta = 0$ (Nash) and $\beta = 1.00$ (Raiffa), respectively. When there is no feasible β within $[0, 1]$, e.g. $\hat{\tau}_1 > 0.89$ in Fig. 5.5, the arbitrator can signal player 1 to decrease its maximum bid according to (5.41) or player 2 to increase its minimum bid according to (5.40).

5.8 Conclusion

We have presented a framework for resource allocation in shared WCDMA networks using the notion of axiomatic bargaining from cooperative game theory. Although our model is of a shared network, the results can be applied to other networks with similar resource allocation problem. The resource allocation problem has been divided into two sub-problems: resource sharing among the operators and resource bargaining among the users and MVNOs of each operator. For the latter, we have derived the symmetric and asymmetric resource bargaining solutions. Unlike conventional schemes that only aim to optimise some system objectives such as throughput; or noncooperative, decentralised solutions that are inefficient and do not consider fairness, our bargaining outcome are Pareto-optimal and fair according to the minimum and maximum requirements of each

5.8 Conclusion

player. In the asymmetric model, the players, i.e. the users of the network operator and the MVNOs, can affect the bargaining outcome by submitting bids to the network arbitrator. The solutions derived are parameterised by β , which quantifies the preference for one's gain and the losses of others. When the solutions are all Pareto-optimal, the selection of β is arbitrary. However, in the asymmetric model, there is an opportunity for operators to maximise their revenue through the optimisation of β . In the resource sharing game, we suggest that the allocation of additional resources to operators who experience resource shortage should depend on their average of their past allocations and contributions.

Chapter 6

Conclusion

6.1 Summary

In this dissertation, we have addressed the problem of efficient management of scarce radio resources and congestion control in wireless networks by proposing an economic framework for dynamic pricing and RRM. The distinction between this research and existing works is illustrated in Fig. 6.1. By way of explanation and as discussed in earlier chapters, existing works focus on the following areas:

- A: static pricing schemes only;
- B: RRM schemes only;
- C: dynamic pricing schemes without RRM, i.e. pricing and RRM strategies are designed separately; and
- E: power control and pricing schemes using the shadow pricing or noncooperative game-theoretic approaches discussed in Chapters 1 and 2.

The pricing and RRM policies implemented in most, if not all, wireless networks today belong to Groups A and B. These policies make pricing and RRM decisions separately and fail to utilise price as an incentive for users to alter their demand and hence shape the network load. The bulk of the pricing literature for wired networks belongs to Group C. We have shown in Chapters 3 and 4 that the proposed integrated policy, labelled as group

6.1 Summary

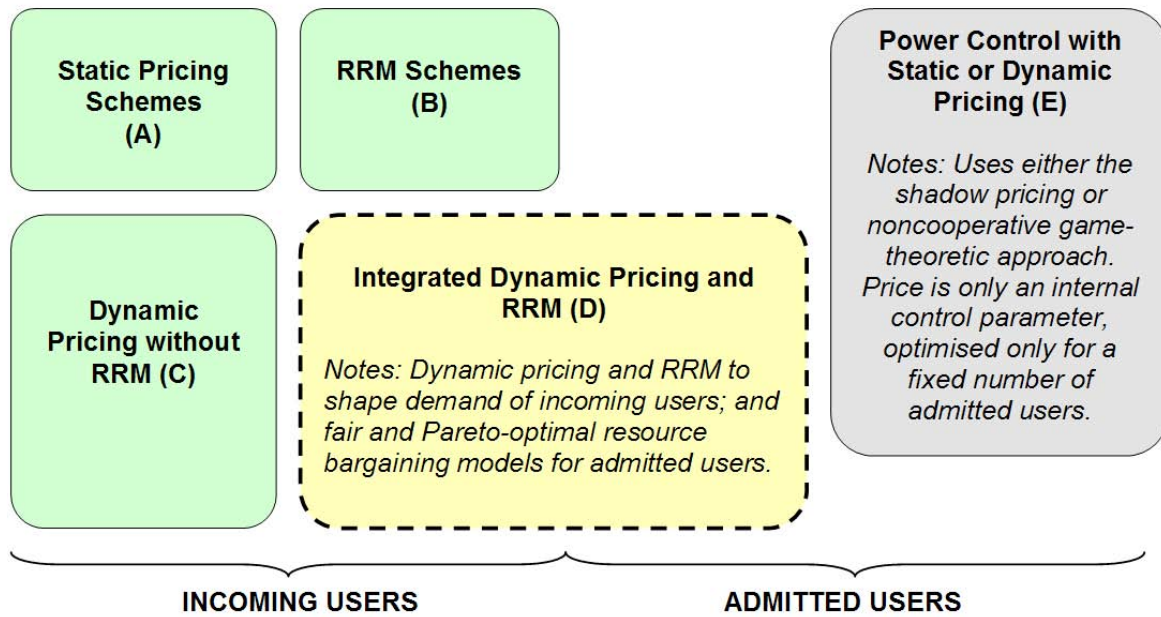


Figure 6.1. Summary of the contributions of this dissertation and key distinctions with existing works.

D in Fig. 6.1, outperforms policies in groups A, B and C in terms of revenue maximisation and congestion prevention. The works in group E are motivated by power control so that the price, which is optimised for a fixed number of users, only serves as an internal control parameter and does not reflect what users might end up paying. Furthermore, the approach using noncooperative game theory results in a non-Pareto-optimal outcome and pricing only provides some Pareto improvements. With those distinction in mind, we now summarise our key results.

In Chapter 3, we presented a stochastic approach to formulate an integrated dynamic pricing and admission control problem for a fixed-capacity, multiservice, cellular network as a dynamic programming problem. The objective is to maximise the long-term expected reward in monetary term. We associate the admission handoff calls with a satisfaction revenue such that the optimal policy places a higher priority on handoff calls. The integrated policy proposed outperforms conventional policies that consider admission control and pricing as separate problems. This policy has the flexibility to reject a new connection request when this is advantageous to the network even though there is sufficient bandwidth to accommodate the call. This strategy provides monetary incentives for low-WTP users to access the network when the load is relatively light and allocates resources

to high-WTP users when the network is relatively congested. The integrated policy is effective in congestion control because it displaces traffic from congested to less congested states. The blocking probabilities of the services decrease because occasions of high price per bandwidth time shift and even out the load, resulting in lower stationary probabilities of the network being in congested states. Finally, a non-discriminatory pricing scheme with a smaller price control space can be used to reduce computational effort. It has been shown to closely approximate the discriminatory pricing policy, which charges different price per unit resource for different services.

In Chapter 4, we extended the work proposed for fixed-capacity, cellular networks in Chapter 3 to interference-limited networks. We formulated an integrated dynamic pricing and resource management problem for WCDMA networks using simulation-based NDP. We proposed to use the Noise Rise Factor, i.e. the amount of interference generated by a call, as a basis for pricing. This parameter captures what existing load-based pricing models cannot, i.e. the nonlinear relationship between resource usage and the interference generated by a call as system loading increases. Our results show that the proposed optimal interference-based dynamic pricing and RRM policy increases average reward and reduces blocking and dropping probabilities compared to load-based schemes. Although the average reward obtained under the optimal integrated policy increases, the average system load factor and noise rise actually decrease. This means that the optimal policy can achieve more with less. Higher reward in terms of actual and satisfaction revenue can be generated without overloading the system and generating high interference. The gain in average reward is also due to the admission and allocation of resources to users with high WTP. By implementing a price sliding window that restricts the size of the price decision space, the sub-optimal policy closely approximates that of the optimal policy with full choice of prices. This idea is based on the observation that price is unlikely to increase or decrease by large amounts during simulation.

In Chapter 5, we analysed the problem of fair and efficient allocation of scarce radio resources to users who have been admitted to the network, based on the optimal integrated policies proposed in Chapters 3 and 4. We considered the case where there are other virtual network operators, i.e. MVNOs, who share the network with the main network operator. Similar to the users of the main operator, these MVNOs can be treated as users with minimum and maximum rate requirements. Using the notion of axiomatic bargaining from cooperative game theory, we derived a set of bargaining solutions that

6.2 Potential Directions for Further Research

are both fair and Pareto-optimal according to the minimum and maximum rate requirements of all users. Our approach is a departure from works using noncooperative game theory that result in inefficient outcomes, i.e. the Nash equilibria; or works using cooperative game theory that focus on one solution on the Pareto-optimal boundary. The main advantage of analysing a class of solutions instead of one is that the best outcome, based on some other objectives such as revenue maximisation, can be selected among these optimal solutions. The bargaining solutions are situated between the Nash and Raiffa-Kalai-Smorodinsky bargaining solutions on the Pareto-optimal boundary. These solutions vary according to a parameter that quantifies the tradeoff between one's gain and the losses of others. For example, the Nash bargaining solution is the point that maximises the Nash product, i.e. a product of the gain of all users. By contrast, the Raiffa-Kalai-Smorodinsky bargaining solution places equal importance on one's gain and the losses of others. When the axiom of symmetry is removed, users' bargaining outcome also depends on their admission price or the bid that they submit to the resource manager. An important contribution of the asymmetric bargaining model is that, unlike noncooperative auction models, the outcome is still Pareto-optimal and fair according to the users' bargaining power. Finally, we provided a model to encourage resource sharing among operators. When these operators have non-coincident peak demands, they will benefit from temporarily sharing their resources or borrowing from others. Although our model is of a shared WCDMA network, it can be easily adapted to networks with similar problems.

6.2 Potential Directions for Further Research

There are several directions in which our models can be extended.

6.2.1 Online Dynamic Pricing Schemes

The problems considered in Chapters 3 and 4 involve offline computation or approximation of the average reward (and parameter vector in the latter), assuming some knowledge of network parameters such as the arrival and departure rates of the users. The dynamic programming approach used does not naturally extend to an *online* setting, where the only model available for the Markov process is the process itself and the lack of robustness to

unknown or slowly time-varying parameters [17]. However, we have provided a model-free and easily extendable neuro-dynamic programming solution to the problem in Chapter 4. Once the approximation obtained offline via learning is satisfactory, it can be used to generate decisions fast enough for use in real time.

To enable online computation of the integrated dynamic pricing and RRM policy, arrival rates λ^n and λ^h can be estimated using parameter estimation techniques such as maximum likelihood estimation in order to detect changes in the arrival rates. The parameters estimated, $\hat{\lambda}^n$ and $\hat{\lambda}^h$, will then be fed to the feature extraction module to compute the feature vector. Fig. 6.2 shows the suggested approach for an online adaptation of this problem. The cost-to-go approximation will then continue to improve as the system operates in real time. In this work, we have only experimented with linear features in terms of future revenue rate of the system. The extraction of nonlinear features might provide better approximation and can be considered in future work.

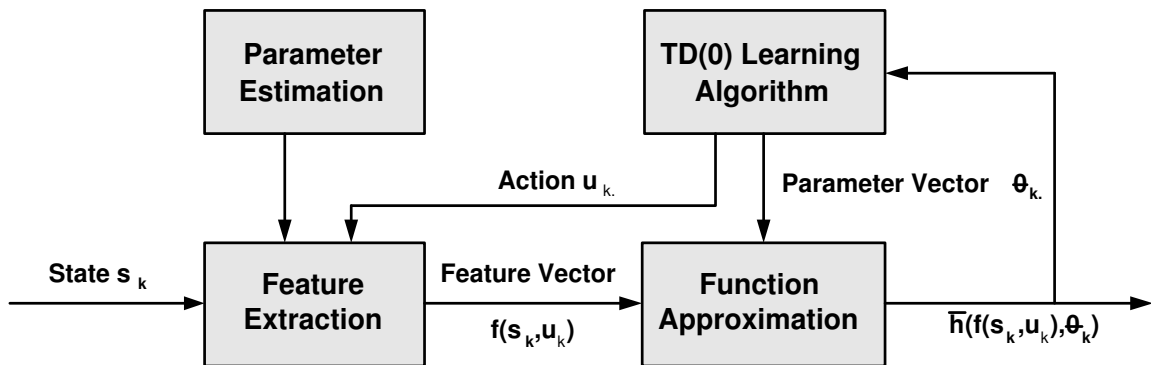


Figure 6.2. Modified feature-based approximation architecture with parameter estimation.

6.2.2 Stochastic Pricing and Resource Allocation Games

The dynamic programming and neuro-dynamic programming techniques used in Chapters 3 and 4 addressed the problem for a single decision maker acting in a stationary environment. In the asymmetric bargaining model introduced in Chapter 5, users' bargaining powers depend on their bids, which can be determined using the admission price from an optimal pricing policy or an auction. Like Kelly's paper on resource allocation [25], we have assumed that network users do not anticipate their effect on the price and the resource allocated. However, when the users recognise that they are not merely "price

6.2 Potential Directions for Further Research

takers”, the problem becomes a game in which the setting of WTP, demand and bids becomes strategic for the network users. Users make self-serving decisions and economists are well aware that these selfish behaviours can lead to inefficiency. Johari and Tsitsiklis showed that the price of anarchy in networks with elastic supply amounts to up to 25% in efficiency loss, which they measured by computing the ratio of the Nash Equilibrium utility function to the socially optimal utility function and showing that it is $3/4$ at worst [57, 58]. Stochastic games are natural extensions of Markov decision processes to include multiple decision makers. An excellent introduction to stochastic game theory for environments with multiple reinforcement learners is [14].

6.2.3 Noncooperative Implementation of the Cooperative Bargaining Solutions in Self-organising Networks

In Chapter 5, we have derived a class of bargaining solutions using the axiomatic bargaining concepts from cooperative game theory. The payoffs, indicated by the amount of resources allocated to the players, are sustained by a binding agreement guaranteed by an outside enforcer or arbitrator, i.e the network operator in our case. However, in networks without the presence of an arbitrator, the enforcement of such payoffs falls outside of the domain of cooperative game theory. In such networks, the decision-making process has to be decentralised, as in a noncooperative game. Examples of such self-organising architecture are Mobile Ad hoc Networks (MANETs) and Wireless Mesh Networks (WMNs) [2].

A MANET is a collection of nodes which forms a network independent of any fixed infrastructure. As opposed to networks which use routers to support network functions, such as packet routing and forwarding, these functions are provided by the nodes (or hosts) themselves. Such a network can operate in a stand-alone fashion or may be connected to the Internet. The interconnections among nodes often change continually and arbitrarily. These networks were initially designed for military operations and play an increasingly important role in many environments, such as ad hoc networking for collaborative and distributed disaster recovery, search-and-rescue and crowd control. More recently, they have been envisaged as able to provide Internet connectivity for nodes that are not in transmission range of a wireless access point. The IEEE 802.11 wireless protocol incorporates an ad hoc networking system when no access points are present.

In WMNs, all routers are capable of organising and auto-reconfiguring themselves wirelessly, meaning that no cabling is needed to connect them. These routers, or nodes, form a rich radio mesh connectivity among themselves that is difficult to provision in wired networks. The principle is similar to the way packets travel around the wired Internet – data will hop from one node to another until it reaches its given destination. While wireless node connectivity significantly reduces the up-front deployment and subsequent maintenance costs, the rich mesh connectivity helps to deliver high levels of reliability and robustness. Mesh networks are self-healing and extremely reliable because each node is connected to several others and if one drops out, due to hardware failure or any other reason, its neighbours simply find another route. Because of these attractive features, WMN is being considered for a wide variety of applications such as backhaul connectivity for cellular radio access networks, defence systems, city-wide surveillance systems and real-time racing car telemetry. It can effectively extend a network by sharing access to higher cost network infrastructure.

Due to the complexity of the mobility and traffic models as well as the infrastructure-less, dynamic topology of these networks, noncooperative game theory is the primary tool for studying individual, independent decision makers whose actions potentially result in efficiency loss. The vast number of works on the application noncooperative game theory in MANETs are surveyed in [112]. Another approach is to study the implementation of the Pareto-optimal bargaining solutions in a noncooperative manner. The distinction between cooperative and noncooperative games is not new. The *Nash Program*, initiated by Nash [78], is an attempt to bridge the gap between these two branches of game theory. This is accomplished by investigating noncooperative procedures that yield cooperative solutions as their equilibrium outcomes. A result in the Nash Program is referred to as the “noncooperative foundation” or “noncooperative implementation” of a cooperative solution [102]. Nash first expressed the bargaining model as a noncooperative game, called the *Nash demand game*, in [78]. Both bargainers demand a utility level simultaneously. If the vector of demands is feasible, it will be implemented. Otherwise, the disagreement point will be enforced, if there is one, or, if not, the players will receive nothing. Another example is Moulin’s [73] implementation of two-person bargaining games. We refer our readers to [64, 100, 102, 119] and the references therein for further information on the noncooperative solutions to the bargaining problem and note their connection with this work.

Appendix A

Appendix

A.1 Infinitesimal Generator

Forward Transitions, $\mathbf{Q}_0^{(x_1)}$ Forward transitions, i.e. $x_1 \rightarrow x_1 + 1$ occur when a new user of service 1 reattempts or a user from service 2 substitutes in. $\mathbf{Q}_0^{(x_1)}$ can be broken into:

$$\mathbf{Q}_0^{(x_1)} = \begin{pmatrix} \mathbf{Q}_{01}^{(x_1,0)} & \mathbf{0} & \mathbf{0} \\ \mathbf{Q}_{02}^{(x_1,1)} & \mathbf{Q}_{01}^{(x_1,1)} & \mathbf{0} \\ & \ddots & \ddots & \ddots \\ & \mathbf{0} & \mathbf{Q}_{02}^{(x_1,X_2)} & \mathbf{Q}_{01}^{(x_1,X_2)} \end{pmatrix} \quad (\text{A.1})$$

where $\mathbf{Q}_{01}^{(x_1,x_2)} = \mathbf{Q}_{01}^{(x_1)} = \lambda_{51}\mathbf{I}$ and $\mathbf{Q}_{02}^{(x_1,x_2)} = \lambda_{42}(x_2)\mathbf{I}$ are $M \times M$ matrices, M is the number of possible allocation of B channels to two services, and \mathbf{I} denotes identity matrix of appropriate size.

Backward Transitions, $\mathbf{Q}_2^{(x_1)}$ Backward transitions, i.e. $x_1 \rightarrow x_1 - 1$ occur when a user in orbit successfully makes a call connection, abandons the network or substitutes out to service 2. $\mathbf{Q}_2^{(x_1)}$ can be broken into:

$$\mathbf{Q}_2^{(x_1)} = \begin{pmatrix} \mathbf{Q}_{21}^{(x_1,0)} & \mathbf{Q}_{20}^{(x_1,0)} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{21}^{(x_1,1)} & \mathbf{Q}_{20}^{(x_1,1)} \\ & \ddots & \ddots & \ddots \\ & \mathbf{0} & \mathbf{0} & \mathbf{Q}_{21}^{(x_1,X_2)} \end{pmatrix} \quad (\text{A.2})$$

A.1 Infinitesimal Generator

where $\mathbf{Q}_{21}^{(x_1, x_2)}$ and $\mathbf{Q}_{22}^{(x_1, x_2)}$ are $M \times M$ matrices defined by

$$\mathbf{Q}_{21}^{(x_1, x_2)} = \begin{pmatrix} \lambda_{31}(x_1)\mathbf{I} & \lambda_{21}(x_1)\mathbf{I} & \mathbf{0} & & \\ \mathbf{0} & \lambda_{31}(x_1)\mathbf{I} & \lambda_{21}(x_1)\mathbf{I} & & \\ & \ddots & \ddots & \ddots & \\ & \mathbf{0} & \mathbf{0} & \lambda_{31}(x_1)\mathbf{I} & \end{pmatrix} \quad (\text{A.3})$$

and

$$\mathbf{Q}_{22}^{(x_1, x_2)} = \lambda_{41}(x_1)\mathbf{I}. \quad (\text{A.4})$$

Local Transitions, $\mathbf{Q}_1^{(x_1)}$ Local transitions, i.e. $x_2 \rightarrow x_2 \pm 1$, $n_1 \rightarrow n_1 \pm 1$ and $n_2 \rightarrow n_2 \pm 1$, include all other transitions. $\mathbf{Q}_1^{(x_1)}$ can be broken into:

$$\mathbf{Q}_1^{(x_1)} = \begin{pmatrix} \mathbf{Q}_{11}^{(x_1, 0)} & \mathbf{Q}_{10}^{(x_1, 0)} & \mathbf{0} & & \\ \mathbf{Q}_{12}^{(x_1, 1)} & \mathbf{Q}_{11}^{(x_1, 1)} & \mathbf{Q}_{10}^{(x_1, 1)} & & \\ & \ddots & \ddots & \ddots & \\ & \mathbf{0} & \mathbf{Q}_{12}^{(x_1, X_2)} & \mathbf{Q}_{11}^{(x_1, X_2)} & \end{pmatrix} \quad (\text{A.5})$$

where $\mathbf{Q}_{10}^{(x_1, x_2)}$, $\mathbf{Q}_{11}^{(x_1, x_2)}$ and $\mathbf{Q}_{12}^{(x_1, x_2)}$ are $M \times M$ matrices defined by

$$\mathbf{Q}_{10}^{(x_1, x_2)} = \lambda_{52}\mathbf{I} \quad (\text{A.6})$$

$$\mathbf{Q}_{11}^{(x_1, x_2)} = \begin{pmatrix} \mathbf{A} & (\lambda_{01} + \lambda_{11})\mathbf{I} & \mathbf{0} & & \\ (\mu_1 + \gamma_1)\mathbf{I} & \mathbf{A} & (\lambda_{01} + \lambda_{11})\mathbf{I} & & \\ & \ddots & \ddots & \ddots & \\ & & N_1(\mu_1 + \gamma_1)\mathbf{I} & \mathbf{A} & \end{pmatrix} \quad (\text{A.7})$$

$$\mathbf{Q}_{12}^{(x_1, x_2)} = \begin{pmatrix} \mathbf{B} & \mathbf{0} & \mathbf{0} & & \\ \mathbf{0} & \mathbf{B} & \mathbf{0} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathbf{0} & \mathbf{B} & \end{pmatrix} \quad (\text{A.8})$$

where $\mathbf{A} = \phi\mathbf{I} + n_2(\mu_2 + \gamma_2)\mathbf{I}_D + (\lambda_{02} + \lambda_{12})\mathbf{I}_U$, $\mathbf{B} = \lambda_{32}(x_2)\mathbf{I} + \lambda_{22}(x_2)\mathbf{I}_U$, the diagonal elements, ϕ , are such that the row sums of \mathbf{Q} are zero and

$$\mathbf{I}_U = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{I}_D = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & 0 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{A.9})$$

\mathbf{I}_U and \mathbf{I}_D are matrices with ones on the upper and lower diagonal respectively and zeroes everywhere.

Bibliography

- [1] Mohamed Hossam Ahmed. Call admission control in wireless networks: a comprehensive survey. *IEEE Communications Surveys* (available at www.comsog.org/pubs/surveys), 7(1):50–69, 2005.
- [2] Ian F. Akyildil, Xudong Wang, and Weilin Wang. Wireless mesh networks: A survey. *Computer Networks*, 47:445–487, January 2005.
- [3] Tansu Alpcan, Tamer Basar, R. Srikant, and Eitan Altman. CDMA uplink power control as a noncooperative game. *Wireless Networks*, 8:659–670, 2002.
- [4] E. Altman, T. Boulogne, R. El-Azouzi, T. Jimenez, and L. Wynter. A survey on networking games in telecommunications. *Computers and Operations Research*, 2004.
- [5] E. Altman, T. Jimenez, and G. Koole. On optimal call admission control in a resource sharing system. *IEEE Trans. on Commun.*, 49(9):1659–1668, September 2001.
- [6] L. Anania and R. J. Solomon. Flat: The minimalist price. In L. W. McKnight and J. P. Bailey, editors, *Internet Economics*. MIT Press, Cambridge, Massachusetts, 1997.
- [7] F. Ashtiani, J.A. Salehi, and M.R. Aref. A flexible dynamic traffic model for reverse link CDMA cellular networks. *IEEE Trans. on Wireless Commun.*, 3(1):159–64, January 2004.
- [8] D. Ayyagari. Optimal admission control in cellular DS-CDMA systems with multimedia traffic. *IEEE Trans. on Wireless Commun.*, 2(1):195–202, January 2003.

Bibliography

- [9] C. Beckman and G. Smith. Shared networks: Making wireless communication affordable. *IEEE Wireless Commun.*, pages 78–85, April 2005.
- [10] R. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, 1957.
- [11] R. Berezdivin, R. Brenig, and R. Topp. Next-generation wireless communications concepts and technologies. *IEEE Commun. Magazine*, 40(3):108–116, March 2002.
- [12] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1 and 2. Athena Scientific, Massachusetts, 1995.
- [13] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Massachusetts., 1996.
- [14] Michael Bowling and Manuela Veloso. An analysis of stochastic game theory for multiagent reinforcement learning. Cmu-cs-00-165, School of Computer Science, Carnegie Mellon University, Pittsburgh, Oct. 2000.
- [15] L. Bright. *Matrix Analytic Methods in Applied Probability*. PhD thesis, The University of Adelaide, 5005, South Australia, March 1996.
- [16] L. Bright and P. G. Taylor. Calculating the equilibrium distribution in level dependent quasi-birth-death processes. *Stochastic Models*, 11(3):497–525, 1995.
- [17] Enrique Campos-Nanez and Stephen D. Patek. On-line tuning of prices for network services. In *Proceedings of IEEE INFOCOM*, 2003.
- [18] Xiren Cao. Preference functions and bargaining solutions. In *Proc. CDC-21, Orlando, FL.*, Dec. 1982.
- [19] Xiren Cao, H. X. Shen, R. Milito, and P. Wirth. Internet pricing with a game theoretical approach: Concepts and examples. *IEEE/ACM Trans. on Networking*, 10(2):208–216, April 2002.
- [20] J. Choi, T. Kwon, Y. Choi, and M. Naghshineh. Call admission control for multimedia services in mobile cellular networks: a markov decision approach. In ***IEEE International Conference on Personal Wireless Communication*, pages 486–490, Feb 2000.

- [21] C.N. Chuah and R.H. Katz. Characterizing packet audio streams from internet multimedia applications. In *IEEE International Conference on Communications*, pages 1199–1203, April 2002.
- [22] R. Cocchi, D. Estrin, S. Shenker, and Lixia Z. A study of priority pricing in multiple service class networks. In *Proceedings of Sigcomm'91*. Oxford University Press, Sept. 1991.
- [23] R. Cocchi, S. Shenker, D. Estrin, and Lixia Z. Pricing in computer networks: Motivation, formulation, and example. *IEEE/ACM Trans. on Networking*, 1:614–627, Dec 1993.
- [24] C. Courcoubetis, F.P. Kelly, V.A. Siris, and R. Weber. A study of simple usage-based charging schemes for broadband networks. *Telecommunication Systems*, pages 323–343, Mar-Apr 2000.
- [25] C. Courcoubetis, F.P. Kelly, and R. Weber. Measurement-based usage charges in communication networks. *Technical Reports, Statistical Laboratory, University of Cambridge*, 1997.
- [26] Costas Courcoubetis and Richard Weber. *Pricing Communication Networks: Economics, Technology and Modelling*. Series in Systems and Optimization. Wiley Interscience, 2003.
- [27] L.A. DaSilva. Pricing for QoS-enabled networks: A survey. *IEEE Communications Surveys*, pages 2–8, Second Quarter 2000.
- [28] Z. Dziong and L. G. Mason. Fair-efficient call admission control policies for broadband networks – a game theoretic framework. *IEEE/ACM Trans. on Networking*, 4(1):123–136, February 1996.
- [29] Ericsson. Shared networks. Ericsson white paper, available at <http://www.ericsson.com>, May 2005.
- [30] Zoheir Ezziane. Charging and pricing challenges for 3G systems. *IEEE Communications Surveys (available at www.comsog.org/pubs/surveys)*, 7(4):58–68, 2005.
- [31] G. Falin. A survey of retrial queues. *Queueing Systems*, 7:127–168, 1990.

Bibliography

- [32] D. Famolari, N. Mandayam, Goodman D., and Shah V. *Wireless Multimedia Network Technologies*, chapter A new framework for power control in wireless data networks: Games, utility and pricing, pages 289–310. Kluwer Norwell, MA., 1999.
- [33] Y. Fang and Y. Zhang. Call admission control schemes and performance analysis in wireless mobile networks. *IEEE Trans. on Vehicular Technology*, 51(2):371–382, March 2002.
- [34] H. Fattah and C. Leung. An overview of scheduling algorithms in wireless multimedia networks. *IEEE Wireless Commun.*, 9(5):76–83, Oct 2002.
- [35] A. R. Fattahi and P. Fernando. New economic perspectives for resource allocation in wireless networks. In *American Control Conference*, pages 3960–65, June 2005.
- [36] E.D. Fitkov-Norris and A. Khanifar. Dynamic pricing in mobile communications systems. In *3G Mobile Communication Technologies, IEE*, volume 1, pages 416–420, 2000.
- [37] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, Cambridge, Mass., 1991.
- [38] R. Gibbens and F. Kelly. Resource pricing and the evolution of congestion control. *Automatica*, 35:1969–1985, 1999.
- [39] D.J. Goodman and N.B. Mandayam. Power control for wireless data. *IEEE Personal Commun.*, 7:48–54, April 2000.
- [40] Zhu Han, Zhu Ji, and K. J. Ray Liu. Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions. *IEEE Trans. on Commun.*, 53(8):1366–1376, August 2005.
- [41] S. Hart. Classical cooperative theory I and II. In S. Hart and A. Mas-Colell, editors, *Cooperation: Game-Theoretic Approach*, NATO ASI, Series F. Springer-Verlag, 1997.
- [42] Nidhi Hedge and Eitan Altman. Capacity of multiservice WCDMA networks with variable GoS. In *IEEE Wireless Communications and Networking Conference 2003*, volume 2, pages 1402–1407, 16-20 March 2003.

- [43] Tina M. Heikkinen. On congestion pricing in a wireless network. *Wireless Networks*, 8:347–354, 2002.
- [44] Siew-Lee Hew and Langford B. White. Interference-based dynamic pricing and radio resource management for WCDMA networks. In *61st IEEE Vehicular Technology Conference (VTC2005-Spring)*, Stockholm, Sweden, May 30 - June 1 2005.
- [45] Siew-Lee Hew and Langford B. White. Optimal integrated call admission control and dynamic pricing with handoffs and price-affected arrivals. In *2005 Asia-Pacific Conference on Communications (APCC'05)*, Perth, Australia, Oct. 3-5 2005.
- [46] Siew-Lee Hew and Langford B. White. Fair resource bargaining solutions for cooperative multi-operator networks. In *IEEE 2006 International Zurich Seminar on Communications (IZS'06)*, Zurich, Switzerland, Feb. 22-24 2006.
- [47] Siew-Lee Hew and Langford B. White. A game theoretical approach for resource bargaining in shared WCDMA networks: Symmetric and asymmetric models. In *First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications (AusWireless'06)*, Sydney, Australia, Mar. 13-16 2006.
- [48] K. Hiltunen and R. De Bernardi. WCDMA downlink capacity estimation. In *Tbd*, volume 2, pages 992–996, 2000.
- [49] Harri Holma and Antti Toskala, editors. *WCDMA for UMTS*. John Wiley and Sons, Ltd., 2001.
- [50] J. Hou, J. Yang, and S. Papavassiliou. Integration of pricing with call admission control to meet qos requirements in cellular networks. *IEEE Trans. on Parallel and Distributed Systems*, 13(9):898–909, September 2002.
- [51] V. Huang and W. Zhuang. Qos-oriented access control for 4G mobile multimedia CDMA communications. *IEEE Commun. Magazine*, 40(3):118–131, March 2002.
- [52] J. Hultell, K. Johansson, and J. Markendahl. Business models and resource management for shared wireless networks. In *Proceedings IEEE VTC-Fall*, Sept 2004.
- [53] Yoshihiro Ishikawa and Narumi Umeda. Capacity design and performance of call admission control in cellular CDMA systems. *IEEE J. Sel. Areas Commun.*, 15(8):1627–1635, Oct 1997.

Bibliography

- [54] Wha Sook Jeon and Dong Geun Jeong. Call admission control for CDMA mobile communications systems supporting multimedia services. *IEEE Trans. on Wireless Commun.*, 1(4):649–659, Oct 2002.
- [55] K. Johansson, M. Kristensson, and U. Schwarz. Radio resource management in roaming based multi-operator WCDMA networks. In *Proceedings IEEE VTC-Spring*, June 2004.
- [56] R. Johari and J. N. Tsitsiklis. Network resource allocation and a congestion game. submitted for publication, May 2003.
- [57] R. Johari and J. N. Tsitsiklis. Efficiency loss in a network resource allocation game. *Math Oper. Res.*, 29(3):407–435, 2004.
- [58] R. Johari and J. N. Tsitsiklis. Efficiency loss in a network resource allocation game: the case of elastic supply. *IEEE Transactions on Automatic Control*, 50(11):1712–1724, 2005.
- [59] E. Kalai and M. Smorodinsky. Other solutions to Nash’s bargaining problem. *Econometrica*, 43:513–518, 1975.
- [60] F. Kelly. Charging and accounting for bursty connections. In L. W. McKnight and J. P. Bailey, editors, *Internet Economics*. MIT Press, Cambridge, Massachusetts, 1997.
- [61] F. Kelly. Charging and rate control for elastic traffic. *European Trans. Telecommunications*, 8(1):33–37, 1997.
- [62] F.P. Kelly. *Notes on Effective Bandwidths*, volume 4 of *Royal Statistical Society Lecture Notes Series*, pages 141–168. Oxford University Press, Oxford, June 1996.
- [63] Neil J. Keon and G. Anandalingam. Optimal pricing for multiple services in telecommunications networks offering quality-of-service guarantees. *IEEE/ACM Trans. on Networking*, 11(1):66–80, Feb 2003.
- [64] Klaus Kultti and Hannu Vartiainen. A Noncooperative Solution to the Bargaining Problem. Discussion papers 592, Department of Economics, University of Helsinki, 2004.

- [65] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modelling*. ASA-SIAM Series on Statistics and Applied Probability, 1999.
- [66] Jang Won Lee, Ravi Mazumdar, and Ness Shroff. Downlink power allocation for multi-class CDMA wireless networks. *IEEE/ACM Trans. on Networking*, 13(4):854–867, Aug 2005.
- [67] P.J. Liu, P.F. Zhang, S. Jordan, and M. L. Honig. Single-cell forward link power allocation using pricing in wireless networks. *IEEE/ACM Trans. on Wireless Commun.*, 3(2):533–543, March 2004.
- [68] J. Mackie-Mason and H. Varian. Pricing congestible network resources. *IEEE J. Sel. Areas Commun.*, 13:1141–48, Sept 1995.
- [69] P. Marbach, O. Mihatsch, and J. N. Tsitsiklis. Call admission control and routing in integrated service networks using neuro-dynamic programming. *IEEE J. Sel. Areas Commun.*, 18(2):197–208, February 2000.
- [70] Peter Marbach and Berry Randall. Downlink resource allocation and pricing for wireless networks. In *Proceedings of IEEE INFOCOM*, 2002.
- [71] Farhad Meshkati, M. Chiang, V. Poor, and S. Schwartz. A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems. *to appear in IEEE Journal on Selected Areas in Commun. (JSAC)*, 2005.
- [72] Christian Montet and Daniel Serra. *Game Theory and Economics*. Palgrave Macmillan, New York, 2003.
- [73] H. Moulin. Implementing the Kalai-Smorodinsky bargaining solution. *Journal of Economic Theory*, 33:32–45, 1984.
- [74] Abhinay Muthoo. *Bargaining Theory with Applications*. Cambridge University Press, Cambridge, 1999.
- [75] M. Naghshineh and S. Schwartz. Distributed call admission control in mobile/wireless networks. *IEEE J. Sel. Areas Commun.*, 14(4):711–717, May 1996.
- [76] J. Nash. The bargaining problem. *Econometrica*, 18, 1950.

Bibliography

- [77] J. Nash. Equilibrium points in N-person games. In *Proc. National Academy Science*, 1950.
- [78] J. Nash. Two-person cooperative games. *Econometrica*, 21, 1953.
- [79] M. F. Neuts. *Matrix Geometric Solutions in Stochastic Models*. Johns Hopkins, 1981.
- [80] M. F. Neuts and B. M. Rao. Numerical investigation of a multiserver retrial model. *Queueing Systems*, 7:169–190, 1990.
- [81] Nokia. Network sharing in 3G. Nokia white paper, available at <http://www.nokia.com>, May 2005.
- [82] D. M. Norakovic and M. L. Dukic. Evolution of the power control techniques for DS-CDMA toward 3G wireless communication systems. *IEEE Communications Surveys (available at www.comsog.org/pubs/surveys)*, 2:2–15, 2000.
- [83] A. M. Odlyzko. A modest proposal for preventing internet congestion. available from URL <http://www.research.att.com/~amo>, Sept 1997.
- [84] Seong-Jun Oh, Danlu Zhang, and Kimberly M. Wasserman. Optimal resource allocation in multiservice CDMA networks. *IEEE Trans. on Wireless Commun.*, 2(4):811–821, July 2003.
- [85] Martin Osborne. *An introduction to Game Theory*. Oxford University Press, New York, 2004.
- [86] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw Hill, 4th edition, 2002.
- [87] C. Parris and D. Ferrari. A resource based pricing policy for real-time channels in a packet-switching network. *Technical Report TR-92-018, International Computer Science Institute, Berkeley, CA*, Mar 1992.
- [88] C. Parris, S. Keshav, and D. Ferrari. A framework for the study of pricing in integrated networks. *Technical Report TR-92-016, International Computer Science Institute, Berkeley, CA*, Mar 1992.

- [89] I.C. Paschalidis and Y. Liu. Pricing in multiservice loss networks: Static pricing, asymptotic optimality, and demand substitution effects. *IEEE/ACM Trans. on Networking*, 10:425–437, June 2002.
- [90] I.C. Paschalidis and J.N. Tsitsiklis. Congestion-dependent pricing of network services. *IEEE/ACM Trans. on Networking*, 8(2):171–184, Apr 2000.
- [91] S.D. Patek and E. Campos-Nanez. Pricing of dialup services: an example of congestion-dependent pricing in the Internet. In *Proceedings of the 39th IEEE Conference on Decision*, volume 3, pages 2296–2301, 2000.
- [92] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Mathematical Science. John Wiley & Sons, Inc., 1994.
- [93] H. Raiffa. Arbitration schemes for generalised two-person games. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Game II*. Princeton University Press, Princeton, NJ., 1953.
- [94] Theodore S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, second edition, 2002.
- [95] Alvin Roth. *Axiomatic models of bargaining*. Springer-Verlag, Berlin, New York, 1979.
- [96] Alvin Roth. *Game-theoretic models of bargaining*. Cambridge University Press, Cambridge, New York, 1985.
- [97] Jukka Ruusunen. On barter contracts in electricity exchange. In Tamar Basar and Alain Haurie, editors, *Annals of the International Society of Dynamic Games, Advances in Dynamic Games and Applications*, volume 1. Birkhauser, 1994.
- [98] C.U. Saraydar, N.B. Mandayam, and D.J. Goodman. Pricing and power control in a multicell wireless data network. *IEEE J. Sel. Areas Commun.*, 19(10):1883–1892, Oct 2001.
- [99] C.U. Saraydar, N.B. Mandayam, and D.J. Goodman. Efficient power control via pricing in wireless data networks. *IEEE Trans. on Commun.*, 50(2):291–303, Feb 2002.

Bibliography

- [100] Roberto Serrano. Fifty years of the Nash Program. Technical report, available at www.econ.brown.edu/faculty/serrano, Brown University, U.S.A., 2004.
- [101] Roberto Serrano. *The New Palgrave Dictionary of Economics*, chapter Bargaining. McMillan, London, 2nd edition, 2005.
- [102] Roberto Serrano. *The New Palgrave Dictionary of Economics*, chapter Nash Program. McMillan, London, 2nd edition, 2005.
- [103] L. S. Shapley. A value for n-person games. In H. Kuhn and A. Tucker, editors, *Contributions to the Theory of Games II (Annals of Mathematics Studies 28)*. Princeton University Press, Princeton, 1953.
- [104] S. Shenker, D. Clark, D. Estrin, and S. Herzog. Pricing in computer networks: Reshaping the research agenda. *ACM Computer Communication Review*, pages 19–43, 1996.
- [105] Siemens. 3G Infrastructure Sharing – The Siemens Perspective. Siemens white paper, available at <http://www.siemens.com>, May 2005.
- [106] A.P. Silva and G.R. Mateus. Performance analysis for data service in third generation mobile telecommunication networks. In *35th Annual Simulation Symposium*, pages 227–234, April 2002.
- [107] S. Singh, V. Krishnamurthy, and Vincent H. Poor. Integrated voice/data call admission control for wireless DS-CDMA systems. *IEEE Trans. on Wireless Commun.*, 50(6):1483–95, June 2002.
- [108] V.A. Siris, B. Briscoe, and D. Songhurst. Economic models for resource control in wireless networks. In *13th IEEE International Personal Indoor and Mobile Radio Comms, Proceedings of*, volume 3, pages 1112–1116, 2002.
- [109] Vasilios A. Siris. Congestion pricing for resource control in wideband CDMA. Technical report, ICS-Forth, Dec 2001.
- [110] A.H. Solana, A.V. Bardaji, and F.C. Palacio. Capacity analysis and performance evaluation of call admission control for multimedia packet transmission in UMTS WCDMA system. In *Tbd*, volume 2, pages 1550–55, 2003.

- [111] M. Sorousnejad and E. Geraniotis. Multi-access strategies for an integrated voice/data CDMA packet radio network. *IEEE Trans. on Commun.*, 43(2/3/4):934–45, 1995.
- [112] V. Srivastava, J. Neel, A. MacKenzie, R. Menon, J.E. DaSilva, L.A. Hicks, J.H. Reed, and R.P. Gilles. Using game theory to analyze wireless ad hoc networks. *IEEE Communications Surveys*, 7(4), Fourth Quarter 2005.
- [113] Chi Wan Sung and Wing Shing Wong. Noncooperative power control game for multirate CDMA data networks. *IEEE Trans. on Commun.*, 2(1):186–194, January 2003.
- [114] R.S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44, 1988.
- [115] R.S. Sutton and A.G. Barto. *Reinforcement Learning*. The MIT Press, 1998.
- [116] John B. Taylor and Lionel Frost. *Microeconomics*. John Wiley and Sons, Australia, 2nd. edition, 2002.
- [117] P. Thomas, D. Teneketzis, and J. K. Mackie-Mason. A market-based approach to optimal resource allocation in intergrated-services connection-oriented networks. In *Proceedings of the Fifth INFORMS Telecommunications Conference*, 2000.
- [118] W. Thomson. Cooperative models of bargaining. In R. J. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 2. Elsevier, New York, 1994.
- [119] Walter Trockel. On the Nash Program for the Nash Bargaining Solution. Working paper 788, University of Bielefeld, Institute of Mathematical Economics, 1999.
- [120] J.N. Tsitsiklis and B. Van Roy. Average cost temporal-difference learning. In *Proc. of the 36th Conf. on Decision and Control*, pages 498–502, Dec 1997.
- [121] B. Van Roy, D. Bertsekas, Y. Lee, and J. N. Tsitsiklis. A neuro-dynamic programming approach to retailer inventory management. In *Proc. of the 36th IEEE Conf. on Decision and Control*, pages 4052–4057, 1997.
- [122] Audrey M. Viterbi and Andrew J. Viterbi. Erlang capacity of a power controlled CDMA systems. *IEEE J. Sel. Areas Commun.*, 11(6):892–900, August 1993.

-
- [123] Audrey M. Viterbi, Andrew J. Viterbi, and E. Zehavi. Other-cell interference in cellular power-controlled CDMA. *IEEE Trans. on Commun.*, 42(2/3/4):1501–04, February/March/April 1994.
- [124] Q. Wang and J. M. Peha. State-dependent pricing and its economic implications. In *Proceedings of 7th International Conference on Telecommunications Systems Modeling and Analysis*, pages 61–71, 1999.
- [125] Langford B. White. A new policy evaluation algorithms for markov decision processes with quasi birth-death structure. *Stochastic Models*, 21(2-3):785–797, 2005.
- [126] Mingbo Xiao, Ness Shroff, and Edwin K.P. Chong. A utility-based power-control scheme in wireless cellular systems. *IEEE/ACM Trans. on Networking*, 11:210–221, Apr 2003.
- [127] H. Yaiche, R. Mazumdar, and C. Rosenberg. A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Trans. on Networking*, 8(5):667–678, October 2000.
- [128] Wen-Bin Yang and Evaggelos Geraniotis. Admission policies for integrated voice and data traffic in CDMA packet radio networks. *IEEE J. Sel. Areas Commun.*, 12(4):654–664, May 1994.
- [129] K. Yao, M.K. Simon, and E. Biglieri. Statistical modeling of fading channel statistics based on SIRP and H-functions. In *International Conference on Communications, Circuits and Systems*, pages 265–268, June 2004.
- [130] J. Zhang, J. Huai, R. Xiao, and B. Li. Resource management in the next-generation DS-CDMA cellular networks. *IEEE Wireless Commun.*, pages 52–58, August 2004.
- [131] Chi Zhou, Peifang Zhang, Michael L. Honig, and Scott Jordan. Two-cell power allocation for downlink CDMA. *IEEE Trans. on Wireless Commun.*, 3(6):2256–2266, Nov 2004.