Project Reference Number	TP/3/PIT/6/I/15927
Project Title	Mobile Network Optimisation Through Advanced Simulation
Project Acronym	MONOTAS
Project Partners	Multiple Access Communications Limited Vodafone Group Services Limited Nortel

Report Title	Final Project Report (Public Version)
Report Reference Number	199.PUB (Deliverable 7.2)
Report Date	22 January 2008
Editor(s)	Peter Gould
Author(s)	Simon Avis, Steve Baines, Damian Bevan, Derek Chandler, Steven Hall, Peter Myerscough-Jackopson, Neil Scully Stefan Thiel
Reviewer(s)	Andy Jeffries, David Lister

Abstract	This project has examined the feasibility of decreasing the time taken to optimise 3G and 4G mobile networks by using advanced simulation techniques to examine the performance achieved with a particular parameter setting and using optimisation algorithms to find the best parameter settings for a given traffic load. In this document we present a summary of the work performed and the key results and conclusions of the project.
Keyword(s)	3G, 4G, Femtocells, Network Optimisation, Simulation, Computing Cluster
Confidentiality	PUBLIC

© 2008 Multiple Access Communications Limited, Nortel Networks, Vodafone Group Services Limited

Document History

Date	Version	Comment	Editor
22 January 2008	1.0	Public Release	Peter Gould

Acknowledgements

The project consortium would like to thank the Department of Trade and Industry and the Technology Strategy Board who provided part of the funding for this project through the Technology Programme. It would also like to thank the external Project Monitoring Officer, Alan Steventon of Judal Associates, who provided valuable direction, advice and guidance during the course of the project.

Executive Summary

Project MONOTAS - Final Project Report (Public Version)

Network operators expend significant amounts of time, effort and money on tuning their networks to meet the needs of their customers. This is currently a very manual process, relying on the knowledge and skills of experienced optimisation engineers. It is also a very challenging process, since the network load can vary dramatically from hour-to-hour and even from minute-to-minute. Typical network optimisation cycles are currently of the order of months and, as a result, network operators must tune their networks to an average loading scenario to ensure that satisfactory performance can be maintained across the range of different loads that may be experienced in practice. In this project we have examined the feasibility of dramatically decreasing the optimisation cycle time by researching techniques for automating the selection of the best parameter settings at a given point in time.

The approach to optimisation investigated as part of this project has been to build representative computer models of a network and the associated traffic load and use simulations based on these models to understand the performance achieved with different parameter settings. By applying optimisation algorithms that use the results of these simulations to determine the most appropriate parameter settings for a given network load, we have shown that it is feasible to use simulation models in the optimisation loop. We have also shown that it is feasible to develop network models and simulations that can be distributed across a large scale computing cluster to limit the amount of time required to identify the most appropriate parameter settings and hence speed up the optimisation process.

Our investigations have shown that it is possible to improve the capacity of a network in hotspot scenarios by up to 80%, for a given set of key performance targets, by adjusting particular parameters within the network. We have also seen that more advanced handover algorithms can increase the overall capacity of a network by 20% and improvements in the cell selection process can provide significant throughput benefits to users in the most disadvantaged locations within the network.

As part of the project, we also examined the deployment of small, self-installed domestic base stations, known as femtocells. There is currently a significant amount of interest within the industry around these devices, but very little analysis of their performance and, in particular, their impact on an existing network, has entered the public domain. In this project we examined the impact on network performance of femtocells operating in closed access mode (ie, they can be accessed by only a small subset of subscribers) and open access mode (ie, they can be accessed by any network subscriber). We found that operating femtocells in open access mode has the potential to improve significantly the quality of service experienced by the femtocell and macrocell users. By contrast, although closed access femtocells have the potential to improve the quality of service experienced by the femtocell users, the macrocell users are likely to experience a dramatic degradation in their quality of service, since they will suffer significant interference from the femtocells.

Table of Contents

List of Abbreviations

3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
CDMA	Code Division Multiple Access
EDR	Effective Data Rate
GSM	Global System for Mobile Communications
HSDPA	High Speed Downlink Packet Access
LTE	Long Term Evolution
MONOTAS	Mobile Network Optimisation Through Advanced Simulation
NGMN	Next Generation Mobile Networks
OFDMA	Orthogonal Frequency Division Multiple Access
PC	Personal Computer
RAN	Radio Access Network
RF	Radio Frequency
SIR	Signal-to-Interference Ratio
SNIR	Signal-to-Noise-plus-Interference Ratio
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
VPN	Virtual Private Network
WiMAX	Wireless Interoperability for Microwave Access

1. Introduction

Large-scale mobile radio networks are complex systems and their behaviour is governed by a range of different configuration parameters. These parameters range from physical factors, such as the location of the base stations and the orientation of their antennas, to parameters that control the operation of software processes running within the network. Network operators expend significant amounts of time, effort and money on trying to find a set of these parameters that ensures the network's performance meets the needs of its users.

One of the key challenges the network operator faces is adapting its network to changes in the traffic load and the mix of services being used, which can vary dramatically from day-to-day and even from minute-to-minute. Some network parameters can only be adapted relatively slowly (eg, the location of the base stations) and, as a result, these will be used to tailor the network to longer term trends in the network traffic. Tools that allow network operators to assess the best site locations and antenna configurations based on a particular traffic load are well established and have been on the market for some years.

On the other hand, many of the network parameters, such as the transmitted power of the base stations or the software variables controlling handover within the network, can be varied very rapidly. However, there are no commercial tools available that will allow a network operator to assess the impact of changing these parameters and, instead, the effects of any changes must be assessed through drive testing, network performance statistics and customer feedback. As a result, these parameters tend to be adjusted relatively infrequently as part of monthly or quarterly network optimisation cycles.

The main aim of Project MONOTAS (Mobile Network Optimisation Through Advanced Simulation) has been to investigate the use of network simulations, based on complex computer models, as a means of rapidly identifying the best parameter settings to cope with a particular traffic load.

In a project of this nature it is important to consider the problem from a range of different perspectives and, for this reason, a project consortium was assembled that contained a network operator (Vodafone Group Services Limited), an equipment vendor (Nortel) and a tools and services provider (Multiple Access Communications Limited). Each project partner dedicated three man-years of effort to the project and, as the project leader, Multiple Access Communications Limited (MAC Ltd) also dedicated a further half of a man-year to project management duties.

In this document we present an overview of the work performed during the project and some of the key results and conclusions. In Section 2 we outline the approach taken in the simulator development and we also provide details of the computing cluster that was assembled to run the simulations. In Section 3 we discuss the work that was performed on the optimisation of Third Generation (3G) systems and in Section 4 we discuss the optimisation of Fourth Generation (4G) systems. The work performed on investigating and analysing femtocells is described in Section 5 and in Section 6 we present the key conclusions of the project. Finally, in Section 7 we highlight ways in which the consortium members plan to exploit the project results.

2. Simulation Approach

One of the key elements of this project was the design and implementation of simulation models that could be used to rapidly assess the performance of a network under a given traffic load. These would be used to determine the performance that could be achieved with specific parameter settings and also to find optimal parameter settings using different optimisation algorithms. In this section we discuss the approach that was taken in the development of the simulators and the implementation of the computing cluster on which the simulators would run.

2.1. High Level Simulator Design

The simulators developed as part of this project were used to investigate the impact of changing different parameters on network performance and to assess the potential performance improvement that could be achieved through parameter optimisation. It was also important to establish the credibility of the simulators by comparing their results, based on predefined test scenarios, with results produced using other independently developed tools and models. This meant that the simulators had to be able to access input data created using a set of known input parameters. However, it was also important as part of the project to demonstrate the feasibility of feeding information gathered from a real network into the simulator and, therefore, the simulator also had to be designed to accept processed data from a real network.

The high-level simulator design is shown in block diagram form in Figure 1. On the left of the diagram we have the two different input blocks. Input Block 1 is used to generate user equipment (UE) trace files to test the simulators and also to investigate specific features of the system operation. A UE is a terminal device that is used to connect to the radio network, eg, a mobile phone or a wireless data card. The UE Trace File generator uses the Coverage Model, which contains the propagation loss information, including fading, for each of the base stations within the system, and the Traffic Map, which contains details of the UE distribution and density, to generate UE traces, which are records of the path loss from the UE to each base station in the network over time. These traces are produced by placing fictitious UEs within the network coverage area, based on the predefined UE distribution and density. The Coverage Model in Input Block 1 can either be based on a simple network arrangement (eg, tessellated hexagons) or it can be based on a real network deployment. As part of this project, part of Vodafone’s network was used to generate UE traces based on a real network deployment.

Input Block 2 is used to convert data collected from a real network into UE traces that can be used in the simulator. The network data consists of call traces generated either within the network or from test mobiles. These call traces are analysed to determine the path loss from the UE to each base station in the network at each point in time and this information is used to create the UE Trace Files. In this case we do not require a Coverage Model or a Traffic Map, since this information will be incorporated into the trace information from real call traces.

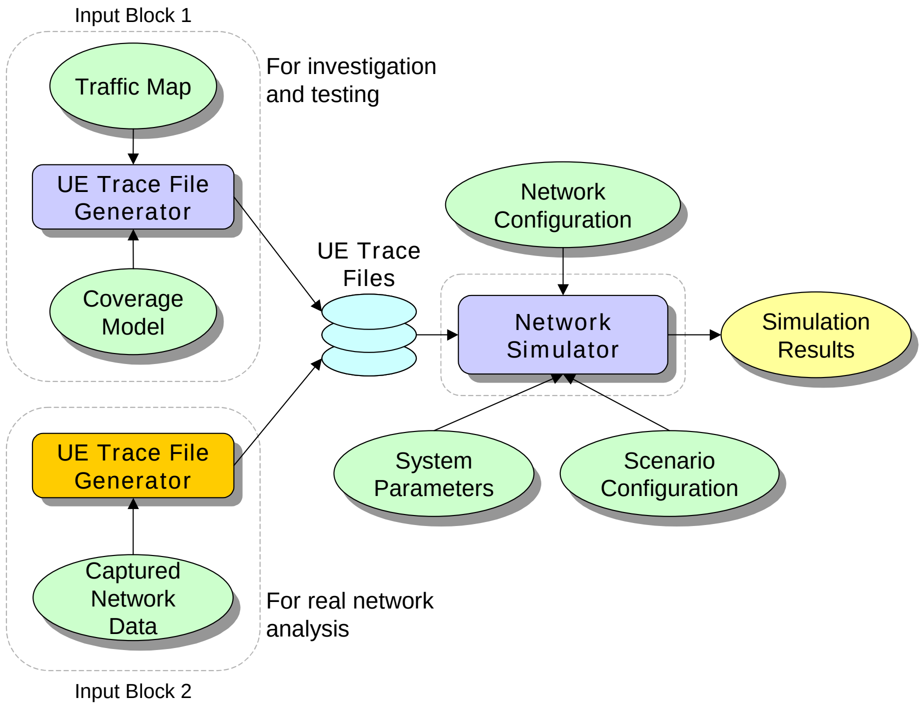

Figure 1 High level simulator design

In addition to the UE Trace Files, the Network Simulator also has a number of other inputs. The Network Configuration data contains information on the manner in which the base stations within the network are configured (eg, the arrangement of sectors and their types). The System Parameters information contains general data on the system being modelled, including the conversion between the signal-to-noise-plus-interference ratio (SNIR) and the achievable data rate. Finally, the Scenario Configuration data provides information on the manner in which the UE Trace Files should be used. For example, it is possible to produce two different sets of UE trace files, each with different user distributions that can be used to represent different times of the day. The Scenario Configuration can be used to define the manner in which the simulator transitions between these two different distributions during the simulation.

In addition to the flexibility offered through this approach, the design also offers other benefits in terms of the simulation speed, since the architecture allows various aspects of the simulation model to be generated once and used in many subsequent simulations, rather than having to generate the information on each run.

2.2. Requirement for Location Information

One important area that was investigated as part of this project was the possible use of UE location information within the optimisation process. This information can be very important when it comes to planning a network or adjusting the physical configuration of the base stations (eg, the orientation of the antennas). For example, when deciding where to place a new base station, it is important to ensure that it is placed in a region with a high concentration of users to ensure that it carries enough traffic to justify its cost. However, when we analysed the potential value of location information in the optimisation of software parameters within the network or base station transmission powers, we concluded that location was less important and, in general, it was used as a means to derive other information such as received signal strength. Therefore, we concluded that mobile location was of limited value in rapid network adaptation and measurement data relating to the radio environment experienced by the UEs was far more important.

2.3. Computing Cluster

As part of this project, the consortium also investigated ways in which the simulations could be made to run as fast as possible. This would allow the results of various investigations to be made available more quickly and, hence, increase the number of investigations that could be performed. It was also important to demonstrate the feasibility of using complex simulations as part of the optimisation process and one aspect of this was to show that simulation results could be generated within reasonable timescales. As we have already explained, the simulators themselves were designed to be as efficient as possible, with as much of the input data processed into a format such that it could be used in multiple simulations and there was no requirement to generate the information for each run.

However, it is also important to ensure that the simulation platform on which the simulations will be run has sufficient processing power and also that this power is used as efficiently as possible. Therefore, as part of this project, the team constructed a large-scale computing cluster consisting of 29 personal computers (PC), each with four processors (ie, a total of 116 processors). The computing resources were distributed roughly evenly across the project partners' three locations and the project team demonstrated the feasibility of linking the cluster together using a virtual private network (VPN), such that each member of the project team could gain access to the full computing power. Figure 2 shows a photograph of the computing cluster installed at one of the project partners' offices.

Figure 2 A photograph of the computing cluster at one of the project partners' offices

During the project, the team also examined different cluster architectures and weighed up the advantages and disadvantages of dedicating particular computing resources to individual tasks. Since scalability was a key requirement for the computing cluster, it was decided that each PC should be configured in the same manner and each processor would be capable of running the full simulation stack, rather than distributing each simulation across several computers, with each computer running a different component of the simulation.

The job management software chosen for the computing cluster was the high throughput computing environment, Condor [1,2]. This allows simulation jobs to be efficiently scheduled across all of the available processors.

2.4. Optimisation Process

Network optimisation was one of the key elements of the project and at the outset a review of various optimisation procedures was performed [3]. Ordinal optimisation was chosen as the key optimisation approach and, along with a refinement that optimally assigned computing time to simulations [4], this method was included into control software that submitted simulation jobs on the computing clusters and processed the results before submitting new simulation jobs. This was important since the simulations were of a stochastic nature and so the performance of a particular network parameter configuration varied between simulation runs. Therefore a means had to be found of ensuring the most promising network configurations were identified and simulated using the available resources.

3. 3G System Optimisation

The optimisation of 3G systems was the initial focus of Project MONOTAS. Three separate areas were researched. The first area was pilot power optimisation described in Section 3.1. In 3G systems, the pilot is used for cell selection and it commonly consumes between 5 and 10% of a cell’s total power. Optimising the pilot power allows the boundaries of a cell to be altered and load balancing to be performed. Additionally, if reductions in pilot power are possible, interference within the system will be decreased and this, in turn, will result in an increase in system capacity.

The second area of research was call admission control, whose main function is to connect as many users as possible to the network, whilst still maintaining an acceptable quality of service for those already connected. 3G systems utilise code division multiple access (CDMA) and are considered to have a ‘soft capacity’, whereby more users can continue to be added to the network, but at the detriment to those already in the system. In Section 3.2 we show how call admission control can use measurable system outputs to consider when users should be blocked from entering a network.

The optimisation of the soft handover process was the final area considered. Soft handover allows a mobile to be connected to multiple base stations simultaneously. This aims to prevent a connection from being lost due to the rapidly changing radio frequency (RF) environment, with the hope that if one connection fails, the others will not. Each additional connection a mobile makes consumes network resources, which could otherwise be used to serve another mobile. The balance between ensuring connectivity through soft handover for individual mobiles and the capacity of the system as a whole is the subject of the optimisation work described in Section 3.3.

3.1. Pilot Power Optimisation

Our investigations into the optimisation of transmitted pilot power at each base station within the network have shown that there is the potential for significant gains over a network with a fixed pilot power at each base station. By adjusting the pilot power, the traffic within the network can be moved from heavily loaded base stations to more lightly loaded base stations and this means the network capacity can be better utilised compared with a network using fixed pilot powers. We found that the largest gains are achieved when there is a non-uniform loading within the network and pilot power optimisation performs best when there is a traffic 'hot spot' in one cell within the network. We have found that with very severe hot spots, the potential capacity gains of pilot power optimisation can be very large compared to the situation where the pilot powers are fixed across the network. In Figure 3 we show the results for the uplink (ie, the link from the UE to the base station), assuming an active set size of six (ie, a UE can be connected to up to six cells at the same time). The *y*-axis shows the network capacity achieved with a particular combination of pilot power settings relative to a network with a uniform pilot power across all of the cells. The *x*-axis represents the different pilot power configurations that were evaluated, ordered by improving performance from left to right. The different curves represent the different hot spot loading factors, where the ratios given represent the relationship between the traffic in the hotspot to the traffic across the network as a whole, ie, a ratio of 1:57 means that for each user placed in the hotspot, 57 users are distributed throughout the network. The results show that increases in uplink network capacity of close to 60% are possible for large traffic hotspots (eg, a hotspot loading factor of 1:1). On the downlink (ie, the link from the base station to the UE), with an active set size of six, Figure 4 shows that capacity gains range from 6% for minor hotspots, up to 30% when a hotspot is very severe.

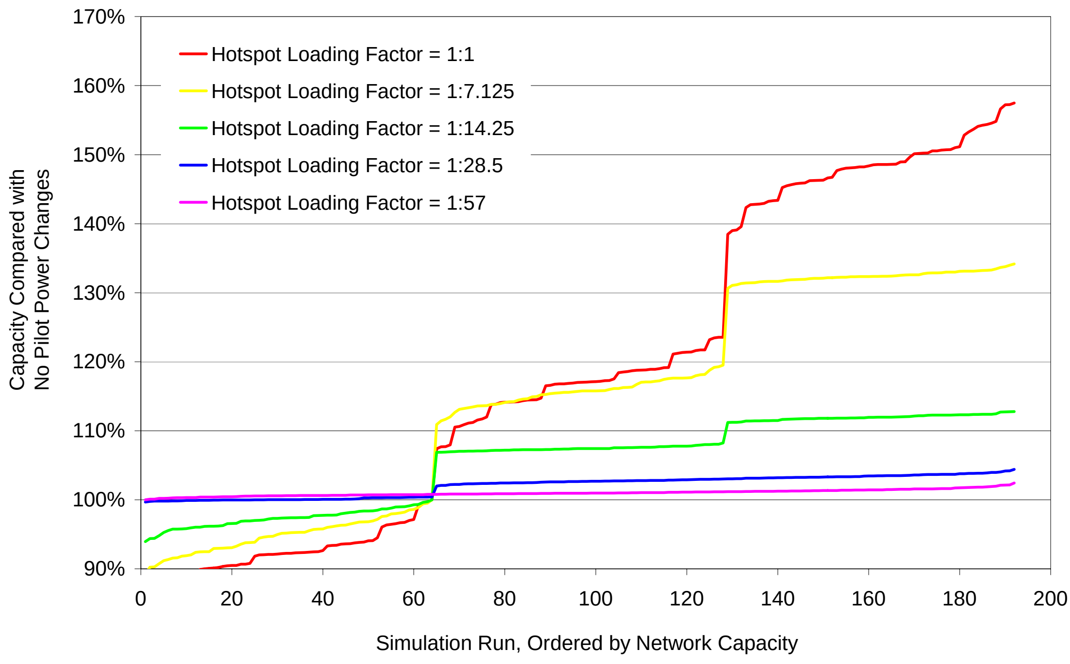

Figure 3 Uplink capacity gain for pilot power optimisation with an active set size of six

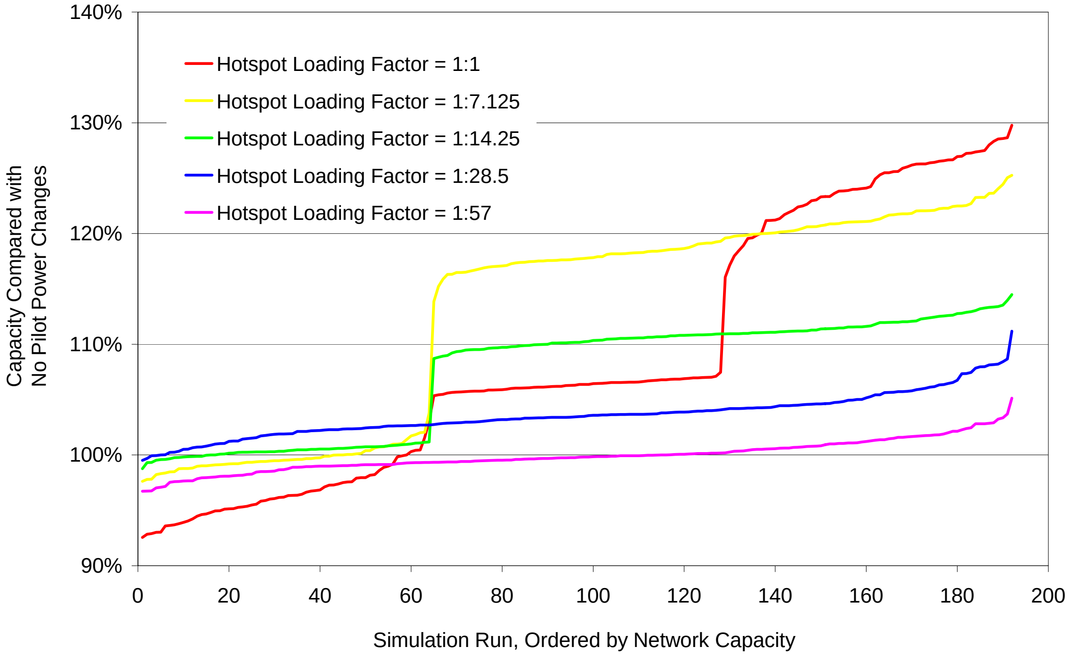

Figure 4 Downlink capacity gain for pilot power optimisation with an active set size of six

Finally, we considered the case for mobiles with an active set size of one to mimic high-speed downlink packet access (HSDPA) users, which are unable to use soft handover. Under these settings, gains of around 30% were seen even in cases where minor hotspots were present in the network.

3.2. Call Admission Control

Call admission control describes the means by which a network decides whether or not to allow a new user to access the network resources. The main aim of the call admission control algorithm is to ensure that as many calls as possible are accepted onto the network, whilst still maintaining the quality of the calls already being served by the network. As part of Project MONOTAS, the project team has investigated whether the parameters controlling the call admission control process should be rapidly adapted in response to the network load to provide a better overall network performance, compared with fixed call admission control parameters.

Simple call admission control schemes tend to use uplink noise rise (ie, the total amount of received power measured at a base station) to set a limit on the amount of traffic that can be served at a particular base station. We examined the potential capacity increase that could be achieved if the noise rise within the network could be increased beyond its usual limit of 6 dB. The gains have been evaluated for different severities of hotspot, as defined previously, and the results are shown in Figure 5. These show that potential gains of over 18% could be achieved, depending on the size of the hotspot, with only the most severe hotspot showing less impressive gains. Figure 5 also shows the results for a network with a uniform traffic load, ie, without a traffic hotspot.

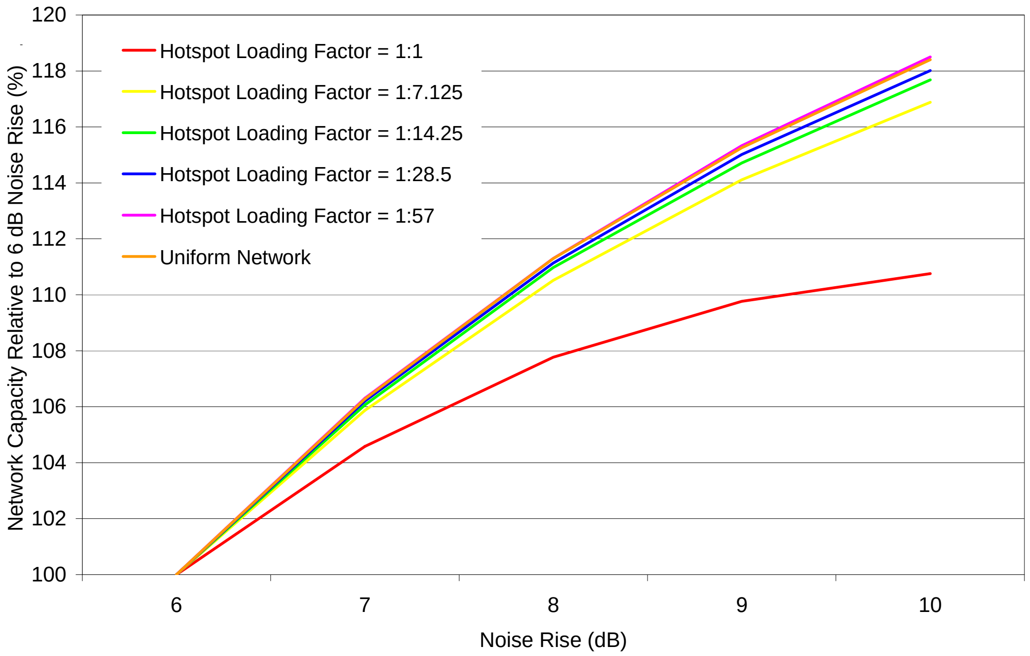

Figure 5 Capacity improvement as noise rise is increased

Our investigations into call admission control also suggested that performance gains may be possible by using more complex algorithms to control the admission of high data rate users. Current algorithms appear to limit the number of high data rate users (ie, those users accessing data rates of 384kb/s and above) to just two per sector, but our investigations showed that using noise rise as a measure of a cell's congestion allowed a third high data rate user to be admitted to some cells. This could give a potential capacity increase of 17%.

During our investigations, we also discovered that around 15% of a base station's power is reserved for handover calls, ie, calls that enter a cell by means of a handover from a neighbouring cell. This power is reserved to decrease the probability that handover calls are prematurely terminated because no resources are available on the target cell at the expense of an increase in the probability that a new call will be blocked. In low mobility situations (eg, traffic jams) where the number of handovers is decreased, this degree of reserved power can be reduced and the overall network capacity can be improved.

3.3. Handover Parameter Optimisation

Handover is the means by which a mobile remains connected to the most appropriate serving base station as it moves around the network. As we have already explained, in 3G systems, the mobile has the ability to communicate with more than one base station simultaneously in a process known as soft handover. Supporting a single mobile using two base stations is inherently less efficient when compared with supporting a mobile on a single, optimum base station, but it can be necessary to prevent a connection from dropping in a rapidly changing RF environment. A number of different parameters are used to control the number of base stations to which a mobile is connected. Project MONOTAS has considered optimising these parameters and, in particular, those concerned with 'adding' and 'dropping' base stations from a mobile's active set, ie, the set of base stations with which the mobile is actively communicating at any point in time. The 'add/drop window' is used to define the threshold power at which cells can be added or removed from the active set, relative to the strongest cell and in Figure 6 we show the relationship between the add/drop window and the relative network capacity for different mobile speeds. This demonstrates that tailoring the add/drop threshold to the mobile speed can lead to a performance improvement, compared with using the same threshold for all mobiles, regardless of their speed. By applying a typical mix of mobile speeds to the results shown in Figure 6, we found that the overall network performance could be improved by 22% through the use of mobile-specific, speed-sensitive add/drop windows.

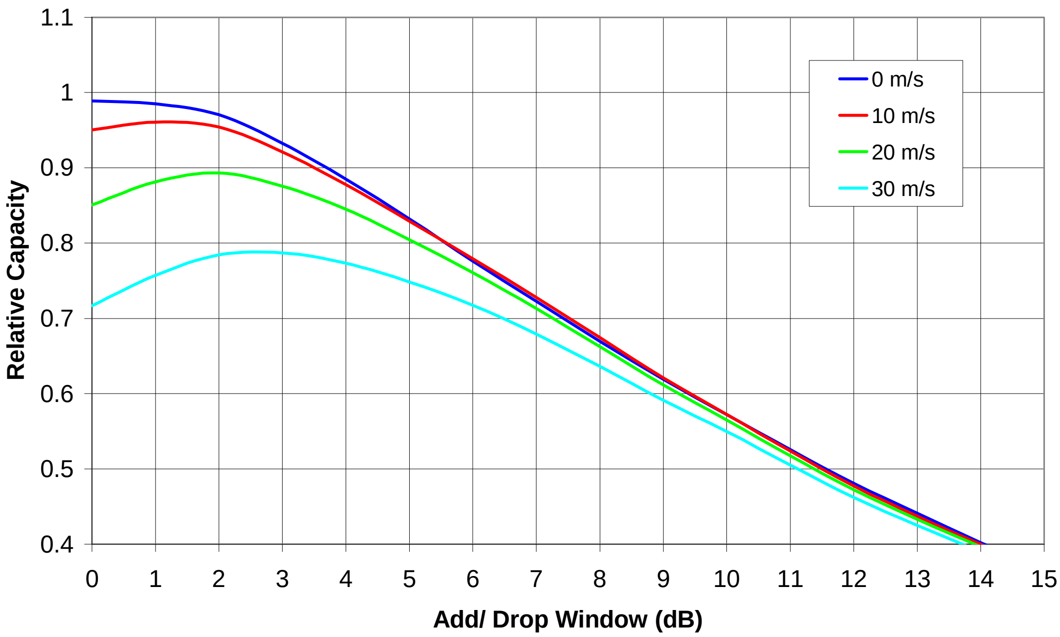

Figure 6 The relationship between network capacity and the handover add/drop window

It is important to note that, although this technique has been evaluated in the context of the 3G technology, it is applicable to any system that supports soft handover and, as such, is likely to be applicable to 4G systems. As the data rates available on 3G and 4G systems increase, it is reasonable to expect that the load generated by lower speed users will increase (ie, high data rate users, which generate a high network load, will typically be moving at a slower speed than lower data rate users) and this will further increase the capacity gains available with this approach.

We note that one of the project partners has filed a patent in the area of mobile-specific, speed-sensitive add/drop windows as a result of the work on this project.

4. 4G System Optimisation

In this section we present the work performed on the optimisation of 4G systems, eg, WiMAX and Long Term Evolution (LTE), which use orthogonal frequency division multiple access (OFDMA). These systems are in the early stages of deployment (WiMAX) or at the standardisation stage (LTE) and, hence, in addition to a requirement to understand how such systems can be optimised, there is also a requirement to understand more about how such systems can initially be designed.

There is a range of different parameters that could be used to optimise the performance of these networks including the resources allocated to each cell within the network, the manner in which the resources are allocated to individual mobiles and the transmitted power of individual sub-carriers within the OFDMA signal. However, it would be a very complex process to tune all of these parameters simultaneously and, in order to make the problem manageable, it was decided to concentrate on three specific areas. These areas were the power with which the preamble is transmitted from the base station (ie, the preamble power), the mechanism used by the network to determine the most appropriate cell to serve a particular mobile (ie, the cell selection algorithm) and the manner in which the available radio resources are subdivided between the different cells within the network and allocated to the users within each cell.

4.1. Scenario Description

The network used in the investigation of 4G systems consisted of a series of regular hexagonal cells generated by two rings of three-sectored base stations around a central base station, as shown in Figure 7. The three red circles represent the centre of the three traffic hotspots that were included in the traffic profile for specific scenarios. The traffic within the hotspots was scaled to represent either an additional 25% or 50% cell load, compared with the situation without the hotspot. The green circle shows the region where users are uniformly distributed, and the yellow rectangle contains the femtocells that were investigated, as described in Section 5. The cells that are coloured blue in Figure 7 are those for which the preamble powers were adjusted in the preamble power optimisation investigation, which is discussed in the next section.

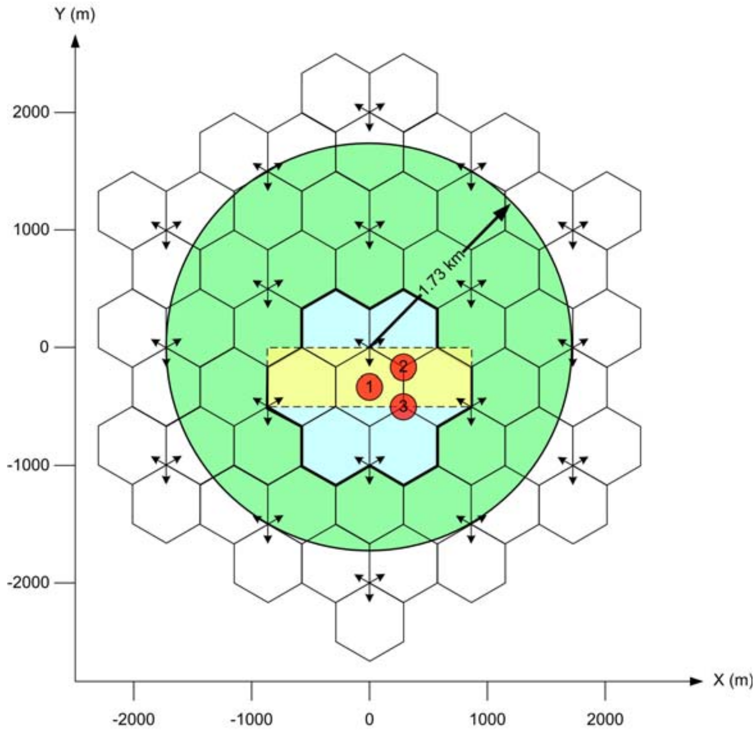

Figure 7 The 4G network used for the simulations

Simulations were also performed based on part of Vodafone’s network. This allowed the project team to investigate the performance of different optimisation approaches on a real network deployment. The layout of this part of Vodafone’s network, which contains a major European city, is shown in Figure 8. The dots represent the locations of the base stations and the arrows represent the orientation of the sectored antennas. Base stations without arrows are omnidirectional microcell sites. The scenarios simulated contained representative user densities and traffic loads.

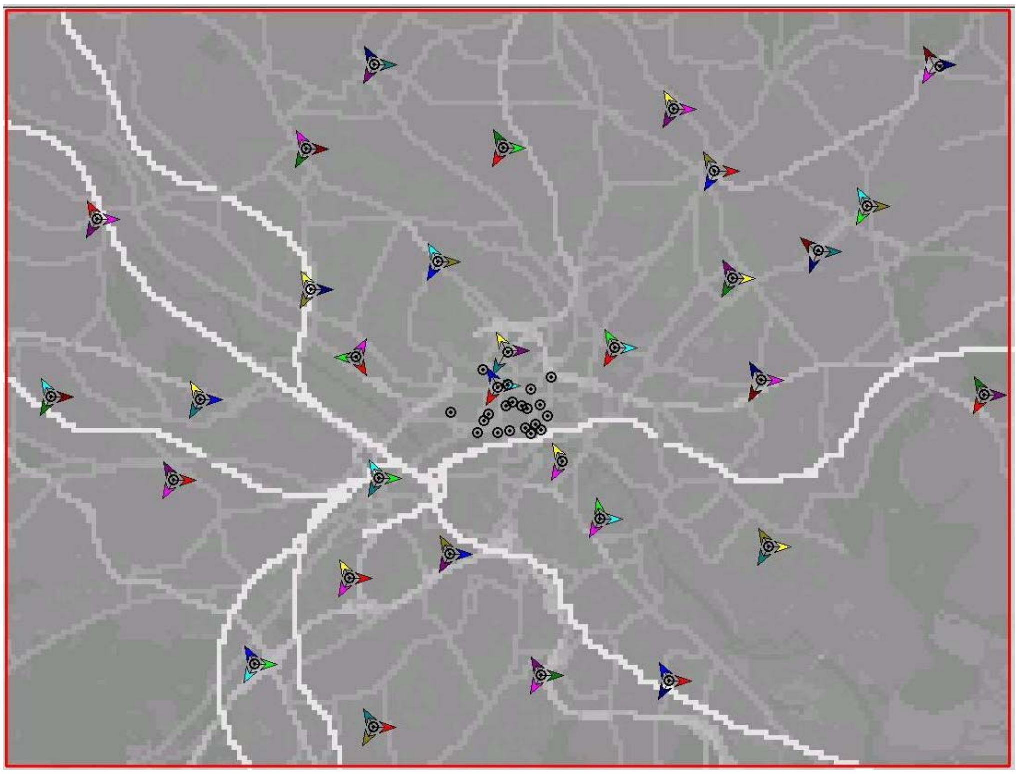

Figure 8 The layout of the part of Vodafone’s network used in project

4.2. Preamble Power Optimisation

Initially, we expect 4G systems to be deployed with a default preamble power of 43 dBm (ie, the default maximum base station transmitted power). However, this value can be adjusted to change the size of a particular cell and move traffic from overloaded cells to more lightly loaded cells. We investigated optimisation algorithms that could decrease the preamble power from its maximum value in regular steps and, in particular, we investigated four different possible power settings for each cell in the network, namely 34 dBm, 37 dBm, 40 dBm and 43 dBm.

In the optimisation process, the preamble power of the cells shaded in blue in Figure 7 was modified, together with the power of all the remaining cells in the network, which were adjusted together (ie, they all had the same preamble power setting). This gave a total of 4^8 = 65,536 possible different combinations of preamble power setting. To reduce the search space, an ordinal optimisation algorithm [4] using an initial 2049 parameter settings was chosen (2048 random settings and the default setting of all cells having a preamble power of 43 dBm). The optimisation algorithm takes the results from simulations of all the 2049 parameter settings and continues to perform further simulations on a reduced subset of the settings, with the members of this subset being selected based upon their mean performance and the variance of the results from previous simulations.

4.2.1. Simulation Results

Having verified the operation of the simulator and the optimisation algorithm, simulations were then performed for the three different hotspot scenarios. The stopping criteria for the simulation were chosen such that a simulation was halted when more than 5% of users on one cell could not be satisfied (ie, they could not be provided with their required data rate) or 2% of all the users across the network could not be satisfied. The performance of the network for each preamble power configuration was taken to be the overall data throughput achieved by the network at the point when the stopping criteria were reached.

In Figure 9 we show the simulation results for a scenario with a uniform distribution of users being placed into the green zone highlighted in Figure 7 and with the equivalent of 25% of a normal sector's load added to form a hotspot at Position 1, again highlighted in Figure 7. The *x*-axis gives the rank of the different scenario settings, sorted by throughput, with the best performing parameter settings on the left and the worst performing parameter settings to the right. In this case, the throughput is the mean throughput over the different simulations performed on the particular parameter setting, and the red dot marks the position of the uniform parameter settings (ie, with all pilot powers the same). The uniform setting is ranked 482 out of the 2049 and it provides a throughput of 26.0 Mbits/s, compared to the result with the best parameter setting with a throughput of 31.6 Mbits/s. This demonstrates that the optimisation of preamble powers provides a throughput gain of over 20% compared with the uniform settings.

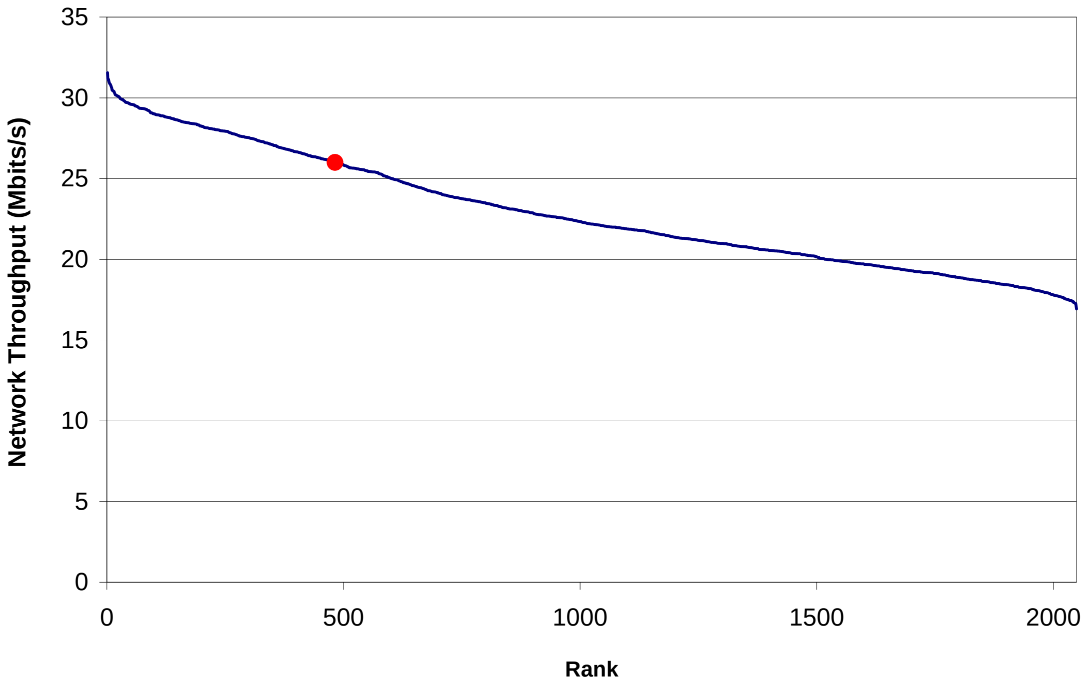

Figure 9 Simulation result for Hotspot 1 with a 25% weighting

Figure 10 shows the average preamble power settings for the 10 best configurations, where each cell has been coloured according to its relative preamble power. The cell containing the hotspot (shown as a black circle and labelled '1') can be seen to have a lower power compared to its neighbours. Therefore, with the optimal power setting, the size of the cell containing the hotspot has shrunk, thereby shedding traffic from this cell to its neighbours. It is also interesting to note the higher power of all the other non-neighbouring cells (ie, the background power). This means that the non-neighbouring cells around the hotspot have also increased their coverage area to take some of the load from the neighbouring cells, which are in turn relieving some of the load from the hotspot cell.

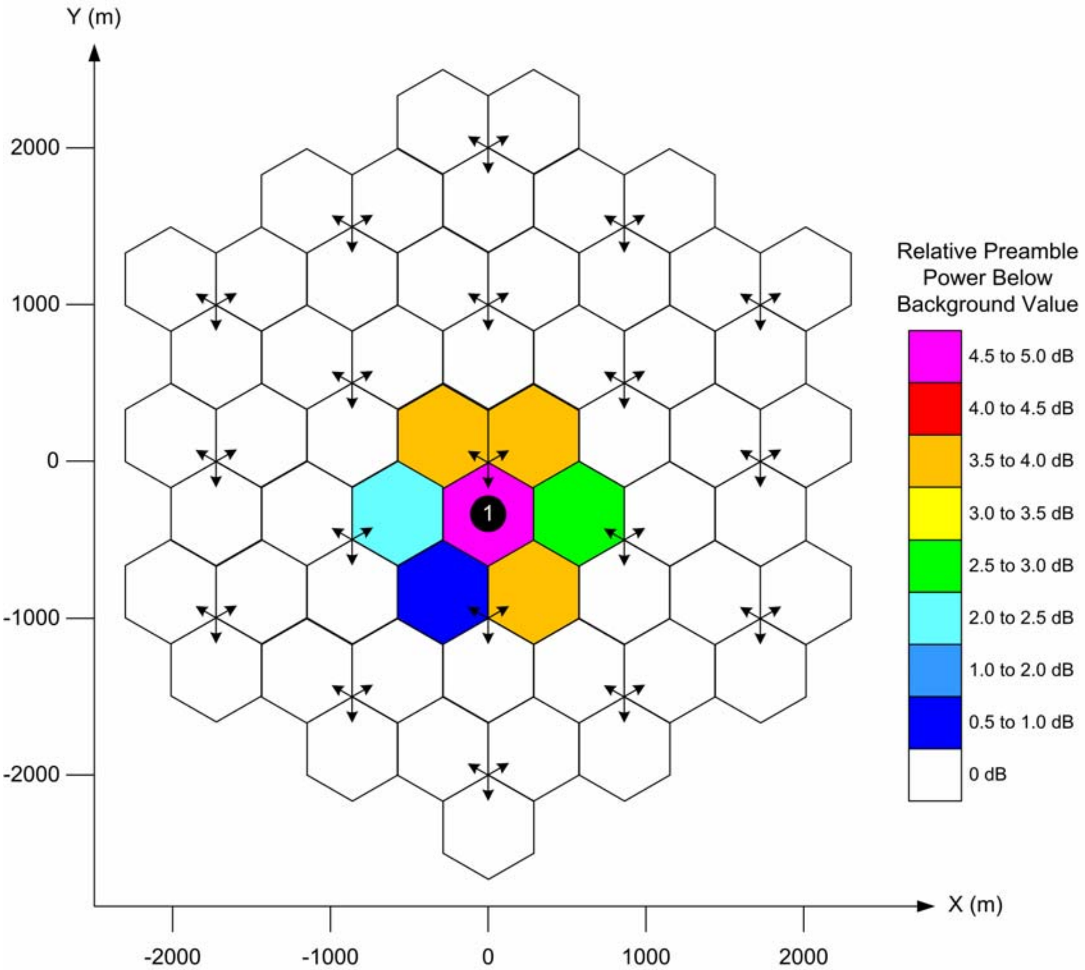

Figure 10 The average preamble power for the top ten configurations

In Table 1, the results of the simulations for the three different hotspot locations (ie, Hotspots 1, 2 and 3) and the two different hotspot scaling factors, ie, 25% and 50% are presented. In each case, the mean throughput achieved with the optimum preamble power settings is compared with the throughput for a uniform preamble power setting. This shows that the performance gains range from between 16% and 80%, with the greatest gains being achieved for Hotspot 2 (see Figure 7). Based on the criteria used, this demonstrates the manner in which the network is able to 'mould' the available capacity to the network load using the preamble power optimisation technique. It should be noted that the total capacity of the network will be at its greatest for a uniform load.

	Hotspot 1		Hotspot 2		Hotspot 3	
	25%	**50%**	**25%**	**50%**	**25%**	**50%**
Best Mean Throughput (Mbits/s)	31.6	31.8	33.4	32.2	31.5	31.1
Mean Throughput with Uniform Settings (Mbits/s)	26.0	27.3	18.5	18.2	27.0	24.5
Performance Gain %	21.4	16.4	80.6	76.7	16.8	27.1

Table 1 Comparison of the best preamble power configuration with the uniform setting

4.3. Optimisation Using EDR-Based Cell Selection

An alternative strategy for optimising macrocellular networks explored under this project is the use of effective data rate (EDR)-based cell selection. This was developed through the work on femtocell deployments where there are great differences in the loading on cells (typically macrocells have loadings in the hundreds of users, whereas femtocell loading is of the order of 1-2 users). The EDR that a user achieves (measured in bps/Hz/user) is the raw bit rate that could be achieved by that user if it were the only user being served by a particular sector (measured in bps/Hz), which is a function of that user’s SNIR, divided by the total number of users served by that sector. Thus, all things being equal, a sector that is lightly loaded will provide higher EDRs to the users it is serving than a sector that is heavily loaded. Conventional cellular systems allocate users to sectors based on the best SNIR that can be achieved from the set of servers. Our novel ‘Best EDR’ cell selection strategy selects the server that offers the best EDR, which is not necessarily the server offering best SNIR. This EDR-based cell selection mechanism, like the preamble power adjustment explored above, is aimed at adjusting the loading on the base stations within the network, and hence maximising the overall distribution of EDRs throughout the network.

For the case of a regular hexagonal network, with randomly distributed users in the network and realistic log normal shadow fading, the distribution of EDRs using the ‘Best SNIR’ and ‘Best EDR’ cell selection algorithms are shown in Figure 11. As the curves show, there is only a modest increase in EDR across the whole range of users, so that the best EDR cell selection method applied to this network makes the quality no worse from an EDR perspective.

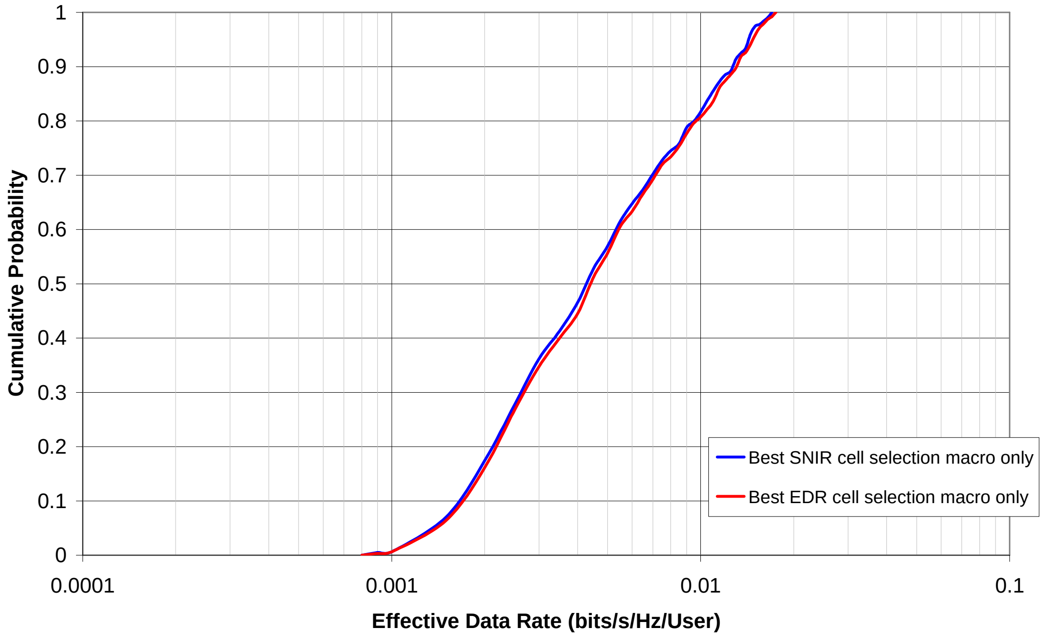

Figure 11 EDR distributions for users in a macrocell-only environment using different cell selection methods, distribution across all users

Whilst the EDR-based cell selection algorithm shows limited benefits for a regular hexagonal grid deployment with a uniformly distributed load, when a traffic hotspot occurs in such a network, the Best EDR cell selection scheme can make a difference, at least from the perspective of moving the load. Figure 12 shows a map which represents the difference in cell loading between the Best SNIR and Best EDR cell selection methods when a traffic hotspot occurs in the south facing sector of the central base station in an otherwise uniformly loaded network (labelled HS1). The hotspot represents an additional 50% of users on top of the uniform cell loading present in the rest of the network. The colour of the cells in Figure 12 reflects the change in number of users, as shown by the scale on the right. There is no log normal fading component in the data in Figure 12 so that the impact of the cell selection algorithm can be more clearly seen.

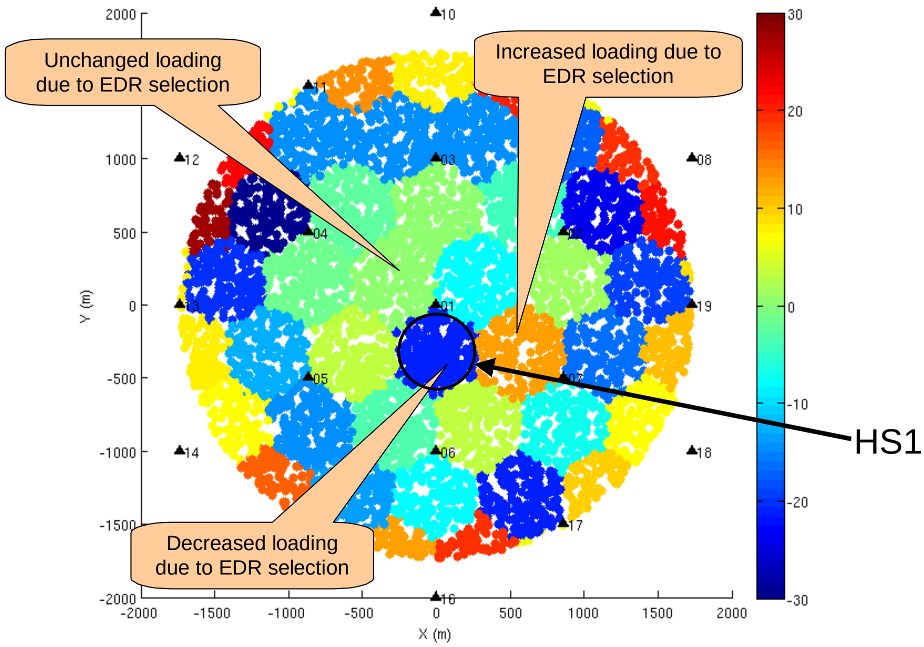

Figure 12 Difference in cell loading when moving from a Best SNIR to Best EDR cell selection scheme

As Figure 12 shows, the use of the Best EDR cell selection scheme results in the loading of the cell with the hotspot being reduced significantly, while a number of adjacent sectors (notably the one due East of the hotspot sector) see an increased load. Thus the loading on the cells is evened out by using the Best EDR cell selection technique. The outer sectors also show increased loading with the tier of sectors adjacent to them seeing reduced load. This is due to the outermost sectors having significantly lower loading due to the physical distribution of users. An EDR distribution for this case is shown in Figure 13.

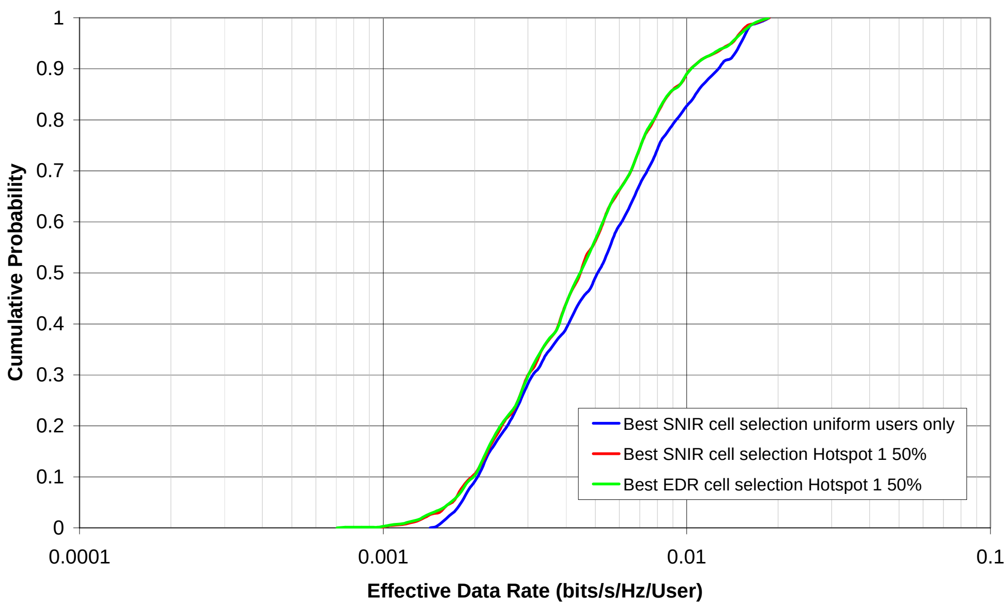

Figure 13 EDR distribution for Hotspot 1 inserted into a uniform macrocell network using different cell selection techniques

The additional users in the network from the hotspot increase the overall network loading, so the EDR distribution is shifted downwards compared with the uniformly loaded case. The difference in EDR performance between the two cell selection techniques is minimal, despite the loading changes in Figure 12. The user data used to form this distribution is concentrated in the centre of the network, which includes the hotspot, and thus excludes the significant contribution from the users at the edge of the network. The use of Best EDR cell selection does not make the EDR distribution worse for this case, despite spreading the load across adjacent cells.

This ability to shift load in non-uniform loading cases is illustrated in the final case shown in Figure 14, where data from a real deployment in part of Vodafone's network has been analysed. Traffic statistics from part of the area have been captured and used to create a realistic user distribution with associated signal levels from the base stations within the area. A theoretical distribution of SNIRs and EDRs for the area was calculated and the resulting data rate distribution is shown in Figure 14, using both cell selection techniques, and this is compared with the equivalent performance from the regular hexagonal network.

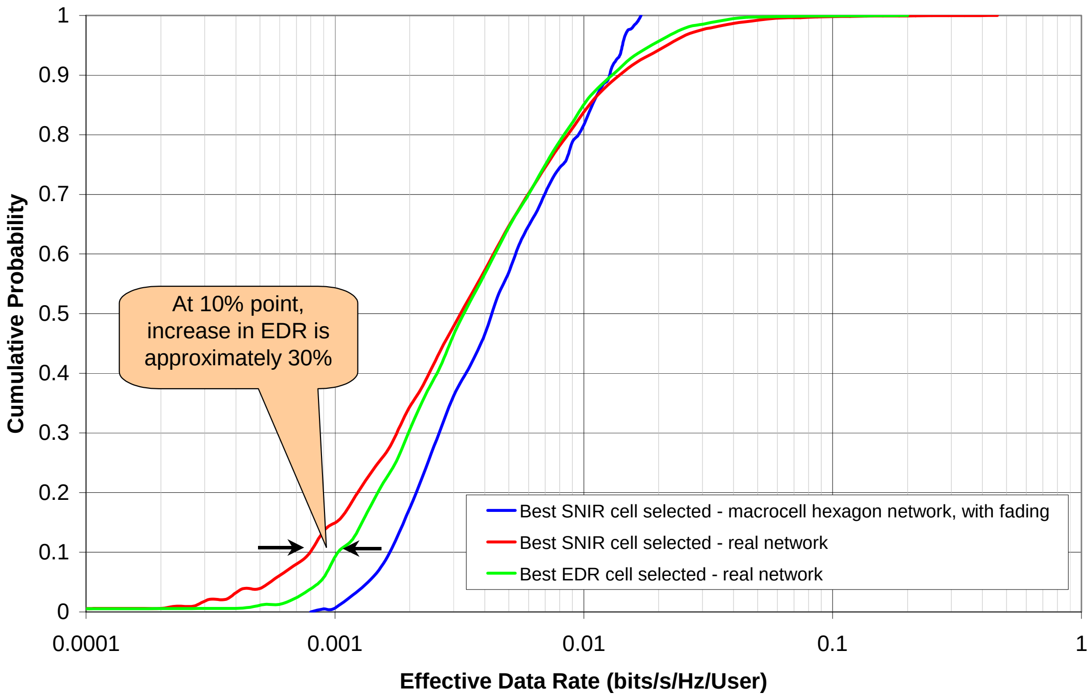

Figure 14 EDR distribution for a realistic user distribution in part of Vodafone's network using different cell selection methods

With the unevenly loaded network based on the real network data, using the Best EDR cell selection method improves the data rates in the lower percentiles of the distribution over those achieved with the Best SNIR cell selection method. EDR improvements of the order of 30% are experienced by the poorest 10% of users. At the upper end of the distribution there is some slight reduction in data rates due to the higher cell loading experienced on some of the cells, although given their high data rates, this is not considered to be a significant issue.

4.4. Frequency Reuse Adaptation

OFDMA-based networks, such as WiMAX, have the ability to support multiple frequency reuse schemes within the same network. This is similar to the fractional reuse concept employed in second generation systems (eg, GSM). However, OFDMA networks provide far more flexibility to allocate different proportions of the radio resources to the different reuse layers on a dynamic basis.

As part of this project we have assessed the potential performance improvement of using an adaptive frequency reuse scheme. By adaptive frequency reuse, we mean that a UE can select, at any given time, the best reuse layer based on its own interference environment. In our investigations, this adaptive reuse scheme has been compared with networks where only a single fixed reuse scheme is in use for all UEs, regardless of their individual circumstances. Adaptability is desirable because when the UE is close to its serving base station it will suffer very little interference from the neighbouring base stations and, as a result, it will be able to use radio resources that are also in use at neighbouring base stations, and thus these resources are utilised to their maximum efficiency. In other words, a mobile close to a base station will be able to use the single-cell frequency reuse layer or the 'N=1' layer. However, as the mobile moves toward the cell boundary it will start to experience interference from the neighbouring cells on the N=1 layer and it must switch to another reuse layer (eg, a three-cell reuse or N=3 layer) where the radio resources are not reused by the immediately neighbouring cells. Thus the appropriate reuse layer for maximum overall system spectral efficiency is a function of the temporary interference environment of each and every mobile station. In our investigation we have considered both single-cell reuse (N=1) and three-cell reuse (N=3), for both fixed (either all N=1 or all N=3) and adaptive (every mobile dynamically switches between N=1 and N=3) reuse schemes.

Typical fixed N=1 and N=3 reuse schemes are illustrated in Figure 15 for the 4G downlink, whereby the colours represent groups of OFDMA frequency tones in use in each sector. For N=1 reuse, all tones are used in all sectors, whereas in N=3 reuse with three-sector base stations, the frequency band is separated into three different groups, which are typically employed as shown (coloured 'red', 'green' and 'blue'). A UE near the cell edge in an N=3 reuse scheme is likely to be able to achieve a downlink signal-to-interference ratio (SIR) much higher than in an N=1 reuse scheme, and so will be able to achieve a significantly higher bit rate on those OFDM tones which it is allocated. On the other hand, for a given size of spectrum allocation, it will only be allocated a third as many tones compared to if it were using a N=1 reuse, since the overall pool of tones has been divided into the three groups. Therefore, whether it receives a lower or higher overall throughput on N=3 reuse compared to N=1 reuse is different for each and every individual UE. This is where the benefit of adaptive reuse comes in, since each UE is allowed to choose whether it would achieve a better overall spectral efficiency using the N=1 or N=3 portion of the band.

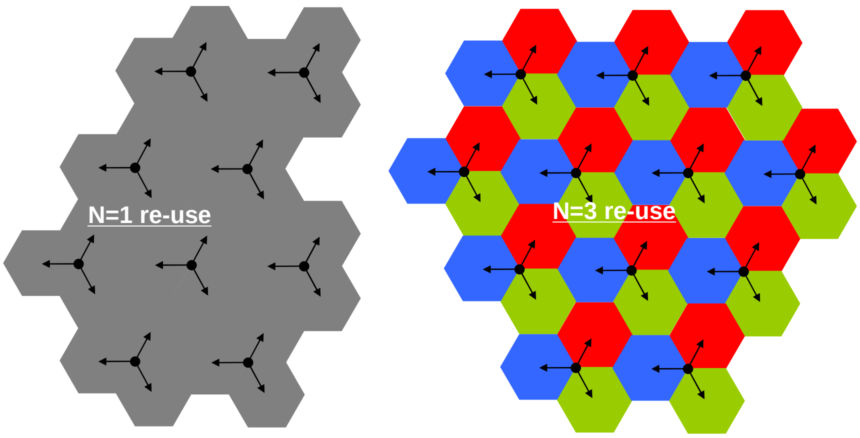

Figure 15 Typical N=1 and N=3 reuse patterns for cellular networks

Figure 16 shows the key results of our investigations in terms of cumulative throughput distributions for N=1 and N=3 reuse, and for hybrid reuse (ie, where each UE can choose whether to be on N=1 or N=3). Results are presented both with and without 10 dB lognormal shadow fading. The key results for the case with shadow fading are summarised in Table 2 which shows that the worst 10% of the users in the network will experience a throughput improvement of 18% and 23% as a result of the introduction of the adaptive frequency reuse scheme, compared with networks employing (fixed) N=1 or N=3 reuse, respectively. This is a significant benefit, which justifies the introduction of such an adaptive reuse scheme for any system that has the signalling capabilities and flexibility to support this form of adaptation.

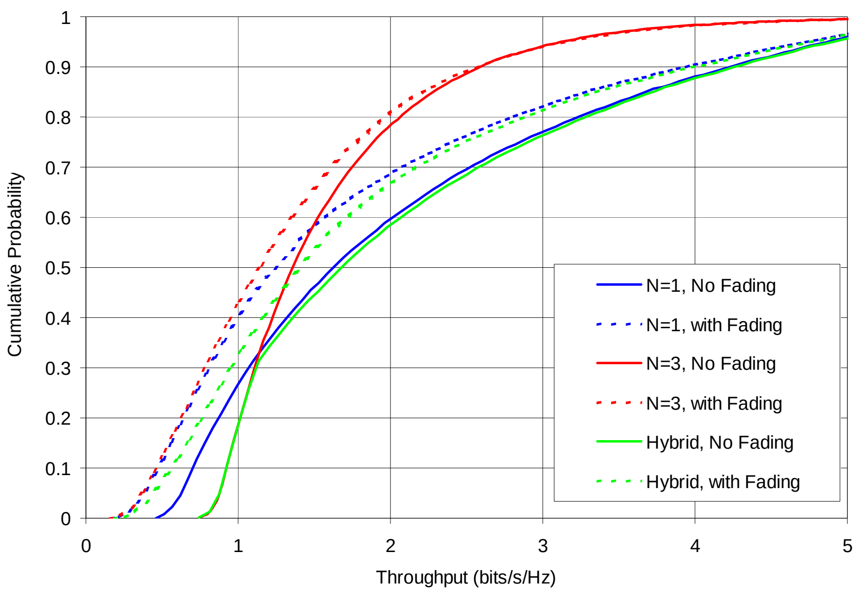

Figure 16 Cumulative distribution function of throughput for N=1 and N=3 reuse, along with a hybrid reuse scheme

Percentile of Interest	Throughput Improvement (%)	
	Over N=1	Over N=3
10%	18	23
50%	13	23
90%	1	55

Table 2 Improvement in throughput offered by an adaptive frequency reuse scheme in WiMAX networks for the 10dB lognormal case

4.5. Summary

In this section we have seen that preamble power optimisation has the ability to mould the capacity available within the network to the offered load by moving capacity from more lightly loaded cells to more heavily loaded cells with significant performance benefits. Depending on the hotspot scenario chosen, we have seen capacity increases of 80% over a network with uniform preamble powers. We have also shown that the EDR-based cell selection algorithm has the ability to improve the data rates experienced by the poorest 10% of users within the network by 30%, without significantly degrading the data rates experienced by other users within the network. This is an important achievement and it shows that the approach has the ability to tailor the network resources to better deliver the key performance targets of the network. Finally, we have shown that by exploiting adaptive frequency reuse in the downlink of 4G systems, we can achieve throughput improvements of between 18% and 23% for the 10% worst-off users, compared to fixed frequency reuse.

Based on this work we believe that preamble power optimisation, the EDR-based cell selection technique and the adaptive frequency reuse feature are complementary and the performance benefits of implementing all technologies will be greater than the performance benefits achieved through implementing only one or two of them.

5. Investigation of Femtocell Deployment Issues

In both 3G and 4G systems, one area of current interest is the use of low power, self-deployed base stations for domestic and enterprise environments. These are sometimes known as ‘Femtocells’, and in the 3GPP standards organisation as ‘Home e-Node Bs’. There is considerable interest in exploring the ways in which these femtocells should be operated and the benefit that can be achieved from using them. From an optimisation perspective, the unplanned nature of the deployment poses challenges to network operators running an existing macrocellular network, where the same radio spectrum is used both by the macrocells and the femtocells. To understand these issues more fully, analysis of a representative deployment of femtocells in a macrocellular 4G network has been simulated.

The simulation model used is based on a regular hexagonal grid of three-sectored base stations, into which a set of femtocells is introduced randomly onto a grid of locations spanning a complete three-sectored base station region (see Figure 7). Statistics are acquired within this region, thus reflecting the full range of locations where a user could be present. Analysis of the impact of femtocells is concerned with the SNIR and the EDR seen by users placed in the femtocell region. The simulation model assumes a conversion from SNIR to data rate used in the 3GPP standardisations activities, and the splitting of spectral resource evenly across the users connected to a server.

Representative single slope median path loss models were used to determine the received signal levels.

5.1. Open and Closed Deployment Modes

There are two key ways in which femtocells can be configured to operate when introduced into an existing macrocellular network. In an ‘Open Access’ configuration, users who are connected to the macrocells can handover to the femtocells seamlessly as if the femtocells were operating in the same way as macrocellular base stations. For femtocells, which are using the same spectrum as the macrocells, their presence increases the interference seen by a user in addition to being a potential server for that user. In a ‘Closed Access’ configuration, only those users that are registered with the femtocell can gain network access through the device. Handover to and from the macrocellular network would still be possible for this restricted set of users, but to other users, the femtocells would appear as additional sources of interference, with the potential to damage their received data rates.

The effective downlink data rates achieved for both configurations are shown in Figure 17, where the EDRs have been categorised according to whether the user is being served by a macrocell or by a femtocell. The EDR distribution for the same user distribution in a macrocell-only deployment scenario is included for comparative purposes. A user is connected to the server that offers the highest SNIR, as is the case in conventional cellular systems.

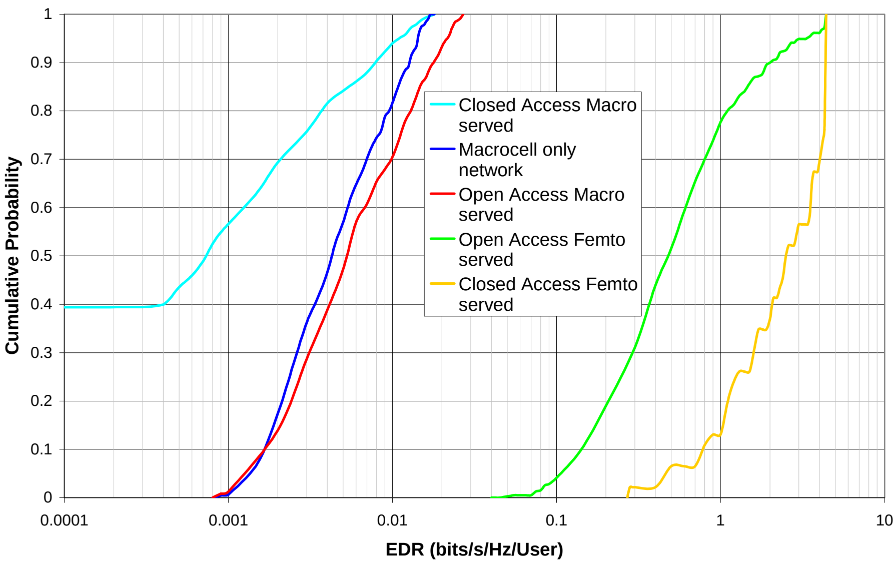

Figure 17 Downlink EDR distributions for Open and Closed Access configurations, with servers selected by the best SNIR

There is an active femtocell density of 20% of households, which translates into around 800 femtocells randomly located within the test region (of size 1732 by 500 metres). A realistic value of 10 dB for the standard deviation of the log normal shadow fading is included in the calculation of received signal powers. A base station power of 43 dBm and a femtocell power of 24 dBm are assumed. The balance between the numbers of macrocell and femtocell served users is approximately 40%:60% (macrocell:femtocell) in the open access case. In the closed access case this becomes 90%:10% (macrocell:femtocell) as fewer users can be handed off to femtocells.

The graph shows that the introduction of open access femtocells into a macrocellular network improves the EDRs for those users who are served by the femtocells and also for those that remain served by the macrocells, even though the femtocells raise the level of interference and so the bulk of users see a reduced SNIR. In terms of EDR, the users served by the femtocells benefit by having a base station almost to themselves and the users served by the macrocells see a reduced loading on their macrocells, which results in them getting a greater proportion of the available time slots, despite a somewhat reduced SNIR. On balance, the loss in SNIR is outweighed by the decrease in macrocell loading and so the data rates that the users experience increase. Closed access operation, by contrast, is excellent for the users served by the femtocells who experience the high data rates associated with having a femtocell in close proximity. However, the users served by the macrocells see poorer data rates due to the significant increase in interference from the femtocells to which they cannot connect. It can be seen that for around 40% of macrocell served users, the interference level is so high that the resulting data rate for those users drops to zero. These results demonstrate that closed access operation of femtocells requires either judicious control of femtocell transmit power, or a separate frequency band for femtocells to operate in. In both cases this reduces the interference seen by the macrocell served users. Open access, by contrast, offers substantial improvements in data rates for all users.

Figure 18 shows the impact of the femtocell transmit power (in closed access femtocell operation) on the proportion of macrocell served users who see too much interference to receive any data, and the proportion of total network users who are served by the femtocells. In order to keep the macrocell users without the ability to receive data down to a figure of around 2%, the femtocell transmit power needs to be reduced to around 3 dBm from the value of 24 dBm, which has been used as the default to generate results in the previous simulations. This is a very low power and may be insufficient to provide reliable coverage within a typical house.

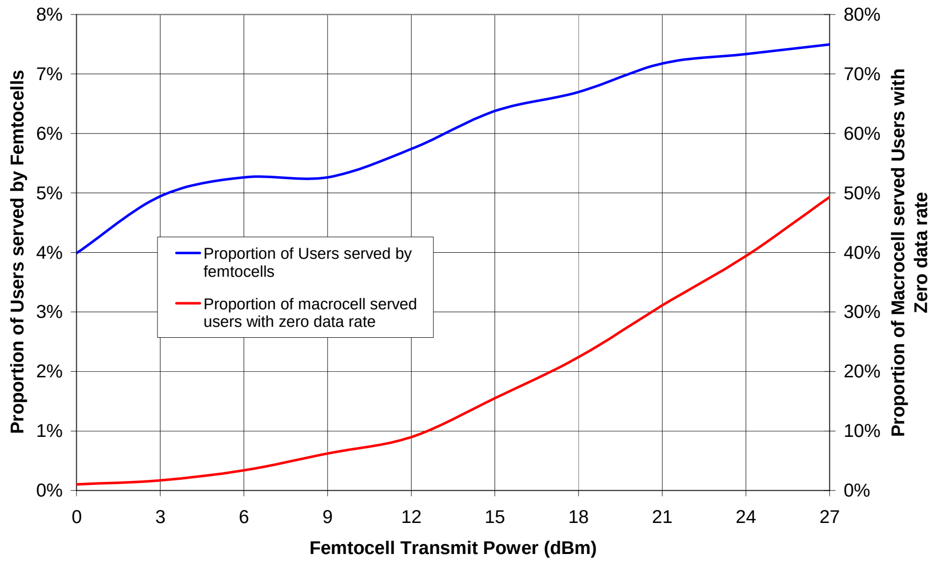

Figure 18 Performance characteristics for Closed Access femtocell operation as femtocell transmit power is varied

In real networks, the proportion of femtocells available within a given area would be expected to grow from 0% up to perhaps 50% of households as the service is marketed. Simulations have shown that femtocell penetration rates of around 10% would be required to off-load half of the users from macrocells to femtocells within an area, thus giving substantial capacity benefits at a relatively early stage of femtocell deployment.

We note that this work on femtocells has formed the basis of a joint standards submission (between Nortel and Vodafone) to the 3GPP RAN group exploring femtocells [5].

5.2. EDR-Based Cell Selection

The open access operation of femtocells benefits both users who are handed off to the femtocells and those that remain on the macrocells because of their reduced loading. This suggests that a cell selection mechanism that preferentially hands over users to the femtocells could potentially further enhance network capacity.

One way to do this is for users to select the server offering the best EDR from the set of servers they can see, rather than simply the one that offers them the best SNIR as is done conventionally. Figure 19 shows a comparison of the downlink EDR distribution using the Best SNIR and Best EDR algorithms for cell selection with the users classified by server type.

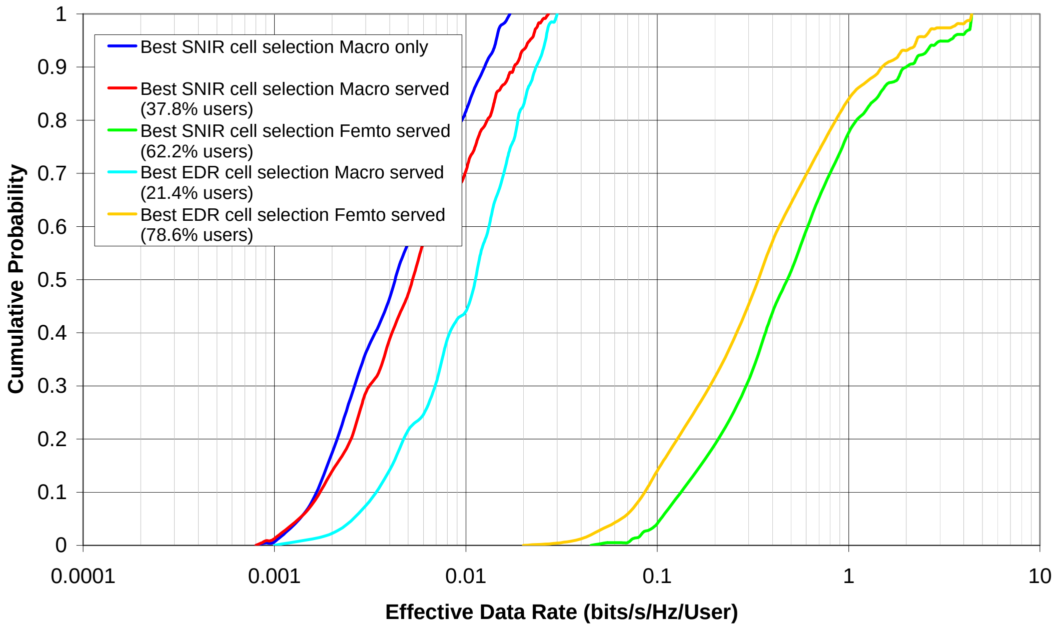

Figure 19 EDR distribution for users in a mixed macrocell and femtocell environment using different cell selection methods with the distributions categorised by server

Figure 20, by contrast, shows the overall EDR distribution of all the users. At the upper end of the distributions, the additional loading on the femtocells resulting from the use of the Best EDR cell selection reduces the very highest throughputs by a small degree. However, this becomes less of an issue when the active femtocell density in the environment is high and most users are served by femtocells even using Best SNIR cell selection.

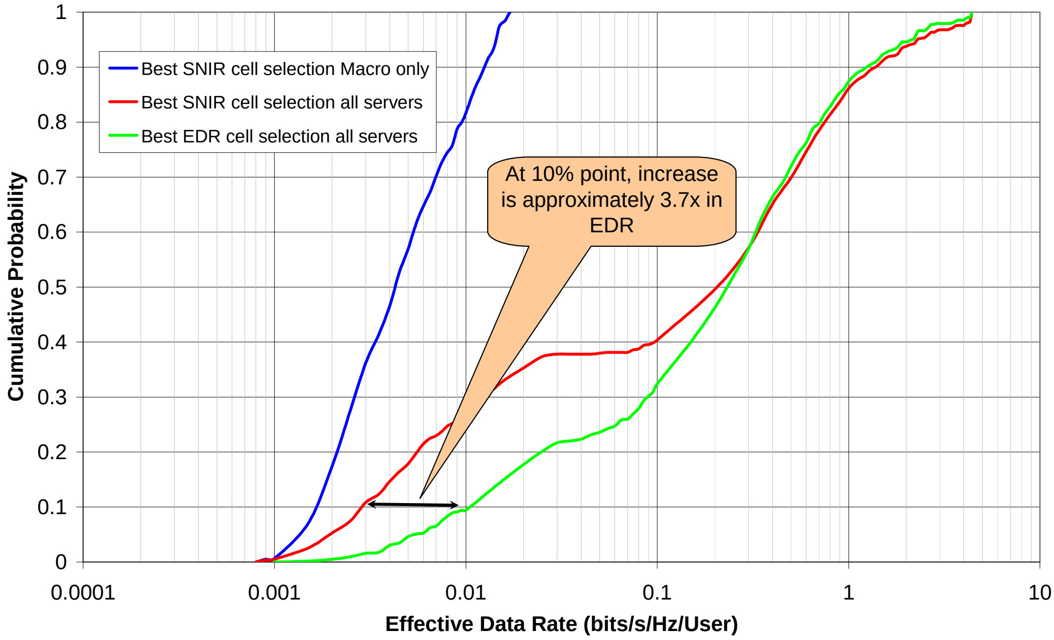

Figure 20 Downlink EDR distribution for users in a mixed macrocell and femtocell scenario using different cell selection methods across all users

Figure 20 also indicates that it is the users with the poorer data rates that see the greatest benefit from the use of the Best EDR cell selection method. Users at the lower end of the distribution are served by the macrocells, and the Best EDR cell selection method shifts a number of these users to the femtocells, leaving those that remain with a greater proportion of the available resource, and hence higher data rates. Improvements in data rates of the order of 3.7 times are seen by the poorest 10% of users.

5.3. Summary

Thus, in summary, the key points from the investigation of femtocell deployment issues are:

- Introducing femtocells into a macrocellular network can provide significant improvements to the network capacity and user data rates. Data rates for users attached to femtocells are two to three orders of magnitude higher than in a macrocell only environment and macrocell served users also benefit due to the reduced macrocell loading.
- Use of femtocells reduces macrocell loading and only a relatively small proportion (around 10%) of femtocells are sufficient to off-load half of the macrocellular users, provided an open access femtocell deployment strategy is used.
- Closed access femtocell operation has more limited benefits and requires careful control of the femtocell transmit power level to prevent degradation of the macrocellular network performance.
- Use of an EDR-based cell selection technique in preference to the more conventional SNIR-based algorithms is a very useful mechanism in open

access femtocell networks to promote users onto femtocells. This is particularly beneficial to users at the lowest part of the EDR distribution, with improvements in data rates of the order of 3.7 times being seen by the poorest 10% of users (over the use of conventional SNIR-based cell selection).

6. Conclusions

This project has demonstrated and quantified the benefits of optimising the ‘soft’ parameters within 3G and 4G networks. The work has shown that capacity gains of up to 60% can be achieved through pilot power optimisation in 3G networks, particularly when large traffic hotspots are present. The work on 3G networks has also shown that improvements in the call admission control process could lead to 18% capacity gains and the rapid adaptation of soft handover parameters based upon mobile speed has the potential to improve the network capacity by around 20%.

In the case of 4G networks, the work concentrated on the optimisation of preamble power, adaptive frequency reuse and EDR-based cell selection. Our results for preamble power optimisation demonstrated that capacity gains of up to 80% are achievable if the optimum preamble power setting can be found. This essentially demonstrates the ability of the network to mould its available capacity to the offered load.

We have shown that by exploiting adaptive frequency reuse in 4G systems we can achieve throughput improvements of 18-23% for the 10% worst-off users compared to fixed frequency reuse.

The EDR-based cell selection technique allows the data rates to be improved for the most disadvantaged users within the network. For mixed femtocell and macrocell deployments, the data rates enjoyed by the poorest 10% of users can increase by almost 4 times, without any significant degradation in the data rates experienced by the other users within the network. For macrocell-only deployments, the improvements in data rate are more modest with only a 30% improvement for the poorest 10% of users, but this is still a very significant performance benefit in terms of meeting the key performance targets set for the network.

Femtocell deployment was also investigated as part of the project and, in particular, the characteristics of open and closed access operation. Our work showed that, when operated in open access mode, femtocells have the potential to dramatically decrease the loading on a macrocellular network, thereby providing significant performance benefits to both the users who continue to be served by the macrocells and also those users served by the femtocells. In closed access mode, we found that although femtocells can provide significant performance benefits to users who are served by these devices, users who remain served by the macrocellular base stations will suffer a heavily degraded quality of service because of the interference caused by the femtocells.

The project has also demonstrated the feasibility of building a simulator to investigate the performance of a 4G network and it has also shown that such a simulator can be distributed across a large-scale computing cluster to improve simulation speeds. During the project we have also shown how network simulations can be incorporated as part of the optimisation process and several optimisation algorithms have been tested.

7. Exploitation of Project Results

In terms of the exploitation of the project results, there are a number of possible applications of the innovations created during the project. In the short term, the knowledge and understanding of 3G and 4G systems will be exploited to provide the telecommunications industry with expert guidance on the best ways to design and optimise these systems. This will be achieved through the provision of training courses and expert consulting services based on the knowledge developed during the project. The detailed system simulator developed as part of this project, and the associated computing cluster, will also be used to support these services by allowing the consortium members to provide advice on the most appropriate parameter settings for a given network. The results of the project will also be used to influence the design of 4G and subsequent equipment through the dissemination of the project results within the various standardisation fora, including 3GPP and the Next Generation Mobile Networks (NGMN) initiative.

Also, the work on femtocells will help to inform the industry on the relative merits of open and closed access deployments. Open access femtocells have the potential to dramatically improve the performance of mobile networks, whilst also leading to a significant reduction in the number of macrocellular base stations required. Following on from this project, the different project partners will use this work as a sound platform for further investigations in this area, leading to more advanced products than those available today that are capable of self-configuring and managing their own interference. The investigations into closed access femtocells and the associated interference issues, the results of which have been disseminated through 3GPP, should better inform the industry of the pitfalls associated with this mode of operation and prevent costly mistakes. In the medium and longer term, the innovative cell selection and channel assignment algorithms developed as part of this project will be developed further and potentially incorporated into future equipment to improve the performance of 4G networks. Using advanced simulations in the optimisation of mobile communications networks has the potential to dramatically improve the performance of these systems, which will become increasingly important as network traffic loads continue to increase with the introduction of higher data rate services. During this project, many aspects of using simulators in the optimisation loop have been addressed and this will form a good basis for stimulating further interest and research in this area.

References

1 Todd Tannenbaum, Derek Wright, Karen Miller and Miron Livny, ‘Condor - A Distributed Job Scheduler’, in *Beowulf Cluster Computing with Linux*, Thomas Sterling [Editor], The MIT Press, 2002, ISBN: 0-262-69274-0

2 The Condor Project, ‘Condor High Throughput Computing’, http://www.cs.wisc.edu/condor/

3 Multiple Access Communications Limited, ‘A Survey of Optimisation Techniques’, Project MONOTAS Document No. 020.PUB.2, 8 March 2006, available from www.macltd.com/monotas.

4 Chun-Hung Chen, Jianwu Lin, Enver Yücesan and Stephen E. Chick, ‘Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimisation’, Journal of Discrete Event Dynamic Systems: Theory and Application, Vol 10, No 3, pp 251-270, July 2000.

5 Vodafone and Nortel, ‘Open and Closed Access for Home Node Bs’, submission to 3GPP TSG-RAN WG4 Meeting 44, Athens, Greece, 20-24th August 2007, Document No: R4-071231.