A Distributed Channel Allocation Scheme for Cellular Networks using Intelligent Software Agents

Eliane Lúcia Bodanese

Submitted for the degree of Doctor of Philosophy

Department of Electronic Engineering
Queen Mary and Westfield College
University of London

November 2000

Abstract

The demand for mobile communications has grown remarkably in the last years. An efficient allocation of communications channels is essential for ensuring good performance of cellular networks, given the limited spectrum available.

Previous work on analogue and second-generation mobile communications has led to several channel allocation schemes being proposed in order to maximise the channel usage and minimise the call blocking probability. However, most of those solutions have an entirely reactive approach, which limits their efficiency. Techniques for increasing flexibility in radio resource acquisition are needed to handle the heterogeneity of services and bit rates to be supported in the forthcoming generations of mobile communications.

This thesis proposes a distributed channel allocation scheme using intelligent software agents for cellular mobile networks. The main reason for using intelligent software agents is to give greater autonomy to the base stations; this autonomy allows an increase in flexibility to deal with new situations in traffic load as well as a decrease in centralised information.

The work in this thesis demonstrates that intelligent software agents acting collaboratively in a multi-agent system are able to increase the robustness of the cellular network as a whole, to distribute the knowledge and to allow negotiation of radio resources.

The major contribution of this work is to exploit the ability of intelligent software agents to perform autonomous and intelligent negotiation to improve resource allocation in a cellular network. Intelligent software agents have not been used for that purpose before.

To God and to my parents

Acknowledgement

I would like to express my gratitude to Professor Laurie Cuthbert for his supervision and for all support and continuous encouragement he gave me in all moments of my stay at Queen Mary and Westfield College.

I also would like to thank Dr. Suthaharan Sivagnanasundaram for his helpful suggestions at the beginning of my research.

Many thanks go to the staff of the Electronic Engineering Department, lectures, secretaries and technicians who helped me during these four years.

During my studies, for many times it was not easy to go on with the work being away from my family, at these times the welcoming environment I found at QMW was very important. I would like to thank the colleagues and friends I met at the department, specially, Husam, Tijana, Lisa, Felicia, Babul, Steven, Sutha, Sammy, Robert, Veselin, Andy and Ho for the good time and for the friendship. In particular, I would like to thank Arif Sultan Al-Hammadi for his priceless friendship and support.

I also would like to acknowledge the financial support I received from CAPES and from CEFET-PR, without which my research at QMW would not have been possible.

Finally, my love and gratitude go to my family for the support and encouragement, specially, to my mother.

Contents

Chapter 5 Simulation Modelling

List of Figures

List of Tables

Glossary

ACO matrix	Augmented Channel Occupancy matrix
AKB	Assertional Knowledge Base
AMPS	Advanced Mobile Phone System
ATM	Asynchronous Transfer Mode
BBL	Behaviour-Based Layer
BCO scheme	Borrowing with Channel Ordering scheme
BDCL scheme	Borrowing with Directional Channel Locking scheme
BDI	Beliefs, Desires and Intentions
BIS	Busy/Idle Status
BS	Base Station
C/I	Carrier-to-Interference ratio
CBWL scheme	Channel Borrowing Without Locking scheme
CC set	Co-channel Cells set
CCS	Circuit Centum Seconds
CDMA	Code Division Multiple Access
CL process	Co-operative Layer process
CNP	Contract-Net Protocol
CP-based DCA scheme	Compact Pattern-based DCA scheme
CPL	Co-operative Planning Layer
CS scheme	Channel Segregation scheme
DAI	Distributed Artificial Intelligence
D-AMPS	Digital AMPS
D-BA scheme	Distributed Borrowing Algorithm scheme
DCA	Dynamic Channel Assignment
DCS-1800	Digital Cellular System 1800
DECT	Digital Enhanced Cordless Telephone
D-LBSB scheme	Distributed Load Balancing with Selective Borrowing scheme
DPS	Distributed Problem Solvers

ES	Extended Spectrum
ETSI	European Telecommunication Standards Institute
FCA	Fixed Channel Assignment
FCC	Federal Communication Commission
FDMA	Frequency Division Multiple Access
FIPA	Foundation for Intelligent Physical Agents
FlCA	Flexible Channel Assignment
FOCC	Forward Control Channel
FPLMTS	Future Public Land and Mobile Telecommunication System
GOS	Grade of Service
GSM	Global System for Mobile Communications
HCC set	Hot Co-channel Cells set
HLR	Home Location Register
HNCC set	Hot Non- Co-channel Cells set
IMT-2000	International Mobile Telecommunications at 2000MHz
INTERRAP	INTEgration of Reactive behaviour and RAtional Planning
LODA scheme	Locally Optimised Dynamic Assignment scheme
LP-DDCA scheme	Local Packing Dynamic Distributed Assignment scheme
LPL	Local Planning Layer
MA network	Multi-Agent network
MAS	Multi-Agent Systems
MIN	Mobile Identification Number
MP scheme	Maximum Packing scheme
MS	Mobile Station
MSC	Mobile Switching Centre
MT	Mobile Terminal
MTSO	Mobile Telephone Switching Office
NA-TDMA	North American TDMA
NCC set	Non-Co-channel Cells set
NES	Non- Extended Spectrum

NMT	Nordic Mobile Telephone
NTT	Nippon Telegraph and Telephone Corporation
OPs	Operational Primitives
PCN	Personal Communication Network
PCS	Personal Communication Services
PDC	Personal Digital Cellular system
PoBs	Patterns of Behaviour
PS process	Planning, Scheduling & execution process
RECC	Reverse Control Channel
RSS	Received Signal Strength
SNR	Signal-to-Noise Ratio
SB scheme	Simple Borrowing scheme
SG process	Situation recognition & Goal activation process
SHCB scheme	Simple Hybrid Channel Borrowing scheme
TACS	Total Access Communications System
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telephone System
VLR	Visitor Location Register

Chapter 1 Introduction

1.1 Introduction

The first cellular networks were implemented using static frequency channel assignment [MacD79]: after careful frequency planning, channels are assigned to cell sites and these sets are not changed except for a new long-term reconfiguration. This frequency channel assignment strategy is known as *fixed channel assignment* (FCA). After configuration the fixed channel assignment is simple to use, but it does constrain channel utilisation. If the traffic demand in a cell is greater than the number of *nominal channels*, (i.e. the frequency channels assigned to that cell) all the excess demand is blocked, regardless of the traffic conditions in other cells. This constraint is too limiting for mobile networks and a more efficient allocation of communications channels is vital for ensuring good network performance, given the limited spectrum available.

Several strategies have been proposed to maximise frequency channel usage and minimise the call blocking probability. The strategies have been divided into three groups: variants of FCA, *Dynamic Channel Assignment* (DCA) and *Flexible Channel Assignment* (FlCA).

Some channel assignment schemes presented in the literature have improved the performance of FCA for different traffic densities (macro/micro/pico cellular networks) over different traffic load conditions. However, most of the solutions proposed have an entirely reactive approach: the response to a series of events follows an algorithm that is prepared to react to specific situations. This limits their efficiency. Even those schemes that contain adaptive features are not ideal: centralised versions can become impractical because of high computational complexity and signalling overhead; distributed versions can find local minima, unaware of global or even neighbouring traffic situations, so degrading the global channel utilisation. The schemes are not able to give sufficient autonomy and flexibility to base stations to allow them to apply different ways of resource management for different situations.

This thesis proposes a distributed channel allocation scheme using intelligent software agents for cellular mobile networks. The main reason for using intelligent software agents is to give greater autonomy to the base stations; this autonomy allows an increase in flexibility to deal with new situations in traffic load as well as a decrease in centralised information.

Agent technology is an interdisciplinary area of research that has received special attention from the research community since the beginning of the 1990s. The definition of an agent is controversial, but in general terms an agent can be described as a hardware or software system with social ability that performs tasks with specific aims in a complex and dynamic environment. Agents are capable of autonomous actions to pursue their objectives, despite the occurrence of expected or limited unexpected events.

This thesis applies agents to the problem of mobile resource allocation in such a way that they do not work in isolation, but as a community. A community of agents is *a multi-agent system*, such a system being defined as a group of agents with specific roles in an organisational structure [Mül96]. The agents interact with the environment and with each other in a co-ordinated way, as collaborators or competitors, seeking to fulfil the local or global aims of the organisation.

The work in this thesis demonstrates that intelligent software agents acting collaboratively in a multi-agent system are able to increase the robustness of the network as a whole, to distribute the knowledge and to allow negotiation of resources.

1.2 Contribution

The major contribution of this work is to exploit the ability of intelligent software agents to perform autonomous and intelligent negotiation to improve resource allocation in a mobile network. *This work is completely new: agents have not been used for that purpose before.*

A particular feature of the approach is to use the special internal architecture of the agent inside each base station and the interaction between agents inside the multi-agent system to perform reactive *and* planning operations.

Other contributions in this work are the development of a special utility function and a new market-based control heuristic that are used during the agent negotiation process.

Results justifying the value of the agent approach are presented.

The author's papers are listed in Appendix A.

1.3 Organisation of the Thesis

Chapter 2 gives a brief overview of the evolution of cellular mobile networks. It shows how frequency planning is performed and how the frequency channels are allocated to cells. This chapter also introduces important traffic engineering concepts that are used throughout this thesis. It describes the different kinds of channel allocation strategies and the features of the best-known channel allocation schemes from the literature.

Chapter 3 is devoted to the theory of intelligent software agents. It describes their main characteristics and the different agent architectures that have been proposed in the literature. The particular agent architecture adopted in this work is then described together with how it can be applied to telecommunications problems. Finally, the chapter briefly describes the main properties of market-based control techniques.

Chapter 4 describes the functional specification of the multi-agent system as it is applied to this problem. The abstract agent architecture described in Chapter 3 is adapted to control the frequency channel assignment and the main functionality of each part of the intelligent agent is fully described.

Chapter 5 discusses how simulation models for all the network control strategies investigated were implemented in OPNET. Verification and validation of these simulations are described in Chapter 6, which also includes the simulation results comparing the multi-agent system implemented in this work with conventional FCA and with a distributed channel allocation scheme described in the literature.

Chapter 7 gives a general discussion and evaluation of the work and also discusses future work and the applicability of this approach in other types of cellular mobile networks. Finally, a conclusion of the work is presented.

Chapter 2 Mobile Networks

2.1 Introduction

In 1974, the Federal Communication Commission (FCC) allocated a 40 MHz band in the 800 to 900 MHz frequency range for cellular communications. The Advanced Mobile Phone Systems (AMPS) standard was introduced in 1979 and adopted by the FCC. Licences were issued in the market in 1982. An additional 10 MHz band was allocated in 1988 and called Expanded Spectrum (ES). The licences were divided into two bands: Band A and band B. Cellular communication is full-duplex and the frequency band is divided between both communication paths: 25 MHz is allocated to the *forward path* or *downlink,* which is the path from the base station transmitter towards the mobile terminal receiver. The other half is for the *uplink* or *reverse path* in the opposite direction. The paths are separated by a 45 MHz guardband in order to avoid interference between the transmission and reception channels. Figure 2.1 depicts the spectrum range occupied by AMPS.

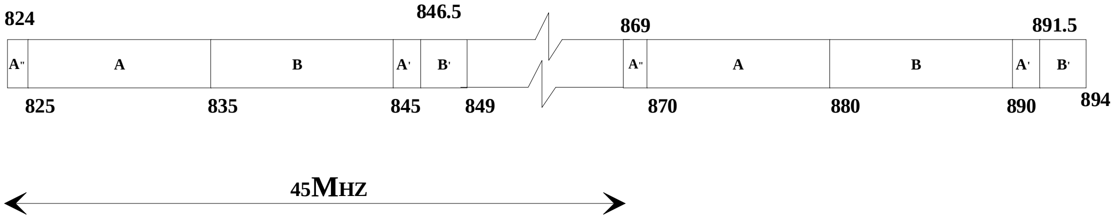


Figure 2. 1: Cellular band allocation

Bands A and B each occupy 12.5 MHz: 10 MHz is Non-Expanded Spectrum (NES) and 2.5 MHz is ES. The 12.5 MHz bands are divided into 30kHz channels, making a total of 416 channels per band. Twenty-one of these channels are used for specific procedures like channel assignment, paging, messaging, etc. They are called *control channels*. The remaining channels are used for conversation and called *voice channels*. In AMPS, each frequency channel corresponds to a frequency carrier and only one mobile can be assigned per channel. Therefore, AMPS is solely Frequency Division Multiple Access (FDMA) and is an analogue cellular system.

In Europe, other analogue cellular systems were also introduced. The two cellular systems mostly used in the European countries were Total Access Communications

System (TACS) and Nordic Mobile Telephone (NMT). TACS was introduced in the United Kingdom in 1982. The system operated at 900 MHz with a band of 25 MHz for each path and a channel bandwidth of 25 kHz. Scandinavian countries in co-operation with Saudi Arabia and Spain built NMT. The system operated at 450 MHz and 900 MHz. The total bandwidth was 10 MHz with channel bandwidth of 25 kHz and 200 channels. Both systems offered handoff and roaming capabilities. However the cellular networks were totally incompatible between countries. Even countries like Austria and Spain using the same NMT system were not compatible because of different frequency allocation, channel bandwidth and protocols. The restriction of the cellular service to mostly national territories made it clear that a future common cellular system over Europe was required.

In Japan, Nippon Telegraph and Telephone Corporation (NTT) developed a system at 800 MHz, similar to AMPS. Tokio received its first mobile telephone system in 1979. In 1985, the system operated over a spectrum of 30 MHz with 600 channels and the channel bandwidth was 25 kHz.

The second generation of mobile communications, i.e. the digital cellular systems, emerged in the 1990's. In North America, additional standards were introduced for digital cellular systems using the same frequency spectrum as AMPS. These standards integrated other multiple access techniques in addition to FDMA.

The IS-54 standard, known as North American TDMA (NA-TDMA) or Digital AMPS (D-AMPS), has integrated the Time Division Multiple Access (TDMA) technique, where each frequency carrier is shared using time division by up to 6 mobile users (currently 3 mobile users).

In 1994, the IS-95 standard introduced the Code Division Multiple Access (CDMA) technique. It is based on the *spread-spectrum* modulation in which multiple users have access to the same band. Each mobile user is assigned a unique orthogonal code, called a *Walsh code*. The 12.5 MHz of a band is divided in 10 CDMA bands of 1.25 MHz. Each CDMA band supports 64 Walsh codes. CDMA can offer about eight times the capacity of analogue [Gar00].

In Europe, the Global System for Mobile Communications (GSM), the Pan-European digital mobile telephony standards specified by the European Telecommunication Standards Institute (ETSI) was introduced in 1982 [Meh96]. It is a TDMA system and operates at 890 MHz to 915 MHz for the uplink and from 935 MHz to 960 MHz for the downlink. The frequency carrier bandwidth is 200 kHz which can be shared by 8 mobile users using time division. In 1992, the first GSM system was deployed in Germany.

Personal Digital Cellular (PDC) is the digital cellular system standard in Japan. The system is a TDMA cellular system operating at 800 MHz and 1.5 GHz. In general, the PDC system is very similar to NA-TDMA. Three mobile users can share each frequency carrier. For the 800 MHz range, the uplink bandwidth is between 810 and 826 MHz and the downlink between 940 and 956 MHz.

The success of the second-generation cellular technology, particularly with the rapidly decreasing costs, has led to the prospect of capacity and service saturation. Not only the growth of the cellular systems, but also the customers' expectations of using a multi-application terminal handset, providing voice, data and multimedia services have made clear the limitations of the available spectrum. In order to supply the personal communications needs of mobile users and in-building communications, another range of the spectrum was allocated, and satellite systems for mobile communication started being developed. In 1989, the United Kingdom promoted a system called “Personal Communication Network” (PCN) to operate at 1.8 GHz. PCN uses a version of GSM originally known as Digital Cellular System 1800 (DCS-1800) and now called GSM-1800. In 1991, the FCC allocated 120 MHz of spectrum into seven bands for the so-called wideband Personal Communication Services (PCS) at 1.85 GHz and narrowband PCS at 900 MHz. In Japan, the 1.5 GHz PDC was in service in Osaka in 1994. In this high spectrum frequency, the wide-area mobility is served by the TDMA or CDMA systems. Low-mobility systems can be served by proprietary radio access technologies [Meh96].

The drastic growth in the use of mobile communications by public and business sectors increased the pressure to integrate fixed and mobile networks. Now, mobile networks are expected to have the same diversity of services offered by fixed networks with the same quality of service and security. Also, full mobility capability is expected. The mobile system needs to have the flexibility to integrate world-wide the different types of mobile communication systems available today, such as public and private cellular systems, data radio and satellite systems. These demands are beyond the technological capabilities of the second generation of mobile communications. These pressures and developments in component technology, network management and service engineering made inevitable the emergence of a third generation of mobile communications. The aim of third generation systems is to provide communication services from any person to any person at any place and at any time through any medium using a compact light-weight terminal with guaranteed quality of service and security. Two systems are being developed: Universal Mobile Telephone System (UMTS) and Future Public Land and Mobile Telecommunication system (FPLMTS) also called International Mobile Telecommunications at 2000 MHz (IMT-2000). The deployment of these systems will be in stages during the first decade of the 2000's.

UMTS will offer a common air interface covering home, office, car, train, aeroplane or as a pedestrian. UMTS will integrate in one service all the services offered by different mobile communications systems such as mobile telephone, cordless telephone, public air radio, satellite radio, etc. It will allow users to roam during an existing connection between different types of communication networks. UMTS will offer broadband services, i.e. it will be possible to transmit voice, text, data and images over one connection.

FPLMTS covers the same aspects as UMTS, however with uniform frequencies world-wide. FPLMTS defines four air interfaces for dealing with different requirements of densely or sparsely populated areas.

One of the key requirements of UMTS/FPLMTS is to have high spectral efficiency and to achieve this requirement there is a need for flexible frequency management and flexible management of radio resources. Frequency management in mobile networks has been a hot topic for research in the past 20 years and the solutions proposed still present a lack of intelligence and flexible behaviour. The technological advances made in software and hardware in the last decade are providing the means to introduce intelligence in control and management of networks. The introduction of more intelligence and flexible behaviour in the management of channel allocation is the objective of the work in this thesis.

The remainder of this chapter provides an understanding of the important aspects involved in frequency channel assignment in mobile networks as introduced by the *cellular concept* [Mac79]. The description of this approach is necessary because it was used in this work and in most of the schemes proposed in the literature. The main characteristics of a cellular system (concerning frequency channel allocation) are described and the different frequency channel assignment strategies are discussed; the main schemes proposed in the literature are also described. Finally, the advantages and disadvantages of the schemes concerning network performance are discussed.

2.2 The Cellular Concept

The cellular concept, conceived by Bell Systems under the AMPS standard in 1979 [MacD79], is a mobile network architecture composed ideally of hexagonal cells. The cells represent geographic areas. Inside the cells, the users, called *mobile stations* (MS) are able to start/receive communications while moving inside the cellular network. Each cell has a *base station* (BS) which supplies frequency channels to the mobile stations. Base stations in AMPS are known as *cell sites*. The cell sites are linked to a *mobile switching centre* (MSC) called *mobile telephone switching office* (MTSO) responsible for controlling the calls and acting as a gateway to other networks. When an active user (i.e. a mobile station using a frequency channel) reaches the boundary of the cell, it needs to change its current frequency channel for another belonging to the neighbouring cell. This network procedure is known as *handoff* or *handover*. An illustration of the AMPS mobile system architecture is given in Figure 2.2.

The main objectives of AMPS for supplying a large-scale mobile-telephone service were [Mac79]:

- Large subscriber capacity
- Efficiency use of spectrum
- Nationwide compatibility
- Widespread availability
- Adaptability to traffic density
- Service to vehicles and portables
- Regular telephone services and special services
- Quality of service in telephony
- Affordability

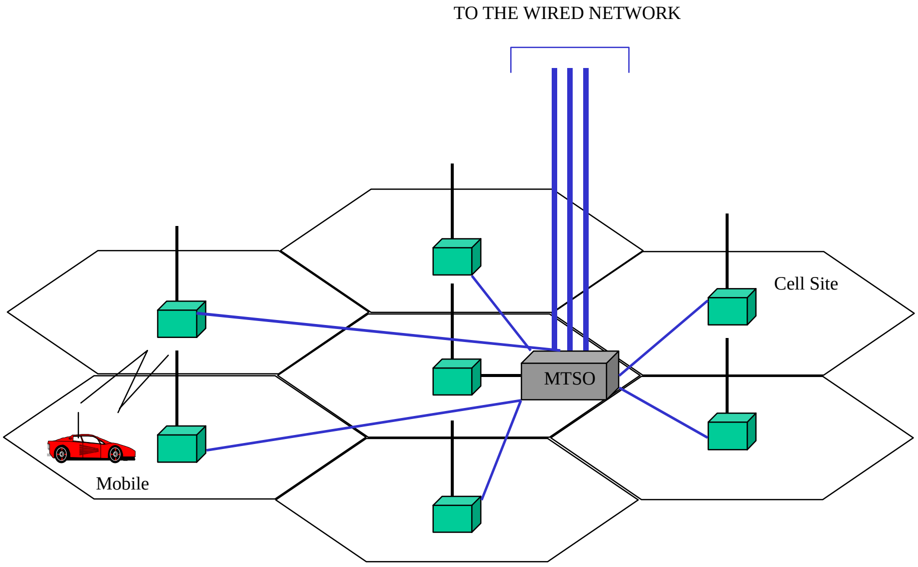


Figure 2.2: Cellular band allocation

The essential features of the cellular system that made possible the achievement of the above objectives were *frequency reuse* and *cell splitting*.

Frequency reuse refers to the use of the same frequency carrier in different areas that are distant enough so that the interference caused by the use of the same carrier is not a problem. The reason for the application of frequency reuse is twofold:

To reduce the cost of the land transmitter/receiver site by placing several moderate power land sites to cover sub-areas (cells) of the designated area for use of the network operator.

To greatly increase the number of simultaneous calls that can be covered by the same number of allocated channel frequencies.

Cell splitting is the reconfiguration of a cell into smaller cells. This feature makes it possible for the same network to service different densities of demand for channels. Larger cells can serve low demand areas and smaller cells high demand areas. Cell splitting is a long-term configuration planning that allows the system to adjust to a growth in traffic demand in certain areas, or in the whole network, without any increase in the spectrum allocation.

2.2.1 Frequency Reuse

The distribution of the frequency channels in a cellular network is dependent on several parameters, such as cellular geometry, signal propagation characteristics and signal interference.

The assignment of frequency channels in the cellular concept is fixed, i.e. a set of frequency channels is statically allocated to a cell. This same set is reused in another cell distant enough to allow the use of the frequency channels with acceptable signal interference. Cells that use the same set of frequency channels are called *co-channel cells* and the distance between them is called *co-channel reuse distance*. The total number of frequency carriers allocated to a network operator is divided in sets and each set is assigned to a cell inside a cluster of cells. The cluster of cells forms a pattern. The pattern is reused according to the co-channel reuse distance. The choice of the number of cells per cluster is mainly governed by co-channel interference considerations. A better understanding about signal propagation and cellular geometry is needed in order to understand how frequency assignment is performed in a mobile cellular system (e.g. AMPS).

The propagation path loss of a signal is a function of several factors, such as environment, antenna type, antenna height, location, etc. Considering onmidirectional antennas, the propagation path loss in a mobile radio environment is normally taken as 40 dB per decade, i.e. the signal will suffer a 40 dB loss each 10 km [Lee95]. The difference in power reception at two different distances d_1 and d_2 would be:

$$\frac{\boldsymbol{C}_2}{\boldsymbol{C}_1} = \left(\frac{\boldsymbol{d}_2}{\boldsymbol{d}_1}\right)^{-4} \qquad (2.1)$$

where

$\boldsymbol{C_1}$ is the received carrier power at receiver 1

$\boldsymbol{C_2}$ is the received carrier power at receiver 2

$\boldsymbol{d_1}$ is the distance measured from the transmitter to receiver 1.

$\boldsymbol{d_2}$ is the distance measured from the transmitter to receiver 2.

Therefore, the signal strength is inversely proportional to the distance to the power 4. In decibel expression Equation 2.1 becomes:

$$\Delta \boldsymbol{C} = 10\log\frac{\boldsymbol{C}_2}{\boldsymbol{C}_1} = 40\log\frac{\boldsymbol{d}_1}{\boldsymbol{d}_2} \qquad (2.2)$$

In the same conditions, but in free space, the propagation path loss would be of 20 dB/10 km. In a real mobile radio environment the propagation path loss will vary as:

$$\Delta \boldsymbol{C} = \boldsymbol{\alpha d}^{-\boldsymbol{\gamma}} \qquad (2.3)$$

Or in decibel:

$$\Delta \boldsymbol{C} = 10\log\boldsymbol{\alpha} - 10\cdot\boldsymbol{\gamma}\log\boldsymbol{d} \qquad (2.4)$$

where

$\boldsymbol{\gamma}$ is the propagation path loss factor.

$\boldsymbol{\alpha}$ is a constant

d is the distance from the transmitter to the receiver

The $\boldsymbol{\gamma}$ parameter usually lies between 2 and 5; it cannot be lower than 2, the free-space condition.

Co-channel interference occurs as a result of multiple uses of the same frequency carrier. The *carrier-to-interference ratio* (***C/I***) is used to measure the amount of interference over a specified carrier.

$$\frac{C}{I} = \frac{C}{\sum_{k=1}^{K_I} I_k} \qquad (2.5)$$

where K_I is the number of co-channel cells interfering in the first tier (the interference of the second tier of co-channel cells can be neglected, see [Lee95]).

Assuming the local noise is much less than the interference level and can be neglected, then the ***C/I*** can be expressed by Equation 2.6:

$$\frac{C}{I} = \frac{R^{-\gamma}}{\sum_{k=1}^{K_I} D^{-\gamma}{}_k} \qquad (2.6)$$

where

$\boldsymbol{\gamma}$ is the propagation path loss factor.

D is the frequency reuse distance

R is the radius of the cell, defined as the distance from the centre of the cell to any of its vertices.

Assuming, for simplicity, that ***Dk = D*** for all $\boldsymbol{K_I}$, the ***C/I*** of a cell site radiating in all directions (omnidirectional antenna) can be represented (dB) by [Far96]:

$$\frac{C}{I} = 10\log\left[\frac{1}{K_I} * \left(\frac{D}{R}\right)^{\gamma}\right] \tag{2.7}$$

After evaluation of the effect of co-channel interference on perceived quality of service, it was decided that to obtain a good transmission quality for a channel, the AMPS system must provide an C/I of 17dB or greater over 90% of its coverage area.

Solving the Equation 2.7 for ***D/R***, with ***C/I*** equals to 17dB and taking the general case of $\boldsymbol{\gamma}$ equal to 4 and the worst case of interference with $\boldsymbol{K_I}$ equal to 6, the result is:

$$17 = 10\log\left(\frac{1}{6} * \left(\frac{\boldsymbol{D}}{\boldsymbol{R}}\right)^{4}\right) \therefore \frac{\boldsymbol{D}}{\boldsymbol{R}} = 4.1642598 \tag{2.8}$$

A distance between the border of the cell and the cell site of the co-channel cell greater than 4.17 cell radii will be enough to guarantee a good transmission quality. After other signal propagation considerations (like Rayleigh fading in UHF), the system designers of AMPS conclude that a radio frequency *signal-to-noise ratio* (SNR) of 18 dB or higher should be applied for a working system. Practical simulations, taking into consideration the specified 75% of the mobile users saying that the voice quality is good or excellent in 90% of the coverage area, led a value of ***D/R*** of 4.6 as illustrated in the Bell lab publication [Mac79].

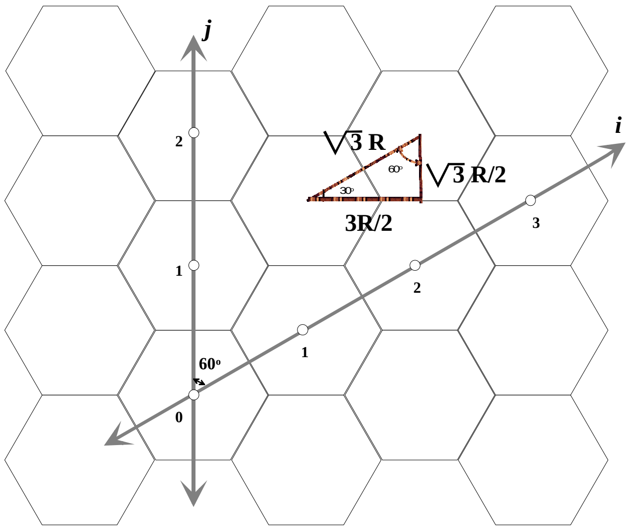


Figure 2.3: Co-ordinates for hexagonal geometry

Assuming the ideal case of hexagonal cells, the frequency reuse distance (***D/R***) can be related to a finite set of cells ***N*** in a hexagonal cellular network. In [Mac79] a convenient set of co-ordinates for hexagonal geometry was introduced. The positive halves of the two axes intersect at a 60-degree angle, and the unit distance along either axis is $\sqrt{3}$ times the radius of the cell (***R***), that corresponds to the distance between the centre of two hexagonal cells (in cell radii) as illustrated in Figure 2.3. The distance between the origin to any cell centre is given by:

$$D = \sqrt{i^2 + ij + j^2} \qquad (2.9)$$

The vectors from the centre of any arbitrary cell and the six adjacent cells are separated from each other by 60 degrees, this same observation is valid for the vectors from a cell to its co-channel cells. Therefore, a cluster of contiguous cells can be visualised as a large hexagon. It is not claimed that all kind of clusters will have a hexagon shape, but a large hexagon can have the same area as any valid cluster. As the distance between centres of adjacent cells is unity, the distance between centres of the large hexagon is $\sqrt{i^2 + ij + j^2}$. The pattern of large hexagons can be visualised as an enlarged replica of the original cellular pattern with a scale factor of $\sqrt{i^2 + ij + j^2}$. Therefore, the number of cell areas contained in the area of the large hexagon is:

$$N = i^2 + ij + j^2 \qquad (2.10)$$

From Equation 2.10, valid number of cells per cluster are 3, 4, 7, 9, 12, 13, 19, etc.

Finally, the relation between the co-channel reuse distance (***D/R***) and the number of cells per cluster can be found by combining the equations 2.9 and 2.10 and replacing the unity by $\sqrt{3}R$:

$$\frac{\boldsymbol{D}}{\boldsymbol{R}} = \sqrt{3 * \boldsymbol{N}} \qquad (2.11)$$

Now, the minimum number of cells per cluster that is needed to meet the system performance requirements can be determined. When considering omnidirectional cell sites and a flat terrain, a cluster with 7 cells gives a 4.58 frequency reuse distance, enough to comply with the performance requirements.

However, in practical systems when onmidirectional antennas are used a cell cluster of 9 or 12 cells are implemented to guarantee the system performance requirements and the reason for that is explained as follows.

A 7-cell cluster does not provide a sufficient frequency reuse distance separation even when an ideal condition of flat terrain is assumed. This happens when the worst case scenario for the mobile station is analysed. In the worst case, the mobile station is at the boundary R, where it would receive the weakest signal from its own base station. The distances from all six co-channel cells are: two distances of D – R, two distances of D, and two distances of D + R.

Following the mobile radio propagation rule of 40 dB/10 km, [Lee95] explains that:

$$\mathrm{C} \propto \mathrm{R}^{-4} \qquad\qquad \mathrm{I} \propto \mathrm{D}^{-4}$$

Then

$$\frac{C}{I} = \frac{R^{-4}}{2(D-R)^{-4} + 2D^{-4} + 2(D+R)^{-4}} \qquad (2.12)$$

For a D/R of 4.6, the value of C/I is 54 or 17 dB, which is lower than 18 dB. In real systems as the site locations are imperfect and the terrain is not flat, the C/I received is always worse than 17 dB and could be lower than 14 dB. Therefore, in an omnidirectional-cell system, a cell cluster of 9 or 12 would be a correct choice, because, even considering the shortest distance of D-R for all six interferes as worst case, the values of C/I would be greater than 18 dB.

The use of directional antennas can improve the C/I without the need to increase the number of cells in the cluster. Figure 2.4 illustrates the worst case situation for a cell site with three 120-degree directional antennas. The front-to-back ratio of a sectored antenna is at least 10 dB, therefore the interference can be considered in only one direction. For the 3-sector cell, the number of interfering cells is reduced to two. Assuming the values of the distance of the interfering cells to the mobile station are D+0.7R and D, then:

$$\frac{C}{I} = \frac{R^{-4}}{(D+0.7R)^{-4} + D^{-4}} \qquad (2.13)$$

Applying in Equation 2.13 the reuse distance for a 7-cell cluster of 4.6 results in a C/I of 24.5 dB that greatly exceeds 18 dB. In real systems, the C/I could be 6 dB weaker in a heavy traffic area as a result of irregular terrain contour and imperfect site locations [Lee95], but in this case the C/I still would be adequate.

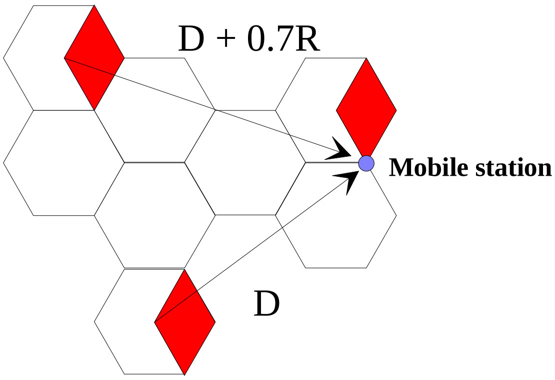


Figure 2.4: Worst case of interference in a 120° directional antenna

For a 4-cell cluster the frequency reuse distance is 3.46, applying to Equation 2.13 gives a C/I of 20 dB that may be unacceptable in situations where the C/I can become 6 dB weaker. A 6-sector antenna may be more appropriate for a 4-cell cluster, because only one interferer at D+07R needs to be considered (Equation 2.14), giving a C/I of 26 dB when a frequency reuse distance of 3.46 is applied.

$$\frac{C}{I} = \frac{R^{-4}}{(D+0.7R)^{-4}} \qquad (2.14)$$

Real systems deployed in less populated areas use onmidirectional antennas and normally have a 12-cell cluster. In more populated areas, systems using 120-degree directional antennas use 7-cell or 4-cell clusters or 60-degree directional antennas with 3-cell or 4-cell clusters.

2.2.1.1 Reuse Plan and Channel Grouping

When the number of cells per cluster is defined it is then necessary to determine which channel set should be assigned to each cell. The frequency reuse layout of the cellular system is easily assembled following a scheme that finds the nearest co-channel cells of any cell of the network.

The scheme uses the hexagonal co-ordinates of Figure 2.3. In the scheme, *i* and *j* are called *shift parameters*. Depending on their values different patterns are formed. The chosen cell and its N-1 surrounding cells form a pattern known as a *compact pattern*. The steps of the schemes are as follow:

- Choose any cell as reference;
- For each side of the hexagon: move ***i*** cells along that side, turn counter-clockwise 60 degrees and then move j cells on this new direction.
- Repeat the scheme to the surrounding cells of the initial reference cell which are not found co-channel cells.

Figure 2.5 shows seven compact patterns in a 7-cell cluster cellular network (where *i*=2 and *j*=1).

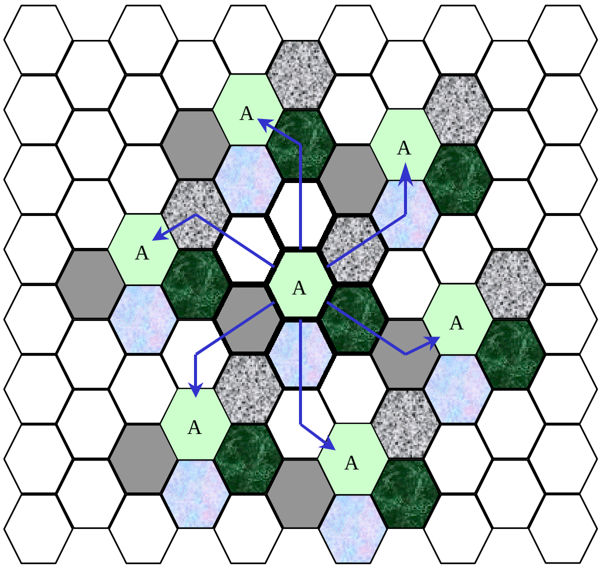


Figure 2.5: Formation of 7-cell compact pattern in anti-clock wise rotation

The co-channel cells can also be found by turning 60 degrees clockwise. Therefore, two different frequency reuse configurations can be achieved when (i 0 and j 0) or (i j).

The second phase is the channel grouping i.e. to determine what channels will compose each set. When planning this distribution other signal interference needs to be considered: the *adjacent channel interference*. In general, frequency filters require a substantial spectral guard band to reject adjacent frequencies adequately. The spacing of 30 kHz and peak deviation of 12 kHz in AMPS are not enough to do this. Therefore, the assignment of channels for the same cell site is kept as far apart as possible. The channels are numbered from 1 upward and the frequency difference between channels is then proportional to the difference of their channel numbers. If $\boldsymbol{N}$ disjoint channel sets are to be deployed, for example seven as in Figure 2.5, the n^{th} set would contain channels n, $n + \boldsymbol{N}$, $n + 2\boldsymbol{N}$, etc. For example, if $\boldsymbol{N}$=7, the first set would contain channels 1, 8, 15, etc. However, there is still a second source of adjacent channel interference produced when adjacent channels are used in geographically adjacent cell sites. Sets with adjacent set numbers have adjacent channels. In order to avoid this source of interference, the sets in the cluster are located in a way that minimises adjacent channel interference. Figure 2.6 shows the set distribution for a 12-cell cluster and for 7-cell cluster.

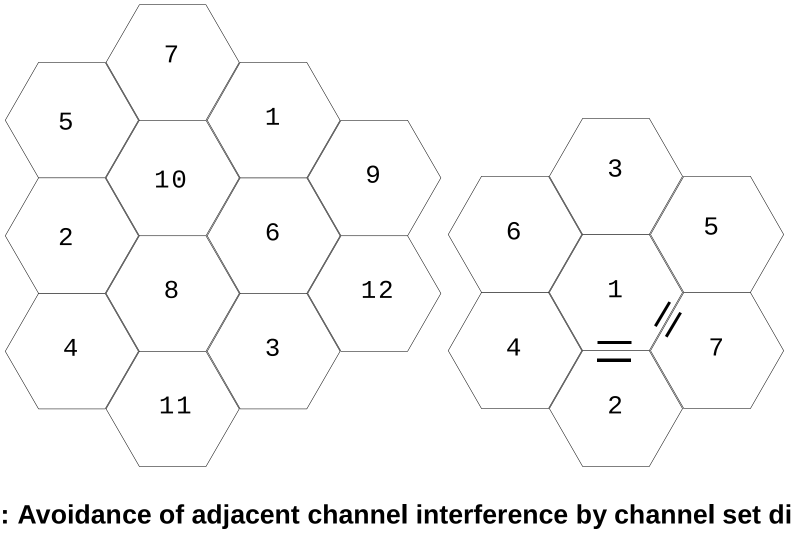


Figure 2.6: Avoidance of adjacent channel interference by channel set distribution

As illustrated in Figure 2.6, it is impossible in a 7-cell cluster (or smaller) not to have adjacent channels at adjacent cell sites. However, the use of directional antennas allows each set be divided in sub-sets and distributed in the cell in such a way that channel adjacent interference is attenuated by the front-to-back ratio of the cell site directional antennas. For more details see [Mac79] and [Far96].

In AMPS, from the 416 available frequency channels, 21 are assigned to be control channels. For a 7-cell frequency plan, it was decided to divide all available frequencies into 21 groups with one control channel assigned to each group. Three groups were assigned to each one of the seven sets.

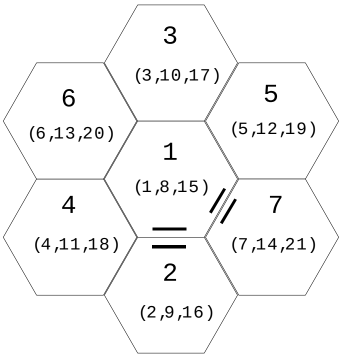


Figure 2.7: Seven-cell frequency reuse plan

Different kinds of frequency reuse plans can be adopted in systems with other numbers of cells per cluster. Other standards are more restricted regarding frequency reuse plans. For example, GSM applies a 4-cell reuse plan and normally uses 120-degree directional antennas [Meh96]. However, the aim of this section is to give a general idea of the principal issues involved in frequency channel assignment in AMPS, more detailed information about this subject can be found in the references.

2.2.2 Cell Splitting

Cell splitting allows the system to grow gradually in response to a growing demand of traffic. It takes place by reducing the radius of a cell by half and splitting the old cell into four new small cells. The reuse frequency can be used more often allowing the traffic to grow four times in the same area where an old cell was placed. The ideal location for new cells is the midway points between neighbouring existing cells. Considering cell splitting in a 7-cell cluster system, the new cluster will present a rotation of 120-degree counter-clockwise in relation to the larger cluster [Mac79]. The channel set assigned to a new cell is the one that makes the new cell lie at a midway distance between two of its nearest co-channel cells in the larger cluster, each one situated approximately a larger cell diameter away from the new cell.

When two or more sizes of cells co-exist in a mobile network, special attention needs to be taken in order to guarantee that the minimum frequency reuse distance is being respected. The use of channels in the new cell sites will not cause interference problems in the larger system because the strength of the signal is smaller and designed to comply with the constant ***D/R***, but taking the radius of the small cell. The problem comes with the channels of the co-channel cell in the larger system, because the small cell is within the frequency reuse distance of the larger cell. One way to deal with this problem is to use the *overlaid-cell concept*. In the overlaid-cell concept the cellular network is seen as a superposition of the smaller cell pattern on top of the complete larger cell pattern, and not only where the smaller pattern is in reality deployed. Each cell face will divide its channels between a larger-cell group and a smaller-cell group. The selected channels in the larger-cell group will be used in all coverage area of a larger cell. The selected channels installed in the smaller cell will compose the smaller-cell group of its larger co-channel cells. The formation and use of the channel groups is then governed by the presence, or not, of real smaller cell neighbours. For example, the use of any channel installed in a small cell must be restricted to the smaller-cell overlay area in the nearest larger co-channels. If a mobile user using a smaller-cell channel goes out of the perimeter of the small cell overlay on the larger cell, it needs to handoff to a channel of the larger-group or to a neighbouring cell if this is the case. Therefore, the presence of two sizes of cells in a network reduces the capacity of some of the larger cells and can force their cell to split even if they would not be split if only the growth of traffic in their areas would had considered.

2.2.3 Other System Parameters

Other important parameters in a cellular network are cell-site position tolerance, and maximum and minimum cell radius [Mac79]. The recommended values for these parameters are based on best tradeoffs between customer capacity, cost restraints and good transmission quality.

In a perfect cellular network the spacing of cell sites should be regular. However, the installation of a cell site in the ideal position is generally not possible. AMPS systems are allowed to position their cell sites up to one-quarter of the nominal radius away from the ideal location [Mac79]. This decision was based on studies of the impact of the cell site position on transmission quality. In general terms, the C/I falls at 10% of the overall C/I distribution. This level decreases gradually as the position of the cell site goes from zero to one-quarter of the ideal position, but it decreases sharply beyond this point. This is the main reason for the imposition on the position tolerance limit.

The value for the maximum cell radius is a decision that needs to be taken in the start up phase of the cellular system. It is a compromise between cost, forecast of ultimate capacity and transmission quality. The transmission quality is defined as a SNR of at least 18 dB in 90% of the coverage area. The maximum power level for mobile terminals does not exceed 20 W (car portable). Considerations of the cost of transmitter power and antenna height, the expected number of cell sites in mature cellular system, signal propagation characteristics of the region being considered and the maximum power level of mobile terminal will define the maximum radius of the cell.

The minimum cell radius in a mature cellular system will have little effect on the system cost per customer or on transmission quality, but it will significantly affect the system capacity. The practical obstacles involved in small cells are the cell-site position tolerance and, most importantly, the processing capacity of the system in dealing with the burden of frequent handoffs.

There are other parameters that influence the frequency planning and cell site engineering of cellular networks; however a deeper study in both subjects is beyond the scope of this thesis.

2.3 Traffic Engineering in Mobile Networks

This section gives a basic understanding of traffic engineering and engineering aspects of cell site provisioning in cellular communication systems.

Telecommunication network resources are limited and they need to be shared by all the network users. The problems that arise from the necessity of sharing resources are addressed by traffic engineering which tries to bring a balance between customer satisfaction and revenue for network operators, i.e. to serve the greatest number of customers with a specified system quality. Traffic engineers need to have a good understanding of traffic distribution, traffic growth and customer requirements. In mobile cellular networks, the results of their calculations will show how many customers will be served in a busy hour, how many subscribers can be taken by the cellular system, how many cells are needed and how many channels per cell are needed.

2.3.1 Traffic Characteristics

The intensity of two kinds of traffic can be measured: the *offered traffic* (the traffic the network receives) and the *carried traffic* (the traffic the network successfully carries).

The offered traffic varies during the day. It is normally low during the night, rises in the opening business hours, goes down during lunchtime and rises again in the afternoon. The two peaks in traffic, measured over an hour, represent the busiest periods and are called *busy-hour traffic*. The daily variation in offered traffic intensity is known as *hourly variations*. The intensity of the busy-hour traffic varies also depending on the day of the week (month of the year) and these variations are known as *daily variations*. The network operator needs to meet the demands of the average busy-hour traffic if it is to be credible to its customers.

Traffic intensity can be measured using two dimensionless units: Erlang and Circuit Centum Seconds (CCS).

The Erlang is a unit named in memory of Anders K. Erlang, the founder of traffic theory. One Erlang is equivalent to one circuit (or trunk) in continuous use, and it can be translated as the number of calls (made in one hour) multiplied by the duration of these calls (in hours). Each call has a different duration or a different *call holding time*; for traffic intensity measurements the *average call holding time* is taken into account. This has different values for business or private subscribers, but the typical values for average call holding time vary between 120 and 180 seconds. Therefore, the traffic intensity in Erlangs can be defined as:

$$T(inErlangs) = \frac{(\text{number of calls in an hour}) * (\text{average call holding time(s)})}{3600} \quad (2.15)$$

Traffic intensity can also be measured in CCSs. One CCS is equivalent to one circuit in continuous use for 100 seconds. The traffic intensity measured in CCS is:

$$T(inCCSs) = \frac{(\text{number of calls in 100s}) * (\text{average call holding time(s)})}{100} \quad (2.16)$$

If a call attempt is made when all circuits (channels in cellular networks) are serving other calls, the call attempt will be blocked. The probability of call blocking in a telecommunication network is called *grade of service* (GOS). The GOS of a telecommunication network varies between zero and one. A GOS of 0.02 is normally taken as acceptable for communication systems.

2.3.2 Assigning the Appropriate Number of Channels per Cell

In 1917, Erlang developed a very important equation, which express the probability of a call being blocked (***P_b***), as a function of the offered traffic (***T***), and a number of circuits (or trunks) (***C***).

$$P_b = {\frac{T^C}{C!}} \Bigg/ {\left(1 + \frac{T}{1!} + \frac{T^2}{2!} + \ldots + \frac{T^C}{C!}\right)} \qquad (2.17)$$

Using Equation 2.17 it is possible to determine the number of channels required to support a certain offered traffic given the desirable GOS. Equation 2.17 is called the *Erlang B formula*. It assumes that blocked calls are cleared and the caller will try again later. Erlang developed another call blocking probability equation assuming that a blocked call will be queued until it is established. That equation is known as *Erlang C formula*.

Another important assumption made in Equation 2.17 is about the pattern of the call attempts or *call arrivals.* The calls occur “individually and collectively at random” [PS96]. The statistical distribution that can be used as a mathematical model for this kind of arrival is the *Poisson distribution*. The Poisson distribution gives the probability that a certain number of events will occur randomly during a particular time interval. Experience over many years shows that this is a good approximation to what happens in practice.

As an example of the use of the Erlang B formula in the determination of the number of channels, assume that the average number of calls per hour in the busy-hour is 880, the average call holding time is 180 seconds and the GOS of 0.02. The offered traffic intensity by Equation 2.15 is 44 Erlang. Looking at Erlang B table for a GOS of 2% (as found in [Lee95]), it is found that 54 channels are needed for an offered traffic of 44 Erlang. Therefore the cell site should have at least 54 channels to cope with an offered traffic intensity of 44 Erlang.

However, special attention needs to be paid to the channel utilisation efficiency when assigning the number of channels per cell. Channel utilisation efficiency or *trunking efficiency* (***TE%***) is defined as the percentile ratio of the offered traffic ***T*** in Erlang and the number of channels ***C***.

$$TE(\%) = \frac{T}{C} * 100 \qquad (2.18)$$

The graph below shows the trunking efficiency curve in relation to the number of trunks and the correspondent capacity in Erlangs of offered traffic that the system can take care of, considering a grade of service of 0.02.

Given a GOS, the trunking efficiency increases as the number of trunks (channels) increases. A cell with less than 15 channels has poor channel utilisation efficiency and consequently is less cost effective and generates less revenue.

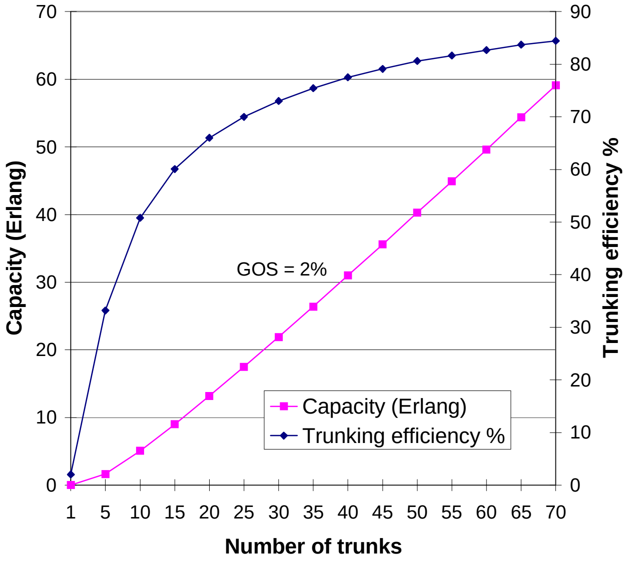


Figure 2.8: Trunking efficiency

When a cell is divided in sectors there is a degradation of channel utilisation efficiency. For example, a cell with 48 channels can receive traffic of 38.4 Erlang (GOS 2%): its efficiency is 80%. If the cell is now divided in three sectors each one having 16 channels, each sector can receive traffic of 9.83 Erlang. The total traffic that the cell can cope now is 29.49 Erlang and the trunking efficiency is only 61.4%

(29.49*100/48). Therefore, it is necessary to allocate more channels to cope with the given offered traffic if the cell is divided in sectors.

2.3.3 Estimating the Number of Subscribers in the Cellular System

The number of subscribers in the system can be estimated assuming the relation between the number of subscribers in the busy hour (η_c) and the number of calls per hour per cell [Lee95]. The maximum number of calls per hour that a cell can take depends on the number of channels allocated for that cell based on traffic conditions under its geographic area. As an example, assume a system with seven cells, the maximum number calls per hour in each cell is 2000, 1500, 500, 1000, 1200, 900, 800. Assuming that 60% of the subscribers will be using their mobile terminals during the busy hour traffic (η_c = 0.6) and one call is made per mobile. The estimated number of subscribers in the system ***M*** is:

$$M = \frac{\sum maximum\, number\ of\ calls\ per\, cell}{\eta_c} = \frac{7900}{0.6} = 13{,}166 \qquad (2.19)$$

2.3.4 Estimating Total Number of Cells

The first step in the deployment of a cellular system is the acquisition of the traffic distribution over the chosen service area. Firstly, the population density per square kilometre is translated into traffic intensity (Erlang). The estimated offered traffic per square kilometre is called *bin*. A grid of bins with different colours or patterns overlays the service area. Each colour or pattern of a bin represents a different level of offered traffic. The system designers superimpose hexagonal cell grids over the entire service area already mapped into bins. The traffic intensity in each cell will be the sum of the contained traffic bins. In this way, traffic engineers and system designers can make all the calculations and select the hexagonal cell grid that provides the best cost/benefit cellular system for the network operator.

2.4 Channel Assignment Strategies

The frequency channel assignment in the cellular concept is static, i.e. after careful frequency planning, channels are assigned to cell sites and these sets will not change except for a new long-term reconfiguration. Cell sites will only make use of the assigned channel set or individual assigned channel sets per sector. This frequency channel assignment strategy is known as *fixed channel assignment* (FCA). After the cellular system has been configured, the fixed channel assignment is simple to use. However, it does constrain channel utilisation. If the traffic demand in a cell is greater than the number of *nominal channels*, (i.e. the frequency channels assigned to that cell) all the excess demand is blocked, regardless of the traffic conditions in other cells. This constraint is very limiting for mobile networks and several strategies have been proposed to maximise frequency channel usage and minimise the blocking probability. The strategies have been divided into three groups: those based on FCA, *Dynamic Channel Assignment* (DCA) and *Flexible Channel Assignment* (FlCA)

Two FCA variant strategies have been proposed: *load sharing* and *channel borrowing* (with or without channel locking).

It is assumed in the *load sharing* strategy [Ekl86] [KE89] that there is an overlapping coverage area between cells where mobiles can obtain a quality of transmission from the neighbouring cell almost as good as that in their own. When there is a call attempt and no more available channels [Ekl86], or when the channel occupancy reaches a pre-defined threshold [KE89], the MSC may advise some mobile users of the cell to check the transmission quality of channels in neighbouring cells. For each one of them that can get acceptable transmission quality from a neighbouring cell, a handoff request will be made to that cell and the mobile moved, provided the cell has enough available channels to allocate one to the requesting mobile user. In this way, the congested cell can have some of its nominal channels freed and use them in the new call requests. This load sharing strategy is also known as *directed retry*.

Schemes using a *channel borrowing* strategy differ from the original FCA concept by allowing a cell to use some of the channels of other channel sets apart from its own. The channel borrowing is performed when there are no more nominal channels to serve call requests (new calls or handoffs) or when the channel occupancy reaches a pre-defined threshold. Borrowed channels normally belong to other sectors of the cell, neighbouring cells and in some cases from the cells of the compact pattern which are not neighbours of the borrowing cell site [DSJ97] [DSJA97]. *Channel borrowing with channel locking* strategy [TJ91] [ZY89] [DSJ97] [DSJA97] borrows a channel from an adjacent cell, but prevents (totally or partially) the use of the borrowed channel in the co-channel cells of the lender that are near to the borrower, taking into account the co-channel interference constraints. Channel locking reduces the traffic capability of the network. To overcome this penalty, the schemes have also adopted channel reassignment strategies. Channel reassignment is an *intracellular handoff,* i.e. a mobile user is asked to change its current frequency channel for another one under the control of the same base station. The channel reassignment is performed in such a way that the blocking probability is decreased and the channel reuse maximised. *Channel borrowing without locking* (CBWL) [JR94] [PAS96] strategy proposes the use of a channel of a neighbouring cell. However, the borrowed channel is used with reduced transmission power to avoid interference with the co-channel cells of the lender that are near the borrower (inside the co-channel reuse distance).

In the DCA strategy [ZY89] [OL97] [YY94] [D-RFR96], there is no pre-assignment of frequency channels to the cells of the cellular network. All frequency channels are kept in a central pool. When there is a channel request in one base station, the MSC chooses the appropriate frequency channel that gives maximum channel efficiency taking into account all the signal interference constraints. The channels are assigned for the duration of a call; after the call has finished, the channel is returned to the central pool or reallocated to a mobile user inside the same cell site that was controlling the channel before.

Finally, the FlCA strategy combines aspects of FCA and DCA strategies.

Some channel allocation schemes, using different channel assignment strategies, have special features. There are schemes that have special policies to prioritise handoff requests over new call requests. Handoff has direct impact on the perceived quality of service. One of the important factors to improve the quality of cellular service is to make handoffs nearly invisible to the user and successful. Unsuccessful handoff requests are one of the main causes of forced call termination. Therefore, some network operators select schemes that reduce the probability of forced call termination at the expense of an increase in the blocking probability of new calls. Other channel assignment schemes are specially designed for hierarchical cellular networks [E-DWS89] [O-GA98]. A hierarchical cellular network is an overlaid cellular system where clusters of microcells are covered by macrocells. The number of channels is divided between the cluster of microcells and the macrocell. In microcells, handoffs occur very often. If there are no available channels in the microcells to perform a handoff, then a channel from the macrocell can be used. The borrowed channel is released as soon as possible, or for call release or for channel reassignment. The drawback of using channel reassignment is an even higher increase in intracellular handoffs.

There is a family of channel allocation schemes that are based on the *reuse-partitioning* concept [KN96]. In the reuse-partitioning concept, each cell in the cellular network is divided into two or more concentric sub-cells called zones. The base station is located in the middle of the cell. The power level delivered to the zone increases proportionally to the distance of the zone from the base station. Therefore, the channel reuse distance for inner zones is smaller than outer zones, resulting in higher spectrum efficiency. With reuse partitioning, the two main design issues are how many channels to allocate to each zone and how the actual channel assignment is performed. Based on these issues, fixed and adaptive reuse partitioning schemes have been proposed [Hal83] [FA93]. Fixed reuse partitioning schemes allow up to three times more traffic as FCA. However, they suffer the same kind of problem than FCA when handling time-variant traffic conditions, so that adaptive schemes have been proposed to try to overcome this drawback. Reuse partitioning schemes prove to be unsuitable for microcellular systems because of the high frequency of handoffs between zones and the separation of the microcell in zones is difficult due to complicated deformed cell shapes.

2.4.1 Performance of Different Channel Allocation Schemes

The performance of a channel allocation scheme, using a determined channel assignment strategy, is measured by the following network characteristics: blocking probability of new call requests, probability of forced termination of ongoing calls, number of handoff requests, delay in channel assignment and total carried traffic.

"*In selecting a channel assignment strategy, the objective is to achieve a high degree of spectrum utilisation for a given quality of service with the least possible number of database lookups and simplest possible algorithms employed at the base station and/or MSC*" [TJ91].

2.4.1.1 FCA with Channel Borrowing and Channel Locking Schemes

The first proposed scheme in this category was the *Simple Borrowing* (SB) scheme [KN96] [TJ91]. In this scheme, when an incoming call request arrives in the cell and there are no more available nominal channels, the base station can borrow a channel from a neighbouring cell to serve the call request, provided this frequency channel does not interfere with the existing calls. The MSC supervises the borrowing procedure, following an algorithm that favours channels of cells with less traffic demand. The cells (within a distance of one or two cell units away from the borrower cell) that have a nominal channel of the same frequency as the borrowed channel will not be able to use it because of the co-channel interference. Therefore, the MSC "locks" the frequency channel in those cells. The MSC keeps a record of free, serving, borrowed and locked channels. The SB scheme gets a lower call blocking probability than FCA under light and moderate traffic conditions of the expense of additional storage requirement at the MSC and the need for database lookups. In heavy traffic conditions, the channel utilisation efficiency in SB is very much degraded because the locked channels reduce the available capacity.

Some variations of SB tried to reduce the number of locked channels by applying an exhaustive and complex search method to find the cell with the best candidate channel. However, the performance results of these schemes proved to be comparable to a much simpler SB variant scheme [KN96].

Hybrid Channel Borrowing Schemes

The main problem with the SB scheme is the absence of control in the number of channels that can be lent by a cell; this is taken into account in the hybrid channel borrowing schemes. In the *Simple Hybrid Channel Borrowing Scheme* (SHCB) [KG78] the set of channels assigned to a cell is divided into two groups, A and B. Group A channels are *local* channels that can only be used to serve call requests inside the cell. Neighbouring cells can borrow channels of group B which are "*borrowable*" channels. The ratio A:B is determined *a priori*, the optimum ratio depends on the percentage increase in the traffic density.

The *Borrowing with Channel Ordering* (BCO) scheme [ESG82] also divides the assigned nominal channels into two groups, but the local to borrowable ratio varies dynamically according to the current traffic conditions. The channels of the cell are ordered such that the first channel has the highest priority to be assigned to the next local call, and the last channel is given the highest priority to be borrowed by neighbouring cells. Each time a call is attempted, the most appropriate channel among all free channels is chosen. If the base station performs this functionality, then the

MSC needs to be informed about the resulting assignment. The MSC uses an adaptive algorithm to calculate and update each channel probability of being borrowed, based on traffic conditions. If the channel frequency is free in the three nearest co-channel cells, only then the channel is suitable for borrowing. In order to increase the availability of channels for borrowing or locking, some versions of BCO offer channel reassignment. When a high-priority channel is released, this channel is reallocated to an existing mobile user using the least-priority serving channel.

Borrowing with Directional Channel Locking (BDCL) [ZY89] is similar to BCO with channel reassignment. However, BDCL uses an efficient way to lock channels. When a channel is locked, it is locked only in the directions that would cause co-channel interference. Cells located toward the free directions can borrow or lock the channel.

The SHCB scheme performs better than FCA with light and moderate traffic. Under heavy traffic load, the point where SHBC still outperforms FCA will be dependent on the A:B ratio. BCO and BDCL schemes outperformed FCA in all kind of traffic conditions under the simulation tests realised in [ZY89]. BDCL outperforms BCO and also a dynamic channel allocation scheme called *Locally Optimised Dynamic Assignment Strategy* (LODA). By incorporating directional locking and channel reassignment BDCL obtains maximum packing of channels, increasing channel reuse. To the knowledge of the author, BDCL is still the best FCA variant scheme described in the literature.

Distributed Channel Borrowing Schemes

The FCA variant schemes with the best results (BCO and BDCL) use centralised control inside the MSC. The MSC has to keep a record of free, serving, borrowed and locked channels and to label them with updated priority. The need for a continuous up-to-date global knowledge of the entire mobile network can lead to a slow response and a heavy signalling load. To alleviate this problem, several authors have proposed modifications to make the schemes more distributed. One example is the *Distributed Load Balancing with Selective Borrowing* (D-LBSB) scheme [DSJA97] that performs better than its centralised version [DSJ97] and also outperforms other existing schemes like direct retry [Ekl86] and CBWL [JR94]. A further description of the D-LBSB scheme is given below, because in the work of this thesis, the channel allocation algorithm implemented inside the agent’s reactive layer is based on this scheme.

D-LBSB scheme

The D-LBSB scheme migrates channels from a cell with available channels (called ***“cold cell”***) to an overloaded cell called a ***“hot cell”***. Together with the borrowing channel algorithm a channel assignment strategy is used, as described later.

Initially, **C** channels are allocated to each cell in the network. The classification of a cell as hot or cold depends on its *degree of coldness* (d_c), i.e., the ratio between the number of available channels and **C**. If d_c is less or equal than a determined threshold **h**, then the cell is hot, otherwise it is cold. The determination of **h** depends on the average call arrival and termination rates of the entire cellular network, **C** and also the probability of channel borrowing rates from other cells. Typical values of **h** are 0.2, or 0.25. The mobile user in a cell is classified as *new*, *departing* or *others* according to the rules shown in Figure 2.9.

The base station periodically monitors the quality of the received signal strength (RSS) from each user through special control channels. If RSS of the user is less than a certain threshold, the user is within one of the shaded peripheral regions in the boundary of a cell.

It is supposed that the cell site transmitter is capable of transmiting information in any of the frequencies of the available spectrum. A cold cell cannot borrow channels and a hot cell cannot lend channels. Three parameters determine the suitability of a cell to be a lender, **L**: degree of *coldness* dc_L, *nearness* D_{BL} and *hot cell blockade* H_{BL}. *Nearness* is the cell distance between the borrower cell **B** and the lender cell **L**. The *hot cell blockade* is the number of hot co-channel cells of the lender that are non-co-channel cells of the borrower.

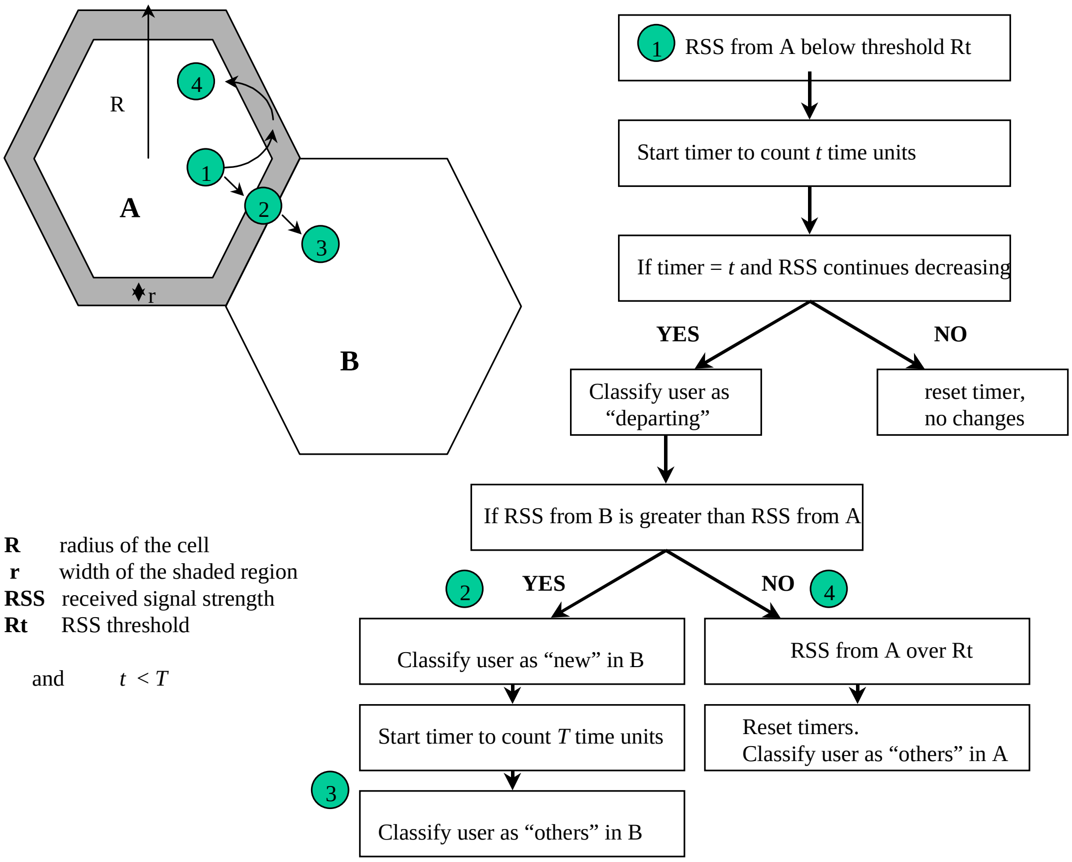


Figure 2.9: Classification of mobile users in a cell from [DSJA97]

The best lender is the cold cell in the compact pattern that maximise the value of the following function [DSJ97]:

$$F_{BL} = \frac{dc_L}{\left[\frac{D_{BL}}{R_{cp}} \cdot \frac{(1 + H_{BL})}{7} \right]} \qquad (2.20)$$

R_{cp} is the radius of the compact pattern in terms of cell distance, which implies $1 \leq D_{BL} \leq R_{cp}$, also $0 \leq H_{BL} \leq 6$ for hexagonal cellular geometry. The factors R_{cp} and 7 are used for normalisation. The value of the function is proportional to the degree of coldness of the cell and inversely proportional to D_{BL} and H_{BL}. Another criterion is the lender **L** should not be hot after lending a channel.

The base station stores the number of *departing* users heading towards the i_{th} cell in the i_{th} element of the array **NumDepart**. The objective of this information is to possibly borrow a channel from the neighbouring cell *i* and assign it to a departing user heading towards that cell. The benefits of this strategy would be: low time of channel locking, a soft handoff (a user would not have to select a different channel) and D_{BL} would have the minimum value, i.e. 1.

A hot cell needs to borrow channels until it reaches the average degree of coldness (dc_{avg}). When a cell becomes hot, the number of available channels is $h \times C$, the number of channels to be borrowed (**X**) to reach dc_{avg} can be calculated by solving the equation:

$$dc_{avg} = \frac{h \times C + X}{C} \qquad (2.21)$$

As seen in Figure 2.9, it can be assumed that **r** << **R** and the shaded portion of the cell has an area approximately given by the product of the perimeter (**p**) of the cell times **r**. The rate of call arrivals is assumed to follow a uniform spatial distribution within the cell. Supposing **K** is the average density of mobile users making calls in a cell, it is easy to derive that the number of departing users is $K \times p \times r$. If the use of the borrowed channel is confined to departing users only, then $K \times p \times r \geq X$ and an approximation to **r** is given by:

$$r \geq \frac{C \times (dc_{avg} - h)}{K \times p} \qquad (2.22)$$

The algorithm is only initialised when the cell becomes hot. The data structure needed for the execution of the algorithm is:

Each cell site keeps as local parameters its **NCC** (set of non-co-channel cells, considering its compact pattern), **CC** (set of co-channel cells), dc, H_{NCC} (set of hot non-co-channel cells), H_{CC} (set of hot co-channel cells). The cell knows the first two parameters by the configuration of the network; dc is updated internally according to the current availability of channels. H_{CC} is updated independently of the algorithm as a cell informs all its co-channel cells when it changes its state (hot or cold). H_{NCC} is computed at the beginning of the algorithm execution.

- As global parameters the cell sites use dc_{avg}, **h** and **C**. As soon as the cell becomes hot, it computes dc_{avg}, by requesting for all cells in the system their dc. The parameter **h** is computed once by the MSC for all cells. **C** is known at the configuration phase. These global parameters will be used in the computation of the number of channels to be borrowed and the width of the shaded area.

Algorithm Body

Initialisation:

- The cell **B** that has just become hot inquires the other cells about their dc, by broadcasting a request message and computes dc_{avg} and H_{NCC}.
- With **C**, dc_{avg}, **h** and **K**, it computes **X** and **r**.
- With **r**, it computes the array **NumDepart**.

Main body:

1. **B** sends messages to the cold neighbouring cells **L** for which **NumDepart**[L] is greater than zero. The message contains **NCC**, H_{NCC}, $D_{BL} = 1$ and requests the cell to compute F_{BL}. Each **L** cell computes H_{BL} and F_{BL} and sends the information back to **B**.
2. **B** orders the **L** cells by decreasing values of F_{BL} and selects the cell with highest F_{BL}. The selected lender computes the set of its co-channels, which are non-co-channel with **B** by comparing **NCC** with its own **CC**.
3. Channels start to be borrowed from the lender cell until the number of borrowed channels is equal to **NumDepart**[L] or the basic criteria is violated, i.e., the lender become nearly hot. After lending the channel, the lender cell instructs its co-channel cells, which are non-co-channel cells of **B,** to lock that frequency channel. The same procedure is executed for the other cells in the listed order until the number **X** of borrowed channels is reached and the algorithm is terminated, or the list of cells is exhausted, whereupon it gives to step 4.
4. **B** sends messages requesting F_{BL}. to the **L’** cells (all cells in its compact pattern excluding the neighbouring cells). The cold **L’** cells will answer.
5. **B** selects **L’** with highest F_{BL}, if the basic criteria is not violated. **L’** computes the set of its co-channels which are non-co-channel with **B**. Steps 4 and 5 are repeated until **X** is reached.

Channel Reassignment Strategy

- The set of available channels can be divided into local and borrowed channels. Hot cells have both, cold cells only local channels. The channel demand in this scheme is classified into four classes:
- Class 1 requests have the highest priority to receive a channel; these are handoff requests. This strategy tries to minimise the probability of disrupting ongoing calls.

- Class 2 requests are the channel requests by originating calls.
- Class 3 requests are requests for channel re-assignments; they are requested by a cell site function that monitors the state of the channels. The re-assignments are divided in two types: The reassignments of type 1 are for re-assigning a new or other user using a borrowed channel to a local channel, if this local channel is not used to satisfy class 1 and 2 demands. The reassignments of type 2 are for reassigning a departing user using a local channel to a borrowed channel, if the borrowed channel is not used to satisfy class 1 and 2 demands. Requests of class 3 are for reassignment of type 1.
- Class 4 requests are re-assignments requests of type 2.

The channel assignment algorithm prioritises the channel requests according to the class that they belong to. The flow of the algorithm can be seen in Figure 2.10. In case of multiple requests of the same class, the algorithm selects one randomly to receive the channel.

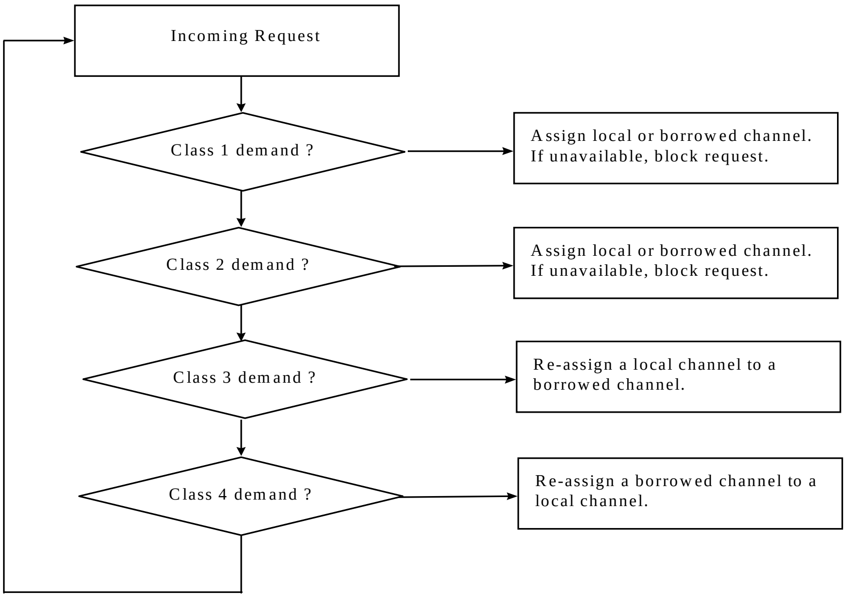


Figure 2.10: Channel assignment algorithm from [DSJA97]

Although the authors of the D-LBSB scheme claim the channel assignment strategy prioritises handoff requests and channel requests are classified differently, the execution of the channel assignment algorithm does not prioritise handoff requests. Each request is treated independently and discarded if blocked. Therefore handoff requests and incoming call requests have the same priority, because they are treated equally inside the algorithm.

In D-LBSB, the borrowing algorithm is not executed every time a call or handoff request is made and there are no more available channels to accommodate the request as is done in BCO and BDCL schemes. It is triggered *before* the nominal channels are all used, once ***h*** is reached. Moreover, it does not get only one channel, but a certain number of channels (**X**), the actual number depending on the average traffic load of the whole network. The authors of the D-LBSB scheme claim that this is a load balancing strategy.

D-LBSB does not perform as well as BCO or BDCL, however it is less complex and it proves to be much faster as the load of traffic increases compared with its centralised version [DSJA97].

2.4.1.2 DCA Schemes

The main idea of DCA schemes is to assign a channel to a call request with minimum cost, but respecting the signal interference constraints. The cost is evaluated by a cost function. The cost function can be formulated taking into account the future blocking probability in the vicinity of the cell, the usage frequency of the candidate channel, the reuse distance, channel occupancy distribution under current traffic conditions, radio channel measurements of individual mobile users, average blocking probability of the system and so on [KN96]. The differentiation factor in DCA schemes is the formulation of the cost function.

Most DCA schemes work on a call-by-call basis, searching the minimum cost channel for the current call request, based on the current traffic conditions. Some DCA schemes are adaptive, i.e. the assignment decision also takes into consideration past traffic conditions. DCA schemes can also be centralised or distributed. Centralised DCA schemes can produce near-optimum channel assignment, but the complexity and the high centralisation overhead can seriously compromise their applicability in real systems. Distributed DCA schemes that provide near optimum channel assignment also present excessive exchange of status information between cells. The distributed DCA schemes, which offer sub-optimum channel assignment and present smaller message exchange overhead, are those based on signal strength measurements with inter-cell information sharing. They are applicable to microcellular systems where inter-cell information sharing by interference measurement is possible.

Centralised DCA Schemes

Locally Optimised Dynamic Assignment (LODA) scheme is a DCA scheme [ZY89] whose cost function is based on the future blocking probability in the vicinity of the cell. In [ZY89] simulation results comparing LODA with BDCL shows that BDCL has better performance.

Several DCA schemes formulate a cost function that maximises the channel efficiency by optimising the reuse channel distance packing. These schemes perform well in light and moderate traffic conditions, but in heavy traffic load they are not able to maximise the channel reuse because the best candidates most probably are already serving call requests. Channel reassignment can be used to pack the co-channels cells. Everit and Manfield [EM89] proposed an ideal DCA scheme called *Maximum Packing* (MP). In MP, a new call will be blocked only if there is no possible reallocation of channels to allow the call to be carried. The MP scheme is impractical for implementation because it requires system-wide information and the complexity of searching all possible reallocations is computationally hard. The *Compact Pattern Based Dynamic Channel Assignment* (CP-based DCA) scheme [YY94] does reduce the complexity of the search and limits the number of channel reassignments, but still presents high centralised overhead. CP-based DCA keeps the co-channel cells of any channel to a compact pattern whenever possible. CP-based DCA consists of two phases: channel allocation and channel packing. Channel allocation is used to assign an optimal idle channel to a new call. Channel packing is responsible for the restoration of the compact pattern and is performed only when a compact pattern channel is released. The number of channels reassigned per released call in channel packing is at most one. CP-based DCA shows better performance than FCA and BDCL in the simulations performed in [YY94].

Distributed DCA Schemes

The distributed DCA schemes are normally cell-based schemes or signal strength measurement based schemes [KN96].

In cell based schemes, the channel assignment is performed by the base station. The base station (and not the MSC as in the centralised schemes) is responsible for keeping information about current available channels in the vicinity. *The Local Packing Dynamic Distributed Channel Assignment* (LP-DDCA) scheme [LC93] uses an augmented channel occupancy (ACO) matrix for channel assignment. The ACO matrix contains all the local and vicinity information needed for the selection of a channel. The base station keeps the ACO matrix updated. *The LP-DDCA with Adjacent Channel Interference Constraint* [LC94] takes into account this kind of channel interference when selecting a channel from the ACO matrix. Both schemes provide near optimum channel assignments, but cause excessive exchange of status information between cells.

The signal strength measurement based schemes, also known as interference adaptation schemes, use the location of the mobile users to maximise the packing of channels. The reason for that is because depending on the location of the mobile user, the reuse channel distance used for the selection of the channel may be greater than that actually needed. The mobile users are able to measure the amount of co-channel interference to determine the reusability of a channel. When mobile users and base stations offer this functionality maximum channel packing could be achieved. The former *Digital European Cordless Telecommunication* and now known as *Digitally Enhanced Cordless Telephone* (DECT) system uses this principle [KN96].

Another signal strength measurement based scheme is the *Channel Segregation* (CS) scheme [FA91] [AA93]. This scheme is a self-organised DCA. Each base station scans channels when selecting an available channel with acceptable signal interference. Each base station will attribute to each channel a probability of channel selection, P(i). The channel “selectability” order is performed independently by each base station and is reviewed through learning methods. For each call request, the base station selects the channel with highest P(i). Then, the base station needs to check if the use of that channel is possible by measuring its power level. If the power level is good enough (acceptable interference) then this channel is considered idle, allocated to serve the call request and its “selectability” increased. If not, the channel is busy and P(i) is decreased. If all channels are busy the call is blocked. The CS scheme is autonomous and adaptive to changes in traffic load. Simulation results [FA91] shows that blocking probability is greatly reduced compared to FCA and DCA schemes and quickly reach a sub-optimum channel allocation. The presence of many local optimum allocations makes the convergence to an optimum channel allocation prohibitive. CS uses the channels efficiently and reduces the need for channel reallocation due to interference. CS is a good solution for TDMA/FDMA microcellular networks [AA93].

2.4.1.3 Flexible Channel Assignment Schemes

In [TI88], the cell sites have a sufficient number of pre-assigned nominal channels to accommodate light traffic. The remaining channels are kept in a central pool and assigned to cell sites in need. The dynamic assignment can have a scheduled or predictive approach. In the scheduled approach, the assignment of channels is made at determined peaks of traffic, following a determined traffic distribution. In the predictive approach, the traffic intensity is measured constantly at all cells and the MSC can reallocate the channels at any time. The ratio of fixed and dynamic channels is a significant parameter that defines the performance of the system. For heavy traffic loads FCA gives better blocking probability than FlCA schemes, again the explanation coming from the fact that FCA will make better use of the minimum reuse distance than FLCA schemes.

Another possible combination of FCA and DCA strategies is to use a DCA scheme for light and moderate traffic loads and FCA in heavy traffic conditions [KN96].

2.4.1.4 Performance Summary of Channel Allocation schemes

Fixed channel assignment (FCA) is too limiting for mobile networks and several strategies have been proposed to maximise frequency channel allocation and minimise call blocking probability. DCA schemes perform better under low traffic intensity; modified FCA schemes have superior performance in high traffic loads. DCA schemes use channels more efficiently (better trunking efficiency) and for the same blocking rate have a lower forced call termination than FCA-based schemes. However, the near-optimum channel allocation is at the expense of high overheads through its use of centralised allocation schemes. This overhead means that such schemes are not practicable for large networks. Distributed DCA schemes with limited inter-cell communication suffer less overhead, but lead to sub-optimum allocations. Such schemes are being proposed for microcellular systems as this cell structure allows inter-cell information sharing by interference measurements and passive non-intrusive monitoring at each base station (busy/idle status of the carriers) [WS94]. For macrocellular systems, where explicit communication is needed, FCA with channel borrowing offers good results and less computational complexity than DCA. However, those FCA variant schemes with best results use centralised control inside the Mobile Switching Centre (MSC). Although they are less complex than DCA schemes, there is still a need to maintain an up-to-date global knowledge of the entire mobile network, leading to a slow response and a heavy signalling load. To alleviate this problem, several authors have proposed modifications to make the schemes more distributed.

Chapter 3 Agents and Multi-Agent Systems

3.1 Introduction

Agent based technology is an interdisciplinary area of research and it has received special attention from the research community since the beginning of the 1990's. The definition of an agent is controversial, but in general terms, an agent can be described as a hardware or software system with social ability that performs tasks with specific aims in a complex and dynamic environment. Agents are capable of autonomous actions to pursue their objectives, despite the occurrence of expected or limited unexpected events. This thesis interests in the application of intelligent co-operative software agents.

The earliest concept of a software agent was dated in the 1970s with the concurrent *Actor model* of Carl Hewitt, 1977 [Hew77]. The Actor model was defined as a self-contained, interactive and concurrently executing object with some encapsulated internal state and capable of answering messages from other similar objects (actors) [Nwa96].

The characteristics of software agents have inherited aspects from distributed computing, such as; modularity, speed (parallelism) and reliability (redundancy), and from Artificial Intelligence (AI) in the knowledge level; easier maintenance, reusability and platform independence.

Agents can present different degrees of complexity, normally known as *granularity*. Simple agents with little or no intelligence regarding their behaviour are called *reactive agents*. More complex agents such as those possessing symbolic internal models, capable of knowing their environment and able to reason about their goals are called *cognitive* or *deliberative agents*. As the complexity of agents can be so broad, it is difficult to find the boundary between an agent and other types of control software. However, agents do present special properties that distinguish them from other software systems. The main properties that software agents exhibit are *autonomy*, *social ability* and *reactivity*.

- *Autonomy* is the ability to have control over its own actions and states. An agent is able to make decisions and complete actions based on its internal representation of the world without direct intervention of a human or a central entity [WJ95a].
- *Social Ability* is the ability to interact with other agents (humans or other homogeneous or heterogeneous software agents) via some kind of communication language in a co-ordinated manner. Agents may co-operate in order to solve a problem or to achieve a task. The terms and conditions of the co-operation may be negotiated at runtime.
- *Reactivity* is the ability to perceive changes in the environment and react timely and appropriately.

Software agents may also present other properties such as *pro-activity*, *learning*, *adaptability* or *mobility*.

- *Pro-activity* is the ability to plan ahead and take the initiative to perform actions that will contribute to the goal achievement without waiting for external instructions or only responding to events in the environment.
- *Learning*, ideally, would be the ability to improve its awareness and to alter actions as the agent reacts with the environment and/or with other agents, in order to avoid past mistakes or increase performance over time.
- *Adaptability* is the ability to adapt to changes in the environment, in order to continue to pursue its objectives. Learning is one of the factors that allow agents to have an adaptive behaviour.
- *Mobility* is the ability to move around a network [WJ95a].

Nwana [Nwa96] has a different approach to defining agents by classifying them using different criteria: mobility, granularity, predominant attributes and roles. In the first two criteria, the agents can be *mobile* or *static* and *deliberative* or *reactive*. Co-operation, autonomy and learning are the three basic attributes of agents and from them three[♣] types of agents can be identified: *collaborative agents*, *interface agents* and *smart agents*, as can be seen in Figure 3.1. Each type giving more emphasis to some attributes. Ideally, agents should be smart agents emphasising the three attributes equally, but the applicability of the agent to certain problem will dictate which type of agent is more suitable for each case.

[♣] In Nwana's classification there is one more type of agent: collaborative learning agent, however it does not present the attribute of autonomy and therefore it is not considered in this thesis as a type of agent.

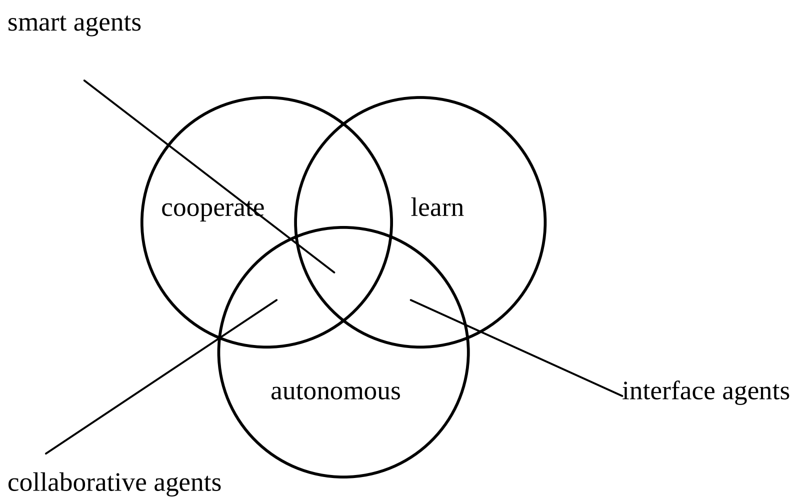


Figure 3.1: Definition of agent types based in their predominant features [Nwa96]

The last criterion to classify agents is by their roles, for example, World Wide Web information agents, which usually exploits Internet search engines such as WebCrawlers, Lycos and Spiders.

Of course, the total classification of an agent would be a combination of these criteria. For example, an agent could be a mobile deliberative information agent or a static reactive interface agent, etc. Moreover, agents can present more than one feature of the same criterion, for example, the existence of a reactive part and a deliberative part inside a single agent. These agents would be classified as hybrid agents.

Nwana’s classification [Nwa96] demonstrates the broad spectrum that agents can fit in. However, his classification is informal and more related to intelligent collaborative agents, which are the type applied in this thesis. Most of the recent research in multi-agent systems considers the application of *self-interested* agents. Self-interested agents are not willing to collaborate or be *benevolent* with each other, but they may negotiate and interact in a co-ordinated way in order to maximise their own expectations or goals and the global behaviour may achieve an optimised solution.

3.2 Agent Architectures

All the properties that have been described that an agent must or may have are embedded inside its architecture. Architecture represents the move from specification to implementation. Agent architecture has also been evolving since the 1970's.

Most of the precursor agent architectures were based on Artificial Intelligence (AI) planning systems like STRIPS [FHN71]. The agent models were based on Simon and Newell's physical symbol system hypothesis [NS76]. The physical symbol system is a physically realisable set of symbols that can be combined to form structures. Symbols represent physical entities. Processes inside the system operate on the symbols according to defined sets of instructions. In this way, the physical symbol system hypothesis states that the system is capable of general intelligent action. The physical symbol system hypothesis is the step stone towards deliberative agents whose complete definition is, now, easier to understand. A *deliberative agent or agent architecture* possesses an explicitly represented, symbolic model of the world, in which decisions (for example about what actions to perform) are made via logical (or at least pseudo-logical) reasoning, based on pattern matching and symbolic manipulation [WJ95b]. Over the past few years, research has explored a deliberative model of agents based on *beliefs, desires and intentions* (BDI) The architectures following this paradigm are known as BDI architectures [Bra87] [RG91].

The capability of general intelligent action towards a goal is very attractive, with it the agent's behaviour could be optimal and highly adaptable to different contingencies. However, logical representation of belief, desire, time and so on tends to be difficult and the reasoning hardly tractable. Symbol manipulation in general demands high computational complexity and consequently long response time. Response time is a key factor in real time systems.

To overcome the problems of deliberative architectures, a new school of thought emerged in mid-1980's, that denies the need of symbolic representation of the world and symbolic reasoning to achieve specific goals. In these architectures, agents have very limited amount of information and their run-time decisions are based on sensory input and simple situation-action rules. Agents and architectures of this type are called *reactive, situated* or *behaviour-based*. The aim of reactive architectures is to have fast response and robust behaviour instead of optimal behaviour. Implementations of reactive architectures were very successful in building robots for room explorations, map building, route planning [Bro90], simple games [AC87] or puzzles [FJ91]. However, these architectures present a limited scope, and they are not able to execute complex tasks that depend on long-term goals or co-operation.

Looking at the drawbacks of reactive and deliberative architectures, researchers have suggested that a combination of both architectures would be more suitable for building agents. Hybrid architectures, as they are known, are designed to respond rapidly to changes in the environment and also to provide means to achieve long-term goals. Most of the hybrid architectures proposed are layered architectures. The different functionalities and goals are arranged in different layers that interact in a well-defined control interface. Ferguson developed the *Touring Machines* architecture [Fer95]. The internal architecture of the agent consists of a *reactive* layer, a *planning* layer and a *modelling* layer. The layers operate concurrently and perform constrained navigation in a dynamic environment. The reactive layer gives fast response to specific environment stimuli. The planning layer is responsible for the generation and execution of the goal of the agent. The modelling layer enables the agent to change its plans when conflict arises, for example, to avoid collision with other agents. This layer includes information about other agents in the environment. However, the knowledge about other agents is used to guide local decisions of the agent. Touring Machine architecture does not support co-operation based on communication. *INTEgration of Reactive behaviour and RAtional Planning* (INTERRAP) [Mül96] is another hybrid architecture, similar to the Touring Machines. It is also composed of three layers, but it supports co-operative planning and problem solving and presents a more elaborate layer control interface. INTERRAP was designed to have proper and timely reaction to unexpected events, long term actions based on goals and to cope with other agent interactions, i.e. to fulfil efficiently the real time requirements of a dynamic environment that can be adapted to the channel allocation problem in mobile networks, as this thesis will later describe. In the next sub-section INTERRAP architecture is described in more detail.

3.2.1 INTERRAP: a Hybrid Agent Architecture

INTERRAP [Mül96] is an agent architecture composed of a set of hierarchical control layers, a knowledge base that supports the representation of different abstraction levels of knowledge and a well defined control architecture that ensures coherent interaction among the control layers. Each layer has different functionality which working concurrently completes the global behaviour of the agent. The three layers in INTERRAP are:

- *Behaviour-Based Layer* (BBL): responsible for reactivity and procedural knowledge for routine tasks.
- *Local Planning Layer* (LPL): produces goal directed behaviour. This layer provides the means to reason about local goals and how to achieve them.
- *Co-operative Planning Layer* (CPL): enables reasoning about global goals and about other agents and supports co-ordinated action with other agents.

Figure 3.2 from [Mül96] assembles the mental state of the agent, which is composed, from different components. The agent's *perception* of the environment is performed by sensors whose information is manipulated by a *belief generator /reviser* that transforms the current perception in new *beliefs*, or changes the attribute values of existing *beliefs*.

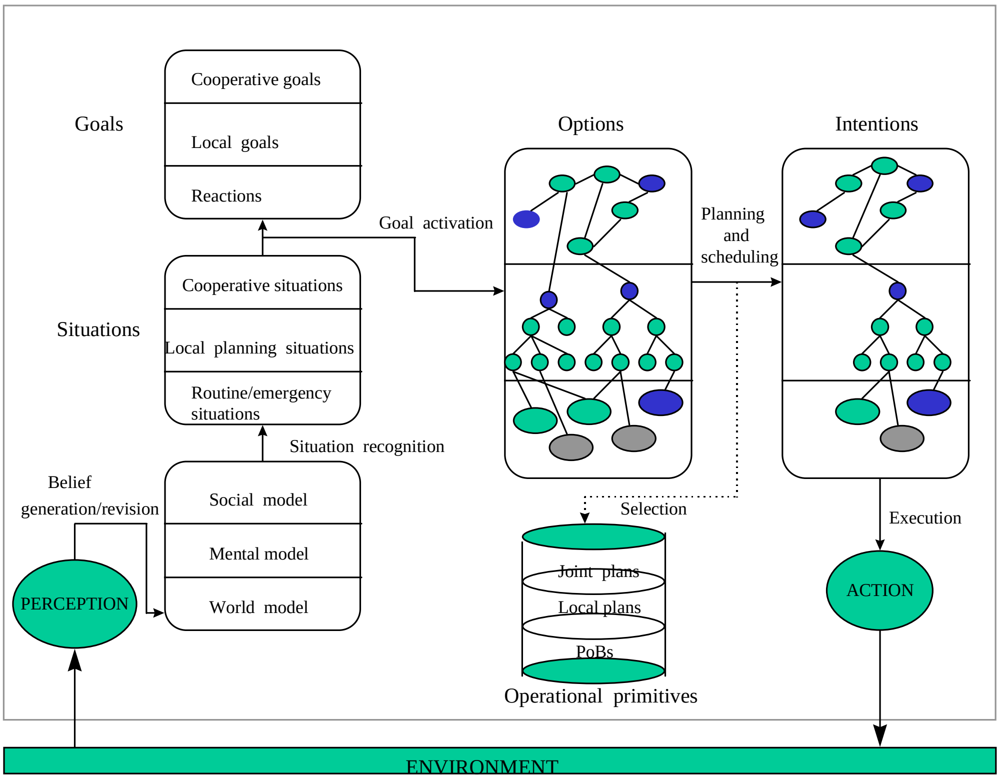


Figure 3.2: The conceptual agent model

The informational state of the agent is its *set of beliefs*. The knowledge base in which the beliefs are placed is divided in three different levels of abstraction: the *world* model, the *mental* model and the *social* model. The world model contains beliefs about the environment. The mental model has beliefs about the agent itself. The social model holds beliefs about other agents.

Beliefs can be suitable for describing the relationship between the informational state and the motivational state of the agent, i.e. what the agent knows and what the agent wants to achieve, however beliefs can be numerous and too unstructured. As actions are taken when specific situations happen, a *situation recognition function* analyses the current beliefs and extracts the subsets of the agent's beliefs that correspond to relevant situations. The situation recognition function enables the agent to identify the need for activity.

Situations are better structured because they take into account physical aspects of the environment and they make the computation of the link between agent's perception and goal easier. Situations are also divided in three layers: *routine/emergency* situations, *local planning* situations and *co-operative* situations.

The agent's *goals* are distinctly classified as *reaction* goals, *local* goals and *co-operative* goals. Reaction goals are short-term goals triggered by external events. Local goals are the local desires or long term goals that the agent pursues. Co-operative goals are the goals of the agent when it is interacting with other agents and normally these goals are shared among agents.

When a situation is recognised the motivational state of the agent changes. The *goal activation function* extracts the goals of the agent, which are the current *options* given determined situations. In the case of the behaviour-based layer, the *options* are basic execution procedures called patterns of behaviour (PoBs). PoBs are hard-wired situation-action procedures highly dependent of the environment domain. For the local and co-operative layers, the *options* are more complex and the goal activation function may need to construct a goal state based on a combination of situations. The goal state is used by the planning mechanism.

The *planning mechanism* is responsible for deciding what to do to achieve the goals. It relies on the execution of *operational primitives* (OPs) for the achievement of the goals. OPs inside the behaviour-based layer are the body of PoBs. At the local planning layer the OPs enable the means-ends reasoning about how to achieve the goal and the consequent generation of the local plan. Similarly, at the co-operative layer OPs generate the joint plan whose execution satisfies the agents' shared goals.

The *scheduler* decides when to do what. It merges sub-plans into one executable schedule. One of the problems in this stage is to cope with the limitations of computational and physical resources and resolve the incompatibilities and constraints among plans. The scheduler uses a priority mechanism to execute PoBs, to assure fast response to urgent reactions. At the Local Planning Layer, it involves the sequencing of non-linear plans, assigning time constraints for the execution of plan steps and modifying plans in case of incompatibility. At the Co-operative Planning Layer, it schedules concurrent negotiations.

The *execution* function is responsible for a correct and timely implementation of the tasks. It also interacts with the situation recognition function to guarantee a consistent evolution of the system. The execution function activates the PoBs that make access to the resources in the environment. Plan steps are mapped into PoBs and executed. It is also responsible for the correct execution of negotiation protocols and commitments of joint plans.

3.2.1.1 Control Architecture

The INTERRAP architecture is based on a layered control, a layered knowledge base, bottom-up activation and a top-down execution as showed in Figure 3.3.

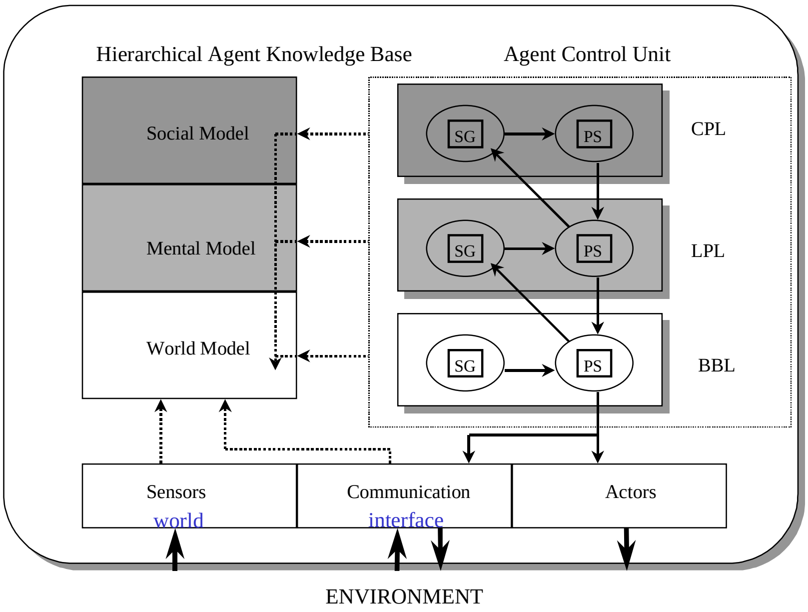


Figure 3.3: The architecture of the agent

The world interface is composed of three subsystems: sensorial subsystem, actor subsystem and communication subsystem.

An object represents the sensorial subsystem: the “sensor”. Its functionality is sensing the environment. The sensorial subsystem has methods to calibrate the sensor, enable and disable sensor activity and for reading the current values in the perception buffer.

All the actions performed in the environment are done by the actor subsystem. The actors control the atomic and continuous physical actions of the agent. Atomic actions are single tasks terminated by success or failure. Continuous actions are control processes that run until they are explicitly finished, suspended or deactivated. An object class called “actor” represents the actor subsystem. Actors have methods to calibrate actions, execute atomic actions and activate, suspend or deactivate continuous actions.

Finally, the communication subsystem provides the functionality of sending messages to or receiving from other agents. It is represented by two classes: *send_queue*, which is a subclass of actor, because sending a message is considered an action on the environment. *Receive_queue* is a subclass of sensor. Receive a message is considered a sensing process. Messages are represented by tuples**:**

Msg = (*Id, Sdr, Recp, Ref, Type, Content*)

where ***Id*** is a unique identifier, ***Sdr*** denotes the sender, ***Recp*** denotes the recipient, ***Ref*** (optional) is a reference to a message-id, ***Type*** is one of the message types, ***Content*** is the actual content of the message.

The knowledge base is layered and represents the structured informational state of the agent with situations, beliefs and goals. It can be represented by an Assertional Knowledge Base (AKB). The elements of AKB are: Concepts (classes of individuals), types, attributes, features (attributes of a concept that cannot be changed, but can be assigned an initial value) and relations (relating concepts each other by cross-product operator #). The AKB offers assertional, retrieval and active information services. Assertional services allow asserting new beliefs into the knowledge base, creating instances of concepts and relations, but it can change attributes of concepts too. Retrieval services provide access to beliefs stored in the AKB. Active information services access information upon demand. When an active information service is requested, it starts a monitoring process that recognises specific changes in the knowledge base and sends this information to the requesting process.

The agent control unit is composed of three layers. Each layer has a uniform structure as illustrated in Figure 3.4. Each layer implements a generic control cycle performing functions of situation recognition, goal activation, planning, scheduling and execution.

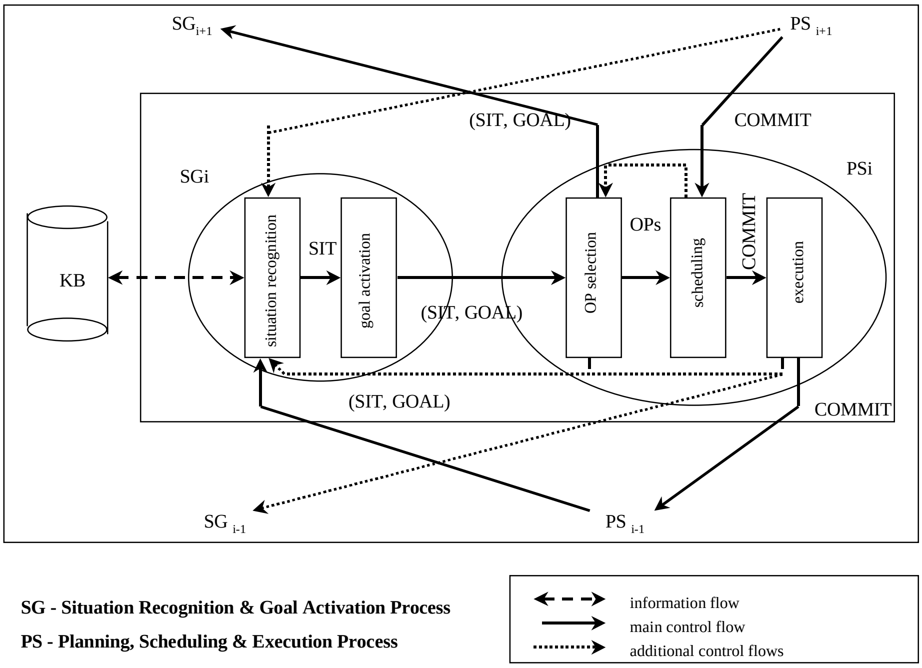


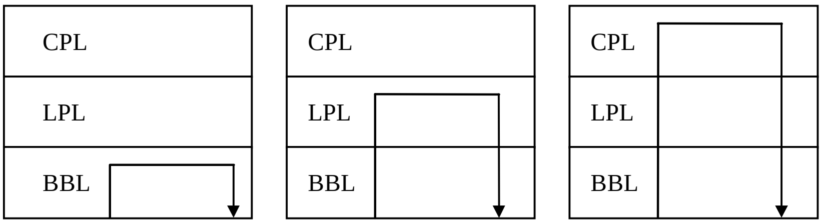
Figure 3.4: INTERRAP control layer

As seen in Figure 3.4, the process ***SG*** covers the situation recognition function and goal activation function; the process ***PS*** covers the planning, scheduling and execution functions. An object defines the generic control cycle of INTERRAP. The main method of the object describes the sense-recognise-decide-act cycle. In each loop of the cycle, the current beliefs are scanned for new situations, the situations and the activation request by the next lower layer computes the new options. The planning, responsible for the decision making, selects operational primitives and updates the intention structure, finally it decides if the goal will be dealt within this layer or an activation request will be issued to the next higher layer to take care of the goal. If the goal is being dealt within the layer, the planning function decides what commitments need to be made to achieve the goal and passes to the scheduling function. After the commitments being scheduled it starts the execution phase. The effects of the actions taking place in the execution phase are monitored by a primitive in the situation-recognition function, as can be seen in Figure 3.4 by the link between execution function and the situation recognition function.

Also planning and scheduling are interleaved. If the planner has created an intention that cannot be scheduled, the scheduler forces re-planning. INTERRAP has two links between situation recognition and planning to deal with the cost of sensor information and disagreement between situation recognition and planning. First from the execution module to the situation recognition, which serves to enable and disable monitoring conditions (in SGi and SGi+1); second, there is a link from the planner to situation recognition, which allows the planner to order additional information from SG when, is necessary.

The behaviour of the agent results from the interplay among individual control layers. INTERRAP presents a bottom-up control direction and a top-down control direction. The first type of control flow is performed by the upward activation request mechanism. This mechanism ensures that situations requiring fast response are handled by the Behaviour-Based Layer, whereas other situations that need more complex planning are shifted upward until they reache a competent layer to solve the problem. Following the representation in Figure 3.4, if *PSi* decides not to be competent for dealing with the *SGi* pair; it sends an upward activation request to the next higher layer. The information provided by layer *i* and the additional information available in layer *i+1* can produce a suitable goal description for PS_{i+1}. The advantage of the bottom-up approach is that lowers layers do not need to know about the capabilities of higher layers. The top-down control flow corresponds to the acting of the agent. The PSi process co-ordinates its activities by posting commitments to the next lower layer. If the commitments can be incorporated in the scheduling process of the lower layer, a successful acknowledgement is sent to the higher layer; if not, a failure report is sent to the higher layer indicating that a re-plan needs to be performed. Commitments between the Co-operative Planning Layer and the Local Planning Layer are partial plans describing the role of the agent in the joint plan. The Local Planning Layer commits to the execution of procedures of PoBs posted to the Behaviour-Based Layer and finally, the Behaviour-Based Layer commitments cause the execution of actions in the world interface.

The two basic control directions determine three generic control paths to deal with different classes of problems or tasks: the reactive path, the local planning path and the co-operative path (Figure 3.5). The reactive path deals with emergency situations and those ones recognised and handled by routine PoBs. The local planning path treats more complex situations that cannot be dealt with the Behaviour-Based Layer. The co-operative control path deals with situations that require co-ordination with other agents.



a) Reactive path

b) Local planning path

c) Co-operative path

CPL – Co-operative Planning Layer
LPL – Local Planning Layer
BBL – Behaviour-Based Layer

Figure 3.5: Generic control paths

Normally, the activities between layers are neither clearly separated nor follow the strict temporal ordering. There are interleavings of planning and execution as illustrated in Figure 3.6. Planning is an ongoing incremental process, and future planning decisions depend on the outcome of current PoB calls.

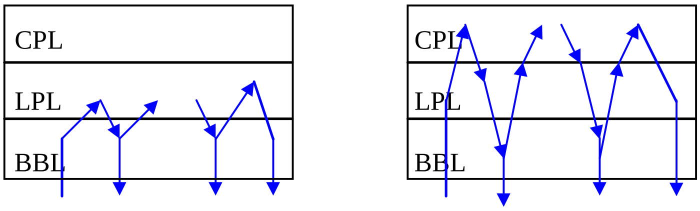


Figure 3.6: Instances of local planning path and co-operative path

INTERRAP has additional inter-layer co-ordination. The Local Planning Layer can enable and disable PoBs as a mechanism to control the activity of the Behaviour-Based Layer, for example, in the case where the plan has changed. Another possibility is for the Behaviour-Based Layer to the Local Planning Layer to devise a plan for a goal-situation pair, to evaluate or interpret a given plan and to stop activity regarding an earlier request. The Co-operative Planning Layer can ask the Local Planning Layer to interpret or evaluate a plan or stop activity. In the same way the Local Planning Layer can ask the Co-operative Planning Layer to devise a plan given a goal situation description and to evaluate and interpret a joint plan, maybe proposed by other agent.

The brief overview of INTERRAP agent architecture can explain why agents have been applied in the solutions of telecommunication problems. Some interesting problems being investigated are related to network management of ATM (*Asynchronous Transfer Mode*) networks [Som96] [HB98], congestion control and load balancing in switched networks using mobile agents [AS94] [SHB97], reestablishment of interrupted connections in telecommunication networks [GM96]. The application of agents in telecommunication problems is an emerging area of research, which is having great repercussions not only in the academic environment but also in industry [WV98].

3.3 Multi-Agent Systems

This thesis is concerned with the application of agents within telecommunications systems that do not work in isolation, but as a community. A community of agents composes a *multi-agent system*. A multi-agent system can be defined, therefore, as a group of agents with specific roles in an organisational structure [Mül96]. The agents interact with the environment and with each other in a co-ordinated way, as collaborators or competitors, seeking to fulfil the local or global aims of the organisation.

A multi-agent system applied to a specific environment needs to be seen on two levels:

- The first is the co-ordination level of the system and defines how the agents communicate between themselves, co-ordinate their activities and negotiate in a joint plan or when conflict occurs.
- The second is the internal level of the components of the system and is related to the internal agent architecture.

Distributed Artificial Intelligence (DAI) has studied the behaviour of groups of intelligent systems. More precisely DAI has originated a sub-area for the study of groups of agents: the *multi-agent systems* (MAS). However, the study of the behaviour of a group of agents interacting among each other in a system demands the consideration of other research areas; like cognitive psychology, control theory and economy studies in market-control. All these areas can be considered as parental disciplines of multi-agent systems.

Distributed Problem Solvers (DPS) can also be considered a parental discipline of multi-agent systems, because many of the features presented in distributed problem solving techniques are also applied to multi-agent systems. However, some researchers make a clear separation between DPS and MAS, considering parallel disciplines derived from DAI. In the point of view of those researchers DPS are only composed of collaborative autonomous software and MAS composed of self-interested agents. Whatever is the case, the fact is that several co-ordination techniques and algorithms used in distributed problem solving have been used in the co-ordination level of multi-agent systems.

Market-based control is a co-ordination technique developed for distributed problem solving and is now being applied in co-ordination of multi-agent systems involved in task and resource allocation in distributed environments. The precursor co-ordination technique is the *Contract-Net Protocol* (CNP) [Smith88]. In the Contract-Net Protocol the agents can play two different roles: manager or contractor. When an agent has a task to perform, but is not capable of doing so, it becomes a manager and looks for the most suitable contractor by announcing a contract to other agents and selecting the best one according to the bids received. The interesting feature is that a contractor can ask for sub-contractors (becoming a manager) to help it to perform its task.

The Contract-Net Protocol when applied to a specific domain can use several different kinds of algorithms to solve a problem or to solve different parts of a problem.

3.3.1 Market-based Control

In the past 40 years, economists have studied how resources in an economy can be optimally shared. They have developed normative models to describe the economy and decentralised optimisation methods to balance the economy. Distributed systems present several similarities with economic systems and several researchers started to apply the models and methods developed within the field of mathematical economics to construct similar models for distributed problem solving [Clearw96] [HSS80]

[FNY89] [KS89] [YA96]. The microeconomic approach is commonly known as *market-based control*. In a market-based control system, the resources to be allocated are the commodities of the economy. The stakeholders are now trading agents. Each agent presents certain preferences and behaviour, agents trade between each other in order to maximise their preferences. The algorithm that agents use to adjust their resource allocation is the auction protocol. Market-based approaches can be divided into two categories: *price-oriented* and *resource-oriented*. In a price-oriented approach the resources are associated with prices. An initial allocation of resources is made and an initial price is chosen. In each iteration of the algorithm, prices change to accommodate the demand for resources. The iterations continue until the total demand reaches the total amount of resource available, at which the resulting final allocation is provably *Pareto optimal*. In a resource-oriented approach, the agent knows how much resource they need each iteration. Each agent computes the marginal value of resource it requires given its current allocation of the resource (mathematically, the agent computes the partial derivative of its utility function with respect of its current amount of resource). The agents send their marginal values to all other agents or to an auctioneer. The agents with an above average marginal utility receive resource, the agents with a below average marginal utility, transfer resource. The price and resource-oriented approaches have different properties. For distributed systems, the most important property is the feasibility of allocation during the convergence process. Price-oriented algorithms may not present feasible intermediate allocations until equilibrium is reached; therefore the pricing process must converge before the resources can be allocated. In resource-oriented the intermediate allocations are feasible if the initial allocation is feasible. Moreover, successive iterations of the algorithm result in increasing system wide utility [KS89], if the formulated problem is feasible and monotonous.

3.4 Multi-Agent Systems Applicability

Frequency channel allocation in mobile networks is a control problem of a complex system with specific goals (minimising the blocking probability of the network and minimising the delay in channel assignment). As described in [HBWC99] the control features of a complex system are commonly agreed to be the:

- Co-ordination of multiple and sometimes conflicting goals.

- Management of multiple inputs some of that may be incomplete or inconsistent.
- Adaptation of the control strategy when the environment changes drastically or parts of its physical or control structure ceases to work normally.

Therefore, this thesis proposes a multi-agent system implementation to control frequency channel allocation in cellular networks. The internal agent architecture is based on INTERRAP because of its model that accomplishes properties of reactive, deliberative and interacting agent architectures. The co-ordination level of the multi-agent system is based on Contract-Net Protocol using resource-oriented algorithms for the co-ordination of joint plans.

Chapter 4 Application of Intelligent Software Agents for Frequency Channel Assignment in Cellular Networks

4.1 Motivation

As the demand for mobile services grows, techniques for increasing the efficiency of channel usage in cellular networks become more important. Resource flexibility is needed to cope with the limited frequency spectrum available for network operators. The work described in Chapter 2 led to several algorithms being proposed to maximise the channel usage and minimise the call blocking probability. As discussed there, some of the channel assignment schemes presented in the literature have improved the performance of the basic fixed channel assignment strategy for different traffic densities (macro/micro/pico cellular networks) over different traffic load conditions. However, most of those solutions have an entirely reactive approach: the response to a series of events follows an algorithm that is prepared to react to specific situations. This entirely reactive approach limits the efficiency.

The aim of the work in this thesis is to propose a channel allocation scheme that improves the efficiency of the frequency channel assignment through a more flexible radio resource acquisition, but also has a sufficiently distributed nature to make its implementation feasible for real systems. The approach adopted is to use a multi-agent system to provide more autonomy and flexibility to the base stations and to increase the robustness of the whole mobile network by allowing base stations to negotiate resources. This work exploits the ability of intelligent agents to perform autonomous and intelligent negotiation in order to improve the acquisition of radio resources in congested areas.

The scenario here is assumed to use macro-cells where base stations are not able to share information by interference measurements, but only by explicit exchange of information; the resources are complete frequency carriers. However, the agent concept applied in this work is generic and can be extended to other types of mobile networks (Chapter 7).

The remainder of this chapter describes the specification of the new channel allocation scheme within the functional agent architecture.

4.2 Specification of the Multi-Agent Based Channel Allocation Scheme

The objective of any frequency channel allocation scheme is to achieve a high degree of channel usage and a low rate of call blocking. However, there will be a trade-off between the network performance and the cost of the network. The cost of the network is normally measured by the signalling load caused by the execution of the channel allocation scheme and by the response time under a defined quality of service. Signalling load and response time are important parameters within the design of a new channel allocation scheme.

In cellular systems, when explicit exchange of information is necessary, the network cost is directly related to the complexity of the scheme. Simple algorithms with the smallest possible number of database lookups are desirable. However, as described in Chapter 2, the schemes that present higher channel efficiency and lower call blocking rates are complex and generate a high signalling load, so increasing their response time. One approach to alleviate this problem is to divide the functionality of the scheme into different layers. The lowest layer allocates channels using a less complex algorithm and leads to a fast response time. The other layers with longer time scales try to maximise the channel usage and decrease the blocking rate, but controlling the signalling load. The interaction between layers can become very complex with the possibility of inconsistent results. The implementation of a layered scheme needs to rely on architecture developed to support layered control.

Channel allocation schemes are implemented to be autonomous and with consistent behaviour; this is done through carefully designed algorithms. However, most of the algorithms in the literature are reactive with a single strategy for adaptation to traffic changes. They lack flexibility because they are not able to monitor their own performance or to change strategies to continue to pursue their objectives efficiently, or to negotiate resources within a group of cells. Flexibility can be achieved through pro-active logical planning and through co-ordinated negotiation.

The desired features of complexity separation, fast response, autonomy, pro-activity and negotiation for a real-time application have been objectives of multi-agent systems research as described in Chapter 3.

Several multi-agent frameworks have been proposed for control management in telecommunication networks. For example, adaptive routing control applying ant-like mobile agents have been studied by [SHB97] [WP98] among others. Also, hierarchical layering of agents for distributed ATM network management was proposed by Fergal Somers in the HYBRID architecture [Som96] and by Hayzelden and Bigham in Tele-MACS [HB98]. The hierarchical arrangement allows levels of co-ordination. Each layer is defined to conduct the control of the network to a certain level of competence. In [Som96], the levels are classified as *local, county, regional* and *national* levels. In [HB98], the hierarchy is composed of *reactive, planning* and *strategic* layers. The adoption of a layered multi-agent control system seems to fit the requirements of distribution, intelligence, robustness and concurrency of broadband network control.

For the channel allocation problem, the main goal is to bring more distribution of control, autonomy and flexibility in resource management. As stated in [Som96] the decision whether to use a hierarchy of agents or to use a set of non-hierarchical agents depends on the control domain. If the interaction between agents to control the resource is limited to the domain of the agents, a non-hierarchical approach would be more straightforward; if the control demands co-ordination of other domains in the network, the hierarchical approach is more suitable. The channel allocation problem can be restricted to one domain, so that a single layer of agents can be placed in the base stations. However, layered control is seen as essential, so the internal architecture of the agent must supply this feature. Moreover, the agent architecture must support planning and negotiation mechanisms.

Internal agent architectures are classified by the degree of reasoning incorporated into the agent, from a completely logical model to a fully reactive model with no symbolic representation. Hybrid architectures combine features of logical and reactive models and are more suitable for real time applications. As described in Chapter 3, INTERRAP [Mül96] is a hybrid agent architecture that also incorporates mechanisms for co-ordination and co-operation among autonomous agents. Internally the agent consists of a set of hierarchical layers, a knowledge base that supports the representation of different abstraction levels of knowledge and a well-defined control architecture that ensures coherent interaction among layers. It was designed to react to unexpected events, to long-term actions based on goals and to cope with other agent interactions. INTERRAP was, therefore, chosen to be the architecture model for the agent implementation in this work. The architecture illustrated in Figure 4.1 was adapted from Figure 3.3 for application in mobile networks.

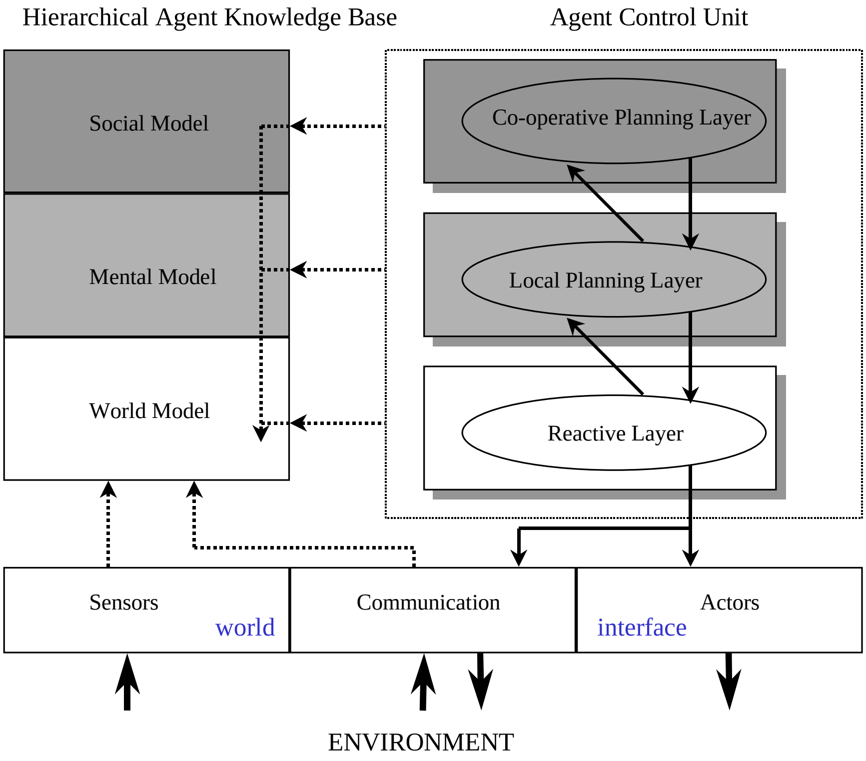


Figure 4.1: The architecture of the agent

In the cellular network scenario, each base station has one agent. The **world interface** presented to the agent includes the **sensor section** responsible for the perception of the environment, which would include requests for channel allocation from new calls, handoff requests, borrowing channel requests and orders for locking channels.

The **communication section** handles message exchanges for channel management and in the negotiation process. The **actor section** is responsible for all execution tasks that actually allocate, release, re-allocate, lock, lend channels, manage normal and supervised handoffs and terminate appropriately unsuccessful requests.

In the **knowledge base**, the **world model** contains the environment information and everything necessary for the operation of the Reactive Layer. The **mental model** contains the complete information about the agent, about the use of frequency channels and possibly history of traffic load in the cell. Finally the **social model** has relevant information about other agents' data. The control unit is structured to include a **Reactive Layer** that is responsible for fast accommodation of traffic demand, a **Local Planning Layer** using other strategies to optimise the local load distribution of channels and the **Co-operative Planning Layer**, responsible for load balancing across a larger area.

In the functional specification of the agent for the cellular scenario, the abstract architecture described in Chapter 3 is embedded inside the control and functional processes that compose the agent. The next section discusses the functionality of each agent layer, and how the agents accomplish the complete channel allocation scheme.

4.3 Functional Specification of the Agent

The channel allocation scheme is the main functionality of each agent. The execution of the complete channel assignment scheme is performed in two parts: one is the channel allocation inside the cell and the other is the load balancing of the network performed by a group of agents. The local channel allocation is mainly executed by the Reactive Layer of the agent, with optimisation of channel usage being done by the Local Planning Layer. The second part of the scheme is a joint plan between agents, executed when the local strategy becomes saturated and the performance of the system starts to degrade. The joint plan is triggered by the Local Planning Layer and co-ordinated by the Co-operative Planning Layer. The different parts of the whole scheme do not conflict, but even so the interaction between layers follows the INTERRAP control architecture, suppressing or changing tasks when channel allocation conflict arises. The Local Planning Layer is responsible for keeping consistency between joint plans and reactive tasks. The approach taken is this work is new in the field of mobile networks.

4.3.1 Reactive Layer

The Reactive Layer is basically composed of a FCA algorithm with channel borrowing and channel locking. The algorithm called *"Distributed Borrowing*

Algorithm" (D-BA) is based on the general behaviour of the D-LBSB scheme [DSJA97]. A description of the D-LBSB scheme is given in Chapter 2.

The decision to build into the Reactive Layer of the agent an algorithm similar to one of the distributed schemes in the literature was taken because it allows a close evaluation of the performance of such algorithms under simulated cellular network conditions and not only under a limited mathematical model. It also allows the benefits of the Local Planning and Co-operative Planning layers of the agent to be evaluated against the same reactive scheme.

The D-BA scheme is similar to the D-LBSB scheme, but not exactly the same. Some design and implementation parameters of D-BA are different as explained below.

4.3.1.1 Differences between D-LBSB and D-BA schemes

Although D-BA classifies mobile users in the same way as D-LBSB, there are important differences.

- In D-LBSB the region ***r*** is given by the equation 2.17 (section 2.4.1.1), relying on more assumptions than the more detailed cellular model built in this work. Therefore, in D-BA, the region ***r*** is fixed in all cells and it is determined based on the signal strength that will be received by the mobile users close to the borders. The decision to calculate the region ***r*** in this way was taken because the main purpose of this region is to indicate which mobile users are more likely to go out of the cell.
- Differently from D-LBSB, in D-BA the mobile user classification routine is independent of the borrowing algorithm. The approach taken is to have the mobile user classification continuously updated, allowing the agent to use this information outside the execution of the borrowing algorithm. It is also faster and more cost-effective to have the mobile user classification being continuously updated than to have to build all the information every time the borrowing algorithm is executed.
- Having the region ***r*** fixed and the classification of each mobile user, the **NumDepart** array is available at all times. This is another difference between the two schemes. D-BA has the advantage of having the **NumDepart** array updated

continuously and not only when the borrowing algorithm is initialised as in D-LBSB.

The cycle in D-BA is shorter and slightly different with regard to threshold decisions, because it considers only the neighbouring cells as sources from which to borrow channels and the threshold decisions are made in order to increase the chance of borrowing channels.

In most of the channel allocation schemes in the literature that use channel borrowing and channel locking, the borrowing algorithm is executed every time a call or handoff request is made and there are no more available channels to accommodate the request. This is not the case in D-LBSB and consequently in D-BA, where the borrowing algorithm is triggered *before* the exhaustion of the nominal channels, when the threshold ***h*** is reached. The borrowing algorithm does not get one channel, but a certain number of channels ***X*** that depends on the average traffic load of the entire network. Therefore, the outcome of each execution of the borrowing algorithm in D-LBSB or D-BA can be *successful* (when all ***X*** channels are borrowed*), partially successful* (when some channels are obtained but not the number ***X*** of channels expected) or *unsuccessful* (when no channels are obtained). However, there is no mention in the references for the D-LBSB scheme about these possible outcomes, and more importantly about the criteria for successive runs of the borrowing algorithm from the same cell. Therefore, in the D-BA scheme, special criteria were introduced for re-executions of the borrowing algorithm as explained later in section 4.3.1.2.

Other differences are found in implementation parameters such as the *degree of coldness* threshold to consider a cell *hot*, the *degree of coldness* threshold for a lender cell to stop lending channels, the *degree of coldness* threshold to release a borrowed channel; these values are not fully described in the references.

4.3.1.2 Description of the D-BA

Mobile users are classified in three categories: *new*, *departing* and *others*. *New* users are those that have successfully received a channel in the cell (new calls and handoff requests) between the time of the channel allocation until time "***t***". After time ***t***, they are classified as *others* or *departing* according to their position inside the cell. Mobile users are classified as *departing* if they are inside a pre-determined region ***r*** close to the borders of the cell. Users are classified as *others* if they are inside the main region of the cell. There is a special routine that re-checks the position of the mobile user periodically in order to keep its classification updated.

When the channel availability in the cell decreases to a certain threshold, it becomes ***"hot"***; cells above the threshold are called ***"cold"*** cells. The classification of a cell as *hot* or *cold* depends on a threshold (**h**) of its *degree of coldness* (**dc**), i.e., the ratio between the number of available channels and the total number of assigned channels of the cell (**C**).

Hot cells are not allowed to lend channels and *cold* cells are not allowed to borrow channels.

Each cell site keeps as local parameters its **NCC** (set of non-co-channel cells, considering its compact pattern), **CC** (set of co-channel cells), **dc**, $\boldsymbol{H_{NCC}}$ (set of *hot* non-co-channel cells), $\boldsymbol{H_{CC}}$ (set of *hot* co-channel cells). The cell knows the first two parameters from the configuration of the network; **dc** is updated internally according to the current availability of channels. $\boldsymbol{H_{CC}}$ is updated independently of the algorithm as a cell informs all its co-channel cells when it changes its state (*hot* or *cold*). $\boldsymbol{H_{NCC}}$ is computed at the beginning of the algorithm execution.

When a cell becomes *hot*, it triggers the execution of the following algorithm:

1. The *hot* cell broadcasts a message requesting the **dc** of all cells.
2. Receiving the responses, the *hot* cell calculates the *average degree of coldness* ($\mathbf{dc_{avg}}$) of the entire network, and $\mathbf{H_{NCC}}$. With this data, it calculates how many channels ($\boldsymbol{X}$) it needs to borrow to reach $\mathbf{dc_{avg}}$, solving the Equation 2.16 from Chapter 2 section 2.4.1.1 (rewritten below):

$$\boldsymbol{X} = \boldsymbol{C} \cdot \left(\boldsymbol{dc_{avg}} - \boldsymbol{h}\right) \qquad (4.1)$$

3. The algorithm starts to execute the cycle below:
    a) The *hot* cell **B** sends messages to the *cold* neighbouring cells **L**. The content of the message allows the **L** cells to compute the utility function $\boldsymbol{F_{BL}}$ (Equation 2.17) for a 7-cell compact pattern as described in section 2.4.1.1 of Chapter 2 (as Equation 4.2). The computations of each **L** cell are sent back to **B**.

$$F_{BL} = \frac{dc_L}{\left[\frac{D_{BL}}{R_{cp}} \cdot \frac{(1 + H_{BL})}{7} \right]} \qquad (4.2)$$

D_{BL} or **nearness** is the measure in terms of cell distance between **L** and **B**, for example, for the neighbouring cells of **B**, D_{BL} is 1.

H_{BL} or **hot cell channel blockade** is the number of *hot* co-channel cells of **L** that are non-co-channel cells of **B**.

R_{cp} is the radius of the compact pattern, also measured in terms of cell distance. R_{cp} and 7 are used for normalisation.

b) **B** orders the list of **L** cells by decreasing values of F_{BL}s and informs the cell with the higher value that it is the current lender.

c) **B** borrows channels from the lender **L** respecting the safety threshold of **L**'s channel availability. Every channel borrowed is locked in **L** and in the co-channel cells of **L** that are non co-channel cells of **B** inside the reuse distance. If the number of channels needed is not reached, the procedure is repeated for the other **L** cells in the list. The algorithm terminates when the number of channels X is reached or when the search in the **L** cells is exhausted.

The possible outcomes of each execution of the borrowing algorithm can be:

1. *Successful*: all X channels are borrowed.
2. *Partially successful*: some channels are obtained but not the number X of channels expected (in the implementation, a partial success is seen as a *failure of type 1*).
3. *Unsuccessful:* no channels are obtained, for one of the following reasons:
    a) *Failure type 2*: the lender cells did have channels to lend, but the channel locking was not possible.
    b) *Failure type 3*: All possible lender cells are *hot*.
    c) *Failure type 4*: the network is so heavily loaded that X is equal to zero.

Borrowing Algorithm Re-execution Criteria

In D-BA, the borrowing algorithm is re-executed if the cell reaches the threshold again, even if the cell has not given back all borrowed channels. However, if the outcome of the algorithm execution is partially successful or unsuccessful, the cell will not be allowed to re-execute the algorithm immediately, but it will be forced to wait for a certain amount of time depending on the cause of the failure of the previous outcome. The introduction of these delays is to avoid an uninterrupted run of the borrowing algorithm, which would overload the data processing in the base station and increase substantially redundant signalling in the network.

Channel Assignment

The D-BA scheme is also responsible for the channel assignment. If the incoming request is a handoff request it will first try to assign the request to a nominal channel; if that is not possible, it will look for a borrowed channel (if any exists in the cell). If both attempts fail, the request is blocked. If the request is an incoming call, it will look firstly for an available borrowed channel and then for an available nominal channel. Failure in both attempts leads to the request being blocked. The decision to use a different order of channel assignment for the two types of channel request is because new users coming to the cell through handoff requests are more likely to stay in the cell. New calls can be generated at any place in the cell and if there are borrowed channels it is better to make use of them in order to keep them for shorter periods, improving the channel usage.

4.3.2 Local Planning Layer

The Local Planning Layer determines the *departing region* ***r*** based on the signal/noise ratio. This layer is responsible for the channel re-assignment scheme. Every time a channel is released, the Reactive Layer requests a re-assignment decision:

- If the channel is a borrowed channel, it will be re-allocated to an appropriately departing user (using a nominal channel) or it will be given back to the owner cell, depending on the channel availability of the current cell.
- If it is a nominal channel, it will be re-allocated to non-departing user that is using a borrowed channel.

The threshold for keeping or releasing a borrowed channel may vary with the traffic load.

This layer monitors the efficiency of the algorithm in the Reactive Layer, and it is responsible for the decision of triggering the Co-operative Planning Layer.

The Local Planning Layer could be much more elaborate than the current version in this work: more reasoning and learning could be added to improve local performance. For example the determination of ***r*** could take in account traffic history conditions. Implications of this are discussed in Chapter 6.

4.3.3 Co-operative Planning Layer

The Co-operative Planning Layer is responsible for the negotiation of frequency resources. When the use of the channel allocation algorithm in the Reactive Layer is not sufficient to keep low rates of call blocking, one way to alleviate the load of a *hot spot* is to move calls to less loaded regions. This can be done through *management handoffs* (or *traffic handoffs* as known in the standards), but this is not an easy task. Only mobile users close to certain borders of the cells can attempt the management handoffs. The handoff attempts must be in a co-ordinated manner to avoid a mobile user being shifted back and forward between two cells. The handoff of users to different cells is a load-balancing problem, so that the co-ordinated control needed to solve this problem depends on the collaboration of a group of cells. The co-ordination of a joint plan using management handoffs is the responsibility of the Co-operative Planning Layer of the agents. The frequency resource negotiation performed by the agents has two phases:

- to find the best region to attempt the movement of calls and to engage in a joint plan.
- the actual execution of the management handoff requests in a co-ordinated manner, i. e. the execution of the joint plan.

In order to engage and collaborate in a joint plan, the agents communicate through a negotiation protocol called Contract-Net Protocol (CNP) [Smith88]. A protocol is a common pattern of conversations used to perform some general task among agents.

In order to perform the Contract-Net Protocol, the agents communicate through communicative acts. Communicative acts correspond to building blocks of dialogue between agents and are based on speech act theory [Aus62] [Sea69]. Each message exchanged between agents is a communicative act. In this work, the agent implementation uses a sub-set of the communicative acts specified in the FIPA Agent Communication Language (ACL) [FIPA97]. Table 4.1 shows the set of ACL communicative acts being used in this work.

Table 4.1 — Categories of communicative acts

Communicative act	Information passing	Requesting information	Negotiation	Action performing	Error handling
accept-proposal			✓		
cancel				✓	
cfp			✓		
failure					✓
inform	✓				
inform-ref (macro act)	✓				
not-understood					✓
propose			✓		
query-ref		✓			
refuse				✓	
reject-proposal			✓		
request				✓	

Figure 4.2 illustrates the CNP as described in the Foundation for Intelligent Physical Agents (FIPA) [FIPA97] and applied in this work.

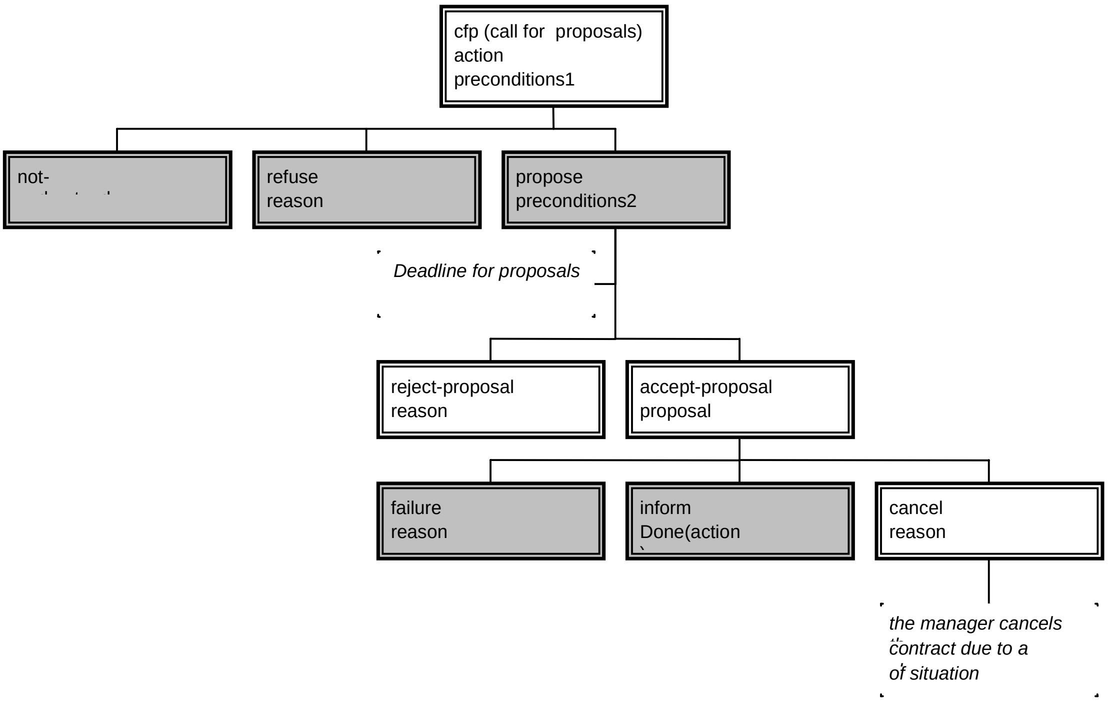


Figure 4.2: FIPA Contract-Net Protocol

In Figure 4.2, boxes with double edges represent communicative actions, white boxes represent actions performed by the initiator, shaded boxes are performed by the other participant(s) in the protocol and italicised text with no box represents a comment [FIPA97].

The execution of the protocol during the two-phase agent negotiation is best explained through an example.

4.3.3.1. First Phase of Negotiation

In the first phase of negotiation, a utility function is used to find the best region to move calls. The best region is the one that maximises such utility function, which takes into account the availability of channels and the proportion of users in the areas where management handoffs may be performed.

Suppose that in the cellular network in Figure 4.3, cell **A** is *hot* and the local channel allocation algorithm is not responding efficiently anymore because the neighbouring cells are also getting *hot*. The Local Planning Layer triggers the Co-operative Planning Layer to start the negotiation with agents in other cells.

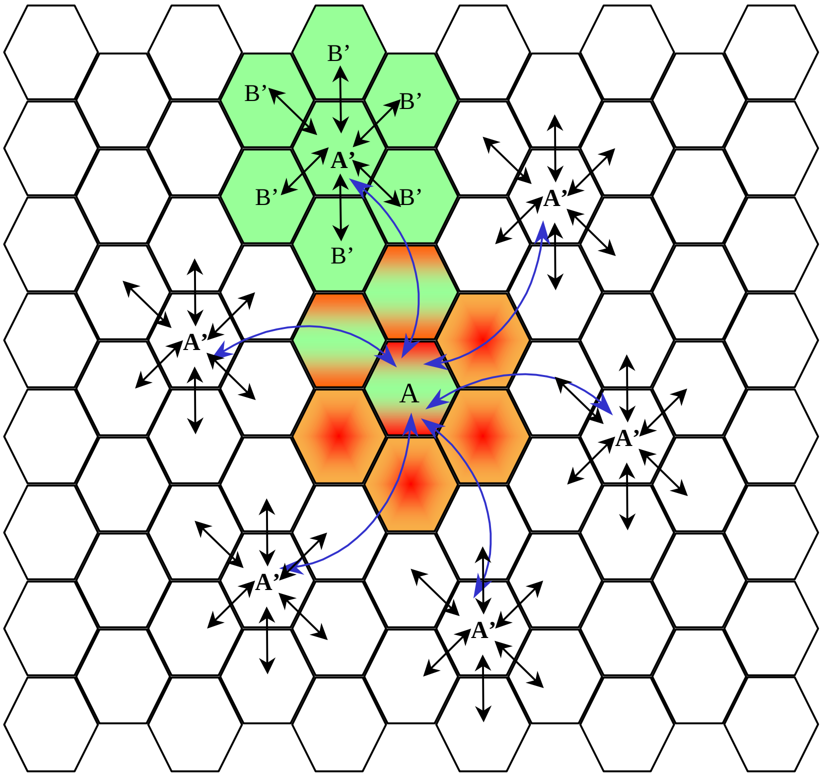


Figure 4.3: Negotiation Strategy

The following actions are taken in the first phase of negotiation:

1. In the *hot* cell, agent ***A*** is a *manager candidate agent* and it sends *call for proposals* (***cfp(0)***) to all its *co-channel cell agents* $\boldsymbol{A'_i}$, where ***i*** can vary from 1 to the current number of co-channel cells of ***A*** (maximum 6).
2. The $\boldsymbol{A'_i}$ agents advertise the need for channel availability assessment to their *neighbouring cell agents* $\boldsymbol{B'_{ij}}$, $1 \leq \boldsymbol{j} \leq 6$ sending ***cfp(1)*** (cell load assessment and plan engagement availability).
3. Each $\boldsymbol{B'_{ij}}$ agent sends to the respective $\boldsymbol{A'_i}$ agent a ***propose(1)*** act if the cell is able to engage in a joint plan in the near future or a ***refuse(1)*** act otherwise. Inside a ***propose(1)*** act the $\boldsymbol{B'_{ij}}$ agent sends its degree of coldness (**dc**) and the percentage of the mobile users in the cell (***Depart***) which are inside the departing region **r** of selected cell borders, called *regions of movement.* The regions of movement in each cell are pre-defined according to the position of the cell of agent ***A***.
4. Each $\boldsymbol{A'_i}$ agent receives the answers of its neighbouring $\boldsymbol{B'_{ij}}$ agents and it computes the value of the utility function given by Equation 4.3. Where ***Reg*** is the minimum set containing $\boldsymbol{A'_i}$ and its neighbours $\boldsymbol{B'_{ij}}$ that have sent ***propose(1)*** acts. In the calculation of ***F'***, ***Reg*** can be less than 7 (even if all $\boldsymbol{B'_{ij}}$ agents have sent ***propose(1)*** acts) because the cellular network is not wrapped, the boundary cells of the network do not have a complete set of neighbouring cells, therefore $1 \leq \left|\mathrm{Re}\,\boldsymbol{g}\right| \leq 7$. The factor α is introduced in order to decrease the influence of ***Depart*** over ***F'*** ($0 < \alpha < 0.25$).

$$\boldsymbol{F}' = \frac{1}{\left|\mathrm{Re}\,\boldsymbol{g}\right|} * \sum_{\boldsymbol{k} \in \mathrm{Re}\,\boldsymbol{g}} \left(\boldsymbol{dc_k} + \boldsymbol{\alpha} \cdot \boldsymbol{Depart_k} \right) \qquad (4.3)$$

5. The $\boldsymbol{A'_i}$ agents that were able to perform the calculation of ***F'***, send the result of ***F'*** to agent ***A*** in ***propose(0)*** acts. The $\boldsymbol{A'_i}$ agents that did not have enough ***propose(1)*** acts from their neighbours in order to calculate ***F'*** send ***refuse(0)*** acts.
6. The received ***propose(0)*** act with biggest ***F'*** value is chosen to be the region for moving the calls (if ***F'*** is greater than a minimum value). Agent ***A*** advertises the result of the auction to the winning *co-channel cell agent* with an ***accept-proposal(0)*** act. If there is no winning region, then agent **A** sends ***reject-proposal(0)*** to all $\boldsymbol{A'_i}$ agents that have sent ***propose(0)*** acts and aborts the joint plan attempt for a specific duration of time.

7. If there is a winning region, then the *co-channel cell agent* of this region sends ***cfp(2)*** (engage joint plan) to its neighbouring **B'_{ij}** agents.
8. Each **B'_{ij}** agent, receiving the ***cfp(2)***, assess its availability to engage the joint plan, considering the number of plans it is already participating in and the regions of movement being already considered in such plans. It sends a ***propose(2)*** act if: the number of current engaged plans is less than two and the regions of movement (if engaged in another plan) match the requesting one. Otherwise, it sends a ***refuse(2)*** act.
9. If the wining *co-channel cell agent* receives back a minimum number of ***propose(2)*** acts from its neighbouring **B'_{ij}** agents, it sends back an ***inform(jp)*** (inform joint plan) act to agent **A** and sends ***accept-proposal(2)*** acts to all of its **B'_{ij}** agents that have sent ***propose(2)*** acts. Otherwise it sends a ***failure(jp)*** (joint plan failure) act to agent **A** and ***reject-proposal(2)*** acts to its **B'_{ij}** agents that have sent ***propose(2)*** acts.
10. The winning *co-channel cell agent* that has just sent an ***inform(jp)*** and its **B'_{ij}** agents will perform all preparatory tasks to engage the joint plan and they will wait for an ***inform(activejp)*** (inform joint plan activation) from agent **A**.
11. If agent **A** receives an ***inform(jp)*** act, it sends a ***reject-proposal(0)*** to all other *co-channel cell agents* that have sent ***propose(0)*** acts before, and a ***request(jp)*** (request joint plan engagement) act to its two *neighbouring cell agents* in connection with the winning region. This request is mandatory. Finally, agent ***A*** will send an ***inform(activejp)*** act to all agents engaged in the joint plan (first joint plan execution act). If agent **A** receives a ***failure(jp)*** act, it selects the next best ***F’*** (if exists) and the actions from 6 to 11 are repeated.
12. An agent receiving a ***request(jp)*** act will perform all preparatory tasks to engage the joint plan and wait for an ***inform(activejp)*** act from agent **A**.
13. End of the first phase of negotiation.

Figure 4.4 shows the flow diagram of a successful first phase negotiation. For clarity, the diagram shows only the manager agent candidate, one of its two neighbouring cell agents, the wining co-channel cell agent and one of its neighbouring cell agents.

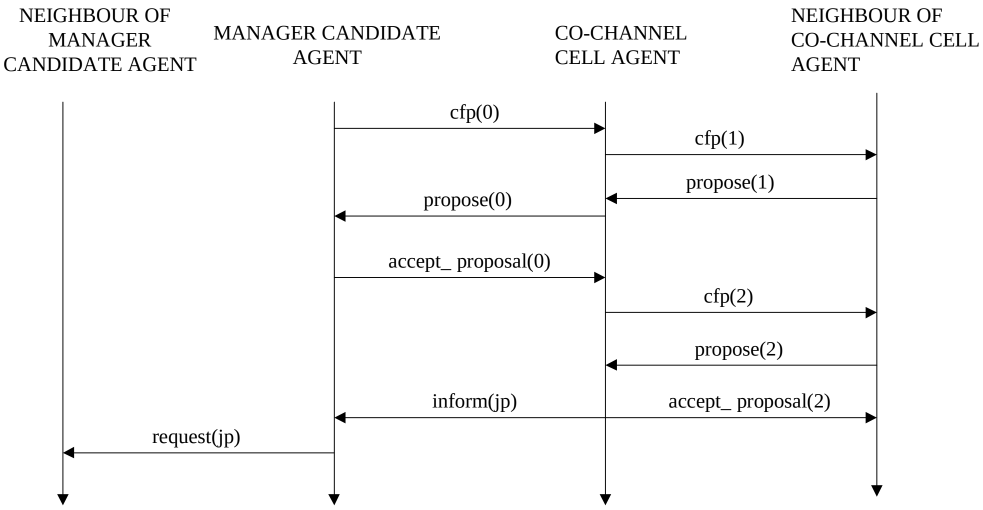


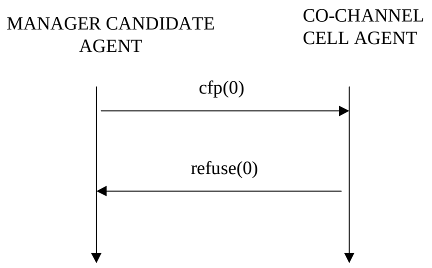
Figure 4.4: Successful first phase agent negotiation

Figure 4.5 shows possible examples of unsuccessful selection of regions for moving calls. The diagrams show the interactions between a manager agent candidate and simplified cell regions (only one of the $\boldsymbol{B'_{ij}}$ agents is shown).

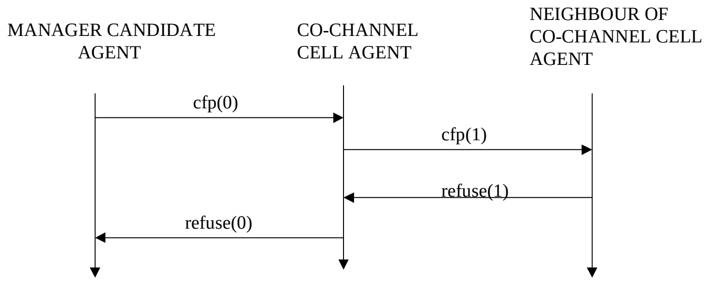
4.3.3.2 Second Phase of Negotiation: the Joint Plan Execution

The second phase of the negotiation starts with the engagement of all agents belonging to the winning region, the manager agent ***A*** and its two neighbouring cell agents (which border the winning region) into the joint plan (as illustrated by the green shaded region in Figure 4.3). Agent ***A*** is the *manager* of the joint plan and the other partner agents are the *contractors* of the plan [Smith88]. The manager has the responsibility to monitor the actions of the contractors and for the termination of the joint plan.

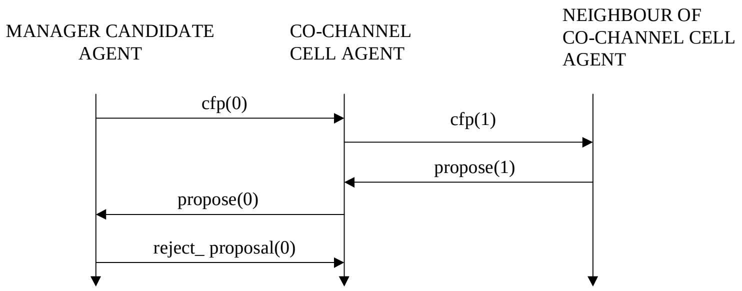
Each iteration of the joint plan needs to be feasible. Therefore, the heuristic developed by the author follows a resource-oriented approach of market-based control [Clearw96]. The aim is to load-balance the whole region so that the difference in degree of coldness of partner cells should be smaller then certain threshold. The heuristic tries to balance the region by distributing users among cells (when possible).



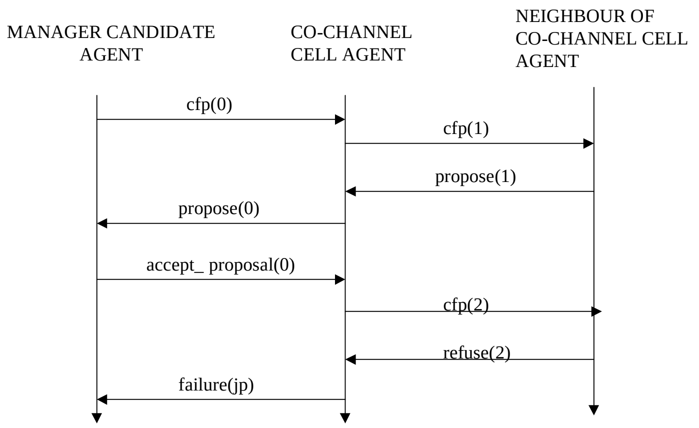
(a) unsuccessful negotiation case 1



(b) unsuccessful negotiation case 2



(c) unsuccessful negotiation case 3



(d) unsuccessful negotiation case 4

Figure 4.5: Examples of unsuccessful region selection

Plan Initialisation

1. The manager agent **A** sends its first act to all partner agents to inform them that the joint plan is in operation (***inform(activejp)*** act).
2. All partner agents receiving the ***inform(activejp)*** act will send an ***inform(ptrnjp)*** (partner cell in the joint plan) act to their manager agent identifying themselves and their neighbouring cells in the regions of movement.
3. Iteration:
    a) The manager agent sends a ***query-ref(0)*** act to all partner agents.
    b) Each partner agent sends its total number of channels and the number of channels in use to the manager agent through an ***inform-ref(0)*** act.
    c) The manager agent computes the rate of change (Δc_i) for each partner agent and itself by calculating the difference between the channel occupancy of the cell (c_i/C_i) and the average channel occupancy of all members (N) of the joint plan (L_{avg}):

$$\Delta c_i = \frac{c_i}{C_i} - L_{avg} \qquad (4.4)$$

$$L_{avg} = \frac{1}{N}\sum_{i=1}^{N}\frac{c_i}{C_i} \qquad (4.5)$$

where $\forall i \in N$

c_i is the total number of channels in use in the cell of agent i.

C_i is the total number of channels (nominal + borrowed) in the cell of agent i.

L_{avg} is the average channel occupancy of all cells of the joint plan.

Δc_i is the rate of change in channel occupancy of the cell of agent i inside the joint plan.

    d) If the cell of agent i has $\Delta c_i > 0$, the manager agent sends to agent i: Δc_i, the Δc of the neighbouring cells having borders with the regions of movement of the cell of agent i and the total number of channels of these cells (C). It also sends L_{avg}. This information is sent through a ***request(jpaction)*** (joint plan action) act.

e) Each agent ***i*** that receives the ***request(jpaction)*** act from the manager agent will try to transfer users in the regions of movement (departing areas) following the algorithm:

   i) Sort the received $\Delta \boldsymbol{c}$ of the neighbouring cells.

   ii) If $\Delta \boldsymbol{c}_i$ is smaller than min $\Delta \boldsymbol{c}$, then no transfers can be made; go to step f). Otherwise, go to step iii).

   iii) Calculate how many users need to be transferred: $\boldsymbol{users} = \Delta \boldsymbol{ci} * \boldsymbol{Ci}$.

   iv) If min $\Delta \boldsymbol{c}$ is greater than L_{avg}, then transfer one user to the neighbouring cell with min $\Delta \boldsymbol{c}$; Go to step viii). Otherwise, go to step v).

   v) Sort only $\Delta \boldsymbol{c}$ that is smaller or equal to L_{avg}. The aim is to transfer mobile users proportionally to the number of channels available in each target neighbouring cell with $\Delta \boldsymbol{c}$ smaller or equal to L_{avg}.

   vi) For all sorted $\Delta \boldsymbol{c}$ find the number of users that the cell can receive. For $\Delta \boldsymbol{c}$ of cell j : $\boldsymbol{us}_j = -\Delta \boldsymbol{c}_j * \boldsymbol{C}_j$

   vii) To find the proportion of users that will be attempted to transfer to each cell, sum all us$_j$: $\boldsymbol{US} = \sum_{j=1}^{m} \boldsymbol{us}_j$. The proportion of users for each cell is: $\min\left(\frac{\boldsymbol{us}_j}{\boldsymbol{US}} * \boldsymbol{users} \right)$.

   viii) Do the handoff attempts.

f) End of the iteration.

4. Repeat this iteration at intervals of ***s*** seconds until the manager decides to terminate the joint plan. When the plan is to be terminated, the manager agent sends a ***cancel(jp)*** (cancel joint plan) act to inform the termination of the joint plan.

The termination of the joint plan can be determined by the completion of certain number of iterations, or when the difference in traffic load between the member cells is smaller than certain value, or by an exception.

Figure 4.6 shows a flow diagram of the execution of a joint plan. Again for clarity, only the manager agent and one of the partner agents are shown. The dashed arrow means that the issue of the act may or may not happen, depending on the outcome of the manager’s heuristic.

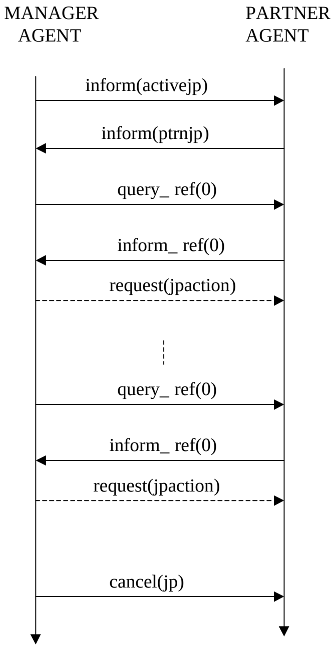


Figure 4.6: Execution of a joint plan

A section of Chapter 5 describes the implementation of the negotiation acts in signalling messages; it also shows the flow diagrams of the signalling message exchanges between agents inside the cellular network model.

4.4 Summary

This chapter described the specification of the proposed multi-agent system applied to the scenario of mobile networks. The abstract agent architecture described in Chapter 3 was adapted to control the frequency channel allocation inside the cellular network environment. The main functionality of each layer of the agent was fully specified.

Chapter 5 Simulation Modelling

5.1 Introduction

In the previous chapter the agent architecture was presented and the specification of the multi-agent system described. The reminder of this thesis is concerned with investigating the benefits of the multi-agent system on channel allocation.

There are two ways of measuring the performance characteristics of a channel allocation scheme: through mathematical analysis or through a simulation model. Most of the performance measurements of channel allocation schemes in the literature were made through mathematical analysis, using Markov chains, the traffic pattern in the cells being modelled using Poisson distributions. Some schemes did not distinguish between incoming calls and handoff requests, as for example in [ZY89][ESG82]. Others represent handoff requests as an additional Poisson traffic load in the cell as in [DSJ97] [DSJA97]. Therefore, they do not model mathematically the possible trajectories of the mobile. The reason for that is the estimation of mobility in communication networks can become too complex for mathematical analysis [Siva98]. Moreover, the variety of signalling messages in real cellular systems is difficult to model mathematically. To be able to make measurements with different user mobility scenarios and greater range of signalling statistics, simulation models are used.

Cellular network systems can be modelled through an event driven simulator. The cellular model can be built using a simulator developed specially for this purpose or using a commercial simulator. Purpose-built simulators are faster in terms of execution time, but the development time can be prohibitive. Commercial simulators are not as fast as purpose-built simulators, because their code is much heavier than a single purpose simulator, but on the other hand, the existence of library models allows the faster construction of more detailed models. In addition, the benefits of documentation and debug facilities improve the quality of the resulting model. Finally the ability to plug in existing models can be beneficial in the development phase, for comparisons or for further work.

OPNET™ is a communication network commercial simulator that can be defined as a general-purpose event driven simulator. A simulation model is defined in four levels (until version 3.1) or five levels. Each level has a different graphical interface. The top level in the later versions is the project level, followed by the network, node, process and code levels. The topology of the network to be simulated is built in the network level. The node level consists of the constituent modules of each network node and their interconnections. The elements are queues, processes, sources, receivers and transmitters. The functionality of each module is built in the process level; a module being defined as a process or a set of processes. Each process is defined as a finite state diagram and the transitions to each state. The code level is the contents of each state, separated by the code to be executed before an event and the code to be executed after an event, before changing state. OPNET is built in the C language, but it has a large number of predefined functions, several of them replacing normal predefined C functions in order to give the expected result inside the simulator.

In this research, the results of the multi-agent system are compared against the conventional mobile network using the FCA and a mobile network using the D-BA scheme under common traffic load scenarios. The network performance measurements used for the comparison between channel allocation schemes are the traffic blocking rate, handoff rejection rate and call dropping rate. To be able to perform the comparisons, a common cellular network needs to be modelled in OPNET and validated against a mathematical model.

This chapter describes the models developed in OPNET to build the cellular network and the different channel allocation schemes, the assumptions made on these models and the implementation parameters used.

5.2 Basic Cellular Model Description

The cellular network was modelled to present much of the functionality of a mobile phone system as specified by AMPS [Mac79]. As explained earlier, AMPS is used because there is more literature that can be used for comparison. The main features implemented in the model are listed below.

- Call set-up/teardown is modelled explicitly.

- Handoffs are performed automatically based on received signal strength and parameterised handoff threshold.
- Emulation of forced call termination on insufficient signal is implemented.
- Mobile stations self locate on cells based on received signal strength from nearest base stations.
- All mobile stations have arbitrary trajectories.
- A cartographic background maps the cell boundaries.
- There can be any number of mobile stations per cell.
- The simulation includes a parameterised call generation distribution and average inter-arrival time.
- A parameterised call length distribution and average call length is implemented.

Modelling assumptions have been made in order to minimise simulation runtime and keep the model maintainable without affecting the investigation of the channel allocation problem. The cellular model makes use of two transmissions mechanisms. The radio interface between the mobile terminal and base stations, and the full-duplex T-carriers between base stations and the MTSO (it is possible to have high speed microwave radio links between base stations and the MTSO, but this transmission mechanism is not considered in the cellular model here). Both mechanisms are divided between voice and data links. In this work, data links are used for signalling purposes only. Propagation delays in transmission links are not considered. The signalling channels were simplified. Mobile stations do not receive calls only generate them, so that paging channels are not necessary. In AMPS, each cell has three forward control channels (FOCC) for a 7-cell cluster, but here, each cell has only one FOCC, represented by a radio data link. Therefore, there is one reverse control channel (RECC) per cell, which is represented by a radio data link and a static variable that emulates the Busy/Idle State (BIS). Control channel bit streams are not modelled explicitly; packets are used instead with relevant signalling information. Voice traffic is also emulated through packets sent periodically by a mobile station during the period that its call is in progress. The base stations and MTSO use the voice packets for signal strength monitoring leading to handoff procedures or forced call terminations if necessary.

The cellular network implemented has forty-nine cells, ten nominal frequency channels per cell in a 7-cell cluster, as shown in Figure 5.1 where frequency channel sets are represent by different colours. The number is used to identify a particular cell.

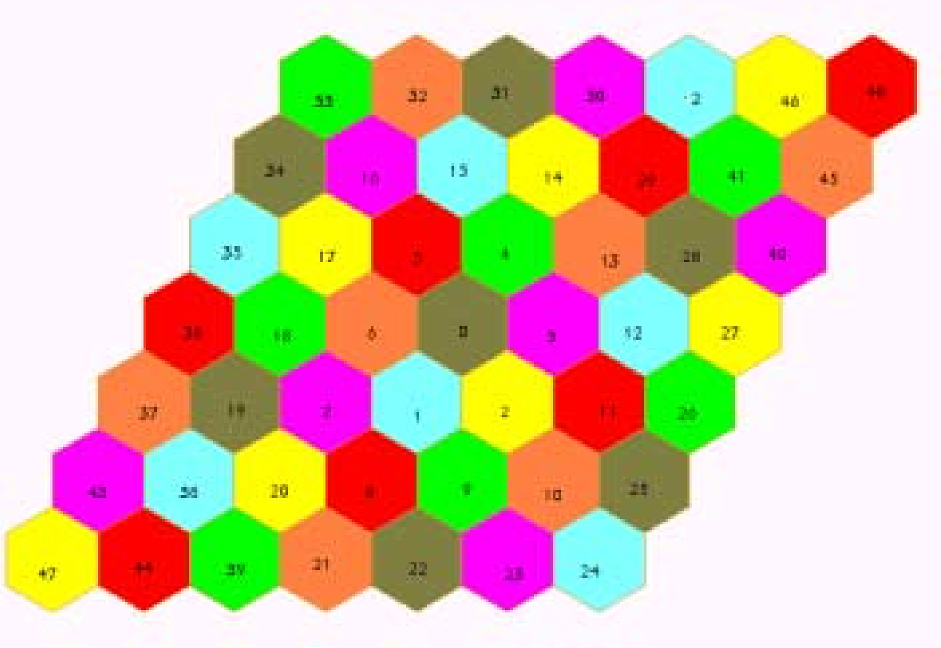


Figure 5.1: Cellular network frequency planning

The cellular network is not wrapped and all configuration settings concerning neighbouring cells and co-channel cells take this fact in account. Mobile users' trajectories also take into consideration the cartographic boundaries of the cellular network.

5.2.1 OPNET Network Model

The network is a macrocellular structure with a cell radius of 3.33207-km (corresponding to 0.03 degrees on the cartographic map at latitude 38^{o} [Rai48]).

In the scenarios used for the simulations, the network model implemented in OPNET consists of the following nodes: one MTSO, 49 base stations and 980 mobile stations. Figure 5.2 depicts part of the OPNET network model where base stations and their full duplex links with MTSO are shown, and mobile stations are distributed inside the cells.

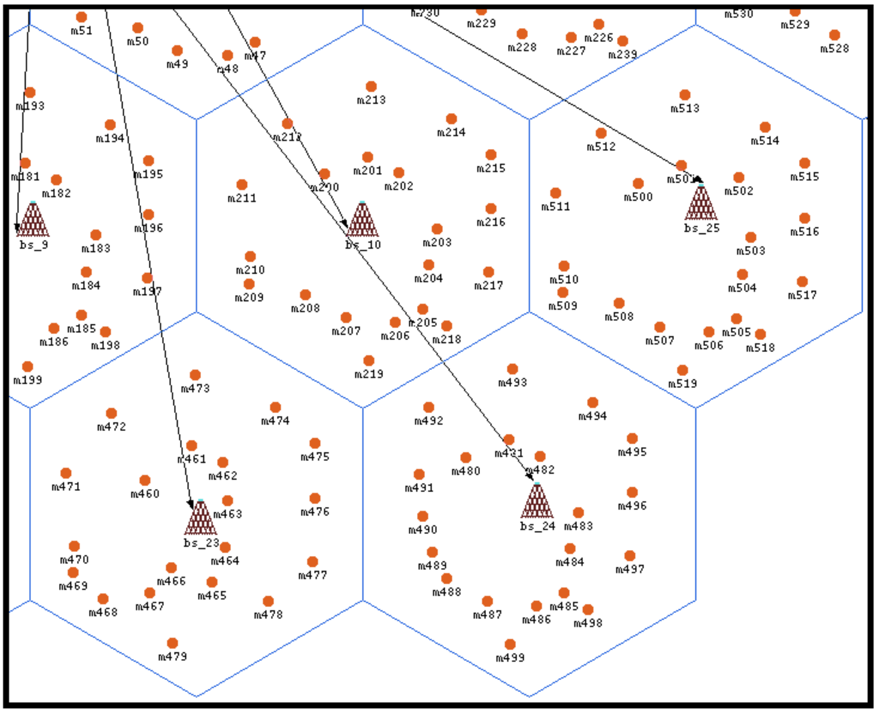


Figure 5.2: OPNET network model for the cellular model

The following sections describe the physical and functional properties of the node and process models of the cellular model. Some of the process models of the cellular model have similar functionalities to the ones presented in the *OPNET Cellular System Example Models* [OPNETb94]. The built-in OPNET cellular system model is a good example, but the code is not error-free. Also, it is missing important features like call dropping and other features like the call generation inside the mobile node, that make it unusable for this work.

5.2.2 Mobile Station

The mobile station has the capability to initiate, maintain and terminate calls, to self-locate in the cell with highest received signal and to move inside the cellular network. Different modules inside the mobile station node implement each one of these functionalities. Figure 5.3 depicts the modules of the mobile station node and their interrelations.

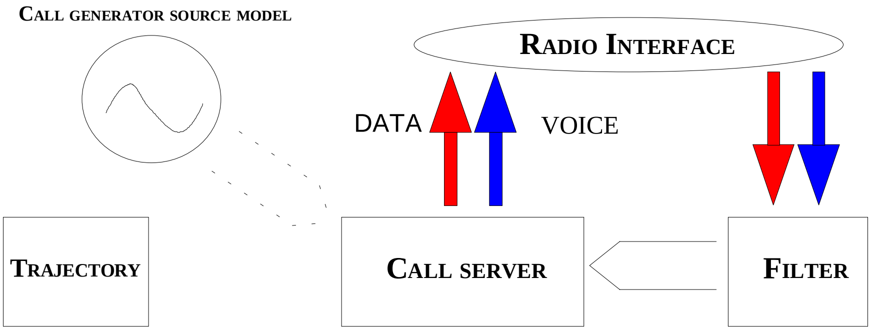


Figure 5.3: Mobile station node model

5.2.2.1 Radio Interface

A simplified model has been built. The three nearest base stations receive the voice packets sent by the mobile station and the signal strength is calculated based on the distance between the mobile station and the receiver base station.

5.2.2.2 Call Generator Source Model

In each cell a source call generator process generates the call requests and randomly chooses an idle mobile station located inside the cell to place the call. The *pdf* for the call inter-arrival time and its mean value are changeable simulation parameters. In the scenarios simulated in this work, a negative exponential distribution was selected for call inter-arrival time with different mean values depending on the desired traffic load in the cell. The call length distribution is also a simulation parameter. Constant and exponential call length distributions were used in the simulations, both with a mean of 180 seconds.

5.2.2.3 Call Server Module

Each mobile station has a unique identification in the system: the mobile identification number (MIN) that is defined by the filter module in the beginning of the simulation and used by the call server module to identify all voice and data packets sent by the mobile station. When the mobile station is idle the call server waits for the arrival of a call request from the call generator; when it gets one, the call set up is initiated. First, the BIS is checked in the RECC of the cell:

- If the BIS is equal to 0: the RECC is busy, the mobile station waits a random time interval between 0 and 200ms to recheck the BIS. The number of times the mobile rechecks the BIS is limited to 10 attempts through a process parameter in order to emulate the *maximum number of busy occurrences* parameter given by the *overhead information* in AMPS systems. When it reaches the limit of recheck BIS attempts the mobile returns to its idle state.

- If the BIS is equal to 1: the RECC is idle and the mobile station waits a random time between 0 and 92ms to send the call request, exactly as in the AMPS system when the *overhead message bit* is equal to zero. After the waiting time, the call request is sent and an *access timer* is set up in order to retransmit the call request if no response is received. The maximum number of call request attempts in the access channel is a parameter (a maximum of 4 attempts was used in the simulations), if the number of attempts reaches the maximum, the mobile returns to idle.

The response of a call request can be unsuccessful forcing the call to be terminated and the mobile to return to its idle state. However, if it is successful, the model deals with the tasks of a call in progress:

- Call update: During the length of the call, voice packets are sent to the voice radio link. The voice packet is used by the base stations and MTSO to perform *signal-to-noise ratio* (SNR) measurements. The periodicity of the voice packets is determined by a simulation parameter.
- Handoff: The mobile station performs a handoff order.
- Call dropping: The mobile station forcibly terminates the call if the signal is too weak. The mobile returns to idle.
- Call termination: The length of the call is reached and the call is normally terminated. The mobile returns to idle after sending a call termination packet in the voice channel.

5.2.2.4 Filter Module

This module is responsible for the self-location of the mobile station in the cell in which the mobile receives the strongest signal. When the mobile station is idle, it performs the self-location procedure periodically each 1 min (this value is also a changeable simulation parameter). During the period a call is in progress, the filter module sends to the call server the packets received from the radio link which are designated for its mobile station and destroys all other received packets.

5.2.2.5 Trajectory Module

This is an independent module inside the mobile station node responsible for the movement of the mobile inside the cellular network. The movement of the mobile station is independent of the status of the mobile (idle or call in progress). During the movement the position of the mobile is updated periodically, in accordance with a position update parameter inside the trajectory process. This parameter was chosen to be smaller than the voice update parameter in order to have an updated position of the mobile station every time a voice packet is sent. The time interval between position updates and the distance between two consecutive positions determine the speed of the movement and is called a *step*.

In this work, three different types of movements were implemented: random, driving on a highway and walking along a street. However any kind of trajectory can be implemented inside the model. The choice of these trajectories was an attempt to create diverse user mobility inside the cellular network.

The mobile station with random movement performs cycles of partially random movement. The start position is always the same (the position determined in the network configuration). The movement consists of vertical and horizontal trajectories (inside the network cartographic grid) in such a way that the mobile returns to its original position in the end of the movement cycle. The first trajectory is always vertical, the decision to go up or down in the grid is determined by a random variable with a uniform distribution between 0 (up) and 1 (down). Secondly, a random variable with a uniform distribution between 0 and a settable maximum number determines the number of steps of this first trajectory. When the mobile finishes its vertical trajectory, it starts the horizontal trajectory. Again a random variable with a uniform distribution determines whether the mobile goes right or left in the grid and another random variable with a uniform distribution determines the number of steps of this horizontal trajectory. The third trajectory is in the opposite vertical direction and with same number of vertical steps. The last trajectory is then in the opposite horizontal direction with same number of horizontal steps. This concludes a cycle of movement; the mobile goes to a stand still state during a random period of time (uniform distribution) before starting a new cycle of movement again. Figure 5.4 shows some examples of possible random movements.

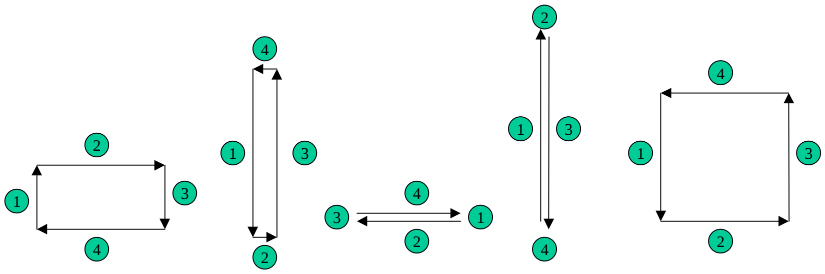


Figure 5.4: Examples of random movements

The parameter values used in the simulations are:

- Speed of the mobile station: 5 km/h.
- Maximum number of steps: 90 (giving a maximum distance of 3.1815 km from the original position).
- Maximum waiting time: 360 s.

The mobile station on a highway performs a horizontal or vertical trajectory depending on the direction of the highway. The movement starts at the beginning of the simulation and only stops at the end of the simulation. Mobile stations on a vertical highway start the movement going up on the cartographic grid until they reach the upper bound of the cellular network, then they change the direction to down, until they reach the lower bound of the cellular network, changing direction to up and so on. Mobile stations on a horizontal highway follow the same process, but left and right instead up and down. The vehicular mobile stations have a speed of 40 km/h; the walking mobile stations have a speed of 2 km/h.

5.2.3 Base Station

The main functions of the base station is to perform SNR measurements on all received voice packets and act accordingly with the result of the measurements and the origin of the voice packet, and to keep an updated profile of each active mobile user in the cell.

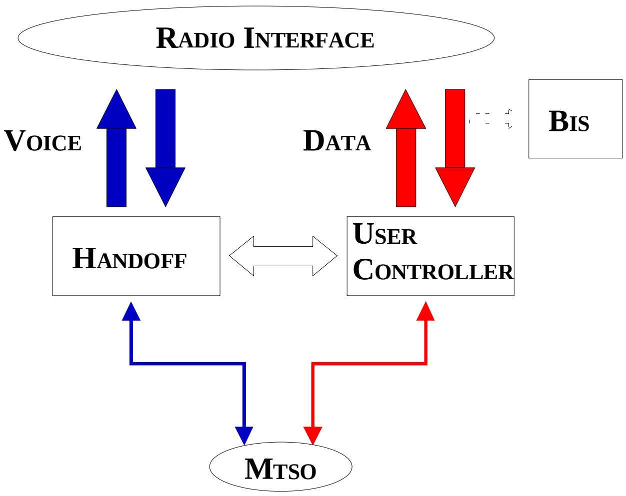


Figure 5.5: Base station node model

5.2.3.1 Handoff Module

The process that implements the handoff module is responsible for monitoring all voice packets it receives. It also keeps an updated list of all active calls within the cell.

When the base station receives a voice packet from an active mobile station within the cell, the signal strength of the received packet is compared against the handoff threshold. If the signal strength is less than the handoff threshold, a handoff request for that mobile station is sent to the MTSO. If the response from the MTSO is successful, the base station forwards the handoff order (inside the voice channel) to the mobile station with the information about the new cell and frequency channel, and also it takes the MIN of the corresponding mobile station off the active calls list. If the handoff is not possible, the MTSO sends a handoff reject message and the base station only updates the handoff reject statistics. If the next voice packet is still under the handoff threshold a new handoff request is made, but if the signal strength of the voice packet continues to decrease below the *call-dropping threshold*, the call is terminated and the call server process in the mobile station is interrupted to forcibly terminate the call.

The base station also receives voice packets from mobile stations that are active in neighbouring cells. The base station compares the signal strength of the received voice packet against a system-defined threshold. If the signal strength is greater than this threshold the base station forwards the voice packet to the MTSO in order to record this cell as a possible handoff cell candidate for the caller who sent the voice packet.

When the base station receives a call termination packet in the voice channel from one of its active mobile stations, it directs this information to the MTSO and to the user controller, and takes the MIN of the corresponding mobile station off of the active calls list.

5.2.3.2 User Controller Module

The user controller module has four main tasks:

- to deal with all signalling messages in the control channels
- to keep updated information about active mobile stations
- to keep, and update when needed, the frequency planning configuration of the cellular network
- to manage the allocation of channels in the cell.

The user controller module consists of five different processes: *perception*, *user controller*, *channel manager*, *frequency manager* and *database*. The processes were created using OPNET's dynamic parent-child process model capability.

The *perception process* works as an environment sensor and dispatcher, receiving all signalling messages and directing them to the responsible processes.

The *user controller process* is responsible for keeping track of all changes occurring in the status of an active mobile station, since the call set up until the call termination or handoff order. Each active mobile station within the cell is classified as *new*, *departing* or *others*. This classification is not related to AMPS standards, it is an extra feature added to the cellular model in order to keep more information about the active mobile stations in the cell. When the user controller process receives a new call or handoff request and a channel is granted, the mobile station identification is inserted in the *user classification list*. If it is a new call, a successful call set-up message is sent to the MTSO with the identification of the caller and the base station. If it is a handoff request, a successful handoff response is sent to the MTSO. The mobile station is classified as *new* when inserted in the *user classification list*. After a period of time ***t*** after the insertion and throughout the duration of the call inside this cell, the mobile station is classified as *others* or *departing*, depending on the SNR measurements made by the handoff module over the voice packets sent by the mobile station. The caller is classified as *others* if the SNR of its voice packets is greater than a threshold called *departing threshold*. If its SNR is less than the *departing threshold*, the caller is classified as *others*, but a timer is started and when this timer expires the caller is classified as *departing* if its SNR is still less or equal than the *departing threshold*. Each time the handoff module receives a voice packet of a caller within the cell, the classification of the caller is checked for update. Therefore a *departing* mobile station can become *others* again if its SNR comes back to be greater than the *departing threshold*. When the call is terminated, the user controller process removes the information related to the mobile station and sends a call termination message to the MTSO.

The *channel manager* and the *frequency manager processes* are responsible for the allocation and release of the assigned channels of the cell. They deal with the channel request for new calls and handoffs and send the appropriate response to these requests. They implement the FCA strategy. The channel manager updates statistics such as the number of a specific request type it receives, and the number of successful and unsuccessful responses. The frequency manager keeps track of the traffic load in the cell, updating the percentage of available channels in the cell every time a channel is allocated or released. It is worth mentioning that in AMPS systems the allocation of the individual channels of each cell is the responsibility of the MTSO. This feature of the system was modified in this cellular model, distributing the task to the corresponding base stations. However this modification does not change the model statistics.

Finally the *database process* is responsible for keeping the information about the cellular system that affects the cell where the base station is placed. It has the identification of its neighbouring base stations, the base stations that are its co-channel cells, all base stations inside its compact pattern, all the base stations inside the channel reuse distance that are not its co-channel cells. For all identified base stations, the database process has the information on the corresponding set of frequency channels being used.

5.2.3.3 BIS Module

This module monitors the busy status of the RECC receiver and sets a BIS flag as the status changes. The mobile station checks this flag reading the object attribute linked to it, therefore emulating the reading of the BIS field in the FOCC performed by a mobile station when trying to seize the RECC.

5.2.4 MTSO

In AMPS systems, the MTSO offers several services like Home Location Register (HLR), Visitor Location Register (VLR), networking services between different cellular networks (IS-41), billing and mobile tracking. In the simplified model here, only the mobile tracking is relevant for the channel allocation investigation. The most important function of the MTSO is to keep the information necessary about a specific active mobile station in order to be able to decide to which cell the mobile station needs to handoff if such request is received.

The MTSO node consists of a module called call manager and the full-duplex (voice and data) links to each base station in the system as shown in Figure 5.6.

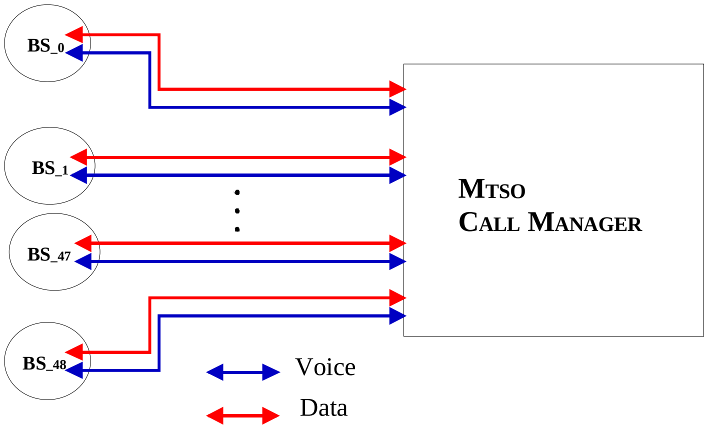


Figure 5.6: MTSO node model

5.2.4.1 Call Manager Module

The *call manager process*, which implements this module, keeps track of each active mobile station by keeping a list of all calls in progress in the cellular network. It is a two-dimensional list (a list of lists). Each element of the master list corresponds to a call in progress. When a new call is established a new element is inserted in the master list. The linked list attached to each element of the master list consists of data structures that contain information on the current signal strength for the call at each base station able to receive above a threshold [OPNETb94].

This process is responsible for the selection of the cell to which an active mobile station needs to be handed over. If a channel is available in the selected cell a handoff order is issued, otherwise a handoff failure is reported to the base station that requested the handoff.

When a call terminates, the corresponding element in the master list is removed and destroyed together with its linked list.

5.3 Cellular Network with D-BA Simulation Model

For the simulations of the D-BA scheme, the basic cellular network model has to be modified. Since the only difference is the channel allocation scheme, the user controller module of the handoff node (section 5.2.3.2) and the call manager module in the MTSO node (section 5.2.4.1) needed alterations. The modifications in the call manager module were minimal, the only feature added is the redirection of signalling messages between base stations. The user controller module, however, needed drastic modifications because the D-BA scheme is implemented entirely inside this module. Therefore the name of this module was changed to D-BA and the base station node now is represented by the model shown in Figure 5.7.

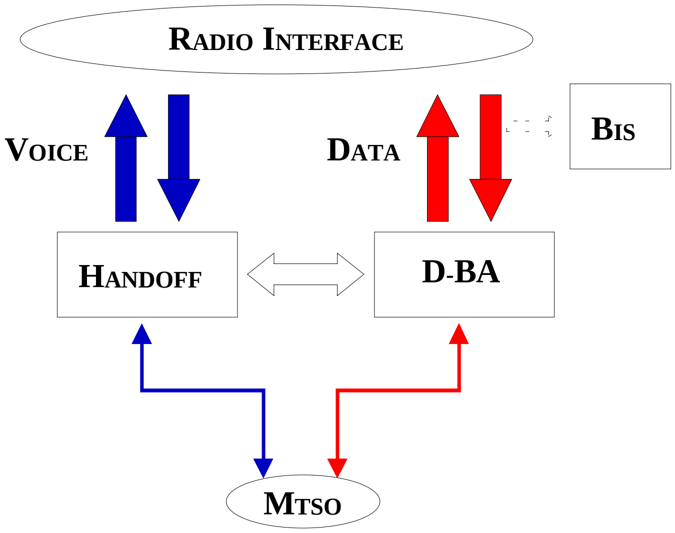


Figure 5.7: Base station node in D-BA cellular network model

5.3.1 D-BA Module

The main tasks of the D-BA module are still the same as the user controller module in the conventional cellular model, but now follow the D-BA scheme for the channel allocation. The D-BA module consists of the following processes: *perception*, *user controller*, *channel manager*, *frequency manager*, *database*, *channel borrower*, *channel lender*, *lender manager* and *channel locker*.

All the processes will interact with each other in order to perform the D-BA scheme exactly as described in Chapter 4, section 4.3.1.2 and also to perform the channel reassignment as described in section 4.3.2.

The functionality of the *perception, user controller* and *database* processes is the same as described in section 5.2.3.2, the only difference being that they are prepared to deal with the signalling messages exchanged when the channel borrowing algorithms are being executed.

The *channel manager* and the *frequency manager* processes keep a consistent and updated record of the status of all channels in the cell. Now the cell can have nominal and borrowed channels. The nominal channels can have the following status: available, in use, lent or locked; a borrowed channel can be available or in use. The channel manager is responsible for triggering the channel-borrowing algorithm, the decision being based on the rules of the D-BA algorithm as explained in section 4.3.1.2. The frequency manager decides what channel will be allocated when a channel request is received. It also deals with channel borrowing and channel locking requests. For the cellular network using only the D-BA scheme, the frequency manager also performs the channel reassignment (Local Planning Layer functionality in the agent model, section 4.3.2) and the channel release when a call is terminated. It also decides when a borrowed channel will be returned to its original cell.

The *borrower* process is the heart of the D-BA scheme. The channel manager invokes the borrower process every time the channel-borrowing algorithm needs to be executed. The borrower process performs the initialisation phase described in section 4.3.1.2 and starts the algorithm cycle to borrow channels. It executes all the steps due to the borrower cell inside the algorithm.

The *lender* process receives the signalling messages of borrower cells. For each borrower cell that has initiated the algorithm cycle and sent a request to this cell, one *lender manager* process is instantiated. The lender manager will be responsible for all requests from only one specific borrower cell. The lender manager process is destroyed when the borrower cell that it was serving advertises that it came back to normal levels of channel availability.

The *locker* process receives the requests of co-channel cells that are lending channels to other cells. The locker process will signal the channel manager process and lock channels if possible, sending the appropriate response to the requesting lender cell.

5.3.1.1 Signalling Message Exchange for Channel Borrowing

Figure 5.8 depicts the signalling flows between the processes that compose the D-BA module. For clarity, the perception and the database processes were omitted. The perception process receives all messages from the environment (mobile station, MTSO, handoff node) and dispatches the messages to the corresponding processes. All processes inside the D-BA module interact with the database. In Figure 5.8 the group of signalling messages exchanged between two processes or between a process and the environment is represented by an arrow and a number.

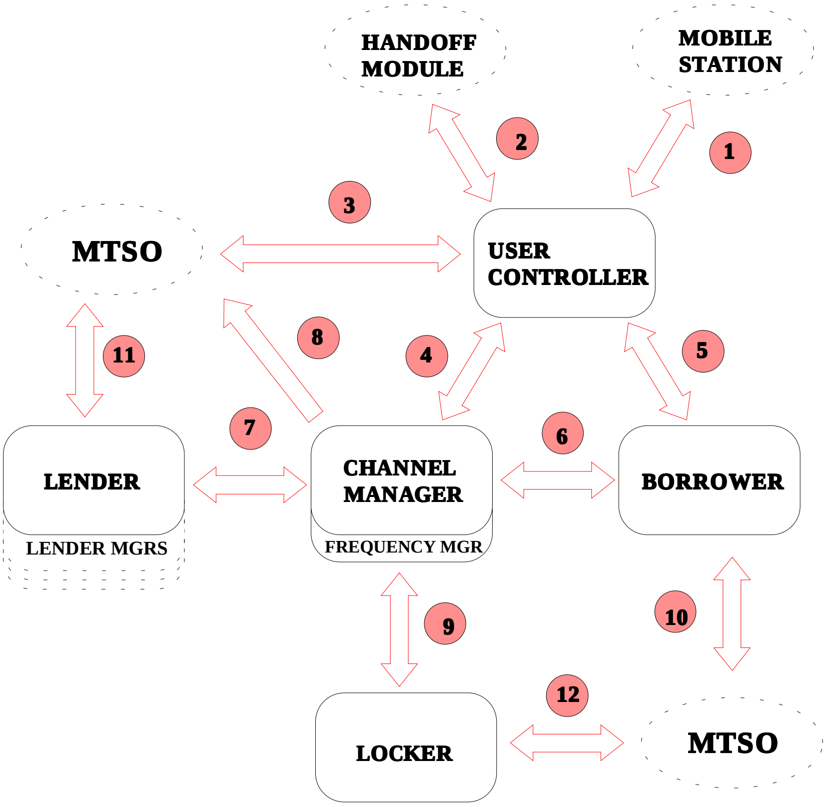


Figure 5.8: Signalling message exchange in D-BA module

A brief description of the signalling messages and its meaning is given below:

Group 1:

MOBILE STATION ▸ USER CONTROLLER PROCESS	
call_ request_ msg	a mobile is asking for an available channel to start a call
USER CONTROLLER PROCESS ▸ MOBILE STATION	
call_ response_ msg	response of the success or failure in a call set up

Group 2:

USER CONTROLLER PROCESS ▸ HANDOFF MODULE	
assigned_ call_ msg	information about a new call that has been set-up
HANDOFF MODULE ▸ USER CONTROLLER PROCESS	
voice_ packet	SNR information about a call being managed by the cell
call_ termination_ msg	information about a call that has been terminated
handoff_ mgs	response of a handoff request or handoff order information

Group 3:

USER CONTROLLER PROCESS ▸ MTSO	
assigned_ call_ msg	informing the MTSO about a new call that has been set-up
handoff_ channel_ rsp	success or failure of a handoff channel request
MTSO ▸ USER CONTROLLER PROCESS	
handoff_ channel_ rq	MTSO is requesting a channel to handoff a call to this cell

Group 4:

USER CONTROLLER PROCESS ▸ CHANNEL MANAGER PROCESS	
newcall_ channel_ rq	channel request for a new call to this cell
handoff_ channel_ rq	channel request to handoff a call to this cell
call_ termination_ msg	information about a call that has been terminated
handoff_ reject_ msg	internal information about a reject handoff request
CHANNEL MANAGER PROCESS ▸ USER CONTROLLER PROCESS	
newcall_ channel_ rsp	success or failure of a new call channel request
handoff_ channel_ rsp	success or failure of a handoff channel request

Group 5:

USER CONTROLLER PROCESS ▸ BORROWER PROCESS	
ms_ depart_ info_ rsp	information about the current departing mobile stations
BORROWER PROCESS ▸ USER CONTROLLER PROCESS	
ms_ depart_ info_ rq	request information about current departing mobile stations

Group 6:

CHANNEL MANAGER PROCESS ▸ BORROWER PROCESS	
start_ borrowalg_ msg	triggers the execution of the borrowing algorithm
BORROWER PROCESS ▸ CHANNEL MANAGER PROCESS	
rcv_ borwdchs_ msg	channels obtained with borrowing algorithm execution

Group 7:

LENDER (LENDER MANAGER) PROCESS ▸ CHANNEL MANAGER PROCESS	
lend_ chs_ rq	request one or more channels to be lent to other cell
ret_ lent_ chs	lent channel(s) being returned to the cell
CHANNEL MANAGER PROCESS ▸ LENDER (LENDER MANAGER) PROCESS	
lend_ chs_ rsp	response about allocation of channels for lending

Group 8:

CHANNEL (FREQUENCY) MANAGER PROCESS ▸ MTSO	
ret_ brwd_ chs	borrowed channel(s) being returned to a specific lender cell

Group 9:

LOCKER PROCESS ▸ CHANNEL MANAGER PROCESS	
lock_ chs_ rq	request one or more channels to be locked
rls_ locked_ chs	locked channel(s) being released to the cell
CHANNEL MANAGER PROCESS ▸ LOCKER PROCESS	
lock_ chs_ rsp	response about locking of channels

Group 10:

BORROWER PROCESS ▸ MTSO	
ch_ availability_ rq	request to inform *degree of coldness* of the receiver cell
calc_ F_ rq	request to a possible lender cell to calculate utility function F
borrow_ chs_ rq	request a lender cell to lend channel(s)
MTSO ▸ BORROWER PROCESS	
ch_ availability_ rsp	cell is informing its *degree of coldness* to requester cell
calc_ F_ rsp	possible lender cell calculated F and is sending the response
borrow_ chs_ rsp	zero or more channels lent to borrower cell

Group 11:

LENDER PROCESS ▸ MTSO	
ch_ availability_ rsp	cell is informing its *degree of coldness* to requester cell
calc_ F_ rsp	possible lender cell calculated F and is sending the response
borrow_ chs_ rsp	zero or more channels lent to a borrower cell
lock_ chs_ rq	request one or more channels to be locked
rls_ locked_ chs	release specified locked channel(s) in the receiver cell
MTSO ▸ LENDER PROCESS	
ch_ availability_ rq	inform own *degree of coldness* to a requester cell
calc_ F_ rq	calculate utility function F
borrow_ chs_ rq	a borrower cell is requesting channel(s)
ret_ brwd_ chs	borrowed channel(s) being returned to this lender cell
lock_ chs_ rsp	response about locking of channels

Group 12:

MTSO ▸ LOCKER PROCESS	
lock_ chs_ rq	request one or more channels to be locked
rls_ locked_ chs	release specified locked channel(s) in the cell
LOCKER PROCESS ▸ MTSO	
lock_ chs_ rsp	response about locking of channels

The following diagrams show the signalling message exchange during the execution of a borrowing channel algorithm cycle: firstly, a successful channel borrowing (Figure 5.9) and then the possible unsuccessful cases (Figure 5.10).

The diagrams describe the signalling messages exchanged between a borrower cell, a lender and/or one of its co-channel cell that can cause interference with the borrower in a successful channel-borrowing algorithm cycle execution. It is needed to bear in mind that, for a successful channel borrowing, the locking of channels need to be successful in all co-channel cells of the lender that can cause interference with the borrower.

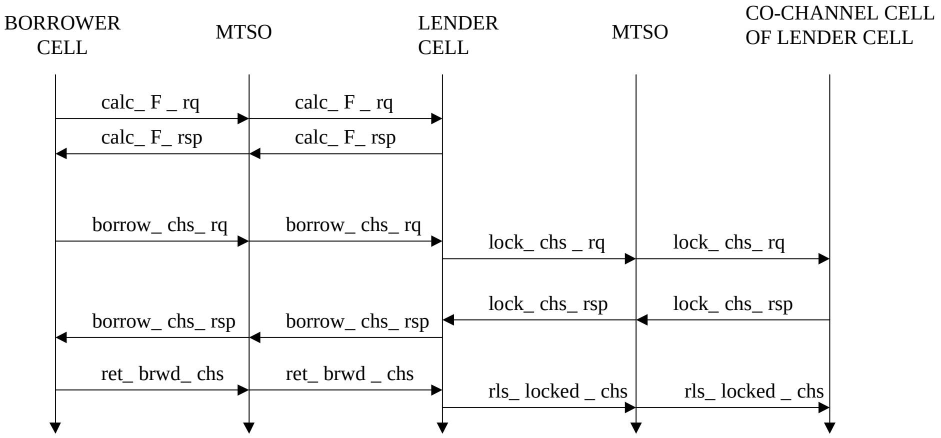
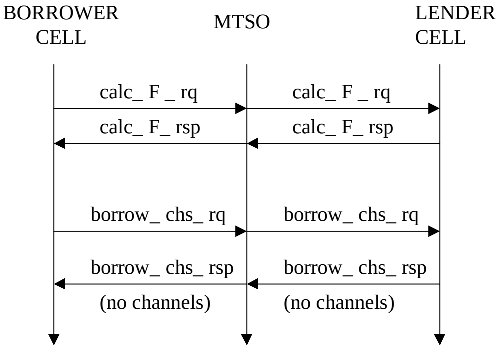
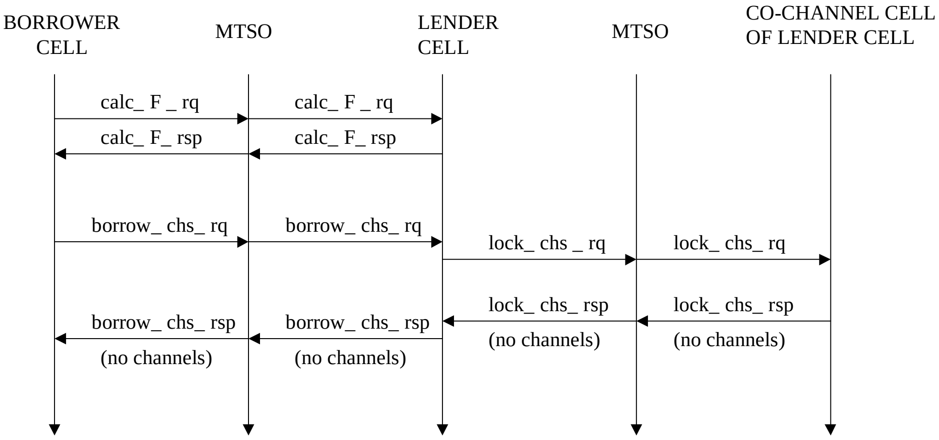


Figure 5.9: Signalling message exchange in a successful channel borrowing



(a) unsuccessful channel borrowing case 1



(b) unsuccessful channel borrowing case 2

Figure 5.10: Signalling message exchange in unsuccessful channel borrowings

5.4 Cellular Network and Multi-Agent System Simulation Model

Another set of modifications leads to the third cellular network model built. This incorporates the multi-agent system. There is one intelligent agent in each base station responsible for the channel allocation strategy. The implementation of the channel allocation strategy using intelligent software agents constitutes the only difference between this network model and the basic network model. However, one important feature has been added in this model, that is the capability of the system to perform *management handoffs*. A management handoff (*traffic handoff* as it is called in the AMPS/GSM specifications) is requested by a base station based on traffic conditions and SNR measurements and not only on SNR measurements as for the normal handoff request (*confinement* handoff) as modelled in the basic cellular model. This new feature was needed for the successful execution of joint plans (section 4.3.3.2).

The implementation changes between the cellular model with intelligent agents and the conventional one are confined in the following modules:

- An agent module, described in the following subsection, replaced the user controller module inside the base station node.
- The handoff module inside the base station node was altered to comply with the management handoffs.
- The call manager module inside the MTSO node was modified to comply with the management handoffs and to redirect signalling messages exchanged between agents.

5.4.1 Agent Simulation Model

The hierarchical structure of the agent architecture was implemented inside a module in the base station node. The different layers were created using OPNET's dynamic parent-child process model capability. Each layer is composed of a different group of processes. Figure 5.11 illustrates how the conceptual agent model (Chapter 3, section 3.2.1) was adapted and implemented inside an OPNET module.

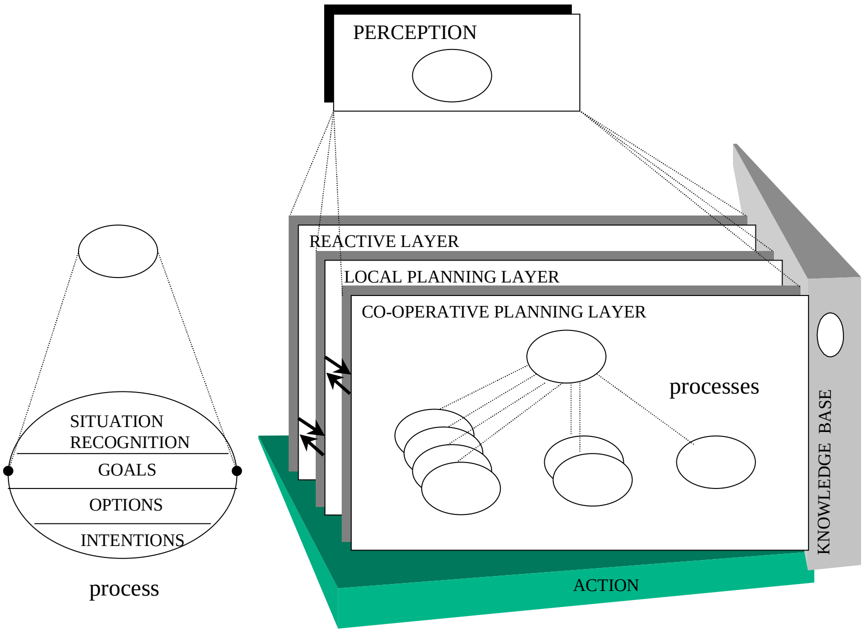


Figure 5.11: Implementation of the conceptual agent model

The perception process comprises the sensor section and part of the communication section of the agent architecture. The action plane represents the execution of tasks performed by the processes on the environment and it comprises the actor section and the remaining communication section of the agent architecture. All control layers intersect the action plane and the knowledge base plane. When a control layer is activated in response to an event in the environment, the overall behaviour of the processes inside the control layer constitutes the execution of the control cycle (section 3.2.1.1). The arrows between layers represent the activation requests and the commitments performed in a control path.

Specifically for the channel allocation strategy, the agent replaces the user controller module in the base station node as depicted by Figure 5.12.

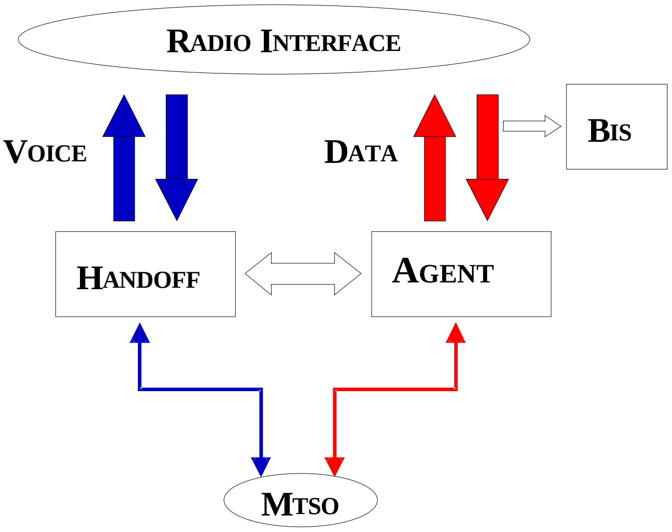
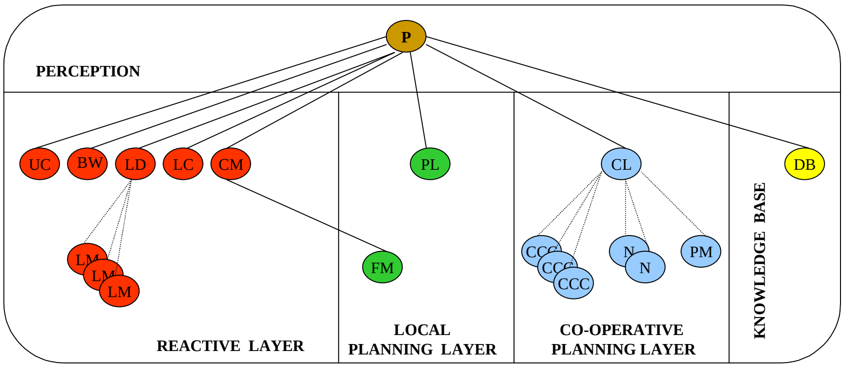


Figure 5.12: Agent module integrated in the base station node

Each layer of the agent consists of a different group of process. Figure 5.13 illustrates the process hierarchy and group division of all processes that make up the agent module.



P : perception process
UC: user controller process
BW: channel borrower process
LD: channel lender process
LC: channel locker process
CM: channel manager process
LM: lender manager process

PL: planning process
FM: frequency manager process
CL: co-operative layer process
CCC: joint plan co-channel cell process
N : joint plan neighbouring cell process
PM: joint plan manager process
DB: data base process

Figure 5.13: Process hierarchy inside agent module

The *perception* process is now capable of interfacing the entries for all three layers of the agent. The *database* process provides information for the three layers.

As explained in Chapter 4, the Reactive Layer makes use of the D-BA scheme as channel allocation algorithm, so that the simulation modelling description made in section 5.3 describes the Reactive Layer of the agent model.

The Local Planning Layer consists of two processes: *local planning* and *frequency manager*. The local planning process monitors the borrower process in the Reactive Layer, analysing the outcomes of the borrowing algorithm runs and making the decision when the Co-operative Planning Layer needs to be triggered to start a joint plan. The frequency manager was moved from the Reactive Layer to the Local Planning Layer because the decision making for reassignment and release of channels depend directly on planning strategies in the this layer.

The Co-operative Planning Layer consists of a master process and possible child processes that will be instantiated depending on the role of this agent in a joint plan negotiation or execution. The master process is called the *co-operative layer* process and is capable of responding to any signalling message belonging to the *Contract-Net Protocol* (CNP) in any kind of role concurrently. The roles the agent can assume are a manager candidate or a manager of a joint plan, a co-channel cell of a manager (or of a manager candidate), or a neighbour (of a co-channel cell or a manager). For each role, different rules are applied concerning the engagement in a joint plan. For example, an agent cannot start a role of a manager candidate (starting a joint plan negotiation, i.e. first phase negotiation) if it is already serving another joint plan (in execution, i.e. second phase negotiation) as a co-channel cell or a neighbour. However, the agent can start a role of a manager candidate if it is not a member of a joint plan in execution, even if it is responding to other agents in other joint plan negotiations. The agent's decision making for each stage of a particular joint plan negotiation depends on the state of all other joint plan negotiations or executions in which this agent is taking part.

The possible child processes in the Co-operative Planning Layer are the *joint plan co-channel cell*, the *joint plan neighbouring cell*, and the *joint plan manager*.

A *joint plan co-channel cell* process is instantiated every time the agent can give continuity to a *call_for_proposal* message from a manager candidate agent in the beginning of a joint plan negotiation. The agent sends a refuse message straight away if it is already engaged in two executing joint plans. The *joint plan co-channel cell* process will perform the actions described in section 4.3.3.1 (first phase of the agent negotiation) for a co-channel cell of a manager candidate. If the joint plan negotiation in which this process is taking part progresses to the second phase (execution), then the *joint plan co-channel cell* process will perform the actions as a partner of the joint plan in execution as described also in section 4.3.3.2. If in some stage of the negotiation (in which this process is taking part), the agent can no longer engage the joint plan or its cell is not in the winning region, then the *joint plan co-channel cell* process is destroyed.

The other two kinds of child processes that the co-operative layer can have are only instantiated during joint plan executions. During the first phase of the agent negotiation, the *co-operative layer* process deals with the actions of a manager candidate or a neighbour of a co-channel cell.

A *joint plan neighbouring cell* process is instantiated if the agent is acting as a neighbour (it does not matter if it is a neighbour of a co-channel-cell agent or of a manager agent) in a joint plan in execution. It will perform the actions of a partner of the joint plan in execution as described in section 4.3.3.2.

A *joint plan manager* process is instantiated if the agent is acting as a manager of a joint plan in execution. It will perform the actions of the manager agent of the joint plan in execution as described in section 4.3.3.2.

In summary, for each joint plan in execution that the agent is engaged in, a special child process in the Co-operative Planning Layer will be instantiated and the type of the child process depends on the role of the agent in the joint plan.

The following subsection describes the signalling messages in the Co-operative Planning Layer during negotiation and execution phases of joint plans.

5.4.1.1 Signalling Message Exchange in the Co-operative Planning Layer

The messages exchanged between agents follow the FIPA-CNP as described in Chapter 4. However, to simplify the implementation in OPNET, some of the CNP act names were modified, but the same semantics were used. This was the case for the acts *inform*, *request*, *failure* and *cancel*, where the context of the message was added to its name, for example *inform_jp*, *reqjp_action*, *failure_jp*, *cancel_jp*, etc. The number inside curled brackets that appears in the end of the messages is the contents of the *content* parameter [FIPA97] and it was introduced in this description to make easier the understanding of the purpose of the message.

Figure 5.14 illustrates the groups of signalling messages exchanged between the co-operative layer processes and other agents (through MTSO) and the signalling messages exchanged inside the agent module.

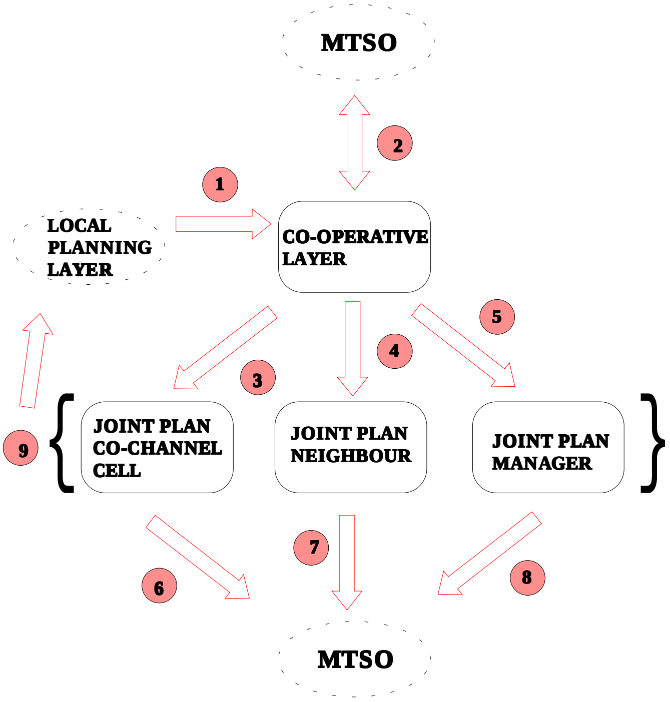


Figure 5.14: Signalling message exchange in the agent module (CPL)

A brief description of the signalling messages and their meaning is given below. For clarity, the role that the agent is performing when sending or receiving signalling messages is underlined.

Group 1:

LOCAL PLANNING LAYER (LPL) ▸ CO-OPERATIVE LAYER (CL) PROCESS	
start_ cf0_ msg	the LPL is requesting the CL process to try to announce a joint plan.

Group 2:

CO-OPERATIVE LAYER PROCESS ▸ MTSO	
cfp(0)	the manager candidate is asking its co-channel cells to assess their regions.
refuse(0)	the co-channel cell cannot propose a joint plan.
propose(0)	the co-channel cell assessed successfully its region and it is sending back a proposition to the manager candidate .
accept_ proposal(0)	the manager candidate is accepting one of the propositions.
reject_ proposal(0)	the manager candidate is rejecting one of the propositions.
request_ jp(0)	the manager is requesting its two neighbours inside the winning region to engage the joint plan.
failure_ jp(0)	the manager candidate accepted the proposition of this co-channel cell, but for some reason the co-channel cell is informing that the joint plan cannot be executed.
propose(1)	a neighbour of a co-channel cell is sending its assessment.
refuse(1)	a neighbour of a co-channel cell cannot send an assessment because it cannot engage any joint plan in near future.
propose(2)	a neighbour of a co-channel cell is informing that it can engage the joint plan.
refuse(2)	a neighbour of a co-channel cell cannot engage the joint plan before proposed by the co-channel cell to the manager candidate.
MTSO ▸ CO-OPERATIVE LAYER PROCESS	
cfp(0)	a co-channel cell is receiving cfp(0) from a manager candidate .
cfp(1)	a neighbour of a co-channel cell is being asked for assessment.

cfp(2)	a neighbour of a co-channel cell is being asked to engage a joint plan.
propose(0), refuse(0)	the manager candidate is receiving a proposition or a refusal from a co-channel cell.
propose(1), refuse(1)	the CL process forwards the message to the respective CCC process.
propose(2), refuse(2)	the CL process forwards the message to the respective CCC process.
accept_ proposal(0), reject_ proposal(0)	the co-channel cell is being informed that its proposition was accepted/rejected by the manager candidate.
accept_ proposal(2), reject_ proposal(2)	the neighbour of a co-channel cell is being informed to engage/not engage the joint plan.
failure_ jp(0)	the manager candidate is being informed that the proposition before accepted it is no longer valid for some reason.
inform_ jp(0)	the manager candidate is being informed that everything is ready to start the joint plan execution.
request_ jp(0)	the neighbour of a manager is being ordered to engage the joint plan.
cell_ jp_ inform(0), query_ ref(0), reqjp_ action(0)	redirect message to the respective N process or to the CCC process.
ptnrjp_ inform(0), inform_ ref(0)	redirect message to the joint PM process.
cancel_ jp(0)	the member of a joint plan is being informed that the corresponding joint plan is terminated.

Group 3:

CO-OPERATIVE LAYER PROCESS ▸ JOINT PLAN CO-CHANNEL CELL (CCC) PROCESS	
propose(1)	the co-channel cell is receiving the assessment of one of its neighbours.
refuse(1)	the co-channel cell is being informed that a neighbour (sender) cannot send its assessment.
propose(2)	the co-channel cell is being informed that a neighbour (sender) is ready to engage the joint plan.
refuse(2)	the co-channel cell is being informed that a neighbour (sender) cannot engage the joint plan.
accept_ proposal(0)	the co-channel cell had its proposition accepted and now it has to send cfp(2) to its neighbours.
cell_ jp_ inform(0)	the co-channel cell is being informed that the corresponding joint plan is starting the execution phase.
query_ ref(0)	the co-channel cell needs to send its load information to the manager .
reqjp_ action(0)	the co-channel cell is being informed (by its manger) that it needs to perform management handoffs.
cancel_ jp(0)	the co-channel cell is being unformed that the corresponding joint plan has finished and it needs to clean up to be destroyed.

Group 4:

CO-OPERATIVE LAYER PROCESS ▸ JOINT PLAN NEIGHBOUR CELL (N) PROCESS	
cell_ jp_ inform(0)	the neighbour is being informed that the corresponding joint plan is starting the execution phase.
query_ ref(0)	the neighbour needs to send its load information to the manager.
reqjp_ action(0)	the neighbour is being informed (by its manger) that it needs to perform management handoffs.
cancel_ jp(0)	the neighbour is being informed that the corresponding joint plan has finished and it needs to clean up to be destroyed.

Group 5:

CO-OPERATIVE LAYER PROCESS ▸ JOINT PLAN MANAGER (PM) PROCESS	
ptnrjp_ inform(0)	the manager is receiving the id and departing regions ids of a joint plan partner agent.
inform_ ref(0)	the manager is receiving load information of a partner agent.

Group 6:

JOINT PLAN CO-CHANNEL CELL PROCESS ▸ MTSO	
cfp(1)	the co-channel cell is requesting a neighbour to assess itself.
propose(0)	the co-channel cell is sending a joint plan proposition to a manager candidate.
refuse(0)	the co-channel cell is sending a joint plan refusal to a manager candidate.
cfp(2)	the co-channel cell is asking a neighbour to engage the corresponding joint plan.
accept_ proposal(2)	the co-channel cell is informing the neighbour to actually perform the engagement to the joint plan.
reject_ proposal(2)	the co-channel cell is informing the neighbour that a joint plan with the manager candidate is no longer possible.
inform_ jp(0)	the co-channel cell is informing the manager that the co-channel cell and its neighbour are ready to start the joint plan execution.
failure_ jp(0)	the co-channel cell is informing the manager candidate that the proposition before sent it is no longer valid for some reason.
ptnrjp_ inform(0)	the co-channel cell is informing its own id and the its departing ids to the manager.
inform_ ref(0)	the co-channel cell is informing its load to the manager.

Group 7:

JOINT PLAN NEIGHBOUR PROCESS ▸ MTSO	
ptnrjp_ inform(0)	the neighbour is informing its own id and the its departing ids to the manager.
inform_ ref(0)	the neighbour is informing its load to the manager.

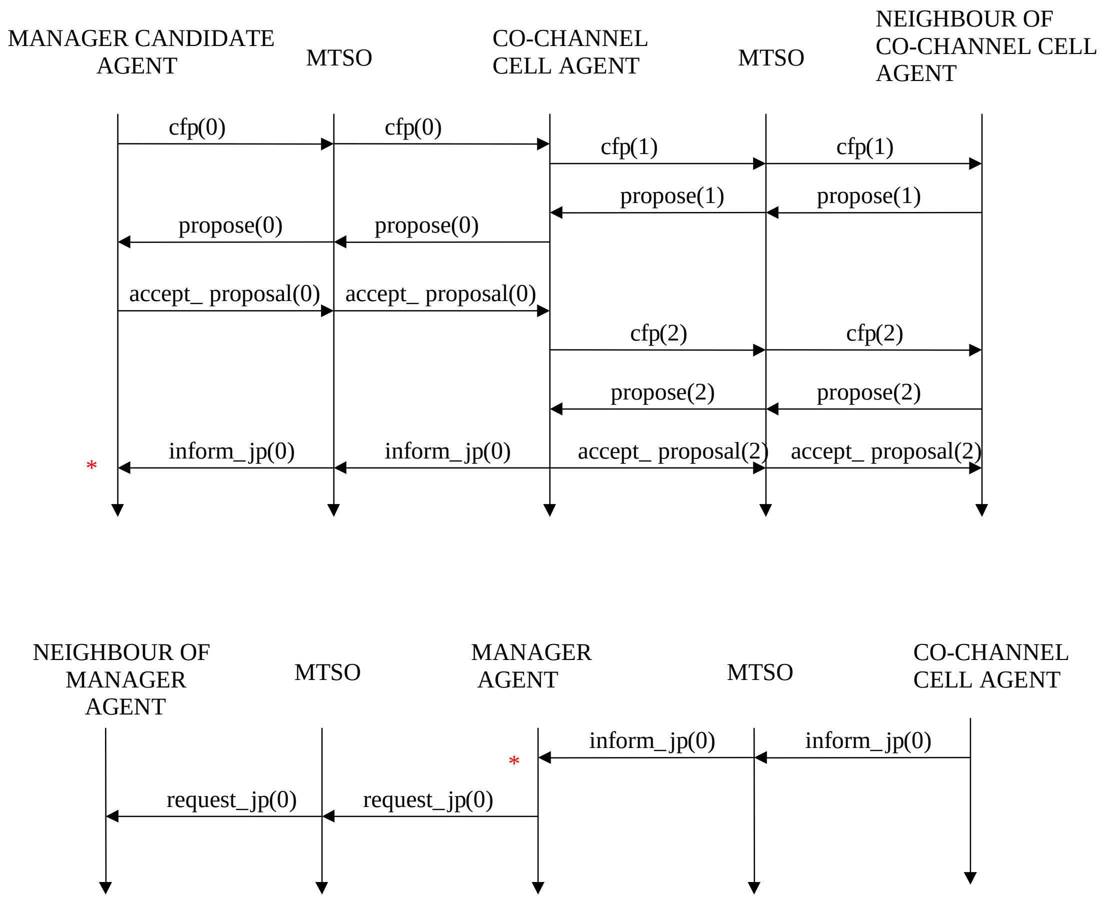
Group 8:

JOINT PLAN MANAGER PROCESS ‣ MTSO	
cell_ jp_ inform(0)	the manager is informing each joint plan partner that the corresponding joint plan is starting the execution phase.
query_ ref(0)	the manager is requesting load information of the receiving partner .
reqjp_ action(0)	the manager is requesting the receiving partner to perform management handoffs.
cancel_ jp(0)	the manager is informing each joint plan partner that the joint plan is being terminated.

Group 9:

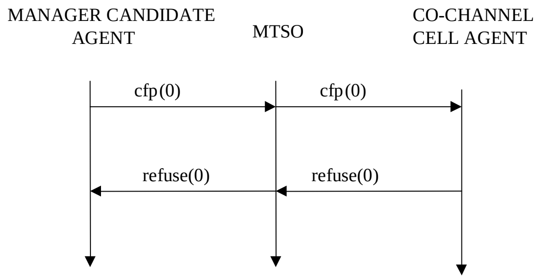
CL CHILD PROCESS ‣ LOCAL PLANNING LAYER ‣ USER CONTROLER PROCESS	
mgmt_ hoff_ rq	Co-operative Planning Layer is signalling the Reactive Layer to request management handoffs accordingly with the information inside the signalling message.

The following diagrams show the signalling message exchanges during joint plan negotiation and execution phases. Firstly, it shows the diagrams during a joint plan negotiation. For clarity, the diagrams show only the interaction of an agent trying to set up a joint plan, the signalling messages exchanged with one of its co-channel cell agents and the latter with one of its agent-neighbours. However, it is necessary to bear in mind that in an agent negotiation, all co-channel cell agents of a manager candidate will receive these messages and also all neighbours of the co-channel cell agents. Examples of possible outcomes of the negotiation phase are shown in separate diagrams in Figure 5.15 and Figure 5.16.

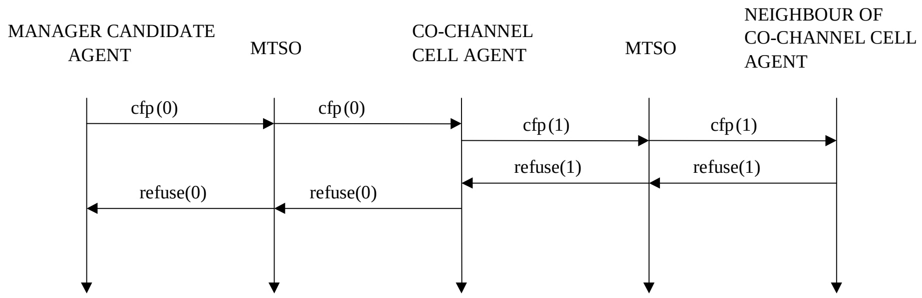


* - continuation of the flow diagram

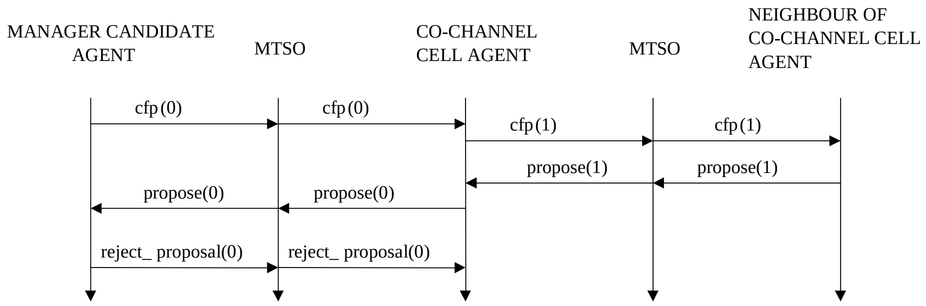
Figure 5.15: Successful joint plan negotiation



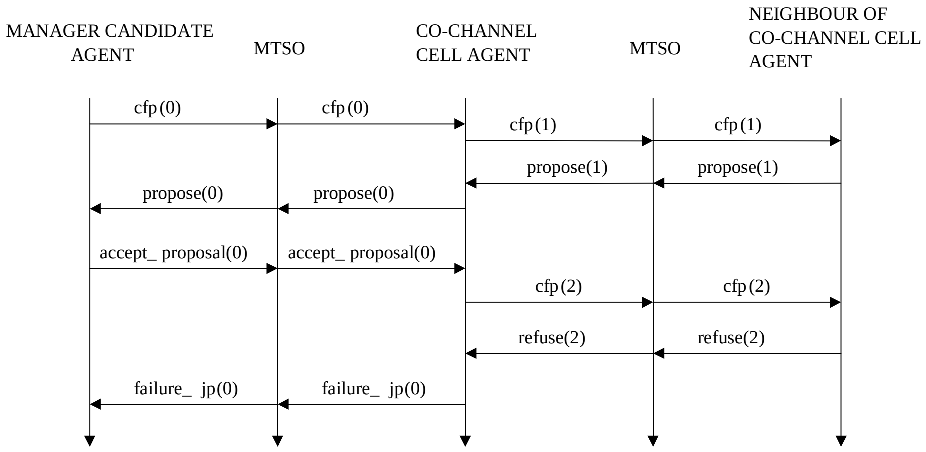
(a) unsuccessful negotiation case 1



(b) unsuccessful negotiation case 2



(c) unsuccessful negotiation case 3



(d) unsuccessful negotiation case 4

Figure 5.16: Different cases of unsuccessful joint plan negotiation

Figure 5.17 shows diagrammatically the signalling message exchanges during the execution of a joint plan. A signalling message with a dashed arrow means that the issue of this message may happen or not depending on the outcome of the manager's heuristic.

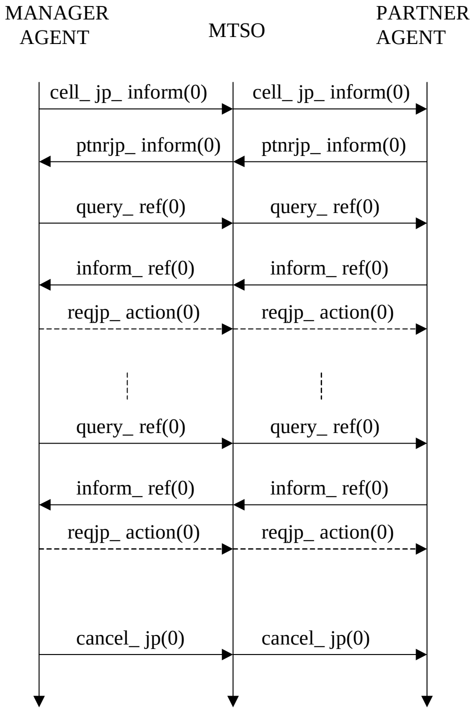


Figure 5.17: Signalling message exchanges during the execution of a joint plan

5.4.2 Handoff Module Modifications

The only modification needed in the handoff module (base station node) was to comply with management handoff requests. The agent module issues management handoff requests, which need to be processed and sent to the MTSO by the handoff module. Once the management handoff request was sent to the MTSO the expected responses are the same as a normal handoff request. The handoff module is responsible for keeping consistency of handoff requests belonging to a same mobile station, i.e. a departing mobile station cannot have a management handoff request and a normal handoff request being processed by the MTSO at the same time. Only the first request will go through towards the MTSO, the other will be discarded by the handoff module if the previous one is still being processed by the MTSO.

5.4.3 Call Manager Module Modifications

The call manager module, inside the MTSO node, processes separately the two types of incoming handoff requests, because the procedure for the selection of the cell for the handoff is different. However, once the target cell is selected, all procedures are the same for both kinds of handoff requests. This was a important implementation issue in order to keep consistency in the statistics being measured in the cellular network model.

Another modification was the capability of the MTSO to redirect accordingly signalling messages exchanged between agents.

5.5 Simulation Modelling Summary

In this chapter the most important simulation modelling issues of this research work were discussed. In order to compare traffic performance characteristics, three cellular network models were built. The first is the conventional cellular model using the FCA strategy. The second cellular network model was based on the conventional model, but using a different channel allocation scheme, the D-BA scheme. A description of the signalling messages used in the execution of the borrowing algorithm was shown. Finally, the third cellular network model was also based on the first basic model built, but now the multi-agent system was introduced to perform the channel allocation strategy. The signalling messages used by the agents were described. A new networking feature was added in the third model: the management handoff procedure.

The verification and validation of the simulation modelling described in this chapter and other important issues about simulation parameters will be discussed in the next chapter.

Chapter 6 Simulation Results and Analysis

6.1 Simulation Modelling Verification and Validation

In order to determine whether the cellular network simulation model developed accurately represents a real system, it needs to be verified and validated. The *verification* determines whether the simulation model performs as intended and the *validation* determines whether the conceptual simulation model is an accurate representation of the system under study. If the simulation model and its results are valid and are used as an aid in making decisions, then the model is said to be *credible* [LK91].

In [LK91], several techniques are described to verify and validate a simulation model; the following techniques were used in this work.

6.1.1 Verification and Validation of the FCA Cellular Network Simulation Model

The FCA cellular network simulation model was verified using the OPNET ODB (debugging) functionality, allowing several kinds of traces and breakpoints to be applied during the execution of the simulation. The overall simulation results and the intermediate results obtained in traces and breakpoints were checked for consistency and coherency. Animation, using a special OPNET package, was also applied to check specific features of the cellular system, such as user mobility. The input probability distributions given by OPNET and used in this work were sampled and their sample mean compared with the desired ones. Finally, run-time checks end the simulation if inconsistencies are detected.

The simulation model was validated by comparing the simulation output data for the call blocking rate of the entire network against the Erlang B formula taking the same mean offered traffic and the same mean call holding time. In this specific validation scenario only call requests are considered as offered traffic in the cell, because mobile users do not have trajectories. Therefore the call-blocking rate is equal to the total traffic offered over the total traffic refused. The decision to use the call blocking rate as a validation statistic is because the Erlang B formula is extensively used for describing telephony systems [DS84] [Bear88], including FCA cellular networks as configured in this scenario [ZY89] [Lee95]. The technique used for the validation was the confidence-interval approach based on independent data for steady-state parameters [LK91].

The simulation model was considered to be in the steady state after 80,000 seconds simulation time. Figure 6.1 shows that this is a valid assumption. The call blocking rate measurements were sampled from 80,000 seconds to 180,000 seconds, following the *initial-data deletion* technique [LK91]. The number of measurements is large enough to consider that its sample mean has closely converged to the *steady-state mean* (*v*) that is equal to the *expected mean* of the output random variable, here the call blocking rate.

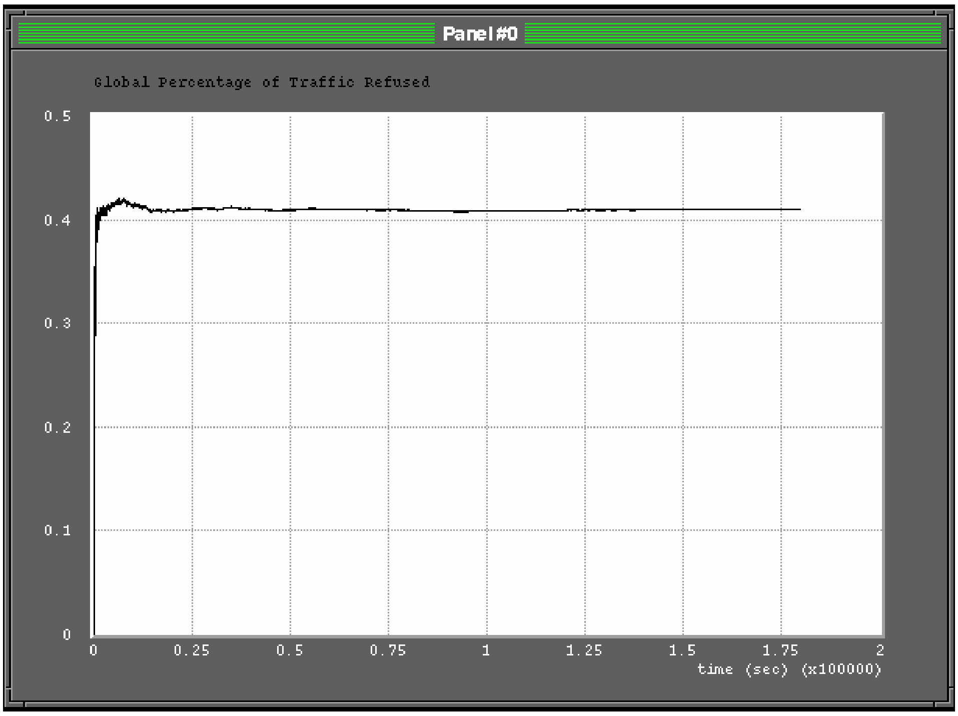


Figure 6.1: Call blocking rate versus simulation time

Figure 6.1 shows the OPNET output graph for the call blocking rate (Global Percentage of Traffic Refused) over the simulation time. Although the graph in Figure 6.1 is for an offered traffic of 15 Erlang, the output graphs for lower offered traffic follow the same general shape only with a lower call blocking rate. Hence the assumption of steady state after 80,000 seconds is reasonable.

For each offered traffic load, the estimation of *steady-state mean* (*v*) for the call blocking rate and its confidence interval was performed using the *replication/deletion* approach [LK91]. In this approach, the observations used for the estimates (call blocking rate measurements) are taken beyond the warm-up period *l* (here, 80,000 seconds). A number *n* ($n \geq 5$) of independent replications of the simulation needs to be performed, each one with length of *m* observations, where *m* needs to be a large number. In this work, five replications of the simulation were performed for each traffic load, using a different *seed* for the random generators in OPNET. Each simulation ran for 180,000 seconds, making the number of observations *m* very large ($m >$ 20,000).

The estimation of *v* is given by first taking the sample mean of the *m* observations (*Y*) of each replication (*i*) of the *n* replications:

$$\bar{Y}_i(m) = \frac{\sum_{j=1}^{m} Y_{ij}}{m} \qquad \text{for } i = 1, 2, .., n \tag{6.1}$$

Second, taking the sample mean of all $\bar{Y}_i(m)$:

$$\hat{v} = \bar{Y}(n) = \frac{\sum_{i=1}^{n} \bar{Y}_i}{n} \tag{6.2}$$

where $\hat{v}$ is the unbiased point estimator for *v*.

Finally an approximate 100(1 –) percent confidence interval for *v* is given by:

$$\hat{v} \pm \left(t_{n-1,1-\alpha/2} * \sqrt{S^2(n)/n} \right) \tag{6.3}$$

The value between brackets is called the *half-length of the confidence interval.* $S^2(n)$ is the *sample variance* given by:

$$S^2(n) = \frac{\sum_{i=1}^{n} \left[\bar{Y}_i(m) - \bar{Y}(n) \right]^2}{n-1} \tag{6.4}$$

The parameter t_n is given by Equation 6.5 and has a *t* distribution with *n*-1 degrees of freedom.

$$tn = \frac{\left\lfloor \overline{Y}(n) - v \right\rfloor}{\sqrt{\frac{S^2(n)}{n}}} \qquad (6.5)$$

Values of $t_{n-1,1-\ /2}$ are given in tables. For the validation of the results of five replications considering a 90% confidence interval, $t_{4,0.95}$ is needed and its value is 2.132 (Table T.1 of the appendix in [LK91]).

For the validation of the FCA cellular model, the overall traffic-blocking rate of the network was taken, as this is the major performance characteristic considered in this work. This is possible because the average traffic blocking probability in an entire cellular network (***PB***) is given by Equation 6.6 (from [ZY91]):

$$PB = \sum_{i=1}^{M} \frac{T_i}{\sum_{j=1}^{M} T_j} Pb(T_i, C_i) \qquad (6.6)$$

M is the total number of cells in the network, T_i is the offered traffic in cell ***i***, and $Pb(T_i, C_i)$ is the Erlang B formula (Equation 2.17) for cell ***i*** considering its offered traffic T_i and the number of channels allocated to the cell C_i among the total number of channels of the compact pattern. Equation 6.6 is general for uniform and non-uniform traffic load distributions in the cellular network. In the case of uniform load distribution and equal number of channels allocated per cell, $T_i = T$ and $Pb(T_i, C_i) = Pb$, therefore :

$$PB = M \frac{T}{MT} Pb = Pb \qquad (6.7)$$

The values for validation were taken from simulations of a 7-cell network. All the process models for the 7-cell network are exactly the same as those for the 49-cell network, but the execution time is much quicker. The average number of calls per hour is the same in all cells, starting from an average of 80 calls/h per cell, with the load in subsequent simulations being increased until it reaches 300 calls/h per cell. The load per cell follows a Poisson distribution with the mean being the selected average of calls/h. The call holding time was constant with a value of 180 seconds, the reason for that is discussed later. The exponentially distribution for the call holding time is not required in order that the results (of the Erlang B formula) hold for this case. The results are *insensitive* to the form of the call holding time distribution in this case [Dai92].

Table 6.1 gives the values of the average calls/h per cell, the resulting overall call blocking rate of the FCA cellular network in each simulation, the sample mean, the sample variation and the 90% confidence interval half length.

Table 6.1: Simulation results and confidence interval calculations

calls/h	seed 4	seed 20	seed 51	seed 75	seed 99	sample mean	sample variance	ci half length.
80	0.005298	0.006257	0.006551	0.004563	0.005834	0.005701	6.21139E-07	0.000751
100	0.01802	0.019183	0.017438	0.01812	0.019193	0.018391	4.04315E-07	0.000606
120	0.041656	0.043055	0.041298	0.0425	0.043713	0.042444	4.97342E-07	0.000672
140	0.078042	0.078368	0.075814	0.080053	0.079381	0.078332	2.29198E-06	0.001443
160	0.121384	0.118123	0.117716	0.120693	0.122337	0.120051	2.57809E-06	0.001531
180	0.165823	0.162686	0.168944	0.168442	0.170037	0.167186	6.28959E-06	0.002391
200	0.21394	0.211171	0.220559	0.21414	0.219273	0.215817	1.20056E-05	0.003304
220	0.25832	0.25871	0.259888	0.260551	0.259825	0.259459	8.01939E-07	0.000854
240	0.302758	0.299958	0.306001	0.301108	0.303881	0.302741	5.19469E-06	0.002173
260	0.341974	0.340113	0.342327	0.341777	0.34172	0.341582	7.25258E-07	0.000812
280	0.376791	0.378053	0.37728	0.376589	0.373985	0.376539	3.99532E-07	0.000603
300	0.408175	0.40987	0.411833	0.41019	0.409602	0.409934	1.68667E-06	0.001238

Figure 6.2 plots in a graph the values of the overall call blocking rate sample mean, and the limits of the half-length of the confidence interval.

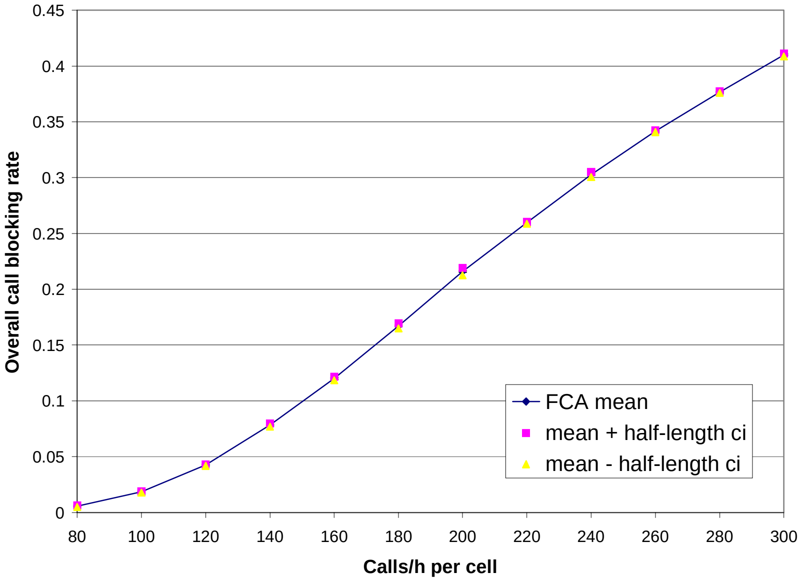


Figure 6.2: Overall call blocking rate with 90% confidence interval

As can be seen the values are very close to each other. To determine how representative the values shown in Figure 6.2 are, Figure 6.3 shows the FCA sample mean versus the values given by Erlang B formula considering the same average of calls/h per cell and the mean holding time of 180 seconds.

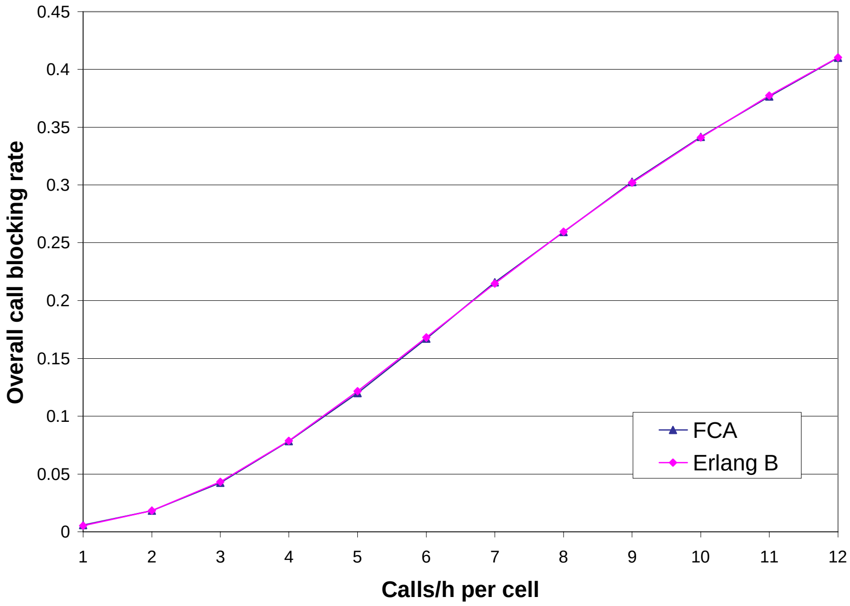


Figure 6.3: Comparison of overall call blocking rate given by the FCA network model simulation and Erlang B formula

Therefore the simulation model developed for the FCA cellular network can be considered valid.

To represent a realistic cellular system, a constant call holding time cannot be used and a negative exponential distribution with mean of 180 seconds should be used instead. This distribution has been used in all simulations where the three systems FCA, D-BA and multi-agent are compared. However, the negative exponential distribution given by OPNET for call holding time yields a lower mean than the expected. This is shown in Table 6.2, where now a negative exponential distribution with a mean of 180 seconds supplied by OPNET was used in each simulation. The minimum call duration was set to be 2 seconds, a value that can be considered negligible because is very close to zero and should not affect the distribution.

Table 6.2: Average call holding time (simulated)

calls/h	seed 4	seed 20	seed 51	seed 75	seed 99	sample mean
80	175.7863	174.3329	175.8036	176.981	174.593	175.4994
100	177.6885	175.8426	175.8606	175.6512	176.9302	176.3946
120	178.3809	178.3288	175.0556	177.9687	174.4097	176.8288
140	178.2147	177.1721	176.8435	174.6213	176.0334	176.577
160	177.681	176.3726	175.8019	174.545	173.032	175.4865
180	177.0934	176.457	175.3934	177.4615	176.7983	176.6407
200	176.4222	176.5711	176.1976	175.3506	175.065	175.9213
220	177.2215	176.3991	176.3565	175.3502	177.8627	176.638
240	176.4282	176.6223	176.0113	176.3661	176.566	176.3988
260	176.6725	176.3998	176.2363	175.7181	176.4263	176.2906
280	176.6449	175.5405	176.4758	176.9534	176.1811	176.3591
300	176.7144	176.3835	175.6368	176.7177	176.0347	176.2974

As can be seen the sample mean of the call holding time given by each simulation and the sample mean of the replications are lower than 180 seconds. The lower call holding time affects the overall call blocking as shown in Figure 6.4.

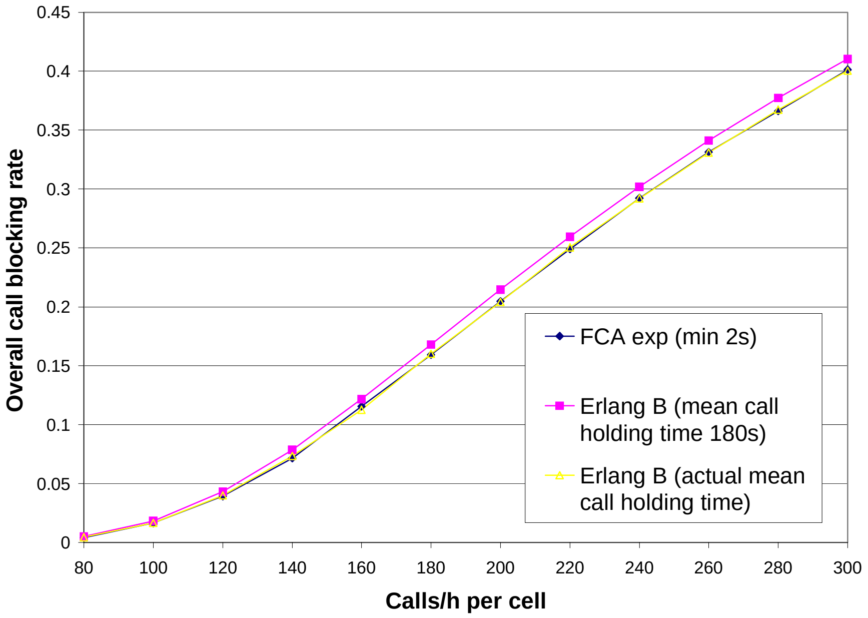


Figure 6.4: FCA network using OPNET negative exponential distribution with mean 180s and a minimum call duration of 2s versus Erlang B

As expected, the overall call-blocking rate is lower than the probability given by the Erlang B formula, because the mean holding time is lower than 180 seconds. In order to prove this statement, Figure 6.4 also shows the blocking probability given by the Erlang B formula, when the offered traffic used in its calculation is given by the average calls/h per cell and the actual average call holding time of the five replications. Now, the values for FCA and Erlang B are very close as expected, showing the validity of the system and also showing the discrepancy within OPNET.

In order to get results of traffic blocking comparable with the ones shown in the literature for a call holding time following a exponential distribution with mean 180s, a minimum call duration of 15 seconds was introduced. The overall call blocking rate for FCA versus the Erlang B formula (with a mean call holding time of 180 seconds) is shown in the graph of Figure 6.5.

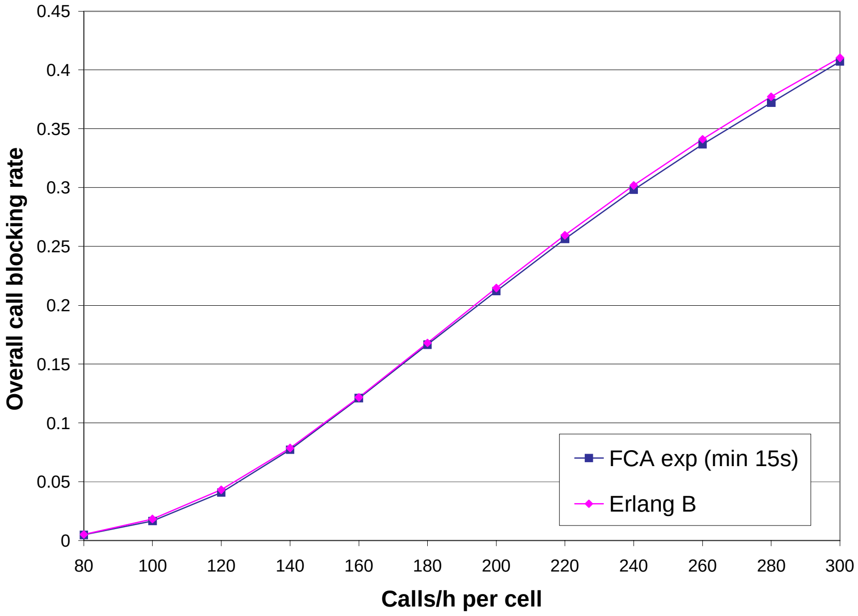


Figure 6.5: FCA network using OPNET negative exponential distribution with mean 180s and a minimum call duration of 15s versus Erlang B

The average call holding time is around 179 seconds when a minimum call duration of 15 seconds is used (as shown in Table 6.3) and the call blocking rates are then closer to the Erlang B formula using 180 seconds. The minimum call holding time is also quite realistic.

The minimum call duration of 15 seconds was used in the simulations of the three different cellular systems. This allows better comparisons with results in the literature.

Table 6.3: Average call holding time (simulated)

calls/h	seed 4	seed 20	seed 51	seed 75	seed 99	sample mean
80	179.549	180.2595	178.9407	178.5605	180.8018	179.6223
100	179.2792	180.3343	178.8608	179.1242	178.5182	179.2233
120	177.2375	178.3094	177.8046	180.9908	178.5276	178.574
140	181.0546	180.3108	177.9094	178.6921	177.4982	179.093
160	181.2971	179.692	179.342	178.9918	178.806	179.6258
180	179.2078	181.4784	179.0102	178.2552	180.1401	179.6183
200	178.4421	178.5511	178.74	177.9716	180.1336	178.7677
220	179.5833	179.6137	179.5018	177.593	178.78	179.0144
240	177.2708	180.0978	179.0394	178.7967	178.994	178.8398
260	179.5722	177.8328	177.6848	178.6288	180.2941	178.8026
280	179.8219	179.8936	179.5534	178.4898	179.7593	179.5036
300	178.9872	178.7327	180.2401	178.4993	179.1894	179.1298

6.1.2 Verification and Validation of the Cellular Network using the D-BA scheme and the Cellular Network using the Multi-Agent System

The two cellular systems were carefully verified using the same techniques used for the FCA cellular network, with special attention being given to the coherency and consistency of the channel allocation behaviour of each system.

Only a partial validation of the two systems was possible, because the traffic blocking rate given by the new channel allocation schemes cannot be compared to theoretical models such as the Erlang B formula. However, all the aspects of the two cellular networks except for the channel allocation schemes are the same as the FCA cellular network and therefore valid in those aspects.

The output data for the D-BA scheme could not be compared directly with the D-LBSB scheme [DSJA 97] because the implementations are not exactly the same and the simulation results shown in [DSJA 97] use very broad assumptions. For example, "*the call arrival in a cell is programmed as a Poisson process with inter-arrival time exponentially distributed with mean 1/ , and the call holding time is programmed as an exponentially distributed random variable with mean 1/* " [Sen97]. Values for are shown, but not for . Also the D-LBSB comparison with FCA and other schemes is performed using a cellular network with 100 cells, but the actual number of channels per cell is not explicitly shown (supposedly 100 channels per cell) and with a specific situation where there are 40 *hot* cells. The output data for the D-BA scheme is not as good as expected from the original description of D-LBSB, but, as it has been stated, the simulation parameters used in the simulations of the D-BA scheme are quite different from those used in the D-LBSB scheme. However, the D-BA output data is consistent and reasonable in all simulation scenarios where it was applied.

The output data given by the cellular network using the multi-agent system is also very reasonable and consistent, performing as expected. The cellular network using the multi-agent system is confidently believed to be a valid model as well.

6.2 Simulation Model Parameters and Traffic Load Scenarios

For simplicity, the three cellular networks being compared will be identified by their channel allocation schemes: FCA, D-BA and MA for the multi-agent system.

The cellular network configuration is the same for all three models. It is a 49-cell network, 10 channels per cell in a 7-cell cluster with 980 mobile stations (section 5.2.1). The call generation follows a Poisson distribution with mean equal to the specified average calls/h. The call holding time follows a negative exponential distribution with mean of 180 seconds. The minimum call duration is 15 seconds as explained in section 6.1.1. The simulation model parameters described below are used in the simulations. All three models use the following model and simulation attribute values:

Mobile Station:

a) Call server module
    - Voice packet update interval = 18 seconds.

b) Filter module
    - Self-location update interval = 60 seconds.

c) Trajectory module
    - Random trajectory maximum number of steps = 90.
    - Mobile station maximum stationary time = 360 seconds.

Base Station:

a) Handoff module
    - Handoff threshold = 18 dBm
    - Call dropping threshold = 10 dBm

b) User Controller process
    - Departing user threshold = 23 dBm

The threshold values chosen are comparable to those used in AMPS systems. The values for updates were selected in order to give good control over the calls or the location of the mobile station without overloading the simulation. The maximum distance performed by a mobile station in random movement is comparable with the cell radius.

The D-BA and the MA models use additionally the following attribute values:

1. Base Station:
    a) D-BA module or Agent module (Reactive Layer)
        - *Degree of coldness* (***dc***) threshold = 0.2
        - Minimum ***dc*** for a cell to lend channels = 0.3
        - Maximum number of channels that can be borrowed per request = 2
        - Threshold variation to return to cold state (above ***dc*** threshold) = 0.2
        - Borrowed channel release threshold = 0.2
        - Delay in executing next borrowing algorithm if outcome was partial success = 300 seconds
        - Delay in executing next borrowing algorithm if outcome was failure 2 = 600 seconds
        - Delay in executing next borrowing algorithm if outcome was failure 3 = 1200 seconds
        - Delay in executing next borrowing algorithm if outcome was failure 4 = 1500 seconds

Finally the MA model uses additionally the following attribute values:

1. Base Station:
    a) Agent module (Co-operative Planning Layer)
        - Number of iterations per joint plan = 10
        - Time interval between iterations in a joint plan = “voice packet update interval”
        - Minimum value of ***F’*** utility function (Equation 4.2) to send an accept-proposal to a region = 0.3
        - Parameter in ***F’*** utility function (Equation 4.2) = 0.08
        - max number of joint plans that an agent can engage simultaneously = 2
        - Delay in executing next the attempt for a joint plan if the current attempt was not successful = 300 seconds

The choice of values for the additional attributes of the D-BA and MA models was performed during the verification process using sensitivity analysis. A further discussion about parameter sensitivity will be given later.

In order to analyse the performance of the MA network in comparison to the D-BA and the FCA networks, four different traffic load scenarios were used. One scenario was selected, where the traffic load was uniformly distributed in the network (all the cells receive the same offered traffic) and three scenarios with non-uniform traffic load distribution. All the mobile stations had trajectories, either random movement or driving/walking. In all scenarios, the models were considered to be in the steady state after 80,000 seconds, as explained earlier. The following sections show the simulation results for each traffic load scenario. The measurements were sampled from 80,000 seconds to 180,000 seconds, following the *replication/deletion* approach [LK91] in the same way as was done in the validation of the simulation model (section 6.1.1). For clarity, the two half-lengths of the 90% confidence interval of the simulation results are not plotted on the graphs, only the sample mean of the five replications.

The main network performance measures used for the comparisons between the three cellular networks are:

- Traffic blocking rate: total number of rejected (call + handoff) requests / total number of (call + handoff) requests.
- Handoff rejection rate: total number of handoffs rejected / total number of handoffs requested.
- Call dropping rate: total number of calls dropped by weakness of signal /total number of accepted calls and successful handoffs.
- Performance of the borrowing algorithm.
- Performance of the management handoffs requested by the Co-operative Planning Layer of the agents engaged in joint plan.

6.3 Simulation Results for the Traffic Load Scenario 1

In this scenario, all cells start with an average of 60 calls/h (offered traffic of 3 Erlang) and the subsequent simulations increase the average by 20 calls/h until the last simulation reaches 200 calls/h per cell (10 Erlang). All mobile stations have random movement. For this scenario where the traffic load distribution was uniform, the simulation parameter “borrowed channel release threshold” was chosen to be 0 instead of 0.2 as explained later.

The simulated call holding time sample mean for the FCA, D-BA and MA networks is shown in Table 6.4. The values are very close to the desired mean of 180 seconds, proving that the choice of minimum call duration of 15 seconds was good.

Table 6.4: Simulated average call holding time

FCA	D-BA	MA
180.6358 s	180.8205 s	180.6243 s

Figure 6.6 shows the overall traffic blocking rate for the three cellular networks. The performance of the D-BA and the MA networks does not show any advantage for scenarios with uniform traffic load distribution; the D-BA network slightly under-performs the FCA. This was expected because when one cell reaches the threshold of channel availability all the other cells are also reaching (or are close to reaching) the same threshold. This makes the probability of a successful channel borrowing attempt low for increasing traffic loads. Moreover, when the channel borrowing attempt is successful, the channel locking in nearby cells increases the blocking rate because these cells are as loaded as the borrower cell. This is the reason why the threshold for releasing borrowed channels was selected to be zero. In uniform load distribution the sooner the borrower cell gives back its borrowed channels the better. For uniform traffic load distribution the conventional FCA is preferable to D-BA and also to MA (which uses the D-BA in its Reactive Layer), although the MA network is slightly better than the FCA because of the load balancing action of the Co-operative Planning Layer. However, the amount of signalling needed in the MA network is prohibitive when compared with the gain in channel allocation.

Schemes that use FCA with channel borrowing and channel locking will not perform well in uniform load distributions. The exception are the BDCL scheme [ZY89] and the BCO schemes [ESG82], where the number of channels that can be borrowed by other cells is adaptively controlled with the current traffic load in the cell, and they rely on intra-handoffs to release quickly borrowed channels. BDCL has better results because it also minimises the channel locking. However they are completely centralised schemes and generate heavier signalling load than the D-BA scheme. Dynamic channel assignment (DCA) schemes are more likely to outperform FCA in uniform traffic load distributions such as in [ZY91] [O-GL-R97].

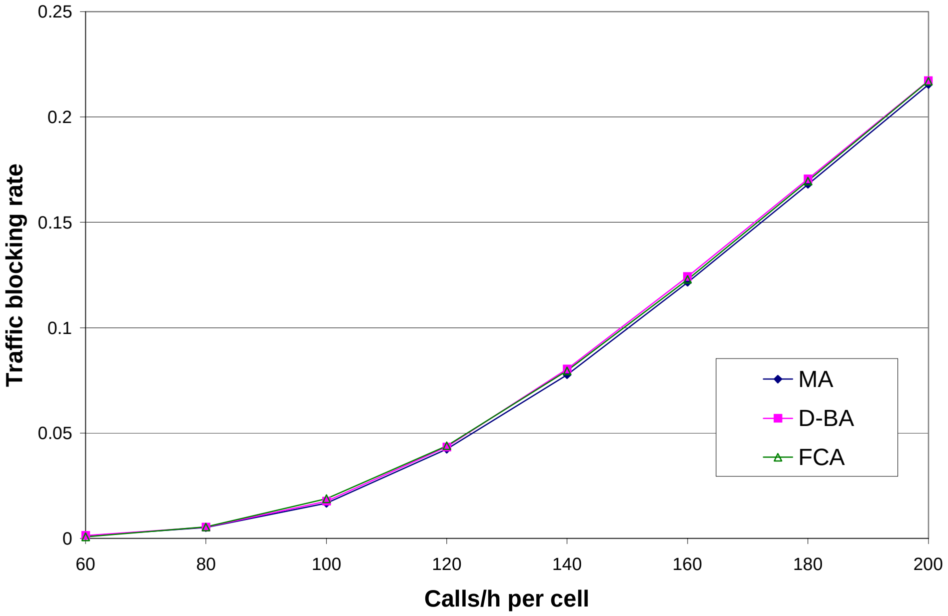


Figure 6.6: Traffic blocking rate for traffic load scenario 1

One good feature of the proposed multi-agent system is that it improves the handoff rejection rate. This is a consequence of the execution of the heuristic proposed by the author inside the agent joint plan and it proves the load balancing feature of the agent negotiation is behaving exactly as expected. Figure 6.7 shows the handoff rejection rate of the three networks.

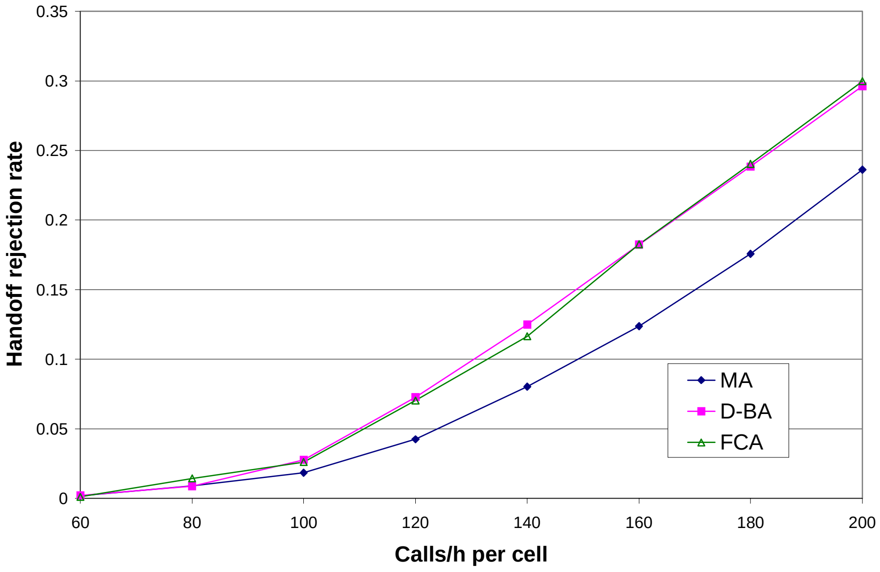


Figure 6.7: Handoff rejection rate in traffic load scenario 1

The call dropping rate graph (Figure 6.8) shows that the number of calls dropped is extremely low. This means that there are insufficient events to be of statistical significance (as evidenced by the rather large confidence intervals plotted for the MA network in Figure 6.8).

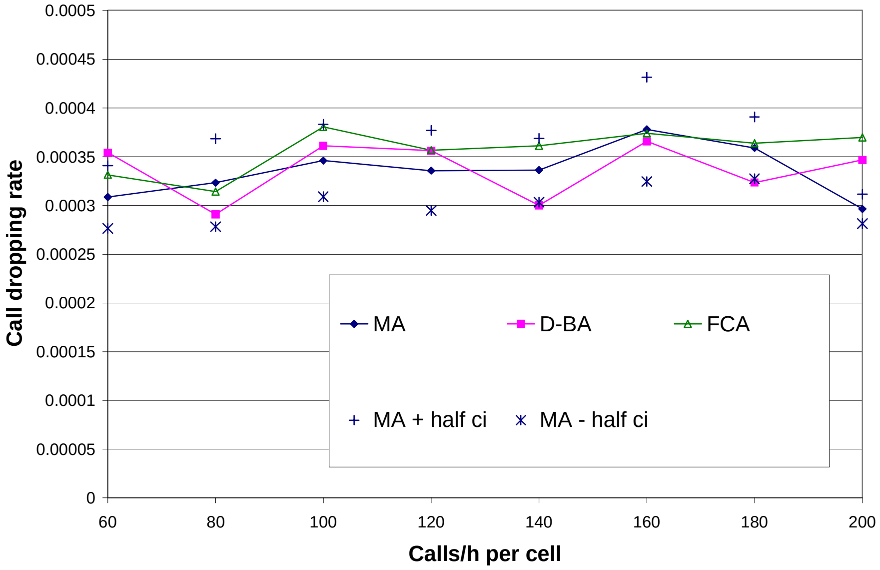


Figure 6.8: Call dropping rate in traffic load scenario 1

Uniform traffic load distributions are not common in real cellular networks and the MA model was actually designed to improve traffic conditions in non-uniform traffic load distributions, mainly in hot spots of traffic load.

The expected behaviour of the MA network is to improve the performance of the D-BA scheme in non-uniform traffic load distributions. The MA network will start to give better improvement than the D-BA network when the latter starts to decrease the efficiency of its borrowing algorithm. The improvement starts to decrease as the traffic load increases, because fewer resources will be available for load balancing. Therefore, the general form of the traffic blocking rate expected in the simulation result is shown in Figure 6.9.

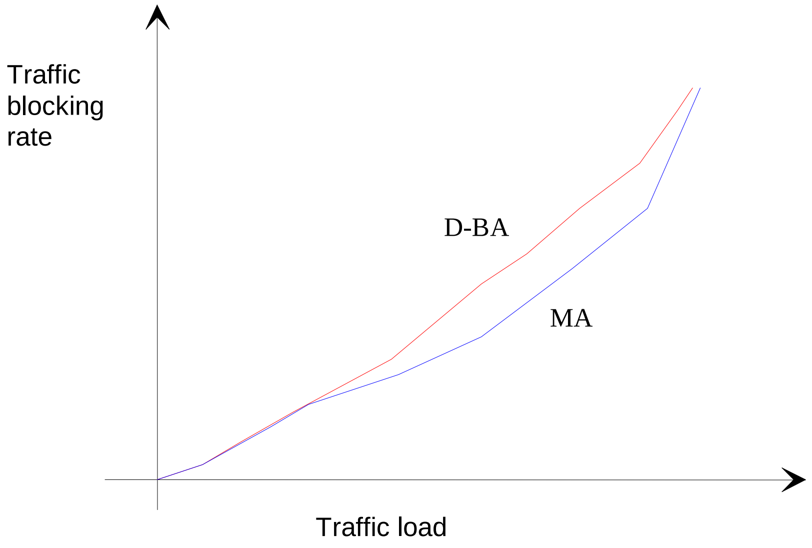


Figure 6.9: Expected behaviour of MA against D-BA for traffic blocking rate

The next three sections show how the MA network behaves with different load distribution scenarios.

6.4 Simulation Results for the Traffic Load Scenario 2

The layout of the network of the second scenario with non-uniform traffic load distribution is shown in Figure 6.10. This cellular layout was first used in the simulation tests in [ZY89], being also used in other papers in the literature. The number in the bottom of each cell is its identification number. The number in the middle represents the average of calls/h the cell is receiving (Poisson distribution). The shaded areas in the network layout represent highways. Mobile stations inside the shaded area can be driving at 40 km/h or walking at 2 km/h and they move in both directions. All the other mobile stations have random movement. The simulation parameters have the values shown in section 6.2.

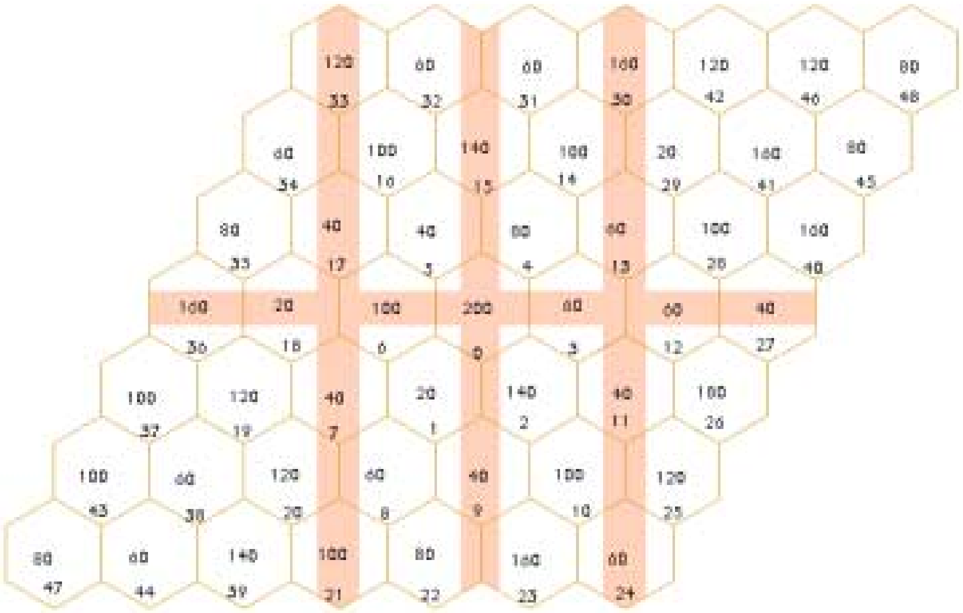


Figure 6.10: Cellular network layout for non-uniform load distribution of scenario 2

In the results, the abscissa of each graph is the percentage of load increase in all cells compared to the traffic load shown in Figure 6.10, called the *base load (0%)*. For example, in the first simulation result, the load in each cell is 40% less than the one shown in Figure 6.10 (e.g. the average of calls/h in cell 0 is 120). Figure 6.11 shows the overall traffic blocking rate of the three networks for different percentages of traffic load increase.

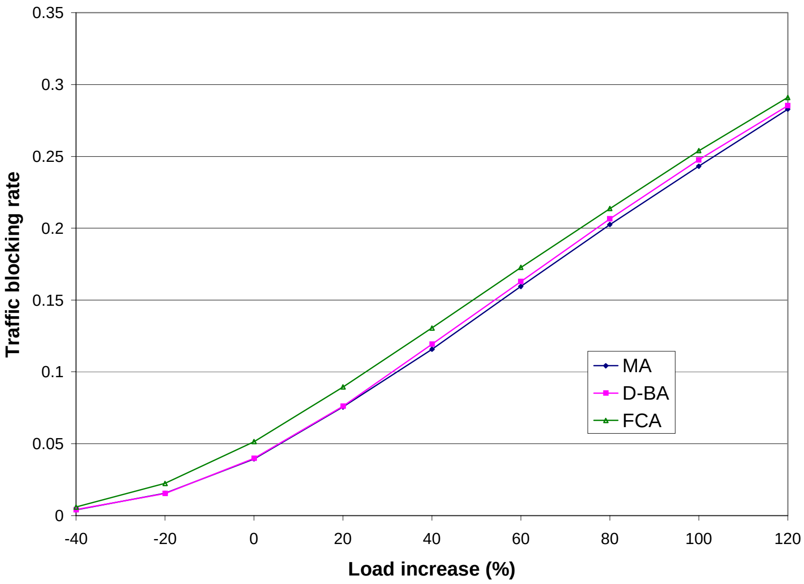


Figure 6.11: Traffic blocking rate for traffic load scenario 2

The MA network outperforms the D-BA and the FCA networks and the expected general behaviour of the MA network is demonstrated. The number of borrowing attempts is kept at the same level as the D-BA network, and at almost the same efficiency ((successful + partial successful outcomes) / total number of borrowing algorithm executions) (Figure 6.12). This result is interesting because the simulations show that actually the MA is not able to improve the efficiency of the borrowing algorithm as supposed. This means the current Local Planning Layer developed in this work could be improved in order to add a more adaptive behaviour to the Reactive Layer. The results show the agent negotiation is playing an important part in load balancing the traffic and consequently reducing the blocking rate, even without an improvement in the efficiency of the borrowing algorithm.

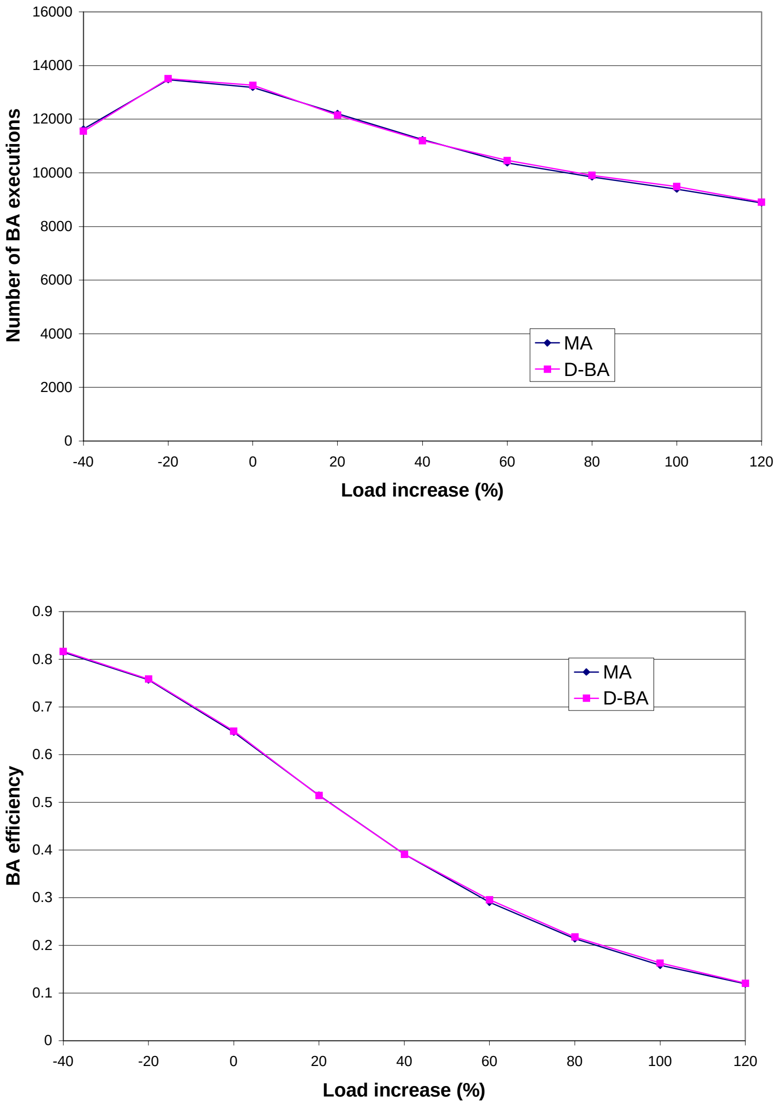


Figure 6.12: MA versus D-BA: Borrowing algorithm (BA) results

The handoff rejection rate is also lower in the MA network, thanks to the action of the agent negotiation (Figure 6.13). This is an important result because it increases the QoS perceived by the mobile user. Finally, Figure 6.14 shows the call dropping rate of the networks. The call dropping rate is small and irregular in all three networks, making it difficult to analyse the significance of the result.

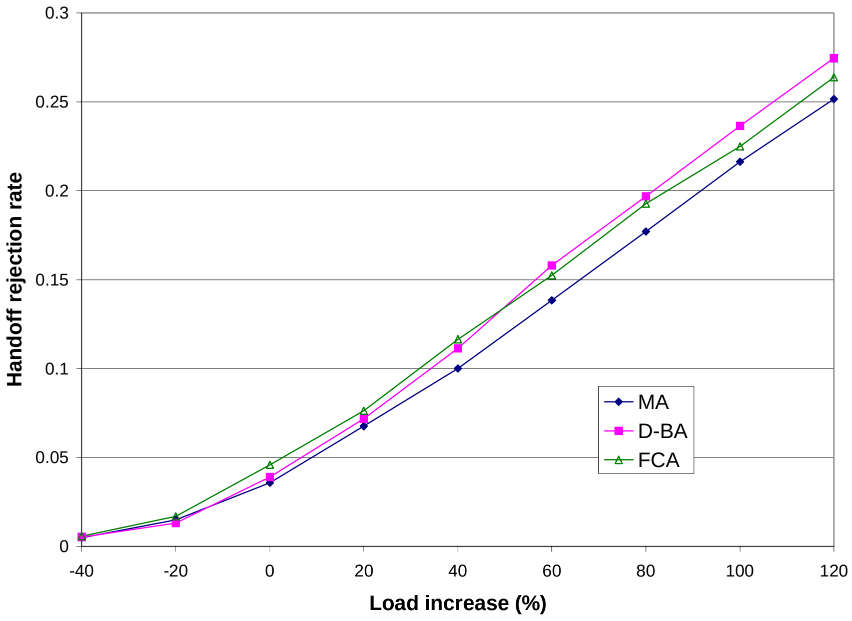


Figure 6.13: Handoff rejection rate in traffic load scenario 2

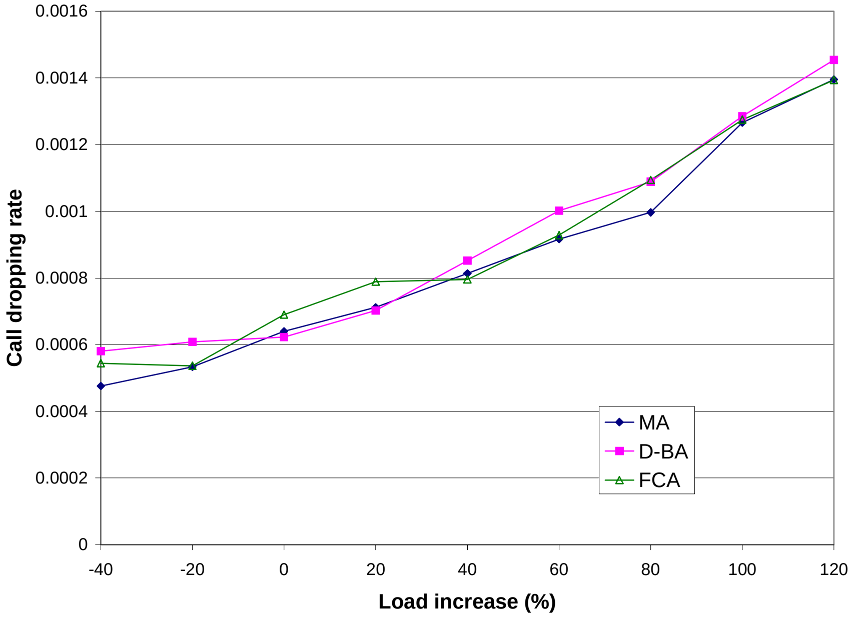


Figure 6.14: Call dropping rate in traffic load scenario 2

Some of the cells with higher traffic load in the network had better improvement in the traffic-blocking rate. The traffic blocking rate of cells 0, 26 and 36 are shown in Figure 6.15.

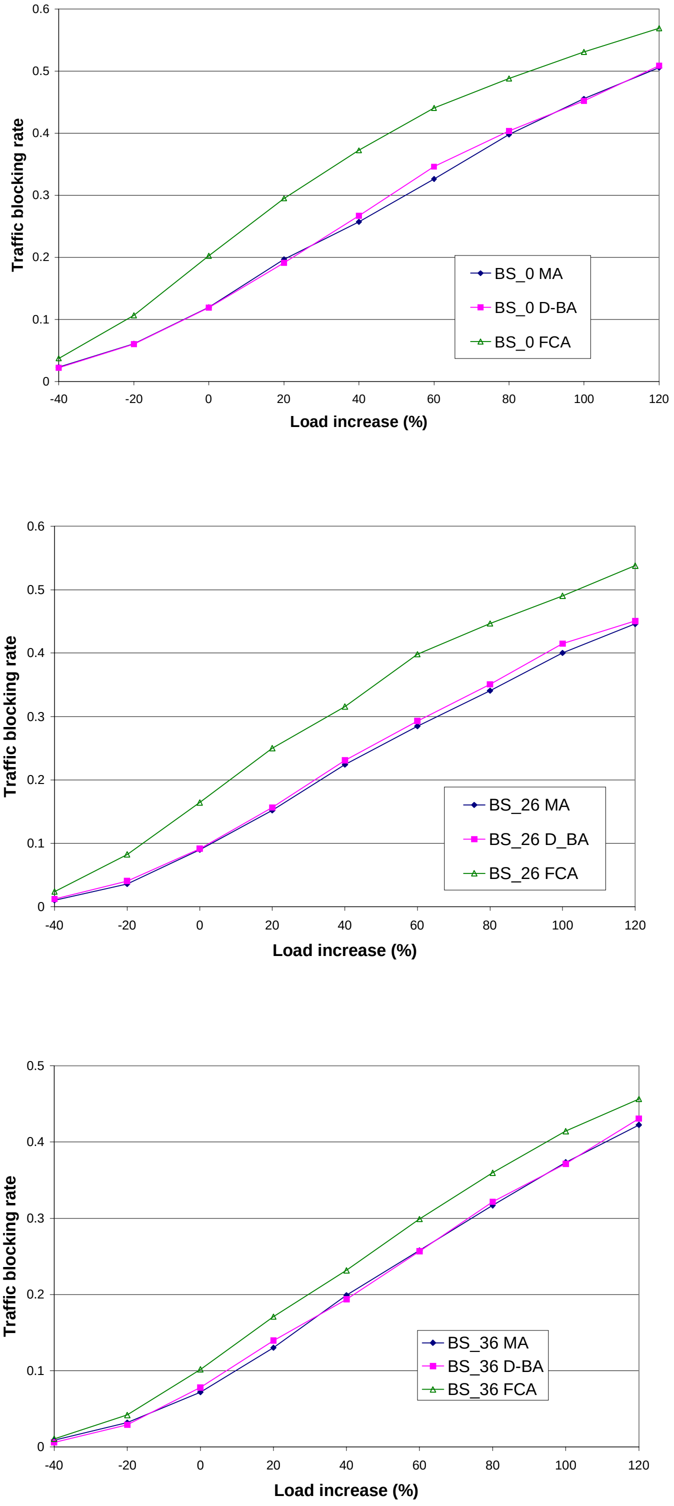


Figure 6.15: Cells 0, 26 and 36: Traffic blocking rate for traffic load scenario 2

The greater improvement in traffic blocking rate of individual highly loaded cells is a good result, because these cells are those that have a greater need of resources. The level of improvement in the more loaded cells is mainly a consequence of the choice of the borrowed channel release threshold. A higher threshold means the loaded cells will keep their borrowed channels for longer. In the next scenario the value of this parameter is changed to show its effect.

6.5 Simulation Results for the Traffic Load Scenario 3

The layout of the network of the third scenario with non-uniform traffic load distribution is shown in Figure 6.16. This cellular layout is the same as used in the simulation tests in [YY94]. Again, mobile stations inside the shaded area can be driving at 40 km/h or walking at 2 km/h and they move in both directions. All the other mobile stations have random movement. The simulation parameters have the values shown in section 6.2, except the release borrowed channel threshold that was decreased to 0.1, as explained later.

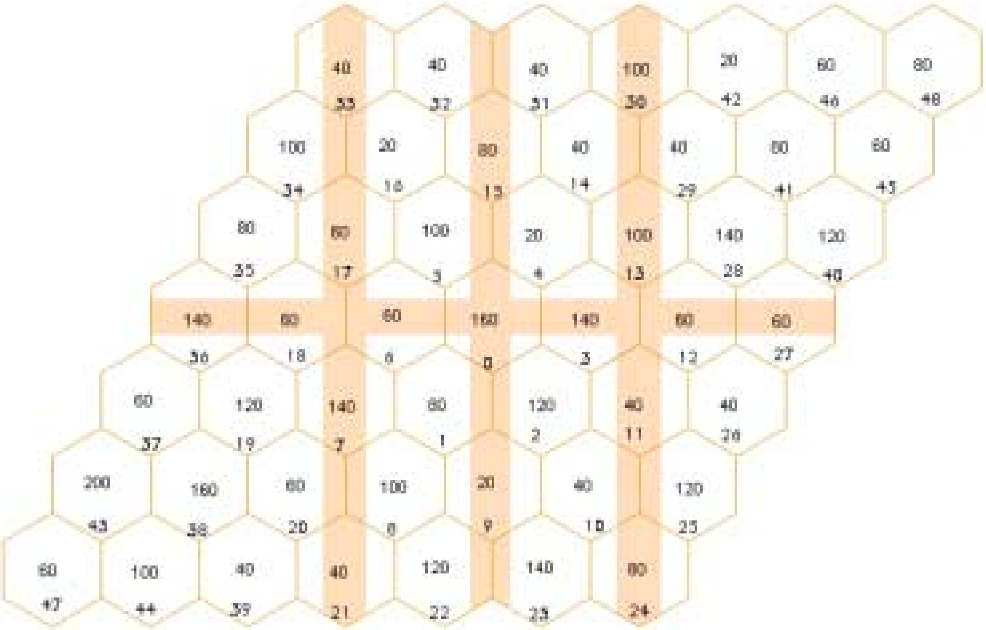


Figure 6.16: Cellular network layout for non-uniform load distribution of scenario 3

Figure 6.17 shows the overall traffic blocking rate of the three networks for different percentages of traffic load increase in scenario 3.

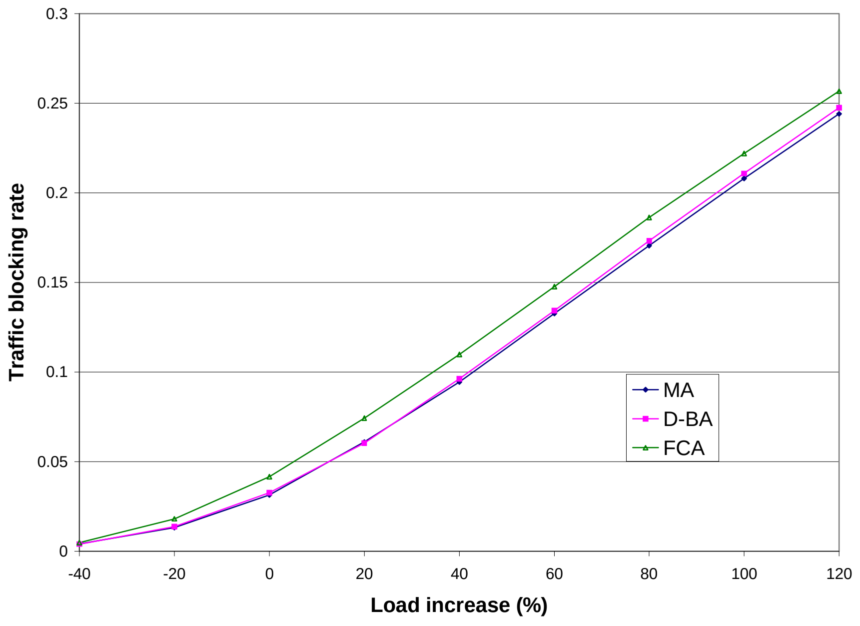


Figure 6.17: Traffic blocking rate for traffic load scenario 3

The MA network outperforms the D-BA and the FCA networks. The improvement is lower than scenario 2, but the MA network behaves as expected. The number of borrowing attempts is also kept at the same level of the D-BA network, with a lower blocking rate. The handoff rejection rate is also lower in the MA network (Figure 6.18).

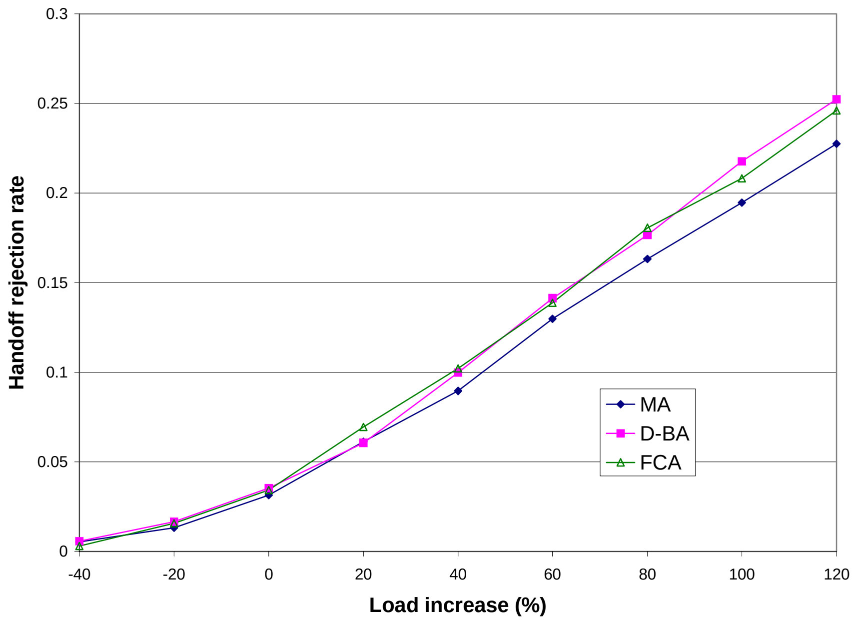


Figure 6.18: Handoff rejection rate in traffic load scenario 3

Figure 6.19 shows the call dropping rate for all three networks.

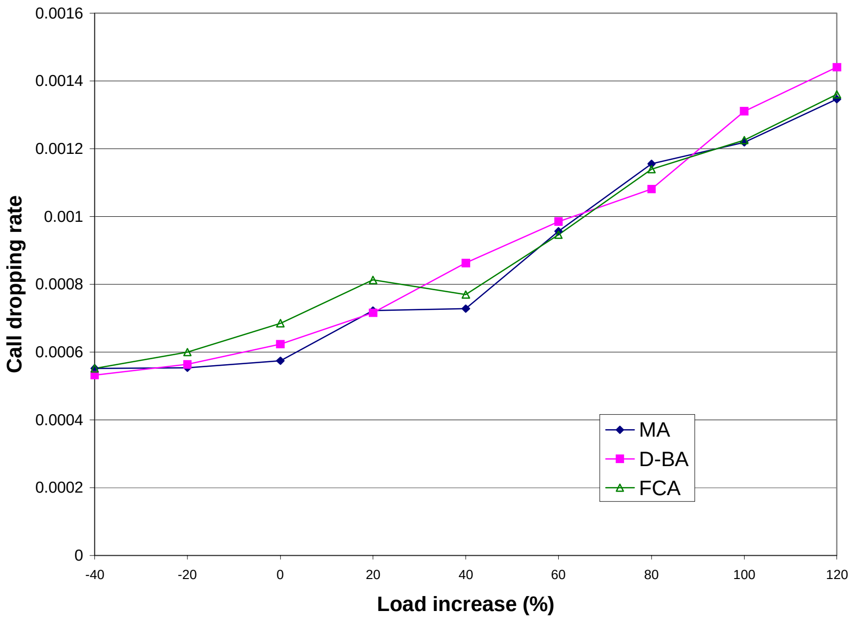
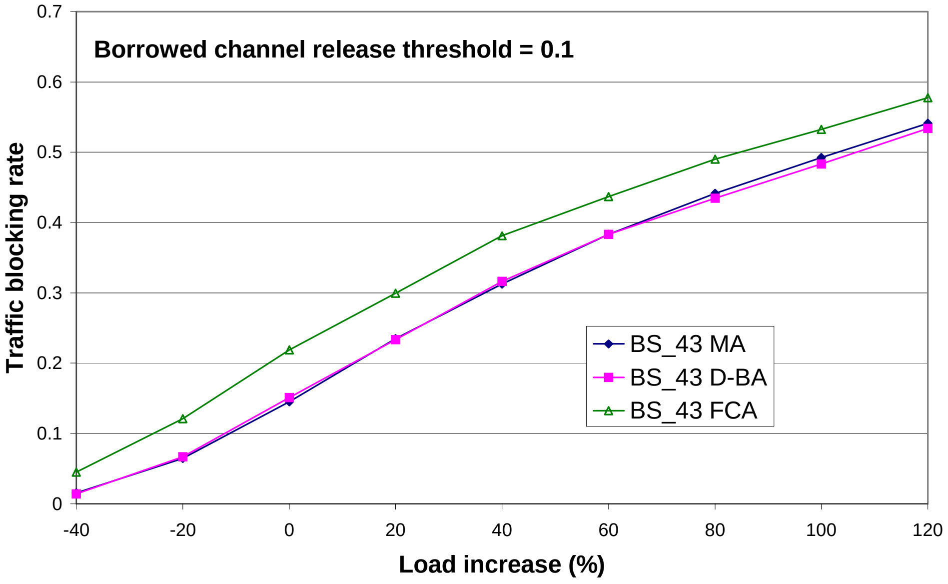


Figure 6.19: Call dropping rate in traffic load scenario 3

Now, the cell with highest traffic load is cell 43; cell 0 also has a high load. The traffic blocking rate of cells 43 and 0 are shown in Figures 6.20 and 6.21.

In these highly loaded cells, the improvement is less when the borrowed channel release threshold is set lower. This is expected because the cells keep the borrowed channels for less time with a lower threshold. However, there is a slight gain in the overall traffic blocking rate because the lender cells will have more channels for their own offered traffic. The decision of what value to use for the borrowed channel release threshold depends on what is more important for the network: a slightly better overall blocking rate or a greater improvement in the more highly loaded cells. Both traffic blocking rates are shown in Figures 6.20 and 6.21.



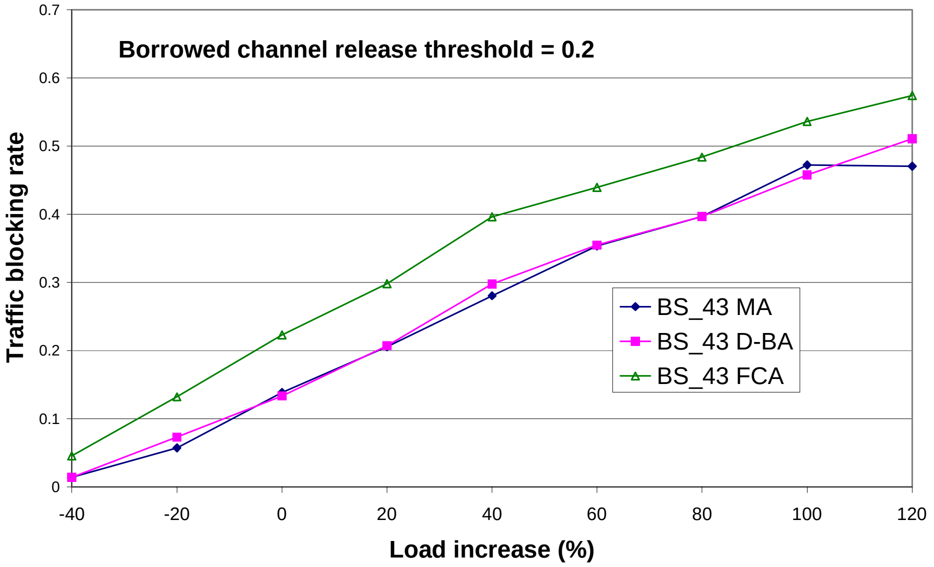


Figure 6.20: Traffic blocking rate in cell 43 for traffic load scenario 3 with different borrowed channel release thresholds

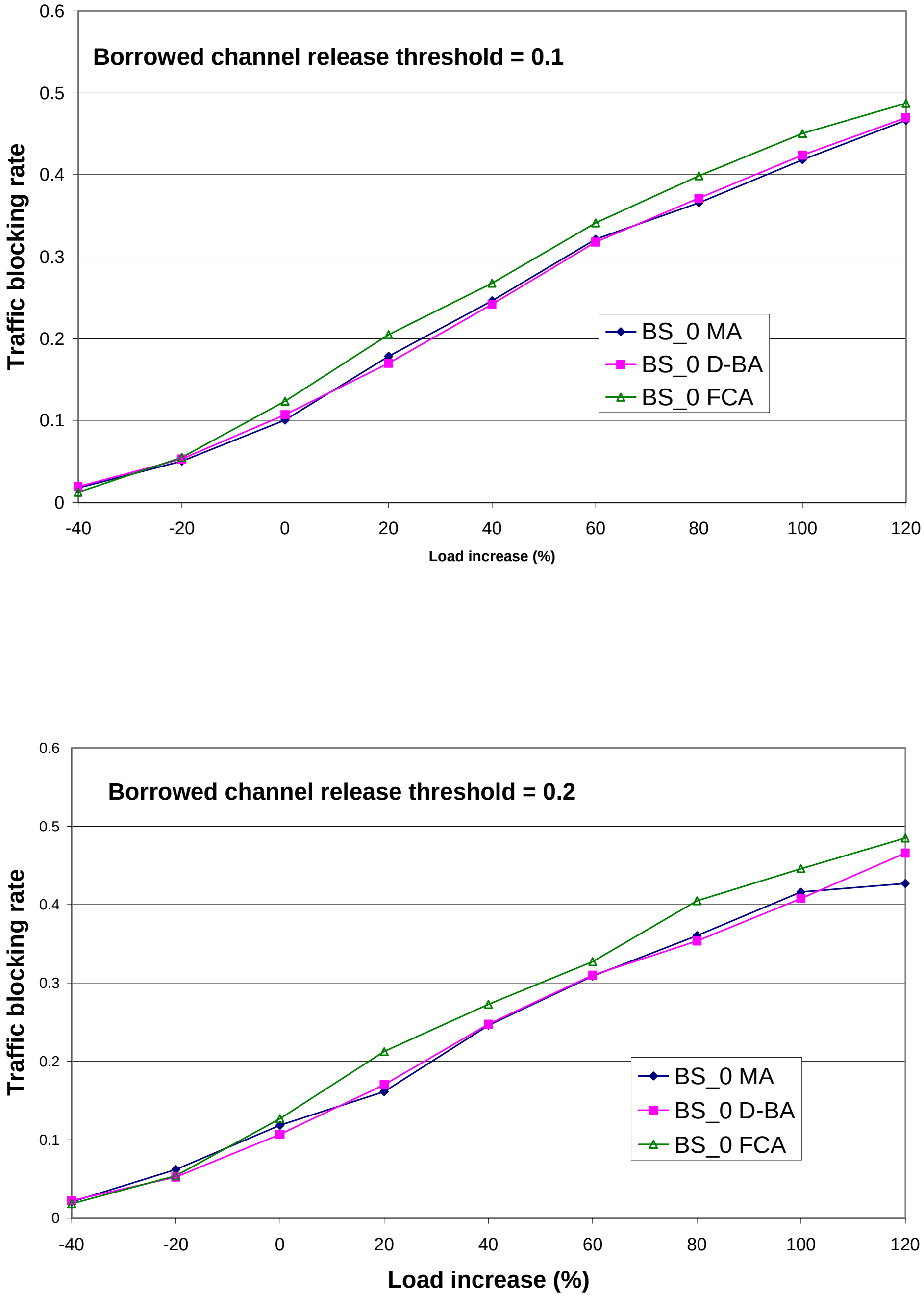
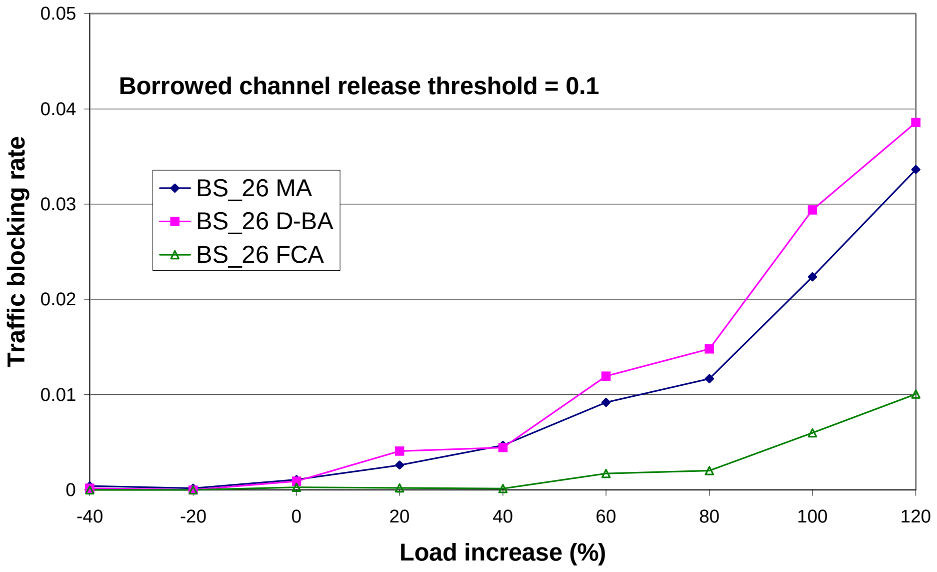


Figure 6.21:Traffic blocking rate in cell 0 for traffic load scenario 3 with different borrowed channel release thresholds

Figure 6.22 shows how a cell with lower traffic load is affected by the three different networks and different borrowed channel release thresholds. Of course, the traffic blocking rate will be lower when the borrowed channel release threshold is set equal to 0.1 because cell 22 will have more available channels to serve its offered traffic. The traffic-blocking rate given by the FCA network is very low as expected;

the blocking is worse with the other schemes because this cell is lending channels to other cells.



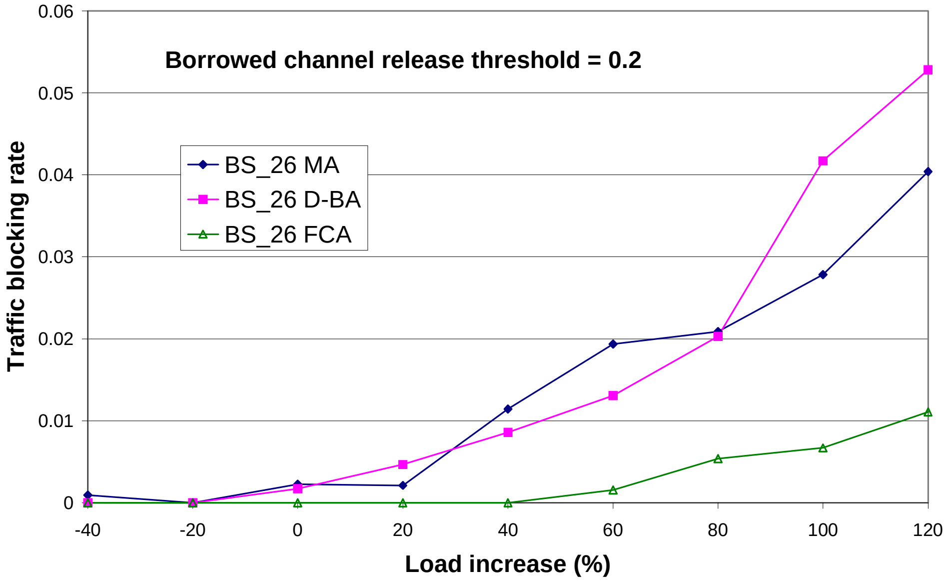


Figure 6.22: Traffic blocking rate in cell 26 for traffic load scenario 3 with different borrowed channel release thresholds

6.6 Simulation Results for the Traffic Load Scenario 4

The last scenario tries to study the effects of call congestion in highways. The non-uniform load distribution layout is shown in Figure 6.23. Now the cells with more traffic load are the ones that contain highways (shaded areas). In this particular scenario a cross point between the highways has the highest traffic load. Mobile stations inside the shaded area can be driving at 40 km/h or walking at 2 km/h and they move in both directions. All the other mobile stations have random movement. The simulation parameters have the values shown in section 6.2.

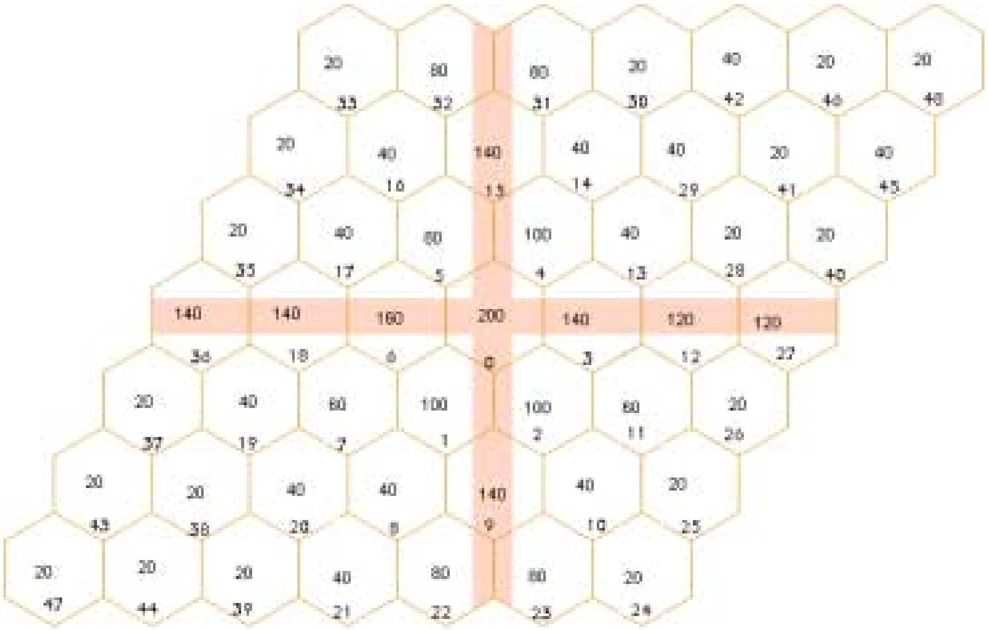


Figure 6.23: Cellular network layout for non-uniform load distribution of scenario 4

Figure 6.24 shows the overall traffic blocking rate of the three networks for different percentages of traffic load increase in scenario 4.

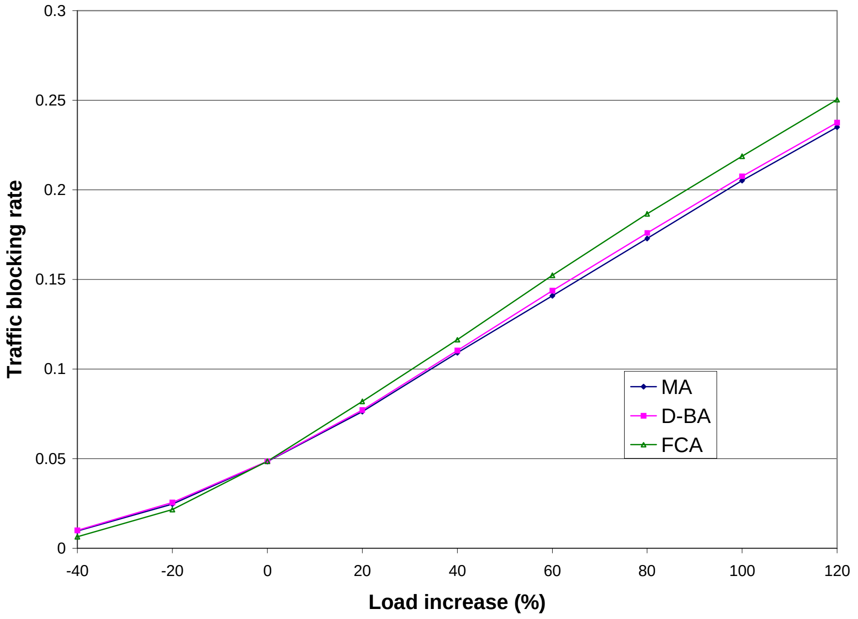


Figure 6.24: Traffic blocking rate for traffic load scenario 4

This scenario illustrates important features in the schemes. FCA outperforms both the other schemes up to the base load, probably because the channel locking is affecting the most heavily loaded central part of the network. The MA scheme is a little bit better than D-BA because of its load-balancing feature. From the base load upward the MA shows a modest improvement over the D-BA because the Co-operative Planning Layer is not being triggered sufficiently to make a noticeable improvement. In the current implementation of the agent, detection of failure types 3 and 4 by the Local Planning Layer is responsible for triggering the agent negotiation. In this scenario, the delays in generating the "failure triggers" have a bigger impact. Failure 3 is numerous because of the traffic condition in the region of the cross point, but failure 4 does not occur until the highest load shown in Figure 6.24. This is expected because several cells have low traffic load. In order to prove the impact of the delays introduced, a simulation was performed with this scenario using the parameters values of section 6.2, but now the values for all borrowing algorithm re-execution delays were lowered by 50 %.

- Delay in executing next borrowing algorithm if outcome was partial success = 150 seconds
- Delay in executing next borrowing algorithm if outcome was failure 2 = 300 seconds

- Delay in executing next borrowing algorithm if outcome was failure 3 = 600 seconds
- Delay in executing next borrowing algorithm if outcome was failure 4 = 750 seconds

Figure 6.25 shows the new traffic blocking rates for the D-BA and MA networks; of course the FCA is not affected by these parameters.

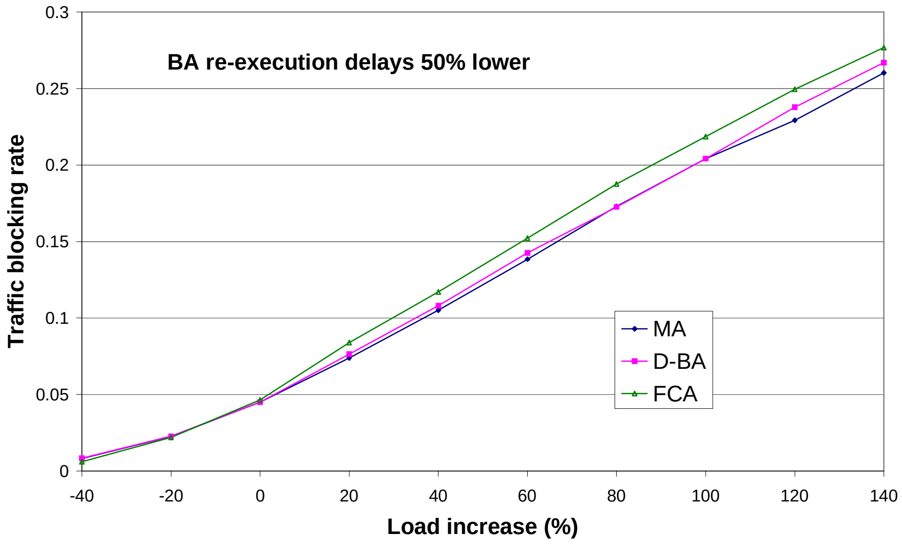


Figure 6.25: Traffic blocking rate for traffic load scenario 4 with different values for the borrowing algorithm re-execution delays

As expected the MA network performs better now, especially as the load gets higher. However, the cost-benefit of the increase in signalling load needs to be further investigated and is not covered in this thesis.

The delay values in section 6.2 have an advantage that the number of runs of the borrowing algorithm does not escalate for higher loads. More specifically for scenario 4, the number of runs of the borrowing algorithm does not change significantly from 40% load increase upwards, as shown in Figure 6.26.

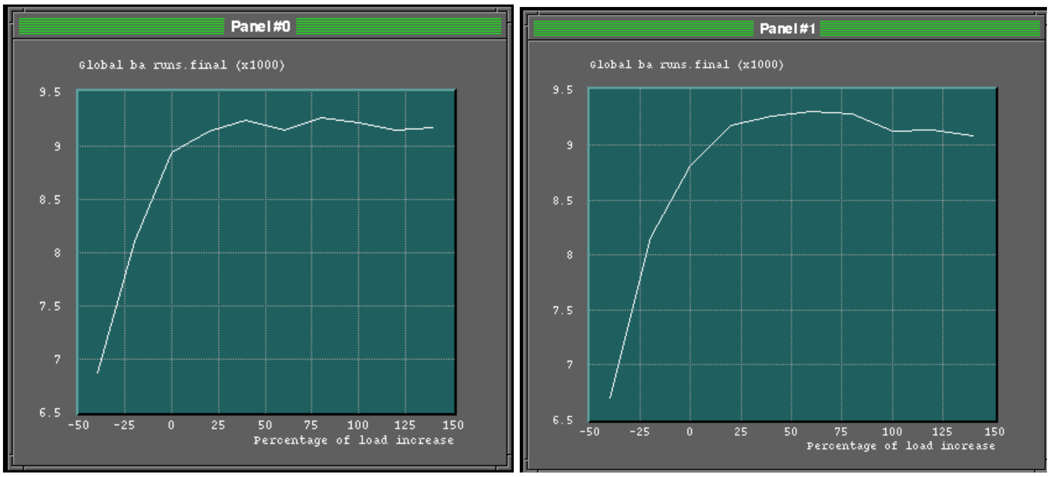


Figure 6.26: Number of borrowing algorithm runs for MA (left) and D-BA (right) networks in scenario 4 versus Percentage of load increase (OPNET graphs)

The following simulation results also used the parameter values shown in section 6.2. The handoff rejection rate given by the MA network is still the lowest from the base load upwards. But for the first time in the scenarios investigated, the handoff rejection rate of the MA network was noticeably worse that the FCA up to the base load, as shown in Figure 6.27. The probable cause is because there were fewer failures of type 3 up to the base load and consequently there were not sufficient executions of the joint plan to improve the rates in the MA network. It is worth remembering that several cells have a very low traffic load up to the base load in scenario 4.

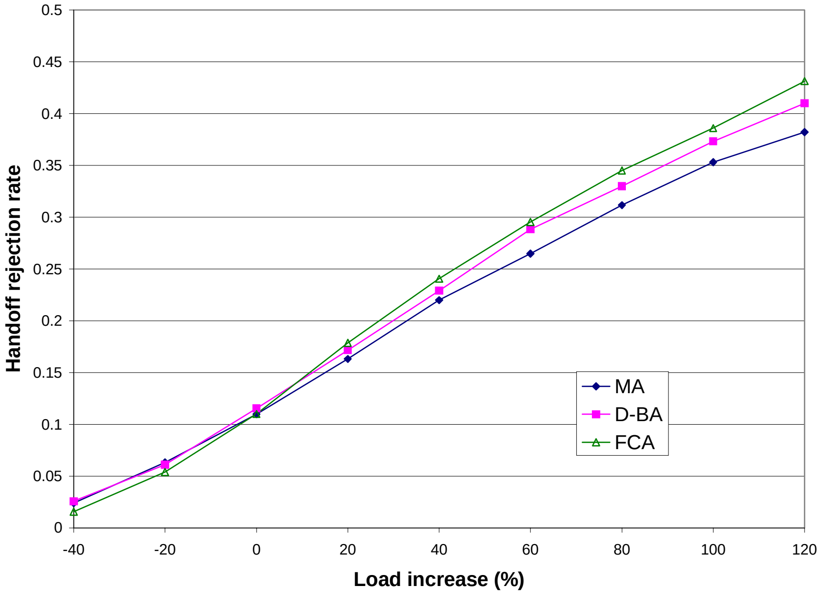


Figure 6.27: Handoff rejection rate in traffic load scenario 4

The call-dropping rate is also irregular for scenario 4 (Figure 6.28).

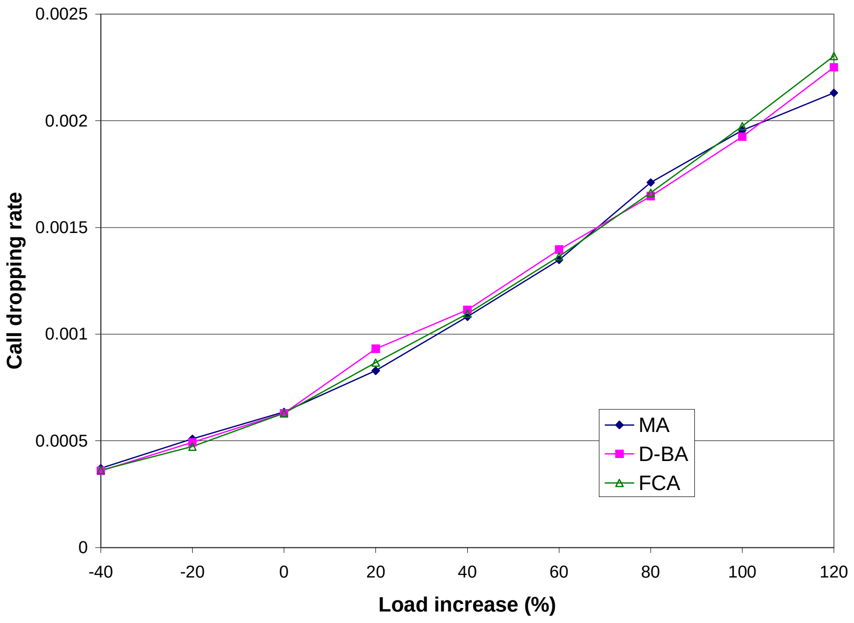


Figure 6.28: Call dropping rate in traffic load scenario 4

The cell with highest traffic load is cell 0 and now this is surrounded by other heavily loaded cells. The traffic blocking rate in this cell is shown in Figure 6.29.

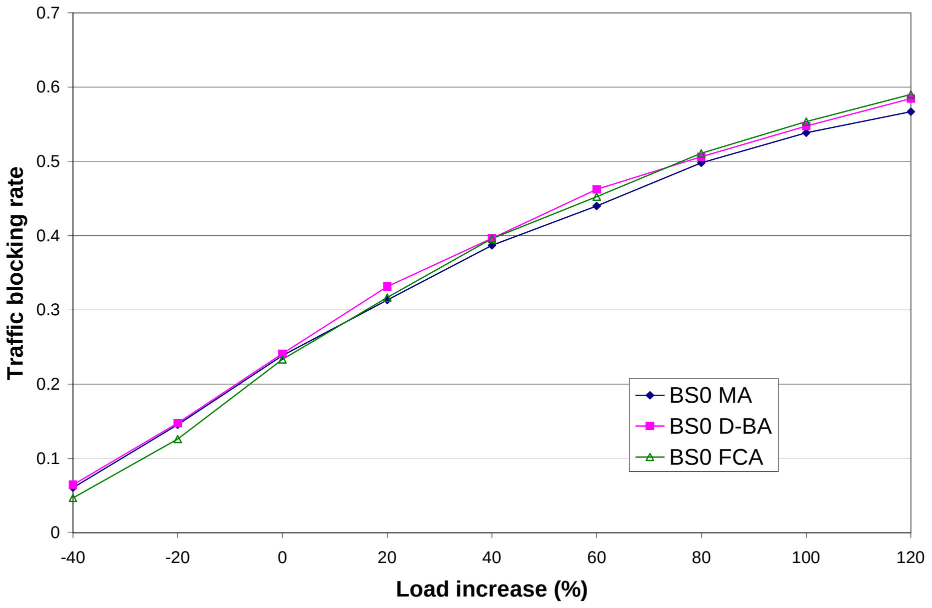


Figure 6.29:Traffic blocking rate in cell 0 for traffic load scenario 4

Here, the D-BA scheme has the highest traffic blocking rate up to 80% load increase. The runs of the borrowing algorithm in cell 0 mostly result in failure type 3. The probability of borrowing channels decreases because of the delays introduced for re-execution of the algorithm. The blocking rate is also degraded because, in the D-BA scheme, there is no channel availability threshold to avoid the locking of channels in a cell. Therefore, if a channel is available in cell 0 it might be locked because a co-channel cell is lending a channel to another cell.

The MA network outperforms the FCA network for traffic loads above the base load. This is a good result, despite the improvement being only 5% on average. However, the poor performance of the Reactive Layer of the agent needs to be considered. As the current Local Planning Layer is simple, the results of the agent system are being "dragged down" by the algorithm in the reactive layer clearly not being suitable for a cell surrounded by congested neighbours. It is the action of the Co-operative Planning Layer of the agent that is responsible for the lower traffic blocking rate.

Finally, the traffic blocking rate of cell 36 is shown in Figure 6.30.

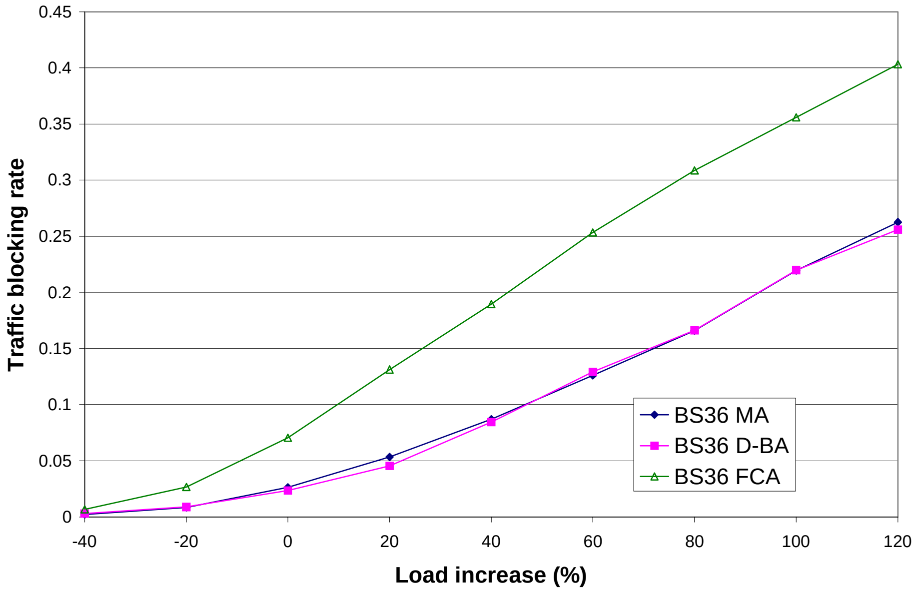


Figure 6.30:Traffic blocking rate in cell 36 for traffic load scenario 4

The D-BA scheme works very well for isolated hot spots, such as cell 36. Consequently the proposed MA network also improves the traffic blocking rate. Although it is sometimes worse than the D-BA because of the action of the Co-operative Planning Layer bringing more users to cell 36 (load balancing) and therefore generating a higher probability of blocking incoming traffic than the D-BA scheme alone.

6.7 Analysis of Simulation Results

The most important feature shown in the simulation results is that the behaviour of the multi-agent system is as expected by the author. There are still resources available when the neighbouring cells (of a borrower cell) also reach the threshold (***h)*** of channel availability. However, at that point the borrowing algorithm in the Reactive Layer is not able to obtain those resources anymore and at this time the agent negotiation has an important role in shifting some of the traffic to less loaded regions. The agent negotiation proved to work well as shown by the simulation results. Almost 100% of the management handoffs were successful in all scenarios when the mobile station had enough signal strength to shift cells. This clearly shows good performance from the heuristic proposed by the author inside the agent negotiation, choosing the right cells to receive mobile users. As an example, Table 6.5 shows the management handoffs for scenario 2, where the success rate is high. In this case the success rate is defined as the ratio of successful handoffs to possible handoffs; possible handoffs exclude those where the signal strength is too low.

The number of rejections for lack of signal strength is high, and perhaps could be reduced if the threshold for a mobile user to be considered as departing were optimised. This is one of several parameters that could be optimised. However, this could never be a firmly-fixed value as different operators would have different views on a suitable value, given its effect on perceived quality of service (QoS): a lower threshold would give less calls dropped (better QoS), but more likelihood of calls failing (worse QoS).

Another factor that was fairly artificial was the simulated user mobility. However, the work performed is sufficient to demonstrate that user mobility directly affects the performance of the agent negotiation. Knowledge of traffic distribution, mobility and geography (taking into account buildings and roads) inside the cell would allow a better and a more precise procedure for building and updating the departing users lists (***NumDepart*** array). This would substantially improve the agent negotiation performance and consequently decrease the traffic blocking rate. It may even be practicable to have different procedures for different types of cell geography.

Table 6.5: Management handoff request performance

Load increase (%)	Total number of management handoff requests	Total number of rejected management handoff requests because lack of signal strength	Total number of successful management handoffs	Success rate (%)
-20	551	331	220	100.0
-40	1,781	1,115	666	100.0
0	3,267	1,889	1,378	100.0
20	4,780	2,846	1,934	100.0
40	5,787	3,409	2,376	99.9
60	6,494	3,735	2,751	99.7
80	6,538	3,778	2,752	99.7
100	6,914	3,843	3,052	99.4
120	6,122	3,301	2,790	98.9

Another interesting result is that it is the agent negotiation that is decreasing the blocking rate when the entire network reaches the channel availability threshold (***h***). At this point, the use of the D-BA scheme is not worthwhile anymore and the load balancing performed by the agents causes the decrease in traffic blocking.

The amount of improvement given by the multi-agent system is fairly modest. Currently the agent implementation has a fairly simple Local Planning Layer. The task of this layer is to optimise the channel allocation algorithm in the Reactive Layer through learning and prediction or traffic monitoring techniques that would make the reactive algorithm more adaptive to the traffic load and consequently perform better. The parameters that affect the performance of the borrowing algorithm (D-BA) most in the Reactive Layer implemented in this work are those shown in section 6.2 (additional parameters for the D-BA model and hence also for the MA model). These parameters should be optimised and adapted to different traffic situations by the Local Planning Layer. Therefore, although the implemented Local Planning Layer is sufficient for the objectives of the work in this thesis, a more elaborate Local Planning Layer could be considered in future work as discussed in Chapter 7.

In section 6.5, the author has shown some of the sensitivity analysis performed, where the borrowed channel release threshold was applied to scenario 3 with different values. One of the effects of this parameter is to shift the congestion of the network. If the value is lower, the congestion of the whole network is shifted to a higher traffic load. This can be easily seen by the output data of failure 4 of the borrowing algorithm, as shown in Table 6.6.

Table 6.6: Output data for BA failure 4 of scenario 3

Load increase (%)	BA failure 4 for scenario 3 (borrowed channel release threshold = 0.1)	BA failure 4 for scenario 3 (borrowed channel release threshold = 0.2)
-20	0	0
-40	0	0
0	0	0
20	0	0
40	0	0
60	0	0
80	0	0
100	0	0
120	0	4
130	4	4
140	32	57

The outcomes from executing the borrowing algorithm are an important source of information about the traffic load in the network, and the Local Planning Layer can use this information to improve the efficiency of the Reactive Layer and also the Co-operative Planning Layer. In the current implementation of the agent, detection of failure types 3 and 4 by the Local Planning Layer are responsible for triggering the agent negotiation. Therefore, the delays introduced in the criteria for re-execution of the borrowing algorithm (section 4.3.1.2) have a direct effect on the number of agent negotiations being triggered. The introduction of these delays is very important to decrease the waste of signalling resources. Continuing to run the borrowing algorithm when the probability of failure is very high only yields signalling with no benefit on traffic handling. However, the actual amount of time set for each delay needs also to be optimised by the Local Planning Layer and better-tuned values for these delays could lead to a better performance of the agent negotiation.

Finally, the parameters belonging to the Co-operative Planning Layer also need to be optimised. The simulation results are satisfactory, but an optimisation of those parameters might improve the traffic blocking and handoff rejection rates.

In summary, the simulation results prove that the agent negotiation is efficient but a more elaborate Local Planning Layer could be implemented to improve the overall results given by the multi-agent system.

Chapter 7 Discussions and Conclusion

7.1 Discussion

This thesis has addressed the challenge to bring more flexibility to cellular networks with the distributed control needed to make the implementation feasible for real systems. The approach adopted has been to use intelligent software agents to provide more autonomy and flexibility to base stations. The major contribution of this work has been the exploitation of the ability of intelligent agents to perform autonomous and intelligent negotiation in order to improve the means of resource acquisition. This approach has not been used for any other channel allocation scheme and so is a new contribution.

The intelligent agents are completely responsible for the frequency channel assignment in the cellular network, which for this work was a macro-cellular network using FDMA access technology. Base stations are not able to share information by interference measurements, but only by explicit exchange of information. The resources are full frequency carriers and each base station has one intelligent agent.

The channel allocation schemes proposed in the literature were designed to be autonomous and they have achieved their design goals. However, most of these algorithms in the literature are reactive with a single strategy for adaptation to traffic change conditions. They lack flexibility because they are not able to monitor their own performance and change strategies to continue to pursue their objectives efficiently or to negotiate resources within a group of cells. Flexibility can be achieved through pro-active logical planning and through co-ordinated negotiation. The features of complexity separation, fast response, autonomy, pro-activity and negotiation for real-time applications have been objectives of multi-agent systems research.

Several multi-agent frameworks have been proposed for control management in telecommunication networks (section 4.2). A layered multi-agent control system approach seems to fit the requirements of distribution, intelligence, robustness and concurrency of broadband network control. The hierarchical arrangement allows levels of co-ordination. Each layer is defined to conduct the control of the network to a certain level of competence and in this work such layering was considered desirable. As one agent is placed at each base station, there is only one level of agents in the cellular network, and therefore the internal architecture of the agent should supply the layered control. The agent architecture also needed to offer planning and negotiation mechanisms in order to provide autonomy and flexibility.

As explained in the body of the thesis, the INTERRAP architecture was adapted to the cellular network environment. The **Reactive Layer** is responsible for fast accommodation of traffic demand; the **Local Planning Layer** uses other strategies to optimise the efficiency of the tasks in the Reactive Layer and the local distribution of channels; the **Co-operative Planning Layer** is responsible for load balancing the traffic across a larger area.

The Reactive Layer is basically composed of a FCA algorithm with channel borrowing and channel locking. The algorithm called “distributed borrowing algorithm” (D-BA) is based on the general behaviour of the D-LBSB scheme [DSJA97]. The decision to use an algorithm similar to a distributed scheme proposed in the literature was made for two reasons: first it allowed a close evaluation of the performance of a distributed reactive algorithm in the cellular model built by the author using a commercial network simulator; second, it allowed measurement of the benefits of the other two layers of the agent against a purely reactive control system.

The Local Planning Layer is responsible for the channel re-assignment scheme. This layer monitors the efficiency of the algorithm in the reactive layer, and it is responsible for the decision of when to trigger the Co-operative Planning Layer.

The Co-operative Planning Layer is responsible for the negotiation of resources. The agent negotiation has two phases: first to find the best region to attempt to shift the frequency channel; second the actual execution of the handoff requests in a co-ordinated manner, i. e. the execution of a joint plan. In the first phase, a utility function (defined by the author) is used to find the best region. In the second phase, all cells belonging to the best region engage in the joint plan. The co-ordination of the handoff attempts is performed through a market-based control heuristic developed by the author using the Contract-Net Protocol (CNP) [Smith88] as the negotiation protocol.

To investigate this approach, a cellular network model was built using the commercial simulator OPNET. The simulation model was verified and validated using the *replication/deletion approach* [LK91] and compared against the call blocking probability given by the *Erlang B formula*. The simulator was extended to incorporate the D-BA scheme and the full agent model.

The simulation results obtained from the multi-agent system were compared with those of the conventional cellular network using FCA and a cellular network using the D-BA scheme.

The network performance characteristics used for comparisons were the traffic blocking rate, the handoff rejection rate and the call-dropping rate. Different traffic load scenarios were set to demonstrate the behaviour of the three different cellular networks facing different situations in traffic load.

The network using intelligent software agents gave the best result for overall traffic blocking rate in all scenarios. Although the improvement of the multi-agent system has been fairly modest compared to the D-BA scheme, the multi-agent system has behaved as expected by the author, shifting some of the traffic to less loaded regions when the Reactive Layer scheme could not obtain the sparse resources still available. The simulation results show that the improvement in traffic blocking at high traffic loads is a direct consequence of the agent-negotiation. The performance of the market-control heuristic proposed by the author is very good with a success rate of almost 100% in management handoffs (Table 6.5) in the scenarios chosen.

In non-uniform traffic load scenarios, individual cells with heavier traffic load show a greater improvement in traffic blocking compared with FCA. In the scenarios used, the results from the multi-agent system have been similar to the D-BA scheme because the borrowing algorithm used in the Reactive Layer of the agent and used in the D-BA scheme has been able to reduce the blocking rate in these individual cells. The author has shown (Chapter 6) some of the sensitivity analysis of the borrowing algorithm parameters, demonstrating how they affect the traffic-blocking rate in individual cells.

The multi-agent system has presented the lowest handoff rejection rate in all scenarios. This is a very interesting result that confirms the efficient action of the agent negotiation. For all scenarios, the D-BA scheme has shown the worse handoff rejection rate so that it can be concluded that the reduction in handoff rejection rate is down to the action of the Co-operative Planning Layer of the agent. This result is important because the multi-agent system is able to improve the user's perceived quality of service.

The call-dropping rate has been very irregular in all three cellular networks for the different scenarios, making it difficult to analyse the significance of the results. However, the rate is so low that the results are not statistically significant.

Overall, the multi-agent system has proved to be feasible and a efficient. It has brought more flexibility to the base stations in obtaining extra resources using a different strategy than that being applied by the reactive channel allocation algorithm. The agent negotiation is well behaved and controlled.

7.2 Evaluation of the Multi-Agent System

The work performed in this thesis gives a good insight into the applicability of multi-agent systems in mobile telecommunications. The architecture chosen is general and sufficiently flexible to be applied to other problems that involve resource management and load balancing. In telecommunications, the use of agents is not restricted to the network layer, but agents can also find a role at higher layers, such as the services layer. The strength of multi-agent systems lies in their social behaviour and in their capabilities to be proactive when pursuing their own goals.

The major strength of the internal agent architecture used in this work is the separation of the complexity, the domain of action/intervention and the different response timescales. Moreover, the work shows that an autonomous negotiation of resources following a well design optimisation heuristic can improve QoS or/and reduce costs.

Agents will be a good solution for problems where automated negotiation can be applied, not only because they are autonomous, but mainly because of their other capabilities: proactiveness, reactivity and learning.

However, the computational cost embedded in internally layered agents needs to be carefully evaluated when applying them in a determined solution. The gathering of information necessary for the execution of the negotiation process needs to avoid excessive numbers of messages being exchanged. Ways to minimise the volume of messages and different approaches for decision making based on incomplete information is part of the research activity into multi agent systems.

Any delays in response because of negotiation between agents would also be a severe drawback. However, careful layering (as demonstrated in this work) between fast reactive layers and longer-term planning processes should avoid this problem. In certain cases, off-line learning or off-line optimisation could improve the response time of the planning layers.

In this work, the co-operative layer of the agent is light in processing cost, giving a good response time. However, in the reactive layer the trade-off between the processing cost and the reduction in blocking rate is not so good. A simpler reactive algorithm coupled with a richer local planning layer might well improve performance, not only in terms of blocking probability, but also in responsiveness.

For real solutions in telecommunications, an approach that keeps simplicity in mind is generally key to success. This is the reason that a careful design of each layer of the agent is extreme important and the implementation of simulation models (as performed in this thesis) can be useful for measuring the real cost-benefit of applying agents in a specific solution.

7.3 Future Work

Currently the Local Planning Layer of the agent implemented is fairly simple. This layer's task is to optimise the reactive channel allocation algorithm in the Reactive Layer and to make it more adaptive to the incoming traffic load. Learning and prediction or traffic monitoring mechanisms should be used by the Local Planning Layer to achieve these goals. The author has shown the parameters that affect the performance of the borrowing algorithm most. The Local Planning Layer should optimise these parameters and should adapt them to different traffic situations. A more elaborate planning layer could also use learning and prediction techniques to optimise the situations when the Co-operative Planning Layer is triggered, saving signalling resources. *A more elaborate Local Planning Layer should be considered in future work.*

Although the Co-operative Planning Layer of the agent implemented is performing well, still there is space for improvement. This layer also has planning features as described in the INTERRAP architecture. Again, optimisation techniques could be applied to this layer in order to determine how many iterations a joint plan should perform given certain traffic conditions. The parameters that affect the performance of the two phases of the agent negotiation (section 6.2) could also be optimised. This should lead to better performance of the agent negotiation. *A Co-operative Planning layer with richer planning features should also be considered in future work.*

Further investigation of the behaviour of the multi-agent system under more realistic traffic load scenarios would be desirable. This investigation could bring important information on how to improve the current multi-agent system.

The implementation of a more realistic user mobility simulation model (taking into account buildings and roads inside the cells) would be a very important feature inside the cellular network. User mobility directly affects the agent performance and more precise procedures for building and updating the user departing lists would substantially improve the agent negotiation performance and consequently decrease the traffic-blocking rate. It could even be worth building different procedures for different types of cells.

Finally, the multi-agent system developed in this work could be adapted to networks that use different access technologies or different cellular structures (e.g. GSM, IS-95 CDMA or 3G) as discussed in the next section.

7.4 Applicability to Other Mobile Networks

The Agent framework proposed in this work is open. Different types of channel allocation algorithms could be implemented in the Reactive Layer. The Local Planning Layer would still have the task of optimising whatever algorithm is placed in the Reactive Layer and decide when to trigger the Co-operative Planning Layer, which in turn could have a broader range of strategies for negotiation.

In a hierarchical structure, the agent of a macro-cell could negotiate resources with agents of micro-cells beneath it or even agents in pico-cells. The selection criteria for the partners in a joint plan could be in terms of class of service required and the spare bandwidth available for that service class. Therefore, although the implementation in this work was performed in FDMA access networks, the framework of the agent is powerful and general enough to be applied to third generation mobile networks. As a matter of fact, the work presented in this thesis was used as the basis for a successful proposal for an IST EU 5th Framework Project (SHUFFLE) for radio resource management using intelligent agents in third generation mobile network [Shuffle00]. Although the access technology characteristics are much more complex than FDMA and very different interference and traffic capacity considerations need to be made, radio resource management using negotiating agents is considered possible by the author. This assertion is being proved in the Shuffle project and that work will have an important impact on the implementation of intelligent agents for resource management in real mobile networks.

7.5 Conclusion

Resource flexibility is one of the most important requirements in the next generation of mobile communications. Techniques for increasing the flexibility of the network to deal with new services and traffic characteristics are a requirement and an implementation challenge.

The aim of the work in this thesis was to provide more autonomy and flexibility to base stations in order to improve their means of radio resource acquisition. The approach proposed by the author to achieve this aim, was to use intelligent software agents to control the channel assignment in the cellular network. A special hybrid agent architecture was adopted consisting of three layers. The first layer (**Reactive Layer)** is responsible for fast accommodation of traffic demand. The second layer (**Local Planning Layer)** uses other strategies to optimise the efficiency of the tasks in the Reactive Layer and the local distribution of channels. Finally the third layer (**Co-operative Planning Layer)** is responsible for load balancing the traffic across a larger area. The specification and implementation of the **Co-operative Planning Layer** is the major contribution of this thesis: the exploitation of the ability of intelligent agents to perform autonomous and intelligent negotiation in order to acquire radio resources. However, the interplay of the three layers proved to be a powerful framework to improve radio resource flexibility and to increase the robustness of the cellular network as a whole.

The performance of a cellular network using the multi-agent system was compared with the conventional cellular network using FCA and another using a reactive distributed borrowing algorithm called D-BA. The simulation results demonstrated that the use of intelligent software agents brought more flexibility in obtaining extra radio resources to the network than the other two approaches.

Overall, the multi-agent system proved to be feasible and efficient. The agent negotiation was an important feature of the system in order to improve perceived quality of service and to improve the load balancing of the traffic.

The multi-agent system proposed in this work is general enough to be applied to different kinds of cellular networks.

References

[AA93] Yoshihiko Akaiwa and Hidehiro Andoh. Channel Segregation – A Self-Organised Dynamic Channel Allocation Method: Application to TDMA/FDMA Microcellular System. *Journal on Selected Areas in Communication*, volume 11, number 6, pages 949-954. August 1993.

[AC87] P. E. Agre and D. Chapman. Pengi: an implementation of a theory of activity. In *Proc. of the Sixth National Conference on Artificial Intelligence (AAAI-87),* pages 268-272, Seattle, WA, 1987.

[AS94] Stephen Appleby and Simon Steward. Mobile Software Agents for Control in Telecommunications Networks. *BT Technology Journal*, volume 12, number 2, pages 104-113. April 1994.

[Aus62] Austin J. L. How to Do Things with Words. *Clarendon Press* 1962

[Bear88] D. Bear. Principles of Telecommunication Traffic Engineering. 3rd Edition. Peter Peregrinus Ltd on behalf of the Institut of Electrical Engineers. 1988.

[Bra87] M. E. Bratman. *Intentions, Plans, and Practical Reason*. Harvard University Press, 1987.

[Bro90] Rodney A. Brooks. A robot that walks: Emergent behaviors from a carefully evolved network. In Patric Henry Winston and Sarah Alexandra Shellard, editors, *Artificial Intelligence at MIT, Expanding Frontiers*, pages 28-39. MIT Press, Cambridge, Massachusets, 1990.

[Bun85] Brian D. Bunday. Basic Linear Programming. Edward Arnold (Publishers) Ltd. ISBN 0 7131 3509 3. London, 1985.

[C-LC93] I. Chih-Lin and Pi-Hui Chao. Local Packing – Distributed Dynamic Channel Allocation at Cellular Base Station. Proceedings of IEEE *Global Telecommunications Conference* (GLOBECOM'93), part 1, volume 1, pages 293-301. New York, NY, USA. 1993

[C-LC94] I. Chih-Lin and Pi-Hui Chao. Distributed Dynamic Channel Allocation Algorithms with Adjacent Channel Constraints. Proceedings of the 1994 *5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* (PIMRC'94), volume 1, pages 169-177. Amsterdam, The Netherlands. 1994.

[Clearw96] Scott H. Clearwater. Market-Based Control: A Paradigm for Distributed Resource Allocation. World Scientific, 1996.

[Dai92] John N. Daigle. Queuing Theory for Telecommunications. *Addison-Wesley Series in Telecommunications*. Addison-Wesley Publishing Company, Inc., pages 93, 1992.

[D-RFR96] Enrico Del Re, Romano Fantacci and Luca Ronga. A Dynamic Channel Allocation Technique based on Hopfield Neural Networks. *IEEE Transactions on Vehicular Technology*, volume 45, number 1, pages 26-32. February 1996.

[DS84] J. Dunlop and D. G. Smith. Telecommunications Engineering. Van Nostrand Reinhold (UK) Co. Ltd.1984.

[DSJ97] Sajal K. Das, Sanjoy K. Sen and Rajeev Jayaram. A dynamic load balancing strategy for channel assignment using selective borrowing in cellular mobile environment. *Wireless Networks*, volume 3, pages 333-347, 1997.

[DSJA97] Sajal K. Das, Sanjoy K. Sen, Rajeev Jayaram and Prathima Agrawal. A Distributed Load Balancing Algorithm for the Hot Cell Problem in Cellular Mobile Networks. Proceedings of the *1997 6th IEEE International Symposium on High Performance Distributed Computing*, pages 254-263. Portland, USA. 5-8 August 1997.

[E-DWS89] Sami A. El-Dolil, Wai-Choong Wong and Raymond Steele. Teletraffic Performance of Highway Microcells with Overlay Macrocell. *IEEE Journal on Selected Areas in Communciation*, volume 7, number 1, pages 71-78. January 1989.

[Ekl86] Berth Eklundh. Channel Utilisation and Blocking Probability in a Cellular Mobile Telephone System with Directed Retry. IEEE *Transactions on Communications*, volume 34, number 4, pages 329-337. April 1986.

[EM89] David Everit and David Manfield. Performance Analysis of Cellular Mobile Communication Systems with Dynamic Channel Assignment. IEEE *Journal on Selected Areas in Communications*, volume 7, number 8, pages 1172-1180. October 1989

[ESG82] Said M. Elnoubi, Rajendra Singh and Someshwar C. Gupta. A new frequency channel assignment algorithm in high capacity mobile communication systems. IEEE *Transactions on Vehicular Technology*, volume VT-31, number 3, pages 125-131. August 1982.

[FA91] Y. Furuya and Y. Akaiwa. Channel Segregation: a Distributed Channel Allocation Scheme for Mobile Communication Systems. *IEICE Transactions*, volume 74, pages 949-954. 1991.

[FA93] H. Furukawa and Y. Akaiwa. Self Organised Reuse Partititoning, a Dynamic Channel Assignment Method in Cellular Systems. Proceedings of the 1993 *43rd IEEE Vehicular Technology Conference*, pages 524-527. New York, NY, USA. 1993.

[Far96] Saleh Faruque. Cellular Mobile Systems Engineering. *Mobile Communication Series*. Artech House Publishers. ISBN 0-89006-518-7.1996.

[Fer95] I. A. Ferguson. Integrated control and co-ordinated behaviour: a case for agent models. In M. J. Wooldridge and N. R. Jennings, editors. *Intelligent Agents – Theories, Architectures, and Languages*, volume 890 of *Lecture Notes in Artificial Intelligence*, pages 203-218. Springer-Verlag, 1995.

[FHN71] R.E. Fikes, P. E. Hart, and N. Nilsson. STRIPS: A New Approach to the Application of Theorem Proving. *Artificial Intelligence,* volume 2, pages 189-208, 1971.

[FIPA97] FOUNDATION FOR INTELLIGENT PHYSICAL AGENTS. FIPA 97 Specification, Version 2.0 - Part 2 - Agent Communication Language Publication. October, 1998. http://www.fipa.org.

[FJ91] J. Ferber and E. Jacopin. The framework of eco-problem solving. In Y. Demazeau and J.-P. Müller, editors, *Decentralized A. I.*, volume 2, pages 181-193. North-Holland, 1991.

[FNY89] D. F. Ferguson, C. Nikolaou and Yechiam Yemini. An Economy for Flow Control in Computer Networks. Proceedings for the *8th Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 110-118, Washington, 1989.

[Gar00] Vijay K. Garg. IS-95 CDMA and CDMA 2000 – Cellular/PCS Systems Implementation. Prentice Hall Communication Engineering and Emerging Technologies. Series Editor – Theodore S. Rappaport. Prentice Hall PTR. 2000

[GM96] Tibor Gyires and Krishnan Muthuswamy. Heuristic Routing Algorithm for Re-establishing Interrupted Connections in Telecommunication Networks. Proceedings of the *ECAI Workshop on Intelligent Agents for Telecoms Applications (IATA'96)* Budapest, Hungary. 11-16 August 1996.

[Hal83] S. W. Halpern. Reuse partitioning in Cellular Systems. Proceedings of the 1983 33rd IEEE *Vehicular Technology Conference*, pages 322-327. New York, NY, USA. 1983.

[HBWC99] Hayzelden, A. L. G., J. Bigham, M. Wooldridge and L. G. Cuthbert. Future Communication Networks using Software Agents. Chapter 1 in *Software Agents for Future Communications Systems*. Hayzelden and Bigham (editors). Published by Springer Verlag, ISBN 3-540-65578-6. 1999. Also available at http://www.agentcom.org/agentcom/

[HB98] Alex L. G. Hayzelden and John Bigham. Heterogeneous Multi-Agent Architecture for ATM Virtual Path Network Resource Configuration. Proceedings of 2nd International Workshop on Intelligent Agents for Telecoms Applications (IATA'98). Lecture Notes in Artificial Intelligence 1437, pages 45-59. Sahin Albayrak and Francisco J. Garijo Editors. July 1998.

[Heal80] G. Heal. Planning without prices. Review of Economic Studies, volume 36, pages 346-362, 1969.

[Hew77] Carl Hewitt. Viewing control structures as patterns of passing messages. *Artificial Intelligence* (Journal), volume 8, number 3, pages 323-364. Netherlands. 1977.

[HSS80] Y.C. Ho, L. Servi and R. Suri. A class of center-free resource allocation algorithms. *Large Scale Systems – Theory and Applications* Journal, volume 1, pages 51-62. Netherlands. 1980.

[JR94] H. Jiang and S. S. Rappaport. CBWL: A new channel assignment and sharing method for cellular communication systems. IEEE *Transactions on Vehicular Technology*, volume 43, number 2, pages 313-322. May 1994.

[JS96] Jonathan M. Pitts and John A. Schormans. Introduction to ATM Design and Performance. John Wiley & Sons. ISBN 0-471-96340-2. 1996.

[KE89] Johan Karlsson and Berth Eklundh. A cellular mobile telephone system with load sharing - an enhancement of directed retry. IEEE *Transactions on Communications*, volume 37, number 5, pages 530-535. May 1989.

[KG78] T. J. Kahwa and N. D. Georganas. A Hybrid Channel Assignment Scheme in Large-Scale, Cellular-Structure Mobile Communication Systems. IEEE Transactions on Communications, volume 26, number 4, pages 432-438. April 1978.

[KN96] I. Katzela and M. Naghshineh. Channel Assignment Schemes for Cellular Mobile Telecommunication Systems: A Comprehensive Survey. *IEEE Personal Communications Magazine*, volume 3, number 3, pages 10-31. June 1996.

[KS89] James F. Kurose and Rahul Simha. A Microeconomic Approach to Optimal resource Allocation in Distributed Computer Systems. IEEE Transactions on Computer, volume 38, number 5, pages 705-717. May 1989.

[Lee95] William C. Y. Lee. Mobile Cellular Telecommunications: Analog and Digital Systems. Second Edition. McGraw-Hill, Inc. ISBN 0-07-038089-9. 1995.

[LK91] Averill M. Law and W. David Kelton. Simulation Modeling and Analysis. 2nd Edition, McGraw-Hill International Series, 1991.

[Mül96] J. P. Müller. The Design of Intelligent Agents: A layered Approach, volume 1177 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, 1996.

[Mac79] V. H. Mac Donald. Advanced Mobile Phone Service: The Cellular Concept. *The Bell System Technical Journal*, volume 58, number 1, pages 15-41, January 1979.

[Meh96] Asha Mehrotra. GSM System Engineering. *Mobile Communication Series*. Artech House Publishers. ISBN 0-89006-860-7. 1996.

[NS76] A. Newell and H. A. Simon. Computer science as empirical enquiry: Symbols and search. *Communications of the ACM*, volume 19, number 3, pages 113-126, 1976.

[Nwa96] H. S. Nwana. Software Agents: An Overview. *Knowledge Engineering Review*, volume 11, number 3, pages 205-244, October/November 1996.

[O-GA98] Lauro Ortigoza-Guerrero and A. Hamid Aghvami. A Distributed Dynamic Resource Allocation for a Hybrid TDMA/CDMA System. IEEE *Transactions on Vehicular Technology*, volume 47, number 4, pages 1162-1178. November 1998.

[O-GL-R97] Lauro Ortigoza-Guerrero and Domingo Lara-Rodriguez. CPMCB: A Suitable Scheme for the Pan-European GSM System. Proceedings of the *1997 47th IEEE Vehicular Technology Conference*, part 3, pages 1679-1683. Phoenix, AZ, USA. 4-7 May 1997.

[OPNETa94] OPNET Online Documentation Version 3.0A. MIL 3, Inc., Washington, USA, 1994.

[OPNETb94] OPNET standard models. Cellular System Example Models. Release 2.4, MIL 3, Inc., Washington, USA, 1994.

[PAS96] Tejaskumar Patel, Dinesh K. Anvekar and B. S. Sonde. DCBWL: A New Channel Borrowing Scheme for Mobile Cellular Communication Systems. Proceedings of the *1996 IEEE International Conference on Personal Wireless Communications*, pages 163-167. New Delhi, India. 19-21 February 1996.

[Rai48] Erwin Raisz. General Cartography. Chapter five: Scales, Parallels and Meridians, pages 54-62. McGraw Hill 1948.

[RG91] A. S. Rao and M. P. Georgeff. Modeling Agents within a BDI-Architecture. In R. Fikes and E. Sandewall, editors, *Proc. of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 473-484, Cambridge, Massachusets, April 1991. Morgan Kaufmann.

[Sea69] Searle J.R. *Speech Acts*, Cambridge University Press, 1969.

[Sen97] Sanjoy Kumar Sen. Efficient Algorithms and Framework for Bandwidth Allocation, Quality of Service Provisioning and Location management in Mobile Wireless Computing. Dissertation for PhD, University of North Texas. Denton, Texas, December 1997.

[SHB97] R. Schoonderwoerd, O. Holland and J. Bruten. Ant-like Agents for Load Balancing in Telecommunications Networks. Proceedings of Agents'97, pages 209-216. ACM Press. Marina del Rey, CA. 1997.

[Shuffle00] Shuffle: an agent based approach to controlling resources in UMTS networks. *Information Society Technologies (IST) 5th Framework*, project number 1999-11014. The European Commission - Community Research. http://www.ist-shuffle.org.

[Siva98] Suthaharan Sivagnanasundaram. GSM Mobility Management using an Intelligent Network Platform. PhD thesis, Department of Electronic Engineering, Queen Mary and Westfield College, University of London, London, UK, 1997.

[Smith88] R. G. Smith. The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. In *Readings in Distributed Artificial Intelligence*, pages 357-366. Morgan Kaufman Publishers, Bond, A. and Gasser, L. editors, 1988.

[Som96] Fergal Somers. HYBRID: Intelligent Agents for Distributed ATM Network Management. Proceedings of First International Workshop on Intelligent Agents for Telecoms Applications (IATA'96) at the European Conference on Artificial Intelligence (ECAI'96). Budapest, Hungary.12-16 August 1996.

[TI88] Jun Tajima and Kenji Imamura. A Strategy for Flexible Channel Assignment in mobile Communication Systems. *IEEE Transactions on Vehicular Technology*, volume 37, number 2, pages 92-103. May 1988.

[TJ91] Sirin Tekinay and Bijan Jabbari. Handover and Channel Assignment in Mobile Cellular Networks. *IEEE Communications Magazine*, volume 29, number 11, pages 42-46. November 1991.

[WJ95a] M. J. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, volume 10, number 2, pages 115-152, June 1995.

[WJ95b] M. J. Wooldridge and N. R. Jennings, editors. *Intelligent Agents – Theories, Architectures, and Languages*, volume 890 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, 1995.

[WP98] T. White and B. Pagurek. Towards Multi-Swarm Problem Solving in Networks. Proceedings of 3rd International Conference on Multi-Agent Systems, pages 333-340. Paris, France. July 1998.

[WS94] Kevin A. West and Gordon L. Stüber. An Aggressive Dynamic Channel Assignment Strategy for a Microcellular Environment. *IEEE Transactions on Vehicular Technology*, volume 43, number 4, pages 1027-1038. November 1994.

[WV98] R. Weihmayer and H. Velthuijsen. Intelligent Agents in Telecommunications. *Agent Technology: Foundations, Applications, and Markets*, pages 203-217 Nicholas R. Jennings and Michael J. Wooldridge editors. ISBN 3-540-63591-2. Springer 1998.

[YA96] Frederick Ygge and Hans Akkermans. Power Load Management as a Computational Market. Proceeding of the *2nd International Conference on Multi-Agent Systems ICMAS'96*, pages 393-400, Kyoto, Japan, December 1996.

[YY94] Kwan Lawrence Yeung and Tak-Shing P. Yum. Compact Pattern Based Dynamic Channel Assignment for Cellular Mobile Systems. IEEE *Transactions on Vehicular Technology*, volume 43, number 4, pages 892-896. November 1994.

[ZY89] Ming Zhang and Tak-Shing P. Yum. Comparisons of Channel Assignment Strategies in Cellular Mobile Telephone Systems. IEEE *Transactions on Vehicular Technology*, volume 38, number 4, pages 211-215. November 1989.

[ZY91] Ming Zang and Tak-Shing P. Yum. The Nonuniform Compact pattern Allocation Algorithm for cellular Mobile networks. IEEE Transactions on Vehicular Technology, volume 40, number 2, pages 387-391. May 1991.