THIS THESIS HAS BEEN ACCEPTED FOR THE AWARD OF THE DEGREE OF MASTER OF ENGINEERING SCIENCE

"Signalling Requirements for Smart Pricing in Mobile Telecommunications Systems"

SIGNALLING REQUIREMENTS FOR SMART PRICING IN MOBILE TELECOMMUNICATIONS SYSTEMS

Thesis submitted by

Dat Tien Vo

for the Degree of *Doctor of Philosophy*

September 2010

School of Electrical and Electronic Engineering

Faculty of Engineering, Computer and Mathematical Sciences

Statement of originality

This work contains no material which 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 my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

Signed

Date

Acknowledgements

First, I would like to thank my supervisors Dr Matthew Sorell and Associate Professor Michael Liebelt for their guidance and support throughout this research. It was a sustained effort and without their commitment, the research could have not been completed.

Second, I would like to thank Dr Andrew Coyle for pointing me to a fruitful direction in the most critical stage of this research, for guiding me until I reached successful outcome, and for making himself available for consulting when I needed him most. His unconditional support will be remembered for many years to come.

Third, I would like to thank Dr Paul Chapman, Professor Lang White, Dr Andrew Metcalfe and Dr David Green for their help or advice at the beginning of this research. I am also grateful to have had the general support from staff in the Electrical and Electronic Engineering School Office, particularly Rose-Marie Descalzi. I am thankful to Angela Wong for proof-reading this thesis.

Fourth, I would like to thank Professor Reginald Coutts and Smart Internet Technology Cooperative Research Centre for awarding me with a scholarship. I would also like to thank my employer, the Australian Communications and Media Authority, for approving my leave so that I could spend time on this research at its final stage.

Finally, I would like to thank Chris Gent for believing in me during difficult situations. Without his understanding, support and effort, this research would have not even commenced. If one doubts about the power of believing in someone, the completion of this research is evidence.

Dedication

No house can be built without a good foundation. I am lucky to have a strong and reliable foundation from my family and wife. They are my grounding and sources of motivation. This research has been a long journey and along the way there are many challenges and obstacles. When encountering those, receiving the support and love from my foundation has given me determination to finish this research. This thesis is dedicated to my extended family - the Viens, Vos and Lees.

Abstract

Smart Pricing can be classified as Dynamic Pricing and bears resemblances to Congestion Pricing. It is a pricing scheme that varies prices according to the current users' responses to rising load. *Smart Pricing* is a solution to the problem of under-utilised network resources or to accommodate growing demand within existing network resources. All three pricing schemes necessitates signalling, however, little is known about the signalling requirements. This thesis makes original contributions in this very area whereby it:

- analyses the current 3G mobile telecommunications systems network architecture and shows how *Smart Pricing* can be implemented;
- proposes two models for implementing *Smart Pricing* in 3G mobile telecommunications systems. In these models, a new network element so-called *Dynamic Pricing Engine* is proposed to be added;
- calculates and reports required signalling requirements for *Smart Pricing*; and
- extends the models to more advanced telecommunications systems.

The first model proposed is the *Monte Carlo Simulation* model in which operation of *Smart Pricing* is simulated and the required signalling is calculated. Both small and large *Smart Pricing* systems[1] are investigated and eighteen simulation scenarios are conducted. Highlights[2] of our findings are as follows. When there are more users in the system, the bidding signalling percentage on the uplink increases but decreases on the downlink and on links between network elements. It is not how the level of congestion is defined, it is the user behaviours that dictate the signalling requirements for *Smart Pricing*.

The second model is the *State Space Analysis* model, in which the *Markov Chain* technique is employed. Highlights[3] of our findings are as follows. In the steady-state

[1] A *Smart Pricing* system is defined as a WCDMA UMTS system which adopts *Smart Pricing*.
[2] A complete set of the findings can be found in Section 3.9.
[3] A complete set of the findings can be found in Section 4.12.

condition, the maximum average signalling loads for the uplink, downlink and links between network elements can be accommodated with existing signalling system capacity. With respect to simulation time, this model is significantly faster than the*Monte Carlo Simulation* model. It is recommended that the *Dynamic Pricing Engine* be collocated with the Billing System.

Applicability of the proposed models to more advanced cellular telecommunications systems, such as HSDPA, HSPA+ and LTE is also demonstrated. Then, estimated average signalling loads[4] are reported. Finally, the models are shown to be able to be applied in other resource-constrained and non-cellular telecommunications systems, particularly *Cognitive Radio*[5].

[4]Summary of details can be found in Section 5.6.
[5]Summary of details can be found in Section 6.6.

Publication

D. Vo and M. Sorell, "Signaling Requirements for Smart Dynamic Pricing in Cellular Networks". in *Proc. of the 1st International Conference on Communications and Networking in China (ChinaCom 2006)*, 25-27 Oct. 2006, pp. 1-5, Beijing, China.

Contents

List of abbreviations

2G	Second Generation
2-WTP	CONF with 2 levels of WTP
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5-WTP	CONF with 5 levels of WTP
16-QAM	16-Quadrature Amplitude Modulation
64-QAM	64-Quadrature Amplitude Modulation
AMR	Adaptive Multi Rate
AMT	Aeronautical Mobile Telemetry
AoC	Advice of Charge
ASP	Application Service Part
ATM	Asynchronous Transfer Mode
AuC	Authentication Centre
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BER	Bit Error Rate
BISUP	Broadband ISUP
BLER	Block Error Probability
BSC	Base Station Controller
BSS	Base Station System
BTS	Base Transceiver Station
CAMEL	Customised Applications for Mobile Network Enhanced Logic
C/I	Carrier to Interference Ratio
CCPCH	Common Control Physical Channel
CDF	Charging Data Function
CDMA	Code Division Multiple Access
CDR	Charging Data Record

CGF	Charging Gateway Function
CONF	Video Conferencing Service
CPCH	Control Physical Channel
CQI	Channel Quality Indicator
CR	Cognitive Radio
CR-BCH	CR Broadcast Channel
CR-DCCH	CR Downlink Control Channel
CRE	CR Engine
cr	cost recoverable signalling
CR-RACH	CR Random Access Channel
CRRM	Common Radio Resource Management
CR-UCCH	CR Uplink Control Channel
CRW	Cognitive Radio Gateway
CS-MGW	Circuit Switched-Media Gateway
DL	Downlink
DL-SCH	Downlink Shared Channel
DPA	Dynamic Pricing Adapter
DPDCH	Dedicated Physical Data Channel
DPE	Dynamic Pricing Engine
EDGE	Enhanced Data rates for GSM Evolution
ECR	Effective Coding Rate
eCell_FACH	enhanced Cell_FACH
EIR	Equipment Identity Register
ENG	Electronic News Gathering
EPS	Evolved Packet System
FACCH	Fast Associated Control Channel
FACH	Forward Access Channel
FCC	Federal Communications Commission
FDMA	Frequency Division Multiple Access
FDD	Frequency Division Duplex
FTAM	File Transfer Access and Management
FTP	File Transfer Protocol
GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switching Centre
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GTP	GPRS Tunnel Protocol
HLR	Home Location Register

HSDPA	High Speed Downlink Packet Access
HS-DPCCH	High Speed-Dedicated Physical Control Channel
HS-DSCH	High Speed-Downlink Shared Channel
HSPA+	High Speed Packet Access Evolution
HS-PDSCH	High Speed-Physical Downlink Shared Channel
HS-SCCH	High Speed-Shared Control Channel
HSS	Home Subscriber Server
IETF	Internet Engineering Task Force
IMT	International Mobile Telecommunications
IN	Intelligent Network
IP	Intelligent Peripheral
ISDN	Integrated Service Digital Network
ISDN-UP	ISDN-User Part
ISUP	ISDN-UP
ITU	International Telecommunication Union
ITU-T	ITU-Telecommunication Standardization Sector
LAPD	Link Access Protocol on D Channel
LAPDm	Link Access Protocol on Dm Channel
LRT	Likelihood Ratio Test
LTE	Long Term Evolution
MAC	Medium Access Control
MAC-hs	Medium Access Control-high speed
MAP	Mobile Application Part
MCS	Monte Carlo Simulation
ME	Mobile Equipment
MIMO	Multiple Input Multiple Output
MMI	Man Machine Interface
MS	Mobile Station
MSC	Mobile Switching Centre
MTP	Message Transfer Part
NBAP	Node B Application Part
Netw2Netw	Between Network Elements
OCS	Online Charging System
OFCS	Offline Charging System
OFDM	Orthogonal Frequency Division Multiplexing
OMC	Operations and Maintenance Centre
OSF	Operations System Function
P-CCPCH	Primary-Common Control Physical Channel

PCH	Paging Channel
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Resource Function
PDN	Packet Data Network
PDN-GW	Packet Data Network Gateway
PDSCH	Physical Downlink Shared Channel
P-GW	Packet Data Network Gateway
PLMN	Public Land Mobile Network
P-SCP	Prepaid Service Control Point
PSTN	Public Switched Telephone Network
PU	Primary User
QAM	Quadrature Amplitude Modulation
QCI	QoS Class Identifier
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RACH	Random Access Channel
RANAP	Radio Access Network Application Part
RAT	Radio Access Technology
RF	Radio Frequency
RLC	Radio Link Layer
RNC	Radio Network Controller
RNS	Radio Network System
RNSAP	Radio Network Subsystem Application Part
RRC	Radio Resource Control
SAE	System Architecture Evolution
SAT	SIM Application Toolkit
SCCP	Signalling Connection Control Part
S-CCPCH	Secondary-Common Control Physical Channel
SC-FDMA	Single Carrier - Frequency Division Multiple Access
SCP	Service Control Point
SCTP/IP	Stream Control Transmission Protocol/Internet Protocol
SDR	Software Defined Radio
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SGW	Signalling Gateway Function
SIM	Subscriber Identity Module
SIR	Signal-to-Interference Ratio
SINR	Signal-to-Interference-plus-Noise Ratio

Abbreviations

SIT CRC	Smart Internet Technology Cooperative Research Centre
SMS	Short Message Service
SMS–GMSC	SMS–Gateway MSC
SMS–IWMSC	SMS–Interworking MSC
SMS–SC	SMS–Service Centre
SNR	Signal-to-Noise Ratio
SPR	Subscription Profile Repository
SR	Software Radio
SS7	Signalling System #7
SSA	State Space Analysis
SSP	Service Switching Point
STP	Signal Transfer Point
TCAP	Transaction Capabilities Application Part
TCP/IP	Transmission Control Protocol/Internet Protocol
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TFTP	Trivial File Transfer Protocol
TOU	Time-Of-Use
TSS	Tariff Setting System
TUP	Telephone User Part
TVOB	Television Outside Broadcast Network
uc	uplink signalling
UE	User Equipment
UHF	Ultra High Frequency
UL	Uplink
UMTS	Universal Mobile Telecommunication Systems
USAT	USIM Application Toolkit
USIM	Universal SIM
USSD	Unstructured Supplementary Service Data
UTRAN	UMTS Terrestrial Radio Access Network
VHF	Very High Frequency
VLR	Visitor Location Register
WCDMA	Wideband Code Division Multiple Access
WIN	Wireless Intelligent Network
WTP	Willingness To Pay
WTP1	Lowest WTP
WTP5	Highest WTP

CHAPTER 1

Introduction

1.1 Why this research?

Many mobile telecommunications networks are still in transition from *Second Generation* (2G) to *Third Generation* (3G) Systems. The move to 3G systems is being driven by increasing numbers of subscribers and by the desire for more and better services. As a result, required data rates are growing substantially. To accommodate these needs, 3G systems are designed to handle many more subscribers and to offer higher bit rates (up to 384 kbits/s), variable bit rate for bandwidth on demand, error rates down to 10^{-6} and coexistence of 2G and 3G systems [1].

Parallel to the evolution of the mobile telecommunications networks, other communication technologies have also been developed. The size and data rate of the applications for these technologies rise with the advances in technology, making congestion a potentially serious problem. At peak times, with subscribers congregating in a small geographical area, concurrently using these advanced applications, a significant burden is imposed on mobile network resources. Thus, the capacity in 3G systems can be quickly exhausted and more network infrastructure investment will be needed. Smart Pricing, which has prices varied according to users' responses to congestion, is a possible response to the problem of congestion. The focus for this research is investigating the signalling required to implement Smart Pricing in mobile telecommunications systems. Smart Pricing describes pricing schemes which help network operators effectively utilize their available network resources, thus reducing the need for expanded network infrastructure investment. In such pricing schemes, subscribers are charged dependent on the level of network congestion. The price of calls varies in response to changes in the willingness to supply and demand [2] and corresponds to economists' notion of efficient spot prices that combine shadow prices to allocate resources efficiently in the presence of congestion externalities [3] [4].

Due to the nature of Smart Pricing, new prices will be generated frequently. These price messages must be delivered to subscribers in each cell differently and appropriately to match with different levels of traffic. Responses from subscribers are also needed to indicate whether or not they accept the new price for their current call. The subscribers might also have the option to initiate a request for an increase in *Quality of Service* (QoS). In addition, signalling between network elements is also necessary for purposes such as reporting congestion levels and updating relevant network elements of new prices. Hence, for Smart Pricing to be deployed, three types of signalling are required. They are: uplink, downlink and inter-network-elements signalling.

The primary aim of this thesis is to report the research undertaken into signalling requirements for the implementation of Smart Pricing in mobile telecommunications networks. The secondary aim is to extend the signalling models proposed for mobile telecommunications networks to a wider range of systems.

1.1.1 Objectives

The objectives of this research are to:

1. identify the signalling requirements for Smart Pricing in mobile telecommunications systems, particularly *Wideband Code Division Multiple Access* (WCDMA) for *Universal Mobile Telecommunication Systems* (UMTS);
2. estimate the signalling loads for Smart Pricing in WCDMA UMTS;
3. provide detailed analysis on how to use the proposed signalling models for implementing Smart Pricing in more advanced mobile telecommunications systems; and
4. propose an approach for implementing Smart Pricing in systems other than mobile telecommunications systems.

1.1.2 Scope

This research will not consider the normal signalling requirements for call setup and teardown in mobile telecommunications systems. This means that the signalling load estimated from this research is incremental to the conventional signalling. This research also does not explore the process of setting new prices. It is presumed simply that the price messages are automatically yielded by the *Tariff Setting System*[1] after the Tariff Setting System receives the congestion information. This presumption also applies to the non-cellular telecommunications and other resource-constrained systems that we consider in this research, e.g. *Cognitive Radio*.

[1]See details in Section 2.4

Hence, the scope of this project is to:

1. focus on the current signalling system for mobile telecommunications systems and take into account the evolution of that signalling system;
2. take into consideration network elements of the system architecture of current mobile telecommunications systems that have connection with the Smart Pricing;
3. consider the signalling requirements for Smart Pricing for one cell;
4. extend the proposed signalling models to more advanced mobile telecommunications systems in which technical specifications are set by the 3rd Generation Partnership Project; and
5. extend the proposed Smart Pricing and signalling models to non-cellular telecommunications and other resource-constrained systems, particularly one that is currently drawing significant research interests, i.e. Cognitive Radio.

1.2 Literature review and gap statement

Literature survey has been undertaken with best effort and no work on Smart Pricing has been found apart from the work in [2] by a research team of which the author is a member. This is expected, as Smart Pricing is a new pricing scheme which the research team proposed. Some detailed work on Smart Pricing has also been reported in [5], but the focus of that work was on resource allocation.

A few similar works to Smart Pricing, i.e. Dynamic Pricing and Congestion Pricing in mobile telecommunication systems, were found but there are a limited number of such publications. This has been confirmed in [6], [7], [8] and [9]. In the works that have been published, there is little consideration regarding the signalling requirements.

In the next two sections, we firstly look at how Dynamic and Congestion Pricing are defined and secondly give a summary of the current works on these two types of pricing schemes in mobile telecommunications systems, particularly the signalling requirements aspect. From the summary, we then identify gaps, which allows us to make original contributions, in the subsequent section.

1.2.1 Dynamic Pricing and Congestion Pricing

Dynamic Pricing describes pricing schemes that have different prices for calls depending on the network situation. Congestion Pricing (also known as Congestion-dependent Pricing) describes pricing schemes that take into account the congestion-dependent component in setting prices of calls [10]. Congestion Pricing not only discourages usage when the network is congested but also generates revenue for capacity expansion [11]. Congestion Pricing is a category of Dynamic Pricing.

A good summary of Dynamic Pricing schemes can be found in [5], whereby the author categorises Dynamic Pricing into three types. In *auction-based* pricing schemes, prices are set based on users' bids for network resources. In *shadow* pricing schemes, prices are set to maximise the aggregate utility of all users. Lastly, in *stochastic control* pricing schemes, prices are set based on network operator knowledge about the nature of arrival and departure of the users. In contrast, in [12], the authors categorise Dynamic Pricing into only two types: *non-differentiated* and *differentiated*. The difference between these two types is whether a Dynamic Pricing scheme offers different services, each with a different level of QoS which is honoured even during congestion periods.

Thus, if based on [5], Smart Pricing falls into the mixed categories of the *auction-based* and *shadow* Dynamic Pricing because prices are shadow prices and set based on users' bids. On the other hand, if based on [12], Smart Pricing falls into the category of *differentiated* Dynamic Pricing because it guarantees best QoS for all calls when network load is light. When congestion occurs, best QoS is guaranteed for highest bidders, and for any other users QoSs are guaranteed to be greater than or equal to a pre-agreed minimum QoS.

1.2.2 Current work on Dynamic Pricing in mobile telecommunications systems

In general, it is recognised in [13] that a flexible charging architecture is required to accommodate Dynamic Pricing because other pricing schemes such as flat-rate and one-off charge per service still exist and not every mobile telecommunications network operator is willing to adopt Dynamic Pricing.

In [14], the authors propose a Congestion Pricing model for general communication networks. An important finding in this work is that static pricing can approach optimality, which indicates that time-of-day pricing (a slow-price-variation Dynamic Pricing scheme) is adequate if demand varies slowly. A large number of aspects are considered in this work, which among others include: revenue maximization, welfare maximization and ways to obtain an optimal Dynamic Pricing policy.

In [11], Congestion Pricing is also the focus. The authors outline the economic theory of pricing a congestible resource, such as a communication network, whereby many users can share but the quality of the resource degrades as the number of users increase. Market power is taken in account in determining ways to set price in order to maximise net social benefits in this work.

In [12], the focus changes to Dynamic Pricing. The authors propose a differentiated Dynamic Pricing scheme in which users are classified into different QoS levels based on QoS parameters. The objective of this work is to use Dynamic Pricing to maximise network operator revenue while assuring resource and QoS constraints are met. In solving such an optimization problem, the *Block-Depth-First Search* algorithm is utilised.

In [15], the authors propose a Dynamic Pricing scheme which is based on a linearly increasing price vector. User demand and call duration are modelled as functions of price and Markov technique is used to obtain an optimal pricing policy. Beside having the spotlight on maximizing network revenue like the above works, the authors also investigate blocking probability and cost per unit time for voice service.

In [8], the author tests the effectiveness of Dynamic Pricing when prices and network load has a linear relationship. The author finds that a linear pricing function should not be used when a target revenue is aimed to achieve. Control theory is used in the test.

In [16], Dynamic Pricing is applied to the same mobile telecommunications system considered in Smart Pricing, i.e. the WCDMA system. The authors attempt to use Dynamic Pricing as a means to allocate the uplink radio capacity of a cell to voice and data calls, each service with a different utility function. Different from any of the above works, the focus of this work is not to maximise the network operator revenue but to attain an optimal signal-to-interference ratio for data calls and transmit rate for voice calls.

In [17], the focus is placed on auction system, which is an element of auction-based Dynamic Pricing schemes. The auction system that the authors propose allows a user to delegate an auction task to a mobile agent and controls the agent through its mobile device. This auction system is similar to the one employed in Smart Pricing, which to a certain extent is captured in Figure 6.1. The idea integrated into that figure is that the *Bidding Agent* (a function of the *Dynamic Pricing Adaptor* in the user's equipment) participates in an auctioning process in the *Auction Room* (a function of the *Dynamic Pricing Engine*) making pre-programmed decisions on behalf of the user. The user still has the power to manually override those pre-programmed decisions if necessary.

In [18], the authors propose two *real-time*[2] payment methods in a form of micropayments. Additional computation cost for the methods compared with the conventional method which is based on *Charging Data Records* is identified. The authors argue that Dynamic Pricing can be accommodated by these real-time payment methods if a pricing contract is established upon call connection.

Although various aspects of Dynamic Pricing are investigated in the above eight works, no consideration is given on the signalling requirements. We are somewhat surprised that such consideration is missing even in [18] and [17] when one would think that it is vital for a payment method or an auction process to be proposed.

[2]In some works, such as [6] and [7], Dynamic Pricing is also named *real-time* pricing.

In [19], it is said that there is no need for complex control mechanism for congestion because Dynamic Pricing is auto-regulating. To a certain extent, this verifies the finding of the work in [8]. In that work the authors propose a dynamic pricing scheme for data services in General Packet Radio Service networks. A similarity between this pricing scheme and Smart Pricing is that high prices will only apply when the network load reaches a certain threshold. The difference is that with this pricing scheme, the user has to stop transmitting if he/she is not willing to pay the high price in the event of congestion, whereas with Smart Pricing the user will still be able to transmit, only the bit rate is reduced. The only information about the signalling requirements in this work is a mention that prices (when the system experiences congestion) will be broadcasted to users.

In [20], the authors propose a Dynamic Pricing scheme which allows users to choose between the conventional fixed price with acceptable QoS degradation and dynamic prices with superior QoS. There are two queues and depending on which pricing option a user chooses, he/she will be placed into the priority queue or the conventional queue. This pricing scheme bears a resemblance to Smart Pricing because a user has the option to pay a higher price to avoid QoS degradation or to accept the QoS degradation and pay a low price. The fundamental difference however is that with this pricing scheme such an option is provided *before* the user is admitted to the network, whereas *after* with Smart Pricing. We argue that *after* is better because, whether the user accepts the dynamic prices or not, the user's call is still continuing, while *before*, the user will have to experience waiting time. Moreover, there is no guarantee in this pricing scheme on how long the user has to wait in the queue. Apart from mentioning that the *pricing block* in its system diagram broadcasts prices to users, no detailed consideration is given on the signalling requirements.

In [9], the authors propose to use Dynamic Pricing to select a Radio Access Technology (RAT) in heterogeneous networks. In a *heterogeneous network*, a number of *Radio Access Networks* connect to the same core network. Such a network allows for a Common Radio Resource Management (CRRM) to be utilised. The author finds that the proposed RAT Dynamic Pricing selection algorithm could help avoid overload in a RAT and balance loads between all involved RATs. Again, the only information about the signalling requirements in this work is a mention of broadcasting price messages to users.

In [6], the authors suggest that call prices are calculated based on the total number of available channels in the circuit switched networks while on average packet delay and throughput in packet switched networks. They then predict users' reaction to changes in price when Dynamic Pricing is applied on the Global System for Mobile Communications (GSM) and UMTS. Load of the network is determined by measuring traffic intensity, however, there is no mention of how this load information is signalled to the Billing System so that new prices can be set. There is also no discussion about how users are signalled with the new prices. The authors state that bidding signalling for Dynamic Pricing will overwhelm the signalling capacity of cellular networks but provide no supporting evidence.

In [7], the authors suggest that call prices are set based on a set of target system performance parameters, which are determined according to traffic intensity, revenue and grade of service. The authors propose a Dynamic Pricing algorithm in a form of a feedback loop and suggest that it be implemented in the Base Station Controller (BSC). Call prices are said to be sent to Mobile Stations (MSs) using the *Broadcast Control Channel* (BCCH). There is however no reasoning as to why the BSC is chosen but not another network element, say in the Billing System. There is also no evidence on why the BCCH is chosen. If, for example, a Dynamic Pricing scheme which is based on users' bidding is adopted then there is no need to send new prices to all users, but only a subset of them. Using the BCCH to broadcast new prices could interfere with users who are not participating in the bidding process. This, in turn, could affect customer satisfaction level.

1.2.3 Gap statement

It can be seen from a summary of the literature on Dynamic Pricing in mobile telecommunications systems in the previous section that focus of the current works is mainly on proposing different Dynamic Pricing schemes, auction systems and payment methods. There are limited discussions about the signalling requirements apart from mentioning that prices will be broadcasted using the BCCH. Many technical aspects required for the implementation of Dynamic Pricing are missing. Those that were not considered, *inter alia*, are:

1. modelling of the signalling requirements for Dynamic Pricing;
2. required signalling components and protocols;
3. required network resources for signalling; i.e. what are the estimated signalling loads on the uplink, downlink and links between network elements;
4. the uplink path which subscribers would use to advise the network operator of their preference for price and QoS; and
5. how a new price is displayed in the MSs.

Without considering, at the very least, those aspects, the claim that Dynamic Pricing helps to reduce the number of blocked calls, increases revenue and increases capacity utilisation can just mean shifting of costs. Some critical questions that must be answered are: how much network traffic needs to be sacrificed for signalling purposes? Are there any changes that need to be made to the current system architecture? Answers to these questions have a significant affect on the decision to bring or not to bring Dynamic Pricing into practice. Furthermore, obtaining these answers beforehand is of great importance for any network operator before committing any resource to implement Dynamic Pricing. This research provides answers to all of those questions and addresses all the technical aspects identified above. Results from this research are not only specifically for Smart Pricing but also serve as a good estimate of the signalling requirements for any other Dynamic Pricing schemes proposed for use in mobile telecommunications systems.

It should be noted that Dynamic Pricing proposals for the Internet are not applicable to mobile telecommunications systems because of the fundamental differences between the two system architectures, even considered in the context of an all-IP core network. The business models and the placement and management of the charing functions are different. Even the *charging*, *accounting* and *billing* terms have different meanings. The Internet systems require direct agreements between a user and each independent service provider, whereas mobile telecommunications systems rely on an operator-centric business model [13].

1.3 Thesis layout and original contributions

In *Chapter 1*, reasons why this research is undertaken and results of the literature review process are provided. Then, gaps for original contributions to be made are identified.

In *Chapter 2*, background knowledge required for this research is summarised, the Smart Pricing scheme is outlined and the current architecture of mobile telecommunications systems is analysed. Then, the first contribution is made by showing how Smart Pricing can be implemented.

In *Chapter 3*, the second contribution is made where the first model for implementing *Smart Pricing* in 3G mobile telecommunications systems is proposed and required signalling requirements are calculated. This model is called the *Monte Carlo Simulation* (MCS) model. The model enables required signalling loads when the Smart Pricing system is in its instantaneous conditions to be calculated. With the MCS model, a new network element, the so-called *Dynamic Pricing Engine* (DPE), is proposed to be added to the current mobile telecommunications network architecture. This network element is responsible for setting prices depending on the congestion level in a cell and specifies the QoS required. The model is investigated thoroughly where three types of systems, how having more types of user's willingness to pay, or how having more users in the system affecting the required signalling loads are examined. For each system, a wide range of parameters are considered from the conditions under which a user is admitted, to the level at which congestion is defined, and users' behaviour when they are already in the system. At the end of the chapter, findings from a large number of simulation scenarios are reported and three recommendations are made.

In *Chapter 4*, the third contribution is made where the second model for implementing *Smart Pricing* in 3G mobile telecommunications systems is proposed and required signalling requirements are calculated. This model is called the *State Space Analysis* (SSA) model. The model is developed using the *State Space* and *Markov Chain* analysis. The model enables signalling requirements when Smart Pricing system is in its steady-state condition to be calculated. With the intention of comparing the MCS model, most of the simulation scenarios conducted for the MCS model are replicated. Preliminary results demonstrate consistency between the two models with much greater computational efficiency, enabling us to analyse a further two aspects of the system. Firstly, how adding more signalling components to an interaction impacting on the overall signalling load is examined. Secondly, different options for locating the DPE are examined. At the end of the chapter, findings are reported showing that the SSA model not only has the same capability as the MCS model of informing the signalling requirements but is also able to predict the long-term signalling requirements and give substantial improvements. A further four recommendations are made.

In *Chapter 5*, the fourth contribution is made where the MCS and SSA models are extended to accommodate more advanced telecommunications systems. Applicability of Smart Pricing and the MCS and SSA models to the 3GPP Release 99 UMTS downlink system and to more advanced 3GPP mobile telecommunications developments are investigated. The systems considered are the High Speed Downlink Packet Access, the High Speed Packet Access Evolution and the Long Term Evolution systems. The intentions for this chapter are to provide mobile telecommunications network operators with a complete solution for the implementation of Smart Pricing in the 3GPP Release 99 UMTS system and to show the effectiveness of the MCS and SSA models. To show the applicability of Smart Pricing, the MCS and SSA models, the characteristics of each system are summarised and the evolution of the capacity and architecture of the systems is analysed. Techniques to calculate the maximum system capacity and the sharing of that capacity by individual users are provided. Smart Pricing system diagrams for the High Speed Downlink Packet Access, the High Speed Packet Access Evolution and the Long Term Evolution system are constructed. Lastly, a technique to estimate the Smart Pricing maximum average signalling loads for all four systems and use it to estimate the maximum average signalling loads is developed.

In *Chapter 6*, the fifth contribution is made where the MCS and the SSA models are extended to accommodate non-cellular telecommunications and other resource-constrained systems. The systems considered are the electricity, Asynchronous Transfer Mode, water and Cognitive Radio system. Among the four system, Cognitive Radio is the one that is given a particular focus. Cognitive Radio is a relatively new concept and currently attracts much research. Two system architectures are proposed for implementing Smart Pricing in non-cellular telecommunications and other resource-constrained systems; a general one for all systems and another one specifically tailored for Cognitive Radio systems. In both system architectures, new network elements are proposed.

Chapter 7 concludes the thesis. A summary of major findings from this research and suggestions for future work is provided.

CHAPTER 2

Background

This chapter provides an overview of the background knowledge for this project. Section 2.1 addresses the nature of Smart Pricing. Section 2.2 discusses the current signalling system, *Signalling System #7* (SS7), and how mobile networks deploy it. Section 2.3 covers the architecture and protocols of the mobile telecommunications systems. The Tariff Setting System is outlined in Section 2.4. Means for displaying a new price in a mobile station are covered in Section 2.5. Section 2.6 examines mobile prepaid service. The Smart Pricing signalling path will be presented in Section 2.7 to illustrate an overall picture of how this background information is combined and fits into this project. Section 2.8 outlines possible signalling means. Finally, Section 2.9 discusses the uplink load and capacity of WCDMA systems.

2.1 Smart Pricing principles

With the current static pricing schemes – single price and, to a lesser extent, peak/off-peak price – mobile network resources are under-utilised during quiet or off-peak hours. The under-utilisation factor of the peak versus off-peak demand can be as high as 20 to 1 [6]. This suggests that a means to allocate better limited network resources is greatly needed. The problem is not simply one of short term efficiency. As the number of subscribers increases, more network resources will be needed to accommodate the growing peak demand. Instead of investing more in network resources to meet this demand, network operators could use dynamic pricing to constrain the demand within the available network resources while allocating them to those users who value them most.

Dynamic Pricing refers to a range of pricing schemes that vary the price of calls as demand on the network fluctuates [7]. Smart Pricing is one of a number of pricing schemes that respond to the state of the network and can be classified as dynamic. Smart Pricing is also a type of congestion pricing where the price varies directly with the load. It is different from a simple congestion price in that it does not posit a simple relationship between price and load such that a given load is associated with a particular price. Instead, Smart Pricing varies that relationship according to the characteristics of the current user group, sometimes increasing prices quickly with load, sometimes more slowly, depending on users' responses to the rising load. It is a pricing scheme which varies price in response to the relative scarcity of network resources and has been developed to be implemented on WCDMA networks, in work initially funded by the *Smart Internet Technology Cooperative Research Centre* (SIT CRC) [2]. That work also includes development of a model to assess various pricing schemes and it will address the implementation issues. In particular, it uses degradation in QoS as an indication to users of network congestion, i.e. as the load increases past a certain point, relative to the total capacity of the cell, the QoS of existing users declines. Users can respond by indicating to the network that they are willing to pay more to avoid the QoS degradation. The network responds by allocating more of the available bandwidth to those users, charging them more and using the higher charges to establish prices for new entrants. Under this approach, the price variation depends on the strength of demand and the estimated congestion costs that one active subscriber imposes on the others. Therefore, the price of a call during the day and possibly even during the course of a call might fluctuate. All active subscribers must be notified of these price changes and there must be a mechanism for the subscribers to respond to the network operator as to whether or not they accept the new combination of price and QoS. All of these implementations impose signalling requirements and the question arises as to how the signalling tasks can best be achieved.

In this section, we have explained that with Smart Pricing new call prices will have to be sent frequently to the active users in the network, and users will also need to be able to communicate with the network via some signalling method. In the next section, we will look at the current signalling system.

2.2 Common channel signalling #7 (SS7)

Whether it is a wireless network or not, signalling within a telecommunications network, in general, allows telephone switches or packet switches to communicate with each other and share the information needed to process any type of call. Signalling System #7 (SS7), in particular, allows data communications on a 56 or 64-kbits/s link between the switches for data, video, voice and audio networks. These communications include call setup and teardown, accessing remote databases to provide routing and billing information for all telephone services.

SS7 uses a packet switching network separated from the voice network. It requires only connectionless services such as those delivered over the Internet. It uses messages, which are encapsulated in packets, to request services from other entities within this or another telecommunications network. The architecture of the SS7 network is shown in Figure 2.1.

NOTE:
This figure is included on page 13
of the print copy of the thesis held in
the University of Adelaide Library.

Figure 2.1: SS7 Network Architecture [21].

There are three kinds of network elements, called signalling points, in the SS7 network [21]. They are:

- *Service Switching Point* (SSP): facilitates the communication with voice switches and creates packets needed for transmission in the SS7 network. These packets are for signalling to connect voice circuits or for database access;
- *Signal Transfer Point* (STP): transfers packets from one SSP to another, interfaces to another networks, performs protocol conversion and routes packets to databases. Its features include: screening to maintain network security, measurements of traffic and usage, and data mining to learn about customers' calling patterns and usage locations; and
- *Service Control Point* (SCP): interfaces to databases of a telephone network.

Mobile telecommunications networks utilise SS7 extensively to establish or disconnect circuit connections within its own network, to other mobile networks, and to *Public Switched Telephone Networks* (PSTNs). SS7 is also used to update the current location of subscribers, access subscriber details in *Home Location Register* (HLR) databases, or share this information with other mobile telecommunications networks in case of, for instance, seamless roaming. These functions are enabled through the connections of the *Mobile Switching Centers* (MSCs) to the SSPs at one end and databases to the SCPs at the other. The protocol stack used for these connections is in Figure 2.2, where:

NOTE:
This figure is included on page 14 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.2: SS7 Protocol Stack [21].

- *Message Transfer Part* (MTP): combines the *Physical*, *Data Link* and *Network* layers into one layer and routes packets from the source all the way to the destination [22];
- *Signalling Connection Control Part* (SCCP): assists the MTP with address translation to route packets to the correct databases in case the database's address is unknown to the sending node;
- *Application Service Part* (ASP): covers the services of the *Transport*, *Session* and *Presentation* layers [22];
- *Transaction Capabilities Application Part* (TCAP): accesses, queries and retrieves information from databases; and
- *Telephone User Part* (TUP), *Integrated Service Digital Network User Part* (ISDN-UP or ISUP) and *Broadband ISUP* (BISUP) are the protocols used to establish or disconnect connections to different PSTNs.

In this section, we have obtained a general idea about the role and functions of a signalling system, how the current signalling system SS7 links the MSCs to each other and to network databases, and also the protocols used in SS7. We have also understood the signalling among the MSCs and between the MSC and the network databases. In the next section, we will look at the mobile telecommunications network architecture to determine the signalling path from the MSCs to the Tariff Setting System (the place where new prices are usually set) and the protocols used.

2.3 Mobile network architecture and protocols

The most widely used 2G (digital voice centric) cellular telephony network is the *Global System for Mobile Communications* (GSM), which has now evolved into a 3G (integrated voice and data) system known as the *Universal Mobile Telecommunications Systems* (UMTS), based on a WCDMA interface. The evolved core network of the 2G and 3G systems are the same [23], the only difference is the access technology, which in GSM is *Time Division Multiple Access* (TDMA) or *Frequency Division Multiple Access* (FDMA), and WCDMA in UMTS [24].

Apart from the MSC and HLR mentioned in the previous section, the core network also includes other network elements. They are:

- *Mobile Station* (MS);
- *User Equipment* (UE);
- *Base Station System* (BSS);
- *Radio Network System* (RNS). It is also called *UMTS Terrestrial Radio Access Network* (UTRAN);
- *Visitor Location Register* (VLR);
- *Equipment Identity Register* (EIR);
- *Operations and Maintenance Centre* (OMC);
- *Billing System*;
- *Gateway Mobile-services Switching Centre* (GMSC);
- *SMS Gateway MSC* (SMS-GMSC);
- *SMS Interworking MSC* (SMS-IWMSC);
- *Short Message Service - Service Centre* (SMS-SC);
- *Authentication Centre* (AuC);
- *Serving GPRS Support Node* (SGSN) – GPRS is *General Packet Radio Service*;
- *Gateway GPRS Support Node* (GGSN);
- *Charging Gateway Function* (CGF);
- *Node B*;
- *Radio Network Controller* (RNC);
- *Base Transceiver Station* (BTS); and
- *Base Station Controller* (BSC).

For definition of these technical terms, refer to [25] and [26], and for functions' descriptions, refer to [23]. The architecture and protocols for mobile telecommunications network (also called *Public Land Mobile Network* (PLMN)) are in Figure 2.3 and Figure 2.4 [1], [23], [27], [28], [29], [30], [31], [32], [33].

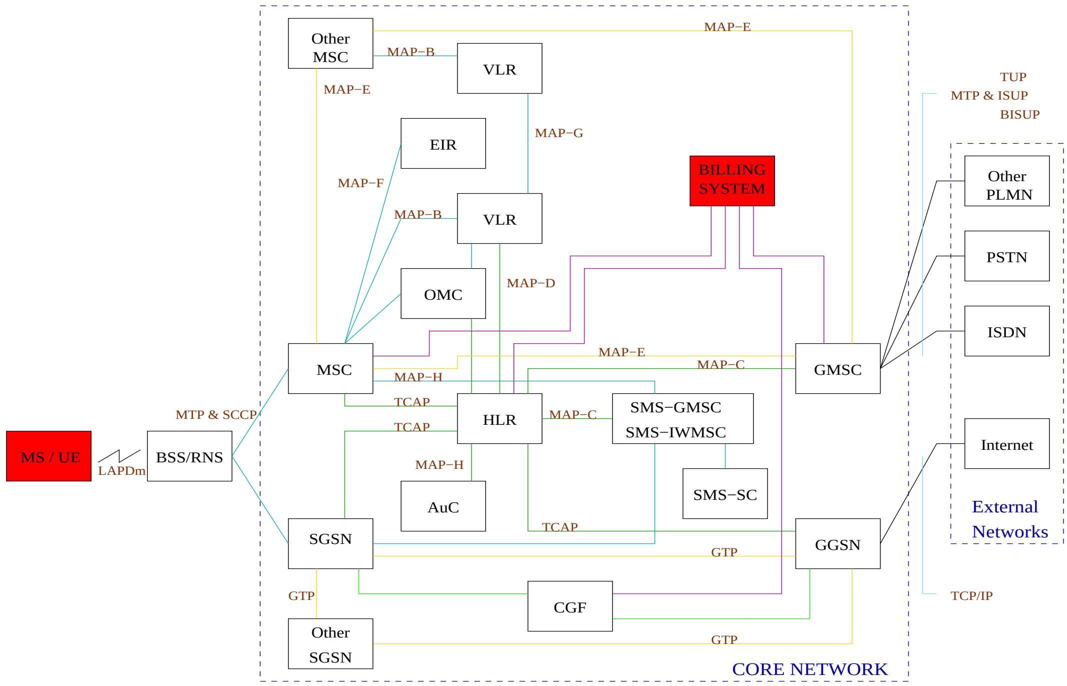


Figure 2.3: Mobile network architecture and protocols.

The protocols used are:

- *Link Access Protocol on D Channel* (LAPD);
- *Link Access Protocol on Dm Channel* (LAPDm);
- *Mobile Application Part* (MAP);
- *GPRS Tunnel Protocol* (GTP);
- *Radio Resource Control* (RRC);
- *Radio Network Subsystem Application Part* (RNSAP);
- *Radio Access Network Application Part* (RANAP); and
- *Node B Application Part* (NBAP).

Also note that the interfaces between the *Universal Subscriber Identity Module* (USIM) and *Mobile Equipment* (ME) is *Cu* and between the *Subscriber Identity Module* (SIM) and ME is *SIM-ME*.

We observe from Figure 2.3 and Figure 2.4 that if a new price message is to be sent to subscribers, its route must be from the Billing System to the MSs. Precisely it is from the Tariff Setting System because, for reasons to be clarified later in Section 2.4, the Tariff Setting System is responsible for setting prices for Smart Pricing and it is located inside the Billing System. There are five network elements that are connected to the Billing System; they are the MSC, GMSC, HLR, SGSN and GGSN. The SGSN and GGSN are connected to the Billing System via the CGF, which can be collocated with the SGSN or GGSN. The protocols for these connections are:

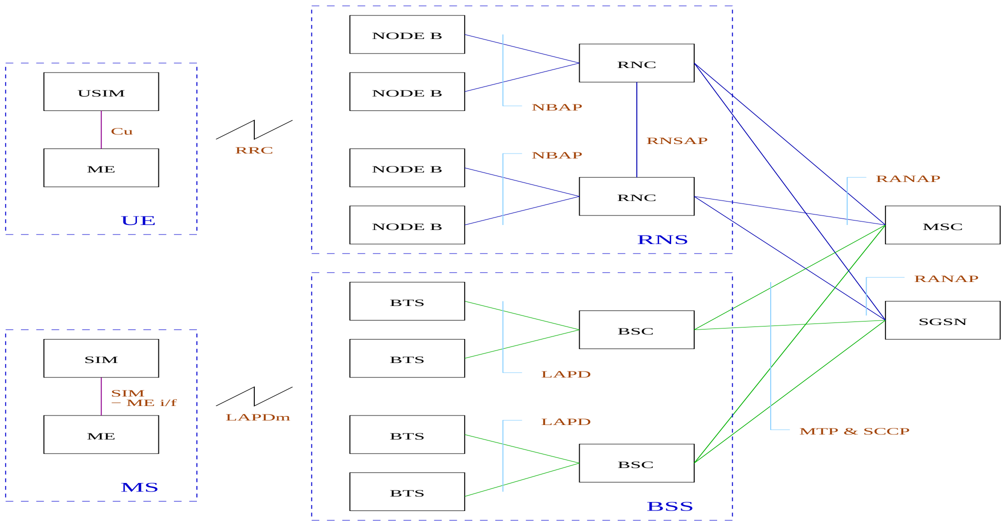


Figure 2.4: Detailed access network architecture of 2G and 3G systems.

- for circuit switched domain: *File Transfer Access and Management* (FTAM), X.25, *File Transfer Protocol* (FTP) (or *Trivial File Transfer Protocol* (TFTP) and *Transmission Control Protocol/Internet Protocol* (TCP/IP); and
- for packet switched domain: GTP.

The previous section explained the SS7 method for signalling among network elements. In this section, we have identified the route and protocols required for the architecture of mobile telecommunications networks. Together, the whole signalling path for a new price to the MSs has been revealed. In the next section, we will see where new transaction prices are set.

2.4 Tariff setting system

Under a Smart Pricing scheme, a function is required which is responsible for not only setting prices depending on levels of congestion of a cell and in accordance with the Billing System rules but also for specifying QoS. We define this as the *Tariff Setting System* (TSS) function.

The original function of a Billing System has been to transform the *Charging Data Records* (CDRs) generated by the charging function into bills required for payment [34]. The charging function in the circuit switched domain includes MSC, GMSC and HLR; whereas in the packet switched domain the CGF [35].

Prices which the Billing System use to charge subscribers are set by the *Tariff Management*. The functions of the Tariff Management are to create, set, get and delete the tariffs defined in the network elements (such as the MSC and GMSC). The Tariff

Management is a function of the *Tariff Administration*. The Tariff Administration, in turn, is a service component of the *Tariff and Charging Administration*. The Tariff Administration provides management functions required to administer the tariffs and parameters required for the application of those tariffs. The Tariff and Charging Administration is an *Operations System Function* (OSF) of the *Telecommunications Management Network*, which is defined by the *International Telecommunication Union-Telecommunication Standardization Sector* (ITU-T). Location of the Tariff and Charging Administration is implementation specific [36] and we assume that it is located inside the Billing System. To minimize possible modifications to the current 3G network architecture, we propose that the TSS also be located inside the Billing System. Such a choice would help reduce the signalling required. The TSS interacts with the Tariff Management function of the Tariff and Charging Administration in setting new prices. Specifically, the Tariff Management sets new prices whenever it is instructed by the TSS and the TSS gives such instructions based on congestion levels of the cell.

Parallel to being responsible for setting new and variable prices dependent on congestion levels, the TSS also sets variable QoS levels appropriate to those prices.

Hence, together with its original function, under a Smart Pricing scheme, the Billing System now has more functions which are those supported by the TSS. In this thesis, we will be concerned about functions of the Billing System provided by the TSS only.

2.5 Displaying a new price in a mobile station

As all price changes must be notified to subscribers, there must be means to make the subscribers aware of the price changes. New price messages must be in an enduring form so that the subscribers will be able to refer to it at any time. There are two modes which subscribers could be in when a new price message is to be displayed, namely the idle and busy modes.

In the idle mode, the display methods could be in text form by using the *SIM/USIM Application Toolkit* (SAT/USAT) in response to network control messages. This could be done by updating an application pre-loaded or remotely loaded after the SIM/USIM card has been issued [37], then by utilising either:

- a field in the MS's display. The sixteen characters of the *Service Provider Indication* [38] field could be used for this purpose; or
- the display inside a menu, such as *Advice of Charge* (AoC). However, this must be preceded by an alerting indicator [38] in the MS, e.g. by sound or flashing light to draw the subscriber's attention.

In the busy mode, an interrupting voice announcement followed by a text message similar to those in the idle mode could be used to inform subscribers of a new price. Others means of signalling new prices in the busy mode may also be adopted.

2.6 Mobile pre-paid service

Beside *Postpaid* mobile subscribers, *Prepaid* mobile subscribers also contribute a significant market for mobile network operators [39]. This payment option presents additional implementation issues for Smart Pricing. For a mobile network to offer *Prepaid* services, additional network elements and signalling were required. This section briefly comments on the choices that network operators have to provide Mobile Prepaid Services and the one believed to be a low risk solution [39] is selected. The extra network elements are then integrated into the mobile telecommunications system architecture to establish an overall signalling path for implementing Smart Pricing.

Network operators have four choices to provide Mobile Prepaid Service [39]:

1. *Handset Based* approach: MS stores the prepaid credit and performs credit deduction by using the AoC and cross-check with the credit limit. There is a fraud risk for this choice because illegal credit modification may occur.

2. *Service Node* approach: includes two extra network elements, the *Service Node* and the *Prepaid Billing Platform*. This approach can accommodate only a small number of subscribers.

3. *Hot Billing* approach: uses charging information in the CDRs generated by the MSC to perform credit deduction at the *Prepaid Service Centre*. However, the CDRs are only sent on the completion of a call. Thus, there is a risk that calls exceeding the credit limit are made.

4. *Wireless Intelligent Network* (WIN) approach: includes two extra network elements, the *Intelligent Peripheral* (IP) and *Prepaid Service Control Point* (P-SCP). The IP is used to advise subscribers of the prepaid credit and the call charging rate. The P-SCP is used to control the call setup at the MSC, perform credit deduction and instruct the MSC to disconnect the call if the credit limit is reached. The communication between the P-SCP and MSC is via the SS7 network. This approach is believed to be a low risk solution since calls exceeding credit limits cannot be made because the P-SCP has immediate control on the MSC to teardown the call. IP is a network element of the *Intelligent Network* (IN) [21]. For simplicity, we assume that the P-SCP can also be integrated into the IN.

In the case which mobile prepaid subscribers use SMS, charging messages will be generated by the SMS-SC. The recipient of these charging messages, via SS7 network, is the P-SCP, which, as in the case of WIN, will perform the credit deduction and instruct the MSC to stop sending any SMS if the credit limit has been reached [39]. Thus, by incorporating the IN and SMS-SC into the the system architecture, mobile pre-paid service can fully be accommodated.

2.7 Signalling paths for Smart Pricing

In Section 2.3, signalling paths from the MS to the charging function (e.g. the MSC) and then from the charging function to the Billing System have been identified. From Section 2.4, we know that the TSS is part of the Billing System. Finally, from Section 2.6 it is known that the IN is required for the prepaid services to be provided. By integrating this information, overall signalling paths for Smart Pricing are revealed and depicted in Figure 2.5.

Since the SGSN can support the CGF internally, the link between the TSS and SGSN may be considered to be direct. The main signalling components to send new price messages to the subscribers are those from the TSS to the MSs. The supporting signalling components are from the MSC to HLR, and others that link the MSC/SGSN, the SMS-SC and the IN. The signalling component to the HLR is required because the MSC needs to be aware of the current location of the subscribers and their subscription details (eg. in the contracted plan). This information is held in the HLR [40]. In Figure 2.5, consistent with the scope set for this research (as outlined in Section 1.1.2), the normal signalling for call setup and teardown components are excluded.

2.8 Possible signalling means

In their work in [7], Fitkov-Noris and Khanifar only identified one possible signalling means, and that is to use the BCCH. In an attempt to find out a possibly more effective means below we identify a range of signalling means which we believe to be feasible for use with Smart Pricing.

2.8.1 The downlink path

The potential signalling means for the downlink path, i.e. from the TSS to the MS, are listed and described below.

1. *Forward Access Channel* (FACH). The FACH is a downlink channel that carries control information to terminals known to be located in a given cell. It is mapped

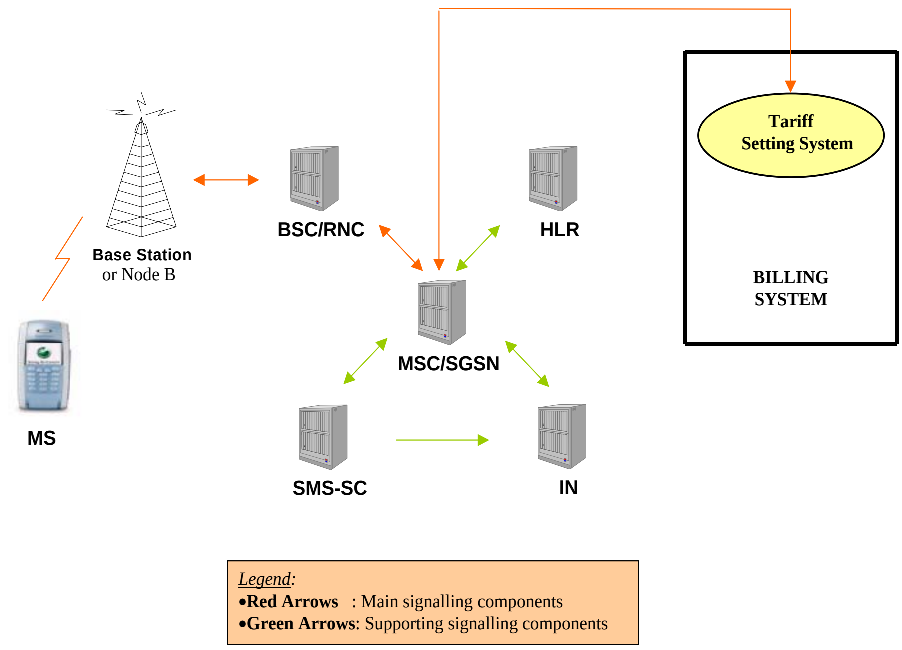


Figure 2.5: Signalling paths for Smart Pricing in mobile telecommunications networks.

onto the Secondary Common Control Physical Channel (S-CCPCH). Messages on the FACH are for intended users only and they must be readable [32]. FACH is also used to carry broadcast and multicast data [41]. With Smart Pricing, when active subscribers wish to signal their intention to pay more to increase or maintain their current QoS, the communications between the network operator and these subscribers should be private. The FACH meets this particular requirement.

2. *Paging Channel* (PCH) in association with S-CCPCH. PCH is a downlink transport channel[1] that is used to initiate the communication between the network operator and a subscriber. The network operator sends paging messages to a group of terminals belonging to a paging group which is identified by a paging indicator. Once a paging indicator is detected, the terminal decodes the next PCH frame transmitted on the S-CCPCH to see if it is intended for it [32]. With Smart Pricing, we will need to set up a communication with a subscriber who wishes to negotiate the QoS of their connection. The function that the PCH

[1]Transport channels are services offered by the physical layer to higher layers. They are capable of being mapped onto physical channels. A transport channel is defined by how and with what characteristics data is transferred over the air interface [42].

together with the S-CCPCH offers matches that requirement.

3. *Broadcast Channel* (BCH). The BCH could be used because it is a downlink transport channel used to transmit information specific to the UMTS radio access network of a given cell. It uses high power to reach all the users in an intended area and contents of the messages are changeable [32]. The BCH is a transport channel that a logical channel, the so-called BCCH, is mapped onto [41]. With Smart Pricing, we need to transmit possibly different price messages to subscribers in each and every cell, with all the subscribers in the cells being able to receive them. The prices are changed from time-to-time so we need the flexibility to alter the contents of the signalling messages. The BCH offers functions matched to Smart Pricing requirements.

4. *Channel Stealing*. Channel Stealing is a phenomenon that occurs when a burst of voice or data is discarded to make capacity within a channel available for other purposes such as a control signalling burst. The loss of voice or data's quality will not be noticeable because the channel coding scheme will make up for it with its redundancy bits [27][28][1]. An example for this is the *Fast Associated Control Channel* (FACCH) in GSM, where the user information is precluded for the transmission of handover orders [28]. The equivalent FACCH in UMTS system is the *Dedicated Transport Channel* [1]. It would be possible to use Channel Stealing to send new price messages to subscribers.

5. *Bit Stealing*. Bit Stealing is a technique used in *Digital Signalling* in the early days before SS7 was deployed. It steals one bit in the voice bit steam and uses it for signalling purpose [21]. With Smart Pricing, there will be cases in which the network operator only needs to give an indication of positive or negative acknowledgement to a subscriber. Given this requirement, we could use the Bit Stealing to send a Yes-or-No response.

6. *Broadcast SMS*. Broadcast SMS is a higher layer signalling method by which information is sent, in a text message, from the network operator to a group or all subscribers in a given area. This method is currently in use, thus it is also a viable option.

7. *Unstructured Supplementary Service Data* (USSD) is a higher layer signalling method used by the network operator to send an unstructured string to a MS. The network operator may explicitly indicate that a response is required and the next input string from the subscriber will be treated as a response [43]. With Smart Pricing, the network operator may use the USSD to send a price message to and then also receive a response from a subscriber.

2.8.2 The uplink path

The potential signalling means for the uplink path, i.e. from the MS to the TSS, are by using the:

1. *Dedicated Channel* (DCH). The DCH is an uplink (or downlink) transport channel and is mapped onto the uplink Dedicated Physical Data Channel (DPDCH) [42]. As dedicated channels use inherent addressing of the UE, every subscriber who is in communication with the network using such a channel will be uniquely identified. This is an important means for Smart Pricing in regard to, for instance, a subscriber submitting a bid. The DCH is transmitted over the entire cell or over only a part of the cell.

2. *Uplink Control Physical Channel* (CPCH). A terminal can reserve several frames in the Uplink CPCH [1]. With Smart Pricing, a subscriber may request the network operator to have their current QoS maintained when the network is becoming more congested. The amount of data needed to send this request requires a few frames and we may use the uplink CPCH for this purpose.

3. *Random Access Channel* (RACH). RACH is an uplink channel which is purposely used to carry one or two frames of control information from a terminal [1]. With Smart Pricing, we may use the RACH, similarly to the Uplink CPCH, to send shorter requests from subscribers.

4. *Channel Stealing*. Similarly to the case in the downlink path, we may use this method to send subscriber request or response information to the network operator.

5. *Bit Stealing*. Similarly to the case in the downlink path, we may use this method to send a Yes-or-No response from a subscriber to the network operator as to whether or not they accept the new price to continue with their call.

6. *USSD*. Apart from the ability to send a response to the network operator when indicated, we may use the USSD to send a request to maintain the current QoS from a subscriber to the network operator. To initiate a request, a subscriber enters a *Man Machine Interface* (MMI) string that should be treated as USSD and sends it to the network operator to initiate an USSD operation [43]. The USSD operation here will be a Smart Pricing related operation; such as request for a QoS increase, which needs to be programmed. An example of the MMI string to initiate a USSD is: *SC*SI# [44].

2.9 Uplink load and capacity of WCDMA systems

In this thesis, we consider the uplink capacity only. The uplink is more tightly constrained than the downlink in *Code Division Multiple Access* (CDMA) systems [45] and sets capacity limits for symmetric and uplink-driven services. However, it should be noted that emerging 3G services include significant downlink-heavy asymmetric services which would need to be considered in a complete analysis. Nevertheless, with no difficulties or loss of generality, our findings for the uplink case can be extended to the downlink.

Like in other CDMA systems, in WCDMA, all users are transmitting in the same radio frequency band. Individual users are separated from each other by orthogonal codes. The performance of a CDMA system is interference limited, meaning that the capacity and quality of the system are limited by the amount of interference in the radio frequency band [46]. Factors that the actual capacity of a CDMA cell is dependent on are [46] [47]:

- receiver demodulation;
- power control accuracy;
- interference power from other users in the cell;
- interference power from users in other cells;
- interference from other sources; and
- background (thermal) noise.

In WCDMA systems, the wideband total received power, I_{total}, can be measured in Node B. It is given by [32] [47]:

$$I_{total} = I_{own} + I_{oth} + N \tag{2.1}$$

where I_{own} is the received power from users in the own cell. Interference power from users in the cell constitutes this value. I_{oth} is the received power from users in the surrounding cells. This is the interference power from users in other cells. N is the total noise power. Background noise, receiver noise, and interference from other sources constitute N.

I_{own} is given by:

$$I_{own} = \sum_{j=1}^{M} P_j \tag{2.2}$$

where P_j is the received signal of user j, and M is the total number of users in the cell.

To find the required P_j at the Node B where a user is connected to, the *energy per bit to noise power density ratio* (E_b/N_0) is used [47]. E_b/N_0 is a link metric in digital communication [46]. E_b/N_0 is related to *signal to noise ratio* (S/R) by:

$$\frac{E_b}{N_0} = \frac{W}{R} \cdot \frac{S}{N} \tag{2.3}$$

where W is *bandwidth*, R is *bit rate*, and S is the *average modulating signal power*.

E_b/N_0 of a user j, $(E_b/N_0)_j$, is given by [32]:

$$\left(\frac{E_b}{N_0}\right)_j = \frac{W}{\upsilon_j \cdot R_j} \cdot \frac{P_j}{I_{total} - P_j} \tag{2.4}$$

where υ_j is the *activity factor* of user j, R_j is bit rate of user j and

$$\frac{W}{\upsilon_j \cdot R_j} = \text{Processing gain of user } j \tag{2.5}$$

A typical value of υ for data call is 1 [32], meaning that the user of that data call is assumed to transmit 100% of the time (i.e. non-stop activity). For voice call, υ is 0.4–0.5 [46] or 0.67 [32]; which is less than that of a data call. The reason is that during a voice call, the user does not speak 100% of the time. When the user does not speak, the vocoder output rate decreases. This, in turn, gives rise to interference reduction [46].

Rearranging Equation (2.4), we have:

$$P_j = \frac{1}{1 + \dfrac{W}{(E_b/N_0)_j \cdot \upsilon_j \cdot R_j}} \cdot I_{total} \tag{2.6}$$

which is equivalent to:

$$P_j = L_j \cdot I_{total} \tag{2.7}$$

where the *load factor*, L, of user j, L_j, is defined to be:

$$L_j = \frac{1}{1 + \dfrac{W}{(E_b/N_0)_j \cdot \upsilon_j \cdot R_j}} \tag{2.8}$$

Summing for all M users in the cell, Equation (2.7) becomes:

$$\sum_{j=1}^{M} P_j = \left(\sum_{j=1}^{M} L_j \right) \cdot I_{total} \tag{2.9}$$

Surrounding one cell, usually there are other cells, each of which serves its own users. Transmitting power from users of these surrounding cells also interferes with transmissions in the cell, even though power control is used in those surrounding cells [46]. We say the cell is *loaded* by users of the surrounding cells. The I_{oth} included in Equation (2.1) is to account for this effect. In an ideal situation, I_{oth} can be assumed to be directly proportional to I_{own} [47]. That is:

$$I_{oth} = i \cdot I_{own} \tag{2.10}$$

where i is called the *loading factor* [46] or *other cell to own cell interference ratio* [47]. The range of values of i is 0 to 1 [46].

Substitute Equation (2.10) into Equation (2.1), we have:

$$I_{total} = (1 + i)I_{own} + N \tag{2.11}$$

Rearranging Equation (2.11), we have:

$$N = I_{total} - (1 + i)I_{own} \tag{2.12}$$

The *Noise Rise* (NR) is defined as:

$$NR = \frac{I_{total}}{N} \tag{2.13}$$

Substitute Equation (2.12) into Equation (2.13), we have:

$$NR = \frac{I_{total}}{I_{total} - (1 + i)I_{own}} \tag{2.14}$$

Substitute Equation (2.2) into Equation (2.14), we have:

$$NR = \frac{I_{total}}{I_{total} - (1+i)\sum_{j=1}^{M} P_j} \tag{2.15}$$

Substitute Equation (2.9) into Equation (2.15), we have:

$$NR = \frac{I_{total}}{I_{total} - (1+i) \cdot I_{total} \cdot \sum_{j=1}^{M} L_j} \tag{2.16}$$

Simplify Equation (2.16), we have:

$$NR = \frac{1}{1 - (1+i)\sum_{j=1}^{M} L_j} \tag{2.17}$$

$$NR = \frac{1}{1 - \eta_{UL}} \tag{2.18}$$

where η_{UL} is the *uplink load factor*, and:

$$\eta_{UL} = (1+i) \sum_{j=1}^{M} L_j \tag{2.19}$$

It can be seen from Equation (2.18) that NR approaches infinity if η_{UL} is equal to 1. This is also when the system reaches its *pole capacity* [32]. Hence, values of η_{UL} are in the $(0, 1)$ range.

Rearranging Equation (2.19), we have:

$$\sum_{j=1}^{M} L_j = \frac{\eta_{UL}}{(1+i)} \tag{2.20}$$

Rearranging Equation (2.18), we have:

$$\eta_{UL} = \frac{NR - 1}{NR} \tag{2.21}$$

The η_{UL} represents the loading of a cell. If high loading is allowed, the coverage of that cell will shrink. On the other hand, allowing high loading means increasing capacity for the cell. This means cell coverage and cell capacity have an inverse relationship. Depending on a certain situation, a network operator may opt for large coverage or for high capacity. To compensate for the loss of signal strength due to *Rayleigh* and *log-normal* fading in order for a received signal to be above the receiver threshold, an *interference margin* in the link budget is used. This interference margin ensures that a target size of a cell coverage area is met. A small interference margin should be used when large coverage area is desired, whereas a large interference margin when high capacity is desired. An interference margin in the link budget must be equal to the maximum planned NR [32].

NR is usually given in unit of dB; i.e. its equivalent linear value is given by:

$$NR = 10^{\left(\frac{NR\ [\text{dB}]}{10}\right)} \tag{2.22}$$

Thus, for a given NR [dB] value, the procedure to find the maximum possible load factor allowed in a cell from all users, η_M, is to:

1. use Equation (2.22) to convert to NR in linear value;
2. substitute NR in linear value into Equation (2.21) to find η_{UL};
3. substitute the η_{UL} into Equation (2.20) to find the maximum possible load factor for all users on the uplink of a cell; and
4. use the above maximum possible load factor as the upper-bound for the sum of load factor values of all M users in the cell.

The load factor of each individual user is determined by using Equation (2.8). Depending on the $(E_b/N_0)_j$, υ_j and R_j parameter values of each user j, the maximum value for M is different. If all users have the same values for those parameter values, the capacity of the cell, M, is given by:

$$M = \frac{\text{Maximum possible load factor from all users}}{L_u} \tag{2.23}$$

where L_u is the load of a user, each of whom has the same E_b/N_0, υ and R parameter values.

CHAPTER 3

Modelling Smart Pricing signalling

3.1 The need for a new network element

To implement Smart Pricing, the loads of all radio covered areas in a network must be reported to the network operator. Depending on this information and responses from subscribers to indications of congestion, a decision will be made by the network operator whether or not to adjust the current prices. The smallest covered area where a MS can uniquely identify the received radio signal from a transmitter is a *cell* [25]. In UMTS, the radio transmitting and receiving functions of a cell are facilitated by *Node B* (equivalent to BTS in a GSM system). Node B, however, does not have the required processing power for resource management. That resides in the next network element in the signalling path from the MS to the Billing System, i.e. the RNC (equivalent to BSC in a GSM system). Therefore, the RNC is the first network element directly responsible for and able to report the congestion level of a cell. Moreover, with soft handover and transmission diversity, multiple Node Bs may be involved in a single connection and since one RNC controls multiple Node Bs, it is able to report precisely the congestion level of every cell. In Section 2.4, we have explained that the TSS in the Billing System is responsible for setting new prices and QoS levels dependent on congestion levels. For the TSS to do that, there must be a network element to receive congestion information for a cell sent from the RNC, to process that information and to decide whether new prices and QoS levels are necessary. If it is deemed necessary, that network element will send a request to the TSS. The TSS will act upon that request and set new prices and QoS levels. In conventional mobile telecommunications systems (see Figure 2.3 for evidence), no such a network element currently exists. Therefore, to bring Smart Pricing into practice, a new network element must be created and undertakes the above-mentioned responsibilities.

Once this new network element is designated, we will be able to see all necessary signalling components for Smart Pricing. Location of this element must be carefully chosen as it will have significant impact on the amount of signalling traffic required.

3.2 Dynamic pricing engine

In Section 3.1, we have identified the need for a new network element. Here, we propose *Dynamic Pricing Engine* (DPE) as that new network element. The DPE can be collocated with the RNC and linked to the TSS. The TSS is to be a distributed function of the DPE.

The DPE is responsible for receiving congestion information of a cell from the RNC as well as for processing that information. Contingent on seeing the need for new prices and QoS levels, the DPE will request the TSS to set new prices. The TSS will act upon such a request, sending information about new prices and QoS levels back to the DPE. The DPE will then relay the new price messages to all active MSs in the cell. The DPE will also need to send the new price messages to the MSC and SGSN as these network elements need that information for generating CDRs. Finally, the DPE needs to instruct the MSC and SGSN to set new QoS levels for MSs' calls, according the information received from the TSS.

Other functions of the DPE are: to communicate with users' responses to new prices, to record bids for maintaining QoS, and to communicate with the MSC to change QoS of users' call in relation to the number of bids. The DPE may also make strategic decisions to vary prices for maximising network revenue. An example of such a decision is to lower the prices when the network is very lightly loaded to encourage subscribers to use the network more.

While being able to reduce the number of signalling links is important, finding a location for the DPE which prevents network elements being overloaded by the signalling traffic is equally important. In this model, although we propose that the DPE be a stand-alone element, it can always be collocated with the RNC or MSC. Having the DPE as a stand-alone element would help us to investigate signalling for Smart Pricing in the worst case scenario. Between the RNC and MSC, we believe that the DPE should be collocated with the RNC because this prevents the MSC from possible heavy signalling traffic imposed by the large number of active users in the cells that the MSC controls.

In Figure 3.1, we incorporate the DPE into the signalling paths shown earlier in Figure 2.5. The complete set of required signalling links for Smart Pricing are now shown. The proposed protocols for the link from the DPE to the:

- RNC, MSC and SGSN is MTP (SS7); and
- HLR is TCAP.

Other required signalling links and protocols are as presented in Figure 2.3 and Figure 2.4.

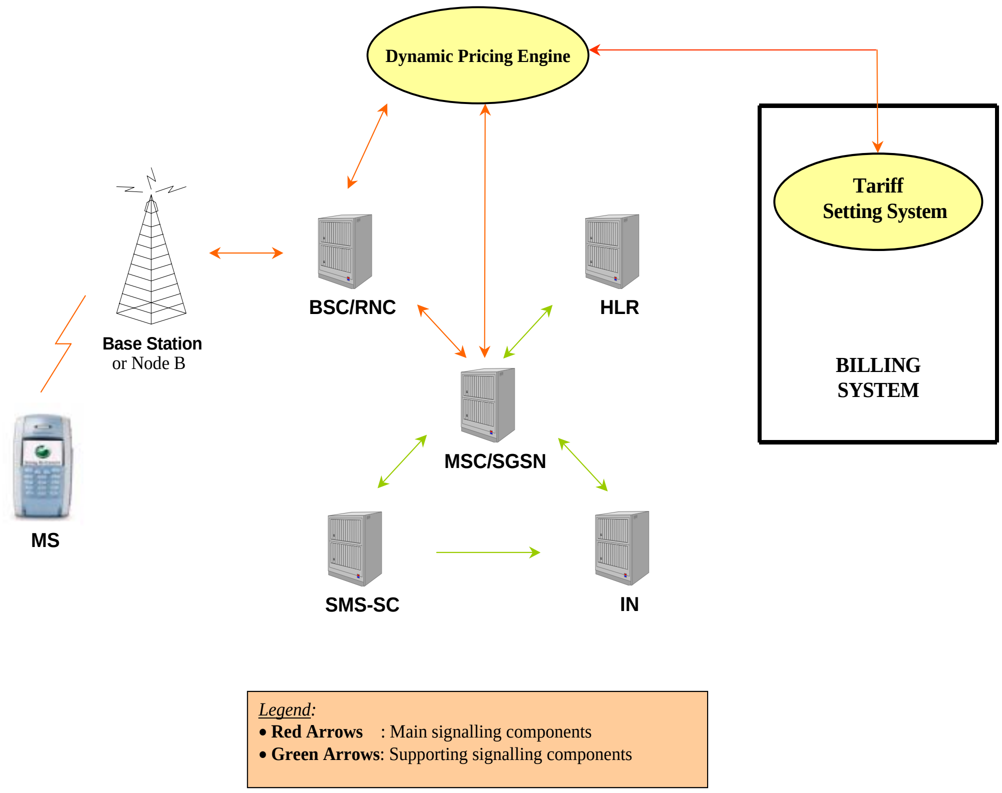


Figure 3.1: Complete signalling paths for Smart Pricing.

3.3 System diagram

For the sake of simplicity for modelling purposes, we will only consider the case where Smart Pricing is applied to conversational QoS class calls only. We also assume that all users subscribe to only post-pay plans, meaning that the SMS-SC and the IN network elements required to accommodate pre-paid services can be omitted. The system diagram in Figure 3.2 reflects these assumptions.

3.4 Smart Pricing signalling algorithm

We propose a signalling algorithm for Smart Pricing as the following:

"*When a new user arrives at the network, depending on the level of congestion of the cell, an appropriate admission price is sent to the user. The user has to send his/her acceptance of the price before the call is connected. Once the user begins his/her call,*

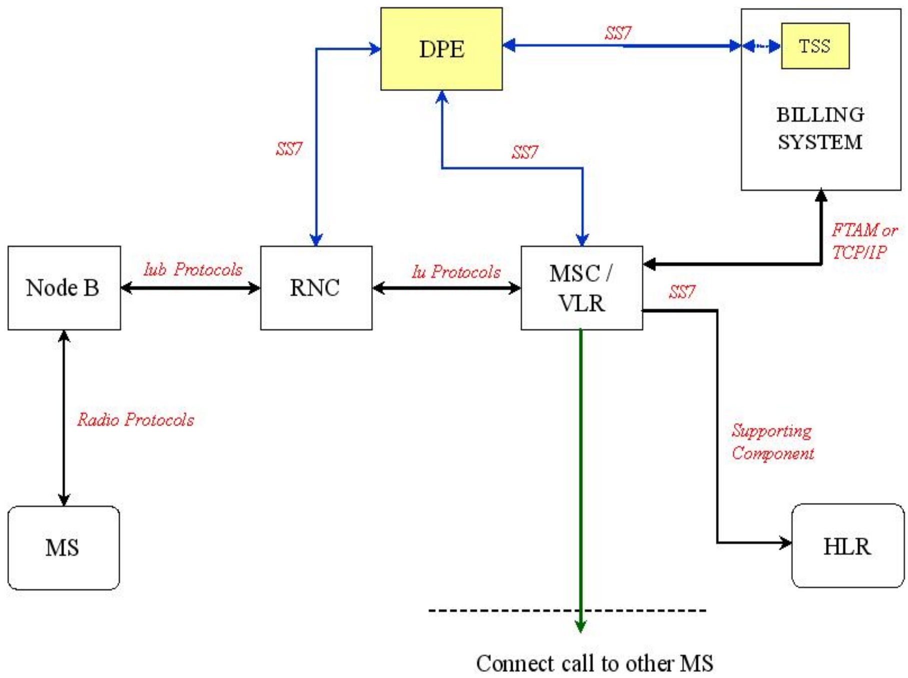


Figure 3.2: Smart Pricing's System Diagram.

the load of the cell is checked to see if the cell then becomes congested. If congested, active users whose guaranteed period of QoS and price have expired must now bid up to prevent degradation of the QoS of their calls. When the guaranteed period of QoS and price of a user's call expires or when a call finishes, the load of the cell is also checked to see if the network then becomes less congested. Upon the result of that check, prices and QoS levels may be recalculated and assigned to users who have their guaranteed period of QoS and price expired."

We illustrate the proposed Smart Pricing signalling algorithm by a flowchart in Figure 3.3. In the flowchart, steps in green are those of a normal call process; the others are Smart Pricing signalling steps. It can seen from the flowchart that the DPE plays a very important role in Smart Pricing. It has control from a call setup to when the call finishes. The DPE ensures that users pay appropriate prices for their calls at any instance of time, and that those prices always shadow the congestion cost imposed on the network by the users.

In Figure 3.3,

- t_{gQ} is the guaranteed period of QoS and price for a call;
- η_T is a point such that if the cell load exceeds the cell is deemed to be experiencing congestion. This is the point of load discussed earlier in Section 2.1; and
- η_M is the maximum possible load factor allowed in a cell from all users as discussed in Section 2.9.

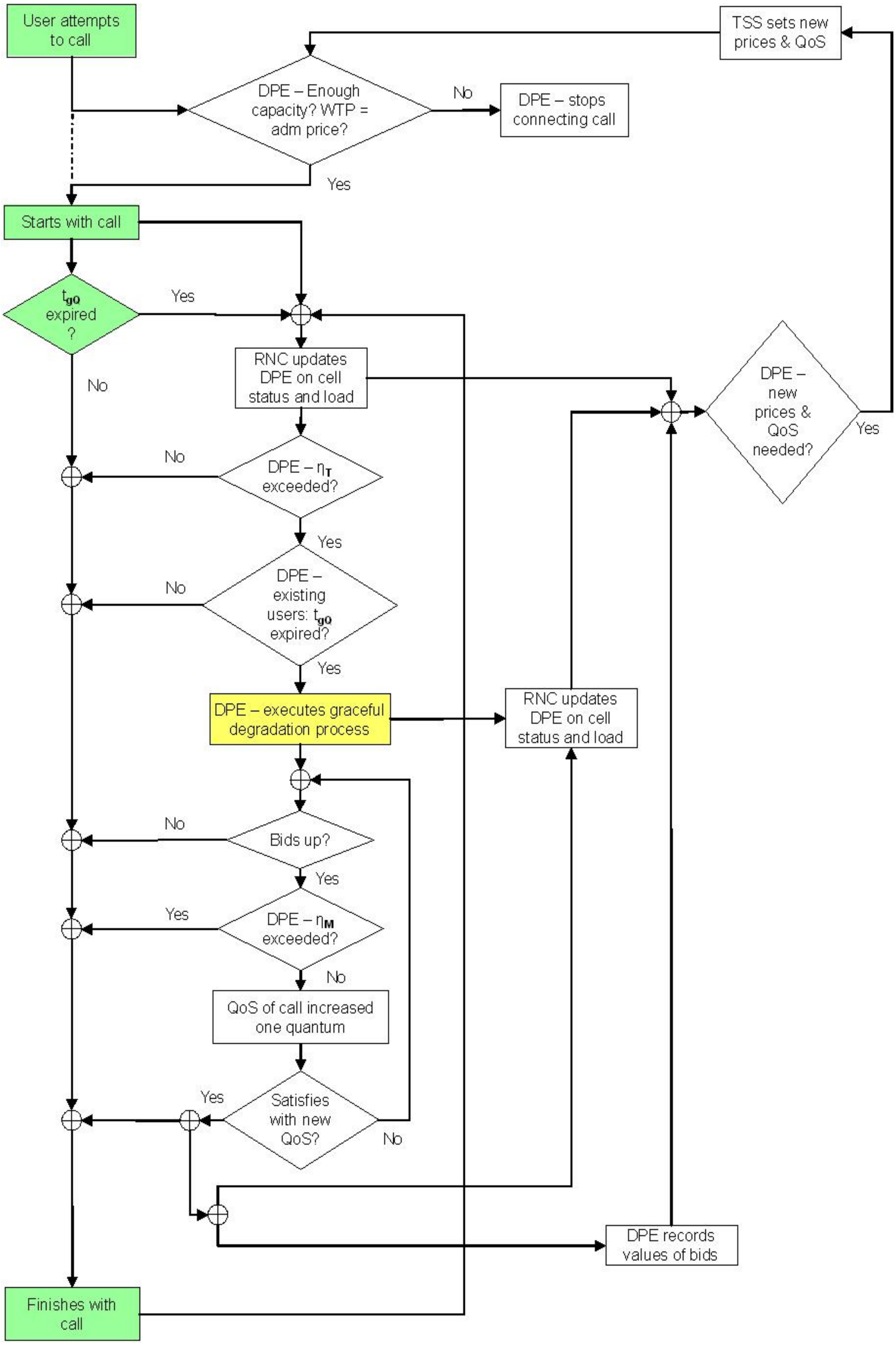


Figure 3.3: Smart Pricing signalling algorithm.

3.5 Modelling Smart Pricing

3.5.1 Smart Pricing scheme

As Smart Pricing is a type of congestion pricing, we need to be able to tell when a cell is congested so that graceful degradation can be enforced. In Section 3.4, we developed a means to do this by defining η_T. We will now call η_T the *Low Load Threshold*.

A network operator can set η_T at their desired level, relative to η_M, by using the Equation 3.1 below:

$$\eta_T = \eta_M \cdot \zeta_T \tag{3.1}$$

where ζ_T is the *Low Load Threshold Factor*. By specifying a value in the [0,1] range for ζ_T, η_T is obtained. Figure 3.4 illustrates position of η_T relative to η_M, as well as to other load factors.

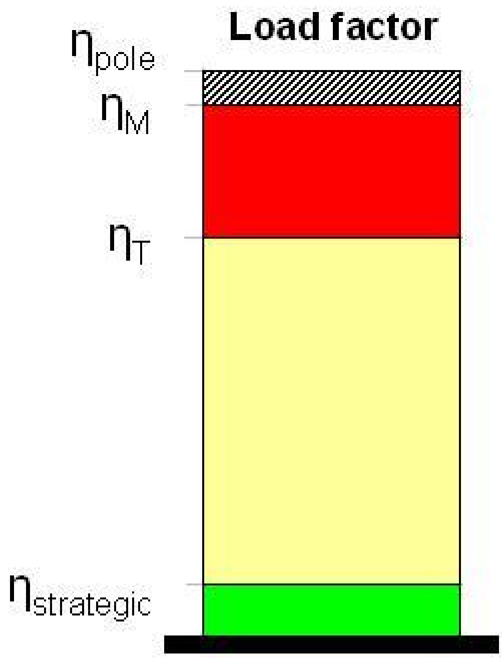


Figure 3.4: Illustration of cell load factors.

In Figure 3.4, $\eta_{strategic}$ is the load factor below which a network operator may make a strategic decision to increase usage, as discussed in Section 3.2. Value for this parameter is set at the network operator's discretion. η_{pole} is the pole capacity of the cell, as discussed in Section 2.9.

Smart Pricing uses degradation in QoS as an indication to users of network congestion. By QoS, we mean such metrics as throughput, *Bit Error Rate* (BER) or response time as appropriate to the application. We will, however, limit our analysis to bit rate. That is, in this thesis, we consider a reduction (or an increase) in QoS a reduction (or an increase) in bit rate. This is because, referring to Equation (2.8) in Section 2.9, if W and υ_j are kept unchanged, reducing (or increasing) L_j by an amount can be achieved by either reducing (or increasing) $(E_b/N_0)_j$ or, equivalently, reducing (or increasing) R_j. For $(E_b/N_0)_j$, reducing (or increasing) its value means reducing (or increasing) *BER*.

Under Smart Pricing, a user is admitted to the network only if:

- the highest bit rate (i.e. QoS) for the service, a call of which the user requests to use, can be offered; and
- the user's WTP is equal to or greater than the admission price.

Otherwise, the user will not be admitted to the network. That highest bit rate is guaranteed for a period t_{gQ} set by the network operator. For every service, t_{gQ} can be set differently, however, for simplicity we assume that it is the same for all services.

Graceful degradation

If the load of a cell is above η_T, a *graceful degradation* process will take place. Graceful degradation is a process during which all users with expired t_{gQ} will have bit rates of their calls reduced towards the minimum bit rates relevant to the services they are using. The minimum bit rate for every service must be set so that an acceptable quality is still guaranteed. Prices of calls will also be reduced in a similar fashion during the graceful degradation process. Users who have bit rates of their calls reduced must indicate to the network operator that they are willing to pay more for increased bit rates. The exact new bit rates such users will get are dependent on how much more (relative to the current prices) the users are willing to pay. As Smart Pricing does not posit a simple relationship between price and load, the price for an increased bit rate is not fixed. The group of users who have t_{gQ} expired will have to compete with each other for an increased bit rate. We model this competition by a bidding process. Users give indications that they are willing to pay more for increased bit rates by submitting bids to the DPE. The highest bidders will get the highest increased bit rate, which is almost always equal to the admission bit rate relevant to that service. The value of the highest bid will be used to establish admission price for new entrants.

Users who bid lower than the highest bidders may still get increased bit rates but to a lesser extent. This, however, will only be possible if the network still has enough capacity to accommodate, after giving the highest bit rates to the highest bidders. In the best case scenario, every user who bids will get an increased bit rate appropriately to the number of bids he/she submits. That means, users with the number of bids equal to the highest minus one bid will get their bit rate increased one quantum lower than that of those highest bidders. Users with the number of bids equal to the highest less two bids will get their bit rate increased two quanta lower than that of those highest bidders, and so forth.

We assume that services are prioritised, meaning that each service has a priority level attached to it, e.g. video conferencing calls have higher priority than voice calls. Capacity of a cell is first used to accommodate the increased bit rate for calls of the highest bidders of the highest prioritised service. Next, the remaining capacity is used to accommodate the increased bit rate for calls of the highest bidders of a lower prioritised service. And so on, until increased bit rates for calls of the highest bidders of all services are accommodated. Then, requests for bit rate increases from users who submit less bids than the highest bidders are handled in the same fashion as that for the highest bidders. This process continues until all requests for bit rate increases are processed, or it will stop when assigning an increased bit rate results in the cell load exceeding η_M. In summary, with two services offered, the order of assigning increased bit rates is:

- highest prioritised service, highest bidders;
- lower prioritised service, highest bidders;
- highest prioritised service, lower bidders; and
- lower prioritised service, lower bidders;
- etc.

Events that triggers graceful degradation

There are three events that triggers graceful degradation, if cell load after any occurance of the events above η_T. The events are:

1. admitting a new user;
2. t_{gQ} of a user's call expires; and
3. a user hangs up.

3.5.2 User

We assume that:

- the number of users that arrive in the network every *unit of time*, hereafter assumed to be second [s], follows a Poisson distribution with mean λ [arrivals/s]. From this *arrival rate* λ we can obtain arrival time for each user. Let the arrival time of a user be $\psi_{AT}[s]$;
- holding time of a user's call is negative exponentially distributed with mean $\overline{ht}$ equal to $1/\mu$ [s], where μ is the *service rate* or *rate at which calls currently in progress end*. Let the hold time of a user be ψ_{HT} [s]; and
- before making a call, a user has already decided a price (*currency* per unit of time, hereafter assumed to be [$/s]) at which the user is willing to pay for his/her call. This is what we have been referring to as WTP [$/s]. We assume that WTP follows a *Weibull* distribution with *scale parameter* δ and *shape parameter* β. WTP is also the maximum price level that the user will bid up to for a bit rate increase every time the user experiences graceful degradation. Let the WTP for a call of a user be ψ_{Γ} [$/s].

Every user is willing to pay to his/her call at a different price. Let ρ be the number of different WTP levels for a service S_i call. A set of WTP levels in ascending order, $\boldsymbol{\Gamma_i}$, is:

$$\boldsymbol{\Gamma_i} = \{\Gamma_{ij};\ j = 1, ..., \rho\} \tag{3.2}$$

where Γ_{i1} is the minimum and $\Gamma_{i\rho}$ the maximum WTP for service S_i. ψ_{Γ} takes values from $\boldsymbol{\Gamma_i}$. Note that not all the time users with $\Gamma_{i\rho}$ are in the system.

During a graceful degradation process, users with expired t_{gQ} must bid to avoid the bit rates of their calls being reduced. We assume that every user will always bid up to his/her WTP or to one bid higher than the immediate lower bidders, whichever is lower.

Let the time at which t_{gQ} of a user's call expires be ψ_{tgQ} [s]. We then have:

$$\psi_{tgQ} = \psi_{AT} + t_{gQ} \tag{3.3}$$

Let the service of a user's call be ψ_S and a set of ρ possible bit rates (one bit rate corresponds to one WTP level) available for service S_i in ascending order, $\boldsymbol{R_i}$, be:

$$\boldsymbol{R_i} = \{R_{ij};\ j = 1, ..., \rho\} \tag{3.4}$$

where R_{i1} is the minimum and $R_{i\rho}$ the maximum bit rate for service S_i.

It is necessary to have one bit rate for each level of WTP because there are cases where a network operator needs to increase bit rates at different quanta according to the numbers of bids that users submit. However, a user with WTP at level j of service i may not get bit rate R_{ij}, but R_{ix} (where $x < j$). One example of such an occurrence might be after giving $R_{i\rho}$ to the highest bidders, the network only had enough capacity to give R_{i2} to users that bid up to WTP level $(\rho - 1)$, then those bidders will only get bit rate increase to R_{i2}, not $R_{i(\rho-1)}$. All users with WTP levels less than $(\rho - 1)$ will not be successful in bidding for a bit rate increase. These bidders together with users that do not bid will get R_{i1}.

From the above, clearly users with different WTP levels could be assigned the same bit rate. Let the vector that stores assigned bit rates (one for each WTP level) for users of service S_i, $\boldsymbol{R_i^a}$, be:

$$\boldsymbol{R_i^a} = \begin{bmatrix} R_{i1}^a & R_{i2}^a & \cdots & R_{i\rho}^a \end{bmatrix} \tag{3.5}$$

where each R_{ij}^a can take any subset of values in $\boldsymbol{R_i}$, and more than one R_{ij}^a elements can have the same value. Highest bidders will almost always get $R_{i\rho}$. Highest bidders are not always those with $\Gamma_{i\rho}$ though, as not all the time these users are in the system. The bit rates in $\boldsymbol{R_i^a}$ are for users with expired t_{gQ} only. Bit rate for users who are still in their t_{gQ}, $R_i^{a,g}$, is admission bit rate and will always be $R_{i\rho}$. That is:

$$R_i^{a,g} = R_{i\rho} \tag{3.6}$$

Let the current bit rate of a user's call be ψ_R [bps]. ψ_R will be equal to $R_{i\rho}$ if the user is still his/her t_{gQ} and equal to an element in $\boldsymbol{R_i^a}$ with index equal to the user's WTP when his/her t_{gQ} expires.

With R_{ij} together with $(E_b/N_0)_j$ and υ_j values, using Equation (2.8), the relevant user load factor, L_{ij}, can be determined. The set of user load factors (one for each possible bit rate) in ascending order for service S_i, $\boldsymbol{L_i}$, will then be:

$$\boldsymbol{L_i} = \{L_{ij};\ j = 1, ..., \rho\} \tag{3.7}$$

Let the set of ρ possible prices for service S_i (one for each possible bit rate) in ascending order, $\boldsymbol{P_i}$, be:

$$\boldsymbol{P_i} = \{P_{ij};\ j = 1, ..., \rho\} \tag{3.8}$$

where P_{i1} is the minimum and $P_{i\rho}$ the maximum price for service S_i.

Let the vector of prices for assigned bit rates (one for each assigned bit rate, which also is one for each WTP level) in vector $\boldsymbol{R_i^a}$, $\boldsymbol{P_i^a}$, be:

$$\boldsymbol{P_i^a} = \begin{bmatrix} P_{i1}^a & P_{i2}^a & \cdots & P_{i\rho}^a \end{bmatrix} \tag{3.9}$$

where each P_{ij}^a can take any values in $\boldsymbol{P_i}$, and more than one P_{ij}^a elements can have the same value. Assigned prices in $\boldsymbol{P_i^a}$ are for users with expired t_{gQ}. Assigned prices for users who are still in their t_{gQ} are their admission prices. Admission price of each user may differ from all other users, even when they have the same WTP.

Let the current price that a user has to pay for his/her call be ψ_P [$/s]. ψ_P will be equal to the user's admission price if the user is still in his/her t_{gQ} and equal to an element in $\boldsymbol{P_i^a}$ with index equal to the user's WTP when his/her t_{gQ} expires.

Let the vector of numbers of users (one for each assigned bit rate, which also is one for each WTP level) of service S_i, $\boldsymbol{U_i}$, be:

$$\boldsymbol{U_i} = \begin{bmatrix} U_{i1} & U_{i2} & \cdots & U_{i\rho} \end{bmatrix} \tag{3.10}$$

The numbers of users in $\boldsymbol{U_i}$ are the numbers of users with expired t_{gQ} only. A vector that stores the numbers of users (one for each WTP level) who are still in t_{gQ}, $\boldsymbol{U_i^g}$, is:

$$\boldsymbol{U_i^g} = \begin{bmatrix} U_{i1}^g & U_{i2}^g & \cdots & U_{i\rho}^g \end{bmatrix} \tag{3.11}$$

In summary, attributes associated to a user are:

1. arrival time: ψ_{AT} [s];
2. hold time: ψ_{HT} [s];
3. service of the call: ψ_S;
4. WTP for the call: ψ_Γ [$/s];
5. current bit rate of the call: ψ_R [bps];
6. current price of the call: ψ_P [$/s]; and

7. time t_{gQ} expires of the call: ψ_{tgQ} [s].

3.5.3 Network

Let the number of services offered by a network be α. The set of services, $\boldsymbol{S}$, will then be:

$$\boldsymbol{S} = \{S_i;\ i = 1, ..., \alpha\} \tag{3.12}$$

Although in practice the number of WTP levels for each service may be different, we assume they are the same. That is, there are exactly ρ WTP levels for every service S_i. Thus, the matrix that stores WTP levels for all services, $\mathbf{\Phi_\Gamma}$, is:

$$\mathbf{\Phi_\Gamma} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} & \cdots & \Gamma_{1\rho} \\ \Gamma_{21} & \Gamma_{22} & \cdots & \Gamma_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{\alpha 1} & \Gamma_{\alpha 2} & \cdots & \Gamma_{\alpha\rho} \end{bmatrix} \tag{3.13}$$

The matrix that stores bit rates for every WTP for each and every service, $\mathbf{\Phi}_{\boldsymbol{R}}$, is:

$$\mathbf{\Phi}_{\boldsymbol{R}} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1\rho} \\ R_{21} & R_{22} & \cdots & R_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ R_{\alpha 1} & R_{\alpha 2} & \cdots & R_{\alpha\rho} \end{bmatrix} \tag{3.14}$$

The matrix that stores assigned bit rates for t_{gQ} expired users for all services, $\mathbf{\Phi}_{\boldsymbol{R^a}}$, is:

$$\mathbf{\Phi}_{\boldsymbol{R^a}} = \begin{bmatrix} R^a_{11} & R^a_{12} & \cdots & R^a_{1\rho} \\ R^a_{21} & R^a_{22} & \cdots & R^a_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ R^a_{\alpha 1} & R^a_{\alpha 2} & \cdots & R^a_{\alpha\rho} \end{bmatrix} \tag{3.15}$$

The column vector that stores assigned bit rates for users who are still in their t_{gQ} (i.e. admission bit rates) for all services, $\mathbf{\Phi}_{\boldsymbol{R^{a,g}}}$, is:

$$\mathbf{\Phi}_{\boldsymbol{R^{a,g}}} = \begin{bmatrix} R^{a,g}_{1\rho} & R^{a,g}_{2\rho} & \cdots & R^{a,g}_{\alpha\rho} \end{bmatrix}^T \tag{3.16}$$

The matrix that stores individual user load factors for every possible bit rate for all services, $\boldsymbol{\Phi_L}$, is:

$$\boldsymbol{\Phi_L} = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1\rho} \\ L_{21} & L_{22} & \cdots & L_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ L_{\alpha 1} & L_{\alpha 2} & \cdots & L_{\alpha\rho} \end{bmatrix} \tag{3.17}$$

The matrix that stores prices for every possible bit rate for every service, $\boldsymbol{\Phi_P}$, is:

$$\boldsymbol{\Phi_P} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1\rho} \\ P_{21} & P_{22} & \cdots & P_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ P_{\alpha 1} & P_{\alpha 2} & \cdots & P_{\alpha\rho} \end{bmatrix} \tag{3.18}$$

The matrix that stores assigned prices for t_{gQ} expired users for all services, $\boldsymbol{\Phi_{P^a}}$, is:

$$\boldsymbol{\Phi_{P^a}} = \begin{bmatrix} P^a_{11} & P^a_{12} & \cdots & P^a_{1\rho} \\ P^a_{21} & P^a_{22} & \cdots & P^a_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ P^a_{\alpha 1} & P^a_{\alpha 2} & \cdots & P^a_{\alpha\rho} \end{bmatrix} \tag{3.19}$$

The matrix that stores the numbers of users whose t_{gQ}'s expire for all services, $\boldsymbol{\Phi_U}$, is:

$$\boldsymbol{\Phi_U} = \begin{bmatrix} U_{11} & U_{12} & \cdots & U_{1\rho} \\ U_{21} & U_{22} & \cdots & U_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ U_{\alpha 1} & U_{\alpha 2} & \cdots & U_{\alpha\rho} \end{bmatrix} \tag{3.20}$$

The matrix that stores the numbers of users who are still in t_{gQ}, $\boldsymbol{\Phi^g_U}$, is:

$$\boldsymbol{\Phi_{U^g}} = \begin{bmatrix} U^g_{11} & U^g_{12} & \cdots & U^g_{1\rho} \\ U^g_{21} & U^g_{22} & \cdots & U^g_{2\rho} \\ \vdots & \vdots & \ddots & \vdots \\ U^g_{\alpha 1} & U^g_{\alpha 2} & \cdots & U^g_{\alpha\rho} \end{bmatrix} \tag{3.21}$$

3.5.4 Calculations of potential loads

In order to demonstrate details of some calculations and graceful degradation process under Smart Pricing scheme, in this section and the next, we consider a simple case. In this simple case, for each service, $(E_b/N_0)_j$ and υ_j are kept unchanged for all users. Users are with two types of WTP levels only, i.e. Γ_{i1} = low, $\Gamma_{i\rho}$ = high, and $\rho = 2$.

From a desired NR [dB] value of a network operator, using:

- Equation (2.22) and Equation (2.21), we can calculate η_M; and
- Equation (3.1), we can calculate η_T.

For service S_i, let:

1. the number of $\Gamma_{i\rho}$ users who are still in their t_{gQ} be g_i. We call these users *g-type* users
$$g_i = U^g_{i\rho} \tag{3.22}$$
and g_i is the last element on row ith of $\mathbf{\Phi}^{\boldsymbol{g}}_{\boldsymbol{U}}$;
2. the number of Γ_{i1} users who are still in their t_{gQ} be y_i. We call these users *y-type* users
$$y_i = U^g_{i1} \tag{3.23}$$
and y_i is the first element on row ith of $\mathbf{\Phi}^{\boldsymbol{g}}_{\boldsymbol{U}}$;
3. the number of $\Gamma_{i\rho}$ users whose t_{gQ} expired be h_i. We call these users *h-type* users
$$h_i = U_{i\rho} \tag{3.24}$$
and $U_{i\rho}$ is the last element on row ith of $\mathbf{\Phi}_{\boldsymbol{U}}$;
4. the number of Γ_{i1} users whose t_{gQ} expired be l_i. We call these users *l-type* users
$$l_i = U_{i1} \tag{3.25}$$
and U_{i1} is also the first element row ith of $\mathbf{\Phi}_{\boldsymbol{U}}$; and
5. the total number of users using the service, M_i, be:
$$M_i = g_i + y_i + h_i + l_i \tag{3.26}$$

Next, we define the *minimum load condition* for service S_i a condition in which all g-, y- and h-type users of service S_i get $R_{i\rho}$, and l-type users get R_{i1}. Therefore, the load factor to guarantee the *minimum load condition* is:

$$L_i^{min} = (g_i + y_i + h_i)L_{i\rho} + (l_i)L_{i1} \tag{3.27}$$

hence, the load factor to guarantee the minimum load condition for the whole system is:

$$L^{min} = \sum_{i=1}^{\alpha} L_i^{min} \tag{3.28}$$

We also define the *maximum load condition* for service S_i a condition in which all M_i users get $R_{i\rho}$. Therefore, the load factor to enable the *maximum load condition* is:

$$L_i^{max} = M_i \cdot L_{i\rho} \tag{3.29}$$

hence, the load factor to guarantee the maximum load condition for the whole system is:

$$L^{max} = \sum_{i=1}^{\alpha} L_i^{max} \tag{3.30}$$

The potential load factor to admit a new service S_i user while guaranteeing the minimum load condition is:

$$L_i^{pot,min} = L_i^{min} + L_{i\rho} \tag{3.31}$$

and, the potential load factor to admit a new service S_i user while guaranteeing the minimum load condition for all services in the system is:

$$L^{pot,min} = L_i^{pot,min} + \sum_{\substack{j=1 \\ j \neq i}}^{\alpha} L_j^{min} \tag{3.32}$$

The potential load factor to admit a new service S_i user while guaranteeing the maximum load condition is:

$$L_i^{pot,max} = L_i^{max} + L_{i\rho} \tag{3.33}$$

and, the potential load factor to admit a new service S_i user while guaranteeing the maximum load condition for services in the system is:

$$L^{pot,max} = L_i^{pot,max} + \sum_{\substack{j=1 \\ j \neq i}}^{\alpha} L_j^{max} \tag{3.34}$$

3.6 System behaviours when cell load changes

The three events that cause the cell load to change are also those that may trigger graceful degradation as discussed in Section 3.5.1. In this section, we provide details about how the system under Smart Pricing behaves when these events occur.

3.6.1 Admitting a new user

The DPE is always able to calculate the $L^{pot,max}$ based on history information (i.e. the numbers of users in the system) and on information about the type of service of the call the new user is requesting to make. Refer to Equation (3.33) and Equation (3.34) for details. Thus, the process to obtain an admission price to advise the new user when $L^{pot,max}$ is below η_T is simple. The DPE simply requests the TSS to set the price to the minimum. No signalling with existing users in the system is required.

On the other hand, when $L^{pot,max}$ is above η_T, in order to calculate $L^{pot,min}$ to see if it exceeds η_M, the DPE needs to know the numbers of highest bidders (i.e. those with high WTP levels) for all services. Refer to Equation (3.31) and Equation (3.32) for details. This means that signalling with existing users, through a graceful degradation process, is required. Once the $L^{pot,min}$ is determined, an admission price for the new user can be obtained. As an admission price set subsequent to a graceful degradation process could be high, if it were beyond the new user's WTP, the new user will leave the network. That makes the disturbance otherwise imposed to existing users in the system unnecessary. The more low WTP users attempt to make calls in such a situation, the more the problem exacerbates, the outcome of which could give rise to customer dissatisfaction, jeopardising the practicality of Smart Pricing. A solution for this problem is for the network to require each user, upon admission to the network, to specify the upper bound of his/her WTP. This, however, will be the focus of future work. At this stage, we simply assume that the DPE is able to predict the numbers of highest bidders of all services. This assumption allows the DPE, in association with the TSS, to be able to set an admission price as in the case when $L^{pot,max}$ is below η_T.

Hence, the system handles cases when a new user requests to make a service S_i call as follows.

If $\boxed{\boldsymbol{L^{pot,max} \leq \eta_T}}$

The system:

- sets admission price to P_{i1};
- upon the new user's acceptance of the admission price (as should usually be the case because admission is at its lowest possible), sets ψ_P to P_{i1};
- sets ψ_R to R_{i2};
- admits the new user to the network, and increases by one $\boldsymbol{U_i^g}(\psi_\Gamma)$[1] and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$[2];
- sets $\boldsymbol{P_i^a}(:)$[3] and $\boldsymbol{\Phi_P^a}(i, :)$[4] to P_{i1}; and
- sets $\boldsymbol{R_i^a}(:)$ and $\boldsymbol{\Phi_R^a}(i, :)$ to R_{i2}.

Elseif $\boxed{\boldsymbol{L^{pot,max} > \eta_T}}$ **and** $\boxed{\boldsymbol{L^{pot,min} \leq \eta_M}}$

The admission price determined by the DPE in association with the TSS should be $P_{i\rho}$. If $\psi_\Gamma < P_{i\rho}$, condition of the system remains unchanged. Otherwise, if $\psi_\Gamma = P_{i\rho}$, the system:

- sets admission price to $P_{i\rho}$;
- sets ψ_P to $P_{i\rho}$;
- sets ψ_R to $R_{i\rho}$;
- admits the new user to the network, and increases by one $\boldsymbol{U_i^g}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$;
- executes graceful degradation process;
- sets $\boldsymbol{P_i^a}(\rho)$ and $\boldsymbol{\Phi_P^a}(i, \rho)$ to $P_{i\rho}$. $\boldsymbol{P_i^a}(1)$ and $\boldsymbol{\Phi_P^a}(i, 1)$ are set to P_{i1} as assigned prices for users who do not bid will remain at the initial price set during the graceful degradation process; and
- sets $\boldsymbol{R_i^a}(\rho)$ and $\boldsymbol{\Phi_R^a}(i, \rho)$ to $R_{i\rho}$. $\boldsymbol{R_i^a}(1)$ and $\boldsymbol{\Phi_R^a}(i, 1)$ are set to R_{i1} as assigned bit rate for users who do not bid will remain at the lowest bit rate set during the graceful degradation process.

---

[1] The (ψ_Γ)th element of $\boldsymbol{U_i^g}$.

[2] Element with index (i, ψ_Γ) of $\boldsymbol{\Phi_U^g}$.

[3] ":" represents all elements. $\boldsymbol{P_i^a}(:)$ is all elements of $\boldsymbol{P_i^a}$.

[4] ":" on the right of the comma represents all elements on a row specified by a value on the left of the comma. $\boldsymbol{\Phi_P^a}(i, :)$ is all elements on row i of $\boldsymbol{\Phi_P^a}$.

Elseif $\boxed{\boldsymbol{L^{pot,max} > \eta_T}}$ **and** $\boxed{\boldsymbol{L^{pot,min} > \eta_M}}$

The system sends a *network-full* message to the user and refuses to admit the user to the network.

3.6.2 A user has his/her t_{gQ} expired

When a service S_i user has his/her t_{gQ} expired, the systems will handle as follows.

If $\boxed{\boldsymbol{L^{max} \leq \eta_T}}$

The system:

- sets the user's ψ_P to P_{i1};
- sets the user's ψ_R to R_{i2};
- decreases by one $\boldsymbol{U_i^g}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$;
- increases by one $\boldsymbol{U_i}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U}(i, \psi_\Gamma)$;
- sets $\boldsymbol{P_i^a}(:)$ and $\boldsymbol{\Phi_P^a}(i, :)$ to P_{i1}; and
- sets $\boldsymbol{R_i^a}(:)$ and $\boldsymbol{\Phi_R^a}(i, :)$ to $R_{i\rho}$.

Elseif $\boxed{\boldsymbol{L^{max} > \eta_T}}$ **and** $\boxed{\boldsymbol{L^{min} \leq \eta_M}}$

The system executes graceful degradation process. At the end of that process, value of the highest number of bids must be at $\Gamma_{i\rho}$, the system then:

- sets $\boldsymbol{P_i^a}(\rho)$ and $\boldsymbol{\Phi_P^a}(i, \rho)$ to $P_{i\rho}$. $\boldsymbol{P_i^a}(1)$ and $\boldsymbol{\Phi_P^a}(i, 1)$ are set to P_{i1};
- sets $\boldsymbol{R_i^a}(\rho)$ and $\boldsymbol{\Phi_R^a}(i, \rho)$ to $R_{i\rho}$. $\boldsymbol{R_i^a}(1)$ and $\boldsymbol{\Phi_R^a}(i, 1)$ are set to R_{i1};
- sets the user's ψ_P to ψ_Γ;
- sets the user's ψ_R to $\boldsymbol{R_i^a}(\psi_\Gamma)$;
- decreases by one $\boldsymbol{U_i^g}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$; and
- increases by one $\boldsymbol{U_i}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U}(i, \psi_\Gamma)$.

3.6.3 A user hangs up

When a service S_i user hangs up, the systems will handle as follows.

If $\boxed{\boldsymbol{L^{max} \leq \eta_T}}$

The system:

- decreases by one:
    1. $\boldsymbol{U_i^g}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$, if the user that hangs up was still in his/her t_{gQ}; or
    2. $\boldsymbol{U_i}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U}(i, \psi_\Gamma)$, if the user that hangs up has had his/her t_{gQ} expired;
- sets $\boldsymbol{P_i^a}(:)$ and $\boldsymbol{\Phi_P^a}(i, :)$ to P_{i1}; and
- sets $\boldsymbol{R_i^a}(:)$ and $\boldsymbol{\Phi_R^a}(i, :)$ to $R_{i\rho}$.

Elseif $\boxed{\boldsymbol{L^{max} > \eta_T}}$ **and** $\boxed{\boldsymbol{L^{min} \leq \eta_M}}$

The system:

- decreases by one:
    1. $\boldsymbol{U_i^g}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U^g}(i, \psi_\Gamma)$, if the user that hangs up was still in his/her t_{gQ}; or
    2. $\boldsymbol{U_i}(\psi_\Gamma)$ and $\boldsymbol{\Phi_U}(i, \psi_\Gamma)$, if the user that hangs up has had his/her t_{gQ} expired;
- executes graceful degradation process. At the end of that process, value of the highest number of bids must be at $\Gamma_{i\rho}$;
- sets $\boldsymbol{P_i^a}(\rho)$ and $\boldsymbol{\Phi_P^a}(i, \rho)$ to $P_{i\rho}$. $\boldsymbol{P_i^a}(1)$ and $\boldsymbol{\Phi_P^a}(i, 1)$ are set to P_{i1}; and
- sets $\boldsymbol{R_i^a}(\rho)$ and $\boldsymbol{\Phi_R^a}(i, \rho)$ to $R_{i\rho}$. $\boldsymbol{R_i^a}(1)$ and $\boldsymbol{\Phi_R^a}(i, 1)$ are set to R_{i1}.

3.7 Smart Pricing signalling

3.7.1 Signalling types

In Section 1.1 we have identified the three types of signalling that are necessary under Smart Pricing. They are:

- the uplink signalling (*uc*);
- the downlink signalling (*dc*); and
- the signalling between networks elements (*nc*).

Each of those types of signalling can be divided into two sub-types:

- signalling for bidding process (with prefix *b*); and
- other signalling.

And, to assist with the implementation of Smart Pricing, we estimate how much signalling volume results in billable events, i.e. signalling that we can directly charge the user who causes the signalling to occur the cost of the signalling, versus signalling which does not. For example, when a new user requests to make a call and the network does not have capacity to admit, the DPE then sends the user a *network-full* message. Those uplink and downlink signalling components are not billable, as the user leaves the network. These two are, respectively, so-called:

- cost recoverable signalling (prefixed *cr*), and
- cost un-recoverable signalling (prefixed *cu*).

Network operators are recommended to take information about cost un-recoverable signalling into account when establishing admission prices for new users and prices for successful bidders when they bid for higher bit rates in the event of graceful degradation.

Altogether, we have: 3 x 2 x 2 = 12 types of signalling. Each time an event triggers a type of signalling to occur, the relevant signalling counter (abbreviated *c* in *uc*, *dc*, and *nc* above) will be increased.

3.7.2 Signalling components

Signalling components specified in this section are those for the case of one service. They are nominal signalling components and a network operator is ultimately to decide which ones are applicable to its network. Some of these signalling components may be omitted and/or additional signalling components be included.

When there is insufficient capacity to admit a new user

The proposed signalling components for this case are in Table 3.5.

	cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
A new MS signals the MSC that he/she is willing to make a call.				1								
RNC sends cell load information to DPE and notifies DPE that a new MS is making a call.						1						
DPE assesses the cell load information, then sends a *network-full* message to the MS.					1							
Subtotals:	0	0	0	1	1	1	0	0	0	0	0	0

Table 3.5: Signalling components when cannot admit a new user.

Let the total number of signalling messages for this case be $\alpha_{blocked}$, we then have:

$$\alpha_{blocked} = 3 \tag{3.35}$$

When there is sufficient capacity but the user's WTP is not sufficient

The proposed signalling components for this case are in Table 3.6.

		cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
The network has enough capacity to admit a new MS:													
	A new MS signals the MSC that he/she is willing to make a call.				1								
	RNC sends cell load information to DPE and notifies DPE that a new MS is making a call.						1						
	DPE assesses the cell load information and requests the TSS to set admission price for the call.						1						
	TSS sets admission price and informs DPE of the price.						1						
	DPE informs the new MS the admission price.					1							
	DPE signals the MSC to hold the connecting of the call.						1						
But the MS's WTP is not sufficient:													
	The MS signals DPE that the admission price is too high for him/her and hangs up.				1								
	DPE signals the MSC to stop connecting the call.						1						
	Subtotals:	0	0	0	2	1	5	0	0	0	0	0	0

Table 3.6: Signalling components when WTP is insufficient.

Let the total number of signalling messages for this case be α_{npa} (npa = no to admission price), we then have:

$$\alpha_{npa} = 8 \tag{3.36}$$

When there is enough capacity and the user's WTP is sufficient

The proposed signalling components for this case comprise of two parts. Signalling components for the first part are in Table 3.7.

	cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
The network has enough capacity to admit a new MS:												
A new MS signals the MSC that he/she is willing to make a call.	1											
RNC sends cell load information to DPE and notifies DPE that a new MS is making a call.			1									
DPE assesses the cell load information and requests the TSS to set admission price for the call.			1									
TSS sets admission price and informs DPE of the price.			1									
DPE informs the new MS the admission price.		1										
DPE signals the MSC to hold the connecting of the call.			1									
And the new MS accepts the price:												
The MS signals DPE he/she accepts the price.	1											
DPE signals MSC to resume the connecting of the call.			1									
DPE confirms with the MS the price it is charging the MS.		1										
DPE informs MSC the price to charge the MS.			1									
Subtotals:	2	2	6	0	0	0	0	0	0	0	0	0

Table 3.7: Signalling components if capacity and WTP are sufficient.

Let the total number of signalling messages for the first part be α_{ypa} (ypa = yes to admission price), we then have:

$$\alpha_{ypa} = 10 \tag{3.37}$$

Depending on the cell load after admitting the new user, different additional signalling components for the second part are required. If, before admitting the new user, $L_i^{pot,max}$ (see Equation (3.33) for details) was found to be below η_T, then after admitting the new user, signalling components in Table 3.8 are required. We call this case *New MS and below* η_T.

	cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
RNC updates DPE on the new cell load after admitting the MS. (This is to keep the DPE informed of the cell load)			1									
Subtotals:	0	0	1	0	0	0	0	0	0	0	0	0

Table 3.8: Signalling components for *New MS and below* η_T.

Let the total number of signalling messages for *New MS and Below* η_T be α_{nh} (*nh* = new user and high bit rate for all users), we then have:

$$\alpha_{nh} = 1 \tag{3.38}$$

If before admitting the new user, $L^{pot,max}$ was found to be above η_T, then after admitting the new user, signalling components in Table 3.9 are required. We call this case *New MS and above* η_T.

		cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
If L_pot,max > η_T													
	RNC updates DPE on the new cell load after admitting the MS.			1									
	DPE requests TSS to set graceful degradation price and bit rate for all t_gQ expired users.			1									
	TSS acts upon DPE's request and informs DPE of the new price and bit rate.			1									
	DPE instructs MSC to assign new price and bit rate to t_gQ expired users.			1									
	DPE broacasts the new prices to t_gQ expired MSs.		1										
	RNC updates DPE on the new cell load after new bit rate is assigned to t_gQ expired users.			1									
Bidding - (if h>0; h is the number of h-type users)													
	High WTP users with t_gQ expired bid up.							h					
	DPE requests TSS to set new price for bid-up MSs.									1			
	TSS acts upon DPE's request and informs DPE of the new price.									1			
	DPE broadcasts the new price to bid-up users.								1				
	DPE instructs MSC to assign new bit rate and informs new price for bid-up users.									1			
	RNC updates DPE on the new cell load after the bidding process.									1			
	Subtotals:	0	1	5	0	0	0	h	1	4	0	0	0

Table 3.9: Signalling components for *New MS and above* η_T.

Let the total number of signalling messages for *New MS and above* η_T be α_{ng} (*ng* = new user and graceful degradation taking place), we then have:

$$\alpha_{ng} = 11 + h \tag{3.39}$$

where h is the number of users with t_{gQ} expired who have high WTP.

When a user has t_{gQ} expired or hangs up

When a user has his/her t_{gQ} expired (or hangs up), if L_i^{max} (see Equation (3.29) for details) is below η_T, signalling components in Table 3.10 are required. We call this case t_{gQ} *expired or HU and below* η_T.

	cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
RNC informs DPE that a MS has hung up (or has t_gQ expired) and updates DPE on the new cell load.						1						
DPE assesses the cell load information and requests the TSS to set a new price and bit rate for t_gQ expired users.						1						
TSS sets and informs DPE of the new price and bit rate.						1						
DPE broacasts the new price to t_gQ expired MSs.					1							
DPE instructs MSC to assign new bit rate and informs new price for t_gQ expired MSs.						1						
RNC updates DPE on the new cell load after assigning new bit rate to t_gQ expired MSs.						1						
Subtotals:	0	0	0	0	1	5	0	0	0	0	0	0

Table 3.10: Signalling components for t_{gQ} *expired or HU and below* η_T.

Let the total number of signalling messages for t_{gQ} *expired or HU and below* η_T be α_{hgh} (hgh = hang up or t_{gQ} expired, and high bit rate), we then have:

$$\alpha_{hgh} = 6 \tag{3.40}$$

When a user has his/her t_{gQ} expired (or hangs up), if L_i^{max} is above η_T, signalling components in Table 3.11 are required. We call this case t_{gQ} *expired or HU and above* η_T.

Let the total number of signalling messages for t_{gQ} *expired or HU and above* η_T be α_{hgg} (hgg = hang up or t_{gQ} expired, and graceful degradation taking place), we then have:

$$\alpha_{hgg} = 11 + h \tag{3.41}$$

	cr_uc	cr_dc	cr_nc	cu_uc	cu_dc	cu_nc	cr_buc	cr_bdc	cr_bnc	cu_buc	cu_bdc	cu_bnc
If load to give all active MSs R_H exceeds low load threshold												
RNC informs DPE that a MS has hungup (or has t_gQ expired) and updates DPE on the new cell load.						1						
DPE assesses the cell load information and requests TSS to set graceful degradation price and bit rate for all t_gQ expired users.						1						
TSS acts upon DPE's request and informs DPE of the new price and bit rate.						1						
DPE instructs MSC to assign new price and bit rate to t_gQ expired users.						1						
DPE broacasts the new prices to t_gQ expired MSs.					1							
RNC updates DPE on the new cell load after the new bit rate is assigned to t_gQ expired users.						1						
Bidding - (conditional on h>0; h is the # of h-type users)												
High WTP users with t_gQ expired bid up.							h					
DPE requests TSS to set new price for bid-up MSs.									1			
TSS acts upon DPE's request and informs DPE of the new price.									1			
DPE broadcasts the new price to bid-up users.								1				
DPE instructs MSC to assign new bit rate and informs new price for bid-up users.									1			
RNC updates DPE on the new cell load after the bidding process.									1			
Subtotals:	0	0	0	0	1	5	h	1	4	0	0	0

Table 3.11: Signalling components for t_{gQ} *expired or HU and above* η_T.

3.8 Simulation and results

3.8.1 Sizes of signalling messages

Sizes of the signalling messages used in the simulation are:

- uplink message = 640 bits. Among the possible signalling means for the uplink identified in Section 2.8, we pick the DPDCH as a nominal means. The data bits in a slot in a radio frame on the DPDCH range from 10 up to 640 bits, when spreading factors ranging from 256 down to 4 are used, respectively [42]. Here we use the maximum value;

- downlink message = 1256 bits. Among the possible signalling means for the downlink identified in Section 2.8, we pick the S-CCPCH as a nominal means as it is the physical channel that both the FACH and PCH are mapped onto. The data bits in a slot in a radio frame on the CCPCH range from 20 up to 1256 bits, when spreading factors ranging from 256 down to 4 are used, respectively [42]. The maximum value is used here; and

- between network elements = 4608 bits. This is the minimum size of an IP packet that any host is required to handle [48].

We assume that every time signalling on a certain link is required, one message/packet gives sufficient room for the signalling content. If network operators deem that this is not enough, simulation results in this chapter (as well as the next) can still be used. This is done by multiplying the average signalling load for the relevant link by the desired factor of the number messages/packets required.

3.8.2 Simulation program and parameters

A simulation program has been written for the Smart Pricing signalling model in this chapter. In this simulation program, we generate random data for the the arrival time, hold time and WTP for each user. The data are generated from Poisson, negative exponential and Weibull distributions with relevant desired parameter value(s) for each distribution as specified in Table 3.12. The three user attributes for which random data are generated are among the seven attributes needed for a user as discussed in Section 3.5.2.

Simulations are conducted for voice services and video conferencing services (CONF), which belong to the *Conversational* class. Conversational class is one of the four UMTS QoS classes. Only varying relevant simulation parameters listed in Table 3.12 simulations for other classes, e.g the *Background* class, can also be conducted. The *Conversational* class is chosen because it is very delay sensitive, which is in contrast with the *Background* class, the most delay insensitive traffic class [49] [50].

Smart Pricing signalling system with CONF is considered a *small system*, as the system can accommodate only a small number of users, whereas with voice, there are relatively more users and so considered a *large system*. For CONF, two cases are considered: two and five WTP levels in the system. Two WTP levels is a simple case and five WTP levels is a more realistic one. Comparing simulation results from these two cases helps us to decide whether modelling of Smart Pricing signalling system in its simple case is adequate to generalise our findings. Great attention will be placed on bidding signalling as more user groups mean greater levels of responding to the rising load, i.e. rising congestion, of the network. We are interested in finding out the effect of that on the system in terms of signalling. On the other hand, comparing voice and CONF simulation results allows us to see how significant an increase of signalling is when having more users in the network.

Table 3.12 summaries values of the main parameters for the simulation.

No.	Parameter	Value				Unit
		Video call (CONF)		Voice call (VOICE)		
		Range when varied	For base case	Range when varied	For base case	
1	η_M	1-20	3	0.1-0.6	0.6	dB
2	λ	0.5-5	3	0.5-5	3	calls/s
3	η_T factor	0-1	0.25	0-0.6	0.6	
4	Bit rates	32-144		4.75-12.2		kbps
5	Bit rate low		32		4.75	kbps
6	Bit rate high		144		12.2	kbps
10	Mean hold time	30-330	180	30-330	180	seconds
11	tgQ	30-180	60	30-180	60	seconds
12	Scale parameter	3.325-10.325	3.325	3.325-10.325	3.325	
13	Shape parameter	3.7				
14	Eb/N0	5				dB
15	i	0.55				
16	W	3.84E+06				chips
17	υ	0.67				

Table 3.12: Parameter values for simulation.

3.8.3 Strategy for simulation

In order to observe how different system characteristics affect the signalling requirements for Smart Pricing system, we will monitor three broad categories of system characteristics. Together with their sub-categories (as applicable), six simulation scenarios will be analysed.

1. Conditions under which a MS is admitted to the network – in this category, signalling requirements when each of the following system characteristics varies are investigated.
   (a) The cell maximum capacity. This is achieved by varying NR;
   (b) Mean users' WTPs. This is achieved by varying the *scale parameter* of the Weibull distribution; and
   (c) Arrival rate.
2. Level of congestion is defined – this is achieved by varying:
   (a) Low load threshold factor; and
   (b) t_{gQ}.
3. User's behaviour when they are already in the system – This is achieved by varying mean hold time of user's calls.

Combined with the three types of simulations – two cases for the small system and one for the larger, mentioned in Section 3.8.2, a total of eighteen simulation scenarios will be simulated. For each simulation scenario, the following parameters' values will be recorded:

- Uplink (UL) average signalling load;
- Downlink (DL) average signalling load;
- Inter-network elements (Netw2Netw) average signalling load;
- Average bidding signalling as percentage of total signalling for UL, DL and Netw2Netw;
- Average cost un-recoverable signalling as percentage of total signalling for UL, DL and Netw2Netw; and
- Average simulation time.

In addition, the following other system performance parameters will also be collected:

- probability of occupancy;
- blocking probability due to insufficient capacity;
- blocking probability due to price (i.e. due to insufficient WTP levels); and
- probability of successful arrivals.

These parameters are among the outputs of the Smart Pricing signalling system and verifying the accuracy of these values is of vital importance. In fact, without first being certain that we have the correct values, none of the above signalling and time parameter values can be validated. The latter three parameter values provide crucial assistance during the analysis process.

For simplicity, not all simulation results will be shown. A subset of results which exhibits distinctive behaviours is included in this chapter.

3.8.4 Determination of simulation warm-up period

In this chapter we model the Smart Pricing signalling system in its instantaneous condition, the so-called MCS model, whereas in the next we will model in its steady-state condition, the so-called SSA model. One of the aims is to compare the time performance (i.e. time taken for one simulation run) of the simulation program for the

SSA model and the one for the MCS model. Since the SSA model is a steady-state model, we need to eliminate the time spent in the transient state for every simulation run of the MCS model. If this is not done, the time performance information collected will not be accurate if the transient state lasts for a relatively long time. This problem is known as the *initialization bias* and the time spent in the transient state is called the *warm-up period* [51].

To eliminate the warm-up period, t_{wu}, although we could employ the *Replication / Deletion* approach or the *Regenerative* method as described in [52], we have chosen to use *Welch*'s method. Welch's method is claimed to be the simplest and most general technique used for determining the warm-up period. We will be following the steps detailed in [51].

1. We make $n = 10$ replications of the simulation for CONF base case (parameter values as in Table 3.12). Run length m is chosen to be 7200 seconds, and with $\lambda = 3$ arrivals/s, we have approximately $m = 21600$ arrival events. Simulation run time was recorded for each arrival event i^{th} from the j^{th} replication, i.e. Y_{ij}.

2. Next, the ensemble averages over the replications are calculated:

$$\overline{Y}_i = \sum_{j=1}^{10} \frac{Y_{ji}}{n} \quad \text{for } i = 1, 2, ..., 21600 \tag{3.42}$$

3. Then, the window w is chosen to be 600 arrival events (which meets the condition $w \leq m/4$) and define a moving average $\overline{Y}_i(600)$:

$$\overline{Y}_i(600) = \begin{cases} \dfrac{\sum_{s=-600}^{600} \overline{Y}_{i+s}}{2w+1} & \text{for } i = 600+1, 600+2, ..., 21600-600 \\ \dfrac{\sum_{s=-(i-1)}^{i-1} \overline{Y}_{i+s}}{2i-1} & \text{for } i = 1, 2, ..., 600 \end{cases} \tag{3.43}$$

4. Finally, in Figure 3.13, we plot $\overline{Y}_i(600)$, for $i = 1, 2, ..., 21600 - 600$.

   The second part of Figure 3.13 is the zoomed-in plot and it is clear from the plot that beyond approximately $t_{wu} = 180$, $\overline{Y}_i(600)$ appears to be following the best fit curve. Hence, we choose 180 arrival events as the simulation warm-up period. The time taken by the this period will therefore be subtracted from all MCS model simulation runs.

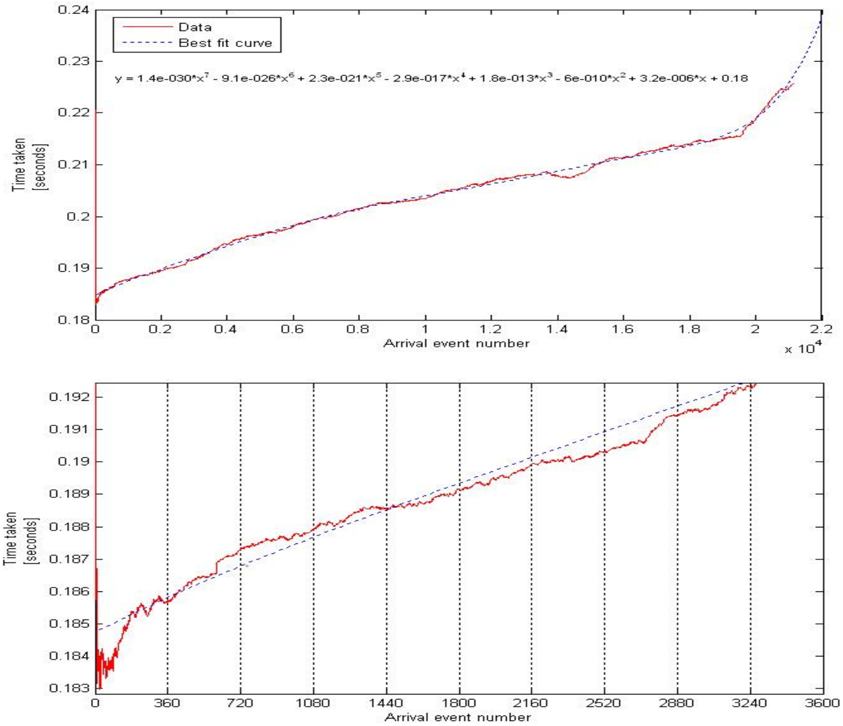


Figure 3.13: Warm-up period.

3.8.5 Small system with two levels of WTP

A system with only two levels of WTP, hence two prices, may make Smart Pricing appear like the conventional peak and off-peak pricing scheme. It is not. With peak and off-peak pricing scheme, prices are fixed and it therefore has, *inter alia*, the following limitations when compared to Smart Pricing:

- the users do not have the option to reject a high admission price. Of course, they can decide not to make the call in the first instance but they will need to remember when high prices apply. The person who makes a call may not always be aware or conscious of off-peak periods;
- the network does not have the option to set the high price to levels that adequately reflect the cost of congestion at different instances of time; and
- the network does not have the option to reduce further the low price to a strategic price level that would give a strong incentive for users to use an under-utilised network during quiet periods.

Probability of occupancy

With parameter values for the CONF case as set out in Table 3.12, we can calculate the number of high WTP users with η_T and η_M, as well as the maximum possible number of users in the system. These values are in Table 3.14 below.

NR [dB]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
η_M	20.58	36.90	49.88	60.19	68.38	74.88	80.05	84.15	87.41	90.00	92.06	93.69	94.99	96.02	96.84	97.49	98.00	98.42	98.74	99.00
High WTP users with η_T	0	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
High WTP users with η_M	1	3	4	5	5	6	7	7	7	7	8	8	8	8	8	8	8	8	8	8
Max users possible	1	3	5	6	6	7	7	8	8	8	9	9	10	10	10	10	10	10	10	10

Table 3.14: CONF case and numbers of users the system can accommodate.

The simulation is run for the base case, $\eta_M = 3$ $[dB]$, and the network occupancy together with its associated probability of occupancy can be seen in Figure 3.15.

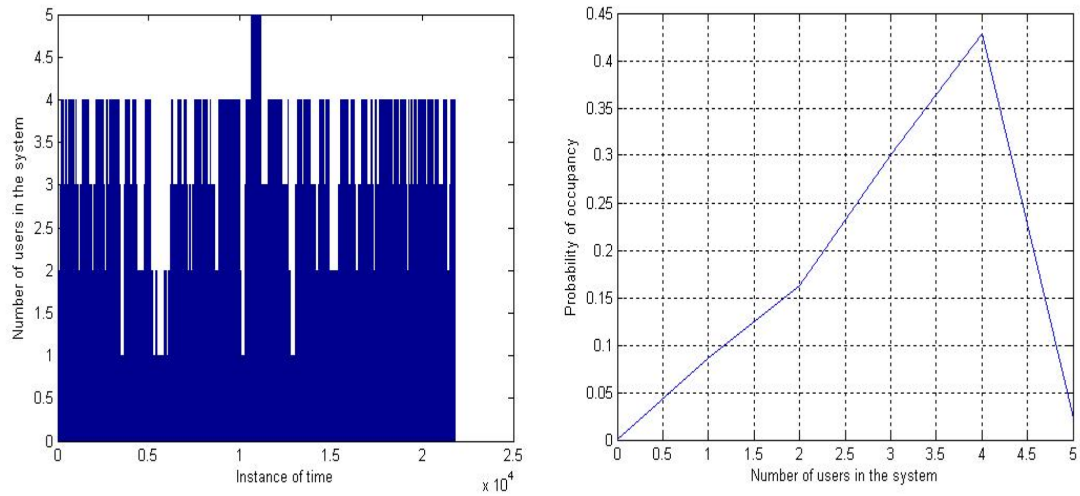


Figure 3.15: Network occupancy - CONF base case 2WTPs.

It is seen that the probability of network occupancy is highest when there are 4 users in the system. This is because the arrival rate is high relative to the network capacity, so all 4 users with a high WTP accommodated with $\eta_M = 3$ $[dB]$ will be able to get onto the system. The probability of 5 users in the system is small because the chances are low for a user with a low WTP arriving at the network when the network is empty.

NR and the required signalling

Figure 3.16 shows the UL, DL and Netw2Netw average signalling loads versus *NR*.

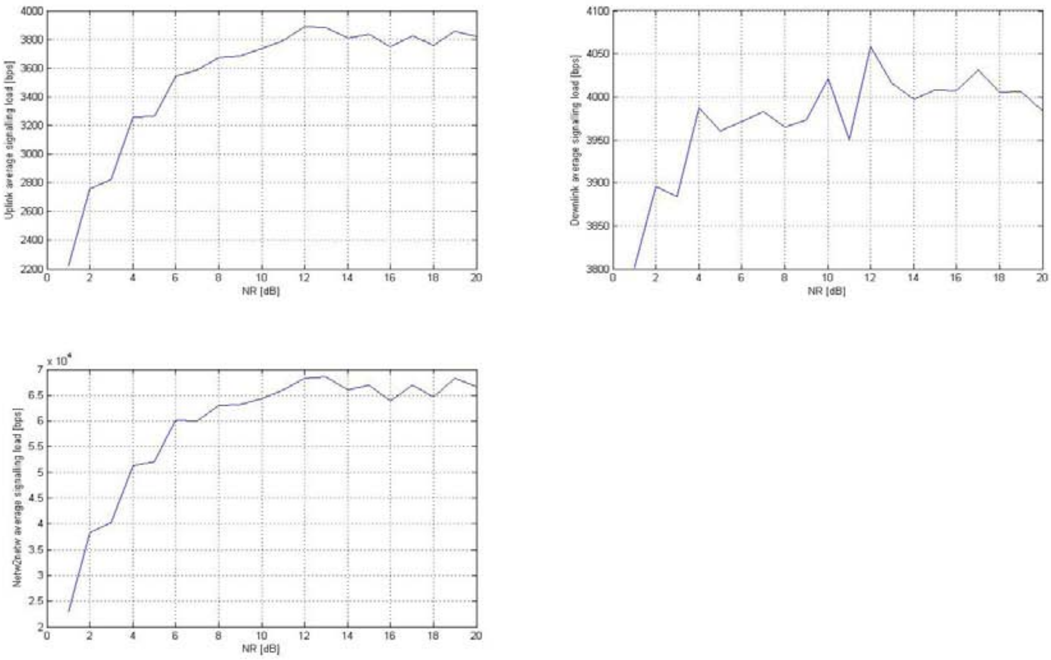


Figure 3.16: Average signalling load vs. NR.

All the signalling loads are seen to increase as *NR* increases up to *NR* = 13 [dB] then remains at this level. Within the 1-13 dB range, if either the number of users with high WTP levels that can be accommodated by η_M or the maximum possible number of users that can be accommodated by the system increases, the signalling loads increase. From 13 dB to 20 dB, the numbers of users remain unchanged. This is when the system reaches 94.99 - 99% of its maximum capacity and an increase in *NR* does not add enough capacity to accommodate an additional user. The average signalling loads have maximum values of 3.89, 4.06 and 68.53 kbps for the UL, DL and Netw2Netw. The UL maximum is the lowest, the DL is 1.04 times greater than the UL, and the Netw2Netw peak is the highest and 17.61 times greater the UL.

Figure 3.17 shows a profile of bidding signalling as percentages of the total signalling. The percentages appear to be influenced by the number of high WTP users that can be accommodated by η_M. These keep increasing until the system reaches *NR* = 11 [dB]; that is when the number of user with high WTP reaches its highest value of 8. UL has the highest maximum percentage of 5.63%, DL peaks at 2.52%, and Netw2Netw has the lowest which peaks at 2.29%.

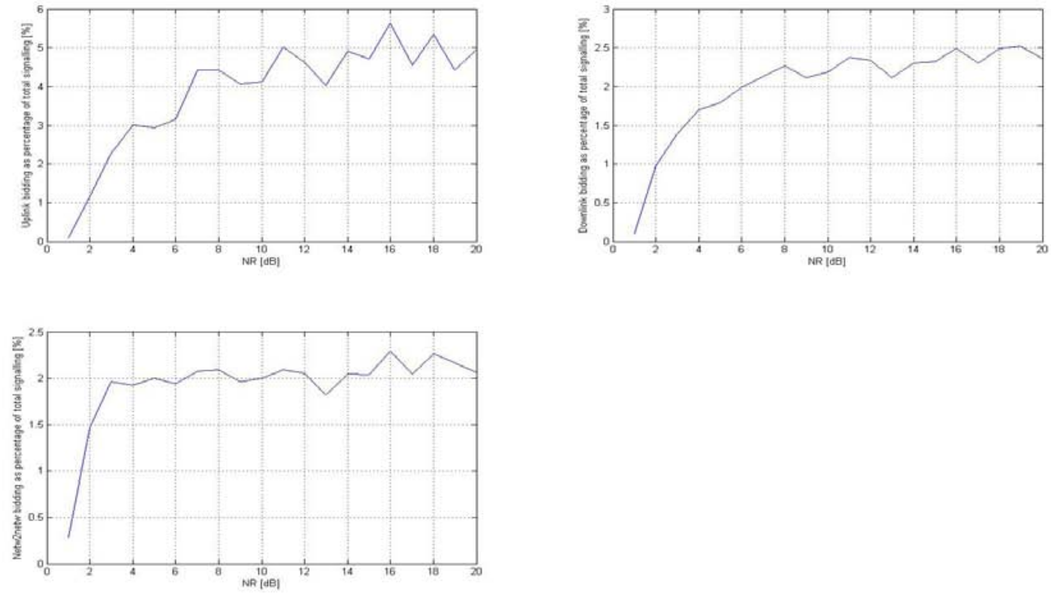


Figure 3.17: Bidding signalling as percentage of total signalling vs. NR.

Figure 3.18 shows a profile of cost un-recoverable signalling as percentages of the total signalling.

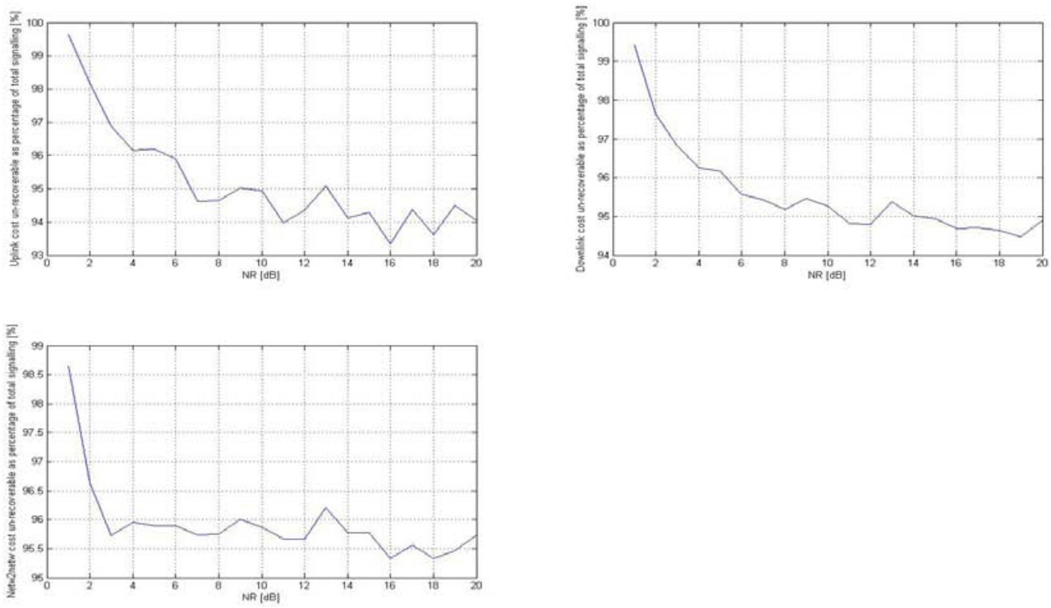


Figure 3.18: Cost un-recoverable signalling as percentage of total signalling vs. NR.

It is seen that the percentages decrease as *NR* increases. DL has the lowest minimum, followed by the UL then the Netw2Netw, however, they are all above 93%.

The percentages also have a relationship with their counterpart bidding percentages. The sum of the average bidding signalling and average cost un-recoverable signalling percentages for each link type is approximately 100%. This is because when the system load goes above η_T, most of the signalling is cost un-recoverable if it is not for bidding. See Section 3.7.2 for details. As the percentages of the average cost un-recoverable are very high, they indicate that a majority of the time the system is either in conditions:

- where it does not have enough capacity to admit a new user, or
- when potentially new users with insufficient WTP levels initiate calls,

as the signalling when the system in those conditions is all cost un-recoverable.

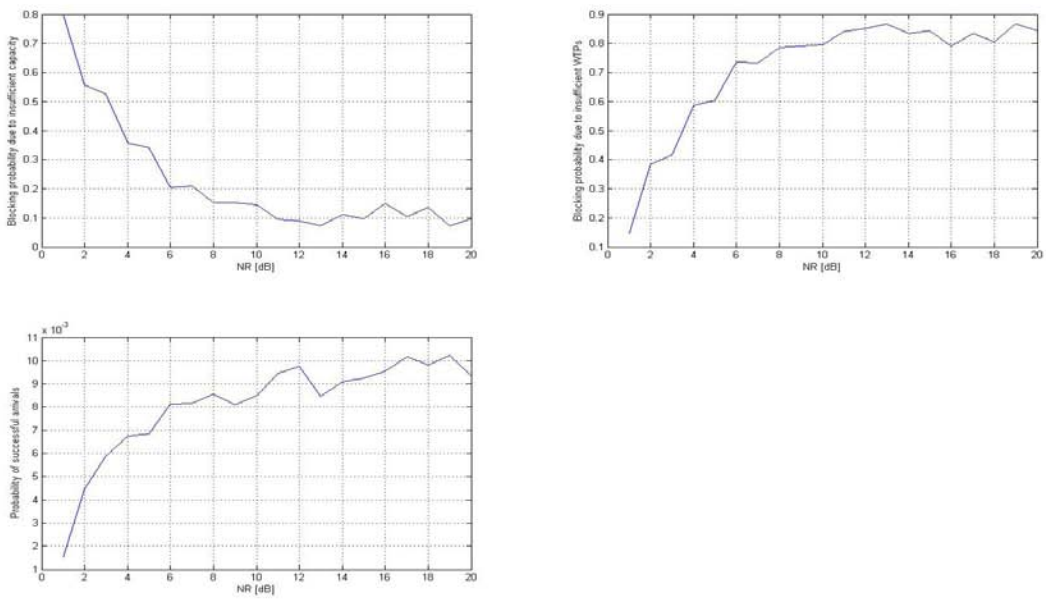


Figure 3.19: Probabilities vs. NR.

Blocking probability due to capacity is seen from Figure 3.19 to decrease whereas probability of successful arrivals increases as *NR* increases. This is expected as more capacity means more users are admitted to the network. The minimum blocking probability due to insufficient capacity is 7.23%. On the other hand, the maximum probability of successful arrivals is 1.02% when *NR* reaches its highest value of 20 [dB]. This is because the probability of high WTP (1.08%) is very close to that value with the scale and shape parameters as specified in Table 3.12. That is also the reason why 86.83% is the maximum blocking probability due to insufficient WTP levels. What constitutes approximately that value can be explained by the sum 1.08% of the users with high WTP levels and around 9% (refer to Figure 3.15) of probability of occupancy of one user in the system. The 9% is an indication the number of users with low WTP levels successfully admitted to the network arrive when the system is empty.

WTP and the required signalling

To see how user WTP levels affect Smart Pricing signalling, we vary the mean value of the users' WTP levels by varying the scale parameter of the Weibull distribution.

Scale parameter (δ)	3.325	3.825	4.325	4.825	5.325	5.825	6.325	6.825	7.325	7.825	8.325	8.825	9.325	9.825	10.325
Mean WTP	3.0007	3.4519	3.9031	4.3543	4.8056	5.2568	5.708	6.1592	6.6105	7.0617	7.5129	7.9642	8.4154	8.8666	9.3178
% of Low WTP user	97.75%	92.55%	81.47%	67.75%	54.51%	43.21%	34.12%	27.03%	21.54%	17.31%	14.03%	11.47%	9.46%	7.86%	6.59%
% of High WTP user	1.08%	6.76%	18.08%	31.95%	45.29%	56.65%	65.77%	72.89%	78.39%	82.64%	85.93%	88.50%	90.52%	92.11%	93.39%

Table 3.20: Scale parameters and corresponding mean and percentages.

As we keep the shape parameter fixed at 3.7, each value of the scale parameter corresponds to a different mean and different percentages of low WTP (WTP1) users and high WTP (WTP5) users. These relationships can be seen in Table 3.20. The 3.7 value of the shape parameter is chosen because, together with the scale parameter value of 3.325, the mean WTP levels is around 3, which is the average value of the price range.

Figure 3.21 shows the variation of signalling with respect to the mean WTP. As the scale parameter increases to 5.835, the mean WTP increases accordingly. This causes blocking probability due to insufficient WTP levels to reduce as more high WTP users than low WTP users will arrive and those users are willing to accept the high price. On the other hand, blocking probability due to insufficient capacity increases because of the higher number of high WTP users who are unable to enter the network. When blocking probability due to insufficient capacity is higher, UL and Netw2Netw signalling loads reduce because UL and Netw2Netw signalling components in Table 3.7 are less than in Table 3.5. This explains why the average UL and Netw2Netw signalling loads in the Figure 3.21 decreases as the scale parameter increases to 5.835.

When scale parameter equals 5.835, the mean WTP is already greater than the high price. Thus, further scale parameter increase will just mean that high WTP users are willing to pay a lot higher than the high price. This, however, will not have much effect on the system as at that point it is already full. That is why we see, from 5.825 to 10.325, the average UL and Netw2Netw signalling loads only slightly fluctuate around their respective values when the scale parameter is equal to 5.835.

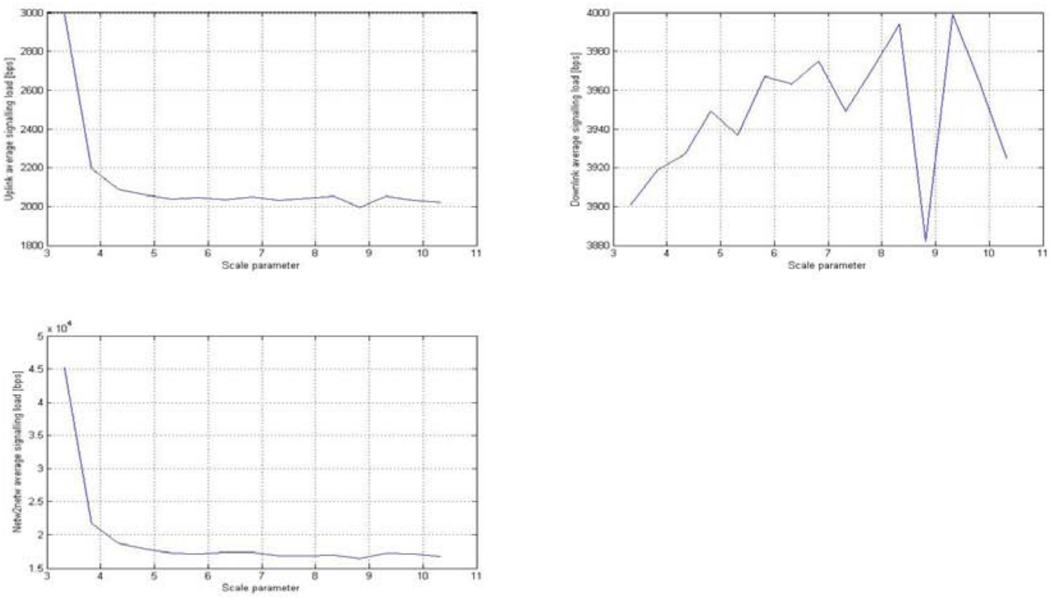


Figure 3.21: Average signalling load vs. WTP.

The DL average signalling load is seen to increase as the scale parameter increases. This is because the blocking probability due to insufficient WTP levels reduces as more high WTP users are admitted to the network. In such a situation, more DL signalling components are needed because the DL signalling components now used are those in Table 3.7, no longer those in Table 3.5.

Arrival rate and the required signalling

From Figure 3.22, the average signalling loads are seen to be directly proportional to the arrival rate. This is due to more users arriving causing more requests to make calls or requests to have their QoS increased (i.e. UL signalling) in the events of graceful degradation. Network elements will need to communicate with each other (i.e. Netw2Netw signallling) to process those requests and then to respond to the users (i.e. DL signalling). At the very least, the increase of signalling are those signalling components described in Table 3.5.

Furthermore, as the DL average signalling load is higher than the UL average signalling load, it indicates that, apart from being in situations when it does not have enough capacity to admit new users, a majority of the time the network is in situations where users have their t_{gQ} expired or have hung up and the load is below η_T. In those situations no UL signalling components are required. Refer to Table 3.10 for details.

As the arrival rate increases, blocking probability due to insufficient capacity in-

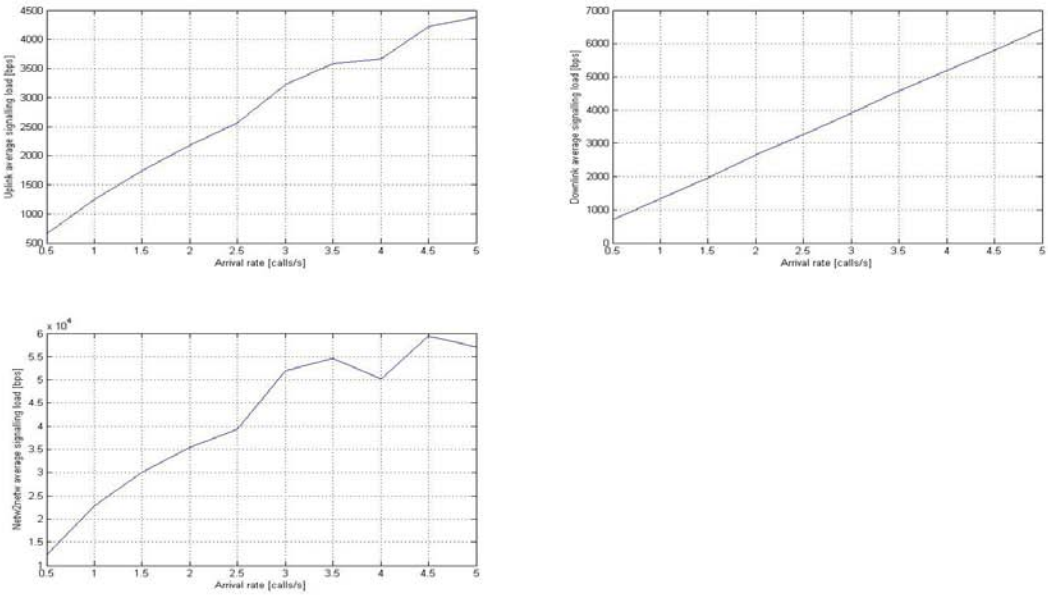


Figure 3.22: Average signalling load vs. Arrival Rate.

creases, whereas blocking probability due to insufficient WTP levels and probability of successful arrivals are seen to increase up to $\lambda = 1.5$ [arrivals/s] then decrease.

Low load threshold and the required signalling

As η_T is calculated based on the low load threshold factor, varying the factor in effect is varying η_T. When the factor equals 0, the network is considered congested at all times, hence high prices will always apply, while at 1 it is considered not congested at all times, hence low prices will always apply. Signalling versus low load threshold factor is shown in Figure 3.23.

Although fluctuating, the trends do show that the average signalling decreases as the low load threshold factor increases. The reductions are clearer for the UL and Netw2Netw (particularly from low load threshold factor = 0.7) than for the DL. The explanation is that as the low load threshold factor increases, more low WTP users will be allowed to enter the network, meaning that probability due to insufficient WTP levels reduces and the probability of successful arrivals increases. In turn, this means the blocking probability due to insufficient capacity increases. The reduction in the DL average signalling is less noticeable because:

- with the high arrival rate of $\lambda = 3$ [arrivals/s] the network will operate most of the time when load is above η_T;
- as long as load is above η_T, graceful degradation will take place;

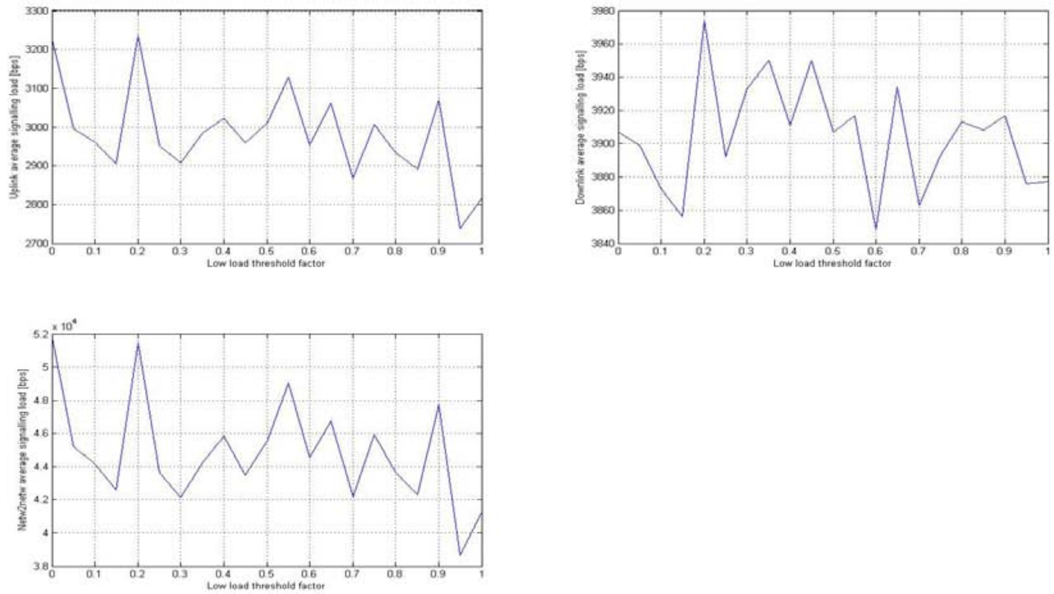


Figure 3.23: Average signalling load vs. η_T factor.

- if graceful degradation takes place, the same amount of DL signalling is required (see Tables 3.9 and 3.11 for details); and
- only when no graceful degradation takes place are DL signalling components in Tables 3.8 and 3.10 required, which is a reduction when compared with Tables 3.9 and 3.11. Such a situation, however, is less likely to occur.

t_{gQ} and the required signalling

Signalling loads versus t_{gQ} can be seen in Figure 3.24. Although fluctuating, the average signalling loads do seem to decrease slightly as t_{gQ} increases. This is because when t_{gQ} increases, admission bit rate of low WTP users are guaranteed longer, which gives less chances for the network to reduce their bit rates in the events of graceful degradation. Consequently, blocking probability due to insufficient capacity increases as less new users are able to be admitted.

Allowing t_{gQ} to assume greater values means less signalling for Smart Pricing, and hence less network resources will need to be dedicated for signalling. This is a benefit for the network operator. However, it also means that the network operator will have less flexibility to respond to network congestion. Conversely, setting t_{gQ} to lower values means more signalling is required for Smart Pricing; at the same time, this means the satisfaction of low WTP users (those that were successfully admitted to the network) will increase. It also means that new users – both low WTP and high WTP users –

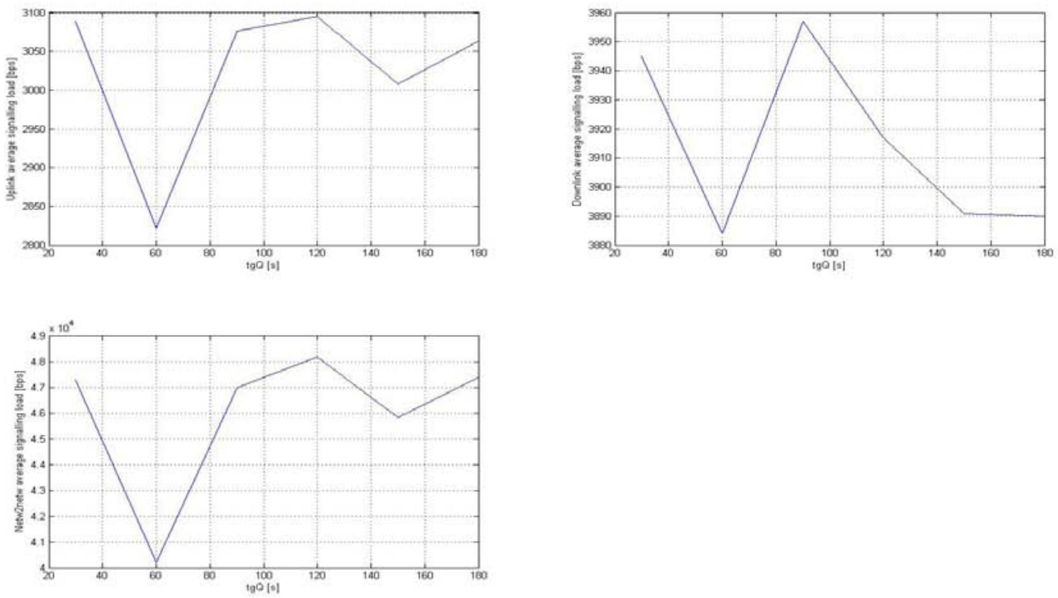


Figure 3.24: Average signalling load vs. tgQ.

will be required to respond to new price messages less often. Thus, the value to which t_{gQ} should be set is a trade-off matter. It is up to the network operators to decide upon which system characteristics they wish to optimize.

Mean hold time and the required signalling

Figure 3.25 shows a profile of signalling versus call holding time. It is seen that the average signalling loads decrease as hold time increases. The explanation for this is that when users stay on the network longer, there are less opportunities for new users to be admitted, therefore less graceful degradation events occur. Signalling required in the event of graceful degradation is greater than for blocking new users. Compare the signalling components in Tables 3.9 and 3.11 with Table 3.5 for details.

Simulation times

A summary of the average simulation times for all six simulation scenarios for the small system (i.e. CONF) with 2 levels of WTP considered in this section is provided in Table 3.26 below. The warm-up period determined in Section 3.8.4 is taken into account in the process of recording the simulation times. All simulation times shown in this thesis are collected from a laptop computer with the following specifications:

- Windows XP Professional Service Pack 2 operating system,

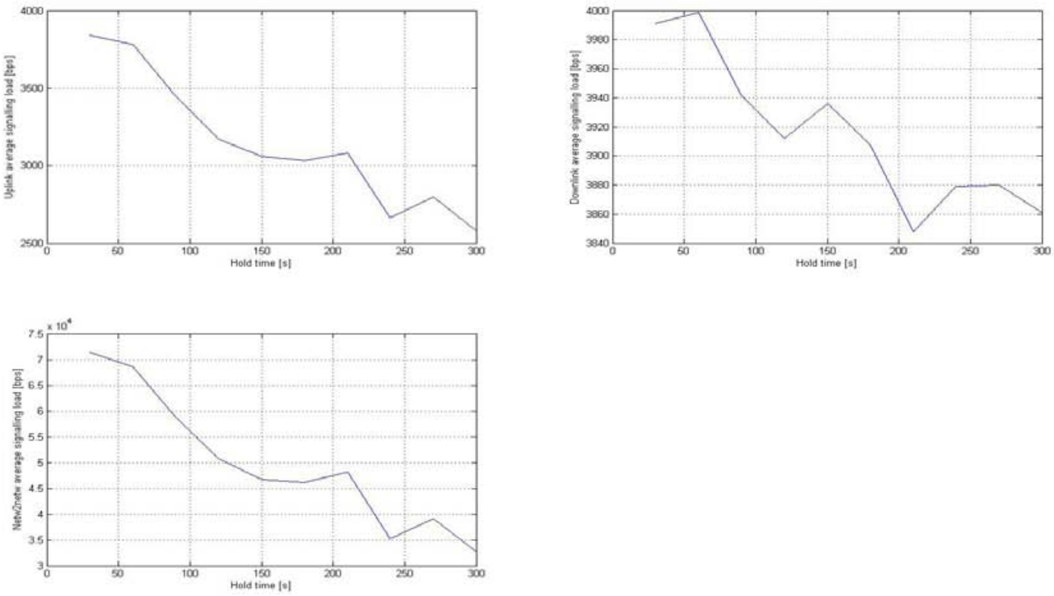


Figure 3.25: Average signalling load vs. hold time.

- 1 GB memory,
- 2.161 GHz processor, and
- Matlab version 7 (Release 14).

The smallest and highest average simulation times are found to be 763 and 1021 seconds, respectively. The overall average simulation time of all six scenarios is 867 seconds.

Simulation scenario	Average simulation time [s]
NR	881
WTP	763
Arrival rate	1021
Low load threshold	848
tgQ	713
Hold time	977
Overall average [s]:	867

Table 3.26: Simulation times - CONF 2-WTP.

3.8.6 Small system with five levels of WTP

In this section, we conduct simulations for a more realistic case in which there are five levels of WTP (5-WTP) in the system. Most of the results are plotted against results of the simple case (2-WTP) in the previous section (Section 3.8.5) for the ease of comparison. For simplicity, we shall be showing results for the *NR* simulation scenario only.

Probability of occupancy

We notice from Figure 3.27 that probabilities of occupancy of the 5-WTP and 2-WTP (see Figure 3.15) cases follow the same pattern. A prominent difference is that with the 5-WTP case, the probability is higher for five users, the maximum possible for the CONF, in the system. This is because firstly there are now three intermediate WTP levels in between the WTP1 and WTP5, and secondly admission prices are set using the highest bid recorded during graceful degradation events (refer to Smart Pricing principles in Section 3.4 for details). Coupled together, these mean greater chances of entering the network for users with lower WTP levels.

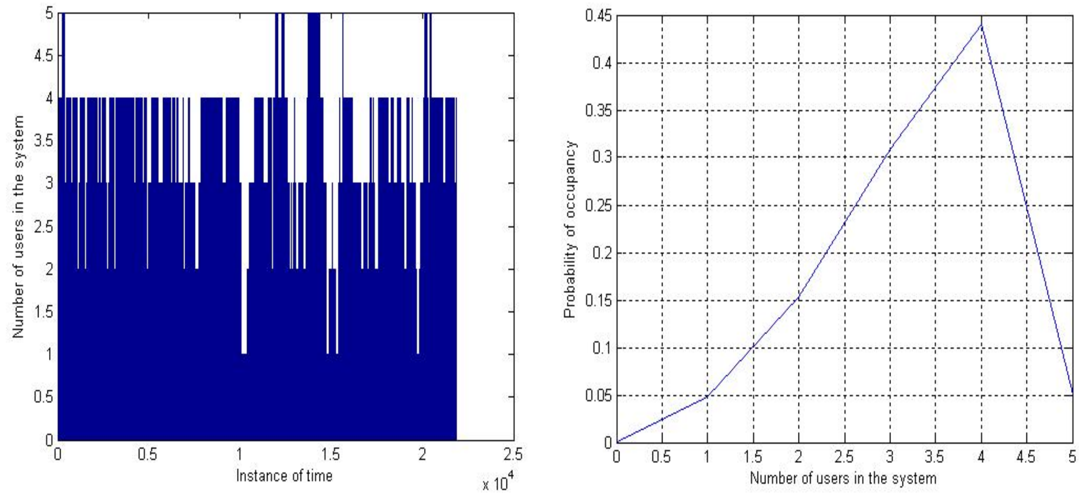


Figure 3.27: Network occupancy - CONF base case 5-WTP.

NR and the required signalling

Although fluctuating, looking closer at the peaks in Figure 3.28, the trends of the average loads do seem to increase for the 5-WTP case when compared with the 2-WTP. The increase is clearer for the DL. The reasons for this will become clearer as we examine the next three figures.

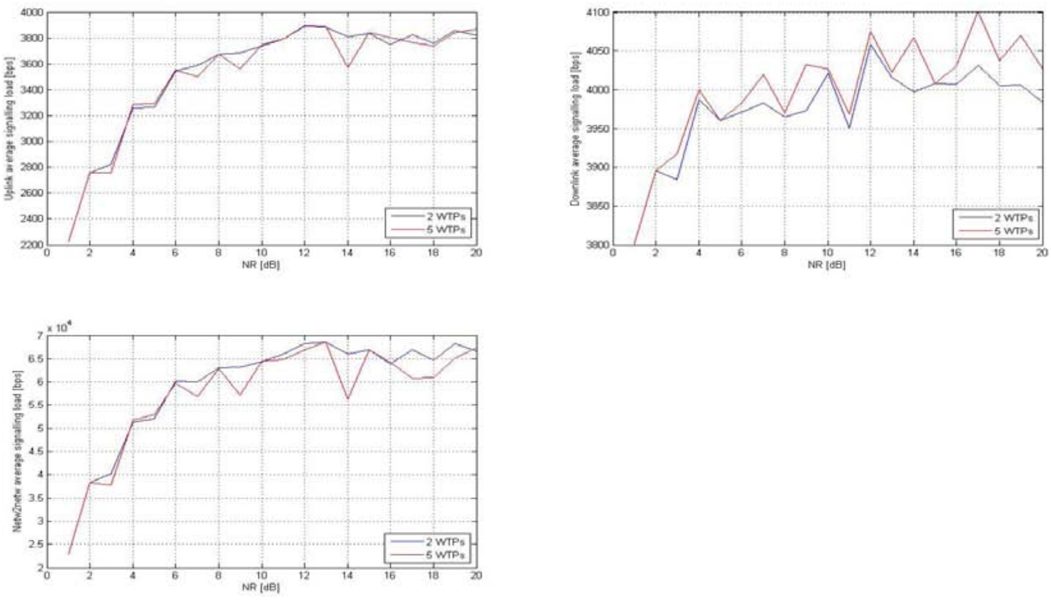


Figure 3.28: Average signalling load vs. NR - 5-WTP.

We see in Figure 3.29 that bidding signalling increases. This is because the bidding process is getting more intense as there are more user groups, each with a different WTP. Users of these groups will each submit a number of bids that is equal to their WTP (or one bid higher than the lower WTP group) in an effort to keep the admission QoS for his/her call.

In the 2-WTP case, as there are only two levels of WTP, WTP5 users only need to submit one bid in order to retain the admission QoS of their calls. Now, WTP5 users potentially need to submit four bids. Apart from the WTP5 users, in the system there are also other users who are with WTP levels 2, 3 and 4. Each of these users will join the contention to keep the admission QoS for their calls and potentially submit 1, 2 and 3 bids, respectively. Even if they do not succeed in keeping the admission QoS, they still try to improve the QoS of their calls as much as possible. These are the reason why the average bidding signalling percentages increase.

As pointed out in Section 3.8.5, the sum of percentages of the average bidding and the average cost un-recoverable signalling for each link type is approximately 100%.

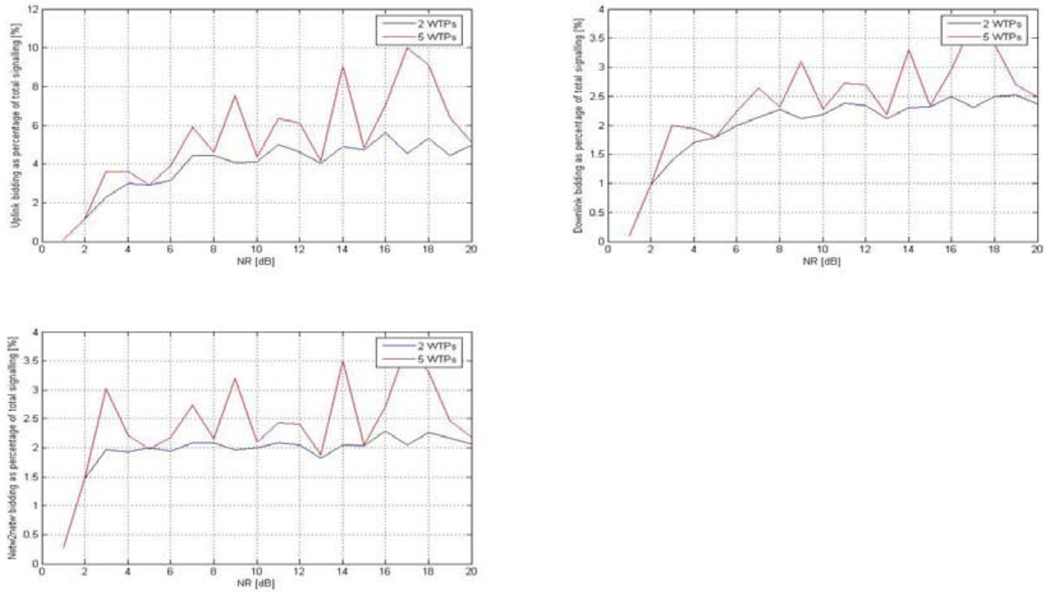


Figure 3.29: Bidding signalling as percentage of total signalling vs. NR - 5-WTP.

The fact that seeing in Figure 3.29 the bidding signalling increases leads us to automatically predict that there will be a decrease in cost un-recoverable signalling. Then, looking at Figure 3.30, our prediction is immediately affirmed.

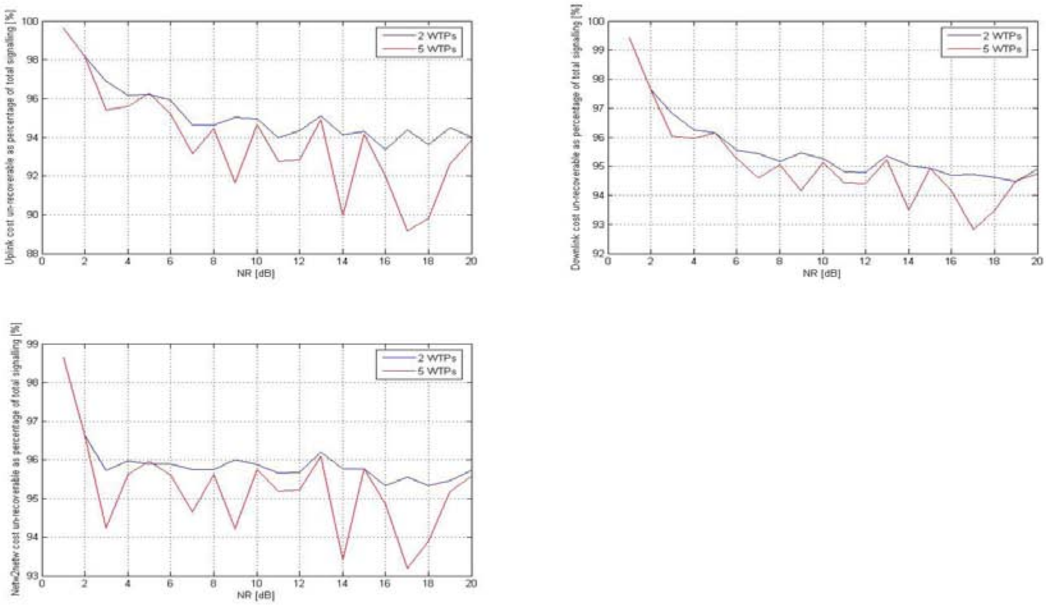


Figure 3.30: Cost un-recoverable signalling as percentage of total signalling vs. NR - 5-WTP.

Figure 3.31 helps explain why the average signalling loads are behaviours as seen in Figure 3.28.

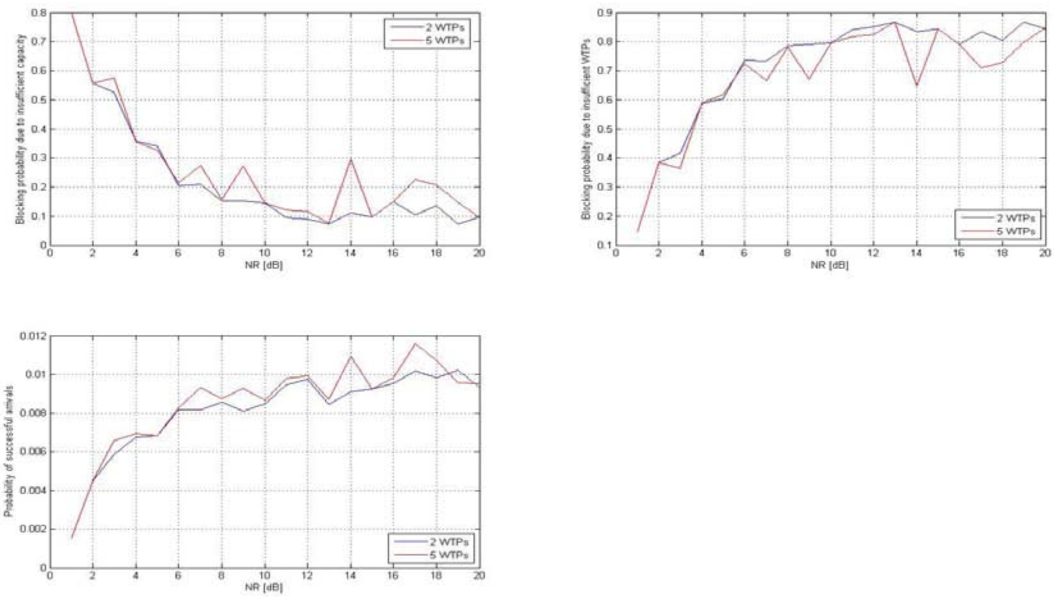


Figure 3.31: Probabilities vs. NR - 5-WTP.

In the figure, the increase in the blocking probability due to insufficient capacity seems to be offset by the decrease in the blocking probability due to insufficient WTP levels. As a result, signalling components in Table 3.5 are employed more, replacing those in Table 3.6. It is obvious that the number of UL and Netw2Netw signalling components in Table 3.5 are less than those of them in Table 3.6. This is what causes the lows in the UL and Netw2Netw average signalling loads when compared with the 2-WTP case. In contrast, the numbers of the DL signalling component in those two tables are the same.

We also see in the figure that the probability of successful arrivals increases, which means signalling components in Tables 3.7, 3.8 and 3.9 are now used more than before. This causes increases in the UL and Netw2Netw signalling loads. Because these increases are offset by the decreases of UL and Netw2Netw signalling loads mentioned in the above paragraph, the peaks are only marginally higher when compared with the 2-WTP case. On the other hand, for the DL, the increase of signalling now being used more due to signalling components in Tables 3.7, 3.8 and 3.9 is not offset by any reduction. Therefore, we see clearer the increase in the DL average signalling load in Figure 3.28.

The average simulation time for this simulation scenario of the small system 5-WTPs case is 1185 seconds. When compared with the 2-WTP case, this is an increase of about 35%.

3.8.7 Large system

In the previous section, we investigate how more types of WTP in the system affect Smart Pricing signalling requirements. In this section, we investigate the effect of having more users in the system on the signalling requirements. The 2-WTP case is chosen for this system.

Probability of occupancy

From parameter values in Table 3.12, the numbers of high WTP users with η_T and η_M and the maximum possible number of voice users in the system are calculated and summarised in Table 3.32. Different from the small system case, the NR range that we target here is a range in which the network capacity is low. This is part of an effort in widening our investigation into the impact of Smart Pricing signalling on the UMTS system.

NR [dB]	0.1	0.2	0.3	0.4	0.5	0.6
η_M	2.28%	4.50%	6.67%	8.80%	10.87%	12.90%
High WTP users with η_T	1	2	3	5	6	7
High WTP users with η_M	2	4	6	8	10	12
Max users possible	2	5	8	11	14	16

Table 3.32: Voice case and numbers of users the system can accommodate.

Figure 3.33 shows the network occupancy and probabilities of occupancy for the voice base case (with $NR = 0.6$ $[dB]$). It can be seen that the probabilities of having zero to six users in the system are virtually zero. This is due to the high arrival rate $\lambda = 3$ [arrivals/s], as the network will almost always be filled with the maximum number of high WTP users accommodated by η_T. The probability of seven users in the system is therefore the highest. It can also be seen that the probability of the system having more than the number of high WTP users accommodated with η_M is very small. In this instance, we see that the maximum possible number of users that the system can accommodate is not reached.

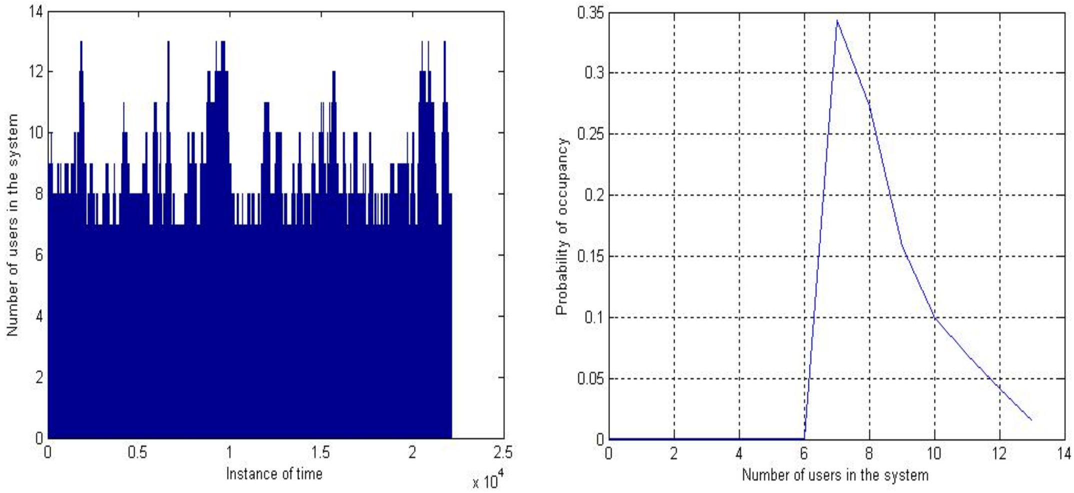


Figure 3.33: Network occupancy - Voice base case.

NR and the required signalling

Figure 3.34 shows the average signalling loads increase as *NR* increases.

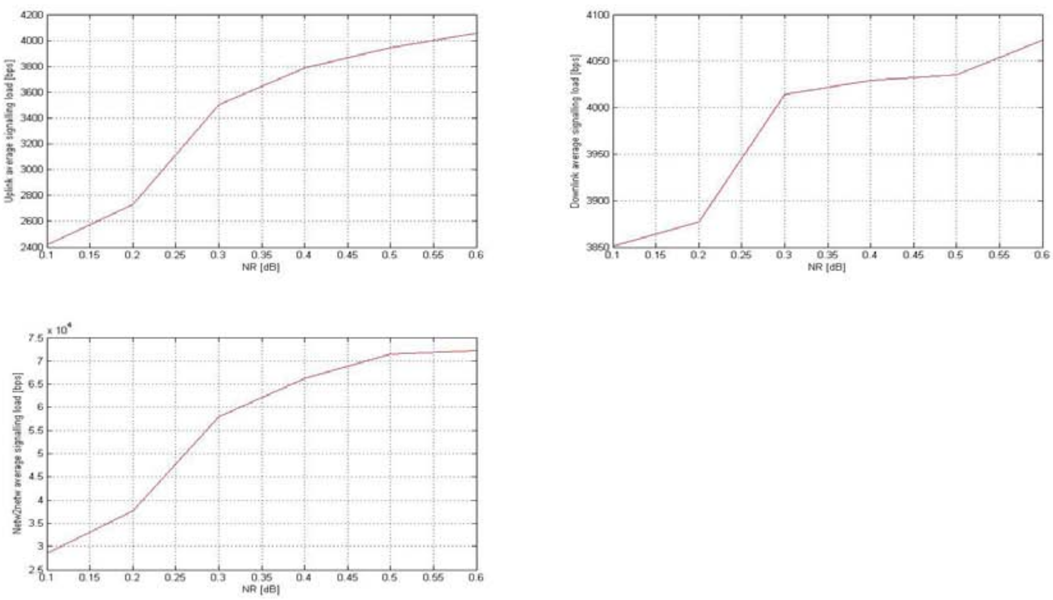


Figure 3.34: Average signalling load vs. NR - Large system.

The increases are continuing throughout the range of *NR* and not becoming steady like the small system case. This is because now with every *NR* increase step, there is enough capacity for accommodating more users. The UL peak is the lowest, the DL is 1.004 times greater than the UL, and the Netw2Netw peak is the highest and 17.79 times greater the UL. All average signalling loads are higher when compared with the small system case. See Figure 3.16 for details.

The percentages of bidding signalling relative to the respective total signalling are seen in Figure 3.35 to increase as *NR* increases.

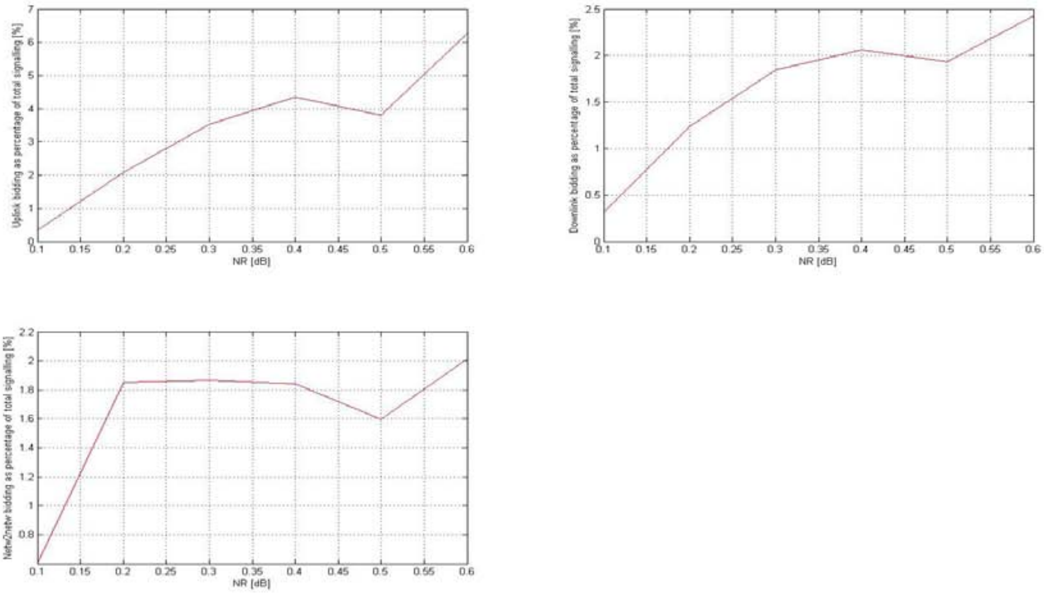


Figure 3.35: Bidding signalling as percentage of total signalling vs. NR - Large system.

However, in the *NR* range of 0.2-0.4 the percentages of Netw2Netw bidding signalling are steady. When compared with the small system case, the percentages are higher, lower and lower, for the UL, DL and Netw2Netw respectively. Refer to Figure 3.17 for comparison. This can be explained as the following.

Looking at Figure 3.37, we see a 4.39% increase in the blocking probability due to insufficient WTP. This means signalling components in Table 3.6 are used more. As these components are non-bidding signalling, hence percentages of bidding signalling are lower because the percentages are relative to the DL and Netw2Netw average signalling. This is the case for the DL and Netw2Netw. For the UL, on the other hand, as bidding signalling are dependent on the number of high WTP users, the increase in the number of high WTP users (due to the change from CONF to voice service) causes significant increase in UL bidding signalling components. This large increase overshadows the non-bidding signalling, hence resulting in an increase in the UL bidding signalling.

In Figure 3.36, it can be seen that the percentages of cost un-recoverable signalling relative to the respective total signalling decrease as *NR* increases. At their minimums, all percentages are above 92% and lower than those of them for the small system case.

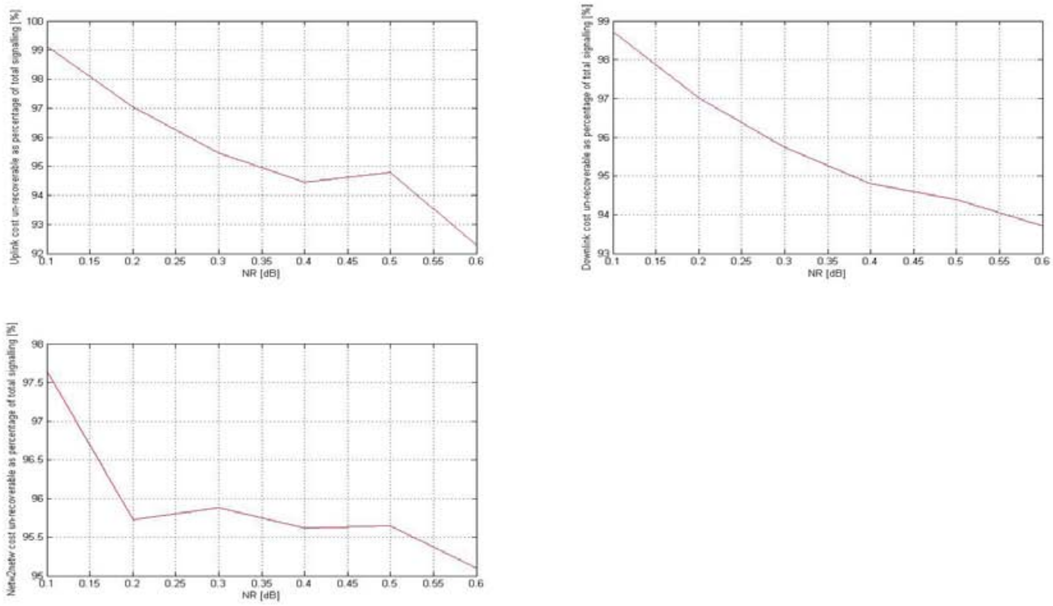


Figure 3.36: Cost un-recoverable signalling as percentage of total signalling vs. NR - Large system.

Figure 3.37 shows that the blocking probability due to insufficient capacity, blocking probability due to insufficient WTP and probability of successful arrivals decreases, increases and increases respectively as *NR* increases. Also, when compared with the small system case, we see that these probabilities are lower, higher and higher, respectively. This is because voice service bit rate is lower than CONF, so the network can accommodate more users. As the arrival rate remains the same, more users admitted into the network means lower blocking probability due to insufficient capacity and higher probability of successful arrivals. With Smart Pricing, when a potentially new user arrives, capacity is checked first to ensure the network can accommodate the user. After passing the first check, the user's WTP will then be checked against the admission price. For low WTP users, it is this second check that they more frequently fail. Hence, an increase in the blocking probability due to insufficient WTP.

The average simulation time for this large system is 1152 seconds. When compared with the CONF 2-WTP case, this is an increase of about 30%.

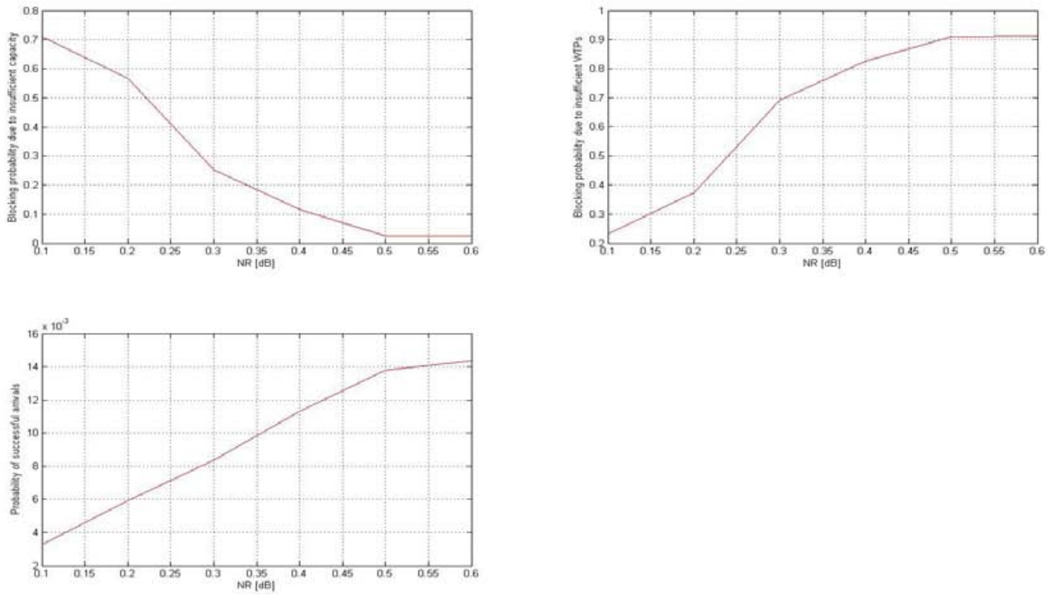


Figure 3.37: Probabilities vs. NR - Large system.

3.9 Findings and discussions

Given:

- the scope of this research project specified in Section 1.1.2;
- the signalling components required for Smart Pricing specified in Section 3.7.2;
- simulation scenarios that we conducted; and
- parameters' values for the simulations as specified in Table 3.12,

in this chapter we develop our first Smart Pricing signalling model. In this model, we investigate the required signalling for the implementation of Smart Pricing system in the UMTS. We model the signalling when the system is in its instantaneous conditions, in which its operation is simulated for a short duration of 2 hours. This first Smart Pricing signalling model is named the *Monte Carlo Simulation* (MCS) model.

The MCS model examines both small and large systems which represent systems that can accommodate small and large numbers of users. CONF is the service chosen for the small system, and voice for the large. We inspect almost the complete range of the system capacity, i.e. *NR*, from 2.28 to 99%. With the small system, two cases are explored: the less (2 WTPs) and the more (5 WTPs) types of users. In effect, this means three systems are examined. Furthermore, for each system, three broad categories of system characteristics are monitored, and together with their sub-categories a total of six simulation scenarios are simulated for each system. This brings the total simulation scenarios that we investigate to eighteen. The number reflects our effort in testing thoroughly our MCS Smart Pricing signalling model. For simplicity we show only results that exhibit distinctive system behaviours.

Table 3.38 provide a summary of the results of the *NR* simulation scenarios for the three systems. Apart from the probabilities, all values are averages.

Parameter	Case									Unit
	CONF 2 WTPs			CONF 5 WTPs			Voice (2WTPs)			
	UL	DL	Netw2Netw	UL	DL	Netw2Netw	UL	DL	Netw2Netw	
Signalling load max	3.89	4.06	68.53	3.90	4.10	68.54	4.06	4.07	72.16	kbps
Signalling-load-max ratio to UL average	1.00	1.04	17.61	1.00	1.05	17.60	1.00	1.004	17.79	
Bidding signalling max	5.63	2.52	2.29	10.01	3.81	3.78	6.29	2.43	2.02	%
Cost un-recoverable min	93.35	94.48	95.33	89.16	92.83	93.19	92.28	93.71	95.09	%
Blocking probability min - insufficient capacity	7.23			7.50			2.40			%
Blocking probability max - insufficient WTPs	86.83			86.67			91.22			%
Probability of successful arrivals max	1.02			1.16			1.44			%
Simulation time	881			1185			1152			seconds

Table 3.38: Summary of results.

It can be seen that:

1. average signalling loads are reasonably low, except for the Netw2Netw link. Average signalling load for the UL is the lowest, followed by the DL then the Netw2Netw. At their maximums, the UL-to-DL average signalling ratio is 1.04, whereas the Netw2Netw-to-DL is 17.79. Overall maximum average signalling loads for the UL, DL and Netw2Netw are 4.06, 4.10 and 72.16 kbps, respectively. The high Netw2Netw signalling we see here is the sum of signalling components between all network elements, not just between one pair of network elements. Consequently, the Netw2Netw average signalling load between one pair of network elements are much less than the value we see. Such a signalling load should easily be handled by the existing 56-64 kbps DS0 channel used in the common channel signalling SS7.

   When there are more types of WTP in the system, average signalling loads increase. The same is true when there are more users;

2. maximum bidding signalling percentage has a range of 2 to 10%. The percentage is highest for the UL, followed by the DL then the Netw2Netw, which is the opposite of the average signalling loads.

When there are more types of WTP in the system, bidding signalling percentages increase for all links. However, when there are more users, only the UL bidding signalling percentage increases, not the DL and Netw2Netw whereby they decrease;

3. maximum cost un-recoverable signalling percentage has a range of 89 to 95%. When there are more types of WTP in the system, the percentage decrease. The same is true when there are more users.

   As discussed in Section 3.7.1, cost un-recoverable signalling is a parameter that we design to assist the network operator in setting prices more effectively. The cost un-recoverable signalling percentages that we have found here should be taken into consideration when setting admission prices and prices for t_{gQ} expired users to maintain admission QoS in the events of graceful degradation. The percentages are also a powerful indication of how profoundly successful arrivals bear the cost imposed by the unsuccessful ones. A more remarkable point to note is who those unsuccessful ones are? Are they those who have immediate needs to make important calls right at those moments. Or, are they those who are opportunistic users probing for cheap-price calls? If the latter is true and the number of such users are large, it could seriously cause congestion on the signalling network;

4. minimum blocking probability due to insufficient capacity has a range of 2.4 to 7.5%. With such a range, high WTP user's customer satisfaction will be met. When there are more types of WTP in the system, the probability increases. However, when there are more users, the probability decreases;

5. maximum blocking probability due to insufficient WTP has a range of 86 to 91%. When there are more types of WTP in the system, the probability decreases. However, when there are more users, the probability increases;

6. maximum probability of successful arrivals has a range of 1 to 1.4%. When there are more types of WTP in the system, the probability increases. The same is true when there are more users;

7. all parameter values in the six items mentioned above are dependent on the signalling components set out in Section 3.7.2 for every situation that Smart Pricing signalling is required;

8. average simulation time has a range of 881 to 1185 seconds. When there are more types of WTP in the system, the simulation time increases. The same is true when there are more users;

9. for both small (2-WTP and 5-WTP cases) and large systems, as *NR* increases, the:

   a) average signalling loads increase;

   b) bidding signalling percentages increase;

   c) cost un-recoverable percentages decrease;

   d) blocking probability due to insufficient capacity decreases;

   e) blocking probability due to insufficient WTP increases; and

   f) probability of successful arrival increases;

   For the small system 2-WTP case, the above parameter values are steady when the system reaches its maximum possible number of users.

10. performance of the MCS model in terms of average simulation times are high. As an example, for the *NR* scenario:

    a) small system 2-WTP case takes 881 seconds;

    b) small system 5-WTP case takes 1185 seconds; and

    c) large system takes 1152 seconds.

11. for the small system, in most cases the results of the 2-WTP case are seen to be the upper bounds for the 5-WTP. Thus, with good prediction about offsets for the average bidding signalling loads, the 2-WTP case suffices to be representative for the 5-WTP.

With respect to our simulation strategy as discussed in Section 3.8.3, for the small system 2-WTP case, we have found that:

1. Conditions a MS is admitted to the network are: as

   a) *NR* increases, average signalling loads increase;

   b) mean WTP increases, the UL and Netw2Netw average signalling loads decrease, while the DL increases; and

   c) arrival rate increases, average signalling loads increase.

   NR, WTP and arrival rate heavily affect Smart Pricing signalling requirements (i.e. the average signalling loads). This suggests that with good admission control scheme, signalling for Smart Pricing can be controlled to a desired level;

2. Level congestion is defined: as

   a) η_T increases, average signalling loads decrease; and

   b) t_{gQ} increases, average signalling loads decrease.

   It is seen that the higher the η_T and t_{gQ} values, the lower the signalling required for Smart Pricing. At what value t_{gQ} should be set is a trade-off matter for network operators. These suggest that network operators need to balance resource dedicated for Smart Pricing signalling and users' satisfaction when defining the congestion level; and

3. Users' behaviour when they are already in the system: as mean hold time increases, average signalling loads decrease. We note that when mean holding time equals 30 seconds (i.e. the shortest period) the Netw2Netw average signalling load is the highest among all six simulation scenarios. Too high a price gives negative incentives to users to make longer calls. Consequently, not only the revenue that are generated from active users is reduced, but also the risk of an overload on the signalling network is high. In turn, the latter causes possible further loss of revenue because new users will not be able to enter the network due to congestion on the signalling network.

These findings mean that it is not how the level of congestion is set by a network operator that matters, but the conditions under which MSs are admitted to the network (i.e. the admission control mechanism) and the behaviour of users. A closer look into aspects of the conditions MSs are admitted to the network reveals that ultimately it is the users' behaviour that dictates the signalling requirements for Smart Pricing.

We recommend that:

Rec.1. a mechanism should be put in place to prevent opportunistic users probing for cheap-price calls;

Rec.2. signalling components for every situation that Smart Pricing signalling is required should be carefully designed; and

Rec.3. care should be taken in setting high prices to avoid the risk of overloading the signalling network.

CHAPTER 4

Modelling Smart Pricing signalling with state space

4.1 Justification for the chosen modelling technique

Signalling requirements for Smart Pricing as a whole are a random process. The process depends not only on:

- deterministic factors such as thresholds η_T and t_{gQ}; but also on
- a discrete random distribution: Poisson distribution (distribution of call arrivals) with parameter λ (arrival rate); as well as
- continuous distributions: negative exponential distribution (distribution of call holding times) with parameter μ, and Weibull distribution (distribution of user WTPs) with scale parameter δ and shape parameter β.

To determine probabilities that these signalling requirements occur, one technique is to combine these probability distributions to form a joint distribution. Determining joint distributions for discrete distributions or continuous distributions are difficult but achievable, at least numerically. However, determining a joint distribution for mixed discrete and continuous distributions is arguably a PhD project in itself. Thus, we have chosen not to follow that direction as this is an engineering research project.

Another technique is to use a state space approach. The technique enables us to incorporate all above-mentioned factors in a simpler fashion. Best of all, it allows us to expand the model but still requires the use of system of linear equations to solve. In addition, equilibrium probabilities found using this technique can be interpreted as the long-term proportion of time the system is in a state that requires certain signalling, or as the probabilities at a arbitrary time point, far away from the initial condition that the system is in a state that requires that amount of signalling [53].

4.2 Limiting the scope of the model for modelling

Signalling requirements for the full version of Smart Pricing have been discussed in Section 3.5 and Section 3.6. To model Smart Pricing in a state space, it is necessary to limit the scope of that full model for the sake of simplicity and convenience of modelling. This level of simplicity is necessary so that core behaviours of the model can easily be revealed and then investigated, without being blurred by other obstructive complexity. In this simplified Smart Pricing model, we:

- consider a system with one service;
- assume that there are two levels of WTP; Low (L) WTP and High (H) WTP. Together with the assumption of one service, Expression (3.2) and Expression (3.13) seen earlier in Chapter 3 become:

$$\boldsymbol{\Gamma} = \{\Gamma_L, \Gamma_H\} \tag{4.1}$$

- assume that there are two types of bit rates. Expression (3.4) and Expression (3.14) then become:

$$\boldsymbol{R} = \{R_L, R_H\} \tag{4.2}$$

- see that Expression (3.5) and Expression (3.15) have now reduced to one expression only, i.e. Expression (4.3) below. R_L^a is the bit rate for Γ_L users with expired t_{gQ}, whereas R_H^a is the bit rate for Γ_H users with expired t_{gQ} and for users who are still within their t_{gQ}.

$$\boldsymbol{R^a} = \{R_L^a, R_H^a\} \tag{4.3}$$

- assume that there are two prices; one for R_L and the other for R_H. Expression (3.8) and Expression (3.18) then reduce to:

$$\boldsymbol{P} = \{P_L, P_H\} \tag{4.4}$$

  We also assume that $\Gamma_L = P_L$ and $\Gamma_H = P_H$.

- see that Expression (3.9) and Expression (3.19) reduce to one expression only, i.e. Expression(4.5) below. P_L^a is the price for Γ_L users with expired t_{gQ}, whereas P_H^a is the price for Γ_H users with expired t_{gQ} and for users who are still within their t_{gQ}.

$$\boldsymbol{P^a} = \{P_L^a, P_H^a\} \tag{4.5}$$

4.3 Initial calculation

From Equation (2.8), the load factor of a user with bit rate R_L is:

$$L_{R_L} = \frac{1}{1 + \dfrac{W}{(E_b/N_0) \cdot \upsilon \cdot R_L}} \tag{4.6}$$

and, keeping the same values for (E_b/N_0) and υ, the load factor of a user with bit rate R_H is:

$$L_{R_H} = \frac{1}{1 + \dfrac{W}{(E_b/N_0) \cdot \upsilon \cdot R_H}} \tag{4.7}$$

From Equation (3.32), the potential load factor to admit a new user while guaranteeing the *minimum load condition* in the system is:

$$L^{pot,min} = (g + y + h + 1)\, L_{R_H} + l \cdot L_{R_L} \tag{4.8}$$

where g, y and h are defined ahead in Section 4.4.

From Equation (3.34), the potential load factor to admit a new user while guaranteeing the *maximum load condition* in the system is:

$$L^{pot,max} = (g + y + h + l + 1)\, L_{R_H} \tag{4.9}$$

where l is defined ahead in Section 4.4.

As λ is the arrival rate, which follows a Poisson distribution, the probability of one or more arrivals to the network at a time instance is:

$$Prob_{Arrival} = 1 - e^{-\lambda} \tag{4.10}$$

As Γ is the willingness to pay (WTP) of a user, which follows a Weibull distribution, Cumulative Distribution Function (CDF) of Γ at a certain γ WTP value is [54]:

$$F_\Gamma(\gamma) = 1 - e^{-\left(\frac{x}{\delta}\right)^{\beta}} \tag{4.11}$$

Hence, the probability of a high WTP user is:

$$Prob(\Gamma \geq P_H) = 1 - F_\Gamma(P_H) \tag{4.12}$$

and probability of a low WTP user is:

$$Prob(P_L \leq \Gamma < P_H) = F_\Gamma(P_H) - F_\Gamma(P_L) \tag{4.13}$$

4.4 State space

From Section 4.2, it is seen that there are four types of users in a cell. Let us call:

- g the number of Γ_H users in a cell who are still in their non-expired period t_{gQ}. We call these users g-type users;
- y the number of Γ_L users in a cell who are still in their t_{gQ}. We call these users y-type users;
- h the number of users with t_{gQ} expired that have high WTP Γ_H. We call these users h-type users; and
- l the number of users with t_{gQ} expired that have low WTP Γ_L. We call these users l-type users.

Let a state i be (g, y, h, l). There are 8 possible states that a state can move to and each of these moves is associated with a transition rate. Figure 4.1 provides a visual representation of all of those states, state transitions and transition rates.

In Figure 4.1:

- q_{g+} is the transition rate from state i to a state in which a new Γ_H user is admitted into the cell;
- q_{y+} is the transition rate from state i to a state in which a new Γ_L user is admitted into the cell;
- q_{g-} is the transition rate from state i to a state in which a g-type user hangs up;
- q_{y-} is the transition rate from state i to a state in which a y-type user hangs up;
- q_{h+} is the transition rate from state i to a state in which a Γ_H user has his/her t_{gQ} expired;

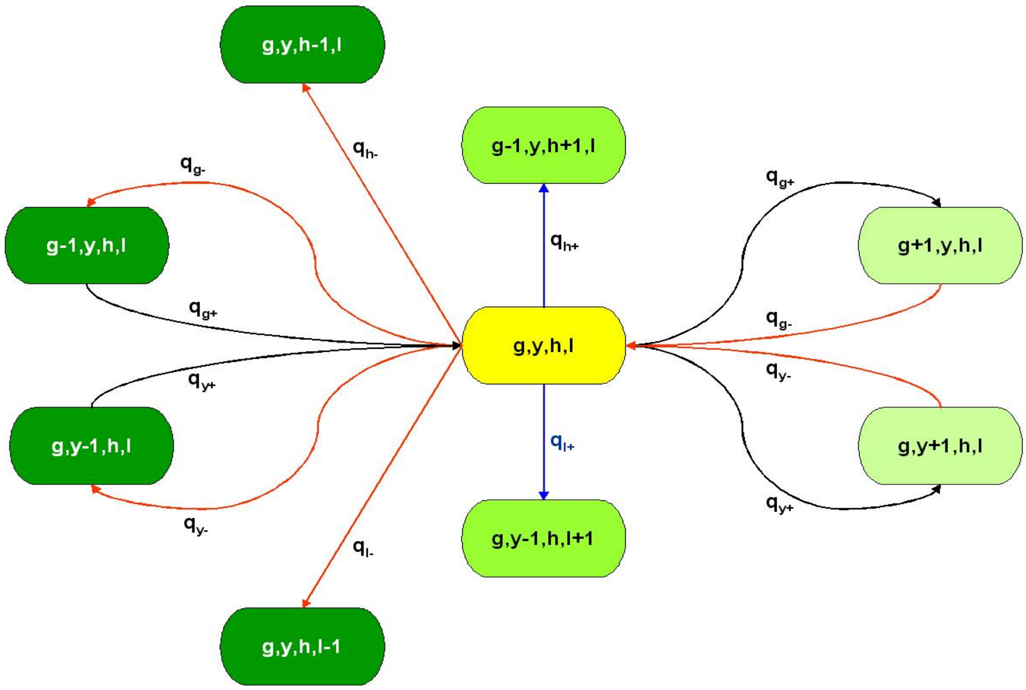


Figure 4.1: State transition digram.

- q_{l+} is the transition rate from state i to a state in which a Γ_L user has his/her t_{gQ} expired;
- q_{h-} is the transition rate from state i to a state in which a h-type user hangs up; and
- q_{l-} is the transition rate from state i to a state in which a l-type user hangs up.

The state space of Smart Pricing system is discussed in Section 4.6. Generally, the total number of states depends on the total number of users in the system and on the number of each of the four types of users that constitute that total number of users. All of these numbers of users are dependent on η_M and η_T.

4.5 Transition rates

Formulas to calculate the eight possible transition rates from any state i discussed in Section 4.4 are provided below. In addition to those eight transition rates, there is a ninth, which is the transition rate when the system remains in a state, q_{ii}. Transition rates other than these nine rates are all zero as there is no possibility that the system could transition to or from any other state.

Rate q_{g+}

$$q_{g+} = I(L^{pot,min} \leq \eta_M) \cdot \lambda \cdot Prob(\Gamma_H \geq P^{adm}(i)) \tag{4.14}$$

where I is an indicator function[1], which allows us to impose a condition for the transition rate to occur. It enables us to model a deterministic aspect of Smart Pricing into this Smart Pricing signalling model. $P^{adm}(i)$ is admission price when the system is in state i, and

$$Prob(\Gamma_H \geq P^{adm}(i)) = Prob(\Gamma \geq P_H) \tag{4.15}$$

Rate q_{g-}

$$q_{g-} = g(i) \cdot \frac{1}{\overline{ht}} \tag{4.16}$$

where $g(i)$ is the number of g-type users when the system is in state i and $\overline{ht}$ is mean hold time of calls as discussed in Section 3.5.2.

Rate q_{y+}

$$q_{y+} = I(L^{pot,min} \leq \eta_M) \cdot \lambda \cdot Prob(\Gamma_L \geq P^{adm}(i)) \cdot I(gyh(i) + 1 <= M_{\eta_T}^{R_H}) \tag{4.17}$$

where:

$$Prob(\Gamma_L \geq P^{adm}(i)) = \begin{cases} Prob(P_L \leq \Gamma < P_H) & \text{if } P^{adm}(i) = P_L \\ 0 & \text{if } P^{adm}(i) = P_H \end{cases} \tag{4.18}$$

and

$$gyh(i) = g(i) + y(i) + h(i)$$

$y(i)$ and $h(i)$ are the numbers of y- and h-type users when the system is in state i. $M_{\eta_T}^{R_H}$ is the number of Γ_H users that can be accommodated with η_T. See Section 4.6.2 for details.

Rate q_{y-}

$$q_{y-} = y(i) \cdot \frac{1}{\overline{ht}} \tag{4.19}$$

[1] $I(f) = 1$ if condition f is true and $= 0$ if false.

Rate q_{h+}

$$q_{h+} = g(i) \cdot \frac{1}{t_{gQ}} \tag{4.20}$$

Rate q_{h-}

$$q_{h-} = h(i) \cdot \frac{1}{\overline{\overline{ht}}} \tag{4.21}$$

Note that q_{h-} could easily be confused as:

$$q_{h-} = h(i) \cdot \frac{1}{\overline{\overline{ht}} - t_{gQ}} \tag{4.22}$$

but that is wrong. If we assume hold times of calls are negatively exponentially distributed, the rate that calls finish at any time is the same, despite the calls having already been in progress for t_{gQ} seconds. This is due to the memoryless property of the negative exponential distribution.

Rate q_{l+}

$$q_{l+} = y(i) \cdot \frac{1}{t_{gQ}} \tag{4.23}$$

Rate q_{l-}

$$q_{l-} = l(i) \cdot \frac{1}{\overline{\overline{ht}}} \tag{4.24}$$

Note that the caveat given in Equation (4.22) also applies to Equation (4.24).

Rate q_{ii}

This transition rate is defined as the negated sum of all the above eight transition rates.

$$q_{ii} = -(q_{g+} + q_{g-} + q_{y+} + q_{y-} + q_{h+} + q_{h-} + q_{l+} + q_{l-}) \tag{4.25}$$

4.6 Maximum possible total number of users

From a target *NR* value, using the procedure in Section 2.9, η_M is obtained. Then, using Equation (3.1), η_T is obtained. The maximum possible total number of users in the system must satisfy the below inequality:

$$(g + y + h) \cdot L_{R_H} + l \cdot L_{R_L} \leq \eta_M \tag{4.26}$$

4.6.1 Number of users in phase 1

According to Smart Pricing principles, when the system load factor is up to η_T, all users in the system get R_H. This means all g-, y-, h- and l-type users get R_H. The load factor of a user with R_H can be determined by using in Equation (4.7). Therefore, the number of users that can be accommodated with η_T, $M_{\eta_T}^{R_H}$, is:

$$M_{\eta_T}^{R_H} = Integer\left(\frac{\eta_T}{L_{R_H}}\right) \tag{4.27}$$

where $Integer(X)$ is the integer part of X.

We name *Phase 1* the phase in which the number of users in the system is within the range [0,$M_{\eta_T}^{R_H}$]. That is, the maximum possible number of users in this phase is $M_{\eta_T}^{R_H}$. A distinct characteristic of this phase is that the maximum possible number of y- and l-type users is also $M_{\eta_T}^{R_H}$.

4.6.2 Number of users in phase 2

The maximum number of users with R_H that can be accommodated with η_M, $M_{\eta_M}^{R_H}$, is:

$$M_{\eta_M}^{R_H} = Integer\left(\frac{\eta_M}{L_{R_H}}\right) \tag{4.28}$$

We name *Phase 2* the phase in which the number of users in the system is within the range ($M_{\eta_T}^{R_H}$,$M_{\eta_M}^{R_H}$]. That is the maximum possible number of users in this phase is ($M_{\eta_M}^{R_H} - M_{\eta_T}^{R_H}$). A distinct characteristic of this phase is that the maximum possible number of y- and l-type users remains at $M_{\eta_T}^{R_H}$. No y-type users will be admitted to the system once the system load factor is above η_T.

4.6.3 Number of users in phase 3

From Section 4.6.2, we know that the maximum possible number of users at R_H that η_M can accommodate is $M_{\eta_M}^{R_H}$. As R_H only needs to be given to g-, y- and h-type users, if l-type users exist in the system, bit rate of their calls can be reduced to R_L once t_{gQ} of the calls are expired. From Section 4.6.1 and Section 4.6.2, we also know that the maximum possible number of l-type users is $M_{\eta_T}^{R_H}$. Thus, Equation (4.26) becomes:

$$(g + y + h) \cdot L_{R_H} + M_{\eta_T}^{R_H} \cdot L_{R_L} \leq \eta_M \tag{4.29}$$

which is equivalent to:

$$(g + y + h) \leq \frac{\eta_M - (M_{\eta_T}^{R_H} \cdot L_{R_L})}{L_{R_H}} \tag{4.30}$$

Therefore, the maximum possible number of users in the system, M_{max}, is:

$$M_{max} = M_{\eta_T}^{R_H} + x \tag{4.31}$$

where x is the maximum integer of $(g + y + h)$ that satisfies Equation (4.30). Note that M_{max} has a minimum value of $M_{\eta_M}^{R_H}$ and it is not always greater than that value. M_{max} is equal to $M_{\eta_M}^{R_H}$ when L_{R_L} is increasing closer to or equal to L_{R_H}. If M_{max} is equal to $M_{\eta_M}^{R_H}$, the possible number of users for this phase will be zero.

We name *Phase 3* the phase in which M_{max} is greater than $M_{\eta_M}^{R_H}$. The number of users in this phase is within the range $(M_{\eta_M}^{R_H}, M_{max}]$ with a maximum value of $(M_{max} - M_{\eta_M}^{R_H})$. Distinct characteristics of this phase are that:

- the maximum possible combined number of g- and l-type users is $M_{\eta_T}^{R_H}$;
- the maximum possible number of users of any type must be less than M_{max}; and
- the maximum possible combined number of g-, y- and h-type users, $M_{gyh,p3}$, is the maximum integer of $(g + y + h)$ that satisfies the below inequality:

$$(g + y + h) \leq \frac{\eta_M - L_{R_L}}{L_{R_H}} \tag{4.32}$$

4.7 Maximum possible number of states

Based on the numbers of users obtained from Section 4.6, we can calculate the maximum number of states in the system's state space. We will be using *composition* for this calculation.

A composition [55] of a positive integer m is a way of writing m as a sum of strictly positive integers. The number of compositions of m into exactly k parts are given by the binomial coefficient:

$$\binom{m-1}{k-1} \tag{4.33}$$

If some of the parts in a composition are allowed to be zero, the composition is referred to as a *weak composition*. For a weak composition, the number of possible combinations are given by:

$$\binom{m+k-1}{k-1} \tag{4.34}$$

Note that:

$$\binom{m}{k} = \frac{m!}{k!\,(m-k)!} \tag{4.35}$$

Thus:

$$\binom{m+k-1}{k-1} = \frac{(m+k-1)!}{(k-1)!\,m!} \tag{4.36}$$

In our model, m represents the total number of users in the system, i.e. the sum of g, y, h and l. As a state i contains four parts, the number of parts, k, is 4.

The maximum possible number of states of the system, N_s, is the sum of the numbers of states in three phases discussed in the next three sub-sections. That is:

$$N_s = N_1 + N_2 + N_3 \tag{4.37}$$

where N_1, N_2 and N_3 are the number of states in phase 1, phase 2 and phase 3, respectively.

4.7.1 Number of states in phase 1

In Section 4.6.1, we have found that the number of users in the system in phase 1 takes values in the $[0, M_{\eta_T}^{R_H}]$ range. Let this number be m. For each value of m in that range, using weak composition, the possible number of states are:

$$\begin{aligned} N_1^m &= \binom{m+k-1}{k-1} \\ &= \binom{\eta_T^{R_H}+3-1}{3-1} \\ &= \binom{\eta_T^{R_H}+2}{2} \end{aligned} \tag{4.38}$$

where N_1^m is the number of states when there are m users in the system.

Hence, the total possible number of states in phase 1 is:

$$N_1 = \sum_{m=0}^{M_{\eta_T}^{R_H}} N_1^m \tag{4.39}$$

4.7.2 Number of states in phase 2

In Section 4.6.2, we have found that the number of users in the system in phase 2 takes value in the $(M_{\eta_T}^{R_H}, M_{\eta_M}^{R_H}]$ range. Let this number be $\hat{m}$. For each value of $\hat{m}$ in that range, using weak composition, we have:

$$\binom{\hat{m}+k-1}{k-1} \tag{4.40}$$

As the maximum possible number of g- and l-type users in phase 2 is $M_{\eta_T}^{R_H}$, we need to exclude compositions found in (4.40) that have the combined value of g- and l-type users greater than $M_{\eta_T}^{R_H}$.

For each value of $\hat{m}$ in the $(M_{\eta_T}^{R_H}, M_{\eta_M}^{R_H}]$ range, let the possible number of states be $N_2^{\hat{m}}$. Hence, the total possible number of states in phase 2 is:

$$N_2 = \sum_{m=M_{\eta_T}^{R_H}+1}^{M_{\eta_M}^{R_H}} N_2^{\hat{m}} \tag{4.41}$$

4.7.3 Number of states in phase 3

The calculation in this section is only required when M_{max} is greater than $M^{R_H}_{\eta_M}$. If M_{max} is equal to $M^{R_H}_{\eta_M}$, N_3 is equal to zero.

In Section 4.6.3, we found that the number of users in the system in phase 3 takes value in the ($M^{R_H}_{\eta_M}$,M_{max}] range. Let this number be $\breve{m}$. For each value of $\breve{m}$ in that range, using weak composition, we have:

$$\binom{\breve{m} + k - 1}{k - 1} \tag{4.42}$$

As the maximum possible combined number of y- and l-type users in phase 3 is $M^{R_H}_{\eta_T}$, we need to exclude compositions found in (4.42) those that have a sum of y- and l-type users greater than $M^{R_H}_{\eta_T}$. In addition, as the maximum possible combined number g- and h-type users in phase 3 is $M^{R_H}_{\eta_M}$, we also need to exclude compositions that a sum of g- and h-type users greater than $M^{R_H}_{\eta_M}$.

For each value of $\breve{m}$ in the ($M^{R_H}_{\eta_M}$,M_{max}] range, let the number of states be $N_3^{\breve{m}}$. Hence, the total number of states in phase 3 is:

$$N_3 = \sum_{\breve{m}=\eta_M^{R_H}+1}^{M_{max}} N_3^{\breve{m}} \tag{4.43}$$

4.8 Q matrix formulation

With a given set of values for parameters discussed in Section 4.4 through to Section 4.7, using the approaches in those sections, we can create a *Q matrix*. A *Q matrix* brings together all the possible number of states of the system and includes relevant transition rates between the states. An example of a *Q matrix* for a system with:

- $M^{R_H}_{\eta_T} = 1$;
- $M^{R_H}_{\eta_M} = 4$;
- $M_{gh,p3} = 4$; and
- $M_{max} = 5$.

is given in the Appendix[2].

[2]In that *Q matrix*, $-q_{\#}$ ("#" is the state number) represents $q_{\#\#}$.

4.9 Equilibrium probabilities

Equilibrium probabilities, also known as *stationary probabilities* [56] or *steady-state probabilities* [57], can be interpreted as the long-term proportion of time that a system is in a certain state [53]. We can calculate equilibrium probabilities for a system using the system *Q matrix*. Smart Pricing system can be modelled as a Continuous Time Markov Chain (CTMC).

A CTMC is called *irreducible* if every state in the state space is reachable from any other state. An *irreducible*, *homogeneous* (i.e. time independent) CTMC is called *ergodic* if and only if the unique equilibrium probability vector, $\boldsymbol{\pi}$, exists. For a CTMC, if $\boldsymbol{\pi}$ exists, elements of $\boldsymbol{\pi}$ are independent of time and we have [58]:

$$\boldsymbol{\pi}\boldsymbol{Q} = \boldsymbol{0} \tag{4.44}$$

where $\boldsymbol{\pi}$ is a [1xN_s] row vector; N_s is the number of states in the system found in Section 4.7. $\boldsymbol{Q}$ is the system *Q matrix* and is an [N_sxN_s] matrix. Thus, in order to use Equation (4.44) for Smart Pricing, we need to prove that Smart Pricing system is *irreducible*.

It can be shown that Smart Pricing system is *irreducible* because any state j can be reached from any state i ($i \neq j$) from a finite number of steps. There may be more than one way to reach state j from state i; one of such a way that guarantees for state i to reach state j is through state (0,0,0,0). This can easily be seen from Figure 4.1 that any state where the number of users in the system is greater than 0, q_{g-}, q_{y-}, q_{h-} and/or q_{l-} transition rates will be greater 0. That means, state (0,0,0,0) can be reached from any other states. This is because any state can go through a sequence of states, each of which the number of users in the system is less than the previous until it reaches a state at which the number of users in the system is 0, i.e. state (0,0,0,0). Clearly, from state (0,0,0,0), any other states can be reached. Hence, we can conclude that Smart Pricing system is irreducible.

To obtain a unique positive solution to Equation (4.44), a normalization condition needs to be imposed because $\boldsymbol{Q}$ is not invertible. We impose a normalization condition by substituting one column of $\boldsymbol{Q}$ by a unit vector. It can be any column of $\boldsymbol{Q}$ [58]. Let the new $\boldsymbol{Q}$ matrix after that substitution be $\boldsymbol{Q_m}$, Equation (4.44) then becomes:

$$\boldsymbol{\pi}\boldsymbol{Q_m} = \boldsymbol{b} \tag{4.45}$$

where $\boldsymbol{b}$ is a row vector $[0, 0, ..., 0, 1]$ with N_s elements.

Equation (4.45) is a linear system of non-homogenous equations in matrix form. To solve this equation, we use a direct numerical method called *Gaussian Elimination*. Direct numerical methods can be used to solve system of linear equations with a general structure. We made this choice because Gaussian Elimination is the most commonly applied among the direct methods. In this method, elementary matrix operations are applied and $\boldsymbol{Q_m}$ is transformed into a product of an *upper triangular matrix*, $\boldsymbol{U}$, and a *lower triangular matrix*, $\boldsymbol{L}$. That is:

$$\boldsymbol{\pi}\,\boldsymbol{U}\,\boldsymbol{L} = \boldsymbol{b} \tag{4.46}$$

Let:

$$\boldsymbol{\pi}\,\boldsymbol{U} = \boldsymbol{y} \tag{4.47}$$

Equation (4.46) then becomes:

$$\boldsymbol{y}\,\boldsymbol{L} = \boldsymbol{b} \tag{4.48}$$

Solve Equation (4.48) for $\boldsymbol{y}$, and then substitute $\boldsymbol{y}$ into Equation (4.47), we obtain the equilibrium probability vector $\boldsymbol{\pi}$. For more details of the Gaussian Elimination method, refer to [58].

4.10 Average signalling messages required

In this Section, we derive formulas for calculating average numbers of signalling messages for all possible scenarios. Important parameters in these formulas are equilibrium probabilities (which are elements of $\boldsymbol{\pi}$ obtained from Section 4.9). Before proceeding with derivation of the formulas, we need to calculate some parameter values first.

Successful arrival rate for state i, $Rate_{SA}(i)$, is:

$$Rate_{SA}(i) = q_{g+}(i) + q_{y+}(i) \tag{4.49}$$

where $q_{g+}(i)$ and $q_{y+}(i)$ are q_{g+} and q_{y+} of state i, respectively.

Blocking rate for state i, $Rate_{Blocking}(i)$, is:

$$Rate_{Blocking}(i) = \lambda - Rate_{SA}(i) \tag{4.50}$$

Hang up rate for state i, $Rate_{HU}(i)$, is:

$$Rate_{HU}(i) = q_{g-}(i) + q_{y-}(i) + q_{h-}(i) + q_{l-}(i) \tag{4.51}$$

where $q_{g-}(i)$, $q_{y-}(i)$, $q_{h-}(i)$ and $q_{l-}(i)$ are q_{g-}, q_{y-}, q_{h-} and q_{l-} of state i, respectively.

The rate that t_{gQ} of calls expires for state i, $Rate_{tgQ}(i)$, is:

$$Rate_{tgQ}(i) = q_{h+}(i) + q_{l+}(i) \tag{4.52}$$

where $q_{h+}(i)$ and $q_{l+}(i)$ are q_{h+} and q_{l+} of state i, respectively.

4.10.1 When a new user is blocked

A user is blocked from entering the network if the network does not have enough capacity to admit the user. The average number of signalling messages required when a new user is blocked from entering the network is:

$$\overline{N_1} = \left(\sum_{i \in \boldsymbol{S_1}} \pi_i \cdot Rate_{Blocking}(i) \right) \cdot \alpha_{blocked} \tag{4.53}$$

where:

- π_i is equilibrium probability of state i, i.e. the ith element of $\boldsymbol{\pi}$;
- $\boldsymbol{S_1}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_1} = \{ i : [g(i) + y(i) + h(i) + 1] \cdot L_{R_H} + l(i) \cdot L_{R_L} > \eta_M \} \tag{4.54}$$

- $\alpha_{blocked}$ can be found from Equation (3.35).

The overall blocking probability due to insufficient capacity of the system, $Prob_{bc}$, is defined as:

$$Prob_{bc} = \left(\sum_{i \in \boldsymbol{S_1}} \pi_i \right) \cdot Prob_{Arrival} \tag{4.55}$$

4.10.2 When a new user's WTP is sufficient

A new user is successfully admitted to the network if both his/her WTP is greater than or equal to the admission price at the time he/she arrives and the network has enough capacity. The average number of signalling messages required when a new user is successfully admitted to the network is:

$$\overline{N_2} = \left(\sum_{i \in \boldsymbol{S_2}} \pi_i \cdot Rate_{SA}(i) \right) \cdot \alpha_{ypa} \tag{4.56}$$

where α_{ypa} can be found from Equation (3.37).

- $\boldsymbol{S_2}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_2} = \{ i : [g(i) + y(i) + h(i) + 1] \, L_{R_H} + l(i) \cdot L_{R_L} \leq \eta_M \} \tag{4.57}$$

- α_{npa} can be found from Equation (3.36).

The overall probability of successful arrivals, $Prob_{sa}$, is defined as:

$$Prob_{sa} = \left(\sum_{i \in \boldsymbol{S_2}} \pi_i \cdot (Prob(\Gamma_H \geq P^{adm}(i)) + Prob(\Gamma_L \geq P^{adm}(i)) \right) \cdot Prob_{Arrival} \tag{4.58}$$

4.10.3 When a new user's WTP is not sufficient

A new user is rejected from entering the network if his/her WTP is lower than the admission price at the time he/she arrives (even if the network has enough capacity to admit). The average number of signalling messages required when the network has enough capacity to admit but the new user's WTP is not sufficient is:

$$\overline{N_3} = \left(\sum_{i \in \boldsymbol{S_3}} \pi_i \cdot Rate_{Blocking}(i) \right) \cdot \alpha_{npa} \tag{4.59}$$

where:

- $\boldsymbol{S_3} = \boldsymbol{S_2}$; and
- α_{npa} can be found from Equation (3.36).

The overall blocking probability due to high price, $Prob_{bp}$, is defined as:

$$Prob_{bp} = \left(\sum_{i \in \boldsymbol{S_3}} \pi_i \right) \cdot Prob_{Arrival} \tag{4.60}$$

The signalling required for this case is a distinct characteristic of Smart Pricing (and for dynamic pricing in general) compared to the flat-rate or peak and off peak pricing schemes.

4.10.4 Admit a new user and load to give $\boldsymbol{R_H}$ to all users not exceeding $\boldsymbol{\eta_T}$

Depending on whether the load to give all existing users in the system plus the new user's R_H exceeds η_T or not, the average signalling required is different. The required average number of signalling messages for the case when that load does not exceed η_T is:

$$\overline{N_4} = \left(\sum_{i \in \boldsymbol{S_4}} \pi_i \cdot Rate_{SA}(i) \right) \cdot \alpha_{nh} \tag{4.61}$$

where:

- $\boldsymbol{S_4}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_4} = \left\{ i : [i \in \boldsymbol{S_2}] \cap \left[M(i) + 1 \leq M_{\eta_T}^{R_H} \right] \right\} \tag{4.62}$$

- $M(i)$ is defined as the sum of users of all types when the system is in state i;
- $M_{\eta_T}^{R_H}$ is as defined in Section 4.6.1; and
- α_{nh} is can be found from Equation (3.38).

4.10.5 Admit a new user and load to give $\boldsymbol{R_H}$ to all users exceeding $\boldsymbol{\eta_T}$

The average number of signalling messages required for the case when the load to give all existing users plus the new user R_H exceeds η_T is:

$$\overline{N_5} = \sum_{i \in \boldsymbol{S_5}} \pi_i \cdot Rate_{SA}(i) \cdot \alpha_{ng}(i) \tag{4.63}$$

where:

- $\boldsymbol{S_5}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_5} = \left\{ i : [i \in \boldsymbol{S_2}] \cap \left[M(i) + 1 > M_{\eta_T}^{R_H} \right] \right\} \tag{4.64}$$

- $\alpha_{ng}(i)$ is α_{ng} when the system is in state i, which can be found from Equation (3.39).

4.10.6 When a user hangs up and load to give $\boldsymbol{R_H}$ to all users not exceeding $\boldsymbol{\eta_T}$

The average number of signalling messages required when a user hangs up and the load to give all remaining existing users R_H does not exceed η_T is:

$$\overline{N_6} = \left(\sum_{i \in \boldsymbol{S_6}} \pi_i \cdot Rate_{HU}(i) \right) \cdot \alpha_{hgh} \tag{4.65}$$

where:

- $\boldsymbol{S_6}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_6} = \left\{ i : 0 \leq M(i) - 1 \leq M_{\eta_T}^{R_H} \right\} \tag{4.66}$$

- α_{hgh} can be found from Equation (3.40).

4.10.7 When a user has his/her $\boldsymbol{t_{gQ}}$ expired and load to give $\boldsymbol{R_H}$ to all users not exceeding $\boldsymbol{\eta_T}$

The average number of signalling messages required when a user has his/her t_{gQ} expired and the load to give all existing users R_H does not exceed η_T is:

$$\overline{N_7} = \left(\sum_{i \in \boldsymbol{S_7}} \pi_i \cdot Rate_{tgQ}(i) \right) \cdot \alpha_{geh} \tag{4.67}$$

where $\boldsymbol{S_7}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_7} = \left\{ i : [g(i) > 0] \cap \left[0 < M(i) \leq M_{\eta_T}^{R_H} \right] \right\} \tag{4.68}$$

4.10.8 When a user hangs up and load to give $\boldsymbol{R_H}$ to all users exceeding $\boldsymbol{\eta_T}$

The average number of signalling messages required when a user hangs up and the load to give all existing users R_H exceeds η_T is:

$$\overline{N_8} = \sum_{i \in \boldsymbol{S_8}} \pi_i \cdot Rate_{HU}(i) \cdot \alpha_{hgg}(i) \tag{4.69}$$

where:

- $\boldsymbol{S_8}$ is a subset of the state space and is defined as:

$$\boldsymbol{S_8} = \left\{ i : M(i) - 1 > M_{\eta_T}^{R_H} \right\} \tag{4.70}$$

- $\alpha_{hgg}(i)$ is α_{hgg} when the system is in state i, which can be found from Equation (3.41).

4.10.9 When a user has his/her $\boldsymbol{t_{gQ}}$ expired and load to give $\boldsymbol{R_H}$ to all users exceeding $\boldsymbol{\eta_T}$

The average number of signalling messages required when a user has his/her t_{gQ} expired and the load to give all existing users R_H exceeds η_T is:

$$\overline{N_9} = \sum_{i \in \boldsymbol{S_7}} \pi_i \cdot Rate_{tgQ}(i) \cdot \alpha_{hgg}(i) \tag{4.71}$$

where:

$$\boldsymbol{S_9} = \left\{ i : [g(i) > 0] \cap \left[M(i) > M_{\eta_T}^{R_H} \right] \right\} \tag{4.72}$$

4.10.10 For some other important cases

Bidding

The average number of signalling messages required for the bidding process is:

$$\overline{N_{bidding}} = \overline{N_{5,bidding}} + \overline{N_{8,bidding}} + \overline{N_{9,bidding}} \tag{4.73}$$

where:

$$\overline{N_{5,bidding}} = \sum_{i \in \boldsymbol{S_5}} \pi_i \cdot Rate_{SA}(i) \cdot [h(i) + 5] \tag{4.74}$$

$$\overline{N_{8,bidding}} = \sum_{i \in \boldsymbol{S_8}} \pi_i \cdot Rate_{HU}(i) \cdot [h(i) + 5] \tag{4.75}$$

$$\overline{N_{9,bidding}} = \sum_{i \in \boldsymbol{S_9}} \pi_i \cdot Rate_{tgQ}(i) \cdot [h(i) + 5] \tag{4.76}$$

Note that 5 in Equations (4.74), (4.75) and (4.76) above are the sum of the cr_bdc and cr_bnc signalling components in Table 3.9 for the former and Table 3.11 for the latter two Equations.

The whole system

The average number of signalling messages required for the whole system is:

$$\overline{N_{system}} = \sum_{i=1}^{9} (\overline{N_i}) \tag{4.77}$$

where $\overline{N_i}$ is the average number of signalling messages required found in the relevant section among the previous nine sections.

4.11 Simulation and results

In order to compare the MCS Smart Pricing signalling model in Chapter 3 and the state space Smart Pricing signalling model in this chapter (so-called the SSA model), we use the same set of parameter values specified in Section 3.8.2. Simulation results of the MCS model's small system 2-WTP and large system cases are used to compare with those of the SSA model. The reason for the small system 5-WTP case to be discounted from this comparison is because in Section 3.9 we have found that the 2-WTP case is representable for the 5-WTP.

For every simulation scenario, results of both models are shown on the same plot for the ease of comparison. Our aim is to demonstrate that the SSA model:

1. is capable of informing behaviours of Smart Pricing signalling requirements like the MCS model;
2. is also able to predict the long-term behaviours; and
3. performs better than the MCS model.

4.11.1 Differences between the MCS and SSA models

Modelling approaches employed by the MCS and SSA models are different. We hereby discuss some key differences.

- For the MCS model, the arrival rate, mean hold time and mean WTP specified in Table 3.12 are used to generate arrival times, hold times and WTP levels. The SSA model, however, uses those parameter values directly;
- With the MCS model, t_{gQ} is a fixed value, whereas it is a probability value plus assumed to be negative exponentially distributed in the SSA model; and
- The MCS model measures the signalling requirements when the system is in instantaneous conditions, while the SSA model measures in the steady-state condition. This is due to the nature of system modelling using state space, whereby the modelling yields stationary probabilities. As the measurements are taken in different system conditions, we expect the results between the two models to be different.

From preliminary simulation results, the SSA model is seen to give better performance. As a result, we decided to conduct a further two simulation scenarios to the six identified in Section 3.8.3. Results from the first allows us to see how imposing more signalling components on a link impacts the link overall signalling requirements.

Results from the second enables us to determine the optimal position for the new network element that we propose to add to the conventional mobile network architecture, i.e. the DPE.

4.11.2 Small system

NR and the required signalling

Figure 4.2 shows average signalling loads versus *NR*. The signalling loads are seen to increase as *NR* increases up to 13 [dB], then remain steady. This is the behaviour that we saw in the MCS model. In the steady-state condition, the signalling loads are seen to have maxima of 3.82, 4.00 and 66.24 kbps for the UL, DL and Netw2Netw, respectively. When compared with results of the MCS model, the deviation is at most 3.34%.

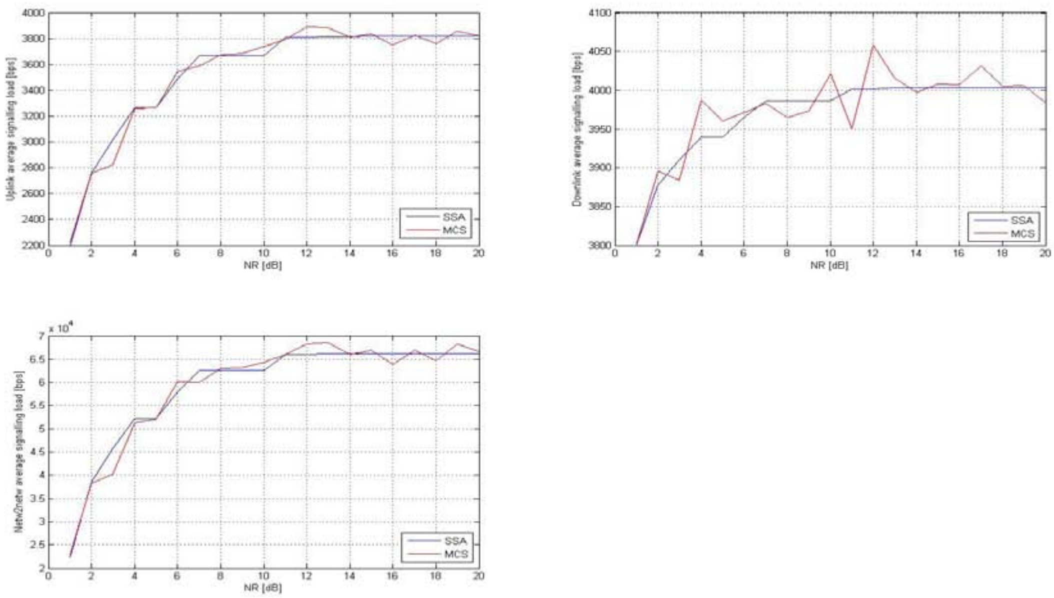


Figure 4.2: Average signalling load vs NR - SSA.

Figure 4.3 shows the bidding signalling percentages vs *NR*. It is seen that the percentages also increase as *NR* increases then becomes steady. On closer examination, we see that it is steady when *NR* equals 15, 13 and 15 [dB] for the UL, DL and Netw2Netw, respectively. Note that from $NR = 13$ [dB], there are no changes to the numbers of high WTP users with η_T and η_M, as well as the maximum possible number of users.

When compared with the MCS model, the bidding percentages behave in similar manners. In the steady-state condition, the percentages are seen to have maximums of 5.11, 2.42 and 2.15% for the UL, DL and Netw2Netw, respectively. The deviation between the two models is at most 9.27%.

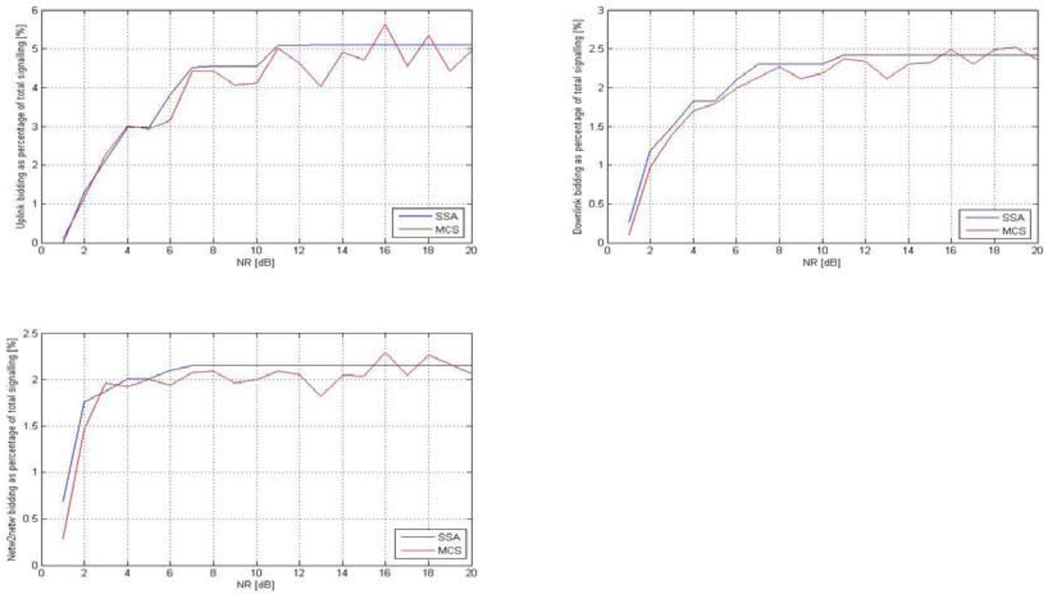


Figure 4.3: Bidding signalling as percentage of total signalling vs NR - SSA.

The cost un-recoverable signalling percentages can be seen from Figure 4.4 to decrease as *NR* increases. The percentages are steady above a *NR* value of 13 [dB]. When compared with the MCS model, the bidding percentages behave in similar manners. In the steady-state condition, the percentages are seen to have minimums of 93.90, 94.83, 95.61% for the UL, DL and Netw2Netw, respectively. The deviation between the two models is at most 0.59%.

It can be seen from Figure 4.5 that as *NR* increases the blocking probability due to insufficient capacity, blocking probability due to insufficient WTP and probability of successful arrivals decreases, increases and increases, respectively. The probabilities are steady above a *NR* value of 13 [dB]. When compared with the MCS model, the probabilities behave in similar manners. In the steady-state condition, the blocking probability due to insufficient capacity has a minimum value of 10.83%, the blocking probability due to insufficient WTP has a maximum value of 84.20% and the probability of successful arrivals has a maximum value of 0.93%. The deviation between the two models is at most 8.82%.

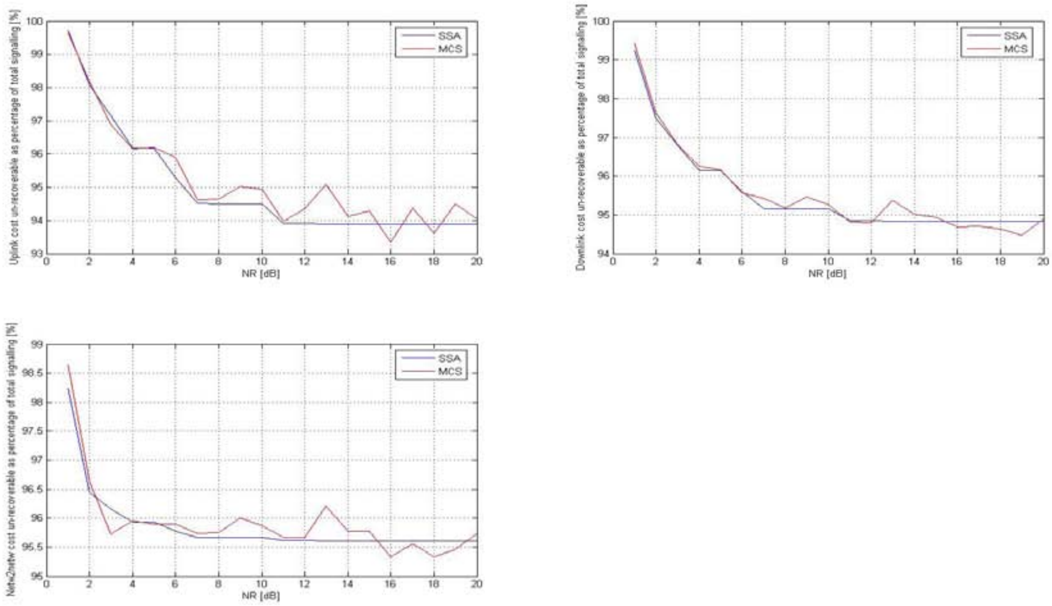


Figure 4.4: Cost un-recoverable signalling as percentage of total signalling vs NR - SSA.

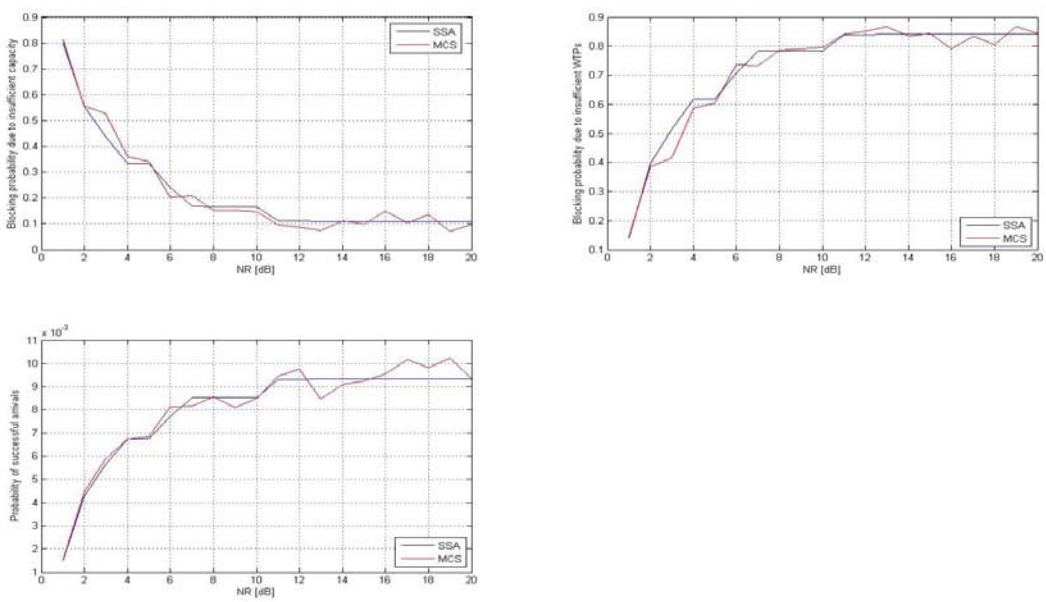


Figure 4.5: Probabilities vs. NR - SSA.

WTP **and the required signalling**

Figure 4.6 shows that as the scale factor increases (equivalent to increasing the mean WTP) the average signalling load decreases, increases and decreases for the UL, DL and Netw2Netw. When compared with the MCS model, the signalling loads behave in similar manners and have a maximum deviation of 3.1%. It is noted that the percentage belongs to the Netw2Netw, not the DL. Big fluctuations seen on the DL is only because of the smaller scale of the Y-axis on that sub-plot.

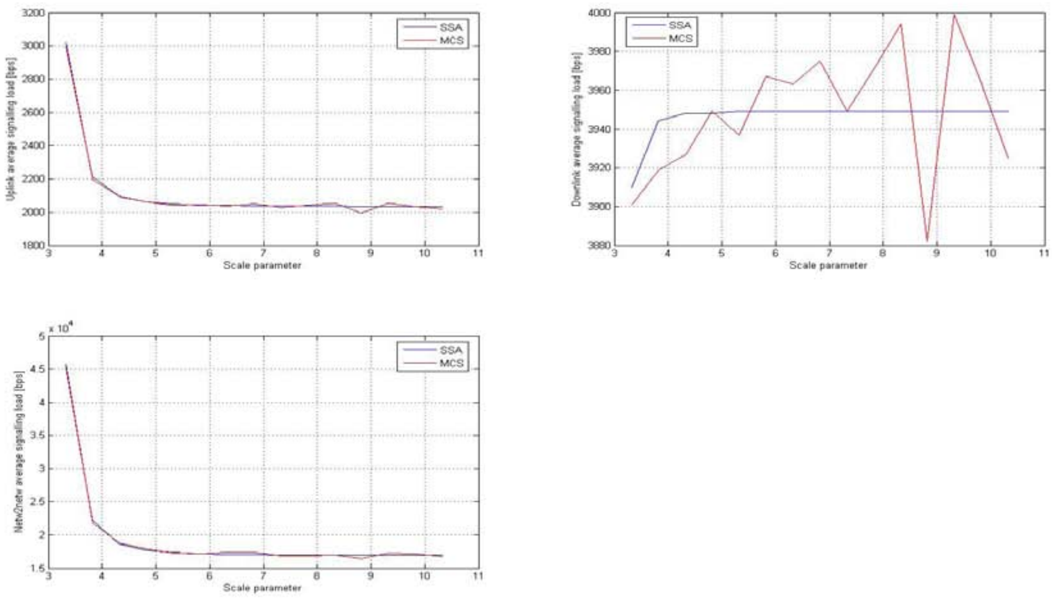


Figure 4.6: Average signalling load vs WTP - SSA.

Arrival rate and the required signalling

Figure 4.7 shows the average signalling loads increase as arrival rate increases. When compared with the MCS model, the signalling loads behave in similar manners and have a maximum deviation of 12%.

Low load threshold and the required signalling

Figure 4.8 shows the average signalling loads decrease as the low load threshold factor increases. When compared with the MCS model, the signalling loads behave in similar manners and have a maximum deviation of 11.81%.

The closer to 1 the low load threshold factor is, the less capacity is reserved for high WTP users, therefore low load threshold factor should not be set too close to 1. With this in mind, looking at Figure 4.8, it can be seen that the factor should be set at 0.7, i.e. 70% of η_M. The advantage of this setting is that the signalling loads are low compared to most other values. In fact, it is second lowest for the UL and Netw2Netw.

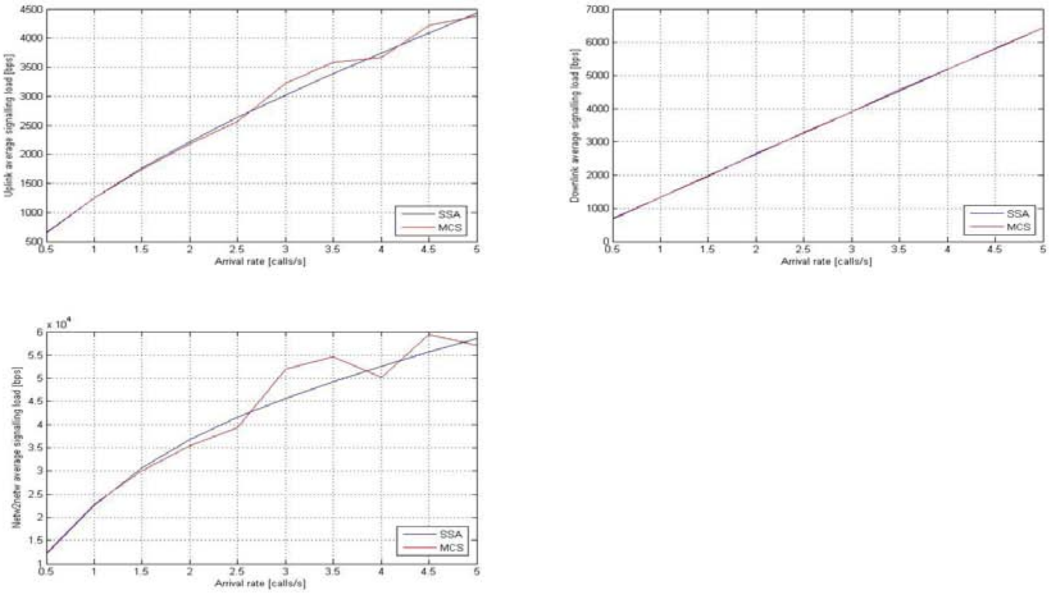


Figure 4.7: Average signalling load vs Arrival Rate - SSA.

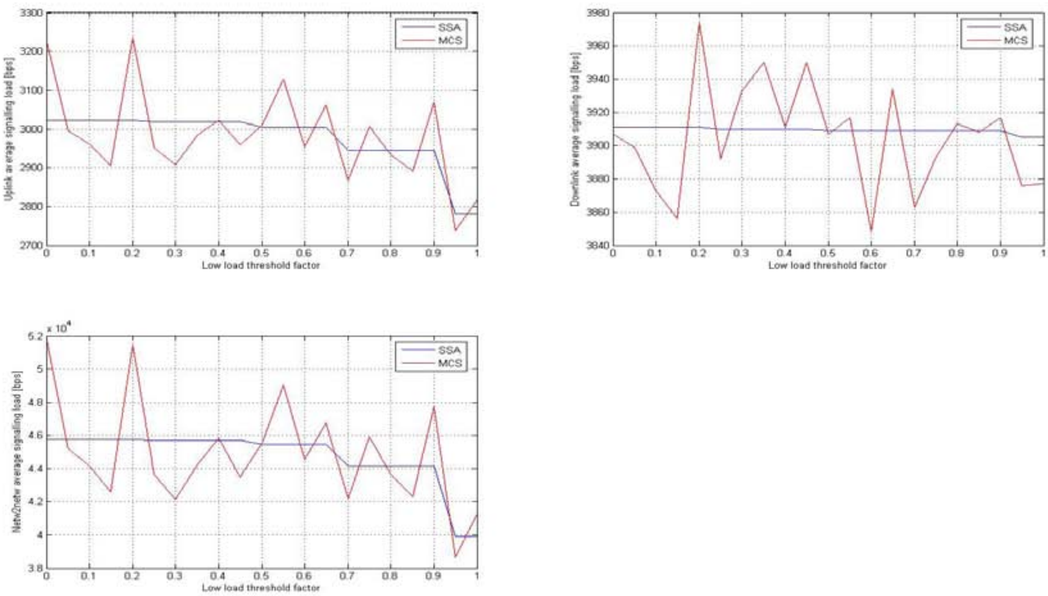


Figure 4.8: Average signalling load vs η_T factor - SSA.

For the DL, however, there is virtually no difference between the signalling, regardless of the value of the low load threshold factor.

t_{gQ} and the required signalling

Figure 4.9 shows that the average signalling loads decrease as t_{gQ} increases. $t_{gQ} = 180$ seconds is when the guaranteed period of the admission quality equals the mean hold time. That is, t_{gQ} lasts for the entire duration of a call.

When compared with the MCS model, the signalling loads behave in similar manners and have a maximum deviation of 13.73%. As mentioned in Section 4.11.1, t_{gQ} is one of the key differences between the MCS and the SSA model. The fact that now we see such a small percentage of difference is assurance that our assumption about the negatively exponentially distributed probability nature of the t_{gQ} is justified.

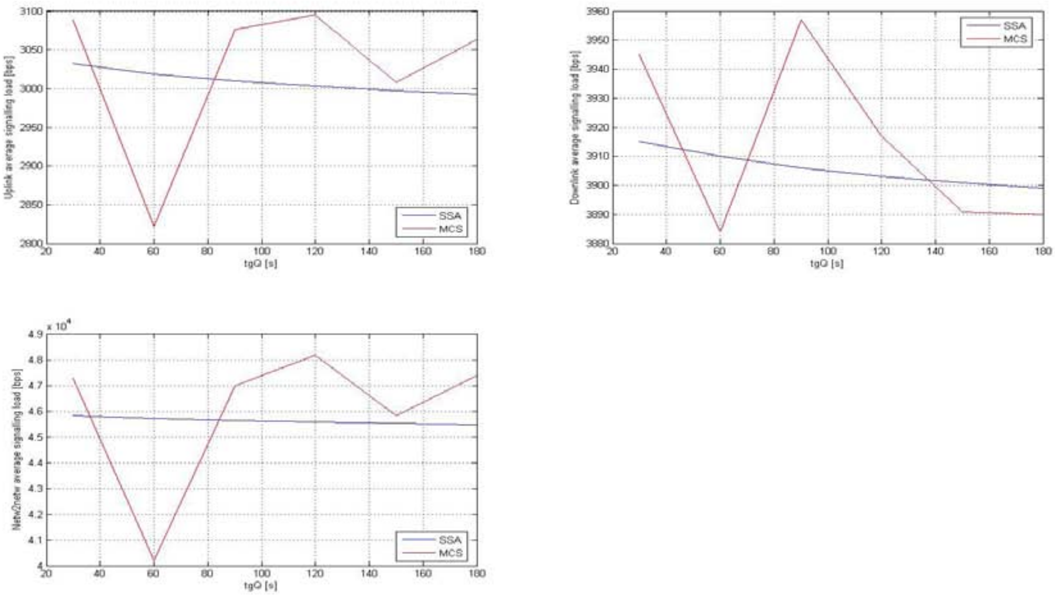


Figure 4.9: Average signalling load vs tgQ - SSA.

Although the results in Figure 4.9 suggests that we should set t_{gQ} higher, we should not. As mentioned in Section 3.8.5, setting t_{gQ} too high reduces the flexibility a network operator has to respond to congestion, which could result in a loss of revenue. On the other hand, setting it too low, as we see in Figure 4.9, will cause the required signalling to increase. The benefit of this, however, is the increase in customers' satisfaction. Taking these factors into account, setting t_{gQ} in the middle of the range, i.e. at 90 seconds, would serve as a good balance.

Mean hold time and the required signalling

Figure 4.10 shows the average signalling loads decrease as the mean hold time increases. When compared with the MCS model, the signalling loads behave in similar manners and have a maximum deviation of 12.29%.

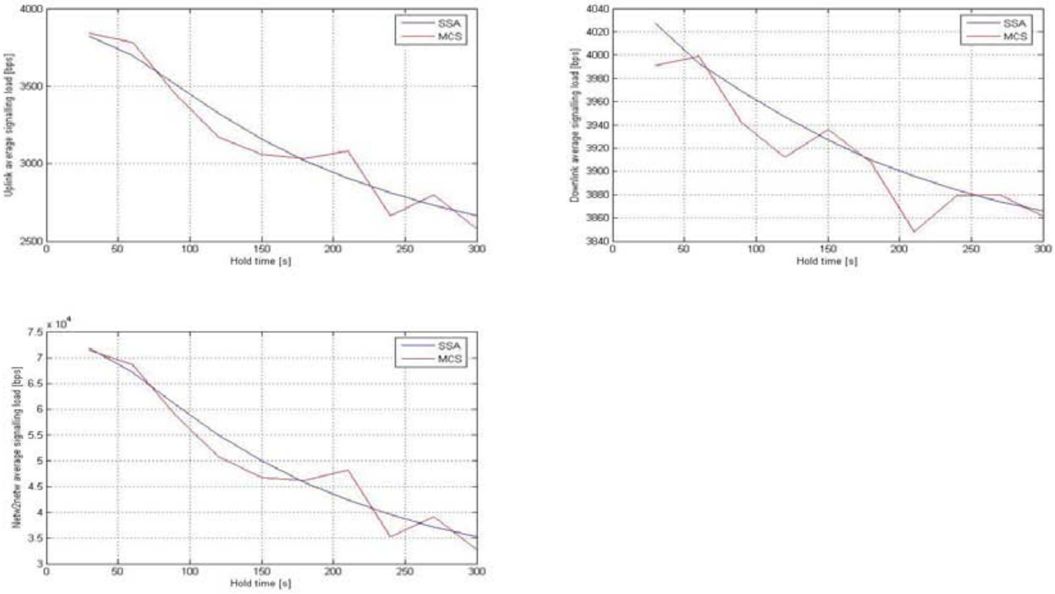


Figure 4.10: Average signalling load vs hold time - SSA.

The effect of increasing signalling components

Measurements of signalling loads for all simulation scenarios so far are based on the signalling components specified in Section 3.7.2. Every network operator who adopts Smart Pricing may add or reduce certain signalling components. Therefore, we see a need to inform the impact of such decisions on the Smart Pricing signalling requirements. As a nominal case, we investigate the impact on the UL signalling load if a decision is made to increase the number of UL bidding signalling messages by a factor, k.

In Tables 3.11 and 3.9, we see that every time a high WTP user bids to maintain his/her admission QoS, an UL signalling message is used. If the bid up factor k is applied, the number of UL signalling messages used will be k. A k factor of 1 means no change, whereas greater than 1 means an increase in the number of UL bidding signalling messages.

Figure 4.11 shows a profile of the increase on the UL signalling load versus the k factor. Also included in the figure is a fitted curve which has a slope of 0.021. The figure shows a linear relationship between the increase on the UL signalling load and the k factor. The slope let us know that the rate of increase on the UL signalling load is 2.1% relative to the k factor. This is not a small number, as it is almost double the percentage of high WTP in the system (1.08% for the CONF base case as seen in Table 3.20). What this tells us is if we increase the UL bidding signalling by a factor of k, the UL signalling load will increase at a rate:

- of $0.021k$, or
- approximately equal to twice the proportion of the number of high WTP users in the system.

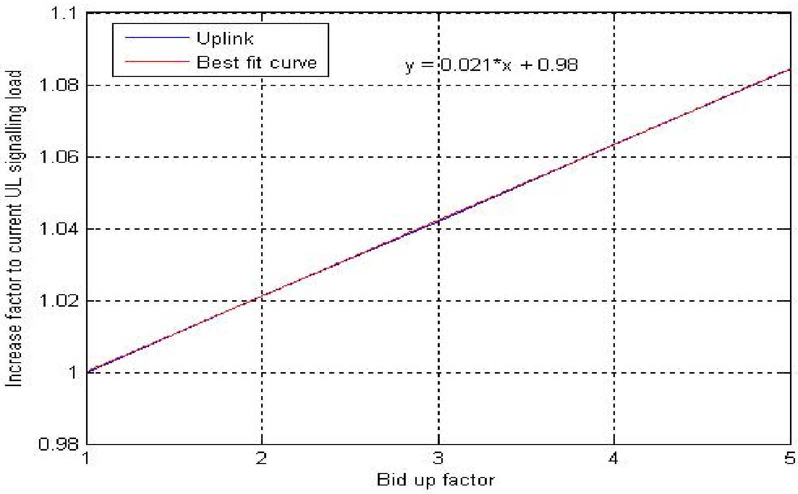


Figure 4.11: Increase of signalling vs bid up factor.

Collocation of the DPE

Although in Section 3.2 we propose that the DPE is to be collocated with the RNC, we have not been able to demonstrate if it is the best place for the DPE. Therefore, so far in all the signalling loads measured, we have considered the DPE as a stand-alone network element.

In order to find out if the RNC is the best place for the DPE to collocate, we conduct simulations in which we measure the Netw2Netw signalling loads when the DPE is collocated with the RNC, as well as with the MSC and the Billing System. Table 4.12 provides a summary of the results. The results demonstrate that our intuition to collocate the DPE with the RNC was a good choice as it gives a reduction in the required signalling only marginally less than other reductions. However, the results also show that collocating with the MSC is a better choice; the Billing System is the best, as signalling reduction is the highest.

DPE	Netw2Netw average signalling [bps]	Reduction [%]
No collocation	45,720	0%
Collocates with RNC	31,231	32%
Collocates with MSC	30,283	34%
Collocates with Billing System	29,926	35%

Table 4.12: Netw2Netw signalling for different places of collocation of DPE.

Simulation times

A summary of the average simulation times for all six simulation scenarios for the small system 2-WTP case is provided in Table 4.13. The smallest, highest and overall average times are 0.13, 0.21 and 0.16 seconds, respectively. When compared with the MCS simulation times, the SSA model is over 4,000 times faster. This is a substantial improvement.

Simulation scenario	Average simulation time [s]	Improvement when compared with C3 model [%]
NR	0.21	419,524
WTP	0.16	476,875
Arrival rate	0.14	729,286
Low load threshold	0.15	565,333
tgQ	0.14	509,286
Hold time	0.13	751,538
Overall average [s]:	0.16	559,462

Table 4.13: Simulation times - CONF SSA model.

4.11.3 Large system

NR and the required signalling

Figure 4.14 shows average signalling loads versus *NR* for the large system. The signalling loads are seen to increase as *NR* increases. This is the behaviour that we saw in the MCS model. In the steady-state condition, the signalling loads are seen to have maximums of 4.04, 4.06 and 72.71 kbps for the UL, DL and Netw2Netw, respectively. When compared with results of the MCS model, the deviation is at most 16.79%.

Figure 4.15 shows the bidding signalling percentages increase as *NR* increases. For Netw2Netw, the signalling appears to be steady from 0.3 [dB]. When compared with the MCS model, the bidding percentages behave in a similar manner. In the steady-state condition, the percentages are seen to have maximums of 5.62, 2.36 and 1.95%

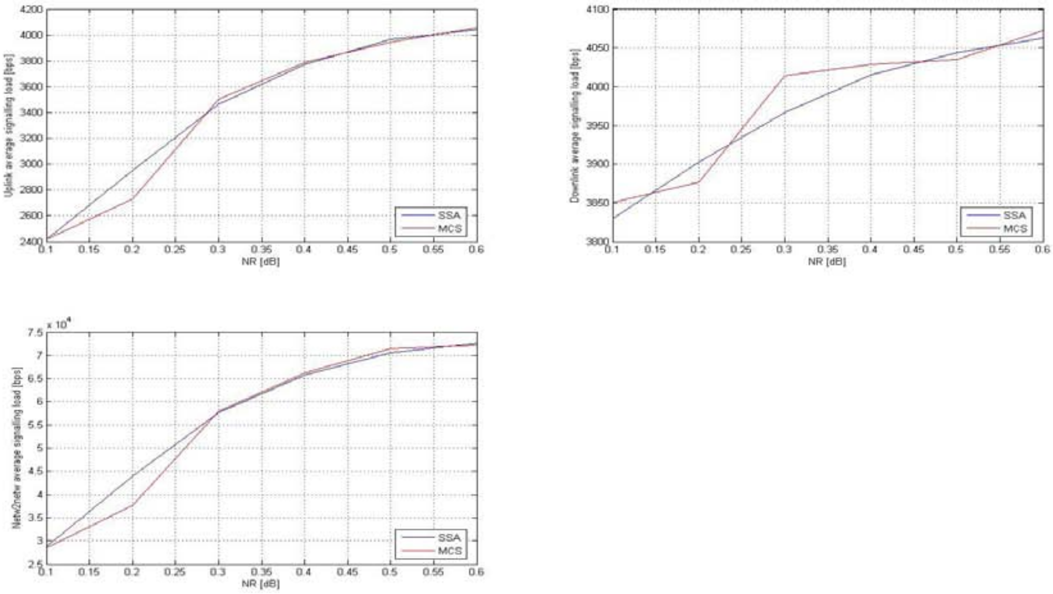


Figure 4.14: Average signalling load vs NR - SSA large system.

for the UL, DL and Netw2Netw, respectively. The maximum deviation from the MCS models is 29%.

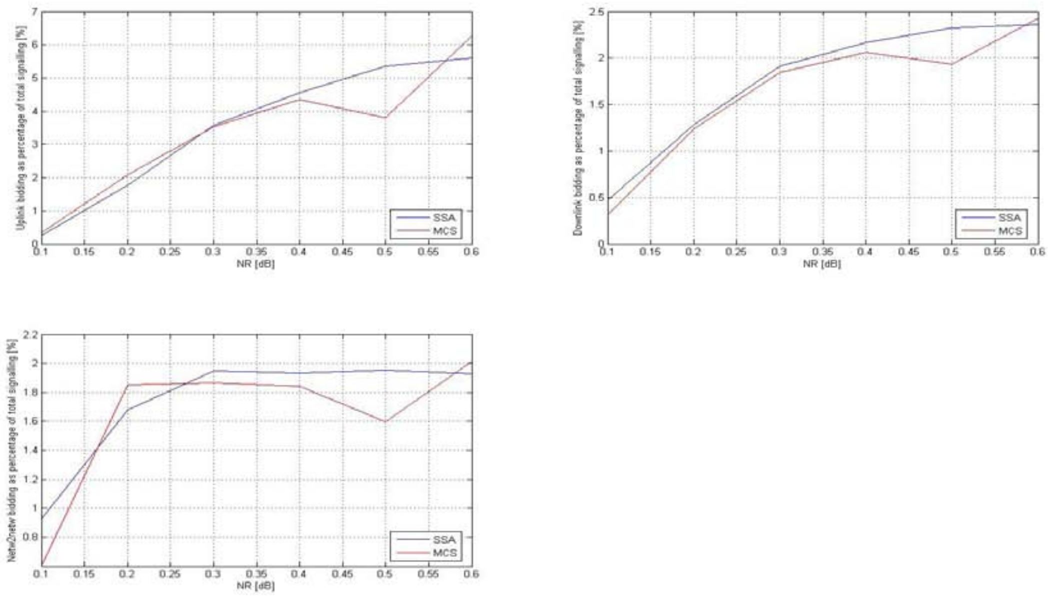


Figure 4.15: Bidding signalling as percentage of total signalling vs NR - SSA large system.

The cost un-recoverable signalling percentages can be seen from Figure 4.16 to decrease as *NR* increases. When compared with the MCS model, the bidding percentages behave in similar manners. In the steady-state condition, the percentages are seen to have minimums of 92, 93, 95% for the UL, DL and Netw2Netw, respectively. The maximum deviation from the MCS models is 1.51%.

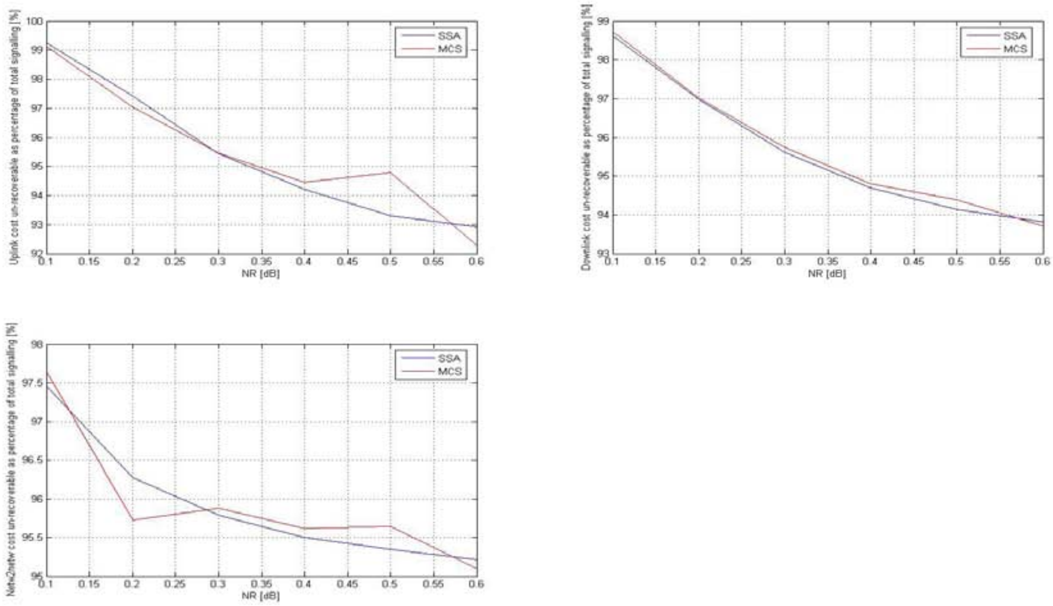


Figure 4.16: Cost un-recoverable signalling as percentage of total signalling vs NR - SSA large system.

It can be seen from Figure 4.17 that as *NR* increases the blocking probability due to insufficient capacity, blocking probability due to insufficient WTP and probability of successful arrivals decreases, increases and increases, respectively. When compared with the MCS model, the probabilities behave in similar manners. In the steady-state condition, the blocking probability due to insufficient capacity has a minimum value of 1.2%, the blocking probability due to insufficient WTP has a maximum value of 93% and the probability of successful arrivals has a maximum value of 1.47%. The maximum deviation from the MCS models is 8.82%.

The average simulation time for this large system (2-WTP) simulation scenario of the SSA model is 1.59 seconds. This is over 700 times faster than the MCS model.

The effect of increasing signalling components

Figure 4.18 shows a profile of the increase on the UL signalling load versus the bid up factor k. The figure shows a linear relationship between the increase on the UL

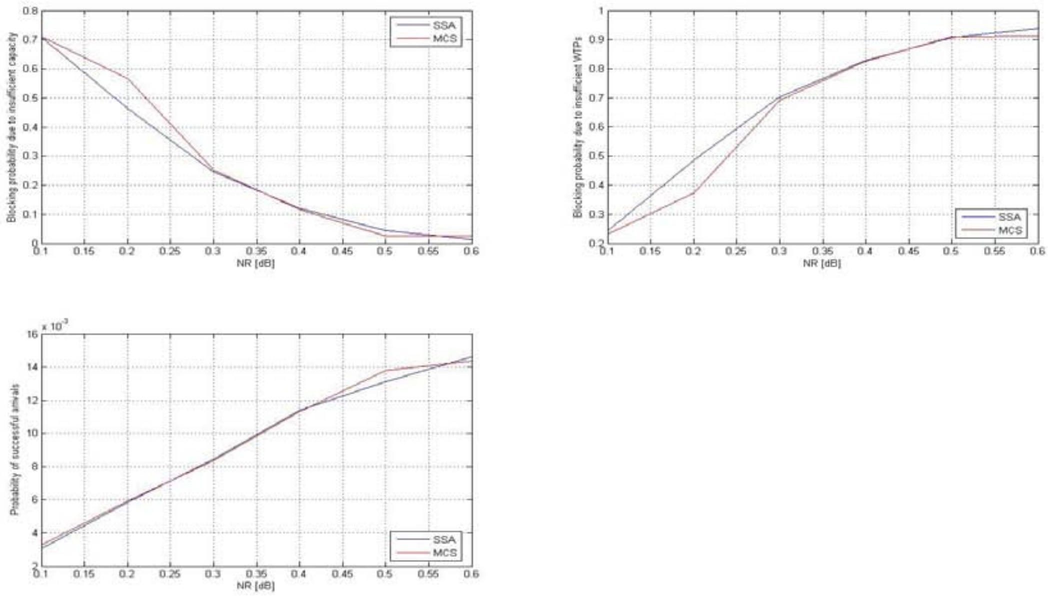


Figure 4.17: Probabilities vs. NR - SSA large system.

signalling load and the k factor. The slope of the best fit curve lets us know that the rate of increase on the UL signalling load is 5.6% relative to the k factor. This means that if we increase the UL bidding signalling, the UL signalling load will increase at a rate of $0.056k$.

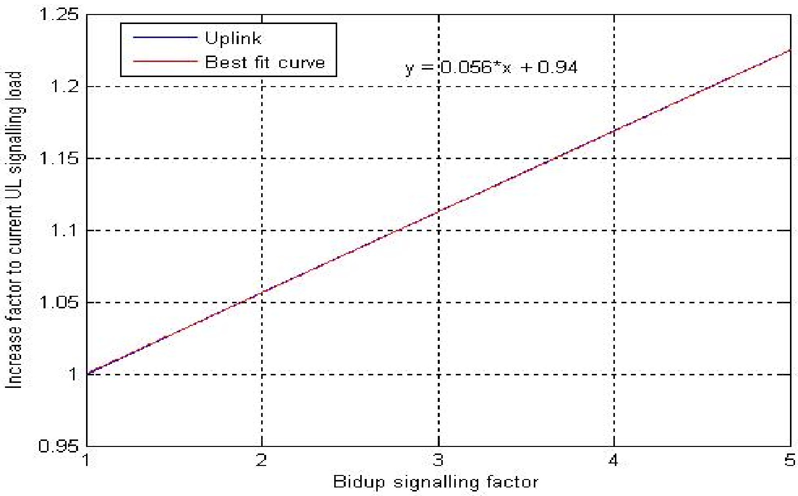


Figure 4.18: Increase of signalling vs bid up factor - large system.

Collocation of the DPE

Table 4.19 provides a summary of the signalling loads for different collocation options for the DPE. The results re-affirm that collocating with the MSC is better than with the RNC, but the Billing System is the best choice. Accordingly, network operators who are considering adopting Smart Pricing should opt to collocate the DPE with the Billing System.

DPE	Netw2Netw average signalling [bps]	Reduction [%]
No collocation	72,705	0%
Collocates with RNC	57,424	21%
Collocates with MSC	44,322	39%
Collocates with Billing System	43,664	40%

Table 4.19: Netw2Netw signalling for different places of collocation of DPE - large system.

4.12 Findings and discussions

In this chapter, we have developed our second Smart Pricing signalling model, using State Space Analysis (SSA). The SSA model allows us to determine the signalling requirements for Smart Pricing when the system is in the steady-state condition.

Keeping in mind the aim set out in Section 4.11, we replicate twelve simulation scenarios of the small system 2-WTP and the large system cases. In addition, for each system, we also conduct an additional two simulation scenarios to investigate two other important aspects of the system. This brings the total of simulation scenarios conducted to sixteen. Again, this is part of our effort in investigating the impact of signalling models as thoroughly as possible.

Table 4.20 provide a summary of the results of the *NR* simulation scenarios.

We have found that:

1. with the SSA model:

   a) as *NR* increases, the:

   - **i.** average signalling loads increase;
   - **ii.** bidding signalling percentages[3] increase. For the large system case, the Netw2Netw bidding signalling appears to be steady when η_M is between 6.67 and 12.90% of the cell pole capacity;

[3] That is bidding signalling loads as percentages of respective total signalling loads.

Parameter	Case						Unit
	CONF 2 WTPs			Voice (2WTPs)			
	UL	DL	Netw2Netw	UL	DL	Netw2Netw	
Signalling load max	3.82	4.00	66.24	4.04	4.06	72.71	kbps
Signalling-load-max ratio to UL average	1.00	1.05	17.35	1.00	1.005	17.98	
Bidding signalling max	5.11	2.42	2.15	5.62	2.36	1.95	%
Cost un-recoverable min	93.90	94.83	95.61	92.93	93.82	95.21	%
Blocking probability min - insufficient capacity	10.83			1.20			%
Blocking probability max - insufficient WTPs	84.20			93.82			%
Probability of successful arrivals max	0.93			1.47			%
Simulation time	0.21			1.59			seconds

Table 4.20: Summary of results - SSA.

iii. cost un-recoverable signalling percentages[4] decrease;

iv. blocking probability due to insufficient capacity decreases;

v. blocking probability due to insufficient WTP increases; and

vi. probability of successful arrival increases;

For the small system, the above parameter values are steady when the system reaches its maximum possible number of users.

b) for the small system case, as the:

i. mean WTP increases, the UL, DL and Netw2Netw signalling loads decreases, increases and decreases respectively. The signalling loads become steady when the mean WTP is close to or above the maximum price;

ii. arrival rate increases, all signalling loads increase;

iii. η_T factor increases, all signalling loads decrease;

iv. t_{gQ} increases, all signalling loads decrease; and

v. mean hold time increases, all signalling loads decrease.

c) all the behaviours mentioned above are the same as in the MCS model. Moreover, for all simulation scenarios, the maximum deviation from the MCS model ranges from 0.59 to 29.70%. Almost 85% of the simulation scenarios have maximum deviations below 20%. This is a strong indication that the SSA model is capable of informing Smart Pricing signalling requirements like the MCS model. In addition, the SSA model also enables us to predict the signalling requirements in the steady-state condition. These values are summarised in Table 4.20. In the steady-state condition:

[4]That is cost un-recoverable signalling loads as percentages of respective total signalling loads.

i. the maximum average signalling load for the UL is the lowest at 4.04 kbps. While the DL is higher at 4.06 kbps and the Netw2Netw is the highest at 72.71 kbps. In terms of ratio, the maximum DL-to-UL average signalling load ratio is 1.05, whereas the Netw2Netw-to-UL is 17.98;

ii. bidding signalling percentages reach their highest maximum value of 5.62%, 2.42% and 2.15% for the UL, DL and Netw2Netw, respectively;

iii. cost un-recoverable signalling percentages reach their highest minimum value of 93.90%, 94.83% and 95.61% for the UL, DL and Netw2Netw, respectively;

iv. blocking probability due to insufficient capacity has minimum values between 1.20-10.83%;

v. blocking probability due to insufficient WTP has maximum values between 84.20-93.82%; and

vi. probability of successful arrival has maximum values between 0.93-1.47%;

2. although t_{gQ} is one of the key differences between the MCS and the SSA model, the maximum deviation of only 13.73% is found. This indicates that our assumption about the negatively exponentially distributed probability nature of t_{gQ} is a valid one.

3. if we increase the UL bidding signalling by a factor of k:

   a) for the small system, the UL signalling load will increase at a rate of $0.021k$. This rate is approximately twice the proportion of high WTP users in the system, and

   b) for the large system, the UL signalling load will increase at a rate of $0.056k$.

4. collocation of the DPE with the Billing System is the best choice, while with the MSC is better than the RNC.

5. performance of the SSA model is far more superior than the MCS model. With respect to simulation times:

   a) for the small system, the SSA model is more than 4,000 times faster than the MCS model; where as

   b) for the large system, the SSA model is more than 700 times faster than the MCS model.

The results indicate that if network operators wish to optimise any system parameter, using the SSA model they will gain substantial benefits.

We recommend that:

Rec.1. the SSA model be used, instead of the MCS model, in determining the signalling requirements for Smart Pricing;

Rec.2. the low load threshold be set at 70% of the cell maximum capacity. This not only reserves adequate capacity for high WTP users during peak periods, but also allows for the second smallest signalling loads to be required;

Rec.3. t_{gQ} be set at 90 seconds for the types of services considered. This provides a good balance in keeping the flexibility for the network operator to respond to congestion and maintaining high customer satisfaction; and

Rec.4. the DPE be collocated with the Billing System. This provides the greatest savings of the required signalling.

CHAPTER 5

Applicability of the proposed models in other cellular telecommunications systems

5.1 Introduction

For Smart Pricing and our proposed Smart Pricing signalling models to be applied to a certain system, three pieces of information about that system needs to be obtained:

1. the maximum capacity of the system;
2. the share of the maximum capacity of an individual user if he/she is admitted to the system; and
3. the system network architecture.

Once that information is attained, the following elements need to be checked and adjusted accordingly if required. They are:

4. system diagram (see Section 3.3); and
5. signalling components (see Section 3.7.2).

5.2 WCDMA downlink

5.2.1 Downlink load and capacity

Although our focus in this thesis in on the uplink only, as demand for downlink capacity to accommodate high data rate multimedia services in 3G networks increases, here we will also discuss load and capacity of WCDMA systems in the downlink. Later in this chapter, we will see how technology in the downlink evolves.

From [59], the total base station transmit power, BS_{TxP}, is given by:

$$BS_{TxP} = \frac{N_{rf} \cdot W \cdot \overline{L} \cdot \sum_{j=1}^{M} \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j}}{1 - \overline{\eta_{DL}}} \tag{5.1}$$

where N_{rf} is the noise spectral density of the mobile receiver front-end, $\overline{L}$ is the average attenuation between the base station transmitter and mobile receiver, and $\overline{\eta_{DL}}$ is the average downlink load factor.

Taking into account the amount of power dedicated for the common channels signalling, Equation (5.1) becomes:

$$BS_{TxP} = \frac{N_{rf} \cdot W \cdot \overline{L} \cdot \sum_{j=1}^{M} \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j}}{(1 - \overline{\eta_{DL}}) \cdot (1 - Pcc)} \tag{5.2}$$

where Pcc is the fraction relative to the BS_{TxP} of the common channel signalling power. As we only model one cell at this stage, the effect of soft handover gain is not considered. If this effect were considered, for instance, for a two-cell scenario, a soft handover gain value of 3 dB [59] should be included.

The N_{rf} is given by:

$$N_{rf} = k \cdot T + NF \tag{5.3}$$

where k is the Boltzmann constant, T is *Kelvin* temperature at the mobile receiver, and NF is the noise figure of the mobile receiver.

The $\overline{\eta_{DL}}$ is given by:

$$\overline{\eta_{DL}} = \sum_{j=1}^{M} \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j} \cdot [(1 - \overline{\alpha}) + \overline{i}] \tag{5.4}$$

where $\overline{\alpha}$ is the average orthogonality factor in the cell, and $\overline{i}$ is the average ratio of other cell to own cell base station power received by user.

Rearranging Equation (5.4), we have:

$$\sum_{j=1}^{M} \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j} = \frac{\overline{\eta_{DL}}}{[(1 - \overline{\alpha}) + \overline{i}]} \tag{5.5}$$

$$\sum_{j=1}^{M} \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j} = \overline{\eta_{mpDL}} \tag{5.6}$$

where $\overline{\eta_{mpDL}}$ is the maximum possible total average downlink load factor and is defined as:

$$\overline{\eta_{mpDL}} = \frac{\overline{\eta_{DL}}}{[(1-\overline{\alpha})+\overline{i}]} \tag{5.7}$$

We define the downlink load factor of a user j, $L_{DL,j}$, as:

$$L_{DL,j} = \upsilon_j \cdot \frac{(E_b/N_0)_j}{W/R_j} \tag{5.8}$$

Substitute Equation (5.5) into Equation (5.2), we have:

$$BS_{TxP} = \frac{N_{rf} \cdot W \cdot \overline{L} \cdot \dfrac{\overline{\eta_{DL}}}{[(1-\overline{\alpha})+\overline{i}]}}{(1-\overline{\eta_{DL}}) \cdot (1-Pcc)} \tag{5.9}$$

Rearranging Equation (5.9), we have:

$$\overline{\eta_{DL}} = \frac{1}{1 + \dfrac{N_{rf} \cdot W \cdot \overline{L}}{BS_{TxP} \cdot (1-Pcc) \cdot [(1-\overline{\alpha})+\overline{i}]}} \tag{5.10}$$

Techniques to calculate $\overline{L}$ are to be discussed in the next section. Apart from $\overline{L}$, other parameter values in Equation (5.10) are usually given, thus $\overline{\eta_{DL}}$ can be determined. Then, substituting $\overline{\eta_{DL}}$ into Equation (5.7), $\overline{\eta_{mpDL}}$ is found.

Note that $\overline{\eta_{mpDL}}$ should not be allowed to exceed the target NR set for the DL[1] by the network operator. By ensuring this condition, as well as the one discussed in the next section in which the maximum allowed path loss for the DL must be equal to the UL maximum allowed path loss, a target coverage area is achieved.

If we assume that all users in the cell having the same values for E_b/N_0, υ and R, using Equation (5.8), the downlink load factor of a user, L_{DLu}, is given by:

$$L_{DLu} = \upsilon \cdot \frac{(E_b/N_0)}{W/R} \tag{5.11}$$

[1]Refer to Equation (2.18) for the relationship between NR and load factor.

The number of users with L_{DLu} that can be support in the DL, M_{DL}, in turn is given by:

$$M_{DL} = \frac{\overline{\eta_{mpDL}}}{L_{DLu}} \tag{5.12}$$

5.2.2 Techniques to calculate $\overline{L}$

To calculate $\overline{L}$, a number of factors need to be taken into account. These include: cable loss, indoor loss (to account for building penetration), building floors attenuation (to account for penetration across multiple levels in a high building), vegetation loss (to account for loss through trees)[2], fast and slow fading margins and path loss.

To calculate path loss, L_{DL}, in general, if the path between the base station and the MS is a line-of-sight one, the free-space propagation model is used, the formula of which can be found in [63]. On the other hand, the diffraction propagation model in [64] can be used to calculate path loss due to diffraction. A comprehensive propagation model in [65] can be used to calculate the combined free-space and diffraction loss, as well other losses such as tropospheric scatter, hydrometeor scatter and surface ducting. Particularly for cellular mobile telecommunications systems, instead of those propagation models, some other models as discussed below are commonly used for urban areas or areas nearby cities. Each of these models is valid for a certain frequency band, ranges of base station and MS antennas' heights, city types and cell types.

The Okumura model is one of the most widely used for frequency band 150-1920 MHz and base station antenna heights 30-1000m [66]. This model is suitable for use with GSM 850 MHz (UL 824-849 MHz and DL 869-894), GSM 900 MHz (UL 880-915 MHz and DL 925-960 MHz) and GSM 1800 MHz (UL 1710-1785 MHz and DL 1805-1880 MHz[3]) and UMTS Time Division Duplex (TDD) systems (1900-1920 MHz). Based on the Okumura model, the Okumura-Hata model is an empirical model. This model was, however, standardised for international use [68]. The model is valid for frequency band 150-1500 MHz and is not suitable for use with the UMTS TDD system. Different from the Okumura and Okumura-Hata models, the COST 231/Walfisch-Ikegami model is a physical model. It is valid for frequency range 800-2000 MHz [68]. All three models are suitable for macro-cell environments. For micro-cell environments, the *two-ray* or *street canyon* models in [68] are recommended.

A propagation model developed by the *International Telecommunication Union* (ITU) and valid for most frequencies used for mobile telecommunications is available

[2] See [60], [61] and [62] for techniques to calculate this loss.

[3] In Australia, a telecommunications operator uses these frequencies for Enhanced Data rates for GSM Evolution (EDGE) as an upgrade to its GPRS system and as an drop back option for its 3G users [67].

in [69]. This model can be used for the Very High Frequency (VHF) and Ultra High Frequency (UHF) bands and for all transmitting antenna heights less than 3000 metres. This model may be used for all systems mentioned above, including the UMTS Frequency Division Duplex (FDD) system (UL 1920-1980 MHz and DL 2110-2170 MHz) and the International Mobile Telecommunications-2000 (IMT-2000)[4] system (2500-2690 MHz) [70],[71]. The readers are referred to [72] for a complete list of UMTS frequency bands identified by the 3rd Generation Partnership Project (3GPP).

For some of the above propagation models, the calculation of path loss is complex and the use of radiocommunications software is recommended. Two examples of software applications that can accommodate these needs are *Visualyse* and *WRAP*. Details about these software applications can be found in [73] and [74] respectively.

Once the path loss is determined, it has to be added to other relevant losses and fading margins mentioned earlier in this section. Next, that resulting loss is subtracted by the base station and MS antenna gains to give the maximum allowed path loss for the DL, L_{maDL}. That is:

$$L_{maDL} = L_{DL} + L_{others} + M_{fadings} - Gain_{BS} - Gain_{MS} \tag{5.13}$$

where L_{others} is the sum of all other losses other than path loss, $M_{fadings}$ is the sum of all fading margins, $Gain_{BS}$ is the base station antenna gain, and $Gain_{MS}$ is the MS antenna gain. All parameter values in Equation (5.13) are in dB.

To meet the target cell coverage area, L_{maDL} must be ensured to be equal to the UL maximum allowed path loss, i.e. L_{maUL}. A technique to calculate L_{maUL} can be found in [59]. Without that assurance, the quality of a user's call cannot be guaranteed. In general, with low load, a network is UL coverage limited, meaning that the L_{maDL} found above is greater than the L_{maUL}. In such a case, to ensure the condition is met, we can reduce the cell size (i.e. the distance between the base station and MS that we use to calculate L_{maDL} using those propagation models discussed above) until L_{maDL} is reduced to equal L_{maUL}[5]. For simplicity, we assume that L_{maDL} obtained from Equation (5.13) is equal to L_{maUL} without the need for any adjustment. Next, L_{maDL} is subtracted by the maximum versus average path loss ratio (in dB) to obtain $\overline{L}$. $\overline{L}$ must be converted to linear value before substituting into Equation (5.10).

Conversely, if a smaller cell size is planned for capacity reasons, the path loss and hence maximum transmission power can be calculated using the appropriate path loss model.

[4] 3G systems are termed IMT systems by the ITU.

[5] Referring to Equation (5.10), reducing L_{maDL} effectively means higher load factor is allowed.

5.2.3 Applicability of Smart Pricing signalling models on the downlink

We notice that the two main differences between the UL and DL are the way the maximum allowed load factor (i.e. the maximum capacity) of the system and the share of the maximum capacity of an individual user (i.e. the user load factor) are calculated. In particular, for:

1. the maximum allowed load factor:
   - on the UL, Equations (2.21) and (2.21) are used; and
   - on the DL, Equations (5.10) and (5.7) are used.
2. the user load factor:
   - on the UL, Equation (2.8) is used; and
   - on the DL, Equation (5.8) is used.

Apart from those two aspects, the system network architecture, the system diagram, the Smart Pricing signalling algorithm and the signalling components are the same. Thus, clearly, Smart Pricing and our proposed Smart Pricing signalling models as detailed in Chapters 3 and 4 can be applied to the WCDMA downlink.

5.2.4 Estimated required signalling for Smart Pricing on the downlink

Simulation results for a case study with a maximum base station transmit power of 43 dBm on the downlink and a load factor limit of 0.75 on the uplink show the downlink mean throughput is 95-98% of the downlink [59]. This means the η_T, η_M and number of users that can be accommodated by the downlink capacity are almost identical to those on the uplink. Also, as mentioned in the previous section, apart from the way load factors are calculated, all other aspects of the downlink and uplink systems are the same. Therefore, we estimate that the signalling requirements for the downlink are equal to those for the uplink.

5.3 High speed downlink packet access

The WCDMA system that we discussed earlier in previous chapters is the 3GPP Release 99 WCDMA system. In this section, we discuss the WCDMA system in Release 5.

On the downlink, included in Release 5 (and enhanced in Release 6), High Speed Downlink Packet Access (HSDPA) is an improvement of the Release 99 WCDMA downlink system while on the uplink, included in Release 6, High Speed Uplink Packet Access (HSUPA) is an improvement. HSDPA and HSUPA as a couple are termed High Speed Packet Access (HSPA). HSPA is critical to the evolution of 3G as a fast multimedia network, by providing data rates comparable with wireline broadband, and dramatically reduced latency times [75].

HSPA can be deployed on top of the Release 5 WCDMA network by using the same or another frequency carrier depending on the capacity required for HSDPA. Either way HSPA and Release 99 WCMA system can share all the network elements, including: Node Bs, RNCs, SGSNs and GGSN, as well as Node B sites, antennas and antenna lines. Upgrading to HSPA, only new software and possibly some new pieces of hardware are required for the Node Bs and RNCs [76].

Chat and peer-to-peer are typically symmetric services, whereas web browsing and streaming are asymmetric services [77]. [76] states that HSPA capacity can support both symmetric and asymmetric services, though a graph in it shows that HSPA enhanced downlink capacity is around 2.4 times that of the uplink capacity. This is consistent with the report in [78] where in 2003, it was concluded that downlink data traffic will exceed that of the uplink by a factor of 2.3 [79]. A point to note is that HSUPA was only included in a Release that is two years after the Release that HSDPA is included. This implies that greater emphasis was placed on capacity of the downlink. In this thesis, although we can consider the applicability of our Smart Pricing signalling models to both HSDPA and HSUPA features of the HSPA, taking into account the greater attention placed on HSDPA, we will only investigate the applicability to the HSDPA.

As Smart Pricing is a solution to the problem of under-utilised network resources or to accommodate growing demand within existing network resources, it is important that we determine accurately the capacity of a system under consideration. In addition, as the focus of this research is on determining the signalling requirements, which are dictated by communications between network elements, it is also important to examine the system architecture. Before these two tasks can be accomplished, an additional task must first be undertaken, namely to understand the system characteristics. In the next three Sections, we will tackle these three tasks[6].

[6] Note that the three tasks will also be carried out for HSPA+, LTE and Cognitive Radio.

5.3.1 Characteristics of HSDPA

HSDPA uses adaptive modulation and Hybrid Automatic Repeat Request[7] (HARQ) to achieve high throughput, high peak rates and to reduce delay [81].

Adaptive modulation, together with adaptive coding, facilitate data rate adjustment according to channel quality [76]. Modulation schemes available for HSDPA in Release 5 are Quadrature Phase Shift Keying (QPSK) and 16-Quadrature Amplitude Modulation (16-QAM). 16-QAM was not utilised in Release 99, which is a contributing factor to its lower data rate. The coding scheme used for HSDPA is turbo coding. The range of Effective Coding Rate (ECR) in Release 5 is from 0.14 to 0.77 [76]. When channel quality is good, high data rates can be achieved by using higher coding rates and/or 16-QAM. If it is bad, low coding rates and QPSK are used to guarantee the target Block Error Probability (BLER). A typical BLER value for HSDPA is 10^{-1}. Channel quality is conveyed by using the Channel Quality Indicator (CQI).

HSDPA uses two retransmission methods for its HARQ scheme, these are: *Chase* code combining[8] (also known as soft combining) and incremental redundancy. The *Chase* code combining method transmits an identical version of an erroneously detected data packet and allows the decoder to combine the received copies weighted by the Signal-to-Noise Ratio (SNR) prior to decoding [83]. This method provides diversity gain and is simple to implement [80]. On the other hand, the incremental redundancy method transmits additional redundant information incrementally if the decoding fails on the first attempt [83]. The incremental redundancy method performs better than the *Chase* code combining, but needs more memory at the MS than the *Chase* code combining [59].

With HSDPA, the first number of retransmissions are handled at physical layer at the Node B. This was not the case with the Release 99 WCDMA. If those retransmissions fail, the conventional higher layer retransmission at the Radio Link Layer (RLC) at the RNC will occur. An example of such a case is due to a cell change in the event of mobility. This means that there are two tiers of retransmission with HSDPA. Employing retransmission at the Node B is a contributing factor that allows for delay to be reduced as it eliminates the 10 ms delay at the Iub interface and another 10 ms delay at the RNC [76].

[7] A term used to describe any combined Forward Error Correction (FEC) and Automatic Repeat Request (ARQ) scheme in which unsuccessful packet transmission attempts are used in FEC decoding instead of being discarded [80].

[8] Named after the scientist who proposed the technique. Code combining represents a technique for combining the minimum number of repeated packets encoded with a higher code rate to obtain a lower code rate. This allows for reliable communications when channel error rates are less than $5\text{x}10^{-1}$ [82].

Another contributing factor that helps reduce delay is the introduction of the Medium Access Control-high speed (MAC-hs) protocol. This sees scheduling and priority handling functions moved from the RNC to the Node B [76] [83]. This move together with a cut of the Transmission Time Interval (TTI) down to a sub-frame (three slots), i.e. a time period of 2 ms [42], are what constitutes the high peak data rates of HSDPA. Not only that the TTI is fixed (at 2 ms), the spreading factor is also fixed at 16. Link adaptation decisions to adjust bit rates based on CQI mentioned above are made by the MAC-hs of the Node B every TTI [76]. When making such decisions, information on the data rate which a MS is able to receive is also taken into account. That information is available in the CQI.

With HSDPA, scheduling of packets can be done based on radio conditions and/or the amounts of information required to send to MSs. If a network operator wish to maximise throughput, maximum Carrier to Interference Ratio (C/I) or throughput scheduling algorithm can be used. If treating every active MS fairly is optimised, the Round Robin scheduling algorithm is recommended. Alternatively, the Proportional Fair algorithm can be used which determines the serve order of MSs based on the highest instantaneous channel quality. Details about the performances of these algorithms, as well as others, can be found in [76] [83] [59].

HSDPA employed three new channels [76] [59]. The first two are on the downlink and the last is on the uplink. Among the three, only one is a user traffic channel. The three channels are:

- the High Speed-Downlink Shared Channel (HS-DSCH). This is the transport channel that carries users' data. The channel is mapped onto the High Speed-Physical Downlink Shared Channel (HS-PDSCH) [84];
- the High Speed-Shared Control Channel (HS-SCCH) This physical channel carries time-critical control information (e.g. modulation); and
- the High Speed-Dedicated Physical Control Channel (HS-DPCCH). This physical channel carries control and feedback information (e.g. CQI).

The spreading factor of 16 mentioned above is the spreading factor of the HS-PDSCH. With that spreading factor, the maximum possible number of channelisation codes (also known as codes) available in the code tree is 16. However, not all of these can be used for the HS-DSCH. At least one of them must be reserved for use by common channels, leaving a maximum of 15 channelisation codes available for use by the HS-DSCH [76]. If HSDPA is deployed on top of Release 99 WCDMA network circuit switched services using the same carrier frequency, associated dedicated channels will consume some of these codes. This means the actual set of channelisation codes reserved for HS-PDSCH transmission could be less than 15. Each channelisation code corresponds to a HS-PDSCH and a MS can be assigned multiple channelisation codes depending on its UE capability [84].

Release 5 defines 12 categories of UE capability. Attributes for each category include: the maximum number of HS-DSCH codes received, the minimum inter-TTI interval, the maximum number of bits of an HS-DSCH transport block received within an HS-DSCH TTI and modulation scheme allowed to use [85]. Using values of attributes together with the subframe structure of the HS-PDSCH in [42], we can calculate the required coding rate and maximum data rate for each categories. See Section 5.3.3 below for details.

It is important to note that circuit switched services such as Adaptive Multi Rate (AMR) audio codec speech calls and video calls are still carried on the Release 99 Dedicated Channel (DCH). These services cannot be mapped on HSDPA [76]. This suggests that capacity for service switched services up to Release 5 remains unchanged comparing to that of it for Release 99.

Voice over IP (VoIP) is not a circuit switched service, thus it can be implemented over HSPA. VoIP is a low-delay real-time service with low bit rate. With scheduling function now at Node B, higher priority can be dedicated to VoIP to accommodate for its low-delay requirement. In addition, with channelisation code-multilplexing, the requirement of serving a large number of users with low bit rates can also be satisfied. Channelisation code-multiplexing is a feature that allows more than one HSDPA user to be scheduled in a single TTI [76].

5.3.2 HSDPA system architecture

With the introduction of HSDPA in Release 5, Node B is now assigned greater responsibilities, such as taking over the scheduling and priority handling functions from the RNC. Nevertheless, if we assume that the Circuit Switched-Media Gateway (CS-MGW) is gathered in the same equipment with the MSC, positions in the architecture of the MS, RNC, MSC, SGSN and HLR remains the same as they were in the Release 99 architecture. This can be seen in [86] and in comparison with Figure 2.3. Moreover, from [87], the MSC and SGSN are seen to still be connected to the Billing System. Hence, the Release 99 architecture is still valid for HSDPA as far as Smart Pricing is concerned because the network elements and links that facilitates Smart Pricing are still the same. As HSDPA is a packet access service, only the SGSN is applicable.

One aspect, however, is potentially different and that is the signalling system. 3GPP does not specifically set whether it is the SS7 or the Internet Protocol based system. It says in [86]: "... The actual [signalling] links may be provided by an underlying network (e.g. SS7 or IP) ...". However, it is likely to be the Internet Protocol when looking at the description of the Signalling Gateway Function (SGW) in which it says: "... the IP based transport of signalling possibly used in post-R99 networks (i.e. between SIGTRAN SCTP/IP and SS7 MTP)". Taking this information and the evolution to an all-IP UMTS network architecture as discussed in [88], we assume that the Stream Control Transmission Protocol/Internet Protocol (SCTP/IP) is the signalling system. Should the SS7 still be used by a network operator, the SGW will perform the protocol conversion. SGW can be integrated into the STP network elements of the SS7.

SCTP is a reliable transport protocol operating on top of a connectionless packet network and is designed to transport PSTN (SS7) signalling messages over IP networks [89], hence the name SCTP/IP. It has been approved by the Internet Engineering Task Force (IETF) as a proposed standard to provide transport layer functions to many Internet applications [90]. SCTP was originally developed in [91] and later was updated by [92]. The current version is available in [89].

5.3.3 HSDPA capacity

Maximum system capacity

In [93], [94] and [95], it was found that the theoretical capacity HSDPA are 7.2 Mbps and 14.4 Mbps. Let us calculate how these figures are derived.

With the Release 5 WCDMA system, there are:

- 2560 chips/slot. As a chip corresponds with a modulation symbol, this is equivalent to 2560 symbols/slot;
- 3 slots/TTI;
- 5 TTIs/frame;
- 100 frames/s; and
- 15 channelisation codes (maximum possible) could be reserved for the HS-DSCH.

As discussed in Section 5.3.1, HS-PDSCH is the physical channel that carries users' data and has a fixed spreading factor of 16. Taking that into account, we have:

$$2560\ symbols/slot : 16 = 160\ symbols/slot$$

The symbol rate per second can then be obtained:

$$160\ symbols/slot\ \text{ x }\ 3\ slots = 480\ symbols/TTI$$

$$480\ symbols/TTI\ \ \text{x}\ \ 5\ TTIs = 2400\ symbols/frame$$

$$2400\ symbols/frame\ \ \text{x}\ \ 100\ frames = 240000\ symbols/s$$

$$2400\ symbols/s\ \ \text{x}\ \ 15\ codes = 3600000\ symbols/s/15\ codes$$

For Release 5 WCDMA system, only QPSK and 16-QAM modulation schemes are used. If QPSK is used, there are 2 bits per symbol. Therefore, the theoretical capacity of HSDPA is:

$$3600000\ symbols/s/15\ codes\ \ \text{x}\ \ 2\ bits/symbol = 7.2\ Mbits/s/15\ codes$$

On the other hand, if 16-QAM is used, there are 4 bits per symbol. Therefore, the theoretical capacity of HSDPA is:

$$3600000\ symbols/s/15\ codes\ \ \text{x}\ \ 4\ bits = 14.4\ Mbits/s/15\ codes$$

which are the 7.2 Mbps and 14.4 Mbps as mentioned in those papers. These values are when HSDPA is deployed on a separate frequency carrier.

In practice, the cell capacity cannot reach those theoretical maximum values. Furthermore, the technique to determine the actual capacity is also different. That technique relies on a link-level simulator. Here, we will show how the actual cell capacity is determined by using the techniques developed in [96] and [76]. As we do not have access to a link-level simulator, we will be using some simulation results from that literature for demonstration during this actual cell capacity determination process. The readers are referred to [96] and [76] to see the graphs where these simulation results are taken from.

User position dictates the geometry factor G, which is defined as:

$$G = \frac{P_{own}}{P_{other} + P_{noise}} \tag{5.14}$$

where P_{own} is the received own cell interference, P_{other} the received other cell interference and P_{noise} the received noise power. Technique to measure G in a live HSPA network in [97] can be employed. Simulation results show that typical values of G are from -5dB to 20 dB.

As G relates to the average power required for the HS-SCCH, $P_{HS-SCCH}$, from the value of G obtained from Equation (5.14), $P_{HS-SCCH}$ is obtained. Simulation results show that $P_{HS-SCCH}$ is approximately 0.1 W for G from -5dB to 20 dB.

Even when HSDPA is deployed using a separate carrier, some of the total power of a Node B, P_{Total}, will need to be dedicated to the common channels such as the S-CCPCH and the Primary-Common Control Physical Channel (P-CCPCH). A typical value of this power allocation is 20% [59], leaving the remaining 80% of power for HSDPA. A typical value of P_{Total} is 40 W [59].

The power for the HSDPA, P_{HSDPA}, is shared between the HS-DSCH and the HS-SCCH, that is:

$$P_{HSDPA} = P_{HS-DSCH} + P_{HS-SCCH} \tag{5.15}$$

where $P_{HS-DSCH}$ is the power used for the HS-DSCH. Rearranging Equation (5.15), we have:

$$P_{HS-DSCH} = P_{HSDPA} - P_{HS-SCCH} \tag{5.16}$$

$P_{HS-DSCH}$ and P_{Total} relate to the HS-DSCH Signal-to-Interference-plus-Noise Ratio, SINR, through the below equation:

$$SINR = 16 \cdot \frac{P_{HS-DSCH}}{P_{Total} \cdot \left(1 - \alpha + \dfrac{1}{G}\right)} \tag{5.17}$$

where α is the orthogonality factor and 16 is the spreading factor of the HS-PDSCH. With HSDPA, the performance evaluation is now SINR, which replaces the E_b/N_0 used in the Release 99 WCDMA system. Bit rate changes every TTI and, for each TTI, different modulation schemes, ECRs and the number of channelisation codes used have rendered SINR a more suitable measure. SINR above 7.5 enables the use of 16-QAM modulation scheme [76].

The actual maximum cell capacity (throughput) depends on the SINR. Because the number of channelisation codes influences the maximum cell capacity (throughput), Figure 5.1 shows the cell capacity values for different numbers of channelisation codes for the case of one user. For each number of channelisation code, the cell capacity also varies with multipath profiles.

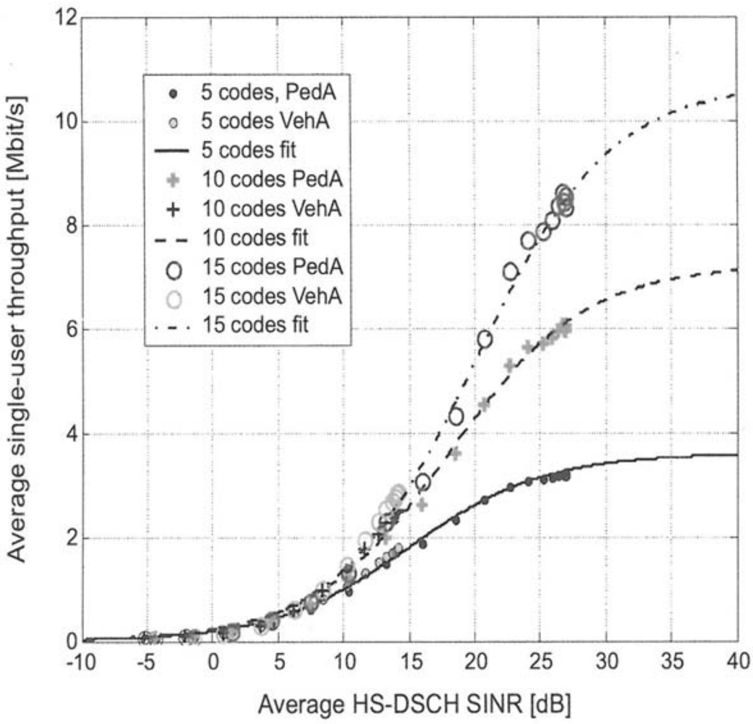


Figure 5.1: HSDPA throughput [76].

From Figure 5.1, it is seen that the capacity reaches around 10.5 Mbps with the 15-codes-fit curve when SINR equals 40 dB. This is around 73% of the theoretical capacity of HSDPA. With any other values of SINR, referring to Figure 5.1, respective cell throughputs can be obtained. For Smart Pricing purposes, η_M should be set to the actual cell capacity. The η_T is set using Equation 3.1.

If there are multiple users in the cell, the actual cell capacity is divided equally among the active users, if:

- Round Robin scheduling algorithm [76] is used;
- packet data transfer activity factors [76] of all users in the cell are the same; and
- all users are located in the same geographical area.

User data rate

The maximum bit rate a MS can achieve is dependent on the UE capability. In Release 5, there are 12 categories of UE capability. Using attributes of these categories, we can calculate the maximum achieved bit rates as well as the required code rates to achieve those bit rates. As an example, let us calculate these two values for a Category 10 UE.

From [85], for Category 10, we have:

- the maximum number of HS-DSCH codes received = 15;
- the minimum inter-TTI interval = 1; and
- the maximum number of bits of a HS-DSCH transport block received within a HS-DSCH TTI = 27952.

From [42], the slot structure of the HS-PDSCH is:

$$M \text{ x } 10 \text{ x } 16$$

where M is the number of bits per symbol and 16 is the spreading factor. With 16-QAM, M equals 4, giving us:

$$4 \text{ x } 10 \text{ x } 16 = 640\ bits/slot$$

$$640\ bits/slot \text{ x } 3\ slots/TTI = 1920\ bits/TTI$$

As the maximum number of HS-DSCH codes received is 15, we have:

$$1920\ bits/TTI \text{ x } 15\ codes = 28800\ bits/TTI/15\ codes$$

As the maximum number of bits of a HS-DSCH transport block received within a HS-DSCH TTI is restricted to 27952 as specified in [85], this means the code rate should be used is:

$$27952\ bits/TTI/15\ codes : 28800\ bits/TTI/15\ codes = 0.97$$

With the maximum number of bits of a HS-DSCH transport block received within a HS-DSCH TTI of 27952, together with the minimum inter-TTI interval is 1 (meaning that all 5 TTIs in a frame can be assigned to the MS), the maximum bit rate a MS can achieve is:

$$27952\ bits/TTI/15\ codes \text{ x } 5\ TTIs = 139760\ bits/frame/15\ codes$$

$$139760\ bits/frame/15\ codes \text{ x } 100\ frames = 13.976\ Mbits/s$$

This MS maximum bit rate is lower than the cell theoretical maximum capacity but higher than the actual cell capacity that we have found in Section 5.3.3.

For Smart Pricing purposes, the above MS maximum bit rate will not be used. Rather, we need to determine the minimum bit rate to which a MS's call will be reduced in the event of graceful degradation (see Section 3.5.1 for explanation), $R_{SP,min}$. We also need to determine the high bit rate, $R_{SP,max}$, for new users and users that submit the highest number of bids in the event of graceful degradation. Nominal values for these two bit rates are proposed as follows. It is at a network operator's discretion to set these at their desired levels by applying the same technique.

$R_{SP,min}$ is proposed to have 1 channelisation code with 1 slot per frame, which is:

$$640 \; bits/frame \;\; \text{x} \;\; 100 \; frames = 64000 \; bits/s$$

$R_{SP,max}$ is proposed to have 15 channelisation codes with 1 slot per frame, which is:

$$640 \; bits/frame \;\; \text{x} \;\; 100 \; frames \;\; \text{x} \;\; 15 \; codes = 960000 \; bits/s$$

Note that the above nominal values are the actual data rates a user will get for his/her call during the respective events. The data rates are guaranteed and will not be shared with other users. For new arrivals, this means the $R_{SP,max}$ is guaranteed until the end of the t_{gQ}.

5.3.4 Applicability of Smart Pricing signalling models to HSDPA

Referring to Section 5.1, we now verify whether we have gathered enough of the required five pieces of information so that Smart Pricing and our proposed Smart Pricing signalling models can be applied to HSDPA.

First, the maximum capacity of the system: in Section 5.3.3, we have shown the technique to derive this value. The η_T has also been addressed.

Second, the individual user load factor: in Section 5.3.3, we have shown the technique to derive two nominal values for the $R_{SP,min}$ and $R_{SP,max}$.

Third, the system network architecture: in Section 5.3.2, we discussed the similarities and differences between the Release 5 and the Release 99 system network architectures.

Taking the above information into account, the fourth piece of information, the Smart Pricing system diagram for HSDPA, is given in Figure 5.2.

A prominent difference between the Smart Pricing system diagram for HSDPA and the Smart Pricing system diagram for Release 99 WCDMA is that the DPE is now also connected to Node B. This is to account for the greater responsibilities Node B now holds.

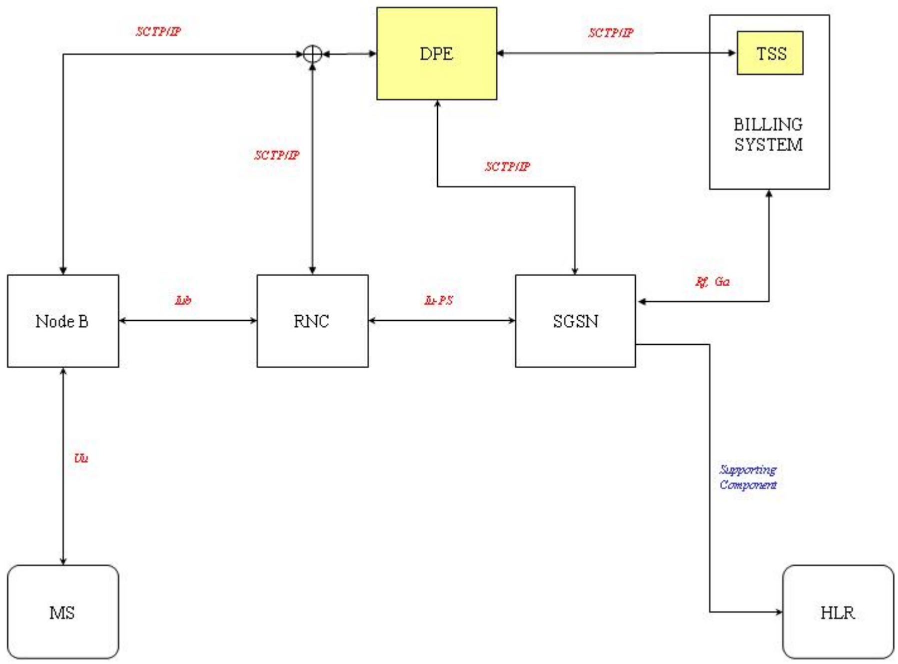


Figure 5.2: Smart Pricing in HSDPA System Diagram.

In Section 3.1, we stated that Node B does not have the necessary processing power to perform the resource management function, particularly reporting the cell load to the DPE. This was because the main function of Node B was then just power control [76]. With HSDPA, Node B, among others, performs scheduling and priority handling functions. This suggests that Node B is now able to report the cell load to the DPE. The benefit of having Node B connected to the DPE is the reduction of signalling delay, which translates to potentially more revenue to the network operator as the decision on adjusting price relative to load is quicker. The more rapid such a decision is made the less risks of losing potential new arrivals are.

The link to and from the RNC to the DPE is still kept. This is in anticipation for the expansion of our Smart Pricing signalling models that we will be proposing in the near future[9]. In those extended Smart Pricing signalling models, as the RNC controls multiple Node Bs, it will report loads of all the cells within its control to the DPE concurrently. With this information, the DPE will be able to make big-picture decisions about prices for all neighbouring cells in a single process. Such decisions will have the implications of discourage movement into congested cells while encouraging movement into under-utilised cells. That advanced feature can be implemented by informing MSs in a cell of prices of neighbouring cells, a feature of which is accomplished by our Smart Pricing signalling models.

[9]This is separate from the current research.

Finally, the fifth piece of information, the signalling components: with two minor modifications, the signalling components that we proposed in Section 3.7.2 can be used for HSDPA. One modification is needed to reflect the change that we mentioned above, i.e. Node B (replacing the RNC) sends cell load information to the DPE. Another modification is to replace the MSC with the SGSN to reflect the change in the circuit switched services to data services.

All the required five pieces of information of HSDPA are addressed, they can be used in their respective places in our proposed Smart Pricing signalling models for Release 99 WCDMA system in Chapters 3 and 4. Thus, we affirm that using our proposed Smart Pricing signalling models Smart Pricing can be applied to HSDPA.

5.3.5 Estimated required signalling for Smart Pricing in HSDPA

To estimate the required signalling when Smart Pricing is adopted in the HSDPA system, we have chosen the 2-WTP small system for demonstration of the technique we used. The maximum average signalling loads will be estimated, however, using the technique all other signalling parameters for this system as well as other systems can also be estimated.

It is noted that the two minor modifications mentioned in the previous section are simply a one-for-one replacement for the purpose of analysis and does not affect the number of signalling components proposed in Section 3.7.2.

In Section 5.3.3, η_M is found to be 10.5 Mbps, and in Section 5.3.3, 64 kbps and 960 kbps are nominated for the $R_{SP,min}$ and $R_{SP,max}$. With η_T factor = 0.25, using Equation (3.1), η_T is found to be:

$$10.5 \text{ x } 0.25 = 2.625 \ Mbps$$

Using Equation (4.27), the number of $R_{SP,max}$ users that can be accommodated with η_T is:

$$M_{\eta_T}^{R_H} = Integer\left(\frac{2.625}{0.064}\right)$$

$$M_{\eta_T}^{R_H} = 41 \ users$$

From Equation (4.30), we have:

$$(g + y + h) \leq \frac{10.5 - (41 \cdot 0.064)}{0.96}$$

$$(g + y + h) \leq 8.2\ users$$

Using Equation (4.31), the maximum possible number of users in the system is:

$$M_{max} = 41 + 8 = 49\ users$$

In Tables 3.14 and 3.32, the maximum number of users in the system for the small and large system are 10 and 16, respectively. This is a difference of 6 users in system. Referring to Table 4.20, an increase of 60% in the number of users results in an increase of 5.76 %, 1.5% and 9.77% in the UL, DL and Netw2Netw maximum signalling load, respectively. In Table 4.20, the UL, DL and Netw2Netw maximum average signalling loads are 3.82, 4.00 and 66.24 kbps, respectively. Assuming a linear relationship between the increase in the maximum number of users and the increase in the signalling load, with 49 users, it is estimated that the Smart Pricing maximum average signalling loads in the steady-state condition for the:

- UL is:

$$\frac{\frac{49-10}{10} \cdot 5.76}{\frac{16-10}{10}} = 37.44$$

$$\frac{37.44 + 100}{100} \cdot 3.82 = 5.25\ kbps$$

- DL is 4.39 kbps, and
- Netw2Netw is 108.31 kbps.

5.4 High speed packet access evolution

5.4.1 Characteristics of HSPA+

High Speed Packet Access Evolution (HSPA+) is an improvement of the HSPA and the technical specifications for this technology was first available in 3GPP Release 7. It is the High Speed Packet Access Evolution in this Release that we study in this section. As HSDPA has been discussed in the previous section, for the improvement to stand out, we have chosen to concentrate in the downlink direction of the High Speed Packet Access Evolution. High Speed Packet Access Evolution is a result of the initiative proposed by 3G America in 2006 and is now called HSPA+ [98]. Only software upgrades are required for HSPA+ to be implemented [99].

Enhancements that HSPA+ brings are summarised in [59] and [99], which include: end-user-performance, network capacity and network architecture. The first enhancement is discussed here and the latter two in the following sections.

With HSPA+, first, a MS can save power by entering the power-saving mode when no data is to sent to it. In ideal situations, it is estimated that this discontinuous downlink reception feature could save 50% of power consumption. Second, the introduction of the enhanced Cell_FACH (eCell_FACH) concept advances the end-user performance by reducing call set up and allocation times. With eCell_FACH, both the HS-DSCH and HS-PDSCH are used for the Cell_FACH and Cell_DCH states. This allows for continuation of data reception during the Cell_FACH to Cell_DCH state transition when larger amounts of data need to be received. In addition, before the introduction of the eCell_FACH, when reception of a small amount of data is required, the FACH can be used. Now, the capacity of the FACH is increased to over 1 Mbps, which allows for the channel to be able to accommodate more types of application.

Finally, *the flexible RLC and Medium Access Control (MAC) segmentation* solution also boosts the end-user performance. Flexible RLC allows for flexible RLC block sizes to be used. Smaller block sizes are advantageous for low-delay real-time application like VoIP, whereas large block sizes are an enabling factor for high bit rates due to reduction in packet processing time. Large RLC block sizes together with segmentation at the MAC brings the total layer-2 overhead down to 1%.

5.4.2 HSDPA+ capacity

Capacity of HSPA+ on the downlink is higher than the HSDPA. Its peak data rate is up to two times that of the HSDPA. This is as a result of the addition of the 64-QAM modulation, the use of multiple antennas for reception and transmission with Multiple Input Multiple Output (MIMO), and the use of advanced receivers [59].

64-QAM helps increase the peak rate because it increases the coding rate to 6 bits/Hz. MIMO helps increase capacity by using more antennas because antenna diversity improves Signal-to-Inference Ratio (SIR). Higher SIR means higher channel CQI, and higher CQI enables higher rates [99]. Capacity gains from using MIMO can be dramatically impacted with more realistic assumptions as flagged in [100]. The enhanced type 3 advanced receiver also helps to increase the SIR, which in turn gives rise to higher capacity. Chip equaliser is utilized in the receiver which removes intra-cell interference from multipath propagation [99]. When interference is less, the SIR is higher.

Maximum system capacity

Using the approach in Section 5.3.3, we can calculate the theoretical capacity of HSPA+ in the downlink direction when HSPA+ is deployed on a separate frequency carrier. In that section, the symbol rate per second is found to be:

$$2400\ symbols/s\ \ \text{x}\ \ 15\ codes = 3600000\ symbols/s/15\ codes$$

With 64-QAM, there are 6 bits per symbol. Hence, the theoretical capacity with 64-QAM is:

$$3600000\ symbols/s/15\ codes\ \ \text{x}\ \ 6\ bits = 21600000\ bits/s/15\ codes$$

In practice, the cell capacity cannot reach this value. The system simulation results in [59] show a maximum bit rate of around 925 kbps per user in a macro cell with 20 users in the cell. If we assume that all users are near to the Node B, the actual maximum cell bit rate is then 18.5 Mbps. This is around 86% of the theoretical capacity.

Again, as in Section 5.3.3, if we assume that:

- Round Robin scheduling algorithm is used;
- packet data transfer activity factors of all users in the cell are the same; and
- all users are located in the same geographical area,

then that actual bit rate can be divided equally among the active users. For Smart Pricing purposes, η_M should be set to the actual cell bit rate. The η_T is set using Equation 3.1.

User data rate

In [99], it is stated the peak bit rate with 64-QAM is 21.1 Mbps. Let us calculate how this figure is derived.

Comparing [42] and [101], we see no change in the sub-frame structure for the HS-PDSCH, therefore we can use the approach in Section 5.3.3 to calculate the HSPA+ downlink user data rate.

The slot structure of the HS-PDSCH [101] is:

$$M\ \ \text{x}\ \ 10\ \ \text{x}\ \ 16$$

With 64-QAM, M equals 6, we then have:

$$6 \text{ x } 10 \text{ x } 16 = 960 \; bits/slot$$

$$960 \; bits/slot \text{ x } 3 \; slots/TTI = 2880 \; bits/TTI$$

With 15 HS-DSCH channelisation codes, we have:

$$2880 \; bits/TTI \text{ x } 15 \; codes = 43200 \; bits/TTI/15 \; codes$$

The maximum number of bits of a HS-DSCH transport block received within a HS-DSCH TTI is restricted to 42192 (without MIMO) as specified in [102], this means the code rate used is:

$$42192 \; bits/TTI/15 \; codes : 43200 \; bits/TTI/15 \; codes = 0.98$$

With the maximum number of bits of a HS-DSCH transport block received within a HS-DSCH TTI of 42192, together with the minimum inter-TTI interval is 1, the maximum bit rate a MS can achieve is:

$$42192 \; bits/TTI/15 \; codes \text{ x } 5 \; TTIs = 210960 \; bits/frame/15 \; codes$$

$$210960 \; bits/frame/15 \; codes \text{ x } 100 \; frames = 21.096 \; Mbits/s$$

which is 21.1 Mbps when rounded up. This user maximum bit rate is lower than the cell theoretical maximum capacity but higher than the actual cell capacity that we have found in Section 5.4.2. The capacity with MIMO is 28 Mbps [99].

For the purposes of demonstrating Smart Pricing, we need to determine $R_{SP,min}$ and $R_{SP,max}$. Nominal values for these two bit rates are proposed as follows. It is at a network operator's discretion to set these at their desired levels by applying the same technique.

$R_{SP,min}$ is proposed to have 1 channelisation code with 1 slot per frame, which is:

$$960 \; bits/frame \text{ x } 100 \; frames = 96000 \; bits/s$$

$R_{SP,max}$ is proposed to have 15 channelisation codes with 1 slot per frame, which is:

$$960 \; bits/frame \text{ x } 100 \; frames \text{ x } 15 \; codes = 1440000 \; bits/s$$

5.4.3 HSDPA+ system architecture

With HSPA+, there are two changes to the system architecture. The changes reflect the evolution towards a flat architecture which is designed to be backward compatible [99]. First, the *GPRS One Tunnel Solution* enables user data to pass directly from the GGSN to the RNC. No change, however, is made on the control plane. The change results in a decrease in latency on the end-to-end transmission path and in cost saving [98]. Second, the RNC functionality is integrated into Node B. The benefit of the whole architecture change is that only two network elements are needed for user data operation. This allows for flexible scalability and is an enabling factor for higher data rate with HSPA+ [59]. The reduction in latency also improves the overall performance of IP-based services [99].

It is noted that, although in [98], [59] and [99], it is shown that the interfaces supporting user traffic is directly passed from the GGSN to the RNC for the first change, and directly from GGSN to Node B for the overall change, no such changes are seen in the system architecture (i.e. the basic configuration of a PLMN) in [103] as Release 7 was frozen.

5.4.4 Applicability of Smart Pricing signalling models to HSDPA+

Referring to Section 5.1, let us verify that we have gathered enough of the required five pieces of information so that Smart Pricing and our proposed Smart Pricing signalling models can be applied to HSDPA.

First, the maximum capacity of the system: in Section 5.4.2, we have shown the techniques to calculate the η_M, as well as η_T.

Second, the individual user load factor: in Section 5.4.2, we have shown the techniques to calculate the $R_{SP,min}$ and $R_{SP,max}$.

Third, the system network architecture: in Section 5.4.3, we have discussed the changes made in the Release 7.

Taking that information into account, the fourth piece of information, the Smart Pricing system diagram for HSDPA+, is given in Figure 5.3. In the figure, the SGSN and GGSN are combined into a single network element. This is to address the the network architecture change in the user plane while keeping the control plane functionality as this is needed for Smart Pricing for assigning variable bit rates.

Finally, the fifth piece of information, the signalling components: the same two modifications to Section 3.7.2 made for the HSDPA are also required for the HSPA+. One modification is to replace RNC with Node B and the other is to replace the MSC with the SGSN/GGSN.

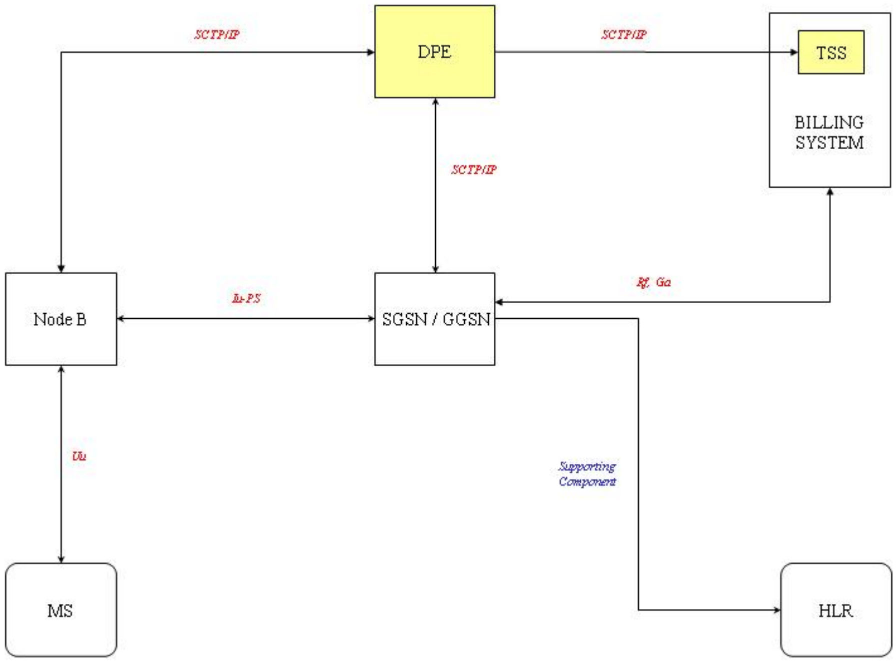


Figure 5.3: Smart Pricing in HSDPA System Diagram.

As all the required five pieces of information of HSPA+ are addressed, they can be used in their respective places the MCS and SSA models in Chapters 3 and 4. Thus, we affirm that using our proposed Smart Pricing signalling models Smart Pricing can be applied to HSPA+.

5.4.5 Estimated required signalling for Smart Pricing in HSPA+

Using the technique in Section 5.3.5, in this section we estimate the Smart Pricing maximum average signalling loads for HSPA+. It can be seen from the previous section that with HSPA+ there are no changes to the number of signalling components proposed in Section 3.7.2.

In Section 5.4.2, η_M is found to be 18.5 Mbps, and in Section 5.4.2, 96 kbps and 1.44 Mbps are nominated for the $R_{SP,min}$ and $R_{SP,max}$. With η_T factor = 0.25, using Equation (3.1), η_T is found to be:

$$18.5 \ \text{x} \ 0.25 = 4.625 \ Mbps$$

Using Equation (4.27), the number of $R_{SP,max}$ users that can be accommodated with η_T is:

$$M_{\eta_T}^{R_H} = Integer\left(\frac{4.625}{0.096}\right)$$

$$M_{\eta_T}^{R_H} = 48 \; users$$

From Equation (4.30), we have:

$$(g + y + h) \leq \frac{18.5 - (48 \cdot 0.096)}{1.44}$$

$$(g + y + h) \leq 9.6 \; users$$

Using Equation (4.31), the maximum possible number of users in the system is:

$$M_{max} = 48 + 9 = 57 \; users$$

Therefore, the estimated Smart Pricing maximum average signalling loads in the steady-state condition for the:

- UL is:

$$\frac{\frac{57 - 10}{10} \cdot 5.76}{\frac{16 - 10}{10}} = 45.12$$

$$\frac{45.12 + 100}{100} \cdot 3.82 = 5.54 \; kbps$$

- DL is 4.47 kbps; and
- Netw2Netw is 116.93 kbps.

5.5 Long term evolution

5.5.1 Characteristics of LTE

Long Term Evolution (LTE) is a new 3GPP packet-only wideband radio network with flat architecture [59]. It is similar to the HSPA+ system not only with a flat architecture but also with the same aim to achieve higher peak rates, greater throughput, lower delays and lower UE power consumption than the 3GPP Release 6 UMTS system. Furthermore, both HSPA+ and LTE use MIMO and the same modulation schemes, as well as ceasing to support soft handover. However, LTE is different from HSPA+ in that the system bandwidth is not fixed. This widens opportunities for LTE to use spectrum that are freed up due to frequency band restructuring to accommodate the ever increase in mobile data usage and number of subscribers.

Freed up spectrum bandwidths vary in sizes, for example the 2x5 MHz with 10 MHz split or 2x7.4875 MHz with 10 MHz split in the 403-520 MHz band identified in [104], or the 2x70 MHz with 120 MHz split in the 2.5-2.69 GHz band identified in [105]. Many countries, including Australia, are switching over from analogue to digital television. As digital television can broadcast at higher quality with less spectrum, such a switch-over process will result in more spectrum available which could be utilised for LTE. LTE can operate with bandwidth ranging from 1.4 to 20 MHz.

Although LTE will likely start by using the new 2.6 GHz band [106], other frequencies can also be used. This is part of the flexibility in bandwidth and frequency offered by LTE. A full list of current frequencies for LTE can be found in [107].

Another major difference between LTE and HSPA+ is that it no longer uses WCDMA as the multiple access protocol. Instead Orthogonal Frequency Division Multiplexing (OFDM) is used in the downlink direction and the Single Carrier - Frequency Division Multiple Access (SC-FDMA) in the uplink direction. OFDM is chosen for the downlink because it is more flexible with different bandwidth and the signals remain orthogonal when the system bandwidth increases. SC-FDMA is chosen for the uplink because it requires low peak to average ratio which helps improve UE power consumption [59].

LTE allows for scheduling in both time and frequency domains, whereby frequency scheduling is enabled by the use of OFDM. Up to 4x4 MIMO can be used for LTE, however with only 2x2 MIMO, peak rates of 50 Mbps for the uplink and 100 Mbps for the downlink set for LTE are already met. Section 5.5.2 proves this for downlink.

LTE minimises latency by using a flat architecture because less network elements means lower round trip time, particularly time spent at the interfaces. It can be seen in [108] and [109] that the radio access network now reduces to only one element. Apart from this change, LTE system architecture also includes new network elements, interfaces and structure. Section 5.5.3 provides further details about LTE architecture.

Lastly, with LTE, there are no dedicated transport channels. Only common transport channels are available and the downlink user data are now transmitted on the Downlink Shared Channel (DL-SCH) [106]. This transport channel is mapped onto the Physical Downlink Shared Channel (PDSCH). The FDD physical channel radio frame structure has a duration of 10 ms which contains 10 sub-frames each 1 ms, equivalent to a TTI [110].

5.5.2 LTE capacity

When compared with HSDPA, the capacity of LTE is significantly higher. This is due to the spectral efficiency gain of over three times that of the HSDPA and other improvements. Factors that constitute to that gain are: OFDM with frequency domain equalization, frequency domain packet scheduling, MIMO and inter-cell interference rejection combining [106].

Maximum system capacity

Each OFDM sub-carrier is 15 kHz and corresponds to 2 slots in a TTI of 1 ms. Hence, the duration of a slot is 0.5 ms. A *resource block* contains 12 sub-carriers and has a total bandwidth of 180 kHz. A physical resource block is defined as the number of consecutive OFDM symbols in the time domain and the number of sub-carriers in the frequency domain. The minimum and maximum number of resource blocks are 6 (for 1.4 MHz bandwidth) and 110 (for 20 MHz bandwidth) [110]. In some literature, a maximum value of 100 is found, for instance in [106].

The number of OFDM symbols per slot depends on the cyclic prefix length and sub-carrier spacing. With a 12 sub-carriers per resource block, the number of OFDM symbols in a slot is 7 for the normal cyclic prefix and 6 for the extended cyclic prefix [110]. These two values are due to the symbol duration and cyclic prefix duration.

The duration of a symbol with a normal cyclic prefix is 66.68 x 10^{-3} ms, and the duration of the normal and extended cyclic prefix are 5.21 x 10^{-3} ms and 16.76 x 10^{-3} ms, respectively [106]. Hence, we have:

$$\frac{0.5\ ms}{66.68 \text{ x } 10^{-3}\ ms} = 7.5$$

for the normal cyclic prefix case, and for the extended cyclic prefix case:

$$\frac{0.5\ ms}{(66.68 - 5.21 + 16.76) \text{ x } 10^{-3}\ ms} = 6.4$$

which are equivalent to 7 and 6 symbols when flooring to the relevant integers. The equivalent number of symbols per sub-carrier is then 14 and 12, respectively. To calculate the maximum system capacity (i.e maximum bit rate), we will be using the number of symbols with normal cyclic prefix per sub-carrier.

Not all of the 14 symbols are used for user data. For every 3 sub-carriers, 3 symbols need to be sacrificed for the PDCCH overhead and 4 for the downlink reference signals when MIMO, with 2 antennas for transmit and 2 antennas for receive (i.e. 2x2 MIMO), is used. Other overhead values of the downlink reference signals for the MIMO-less and 4x4 MIMO cases can be found in [106]. The total is then 7 symbols/3 sub-carriers, which is equivalent to:

$$7 \text{ x } \frac{12}{3} = 28 \ symbols/resource\ block/TTI$$

If the maximum number of resource blocks of 100 for the 20 MHz bandwidth is used, we have:

$$28 \text{ x } 100 = 2,800 \ symbols/20\ MHz/TTI$$

which is equivalent to:

$$2,800 \text{ x } 1,000 = 2,800,000 \ symbols/20\ MHz/s$$

The total number of symbols in the 20 MHz bandwidth is given by:

$$14 \text{ x } 12 \text{ x } 100 \text{ x } 1,000 = 16,800,000 \ symbols/20\ MHz/s$$

In addition to the two overheads mentioned above, another overhead needs to be accounted for, which is dependent on the bandwidth used. This overhead is required for the synchronization signal, Physical Broadcast Channel, Physical Control Format Indicator Channel and one group of Physical Hybrid Automatic Repeat Request Indicator Channel. With a bandwidth of 20 MHz, around 1% of the number of symbols needs to be sacrificed for this overhead [106]. The net number of symbols for user data therefore is:

$$16,800,000 \text{ x } (100 - 1) - 2,800,000 = 13,832,000 \ symbols/20\ MHz/s$$

With the 6 bits/symbol of the 64-QAM modulation and without coding we then have:

$$13,832,000 \text{ x } 6 = 82.992\ Mbps/20\ MHz/s$$

Finally, with the use of the 2 antennas for 2x2 MIMO, the bit rate is doubled (because the number bits per symbol is twice the usual amount) [106]. Hence, the theoretical maximum system capacity is:

$$82,992,000 \text{ x } 2 = 165.984\ Mbps/20\ MHz/s$$

In practice, the above value is unlikely to be reached because the value assumes that the channel condition is ideal, in which the need for coding is not required. As pointed out in Sections 5.3.3 and 5.4.2, the actual system capacity of a system can be determined using simulation programs. In those two sections the actual capacity was found to be 73-86%. As LTE is available from 3GPP Release 8, which is later than the Releases that HSDPA and HSPA+ were first included, we would expect that the percentage to be at least the same for LTE. At 86%, the actual system capacity is:

$$165.984 \text{ x } 0.86 = 142.746\ Mbps/20\ MHz/s$$

For Smart Pricing purposes, the η_M should be set to the above actual system capacity value. The η_T is then set using Equation 3.1.

User data rate

With LTE, user data is transmitted on the PDSCH. A PDSCH frame contains 10 subframes, each of which has a duration of 1 TTI [110]. In [111], the maximum number of bits of a PDSCH *transport block* received within a TTI for UE Category 4 is 75,376 bits. A transport block is defined as the basic data unit exchanged between the physical layer and the MAC [112]. Referring to [113], 75,376 bits corresponds to the transport block size for the number of physical resource block ranging from 100 to 110, the range of the maximum number of resource blocks for 20 MHz bandwidth as discussed in Section 5.5.2. We also see in [113] that 75376 bits and 100-110 resource blocks correspond to a modulation order of 6, which is 64-QAM (i.e. 6 bits/symbol). In addition, in [111], the maximum number of supported layers for spatial multiplexing in the downlink is found to be 2, which corresponds to 2x2 MIMO, hence the maximum number of bits received within a TTI is: $2 \text{ x } 75376 = 150,752$. The maximum user rate is then 150.752 Mbps.

In the previous section, for a system with a 20 MHz bandwidth, 64-QAM and 2x2 MIMO, the theoretical maximum capacity was found to be 165,984, and the actual capacity was 142.746 Mbps. Compared with these values, the maximum user rate of 150.752 Mbps is around 91%. This is consistent with the HSDPA and HSPA+ cases, in which the maximum user data rate is lower than the theoretical maximum system capacity but higher than the actual maximum system capacity.

For Smart Pricing purposes, we nominate the two bit rates $R_{SP,min}$ and $R_{SP,max}$ as the following. First, we define a *resource unit* as 12 sub-carriers in the frequency domain and 1 slot in the time domain. Thus, we have:

$$2 \text{ x } 1,000 = 2,000 \ \textit{resource units/s}$$

As found in the previous section, with a 20 MHz bandwidth the net number of symbols for user data is 13,832,000. This corresponds to:

$$\frac{13832000}{2000} = 6,916 \ \textit{symbols/resource unit/s}$$

and with 64-QAM and 2x2 MIMO, it is equivalent to:

$$6916 \ \text{ x } \ 6 \ \text{ x } \ 2 = 82.992 \ \textit{kbps/resource unit}$$

The $R_{SP,min}$ is proposed to have 2 resource units, which is:

$$82.992 \ \text{ x } \ 2 = 166 \ \textit{kbps}$$

and the $R_{SP,max}$ is proposed to have 20 resource units, which is:

$$82.992 \ \text{ x } \ 20 = 1.66 \ \textit{Mbps}$$

Both bit rates are higher than those for HSDPA and HSPA+ to reflect the capacity increase due to system evolution. Even with $R_{SP,min}$, it is proposed to be over three times the dial-up bit rate.

5.5.3 LTE system architecture

Evolution in the 3GPP UMTS system capacity and radio interface necessitates a System Architecture Evolution (SAE). Targets of the SAE includes: optimization for packet switched services, improvement in the packet delivery delays and simplification of the system compared to the existing 3GPP and other cellular systems. A flat system architecture was deemed necessary as it facilitates reduction in delay and improvement in performance [106].

The flat LTE system architecture looks similar to the flat HSPA+ architecture, at least to the extent of the Evolved Packet System (EPS). Nevertheless, most of the network elements have different names and modified and/or additional functions. There are also new network elements and interfaces. Some similarities are that the Evolved-UTRAN Node B (eNB) replaces of the combined RNC and Node B, the Mobile Management Entity (MME) replaces of the SGSN, and the Serving Gateway (S-GW) and the Packet Data Network Gateway (PDN-GW) combined replace of the GGSN. In some literature such as [114] and [106], instead of PDN-GW, the abbreviation P-GW is used. Below, we provide a brief description of the main network elements together with their functions and the interfaces between them. The reader is referred to [108], [114], [109] and [115] for details.

The eNB's functions include: radio resource management, selection of an MME at UE attachment, routing user plane data to the S-GW, and scheduling [109]. The interface between the eNB and the UE is LTE-Uu [108]. Being responsible for these functions, clearly the eNB has the processing power to report the cell load to the DPE like it is in the case of Node B of the HSDPA and HSPA+.

The MME's functions include: authentication, authorization, S-GW and PDN-GW selection, bearer management, dedicated bearer establishment, MME selection for handovers with MME changes, roaming and lawful interception of signalling traffic [109]. Performing functions such as the former two, the MME needs to connect to the Home Subscriber Server (HSS) where subscriber data are stored. The HLR is a subset of the HSS functionality [116]. The interface between the MME and the eNB is S1-MME, and between the MME and the HSS is S6a [108]. Overall, the MME is comparable to the MSC in the circuit switched domain and the SGSN in the packet switched domain of non-LTE 3GPP systems. With Smart Pricing, the MME will be responsible for checking the subscriber's profile, holding the connection until the subscriber accepts the admission price and sending CDRs to the Billing System.

The Policy and Charging Resource Function (PCRF) is a new network element when compared with non-LTE 3GPP systems. Its functions include: controlling service data flow detection, gating, QoS and flow-based charging towards the Policy and Charging Enforcement Function (PCEF). The PCRF specifies the treatment the PCEF needs to give to a service data flow. If the service data flow is tunnelled at the Bearer Binding and Event Reporting Function (BBERF), the PCRF also provides the BBERF with necessary information for it to perform that function [115]. As the PCRF may use user's subscription information as a basis for policy and charging decisions, a connection to Subscription Profile Repository (SPR) is required. Until the present time, the relationship between the SPR and the HSS, which is the existing subscriber database, is not specified [117]. However, it is said that the SPR can be combined with other databases, thus we consider it part of the HSS. The interface between the PCRF and SPR is Sp [115]. With Smart Pricing, a connection between the PCRF and DPE is necessary for the DPE to control the user's data rate.

The S-GW's functions include: packet routing and forwarding as well as UL and DL charging per UE, PDN and QoS Class Identifier (QCI) [109]. The BBERF is part of the S-GW [106]. With the charging function, the S-GW needs to be connected to the Charging Data Function (CDF). The interface between the S-GW and:

- the MME is S11;
- the CDF is Rf;
- the eNB is S1-U; and
- the PCRF is Gxc.

With Smart Pricing, the DPE does not need to connect to the S-GW. This is because the S-GW's functions relate to user's data, the bearers of which are managed by the MME and PCRF, therefore only connections from the DPE to MME and PCRF are required.

The PDN-GW's functions include: UE IP address allocation, packet filtering, packet marking, and UL and DL service level charging, gating and rate enforcement [109]. The enforcement function in the PDN-GW is the PCEF. Like the S-GW, with the charging function, a connection between the PDN-GW and CDF is required. The interface between the PDN-GW and:

- the S-GW is S5, however, [106] additionally also specifies S8 as the interface. S8 is said to be the inter PLMN variant of S5 [108];
- the CDF is Rf; and
- the PCRF is Gx.

Again, with Smart Pricing, the DPE does not need to connect to the PDN-GW for similar reasons as explained for the S-GW.

Some of the above LTE network elements can be combined with another one to form a single network element. Options are for the S-GW to be combined with either the MME or the PDN-GW [108]. We choose the latter and adopt the name SAE Gateway (SAE-GW) in [106] for the combined network element. In the next section, we will show visually how LTE EPS network elements interconnect. For simplicity, we assume that both the CDF and CGF are collocated with the SAE-GW, making the Rf and Ga interfaces to the Offline Charging System (OFCS) [106] and the Gy to the Online Charging System (OCS) [114] internal.

5.5.4 Applicability of Smart Pricing signalling models to LTE

Referring to Section 5.1, let us verify if we have gathered enough of the required five pieces of information so that Smart Pricing and our proposed Smart Pricing signalling models can be applied to LTE.

First, the maximum capacity of the system: in Section 5.5.2, we have shown the techniques to calculate the η_M and η_T.

Second, the individual user load factor: in Section 5.5.2, we have shown the techniques to calculate the $R_{SP,min}$ and $R_{SP,max}$.

Third, the system network architecture: in Section 5.5.3, we have outlined the evolution of the LTE architecture and the relevant network elements to be used with Smart Pricing.

Incorporating the DPE into the LTE architecture, the fourth piece of information, the Smart Pricing system diagram for LTE, is given in Figure 5.4. To emphasize main Smart Pricing signalling links, broken lines are used for other links.

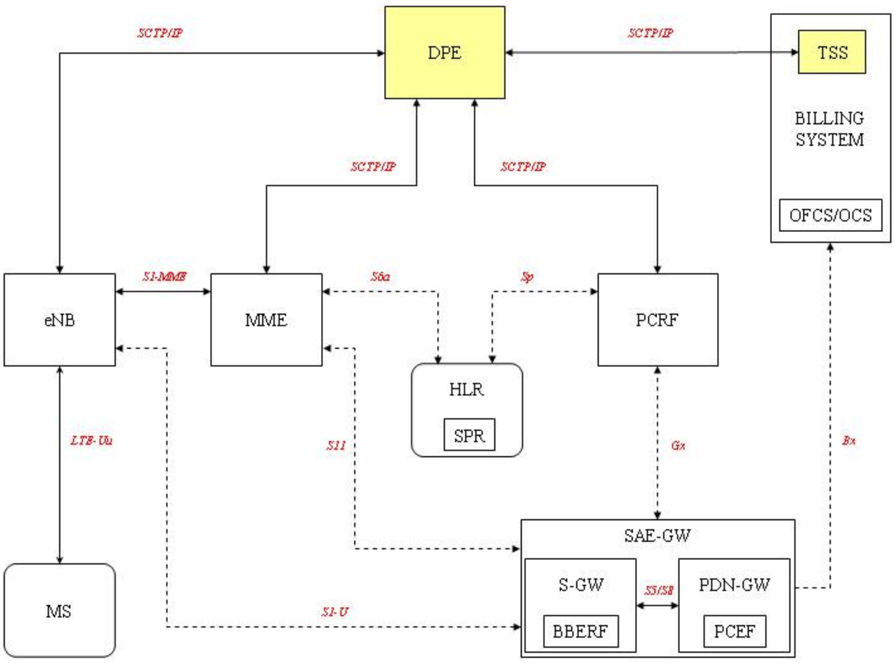


Figure 5.4: Smart Pricing in LTE System Diagram.

Finally, the fifth piece of information, the signalling components: the following modifications to Section 3.7.2 should be made:

- eNB replaces RNC; and
- MME and PCRF replace MSC. This is because both the MME and PCRF have control over the setting up of a call and specifying the call's bit rate.

As all the required five pieces of information of LTE are addressed, they can be used in their respective places the MCS and SSA models in Chapters 3 and 4. Thus, we affirm that using our proposed Smart Pricing signalling models Smart Pricing can be applied to LTE.

5.5.5 Estimated required signalling for Smart Pricing in LTE

Using the technique in Section 5.3.5, in this section we estimate the Smart Pricing maximum average signalling loads for LTE.

As pointed out in the previous section, the MSC is replaced by the MME and PCRF. This means any MSC-related signalling components proposed in Section 3.7.2 are now doubled. For the small system case, we see in Table 4.20 that the sum of the blocking probabilities is over 0.95, which means the majority of the signalling will be used to block new arrivals. Thus, we will focus on doubling the MSC-related signalling components due to blocking. In Table 4.20, it is seen that the blocking probability due to insufficient capacity is 0.1083 and blocking probability due to insufficient WTP is 0.842.

In Table 3.5, there is only one UL MSC-related signalling component and in Table 3.6 there are one UL MSC-related signalling component and two NetwNetw MSC-related signalling components. There is no MSC-related signalling component on the DL. Therefore, by replacing the MSC with the MME and PCRF, the revised maximum average signalling loads on the:

- UL is:

$$3.82 \cdot \left[(2 \cdot 0.1083) + \left(\frac{3}{2} \cdot 0.842\right) + (1 - 0.1083 - 0.842)\right] = 5.84 \ kbps$$

- DL remains at 4 kbps; and

- NetwNetw is:

$$66.24 \cdot \left[0.1083 + \left(\frac{9}{5} \cdot 0.842\right) + (1 - 0.1083 - 0.842)\right] = 110.86 \ kbps$$

In Section 5.5.2, η_M is found to be 142.746 Mbps, and in Section 5.5.2, 166 kbps and 1.66 Mbps are nominated for the $R_{SP,min}$ and $R_{SP,max}$. With η_T factor = 0.25, using Equation (3.1), the η_T is found to be:

$$142.746 \text{ x } 0.25 = 35.69 \ Mbps$$

Using Equation (4.27), the number of $R_{SP,max}$ users that can be accommodated with η_T is:

$$M_{\eta_T}^{R_H} = Integer\left(\frac{35.69}{0.166}\right)$$

$$M_{\eta_T}^{R_H} = 214\ users$$

From Equation (4.30), we have:

$$(g + y + h) \leq \frac{142.746 - (214 \cdot 0.166)}{1.66}$$

$$(g + y + h) \leq 64.59\ users$$

Using Equation (4.31), the maximum possible number of users in the system is:

$$M_{max} = 214 + 64 = 278\ users$$

Therefore, the estimated Smart Pricing maximum average signalling loads in the steady-state condition for the:

- UL is:

$$\frac{\frac{278 - 10}{10} \cdot 5.76}{\frac{16 - 10}{10}} = 257.28$$

$$\frac{257.28 + 100}{100} \cdot 5.84 = 20.87\ kbps$$

- DL is 6.68 kbps; and
- Netw2Netw is 594.65 kbps.

5.6 Findings and discussions

In this chapter, we have investigated the applicability of Smart Pricing and our proposed Smart Pricing signalling models (the MCS and SSA models) to the 3GPP Release 99 UMTS downlink system and to other more advanced 3GPP mobile telecommunications systems, those are: the HSDPA, HSPA+ and LTE. For the downlink, the intention is to provide mobile telecommunications network operator with a complete solution for implementing Smart Pricing in the Release 99 UMTS system, whereas, for the HSDPA, HSPA+ and LTE systems, our intention is to show how robust the MCS and SSA models are. The models can be applied not only to the existing system but also to more advanced systems that are currently being rolled out or will be rolled out in the future.

To show the applicability of Smart Pricing and the MCS and SSA models, we examine each system in details. We summarise each system's characteristics and analyse the evolution of each system's capacity and architecture. We provide techniques to calculate the maximum system capacity and the share of that capacity by an individual user. We nominate two bit rates to be assigned to the users for Smart Pricing purposes. We also construct Smart Pricing system diagrams for the HSDPA, HSPA+ and LTE systems. Finally, we develop a technique to estimate the Smart Pricing maximum average signalling loads for all four systems and use it to produce our estimated results.

In summary, we have:

1. demonstrated through our 5-step verification approach that Smart Pricing and the MCS and SSA models can be applied to the 3GPP Release 99 UMTS downlink, HSDPA, HSPA+ and LTE systems; and

2. estimated Smart Pricing maximum average signalling loads for the 3GPP Release 99 UMTS downlink, HSDPA, HSPA+ and LTE systems. These results are summarised in Table 5.5 below.

System	Maximum Average Signalling Load			Unit
	UL	DL	Netw2Netw	
Release 99 Downlink	3.82	4.00	66.24	kbps
HSDPA	5.25	4.39	108.31	kbps
HSPA+	5.54	4.47	116.93	kbps
LTE	20.87	6.68	594.65	kbps

Table 5.5: Estimated maximum average signalling loads.

It can be seen from the results that the higher system capacity is, the more signalling resources need to be dedicated for Smart Pricing. Although this is expected, we have quantified the estimated increases. The results provide a good indication of the required signalling when Smart Pricing is adopted before comprehensive simulations take place.

The results in Table 5.5 confirm that adopting Smart Pricing will not impose significant signalling traffic on the network. Note that, although the Netw2Netw maximum average signalling loads seem to be high, they are shared between a number of links, not just one link.

CHAPTER 6

Applicability of the proposed models in non-cellular telecommunications and other resource-constrained systems

In this chapter, we address how Smart Pricing and the MCS and SSA signalling models can also be applied to non-cellular telecommunications and other resource-constrained systems. Four systems are examined. The first three systems are briefly investigated to see if Dynamic Pricing has been adopted or contemplated to be adopted. We then proposed a general architecture for implementing Smart Pricing that could be used by any of the systems. We place particular attention on the Cognitive Radio system, which is currently a hot research topic. We investigate the system in details and proposed a detailed and specific plan for implementing Smart Pricing.

6.1 Electricity systems

Dynamic Pricing has won acceptance in electricity utilities in many countries and has been successfully tested. There have been different types of Dynamic Pricing applied in electricity. In one application there were three types of peak days and six rates but only two were used in any particular day [118]. Rates were fixed and known to consumers in advance. Peak days were notified a day in advance. Rate changes were at different periods which could be every hour [118] or every few hours [119]. Explicit messages notifying either a type of peak day [118] or when a peak price becomes active [119] [120] were sent to consumers. Readings of consumers' consumption were taken hourly and remotely through the power lines. Advanced metering devices which allowed the electricity system operator to communicate with the customers and to measure the customers' electricity usage were required [121]. Data communications and mains signalling by means of low voltage power lines has been investigated with some success. However, there were also several disadvantages with this method. Communication systems, data management and Billing Systems have still been the barriers to the implementation of Dynamic Pricing [122].

Smart Pricing is a rather more sophisticated Dynamic Pricing scheme and therefore can also be adopted in electricity systems in the same way as a general Dynamic Pricing scheme. As seen in Chapters 3 and 4, with Smart Pricing, multiple types of rates are sent frequently, in the order of seconds, and to different groups of users. Usage of system resources are also measured even faster than a second interval. Under the Smart Pricing signalling models, not only is the signalling from the electricity system operator to the users (i.e. the downlink signalling) addressed, but also the uplink and inter-network elements signalling. In addition, network elements responsible for billing and billing-related signalling are also dealt with under Smart Pricing. Furthermore, users also have the option to bid to obtain the required amount of resource they need if they are willing to pay higher prices than other bidders. Many of these features of the Smart Pricing model are not available in the Dynamic Pricing schemes that were proposed for electricity systems.

6.2 ATM systems

Dynamic Pricing has been proposed for use in converged fixed line telecommunications protocols such as Asynchronous Transfer Mode (ATM) in combination with or instead of congestion control mechanisms. There are three types of packet streams: guaranteed, packet-oriented best effort and stream-oriented best effort. Packet streams are either in the transmitting or not-transmitting states. In [123], two types of Dynamic Pricing schemes were examined: slow-reacting and fast-reacting pricing. For slow-reacting pricing, price is not changed when streams alternate between transmitting states. For fast-reacting pricing, a change in the arrival rate of a stream triggers a price change for that stream. An *Intelligent Agent* is used to send congestion control or pricing messages back to the users. With packet-oriented, best effort streams, users must specify the value of each packet and only transmit packets when the value of the packets is higher than the current price. Congestion control mechanisms in accordance with Dynamic Pricing was an option to control users' transmission rates [123][124].

The Smart Pricing model can address all of the above implementation issues. The model can handle all kinds of traffic and prices as each traffic packet can be calculated at the TSS based on any set of system parameters. The notion of value of each packet in ATM is equivalently captured by the WTP in the Smart Pricing model. The DPE is designed to receive congestion information, and in association with the TSS set prices in real-time based on the congestion information, then send those prices to the users.

The DPE has control and only allows users to start consuming system resources upon their acceptance of the current price. The DPE is also linked to the network resource controller and therefore it is relatively easy to control users' transmission rates. Thus, it can accommodate fast-reacting pricing as well as slow-reacting pricing schemes since slow-reacting pricing is just a relaxed version of fast-reacting pricing. In addition, the DPE is equipped with a bidding function which allows users to maintain their maximum transmission rates if they are willing to pay more.

6.3 Water systems

No literature on Dynamic Pricing in water systems has been found. The three pricing schemes that are currently used in the water systems are the uniform, decreasing and increasing block rates pricing schemes [125]. We believe the Smart Pricing model could be applied to water systems for two reasons. Firstly, Smart Pricing could be used as a usage management tool to conserve the water resource for countries such as Australia that do not have much water from rivers, a dry summer season and little rainfall. It was found in electricity systems that using Dynamic Pricing, the reduction in consumption between peak periods of peak days and peak periods of off–peak days could be more than one-third [118]. Thus, Smart Pricing could help water authorities to reduce the depletion of water resources by giving negative incentives to users during drought periods. Secondly, with the Smart Pricing model, there are only two new network elements adding to the current water system models: one at the consumers' side and the other at the water system operator's side. During the initial phase of Smart Pricing deployment, these network elements are for communication, reporting users' consumption levels and sending new prices. The bidding options may be used when the system is more mature. New metering devices will be needed. As for signalling networks, the fixed telephone system or emerging consumer broadband can be used.

6.4 Proposed system architecture

Based on the Smart Pricing system diagram for mobile telecommunications systems shown in Section 3.3, we propose a general system architecture for implementing Smart Pricing in the electricity, ATM and water systems in Figure 6.1. The proposed DPE system element in our approach is a more sophisticated version of the *National Dispatching Centre* in the electricity systems [118] and the *Intelligent Agent* in the ATM systems [124]. The DPE interacts with the users to ensure the current price is agreed upon before any consumption of network resources. It also has an optional *Auction Room* feature which allows the network operator to establish prices based on users' bids

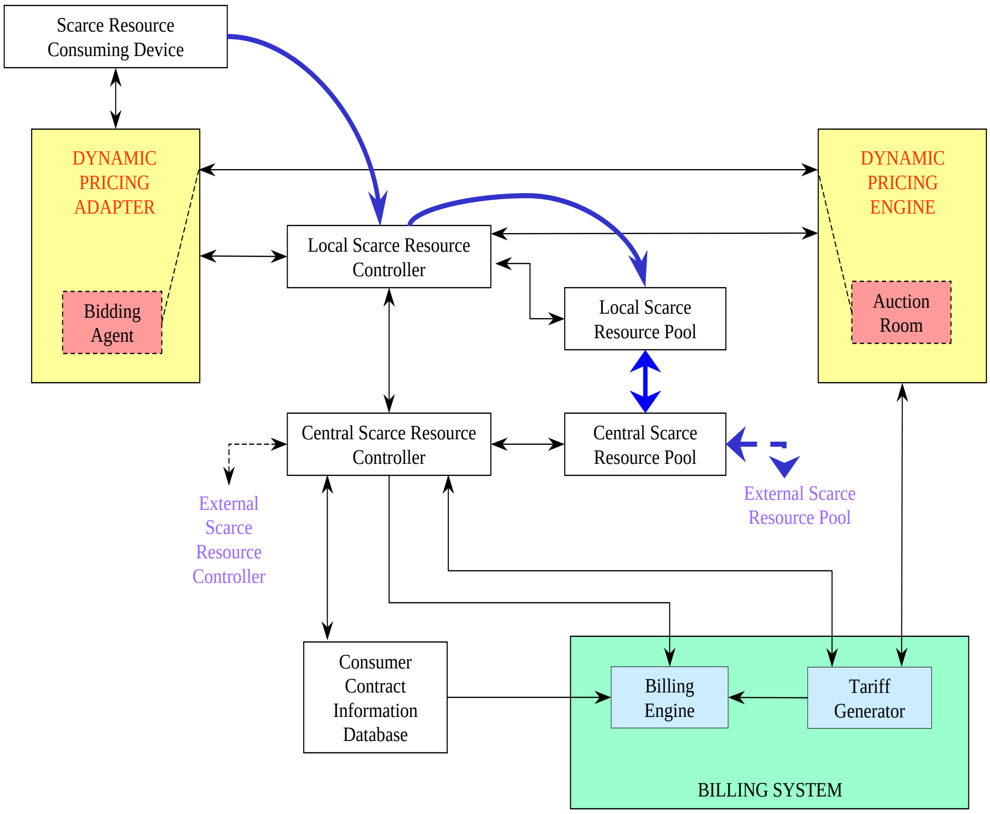


Figure 6.1: Proposed Smart Pricing implementation model for other systems.

for maintaining or increasing their consumption rate. The *Dynamic Pricing Adapter* (DPA) is a new designated network element to help users interact with the network operator for Smart Pricing purposes. The *Bidding Agent* in the DPA is to be coupled with the DPE and facilitate the bidding option for the users. While the DPE and DPA are, to an extent, somewhat captured in Dynamic Pricing in the Electricity and ATM systems, they must be added in the Water systems for Smart Pricing to be implemented. Note that we include in the proposed model the *Consumer Contract Information Database* system element. This is because for ATM systems to be able to determine the price for a connection the user's contract information may be needed [126]. Note also that with the auctioning feature of Smart Pricing, basic tasks and parameters of a bidding system mentioned in [127] will be employed.

6.5 Cognitive radio systems

6.5.1 Motivation

Smart Pricing is a proposed solution to the under-utilised network resources problem. Such a problem arises when more infrastructure is invested only to accommodate peak demand. In the previous chapters, network capacity is in bits/s and dictated by the amount of spectrum a network operator is assigned under its licence. For the Release 99, HSDPA and HSPA+ systems, the amount of bandwidth needed for operation is fixed at 5 MHz, whereas for LTE, it varies between 1.4 and 20 MHz. Understandably, the larger the available bandwidth is, the greater the network capacity. However, spectrum is a scarce resource, managed by government agencies and shared between many services. There are rules which specify what frequency bands can be allocated for a certain service. There are also limits on the channel bandwidths that can be assigned for a service. Spectrum is said to be one of the most tightly regulated resources of all time [128].

A form of spectrum assignment is through issuing spectrum licenses. Such a licence authorises a licensee to use a particular frequency band within a particular geographical area for a fixed period (e.g. 15 years). For some frequency bands, the demand exceeds supply and in those situations, spectrum licenses are generally offered at auction [129]. Hence, it could be very costly to obtain a spectrum licence. However, not all allocated spectrum is effectively utilised, for instance in the 50-950 MHz band as shown in [130] or the 2-6 GHz band in [131] where many parts of frequencies are not used at all or are very lightly used in certain periods of a day. Contradictively, some under-utilised spectrum are eagerly sought after by other users. An example of this can be found at the 2.5 GHz band in [132] whereby the mobile telecommunication industry supports the government's plan to make frequencies in part of the band available for their use. The support is a result of a foreseeable significant spectrum deficit to meet the industry forecasted demand which is growing substantially every year. Even if certain under-utilised spectrum is not sought after by a strong industry like mobile telecommunications, it still makes sense for the spectrum to be shared, at least with opportunistic users. Allowing that to happen will see spectrum efficiency improve, which is consistent with principles of scare resource management.

As such, spectrum sharing and dynamic spectrum allocation are clearly options for addressing the problem of under-utilised spectrum and for improving spectrum efficiency. Sharing of spectrum is also an option for addressing shortages of frequency allocations for some services. A table of frequency allocations for all services can be found in [133]. Sharing a frequency band, instead of assigning to one exclusive licensee, will allow access by multiple licensees using techniques like time division, separation of geographical area or orthogonal codes. Inherently this problem is similar to the problem of under-utilised network resources during quiet periods in mobile telecommunications that we have addressed in previous chapters. Therefore, we see a need to extend Smart Pricing and our proposed signalling models, the MCS and SSA signalling models, to resolve the under-utilised spectrum problem too.

Spectrum bands not in use are called *white spaces* and could be exploited by the Cognitive Radio (CR) approach [130]. In other words, CR improves spectrum efficiency [134]. As CR is said to be a particular extension of software radio [135], we start with an outline of software radio.

6.5.2 Software defined radio and software radio

A receive Software Defined Radio (SDR) is a radio in which the digitization is performed at some stage downstream from the antenna, for instance after the Radio Frequency (RF) section. Conversely, a transmit SDR is one with a reverse process for transmit digitization. Digital signal processing in flexible and reconfigurable functional blocks defines the characteristics of the radio. With SDR, a handset is transformed from a single-frequency (e.g. 900 MHz) single-mode (e.g. TDMA) to a multiband (e.g. 900, 1800 and 2100 MHz) multimode (e.g. TDMA and WCDMA) handset [136]. A detailed treatment of SDR technology and system considerations can be found in [137]. A SDR is said to be a practical version of a Software Radio [138].

A Software Radio (SR) is a SDR in which the software control processing engine is placed at the antenna and all the processing required for the radio is performed by software residing in high-speed digital signalling processing elements [136]. As software can be updated and modified without much effort, the flexibility of SRs is extensive. In turn, components of a SR can be reconfigured. SRs are seen as an essential component of Fourth-Generation (4G) mobile communication systems [139]. They are multi-band radios capable of supporting multiple interfaces and protocols [135].

6.5.3 Characteristics of cognitive radio

CR [135] is a radio which autonomously observes the radio environment, infers context, assess alternatives, generates plans, supervises multimedia services, allocates computational and radio resources to conventional radio software, initiates tasks and learns from its mistakes. Those actions reflect in the CR's 6-element cognitive cycle: observe, orient, plan, decide, act and learn.

CR enables a frequency band to be shared by being able to sense and avoid parts of the band that are in use to avoid interference, then use the other parts. A condition to share the spectrum is that *secondary users*, that is those who are allowed to use the band (or parts of it) in fixed periods, dynamically or opportunistically, must not cause harmful interference to the *primary users*, that is those who own the license. In addition, secondary users must not interfere with each other beyond certain thresholds. Primary users do not need to use a CR. Secondary users, on the other hand, must use CRs. For simplicity, hereafter, we call secondary users who use CRs simply CRs.

Where primary users of a frequency band do not constantly use their licensed spectrum, they can share with CRs with or without charging them a fee. Spectrum available for sharing can be from a singular licensed spectrum, or from a combination of a licensed spectrum. The latter is termed *spectrum pooling* in [135].

By nature, in order to operate, a CR needs information about the radio environment it is in. That information can be acquired autonomously by the CR through sensing or it could be provided to the CR through pre-installed software or updates over the air interface.

When issued, a spectrum licence may be associated with a rich set of conditions that need to be complied with. These conditions include limits on the license frequency band, latitude and longitude coordinates, radiated power and spurious emission limits. A full list of such conditions can be found in [140]. If licensed frequencies are shared, those conditions must be informed to the CRs and the CRs must comply with the conditions the same way as the primary users do. In addition, as owners of the bands, primary users may set additional conditions for CRs.

A CR must know its location [141]. As mentioned in Section 6.5.1, a spectrum licence may authorise a licensee to use a particular frequency band only within a particular geographical area. Thus, without knowing its coordinates, CRs will not able to operate. Another reason why a CR needs to know its location is that if it realises it is on top of a hill and has a line-of-sight signal propagation with its receiver, it can use a higher order modulation scheme plus a higher coding rate. That combination will see an increase in the received bit rate.

A CR must have spectrum awareness [141]. It must know what parts of the shared frequency bands are not in use before using them. Of course, many CRs close to each other in a small geographical area may sense the same yet-to-be-occupied parts, so a mechanism has to be put in place to regulate this potential contentions. If two such CRs transmit at the same time on the same frequency, both of their respective receivers most certainly will not be able to successfully receive the data sent because of the interference caused to them by their unintended CR transmitters.

Transmission power of CRs must be controlled [141]. A main factor that determines the magnitude of the interference is the transmit power. If a CR transmits with an unnecessarily high power, not only does it cause co-channel interference but also adjacent channel interference. That means primary users of spectrum that are not shared by the CR will be interfered with as well. The use of directional antennas, small beamwidths and high elevation angles can help reduce CR transmission power.

The following applications are deemed to be beneficial with CR technology [134]: mobile multimedia downloads, emergency communications systems, broadband wireless services and multimedia wireless networking. This is due to their delay tolerance, lower power requirement and/or low mobility.

6.5.4 Challenges that will need to be overcome

For CR to be brought into use, some regulatory, security and technical issues need to be addressed beforehand.

Regulatory

Shared frequency bands for CRs can come from many sources. It could, for example, be bands allocated by an administration for use primarily for Electronic News Gathering (ENG) or Aeronautical Mobile Telemetry (AMT) services. For ENG in Australia, the band is so-called the *2.5 GHz band* with the frequency range of 2.5-2.69 GHz and a Television Outside Broadcast Network (TVOB) licence is required for a broadcaster to use this band. For AMT in Australia, the band is 2.2-2.3 GHz and is used for operations lasting short durations on a limited number of days a year [105]. Clearly, the two radiocommunications services only use allocated frequency bands intermittently. The problem is that the events are not known in advance. If the bands are not shared, it is wasteful; however, if they are shared, CRs must stop transmitting when the ENG and AMT radios are in use, otherwise the interference could have serious impacts. In such cases, times and durations when these services are in use must be informed to the CRs with short notices in the order of possibly seconds. Therefore, the *policy module* [130] in the CRs must be developed robustly enough to accommodate such rapid regulatory changes.

If the shared spectrum is a frequency band pooled from with different licensees in adjacent jurisdictions or countries, the *policy module* must also adapt quickly enough when CRs change geographical areas, particularly when movements are with high speeds.

In order for frequency bands to be shared, governments may need to change their licensing regimes. The policy that spectrum licences are issued with rights for exclusive use is no longer appropriate and needs to be changed. In 2004, the U.S. Federal Communications Commission (FCC) made secondary markets for spectrum legal, which allows a licensee to lease rights to use spectrum for the duration of the license [142].

Other countries are also in the process of making similar changes to their spectrum regulatory regime. One of such countries is Australia. This reflects in [143] in the RadComms 2010 conference.

Spectrum for sharing can be *spectrum commons* or with *property rights*. QoS cannot be guaranteed under spectrum commons approach but is specified under a property rights approach. For shared spectrum with property rights, primary users can dynamically allocate small portions of the band to CRs or only allow CRs to access it opportunistically [142]. In Australia, spectrum commons are a limited set of common frequencies which people can use if they comply with the conditions specified by the Government, specifically the Australian Communication and Media Authority, in the relevant *class licences* free of charge [144]. Radios operating under a class licence are not protected from interference from their peers who share the frequency band with them. An example of a class licence can be found in [145]. On the other hand, spectrum with property rights are equivalent to spectrum authorised for use under *spectrum licences* in Australia. Holder of spectrum licences are required to register their devices which enable the devices to be protected from interference [146]. If a spectrum licence holder shares its band with CRs but the licence holder does not register the CRs, the CRs will not be protected from interference. This is against the spectrum licence policy. Therefore, some form of protection must be developed to protect the CRs. A result of this may require that the Radiocommunications Act, as in [147], be modified. This is a big challenge and will require tremendous effort.

Security

Among the others, CR are subject to the following security issues [130]:

- malicious abuse which can cause widespread denial-of-service. When sharing spectrum with CRs, primary users rely on CRs not causing harmful interference to them when they use the spectrum. CR should also not cause excessive harmful interference to their peers. Not having the interference issue enforced could render the shared spectrum unusable because data communications will not be successful. If controlling CRs' behaviour is via software, then modification made to the software incorrectly could be the source of the abuse. This could be due to a genuine error or harmful intention to corrupt the system. To ensure the integrity and origin of reconfiguration software, a signed content technique could be used [148]; and

- unauthorised software or policy update downloads. As discussed in the previous section, it may be necessary that policy updates and software downloads occur over the air interface. Without a reliable authentication system, this could open the door for identify theft. Consequences are that some CRs could be incorrectly charged (which is a source of dispute) or prevented from continuing to access the spectrum.

Technical

If primary users cannot be protected from interfering power from CRs, licensed spectrum may not be shared. To meet this requirement, CRs need to be able to detect what channels are used by primary users and not to use those. To detect unused channels, certain functions of a spectrum analyser should be made available in CRs. Cyclic scanning or filter bank are proposed in [141]. Once detected, information about unused channels can then be stored in a spectrum database, which could be centralised or de-centralised and may require updating [149], for access by CRs. As for primary users, to ensure that they can access the shared spectrum when and where they need, a *Polite Backoff* protocol can be used [135].

A well known problem with CR is the *hidden node* problem. It occurs when CRs cannot detect primary users' transmissions and transmit. This interferes with primary users' transmissions, which may cause disruption because SINRs at the intended receivers of those primary users' transmissions is lower than the sensitivity. Two possibilities for such a problem to happen are: distance between the primary users and CRs are great and primary users' transmissions are blocked by terrain [130]. The hidden node problem can be made less severe by improving the sensitivity of CR monitoring instead of just deploying more CR nodes [134] [149].

Before CRs with all functions outlined in Section 6.5.3 are realised, development of technology centric CRs that can monitor their spectral environments and locate their geographical positions should be the first step [138].

6.5.5 Capacity of cognitive radio systems

Capacity of a CR system depends on the size of the shared spectrum. The spectrum can come from a frequency band of only one licensee or multiple frequency bands of multiple licensees. The latter can be call a spectrum pool. Frequency bands in the spectrum pool do not have to be adjacent to each other. OFDM is found to be the best physical layer for CR system because it allows the use of discontinuous and arbitrary-sized frequency bands [150].

Apart from the size of the spectrum pool, how significant interfering power in a wireless channel and whether CRs have knowledge of the primary users' transmissions also affect the capacity of a CR system. Under weak interfering power and non-causal knowledge of the primary users' transmissions, capacity is found to be higher than when spectrum is time-sharing [151]. Capacity under strong interference powers are discussed in [152] and [153], as pointed out in [151].

6.5.6 Applicability of Smart Pricing in cognitive radio systems

In this first attempt to show that Smart Pricing and the MCS and SSA models can be applied to CR systems, we limit ourselves to consider CRs which are yet to have the full potentials as defined in Section 6.5.3. These CRs can still autonomously observe the radio environment, infer context and assess alternatives. They can then propose plans to use, but must not yet use, the radio resources. The plans are signalled to the *CR system operator*. Upon receipt of these plans, the CR system operator can either approve, adjust or reject the plans after taking into consideration the instantaneous policy at that moment and/or any strategic decisions. Resting such final decisions on the CR system operator is believed to only enhance the spectrum efficiency due to the following reasons:

i. the CR system operator has an overall view of the radio resource usage and the number of spectrum users in the system. Further, if scheduling is used, no doubt efficiency will increase;

ii. in unexpected events (e.g. CRs unwillingly ceasing contention for a certain channel), intervention is necessary; and

iii. CRs will have to make less intensive decisions which may reduce the amount of information needed by CRs. This could potentially help reduce CRs' power consumption.

The idea behind the above choice is to keep the admission and resource allocation control at the CR system operator. As Smart Pricing aims to reserve certain capacity for high WTP users as well as to give guaranteed bit rates, even in the event of graceful degradation, a CR system operator needs to have such a control to meet Smart Pricing's objectives. Without such a control power, guaranteed QoS may not be achieved.

The *dynamic spectrum leasing* and *interruptible spectrum leasing* models discussed in [151] and the spectrum with *property rights* model as discussed in [142] are deemed suitable for Smart Pricing. This is because these models facilitate a mechanism in which the CR system operator control the admission to system and the allocation of frequencies and bandwidth within the spectrum pool.

If OFDM and time-sharing techniques are chosen, the maximum capacity a CR system, η_M, can be calculated in a similar manner as with the LTE system. That is shared spectrum is divided into *spectrum units*, each with a nominal value of 15 kHz in the frequency domain and 0.5 ms in the time domain. Then using Equation (3.1), η_T is obtained. The user *spectrum rates*, $R_{SP,max}$ and $R_{SP,min}$, can also be determined based on relevant ratios relative to the η_M specified by the CR system operator.

We propose that:

1. a network element called the Cognitive Radio Gateway (CRW) to host the *policy database* from which the CR's policy module obtains update. The CRW to host the *spectrum database* from which CRs obtained master information about unused frequencies that CRs could use. The CRW is also responsible for making decisions in response to plans to use the spectrum pool proposed by the CRs and on admission and resource allocation. Finally, the CRW is responsible for reporting load of the CR system to the DPE;

2. DPE is now responsible for not only receiving congestion information of the CR system from the CRW, but also for setting prices depending on levels of congestion in the CR system. With the latter, in effect, it means the TSS is now incorporated into the DPE;

3. four channels, 15 kHz each, are set aside from the spectrum pool for the purposes of CR system signalling. We adopt the idea of having these channels from 3GPP mobile telecommunications systems, which has resulted in proven success. The:

    a. CR Random Access Channel (CR-RACH) is an uplink control channel for transmit CRs to signal their intention to communicate with the CRW about their plans using the spectrum pool;

    b. CR Uplink Control Channel (CR-UCCH) is for direct communications negotiating spectrum pool access plan between the transmit CRs and the CRW;

    c. CR Downlink Control Channel (CR-DCCH) is for receive CRs to send feedback of channel quality to the CRW. Like in HSDPA, HSPA+ and LTE, adaptive modulation and coding necessitates this channel. If the CQI is good (i.e. SINR is high) higher modulation and coding schemes are used to improve bit rates. The idea of having a channel for feedback is also flagged in [141]; and

    d. CR Broadcast Channel (CR-BCH) is for the DPE in conjunction with the CRW to notify CRs to take advantage of periods the CR system is under-utilised. Understandably, prices during these periods are significantly low.

The idea of having a broadcast channel is also flagged in [141], however, there the broadcast channel is for use in the case of a non-licensed frequency band, which is different from ours where spectrum pools are from licensed bands.

The effectiveness of having either (or all) of the four channels that we propose here is the focus of our further research;

4. CRs contain the *CR Engine* (CRE) proposed in [154]. The CRE accepts input parameters of three types: transmission, environment and QoS. In addition, the CRs must host the DPA interacting with the DPE for Smart Pricing purposes, which includes bidding, performed by the *Bidding Agent* as discussed in Section 6.4; and

5. the protocols used in the links from CRs and primary users (PUs) to the CRW is CR-Uu, and between the CRW and DPE is SCTP/IP. Details of the former protocol is a subject of our future research.

Taking the above proposed items into account, a proposed Smart Pricing system diagram for CR system is given in Figure 6.2. In the diagram, the CR-Tx and PU-Tx are the transmit CR and transmit PU stations, respectively. The CR-Rx and PU-Rx are the receive CR and receive PU stations, respectively. The PU-Tx and PU-Rx are included only to show the complete CR system, these two stations are not involved in Smart Pricing signalling. The reason is that they are primary users and should have exclusive access to the spectrum pool whenever they need.

As for the signalling components: the following modifications to Section 3.7.2 should be made:

- CR-Tx replacing MS;
- CRW replacing RNC and MSC; and
- setting the numbers of signalling components between the TSS and DPE to 0 as the TSS is now collocated with the DPE.

Referring to Section 5.1, above we have gathered enough of the required five pieces of information so that Smart Pricing and our proposed Smart Pricing signalling models can be applied to the CR system. They can be used in their respective places in the MCS and SSA models in Chapters 3 and 4. Thus, we affirm that using our proposed Smart Pricing signalling models Smart Pricing can be applied to CR system.

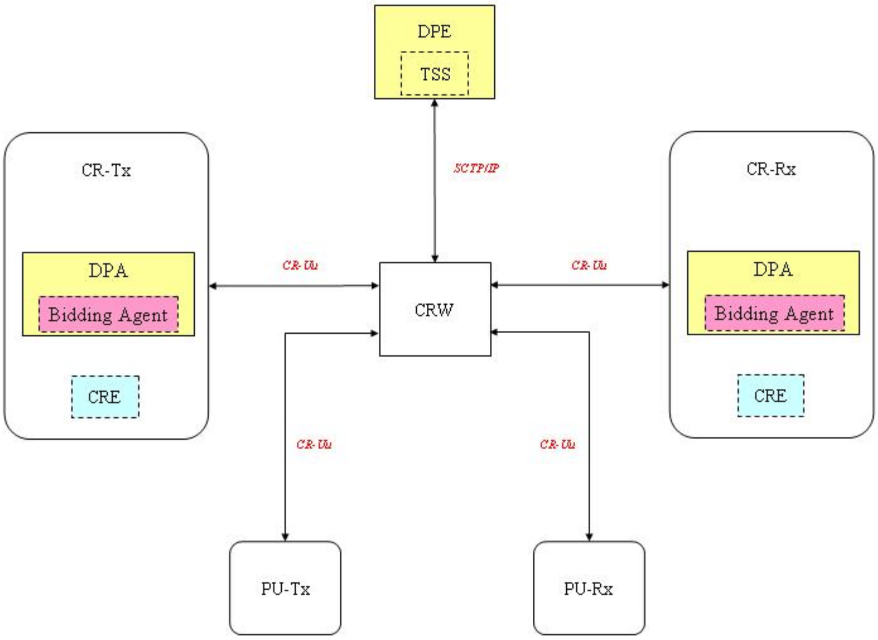


Figure 6.2: Smart Pricing in CR System Diagram.

6.6 Findings and discussions

In this chapter, we have investigated the applicability of Smart Pricing and the MCS and SSA models to non-cellular telecommunications and other resource-constrained systems: the electricity, ATM, water and CR systems. Our intention is to show the robustness of the models.

In summary, we have:

1. briefly investigated the electricity and ATM systems and found indications that Dynamic Pricing schemes have been proposed for use in these systems. There are some similarities between the proposed Dynamic Pricing schemes and our Smart Pricing scheme. However, the Smart Pricing scheme is more sophisticated because of its bidding function. We believe that Smart Pricing can not only be applied in the two systems but also adds more features into the existing proposed Dynamic Pricing schemes;

2. briefly investigated water systems and found no indication that Dynamic Pricing has been proposed for use. We believe that Smart Pricing can be applied in water systems as a tool to regulate the use of water resource in countries such as Australia where water is limited and which frequently suffers long drought periods. New metering devices need to be developed before Smart Pricing can be

deployed. Feasibility study of using the fixed telephone or TCP/IP as a signalling network also needs to be conducted;

3. proposed a general system architecture for implementing Smart Pricing in electricity, ATM and water systems. The main change to the conventional system architectures is the addition of the DPE and DPA network elements which are responsible for setting prices and allocating resources dependent on the system load and bids from users; and

4. conducted detailed investigation on CR systems and affirm that Smart Pricing and the MCS and SSA models can be applied to CR systems. These models are the potential candidate for facilitating the dynamic spectrum allocation in spectrum-licensed frequency bands. We propose a specific system architecture for implementing Smart Pricing in CR systems, in which:

    i. the CRW is the main system element controlling the CR system;

    ii. the DPE is responsible for receiving congestion information sent from the CRW and for setting prices and allocating spectrum depending on levels of congestion in the CR system. The TSS is incorporated into the DPE;

    iii. the CRs have the DPA function incorporated into it to facilitate Smart Pricing functions;

    iv. four new channels are to be created for use with CR related signalling; these are the CR-RACH, CR-UCCH, CR-DCCH and CR-BCH; and

    v. protocols used for the links are proposed to be CR-Uu and SCTP/IP.

CHAPTER 7

Conclusion

7.1 Summary of findings and contributions

Smart Pricing is a proposed solution to the problem of under-utilized network resources and to accommodate growing demand within available network resource. As such, Smart Pricing is an alternative approach to investing more infrastructure to handle only peak demand. It is a pricing scheme that responds to the state of the network and a congestion pricing scheme that has price varying directly with the load. Smart Pricing varies price and load relationship according to the characteristics of the current user group. The process of setting prices involves communications between network elements and once set, new price messages are sent to users. In the event of QoS graceful degradation, users who are outside their t_{gQ}, can bid to maintain their admission QoS. These indicate that three types of signalling are required; those are: signalling on the UL, DL and Netw2Netw.

For Smart Pricing to be deployed, the DPE, a new network element, is proposed to be added to the current mobile telecommunications system architecture. The DPE is responsible for receiving congestion information, processing it and then requesting the Billing System set new prices if necessary. Based on the system diagram with the DPE added, two signalling models, the MCS and SSA models are developed. The models are examined thoroughly by investigating the impact on the required signalling due to factors such as: size of the system, conditions a user is admitted into the network, the way network congestion is defined and user behaviour.

The MCS model is one in which operation of Smart Pricing is simulated. The model enables the required signalling loads when Smart Pricing system is in its instantaneous conditions to be measured. Both small and large Smart Pricing systems are investigated and eighteen operation scenarios are simulated. It is found that when there are more types of user's WTP in the system: the average signalling loads increase, bidding signalling percentages increase, and cost un-recoverable signalling percentages decrease. On the other hand, when there are more users in the system, the first and the last are the same, except the second. In such a case, the UL bidding signalling percentage increases but DL and Netw2Netw decrease. It is also found that it is not how the level of congestion is defined by a network operator, it is the user behaviour that dictates the amount of signalling resources needed to be set aside for Smart Pricing. More findings from this model can be found in Section 3.9.

The SSA model is developed using the State Space and Markov Chain technique. Comparison of results from this model with those from the MCS model is planned and thus most of the simulation scenarios conducted for the MCS model are replicated. Comparison reveals that almost 85% of the time, results from the SSA model deviate by less than 20% from the MSC model. Compensating for that low deviation is a substantial improvement in simulation time. The average simulation time for the SSA model is more than 700 times faster than the MCS model. It is found that, in the stead-state condition, the maximum average signalling loads for the UL, DL and Netw2Netw are: 4.04, 4.06 and 72.71 kbps, respectively. It is recommended that the low load threshold is set at 70% of η_M and the t_{gQ} at 90 seconds. More findings and recommendations from this model can be found in Section 4.12.

To provide mobile telecommunications network operators with a complete solution for implementing Smart Pricing in the 3GPP Release 99 UMTS system, applicability of the MCS and SSA models to the downlink direction is shown. How the two models can further be applied to the HSDPA, HSPA+ and LTE systems is also demonstrated. It is ascertained that adopting Smart Pricing will not impose significant signalling traffic on a network with our estimated signalling loads as supporting evidence. These figures can be found in Section 5.6. Nevertheless, informing precise figures by conducting detailed simulations is intended in the immediate future.

Finally, a further step is taken and Smart Pricing is shown to be able to also be implemented in the electricity, ATM, water and CR systems. Results from a detailed investigation into CR systems allows affirmation that the MCS and SSA models can be applied to the CR system to facilitate dynamic spectrum allocation in spectrum licensed frequency bands. Two system architectures for implementing Smart Pricing in those four systems are proposed: a general one for all systems and another one specifically tailored for the CR systems. Details of the system architectures can be found in Section 6.6. Some protocols used for the links in the architecture for the CR systems are subject to further research.

7.2 Possible future work

In the above section, some possible work for this research in the immediate and near future have been identified. Other possible future work are as follows.

Firstly, extending the MCS and SSA models to a multi-cell system which includes intra-network and inter-network handovers.

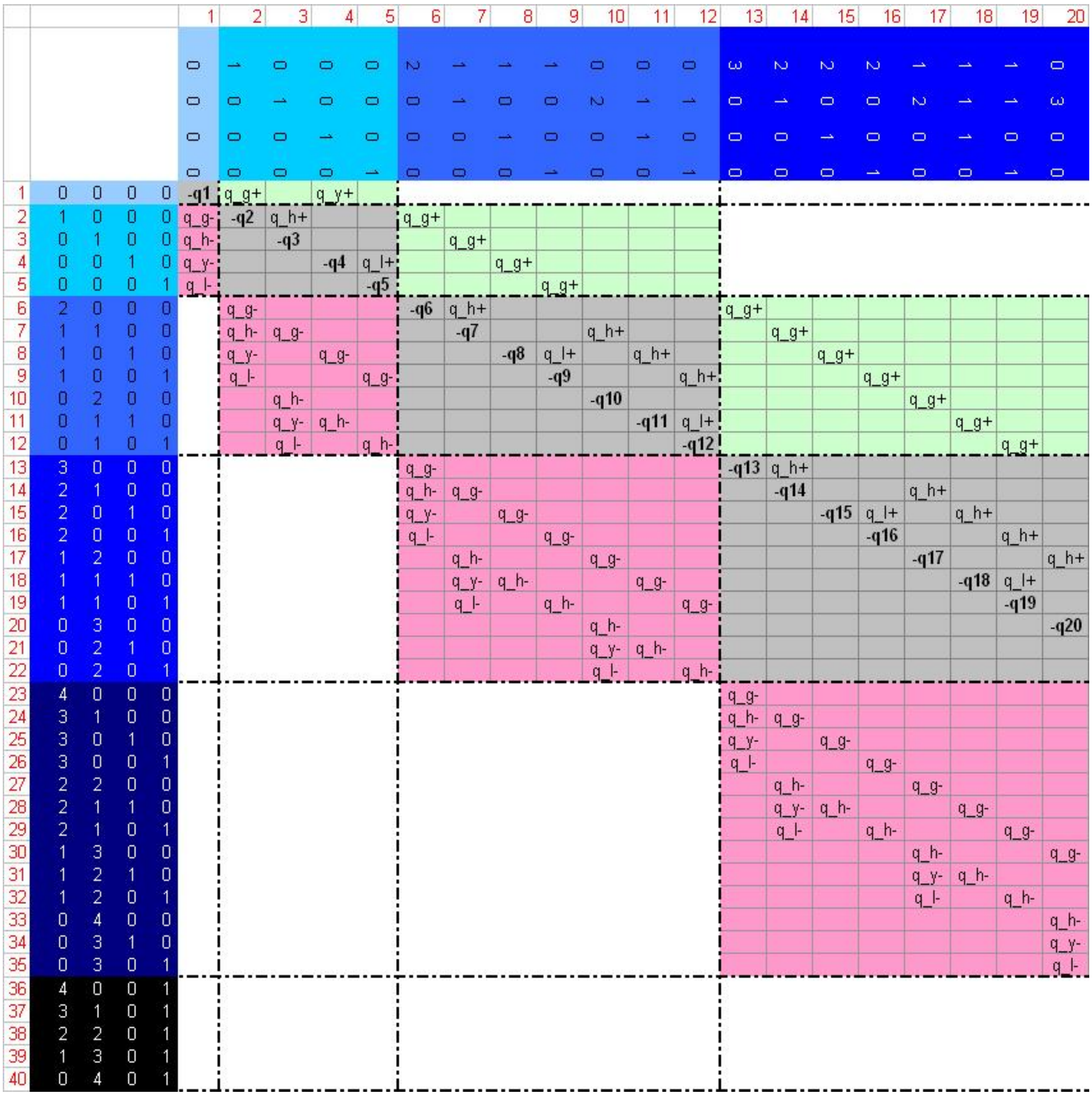
Secondly, for modelling of Smart Pricing in CR systems, investigating the effectiveness of having the CR-RACH, CR-UCCH, CR-DCCH and CR-BCH channels, and developing the CR-Uu protocol.

Thirdly, using *Control Theory* to develop means to cap the required signalling loads to planned signalling resource for Smart Pricing. This is to address cases where arrival rate increases unexpectedly high, as can be seen from the results in Chapters 3 and 4, that the average signalling loads have a linear relationship with the arrival rate. Increases in such a fashion need to be controlled to prevent signalling network overload.

Fourthly, implementing optimal signalling system by using uplink and downlink signalling means identified in Section 2.8. An optimal signalling system is one that is able to deliver new price messages to MSs and transfer the responses back to the DPE within desired time constraints. For notifying new price messages to the users, a simple and easily visible or audible display method on a MS is required. To help hide the complexity of the system, means to allow subscribers to interact automatically with Smart Pricing systems is needed. To help de-emphasise the consciousness of price in subscribers' minds, it is essential that prices are not expressed in monetary units but by some other means such as tokens. The effectiveness of this concept is proven in casinos. To devise a signalling method metric for helping network operators to choose a suitable Smart Pricing signalling method for its network, a metric which is capable of ranking criteria of different signalling methods is required. Lastly, to help determine the practicality of Smart Pricing, a method which is capable of calculating the cost for the implementation of Smart Pricing is necessary.

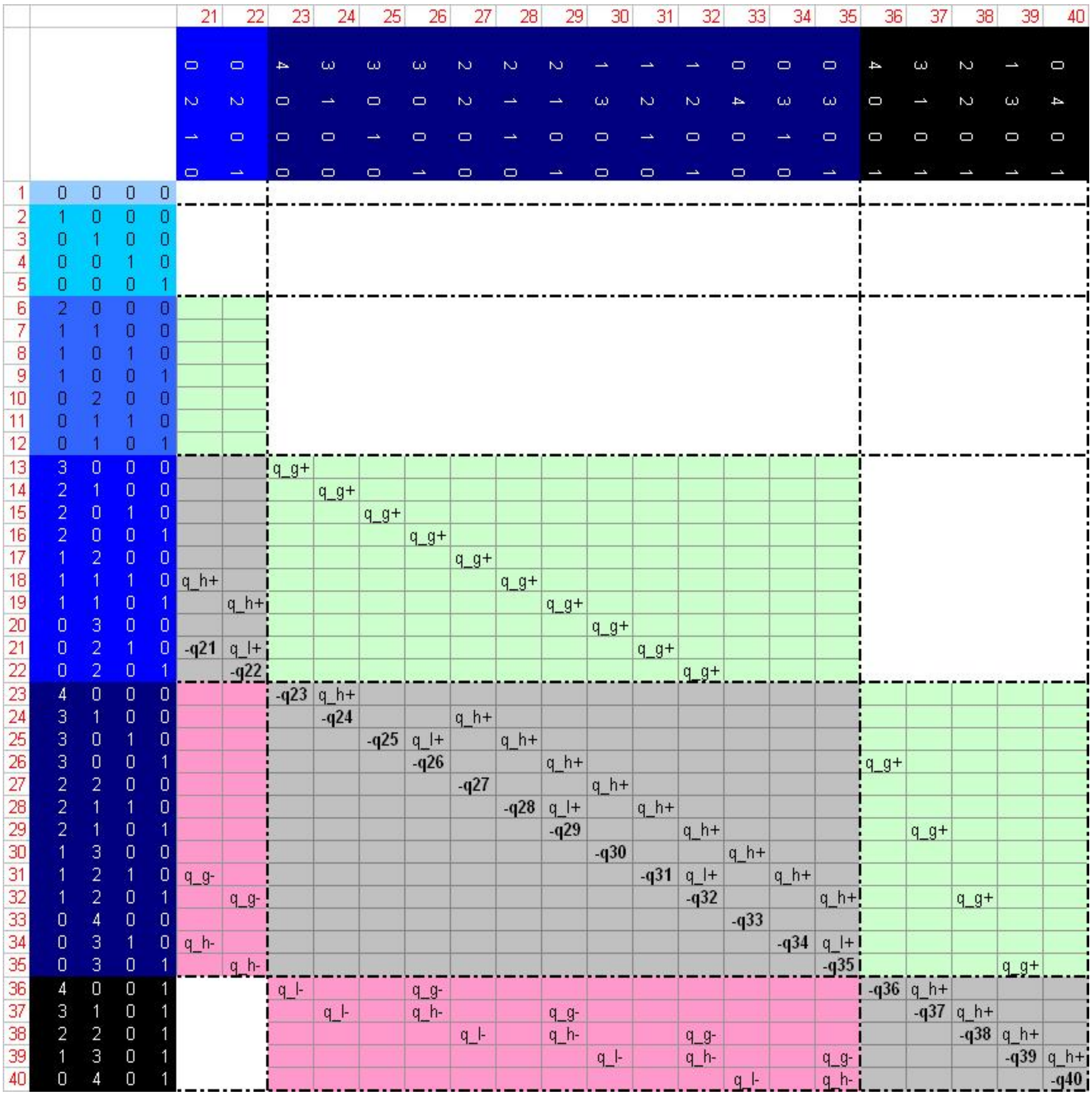
Finally, extending Smart Pricing to include an additional parameter called *Willingness To Bid* (WTB). Different to the WTP, which is the price level that a user prepares pay for his/her call prior to being admitted to the network, WTB is an offset from the WTP. Such an offset may assume a positive or negative value. The sum of a user's WTP and WTB will be the price level that a user will bid up to when he/she experiences QoS graceful degradation. The concept of WTB is introduced to capture the difference in the level of importance of a call in a subscriber's mind before and just after the call is made.

Appendix - Transition diagram



					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
					0 0 0 0	1 0 0 0	0 1 0 0	0 0 1 0	0 0 0 1	2 0 0 0	1 1 0 0	1 0 1 0	1 0 0 1	0 2 0 0	0 1 1 0	0 1 0 1	3 0 0 0	2 1 0 0	2 0 1 0	2 0 0 1	1 2 0 0	1 1 1 0	1 1 0 1	0 3 0 0
1	0	0	0	0	-q1	q_g+		q_y+																
2	1	0	0	0	q_g-	-q2	q_h+			q_g+														
3	0	1	0	0	q_h-		-q3				q_g+													
4	0	0	1	0	q_y-			-q4	q_l+			q_g+												
5	0	0	0	1	q_l-				-q5				q_g+											
6	2	0	0	0		q_g-				-q6	q_h+						q_g+							
7	1	1	0	0		q_h-	q_g-				-q7			q_h+				q_g+						
8	1	0	1	0		q_y-		q_g-				-q8	q_l+		q_h+				q_g+					
9	1	0	0	1		q_l-			q_g-				-q9			q_h+				q_g+				
10	0	2	0	0			q_h-							-q10							q_g+			
11	0	1	1	0			q_y-	q_h-							-q11	q_l+						q_g+		
12	0	1	0	1			q_l-		q_h-							-q12							q_g+	
13	3	0	0	0						q_g-							-q13	q_h+						
14	2	1	0	0						q_h-	q_g-							-q14			q_h+			
15	2	0	1	0						q_y-		q_g-							-q15	q_l+		q_h+		
16	2	0	0	1						q_l-			q_g-							-q16			q_h+	
17	1	2	0	0							q_h-			q_g-							-q17			q_h+
18	1	1	1	0							q_y-	q_h-			q_g-							-q18	q_l+	
19	1	1	0	1							q_l-		q_h-			q_g-							-q19	
20	0	3	0	0										q_h-										-q20
21	0	2	1	0										q_y-	q_h-									
22	0	2	0	1										q_l-		q_h-								
23	4	0	0	0													q_g-							
24	3	1	0	0													q_h-	q_g-						
25	3	0	1	0													q_y-		q_g-					
26	3	0	0	1													q_l-			q_g-				
27	2	2	0	0														q_h-			q_g-			
28	2	1	1	0														q_y-	q_h-			q_g-		
29	2	1	0	1														q_l-		q_h-			q_g-	
30	1	3	0	0																	q_h-			q_g-
31	1	2	1	0																	q_y-	q_h-		
32	1	2	0	1																	q_l-		q_h-	
33	0	4	0	0																				q_h-
34	0	3	1	0																				q_y-
35	0	3	0	1																				q_l-
36	4	0	0	1																				
37	3	1	0	1																				
38	2	2	0	1																				
39	1	3	0	1																				
40	0	4	0	1																				

Figure 1: *Q-matrix* for a system with $\eta_T^{R_H} = 1$, $\eta_M^{R_H} = 4$, and $M_{max} = 5$ (part 1).



					21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
					0 2 1 0	0 2 0 1	4 0 0 0	3 1 0 0	3 0 1 0	3 0 0 1	2 2 0 0	2 1 1 0	2 1 0 1	1 3 0 0	1 2 1 0	1 2 0 1	0 4 0 0	0 3 1 0	0 3 0 1	4 0 0 1	3 1 0 1	2 2 0 1	1 3 0 1	0 4 0 1
1	0	0	0	0																				
2	1	0	0	0																				
3	0	1	0	0																				
4	0	0	1	0																				
5	0	0	0	1																				
6	2	0	0	0																				
7	1	1	0	0																				
8	1	0	1	0																				
9	1	0	0	1																				
10	0	2	0	0																				
11	0	1	1	0																				
12	0	1	0	1																				
13	3	0	0	0			q_g+																	
14	2	1	0	0				q_g+																
15	2	0	1	0					q_g+															
16	2	0	0	1						q_g+														
17	1	2	0	0							q_g+													
18	1	1	1	0	q_h+							q_g+												
19	1	1	0	1		q_h+							q_g+											
20	0	3	0	0										q_g+										
21	0	2	1	0	-q21	q_l+									q_g+									
22	0	2	0	1		-q22										q_g+								
23	4	0	0	0			-q23	q_h+																
24	3	1	0	0				-q24			q_h+													
25	3	0	1	0					-q25	q_l+		q_h+												
26	3	0	0	1						-q26			q_h+							q_g+				
27	2	2	0	0							-q27			q_h+										
28	2	1	1	0								-q28	q_l+		q_h+									
29	2	1	0	1									-q29			q_h+					q_g+			
30	1	3	0	0										-q30			q_h+							
31	1	2	1	0	q_g-										-q31	q_l+		q_h+						
32	1	2	0	1		q_g-										-q32			q_h+			q_g+		
33	0	4	0	0													-q33							
34	0	3	1	0	q_h-													-q34	q_l+					
35	0	3	0	1		q_h-													-q35				q_g+	
36	4	0	0	1			q_l-			q_g-										-q36	q_h+			
37	3	1	0	1				q_l-		q_h-			q_g-								-q37	q_h+		
38	2	2	0	1							q_l-		q_h-			q_g-						-q38	q_h+	
39	1	3	0	1										q_l-		q_h-			q_g-				-q39	q_h+
40	0	4	0	1													q_l-		q_h-					-q40

Figure 2: *Q-matrix* for a system with $\eta_T^{R_H} = 1$, $\eta_M^{R_H} = 4$, and $M_{max} = 5$ (part 2).

Bibliography

[1] Ericsson White Paper, "Basic Concepts of WCDMA Radio Access Network", *Ericsson Radio Systems AB*, 2001. Available on October 9, 2003, at http://www.ericsson.com/products/white_papers_pdf/e207_whitepaper_ny_k1.pdf

[2] Centre for Internet Research (University of Adelaide), "Smart Pricing for Advanced Wireless Networks", SIT CRC 2003 Annual Conference, Sydney, Australia, September 2003.

[3] J. K. MacKie-Mason and H. R. Varian, "Pricing Congestible Network Resources", *IEEE*, Journal on Selected Areas In Communications, vol. 13, No. 7, 1995.

[4] K. Arrow and G. Debreu, "The existence of equilibrium for a competitive economy", Econometrica, vol. 22, pp. 265–290, 1954.

[5] S. L. Hew, "Wireless Optimisation Based on Economic Criteria", PhD thesis, The University of Adelaide, Australia, 2006.

[6] E. D. Fitkov-Norris and A. Khanifar, "Dynamic Pricing in Cellular Networks, A Mobility Model With a Provider-Oriented Approach ", *3G Mobile Communication Technologies, IEEE*, Conference Publication No. 477, pp. 416–420, 2001.

[7] E. D. Fitkov-Norris and A. Khanifar, "Dynamic Pricing in Mobile Communication Systems", *3G Mobile Communication Technologies, IEEE*, Conference Publication No. 471, pp. 416–420, 2000.

[8] E. Fitkov-Norris, "Control Theory Approach for Implementation of Dynamic Pricing in Mobile Cellular Networks", IEE, 2004, pp. 160-164.

[9] R. Piqueras, J. Perez-Romero, O. Sallent and R. Agusti, "Dynamic Pricing For Dencetralised RAT Selection in Heterogeneous Scenarios", 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06), 2006.

[10] P. Liao and A. Wang, "A survey on Dynamic Pricing", Jul. 18, 2004 [Aug. 7, 2010]. Available at: `http://www.cs.ucdavis.edu/~liu/289/Presentation/DynamicpricingFinalversion.doc`

[11] J. MacKie-Manson and H. Varian, "Pricing Congestible Network Resources", IEEE Journal on Selected Areas in Communications, Vol. 13, No. 5, September 1995, pp. 1141-1149.

[12] S. Mandal, D. Saha, and A. Mahanti, "A technique to support dynamic pricing strategy for differentiated cellular mobile services", IEEE Globecom, IEEE, pp. 3388-3392, 2005.

[13] M. Koutsopoulou, A. Kaloxylos, A. Alonistioti, L. Merakos and K. Kawamura, "Charging, Accounting and Billing Management Schemes in Mobile Telecommunication Networks and the Internet", IEEE Communications Surveys & Tutorials, First Quarter 2004, Vol. 6, No. 1, pp. 50-58.

[14] I. C. Paschalidis, and J.N. Tsitsiklis, "Congestion-dependent pricing of network services", IEEE/ACM Transactions on networking, IEEE, Vol. 8, no. 2, pp. 171-184, 2000.

[15] E. Viterbo, and C.F. Chiasserini, "Dynamic pricing in wireless networks", International Symposium on Telecommunications, pp. 385-388, 2001.

[16] L. Zhang and D. Li, "Resource Allocation Models Based on Pricing for MultiService WCDMA Networks", 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication Proceedings, 2003, pp. 1466-1470.

[17] H. Chan, I. Ho and R. Lee, "Design and Implementation of a Mobile Agent-based Auction System", IEEE, 2001, pp. 740-743.

[18] M. Peirce and D. O'Mahony, "Flexible Real-Time Payment Methods for Mobile Communications", IEEE Personal Communications, December 1999, pp. 44-55.

[19] S. Yaipairoj and F. Harmantzis, "A Dynamic Pricing Model for Data Services in GPRS Networks", IEEE Communications Society, Globecom 2004 Workshops, pp. 453-458.

[20] S. Yaipairoj, and F.C. Harmantzis, "Dynamic pricing with 'Alternatives' for mobile networks", IEEE Communications Society, IEEE, pp. 671-676, 2004.

[21] T. Russell, "Signalling System #7". *Mc-GRAW-HILL Telecommunications*, 3rd ed., USA, 2000.

[22] A. S. Tanenbaum, "Computer Networks". *Prentice-Hall*, New Jersey , USA, 3rd ed., 1996.

[23] 3GPP TS 23.002, "Network Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. V5.5.0, 2002.

[24] W. Stallings, "Wireless Communications and Networks". *Prentice Hall*, New Jersey, USA, 2002.

[25] 3GPP TR 21.905, "Vocabulary for 3GPP Specifications". *3GPP Organizational Partners*, Valbounne, France, Rel. V5.2.0, 2001.

[26] ETSI TR 101 748, "Abbreviations and Acronyms". *European Telecommunications Standards Institute*, Sophia Antipolis Cedex, France, Rel. V8.0.0, 2000.

[27] V. K. Garg and J. E. Wilkes, "Principle & Applications of GSM". *Prentice-Hall*, New Jersey , USA, 1999.

[28] M. Mouly and M. Pautet, "The GSM System for Mobile Communications". *Michel Mouly and Marie-Bernadette Pautet*, Palaiseau , France, 1992.

[29] M. Sorell, "Mobile Communication Networks", lecture notes. *Convergent Communications Research Group*, the University of Adelaide, Adelaide, Australia, lecture 5, slide 90, 2003.

[30] International Engineering Consortium On-Line Education, "Universal Mobile Telecommunications System (UMTS) Protocols and Protocol Testing", tutorial material. Available on October 8, 2003 at `http://www.iec.org/online/tutorials/umts/topic03.html`

[31] S. Tafvelin and T Olovsson, "Advanced Internet Technology", course material. *The Department of Computer Engineering*, Chalmers University of Technology, Goteborg, Sweden. Available on October 8, 2003 at `http://www.ce.chalmers.se/undergraduate/IMP/EDA435/lastyear/pictures/F8_4_6perpage.pdf`

[32] H. Holma and A. Toskala, "WCDMA For UMTS". *John Wiley & Sons*, 2nd ed., Chichester, England, 2002.

[33] J. Korhonen, "Introduction to 3G Mobile Communications". *Artech House*, 2nd ed., Boston, London, 2003.

[34] 3GPP TS 32.200, "Charging Principles". *3GPP Organizational Partners*, Valbounne, France, Rel. V4.0.0, 2001.

[35] 3GPP TS 32.015, "Charging and billing; 3G call and event data for the Packet Switched (PS) domain". *3GPP Organizational Partners*, Valbounne, France, Rel. V3.8.0, 2001.

[36] 3GPP TS 32.005, "Charging and billing; 3G call and event data for the Circuit Switched (CS) domain". *3GPP Organizational Partners*, Valbounne, France, Rel. V3.5.0, 2001.

[37] 3GPP TS 22.038, "USIM/SIM Application Toolkit (USAT/SAT); Service Description". *3GPP Organizational Partners*, Valbounne, France, Rel. V5.2.0, 2001.

[38] ETSI TS 100 906, "Mobile Stations (MS) features". *European Telecommunications Standards Institute (ETSI)*, Sophia Antipolis Cedex, France, Rel. V7.1.0, 2000.

[39] Y. Lin, H. Rao and M. Chang, "Mobile Prepaid Phone Services". *IEEE-Personal-Communications*, Vol. 7, No. 3, pp. 6–14, 2000.

[40] ETS 300 524, "Signalling requirements relating to routing of calls to mobile subscribers". *European Telecommunications Standards Institute (ETSI)*, Sophia Antipolis Cedex, France, 1994.

[41] 3GPP TS 25.301, "Radio Inteface Protocol Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. 6, V6.6.0, 2008.

[42] 3GPP TS 25.211, "Physical channels and mapping of transport channels onto physical channels (FDD)". *3GPP Organizational Partners*, Valbounne, France, Rel. 6, V6.10.0, 2009.

[43] ETSI TS 100 625, "Unstructured Supplementary Service Data (USSD)". *European Telecommunications Standards Institute (ETSI)*, Sophia Antipolis Cedex, France, Rel. V7.0.0, 1999.

[44] 3GPP TS 22.030, "Man Machine Interface (MMI) of the User Equipment (UE)". *3GPP Organizational Partners*, Valbounne, France, Rel. V4.0.0, 2001.

[45] H. Chen, S. Kumar and C. C. J. Kuo, "Dynamic call admission control and resource reservation with Interference Guard Margin (IGM) for CDMA systems". *Proc. Wireless Communications and Networking Conference (WCNC2002)*, 7-21 March 2002, Vol. 1, pp. 114-118.

[46] S. C. Yang, "CDMA RF System Engineering". Norwood, MA: *Artech House*, USA, 2002.

[47] J. Laiho, A. Wacker and T. Novosad, "Radio Network Planning and Optimisation for UMTS". Chichester, *John Wiley & Sons*, 2002.

[48] D. Lee, D. Lough, S. Midkiff, N. Davis IV, and P. Benchoff, "The Next Generation of the Internet: Aspects of the Internet Protocol Version 6", IEEE Network, January/February 1998.

[49] 3GPP TS 23.107, "Quality of Service (QoS) concept and architecture". *3GPP Organizational Partners*, Valbounne, France, Rel.6 V6.1.0, 2004.

[50] 3GPP TS 23.107, "Quality of Service (QoS) concept and architecture". *3GPP Organizational Partners*, Valbounne, France, Rel.9 V9.0.0, 2004.

[51] P. Mahajan and R. Ingalls, "Evaluation of methods used to detect warm-up period in steady state simulation", Proceedings of the 2004 Winter Simulation Conference, 2004.

[52] C. Alexopoulos, "A comprehensive review of methods for simulation output analysis", Proceedings of the 2006 Winter Simulation Conference, 2006.

[53] A. Coyle, "Teletraffic Models", Course Notes. *TRC Mathematical Modelling*, the University of Adelaide, Adelaide, Australia, p.87, July 2002.

[54] D. C. Montgomery and G. C. Runger, "Applied Statistics and Probability for Engineers". *John Wiley & Sons, Inc.*, 2003.

[55] A. Tucker, "Applied combinatorics". *Wiley*, New York, 1980.

[56] J. R. Norris, "Markov Chains". Cambridge, *Cambridge University Press*, 1997.

[57] D. P. Heyman and M. J. Sobel, "Handbooks in Operations Research and Management Science", Vol. 2 "Stochastic Models". Amsterdam, *Elsevier Science Publishers B.V.*, 1990.

[58] G. Bolch, S. Greiner, H. Meer and K. Trivedi, "Queueing Networks and Markov Chains". New Jersey, *John Wiley & Sons*, 2nd ed., 2006.

[59] H. Holma and A. Toskala, "WCDMA For UMTS - HSPA Evolution and LTE". *John Wiley & Sons*, 4th ed., Chichester, England, 2007.

[60] International Telecommunication Union, "Recommendation ITU-R P.833-6: Attenuation in Vegetation", Geneva, 2007.

[61] J. Goldhirsh and W. Vogel, "Chapter 2 - Attenuation Due to Trees: Static Case", Internet: `http://www.utexas.edu/research/mopro/papercopy/chapter02.pdf`, Jan. 19, 2010 [Jan. 3, 2001].

[62] J. Goldhirsh and W. Vogel, "Chapter 3 - Attenuation Due to Roadside Trees: Mobile Case", Internet: `http://www.utexas.edu/research/mopro/papercopy/chapter03.pdf`, Jan. 19, 2010 [Jan. 3, 2001].

[63] International Telecommunication Union, "Recommendation ITU-R P.525-2: Calculation of free-space attenuation", Geneva, 1994.

[64] International Telecommunication Union, "Recommendation ITU-R P.526-11: Propagation by diffraction", Geneva, 2009.

[65] International Telecommunication Union, "Recommendation ITU-R P.452-14: Prediction procedure for the evaluation of interference between stations on the surface of the Earth at frequencies above about 0.1 GHz", Geneva, 2009.

[66] T. S. Rappaport, "Wireless Communications: Principles and Practice". New Jersey, *Prentice Hall*, 1996.

[67] P. Wright. "Telstra launched the Next G network in October after announcing that the CDMA EVDO networks would be turned off in January 2008." Internet: `http://apcmag.com/telstra_nextg__telstra_mobile_vs_bigpond_wireless.htm`, Jul. 16, 2007 [May. 21, 2010].

[68] L. Barclay, "Propagation of Radiowaves". London, *The Institution of Electrical Engineers (IEE)*, 2nd ed., 2003.

[69] International Telecommunication Union, "Recommendation ITU-R P.1546: Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz", Geneva, 2009.

[70] Australian Communications and Media Authority: "Australian Radiofrequency Sprectrum Plan", Australia, 2009.

[71] International Telecommunication Union, Radio Regulations, Vol. 3, "Resolution 223 (Rev.WRC-07): Additional frequency bands identified for IMT", Geneva, 2008.

[72] 3GPP TS 25.104, "Base Station (BS) radio transmission and reception (FDD)". *3GPP Organizational Partners*, Valbounne, France, Rel. 9, V9.3.0, 2010.

[73] Transfinite Systems, "Visualyse Professional - make life easier, improve your output" Internet: `http://www.transfinite.com/content/Professional.html`, [May. 21, 2010].

[74] WRAP Intenational, "System overview" Internet: `http://www.wrap.se/index.php/products/system-overview-mainmenu-55`, [May. 21, 2010].

[75] D. Mulvey, "HSPA", Communications Engineer, IET Journals, vol. 5, iss. 1, pp. 38-41, 2007.

[76] H. Holma and A. Toskala, "HSDPA/HSUPA for UMTS", West Sussex, 2006.

[77] J. Lopes, L. Correia and L. Santo, "Performance Analysis of UMTS/HSPA at the Cellular Level", VTC Spring 2009, IEEE 69th, Barcelona, Spain, 2009.

[78] UMTS Forum, "3G Offered Traffic Characteristics", Report No. 33, 2003.

[79] J. Korhonen, "HSDPA An Introduction", TTPCom White Paper, 2004. Internet: `http://www.dcc.fc.up.pt/~mrodrigues/teaching/wn0809/wn0708_additions2.pdf`, [May 24, 2010].

[80] TSG-RAN WG1, "Performance Comparison of Hybrid-ARQ Schemes", Agenda Item: Adhoc#24, 2000. Internet: `http://www.3gpp.org/ftp/tsg_ran/wg1_rl1/TSGR1_17/Docs/PDFs/R1-00-1396.pdf`, [June. 4, 2010].

[81] 3GPP TS 25.308, "High Speed Downlink Packet Access (HSDPA) - Overall description Stage 2". *3GPP Organizational Partners*, Valbounne, France, Rel. 9, V9.1.0, 2009.

[82] D. Chase, "Code Combining - A Maximum-Likelihood Decoding Approach for Combining an Arbitrary Number of Noisy Packets", IEEE Transactions on Communications, No. 5, Vol. Com-33, 1985.

[83] T. Kolding, K. Pedersen, J. Wigard, F. Frederiksen and P. Mogensen, "High Speed Downlink Packet Access: WCDMA Evolution", IEEE Vehicular Technology Society News, IEEE, 2003.

[84] 3GPP TS 25.211, "Physical channels and mapping of transport channels onto physical channels (FDD)". *3GPP Organizational Partners*, Valbounne, France, Rel. 9, V9.1.0, 2009.

[85] 3GPP TS 25.306, "UE Radio Access capabilities". *3GPP Organizational Partners*, Valbounne, France, Rel. 6, V6.13.0, 2009.

[86] 3GPP TS 23.002, "Network Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. 6 V6.10.0, 2005.

[87] 3GPP TS 32.250, "Circuit Switched (CS) domain charging". *3GPP Organizational Partners*, Valbounne, France, Rel. 6, V6.7.0, 2007.

[88] L. Bos and S. Leroy, "Toward an All-IP-Based UMTS System Architecture", IEEE Network, January/February 2001, pp. 36-45, 2001.

[89] S. Stewart Ed., "Stream Control Transmission Protocol", RFC 4960, The Internet Society, 2007.

[90] L. Ong, "An Introduction to the Stream Control Transmission Protocol (SCTP)", RFC 3286, The Internet Society, 2002.

[91] S. Stewart, Q. Xie, K. Morneault, C. Sharp, H. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang and V. Paxson, "Stream Control Transmission Protocol", RFC 2960, The Internet Society, 2000.

[92] J. Stone, S. Stewart and D. Otis, "Stream Control Transmission Protocol (SCTP) Checksum Change", RFC 3309, The Internet Society, 2002.

[93] M. Antunes, L. Correia and C. Caseiro, "Impact of HSDPA Implementation on UMTS Capacity and Cell Coverage", 14th European Wireless Conference, Prague - Czech Replublic, pp. 1-5, 2008.

[94] T. Heikkila, "HSDPA radio network planning", S-72.4210 PG Course in Radio Communications, 2006. Internet: `http://comlab.hut.fi/opetus/4210/presentations/26_hsdpa_rnp.pdf`, [Jun. 7, 2010].

[95] J. Landre and A. Saadani, "HSDPA 14,4 Mbps mobiles - Realistic throughput evaluation", VTC Spring 2008, IEEE, Singapore, pp. 2086-2090, 2008.

[96] P. Zanier and D. Soldani, "A Simple Approach to HSDPA Dimensioning", IEEE 16th International Symposium on Personal, Indoor and Mobile Communications, IEEE, 2005.

[97] S. Tenorio, Y. Le Pezennec and M. Sierra, "3G HSDPA evolution: MIMO and 64QAM performance in macro-cellular deployments", 14th European Wireless Conference, Prague - Czech Replublic, pp. 1-5, 2008.

[98] P. Tapia, J. Liu, Y. Karimli and M. Feuerstein, "HSPA Performance and Evolution - A Practical Perspective". *John Wiley & Sons*, 1st ed., Chichester, England, U.K, 2009.

[99] H. Holma, A. Toskala, K. Ranta-aho and J. Pirskanen, "High-Speed Packet Access Evolution in 3GPP Release 7", Communications Magazine, IEEE, vol. 45, iss. 12, pp. 29-35, 2007.

[100] A. Goldsmith, S. Jafar, N. Jindal and S. Vishwanath, "Capacity Limits of MIMO", IEEE Journal on Selected Areas in Communications, vol. 21, no. 5, pp. 684-701, 2003.

[101] 3GPP TS 25.211, "Physical channels and mapping of transport channels onto physical channels (FDD)". *3GPP Organizational Partners*, Valbounne, France, Rel. 7, V7.9.0, 2009.

[102] 3GPP TS 25.306, "UE Radio Access capabilities". *3GPP Organizational Partners*, Valbounne, France, Rel. 7, V7.10.0, 2009.

[103] 3GPP TS 23.002, "Network Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. 7 V7.6.0, 2008.

[104] Australian Communications and Media Authority, "Spectrum Proposals: 403–520 MHz", March 2009. Available at: `http://www.acma.gov.au/webwr/_assets/main/lib310832/ifc08-2009_spectrum_proposals_403-520_mhz.pdf`

[105] Australian Communications and Media Authority, "Review of the 2.5 GHz band and long-term arrangements for ENG - Discussion Paper", January 2010. Available at: `http://www.acma.gov.au/webwr/_assets/main/lib311275/2.5ghz_discussion_paper_ifc01-10.pdf`

[106] H. Holma and A. Toskala, "LTE For UMTS - OFDMA and SC-FDMA Based Radio Access". Chichester, UK, 2009.

[107] 3GPP TS 36.101, "User Equipment (UE) Radio Transmission and Reception". *3GPP Organizational Partners*, Valbounne, France, Rel. 9, V9.4.0, 2010.

[108] 3GPP TS 23.401, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.10.0, 2010.

[109] 3GPP TS 36.300, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Overall Description - Stage 2". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.12.0, 2010.

[110] 3GPP TS 36.211, "Physical Channels and Modulation". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.9.0, 2009.

[111] 3GPP TS 36.306, "User Equipment (UE) Radio Access Capabilities". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.7.0, 2010.

[112] 3GPP TR 21.905, "Vocabulary for 3GPP Specifications". *3GPP Organizational Partners*, Valbounne, France, Rel. 10, V10.2.0, 2010.

[113] 3GPP TS 36.213, "Physical Layer Procedure". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.8.0, 2009.

[114] 3GPP TS 32.240, "Charging Architecture and Principles". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.6.0, 2009.

[115] 3GPP TS 23.203, "Policy and Charging Control Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.10.0, 2010.

[116] 3GPP TS 23.008, "Organisation of Subscriber Data". *3GPP Organizational Partners*, Valbounne, France, Rel. 8, V8.10.0, 2010.

[117] 3GPP TS 23.203, "Policy and Charging Control Architecture". *3GPP Organizational Partners*, Valbounne, France, Rel. 10, V10.0.0, 2010.

[118] C. Aubin, D. Fougère, E. Husson and M. Ivaldi, "Real-Time Pricing of Electricity for Residential Customers: Econometric Analysis of an Experiment". *John Wiley & Sons, Ltd.*, Journal of Applied Econometrics, vol. 10, S171–S191, 1995.

[119] M. Räsänen, J. Ruusunen, and R. P. Hämäläinen, "Customer Level Analysis of Dynamic Pricing Experiments Using Consumption-Pattern Models". *Elsevier Science Ltd*, Energy vol. 20, no. 9, pp. 897–906, 1995.

[120] M. Räsänen, J. Ruusunen, and R. P. Hämäläinen, "Identification of Consumers' Price Responses in the Dynamic Pricing of Electricity". *IEEE*, pp. 1182–1187, 1995.

[121] DR Enabling Technology Development Project, "Technology to Support Dynamic Pricing and Demand Response – Metering, Communication, and Data Management". *California Energy Commission*, 28/10/2002.

[122] J. Newbury, "Communication Requirements and Standards for Low Voltage Mains Signalling". *IEEE*, Transactions on Power Delivery, vol. 13, no. 1, pp. 46–52, January 1998.

[123] J. M. Peha, "Dynamic pricing and congestion control for best-effort ATM services". *Elsevier Science B.V*, Computer Networks 32, pp. 333–345, 2000.

[124] J. M. Peha, "Dynamic Pricing as Congestion Control in ATM Networks". *EEE*, pp. 1367–1372, 1997.

[125] P. Pashardes, P. Koundouri and S. Hajispyrou, "Household Demand and Welfare Implications for Water Pricing in Cyprus". *Department of Economics – University of Cyprus*, Discussion Paper, 2000.

[126] C. Courcoubetis, V. Siris and G. D. Stamoulis, "Charging Issues for Services in Broadband Networks". *Institute of Computer Science, FORTH, and Dept. of Computer Science – University of Crete*, 1995.

[127] P. R. Wurman, "Dynamic Pricing in the Virtual Marketplace". *IEEE*, IEEE Internet Computing, pp. 36–42, 2001.

[128] G. Staple and K. Werbach, "The end of spectrum scarcity - spectrum allocation and utilization", IEEE, Spectrum, Vol. 41, Iss. 3, pp. 48-52, 2004.

[129] Australian Communications and Media Authority, "Spectrum licences", Jul. 3, 2009 [Nov. 26, 2009]. Available at: `http://www.acma.gov.au/WEB/STANDARD/pc=PC_300172`

[130] A. Shukla, E. Burbridge and I. Usman, "Cognitive Radios - What are they and why are the military and civil users interested in them", EuCAP 2007, Edinhurgh - UK, pp. 1-10, 2007.

[131] R. Brodersen, A Wolisz, D. Cabric, S. Mishra and D. Willkomm, "A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum", CORVUS White Paper, Jul. 29, 2004.

[132] Telstra Corporation Limited, "Submission to the Australian Communications and Media Authority - Review of the 2.5 GHz and long-term arrangements for ENG", Mar. 12, 2010. Available at: `http://www.acma.gov.au/webwr/_assets/main/lib311275/ifc01-2010_telstra.pdf`

[133] International Telecommunication Union, "Radio Regulations - Articles", Vol. 1, Geneva, 2008.

[134] A. Shukla, "Cognitive Radio", Internet: `http://www.ofcom.org.uk/research/technology/events/rd2006/cograd.pdf`, Jul. 3, 2010 [Nov. 23, 2006].

[135] J. Mitola III, "Cognitive Radio for Flexible Mobile Multimedia Communications", MoMuC '99, California - USA, pp. 3-10, 1999.

[136] W. Tuttlebee, "Software Defined Radio - Enabling Technologies", John Wiley & Sons, Chichester, England, 2002.

[137] S. Gultchev, K. Moessner, D. Thilakawardana, T. Dodgson and R.Rafazolli, "Evaluation of Software Defined Radio Technology", Jul. 4, 2010. Available at: `http://www.ofcom.org.uk/research/technology/reserach/emer_tech_sdr/eval.pdf`

[138] F. Jondral, “Software-Defined Radio - Basics and Evolution to Cognitive Radio”, EURASIP Journal on Wireless Communications and Networking, Vol. 2005, pp. 275-283.

[139] W. Tuttlebee, “Software Defined Radio - Origins, Drivers and International Perspectives”, John Wiley & Sons, Chichester, England, 2002.

[140] Australia Communications Authority, “Spectrum Licence - Hutchison 3G Australia Pty Limited”, Jul. 4, 2010. Available at: http://web.acma.gov.au/pls/radcom/licence_search.licence_lookup?pLICENCE_NO=1137962

[141] F. Jondral, “Cognitive Radio: A Communications Engineering View”, IEEE, Wireless Communications, Vol. 14, Iss. 4, pp 28-33.

[142] J. Peha, “Approach to Spectrum Sharing”, Conference Preview, Aug. 1, 2008. Available at: http://www.comsoc.org/ci1/Public/2005/feb/cireg.html

[143] D. Goggin, “Cognitive, ultra-wide band and whitespace”, RadComms2010, Jul. 5-6, 2010. Available at: http://web.acma.gov.au/webwr_assets/main/lib311405/david_goggin_radcomms2010.pdf

[144] Australian Communications and Media Authority, “Class licensing”, Jul. 3, 2010 [Jul. 24, 2009]. Available at: http://www.acma.gov.au/WEB/STANDARD/pc=PC_1612

[145] Comlaw, “Radiocommunications (Citizen Band Radio Stations) Class Licence 2002”, Jul. 3, 2010. Available at: http://www.comlaw.gov.au/ComLaw/Legislation/LegislativeInstrument1.nsf/0/24642FD88C9A7B39CA256FC700798969/$file/cbrsCL.pdf

[146] Australian Communications and Media Authority, “Device registration”, Jul. 3, 2010 [Jan. 22, 2010]. Available at: http://acma.gov.au/WEB/STANDARD/pc=PC_310738

[147] Comlaw, “Radiocommunications Act 1992”, Jul. 3, 2010 [Feb. 28, 2010]. Available at: http://www.comlaw.gov.au/ComLaw/Legislation/ActCompilation1.nsf/0/8CBDF071CA0796D6CA2576EE001C6A2E/$file/RadiocommnsAct92.pdf

[148] E2R, “Authorisation for Download and Activation of Reconfiguration Software”, End to End Reconfigurability White Paper, Dec. 2005.

[149] QinetiQ, “Cognitive Radio Technology - A study for Ofcom”, Vol. 1, 2007. http://www.comlaw.gov.au/ComLaw/Legislation/ActCompilation1.nsf/0/8CBDF071CA0796D6CA2576EE001C6A2E/$file/RadiocommnsAct92.pdf

[150] C. Hwang, G. Lai and S. Chen, "Spectrum Sensing in Wideband OFDM Cognitive Radios", IEEE Transactions on Signal Processing, Vol. 58, Iss. 2, pp. 709-719, 2010.

[151] N. Devroye, P. Mitran and V. Tarokh, "Limits on communications in a cognitive radio channel", IEEE Communications Magazine, Vol. 24, Iss. 6, pp. 44-49, 2006.

[152] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel", IEEE Transactions on Information Theory, Vol. 27, Iss. 1, pp. 49-60, 1981.

[153] A. Carleial, "Interference Channels", IEEE Transactions on Information Theory, Vol. IT-24, Iss. 1, pp. 60-70, 1978.

[154] Y. Xiao and F. Hu, "Cognitive Radio Networks", Taylor & Francis Group, Florida, U.S., 2009.

Notes

— END OF THESIS —