Traffic Dimensioning for Multimedia Wireless Networks

by
Leila Zurba Ribeiro

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Electrical Engineering

Committee:
Dr. Luiz A. DaSilva (Chairman)
Dr. Jeremy E. Allnutt
Dr. Ing-Ray Chen
Dr. Scott F. Midkiff
Dr. Jeffrey H. Reed

April 17, 2003
Falls Church, Virginia

Keywords:
Demand Characterization, Traffic Simulation, Third-Generation, Mobile Multimedia Traffic, Wireless Network Optimization, Wireless Network Dimensioning.

Traffic Dimensioning for Multimedia Wireless Networks

Leila Zurba Ribeiro

(ABSTRACT)

Wireless operators adopting third-generation (3G) technologies and those migrating from second-generation (2G) to 3G face a number of challenges related to traffic modeling, demand characterization, and performance analysis, which are key elements in the processes of designing, dimensioning and optimizing their network infrastructure.

Traditional traffic modeling assumptions used for circuit-switched voice traffic no longer hold true with the convergence of voice and data over packet-switched infrastructures. Self-similar models need to be explored to appropriately account for the burstiness that packet traffic is expected to exhibit in all time scales. The task of demand characterization must include an accurate description of the multiple user profiles and service classes the network is expected to support, with their distinct geographical distributions, as well as forecasts of how the market should evolve over near and medium terms. The appropriate assessment of the quality of service becomes a more complex issue as new metrics and more intricate dependencies have to be considered when providing a varying range of services and applications that include voice, real-time, and non-real time data. All those points have to be considered by the operator to obtain a proper dimensioning, resource allocation, and rollout plan for system deployment. Additionally, any practical optimization strategy has to rely on accurate estimates of expected demand and growth in demand.

In this research, we propose a practical framework to characterize the traffic offered to multimedia wireless systems that allows proper dimensioning and optimization of the system for a particular demand scenario. The framework proposed includes a methodology to quantitatively and qualitatively describe the traffic offered to multimedia wireless systems, solutions to model that traffic as practical inputs for simulation analysis, and investigation of demand-sensitive techniques for system dimensioning and performance optimization.

We consider both theoretical and practical aspects related to the dimensioning of hybrid traffic (voice and data) for mobile wireless networks. We start by discussing wireless systems and traffic theory, with characterization of the main metrics and models that describe the users' voice and data demand, presenting a review of the most recent developments in the area. The concept of service class is used to specify parameters that depend on the application type, performance requirements and traffic characteristics for a given service. Then we present the concept of "user profile," which ties together a given combination of service class, propagation environment and terminal type. Next, we

propose a practical approach to explore the dynamics of user geographical distribution in creating multi-service, multi-class traffic layers that serve as input for network traffic simulation algorithms. The concept of quality-of-service (QoS) is also discussed, focusing on the physical layer for 3G systems. We explore system simulation as a way to dimension a system given its traffic demand characterization. In that context, we propose techniques to translate geographical distributions of user profiles into the actual number of active users of each layer, which is the key parameter to be used as input in simulations.

System level simulations are executed for UMTS systems, with the purpose of validating the methodology proposed here.

We complete the proposed framework by applying all elements together in the process of dimensioning and optimization of 3G wireless networks using the demand characterization for the system as input. We investigate the effects of modifying some elements in the system configuration such as network topology, radio-frequency (RF) configuration, and radio resource management (RRM) parameters, using strategies that are sensitive to traffic geographical distribution.

Case study simulations are performed for Universal Mobile Telecommunications System (UMTS) networks, and multiple system variables (such as antenna tilts, pilot powers, and RRM parameters) are optimized using traffic sensitive strategies, which result in significant improvements in the overall system capacity and performance. Results obtained in the case studies, allied to a generic discussion of the trade-offs involved in the proposed framework, demonstrate the close dependence between the processes of system dimensioning and optimization with the accurate modeling of traffic demand offered to the system.

Acknowledgements

First and above all, I would like to thank God for the strength, health, hope and opportunity He has provided to me, which allowed me to get this dissertation started and completed.

Next, I would like to thank my advisor, Dr. Luiz DaSilva, for his wisdom, knowledge, patience, dedication, and, most of all, for making this entire experience feel as a teamwork with a reliable friend, rather than a solitary journey.

I would also like to thank my committee members, Dr. Jeremy E. Allnutt, Dr. Ing-Ray Chen, Dr. Scott F. Midkiff and Dr. Jeffrey H. Reed, for their invaluable feedback and advice.

My thanks go also to CelPlan Technologies, Inc., my employer during the whole period of this research, for tuition reimbursement provided during most of the period, as well as, constant support and understanding; and to my direct supervisor, Mr. Leonhard Korowajczuk, for the long hours of brainstorming and challenging discussions.

I also wish to express my gratitude to the National Science Foundation (NSF), for partially funding this research through the Integrated Graduate Education and Research Training (IGERT) program, and to all members of the Integrated Research and Education in Advanced Networking (IREAN) program, which through an intense exchange of information and insights, has helped me develop a broader, multidisciplinary view of the many aspects that surrounded my research subject.

I would also like to warmly thank all my friends and colleagues at the Alexandria Research Institute (ARI), for making me feel so welcome and creating such a motivating and friendly environment to work at. My thanks go also to the staff of the ARI and the ECE department, especially to Latricia Nell and Michelle Clark, for their support in many occasions.

I also thank my parents who have taught me the importance of reading and learning from the beginning, and who, despite the distance, together with my sisters and brother are always present and participating.

Finally, I would like to thank my husband Aluisio, for his continuous encouraging, understanding, unfailing support, and loving care throughout all this period, and to our yet unborn baby, who has for the past seven months accompanied us on this experience, and to whom we have already grown so attached and grateful.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Mobile Wireless Evolution

Wireless mobile communications have evolved from the original analog systems in the early 1980's to what was called the Second-Generation" (2G) in the early 1990's. This second generation was characterized mainly by the migration from analog to digital technologies, including those adopting Time-Division-Multiple-Access (TDMA) such as Global System for Mobile communication (GSM) and Interim Standard (IS) 136 and systems adopting Code-Division-Multiple-Access (CDMA) technology such as IS-95.

These changes greatly improved the spectral efficiency of cellular systems, allowing operators to support more users with the same spectrum, but they did little more from the user's point of view than reducing costs and extending battery life, which by themselves were significant benefits. Until these days, the main use of mobile telephony is voice, with occasional data connections that have access to little bandwidth and very low reliability due to the severe limitations imposed by the wireless environment.

The evolution towards Third-Generation (3G) systems addresses the increasing user needs for mobile data access, mainly due to the growth of Internet use, with mobile subscribers becoming more and more demanding of a system that supports services such as wireless e-mail, web-browsing, wireless telecommuting, and streaming video, among others. This evolution includes a migration path through 2.5G systems, such as the General Packet Radio System (GPRS), which introduces support for variable bit rates and the creation of an alternate path for packet switched data in the core network. Third-generation systems will expand the concept of 2.5G by supporting larger bandwidth and quality of service (QoS) provisioning.

Global harmonization efforts are being led by the Third-Generation Partnership Project (3GPP) that standardizes Wideband-CDMA (WCDMA), Enhanced Data rates for GSM Evolution (EDGE) and Time Division-Synchronous CDMA (TD-SCDMA) technologies as the 3G evolution from current GSM and TDMA systems. Additionally, 3GPP2 was created to handle the standardization process of CDMA2000 systems as the 3G evolution path from existing IS-95 systems.

Several wireless networks have started operating 3G systems already, mostly using CDMA2000 such as 1xRTT in the United States, China, and other countries in Asia, and following the GSM-GPRS-UMTS evolution. Additionally, despite the slow down in the investments for new 3G networks

and the spectrum licensing process particularly in the US, many new systems are being planned for implementation in the near future.

1.2. Scope of Work, Motivation and Contribution

In this research we consider the theoretical and practical aspects related to the dimensioning of mobile wireless networks, such as the third-generation systems, to support hybrid traffic mixes (voice and broadband data).

The process of dimensioning a network system is defined here as the appropriate allocation of resources that will satisfy certain performance requirements for the system given the knowledge of the traffic demand expected to be offered to the system. Wireless systems migrating from 2G to 3G technologies, in order to fully support heterogeneous traffic, face a number of new challenges with regards to the design and dimensioning of network infrastructure. These challenges can be summarized in three distinct areas, as follows:

- *Traffic characterization*: Basic, well established assumptions in modeling circuit-switched voice traffic no longer hold true with the convergence of voice and data over a packet-switched infrastructure. Self-similar models may have to be adopted to appropriately take into account the burstiness that packet traffic is expected to exhibit in all time scales.
- *Demand characterization:* For an accurate assessment of resource requirements, one must consider the distinct demand profiles of residential users, business users, and road traffic, as well as forecasts of how the market for mobile services will evolve over the near and medium terms.
- *Achieving QoS in 3G systems*: QoS requirements at the application layer must be mapped into parameters that can be controlled at the physical (spreading factors, power) and/or data link layers (medium access control mechanisms).

The main objective of this work is the proposal of a practical methodology to characterize the traffic offered to multimedia wireless systems in a way that allows proper dimensioning and optimization of the system for that particular demand. The **motivation** for this research is demonstrated through the main benefits that it is expected to provide:

- Proper dimensioning and therefore a more reliable business case analysis for operators planning to migrate to 3G technologies.
- The ability for the operator to design a proper rollout plan for system deployment.

- Efficient resource allocation based on expected demand and growth.
- Demand-sensitive system optimization.
- Customization or selection of resource management algorithms.
- Traffic shaping for capacity maximization.

The main **contribution** of this work is the proposal and realization of a framework that includes a methodology to describe quantitatively and qualitatively the traffic offered to a multimedia wireless system, solutions to model that traffic as practical inputs for simulation analysis and investigation of demand-sensitive techniques for system dimensioning and performance optimization.

The **main steps** explored to achieve these objectives included:

- Review of traffic theory background and recent developments on the fields of voice and data traffic and wireless systems dimensioning.
- Development of a framework for demand characterization that includes the creation of multi-service, multi-class traffic layers associated to user profiles, to describe quantitatively and qualitatively the traffic offered to a multimedia wireless system.
- Proposal of traffic modeling techniques that allow proper representation of the traffic offered to the system while focusing on reducing simulation complexity.
- Detailed analysis of selected system dimensioning scenarios to validate the proposed methodology.
- Investigation of methods of optimizing system performance for a given demand characterization scenario.

1.3. The Mobile Environment

The mobile environment presents some specific challenges to network design. One of the most important is the fact that the spectrum used for any wireless system is limited and it has to be reused in a cellular architecture to satisfy capacity requirements. In systems that use spread-spectrum techniques, this reuse is done in every cell, while in traditional TDMA and analog systems certain reuse patterns are followed, with frequencies being reused in groups of cells as illustrated in Figure 1-1. In both cases, cellular architectures have to deal with one important limiting factor, which is the interference from other cells and users. As more users are added to the system, the signal-to-interference ratio decreases and the signal quality degrades. Techniques, such as closed-loop power

control, dynamic channel allocation, smart antennas and admission control, are used to mitigate the interference effects.

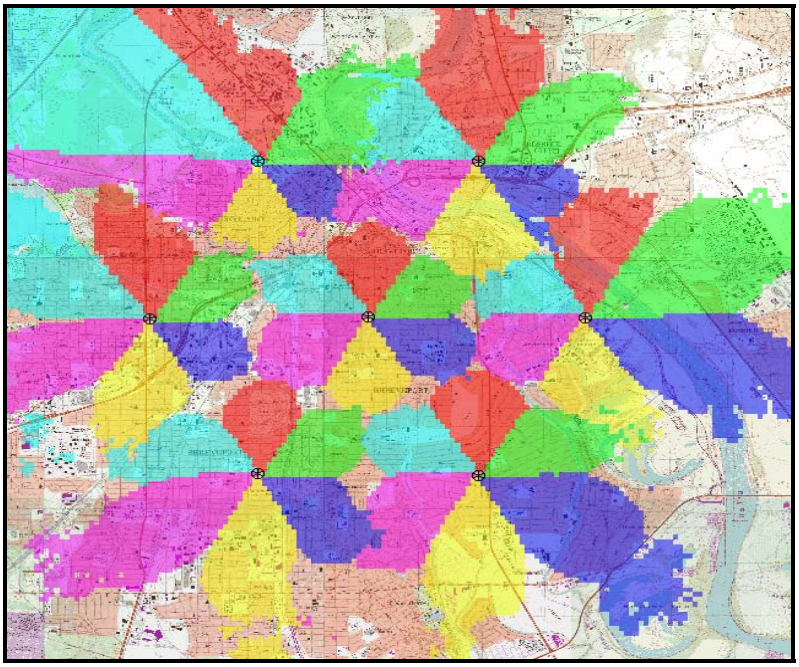


Figure 1-1. Example of cellular architecture with frequency reuse pattern on six-sector cells.

Another important restriction related to mobile systems is the hostile transmission environment due to the ever-changing medium. User movement as well as objects, such as vehicles and people. moving around the user, cause the transmitted signal to experience unpredictable reflection and diffraction effects. Most of the time the signal received by a user consists of many samples of the transmitted signal that arrive with different delays (delay-spread) and with different strengths due to the multiple reflected and diffracted paths they have traveled (multipath). An illustration of typical signal oscillations due to fading is shown in Figure 1-2. Typically, there is no line-of-sight between the transmitter and receiver and the indirect components are the only clue the receiver has about the signal that was transmitted. To cope with this randomly disturbed environment and guarantee a certain confidence level, the link-budget must include link margins. Statistical distributions describe the signal behavior and the most common approaches consider Rayleigh distribution for situations where the indirect components predominate. In circumstances where there is line-of-sight and the direct path predominates, Rician distribution is typically used.

When using fading margins for propagation analysis, the design may compensate in a simplistic way for the impact that multipath has over the signal strength and this will impact the system capacity

calculation, as less bandwidth would be considered available to the user. However, another important effect of multipath must be considered when we talk about data transmission and that is the impact of occasionally deep fades, which cause data to be lost and force retransmission in the case of non-real-time applications. This implies that there is an additional bandwidth trade-off for data transmissions over multipath channels. Forward error correction, which is standard in all 3G technologies, is expected to partially compensate for this problem, at the cost of additional bandwidth overhead. Other actions typically taken in the design to mitigate negative effects of multipath include channel monitoring, link adaptation, incremental redundancy, power control and beamforming smart antennas.

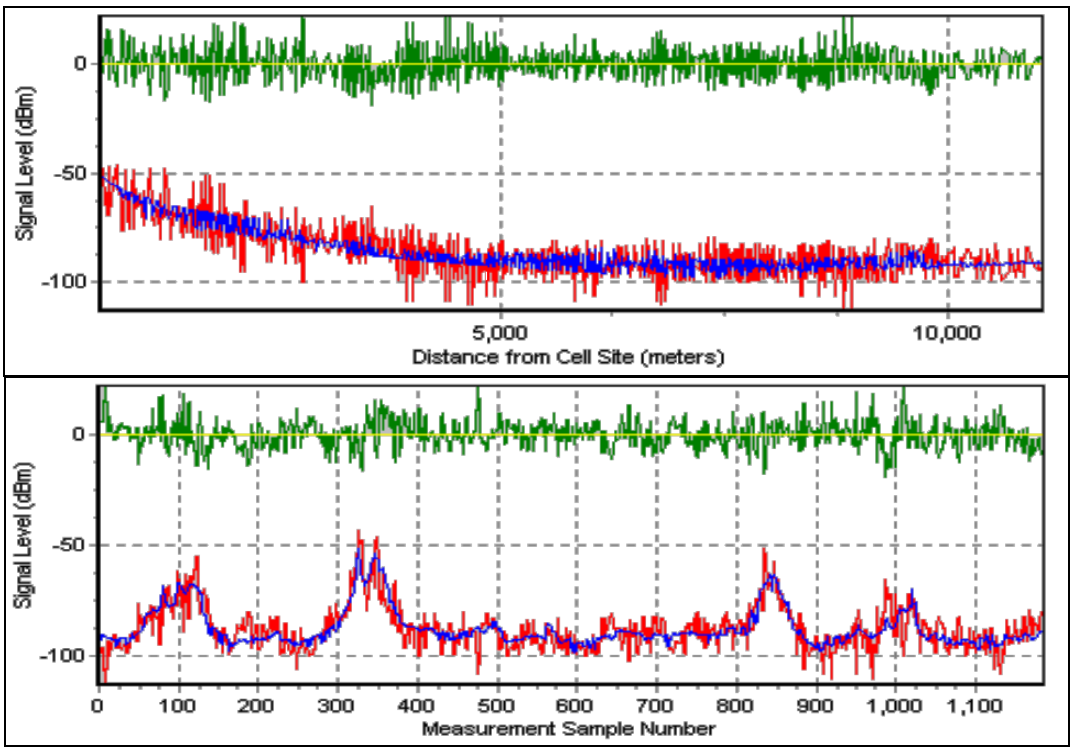


Figure 1-2. Example of measured signal strength from one cell.

Another aspect related to user mobility is that the system needs to be able to support features such as location, registration and authentication. Furthermore, during a call, it also needs to be able to monitor signal quality and transition the service from one cell to another, providing virtually seamless handoff to neighboring cells. This need imposes additional overhead to the system, as overlap between cells must exist to allow soft handoff while maintaining a reasonable signal quality from the user's point of view. Besides, handoff may force the cell to temporarily handle more traffic than that associated to the cell's best server area. This happens because a mobile that originates a call inside

one cell's coverage area may have moved to points inside the best-server area of a neighbor cell, but still retain the call from the original server, due to its configuration of handoff thresholds and hysteresis values. Figure 1-3 illustrates this situation.

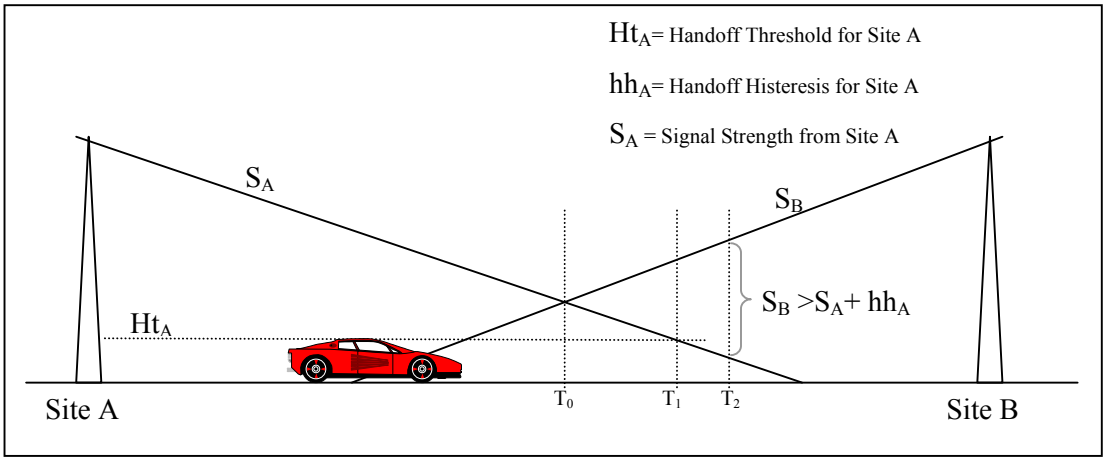


Figure 1-3. Example of a handoff situation when the mobile travels from Site A to Site B: The mobile crosses the best-server border at time T0, but requests handoff only at time T1 (when $S_A<Ht_A$) and the handoff conditions are satisfied only at time T2, when the signal strength from site B is above that from Site A plus the hysteresis (hh_A).

The situation described above illustrates the need for traffic under handoff areas to be accounted for more than once, as a dimensioning redundancy that will improve mobility management. This overhead caused by handoff provisioning is even more explicit in CDMA systems, where soft-handoff is supported. In those cases, the same call may have multiple servers simultaneously. While this means a real overhead in terms of channel utilization, the overlap areas have a vital function to provide continuous service for fast moving users. To consider this effect in its correct proportion the traffic description would need to take into account the expected overlap between cells, as well as some information about the expected mobility of the users, which depends on the location of roads, vehicle traffic intensity and other variables that are difficult to estimate. Figure 1-4 shows an example of service overlap (soft handoff areas) in a CDMA system.

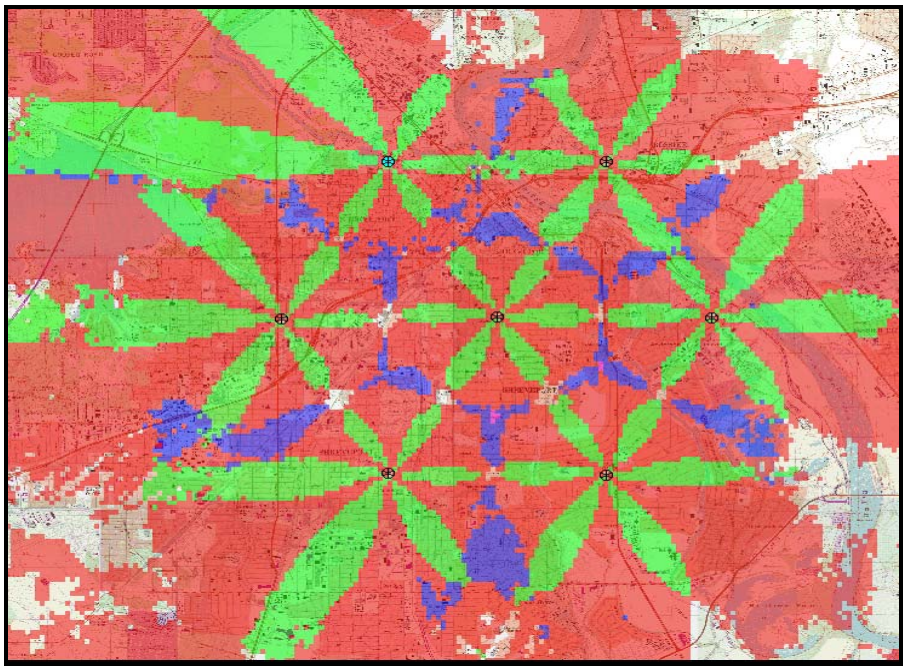


Figure 1-4. Expected soft-handoff areas in a CDMA system.

1.4. Document Overview

This document presents the theoretical and practical aspects related to this research and the proposed framework, as well as results obtained from it.

Chapter 2 presents a brief review of traffic theory, discussing the models available for both voice and data traffic, as well as the most recent and fundamental developments in this area. Some performance metrics are presented and the concept of service class is used to distinguish different types of traffic according to their application and performance requirements for a given service.

Chapter 3 details a practical methodology suggested for demand characterization. We introduce the concept of "user profile," which ties together a given combination of service class, propagation environment and terminal type. Next we present a practical approach to explore the dynamics of user geographical distribution in creating multi-service, multi-class traffic layers that will serve as input for network traffic simulation algorithms.

Chapter 4 discusses the concept of Quality of Service (QoS), with a focus on the physical layer for Third-Generation systems and how these QoS requirements map into higher layers. Mechanisms of admission control and congestion control are also discussed in this section. In this chapter we also explore how system capacity (expressed as total throughput) depends on the user mix, i.e., the proportions of different services that are combined to achieve a certain total traffic demand. This fact,

although rather intuitive, demonstrates the importance of an accurate characterization of different user mixes with their appropriate model parameters, as for the same overall demand different system efficiencies will be achieved with different user mixes.

In Chapter 5, we discuss simulation techniques that may be used to dimension multimedia systems. We start the chapter by presenting the multiple types of system-level simulations, then we select one method (static simulations) and discuss its main limitations, namely the difficulty in correctly estimating the snapshot sizes for simulations. Next, we propose an approach to overcome that challenge, present and discuss results from this approach and address sensitivity and reusability issues. At the end of the chapter, a broader view of the entire system simulation process for wireless systems that support multimedia traffic is presented and discussed.

In Chapter 6, we perform UMTS case study simulations using the proposed framework. We start by discussing the selection of a suitable WCDMA simulator. Multiple simulation scenarios are performed, and results are used to validate the proposed methodology against baseline results of simplified models. The main metrics of system capacity achieved for each scenario are evaluated, and for cross-validation purposes, we compare results to those from a second traffic simulator available in the public domain. Next we evaluate the sensitivity of dimensioning results to main assumptions in traffic modeling and demand characterization and verify the impact of different demand scenarios on the selected capacity metrics.

In Chapter 7, we investigate techniques that could be used to improve capacity, in light of all the previous topics presented and using the proposed framework as means to simulate and verify achievable results, while at the same time demonstrating the importance of demand characterization in this process. The techniques for achieving optimal system performance include modifications in the radio network topology, site configurations and radio resource management parameters. For most of the variables discussed we present example simulations and experiments, whenever found relevant.

Finally, in Chapter 8, we present a summary of results and conclusions obtained from this work and discuss related areas of research.

Chapter 2. Multimedia Traffic

The theory behind Erlang's analysis relies on the assumption of circuit-switched connections and relatively stable call holding times. However, in recent years, the applicability of Erlang theory to dimensioning even the narrowband fixed networks has been challenged, as more and more subscribers are using the telephony network to access the Internet (through dial-up). The dispersion of hold times has increased in magnitude, and networks dimensioned for average holding times of the order of minutes are now confronted with subscribers connected for hours. Due to the circuit-switched nature of these networks, there is a real inefficiency in using such systems for bursty data communications, as the resources allocated to one subscriber cannot be shared with others even if they are not being effectively used. To avoid this type of situation in such a resource limited media as the wireless environment, operators are being urged to offer broadband wireless access, where the resources can be effectively allocated and shared on demand by supporting packet switching.

An appropriate description of the user profile is the first step in the process of modeling the traffic offered to a network and determining the system demand. In this document this is done with focus on system dimensioning, i.e., the qualitative and quantitative estimation of the resources needed to provide service at a certain desired performance, given an estimation of the traffic offered. The most important resources available in the dimensioning process [1] are: (i) bandwidth; (ii) buffers and (iii) time constraints such as transmission delays, computing time, and signaling delays. We may also add to this list the cell site and network equipment itself as for the same geographical area and available spectrum more equipment is needed (namely radio base stations) depending on the traffic load. For wireless systems, it is clear that the most expensive and restrictive of such resources is bandwidth, since spectrum is limited and has to be reused among many cells. Therefore, it is fair to say that the biggest challenge for system dimensioning in wireless multimedia applications is to allocate just enough bandwidth and equipment to satisfy the system requirements at the desired performance.

This chapter describes ways to characterize the broadband traffic source and discusses some traffic models used to correlate the source description with the required network capacity and desired quality of service. We start with a review of voice traffic modeling (Erlang analysis) and then we present some possibilities to model data traffic sources.

2.1. Voice Traffic

For the past decades, the traffic modeling and dimensioning of voice networks has relied on the work developed by Erlang in 1917 [2]. The basic assumptions of Erlang theory consider two important user parameters. Those parameters are the call arrival rate (in calls/time unit) and the average call holding time (in time unit/call). The product of these two parameters is a metric of traffic load, which is a dimensionless quantity with units, however, traditionally referred to as "Erlangs" in the telephony universe. For further insight into the concept of one Erlang, one can think from the point of view of the circuit usage. A traffic load given in Erlangs corresponds to the usage time of a certain resource over a certain time scale. For instance, if the resource is used 80 % of the time, it is said to be carrying 0.8 Erlangs.

To calculate the total traffic offered to one cell, one needs to take into account the forecast (or switch-measured) values for the parameters described above, as an average over the number of subscribers under the cell coverage. Therefore, the following relations hold.

Traffic per subscriber [Erlangs] = Call arrival rate per subscriber * Average holding time per call

Total traffic offered [Erlangs] = Traffic per subscriber * Number of subscribers under service area

The statistical distribution for the call arrivals is assumed to follow a Poisson distribution, with the inter-arrival times following an exponential distribution. The system is assumed to have a finite number of channels and, as a consequence, to support a maximum number of simultaneous calls or connections. Since the maximum supported traffic can be smaller than the traffic offered (which is dependent only on the user parameters), it is possible that the system fails to assign channels when a user attempts to make a call. The probability of this type of failure is referred as *blocking* or Grade of Service (GOS). For a certain traffic offer, A, in Erlangs, the number of simultaneous channels needed by the system, N depends on the Grade of Service to be achieved. For system dimensioning purposes, we use well known mathematical relationships that relate those three variables [2].

Furthermore, there are two ways to define call attempt failure, depending on the system behavior when a call attempt is received and all circuits are busy. In the first case, adopted by most traditional fixed telephony systems and in many cellular systems, such as Advanced Mobile Phone Service (AMPS) and IS-136, the attempt is immediately refused and the user immediately receives an indication of failure and could try again if desired. This type of trunking system is usually referred to as *blocked calls cleared*. It is modeled by the Erlang B formula, which can be obtained from an

M/M/c/c queueing system [3] as the probability that all servers are busy. In the second case, used for instance in GSM, the system puts the user on a queue and allows some waiting time (delay). If the waiting time exceeds a certain pre-defined limit, the call attempt has failed and the user will be free to try again. This mechanism is known as *blocked calls delayed* and the GOS measurement is defined as the probability of the waiting time exceeding the maximum allowed delay. This probability is modeled by the "Erlang C" formula, obtained from an M/M/c queueing system [3]. Both formulas are shown below. They allow direct mathematical evaluation of the GOS parameter as a function of the number of channels and offered traffic, but requires numerical methods for situations when the target GOS and the offered traffic, A, forecast are the independent variables and one needs to calculate the required number, N, of channels to serve the system, which is the typical dimensioning problem. For CDMA systems, the same formulas apply, with the exception that the definition on whether a server is available is based on link quality level, which varies with the system load.

Erlang B Model: Probability[No server available]

$$GOS = \frac{\frac{A^N}{N!}}{\sum_{k=0}^{N} \frac{A^k}{k!}} \qquad \textbf{Equation 2-1}$$

Erlang C Model: Probability[Delay>T]

$$GOS = \frac{A^N}{A^N + N!\left(1 - \frac{A}{N}\right)\sum_{k=0}^{N-1} \frac{A^k}{k!}} e^{-\frac{(N-A)T}{\tau}} \qquad \textbf{Equation 2-2}$$

Where: T = Maximum allowed delay before attempt is considered as failed

τ = Average call holding time

A = Total traffic supported by the system

N = number of channels

GOS = probability of call attempt failure

For system dimensioning purposes, it is usually assumed that the user behavior parameters (call arrival rate and holding time) have normalized standard deviation of small order.

2.2. Data Traffic

The main characteristic of broadband data networks is use of packet switching as opposed to the circuit-switched architecture of voice telephony networks. In packet switched networks we can further distinguish two main groups: connectionless (or datagram-oriented) networks, where every information packet contains enough information to enable their correct delivery, and connection oriented (virtual-circuit switched) networks, which includes an initial negotiation of a virtual path from source to destination (connection setup) followed by data transfer.

2.2.1. Traffic Descriptors

For data communications, some typical parameters used to characterize a traffic source are the peak and sustained (long-term average) data rates as well as the traffic burstiness. These variables are explained below.

- *Average data rate*: Amount of data generated by the source over a certain time interval, usually expressed in bits per second or packets per second. The measured interval is often a large time scale, such as hours, or days.
- *Peak data rate*: Maximum instantaneous data rate generated by one source during a typical time intervals.
- *Burstiness*: Describes how infrequently a source sends traffic. A simple equation to define burstiness is given by [4] as:

$$Burstiness = \frac{PeakDataRate}{AverageDataRate}$$

Equation 2-3

If a source sends data at a constant rate, it is considered as not bursty (Burstiness =1), while it is considered very bursty if it allows long periods of "silence" mixed with periods of transmission close to peak rate (Burstiness >>1).

It should be noticed that the definition above is just a simplistic measure of burstiness. More complex approaches suggest measuring burstiness on multiple time scales [5]. In this case, the burstiness is expressed by its sustainability over different times scales.

The diversity of services offered in broadband connections leads to a big dispersion of burstiness, holding time and peak data rates.

2.2.2. QoS Requirements

For network dimensioning, it is important to obtain models that describe multiple types of data sources, with distinguished behavior characteristics and performance requirements.

As discussed in Section 2.1, the measure of QoS in voice networks is reasonably defined through the use of one single metric, which is the probability of call blocking. For data networks, the definition of QoS involves a larger set of metrics, some more important than others depending on the service class and type of application. The most important data traffic metrics can be described as [1]:

- packet loss rate,
- flow capacity (throughput),
- availability (flow blocking probability),
- delay, and
- delay variability (jitter).

In multimedia service, such as third-generation wireless systems aim to provide, different types of applications have different characteristics and performance requirements. The larger set of possible applications is grouped into four main categories of service classes according to the UMTS definition, as follows [7].

- *Conversational class*: includes voice, video-conference, video-games, etc. The most important performance requirement for this class is consistency in time relations, including both low delay and low jitter requirements. This includes preservation of the source data rate. Data integrity (loss rate) is not as critical.
- *Streaming class*: includes streaming audio and/or video, such as video on demand. It requires preservation of the time relation as to low jitter effect but is not as critical as the conversational class with regards to low delay requirements. Data integrity (loss rate) is not critical.
- *Interactive class*: includes Web-browser, database retrieval, and remote LAN access. An important parameter is the round-trip delay, which characterizes the request-response time. Data integrity (low loss rate) is very important for those types of applications.
- *Background class*: Applications include non real-time background download of e-mails, file transfers, etc. There are no strong restrictions with regards to time relations or delay, but data integrity is critical.

The categories described above could be further grouped in two main sets: real-time traffic, which includes the classes where time and rate variation characteristics must be preserved (conversational and streaming classes); and elastic or non real-time traffic, for those classes less critical to time variations but where data integrity is more critical (interactive and background classes). This last category fits the model of Transmission Control Protocol (TCP), where closed-loop mechanisms are applied for congestion control by adapting the average data rate per flow to react to congestion.

The issues involved in incorporating QoS mechanisms into mobile environments have been receiving considerable attention from the industry as well as the research community [8], [9] and [4]. Renegotiation of performance guarantees during handoff, differentiated allocation of resources through power control and prioritized access to the medium are some of the mechanisms through which QoS may be incorporated into 3G systems.

2.2.3. Traffic Modeling Assumptions

Due to the layered nature of data networks, most commonly exemplified by the Open Systems Interconnection (OSI) network architecture, there are many levels at which one can be interested in describing traffic behavior. Assuming that we are looking for a model capable of describing connectionless networks, it would be natural to choose the packet (network) layer as our focus for analysis. However, while the definition of connectionless packet-switched networks permits packet independence, in practice it is reasonable to consider that packets sent from the same instance of an application to the same end destination are indeed very correlated [10]. Based on this idea, it is useful to define, for the purpose of resource allocation, the concept of flow, as a sequence of packets sent between the same source/destination host pair.

In QoS-based systems, such as the model assumed for third-generation networks, both the real-time as well as the elastic classes can be implemented with some sort of admission control mechanisms, where flows would be considered much like “connections” at their setup phase, with the network deciding on whether to refuse or accept new flow requests after estimating the performance impact that it would cause on already established flows.

There are many recent theories to model the flow arrival and flow size processes in broadband multimedia applications. Although many differences can be observed among them, it is clear that these modeling approaches can be divided in two major groups. The first and more traditional group approaches data traffic description using Markovian models. The second group uses a more recent approach, based on the self-similar nature observed from multimedia data sources. We will discuss each of them next.

The Markovian models, from which the Poisson-related are the main subset, constitute the fundamental pillar of existing queueing theory. Erlang theory is one example of such an approach, where calls are modeled as having Poisson arrivals. Other models exist in this group to model packet arrivals, for instance using Poisson-related such as Poisson-batch or Markov-Modulated Poisson processes. For further description of such models the reader is referred to [11].

In the second group stands a more recent set of models that rely on the observation that bursty traffic patterns generated by data sources and variable bit applications tend to exhibit certain degrees of correlation between flow arrivals and show long-term dependence in time (self-similar traffic) [1]. The seminal paper on the study of self-similar traffic was published in 1993 and further expanded in 1994 [12]. It shattered the basic assumptions of queueing analysis using Poisson models. Based on massive traffic measurements on an Ethernet network, this work and many others that came after it show that data traffic displays structural similarities across a wide range of time scales.

2.2.4. Self-Similarity

Self-similarity is a property associated with "fractals," which are objects whose appearances are unchanged regardless of the scale at which they are viewed [13]. A self similar stochastic process looks or behaves similarly when viewed at different degrees of magnification or different scales on a dimension (time or space).

A mathematical definition of such behavior would show that such processes present approximately the same statistics the magnified (or aggregated) versions of the same process. In other words, by compressing a discrete-time stationary process of a factor *m*, the mean, variance and correlation are preserved. This suggests that the burstiness would be kept at different time scales. A weaker condition is asymptotical self-similarity, which states that autocorrelation functions keep the same form of the non-compressed process. An interesting consequence of this property is that as the compressing factor tends to infinity, the correlation does not tend to zero, as in stochastic data models previously used for packet data. This characteristic is referred to as "long-range dependence."

Self-similarity and long-range dependence are not exactly the same concept and do not necessarily imply each other, but in the context of traffic dimensioning those terms are often used interchangeably, as well as "scale-invariant" burstiness.

It is often observed that file sizes, connection durations, or, more broadly, the duration of on/off periods exhibit a heavy-tailed distribution.

Although the observation of self-similar behavior came to prove wrong the traditional assumption of exponential distribution for flow durations, the flow arrival process for many types of data

applications could still be assumed as following Poisson processes for each individual source [14]. In this case, the candidate model to describe the busy server would the M/G/∞ process, which assumes Poisson arrivals with generic service time following a heavy-tailed distribution. However, this model assumes independence between the sources and this assumption is not possible for all types of data applications. Particularly World Wide Web (WWW) connection arrivals have been shown to present self-similar behavior [15], which is explained by the correlation between multiple flows, including both the correlation among flows generated by the same source as well as the correlation among flows from different sources but that are governed by congestion-control mechanisms such as those implemented by TCP.

As mentioned before, the description of multimedia traffic as self-similar has brought much debate over the traditional assumptions used for decades in traffic theory. Much research has been done on the subject as a result of those initial findings. Some of those topics are listed below [15].

One of the research topics on self-similarity refers to measurement-based traffic modeling (source modeling), where different types of data sources are measured with the intention to quantify their self-similar characteristic. The measure of how much a source behavior is self-similar is expressed by the Hurst parameter, which ranges from 0.5 (non-self-similar) to 1 (completely self-similar). Examples of sources that have already been investigated include [15]: Local Area Network (LAN), Wide Area Network (WAN), Internet Protocol (IP), File Transfer Protocol (FTP), copper, fiber optic, WWW WAN traffic, etc.

Another line of recent research focuses on physical modeling of traffic sources in an attempt to understand the physical causes of self-similarity in network traffic. The idea is that explaining the causes of such effects may help selecting between models that fit equally well and give additional insight into the model description. Two main causality possibilities are considered [15]. In single source causality, the arrival pattern of a single data source, e.g. the variability found at multiple time scales on a variable-bit-rate (VBR) video stream (Moving Picture Experts Group - MPEG), could be explained by the variability in time duration between two successive scene changes. Structural causality is attributed to the heavy-tailed distribution of files or object sizes. “If end hosts exchange files whose sizes are heavy-tailed, then the resulting network traffic at multiplexing points in the network layer is self-similar” [15]. The most typical physical model of traffic sources is the ON/OFF model [16], typically using heavy-tailed distributions for both active (ON) and idle (OFF) periods.

Researchers have also been focusing on the impact of self-similar sources on queueing behavior. An important point to consider is that infinite buffer systems with non-self-similar traffic (short-range dependent) input, generate an “exponentially decreasing” queue length distribution. On the other hand, self-similar traffic (long-term dependent) input generates slower-than-exponential (or sub-

exponentially) decreasing queue length distribution, sometimes with polynomial decreasing behavior. This implies that increasing buffering results in little improvement in the packet-loss rate, at the cost of the queueing delay penalty it imposes to the system. For this reason, there have been proposals advocating small buffer capacity/large bandwidth resource provisioning strategies [15].

Traffic control research explores techniques to improve the efficiency of the resource allocation, based on the knowledge of the traffic nature. For instance [15], by controlling the duration of connections, it is expected, based on the assumption of "heavy-tailedness," that long connections will last even longer, in contrast to the memoryless property of Poisson processes. Based on that expectation, techniques to shape the traffic or dedicate more efficient resources to those flows could be implemented.

While research currently focuses more on self-similar traffic modeling, the recognition of limitations in the previously used Poisson-based approach does not necessarily voids its use. For instance, Poisson assumptions may provide useful bounds on performance metrics of interest for the system. The particular property of those traditional methods that makes them attractive is the simple mathematical formulation for the quality of service in terms of offered traffic and available network capacity.

2.2.5. Multiple Time Scales

One way to describe broadband traffic is to model its different time scales with appropriate stochastic processes used for each level. A simplified summary of the most relevant time scales for the resource allocation problem is adapted from [6] and shown in Table 2-1.

Table 2-1. Traffic Modeling: Scales in Time

	Stochastic phenomenon	*Stochastic models*	*Traffic impairments*
Packet/burst scale	Statistical multiplexing of packets or groups of packets from different sources	Various queue and traffic models	Packet loss, delay, jitter
Flow/Session scale	Admission of connections	Various population models	Blocking
Human activity scale	Correlated human activity giving daily, weekly and yearly cycles	Statistical inference of cycles and correlations	Blocking
Planning scale	Forecasting of traffic demand after planning lead time	Numerous forecasting techniques	Blocking

A brief description of each of these scales as well as possible models for each one are provided next.

2.2.5.1. Packet/Burst Scale

This scale models the arrivals of packets or groups of packets from multiple channels that will be multiplexed using a specific queue model. One phenomenon of interest is the packet-grouping that occurs due to packet/framing performed at the upper layers or due to the packet rate variations existing in each flow due to application demand variation, such as real-time applications on video transmissions.

The main parameter to be modeled at this time scale is the queue buffer size, which should be chosen based on the desired transmission quality, typically expressed by the probability of packet loss.

2.2.5.2. Flow/Session Scale

The flow scale layer would be the initial scale for circuit-switched systems, as no packet or burst scale exists. In the old narrowband telephony systems, a flow would be the equivalent of a call, and failure would be described as a refused call attempt. In an ATM data network, the flow scale could be described as the level at which a connection setup (virtual circuit) is established. At the more general level, a flow can be defined as [17] the unidirectional succession of packets relating to one instance of an application. Still, according to [17], the packets belonging to a flow have the same identifiers (source and destination addresses or port numbers) and occur with a maximum separation of a few seconds.

In the case of circuit-switched networks, such as the traditional voice telecommunications networks, the basic model to look at for the purpose of predicting the number of simultaneous users in the system is the M/G/∞ queue. In this scenario, the system is modeled as having a Poisson process for the arrivals (i.e. flow setup attempts), a generic type of service and an unlimited number of servers. For practical systems, the number of servers is physically limited and the system quality is measured as the probability of the number of servers in use being larger than the number supported by the system.

In the more general packet-switched data networks, the typical variable of interest in dimensioning a system is the probability of the total effective bandwidth required being greater than

the channel capacity, given the statistical description of individual flows behaviors. Various models exist to obtain this relationship and the probability distributions to be used for the flows' effective bandwidths modeling are still the object of research and discussion, mainly with respect to the self-similarity traffic behavior as described before. One interesting observation is that the circuit-switched case is usually a simplification of the general case, if we assume that all flows are at constant bit rate and we express the total link capacity as the total number of supported flows at that fixed rate.

2.2.5.3. Human Activity Cycle Scale

At the human activity cycle scale, we should look at the hourly variations of the parameters used as input at the flow level. Parameters such as number of flow setup attempts, holding times, mean and peak data rates, used in the dimensioning dependencies described at the flow level have local averages that vary along the hours of the day, the days of the week, seasons, etc. For the task of system dimensioning, one needs to consider the strong correlations in human traffic patterns, in order to predict when traffic demand will hit its peaks and planning the capacity for those situations. A typical example of this time scale is the concept of the "busiest hour," used in typical voice telephony systems.

A wireless scenario must account for human cycle variations, which include not only time but also strong spatial fluctuations. In this case, dynamic dimensioning techniques can be used, such as dynamic channel allocation, where the spectrum distribution among cells and sectors can be changed to accommodate the user demand concentration at different times of the day.

A suggestion of a practical approach in considering the human activity cycle scale for broadband networks will be discussed later in this document.

2.2.5.4. Planning Scale

At the planning scale the main aspect consider is system growth, which depends on variables such as the progression of market share and market penetration and the long-term changes in the user behavior.

The basic trade-offs in designing a system to accommodate long-term growth are the risk of over investing in a network and having it under-utilized if too much margin is allowed, versus the risk of sub-dimensioning the network and losing revenue as well as market-share. The appropriate demand characterization for optimal system dimensioning needs to take into account variables such as latency between the time a requirement for growth is perceived and expansion can be implemented

and costs involved in system expansion as opposite to planning ahead for a large system. In the particular case of wireless networks, system deployment is known as allowing a very quick response time, due to the cellular architecture that allows the operator to "drop" sites in busy areas without really building expensive infrastructure to get to those places by land. However, resources are usually wasted when cell splitting techniques have to be used to solve traffic problems compared to the design that would have been done if the same traffic had been considered at the initial deployment. This is due to the fact that cell splitting usually requires deactivating and moving sites or keeping under-utilized cells in some areas. Some additional practical aspects of the planning scale will be discussed more later in this document.

2.3. Traffic and Link Asymmetry

Implicit in the problem of user profile characterization is the fact that data demand is usually asymmetric, i.e., traffic is typically more dense in one direction (e.g. base-station to mobile) than in the other (mobile to base-station). Many of the multimedia services to be offered in broadband wireless systems are heavily unbalanced towards the downstream direction. Examples include video and audio streaming, web browsing, and simulcast. For those applications, there is usually little interaction and transmission from the user end. Other applications like real time video-conferencing, voice and data transfer, present similar proportions between downstream and upstream traffic volumes, closer to symmetry. There are still a few applications, mostly related to telemetry, that present heavier upstream traffic.

The process of traffic modeling described in the previous sections could be initially thought of as two separate problems, the separate demand characterization of the downstream and upstream user behavior. However, those two problems are strongly correlated and the probability distribution of a subscriber sending data is in many ways related to that of the user receiving data, depending significantly on the type of application. The correct characterization of the user asymmetry and its evolution as more and more broadband applications are available is a challenging traffic-engineering problem and has strong consequences on the way a system is dimensioned.

While traffic on multimedia applications is typically heavier on the downlink, the wireless mobile coverage is typically limited on the uplink due to antenna, size and battery restrictions at the mobile unit and that link asymmetry towards a weaker uplink may partially counterbalance the traffic asymmetry on the downlink.

In wireless systems there are many techniques to deal with the asymmetry and they may be grouped based on the system's duplexing mode, i.e., whether it uses frequency-division-duplexing (FDD) or time-duplexing-division (TDD).

In FDD systems, slices of the spectrum are allocated separately for the uplink and downlink. Support for asymmetry may include differentiated spectrum allocations for the downstream and upstream and adaptive coding/modulation techniques (link adaptation). The first type of solution is a resource typically used in fixed-wireless systems (MMDS, LMDS), where technologies do not follow strict standards and operators have flexibility to distribute spectrum allocation. For mobile systems (cellular, PCS and 3G) spectrum allocation is typically standardized and symmetric (paired channels) and, as a consequence, link adaptation solutions are the most common techniques used to compensate for traffic asymmetry.

Link adaptation relies on the use of different coding and/or modulation schemes in each direction, with stronger schemes being used on the upstream to compensate for lower signal strength requirements due to the power restrictions of the mobile. On the downlink, more spectral-efficient schemes may be used to provide for heavier traffic demand, at the expense of higher power requirements that may be satisfied by the base-station's stronger power capabilities. Although link adaptation works nicely to compensate unbalances, depending on the intensity of traffic asymmetry one side of the spectrum may still end up being used less efficiently than the other.

In TDD systems the same spectrum is allocated for both downlink and uplink, but different time slots are used in each direction. To support asymmetry, the system can use uneven distribution of time-slots between downstream and upstream. This option provides more flexibility to the system in terms of accommodating future changes in the user asymmetry proportions, as asymmetry may be adapted on a dynamic basis as traffic behavior changes, with the objective of balancing spectrum efficiency in both directions.

2.4. Summary

In this chapter, we presented a brief review of traffic theory, discussing the models available for voice and data traffic, as well as the most recent and fundamental developments in this area. Some performance metrics were presented and the concept of service class was used to distinguish different types of traffic according to their application and performance requirements for a given service.

In the next chapter, we develop a practical methodology for demand characterization. We introduce the concept of "user profile," which ties together a given combination of service class,

propagation environment and terminal type. Next, we present a practical approach to explore the dynamics of user geographical distribution in creating multi-service, multi-class traffic layers that will serve as input for network traffic simulation models.

Chapter 3. Practical Aspects of Wireless Demand Characterization

The previous chapter dealt with the characterization of traffic that may be produced by distinct applications generated by each user. In dimensioning the system, we must be aware of the different *user profiles* supported, which are distinguished by a combination of attributes defining their service class, propagation environment and equipment characteristics. The concept of user profile is explained in detail in the following paragraphs.

The service class attributes specify the nature and volume of traffic of a user profile, as well as its QoS requirements. Looking at the many possible types of users for mobile broadband systems and their most likely applications, such as Internet access, voice, video conferencing, etc., it seems natural that we should try to group users into categories and describe each category by its typical parameters.

The mobility characteristics of the system will influence the user profile by having an impact on propagation behavior, which in turn limits data rates available to the user. The mobility also determines the user equipment alternatives for each profile. With regards to the mobility characteristic of a user profile, we divide users into three major groups and present below the main propagation and terminal differences among them.

Indoor users: Building construction attenuation presents the major propagation impairment, having an impact on link strength and limiting coverage area and bandwidth. On the other hand, users are mostly fixed or at very low speed, which reduces fast fading effects. Indoor signal boosters and microcells can be used to compensate for wall attenuations. Typical applications for this type of user are Internet access and connection to office LANs using laptops.

Pedestrian users: Terminal types include handheld phones and palmtops. Typically no building attenuations are expected, but the propagation path is stressed more due to larger cell radius allowed for outdoor cells in the system design. Voice and Internet access and video applications are the primary focus for these users.

Vehicular users: Typical terminal types are on-board car computers, hands-free terminals and handset phones. A fast-moving subscriber unit means that the effects of multipath fading may be severe. Expected applications include voice, video conferencing, Internet access and navigation systems.

Of course, each of these types of users may generate multiple simultaneous flows of different service types. In turn, the performance requirements within an application type may differ depending

on the propagation environment and terminal type. For instance, Internet access from a laptop within a building will allow higher data rates than Internet access by a pedestrian carrying a palmtop.

To characterize demand, we propose forming what we term a *user profile* by taking the Cartesian product of the service class (which specifies an application, performance requirements and traffic characteristics), propagation environment and terminal type. For instance, Video Conference-Indoor-Laptop and Web Browsing-Pedestrian-Palmtop would be two possible user profiles. Besides, traffic models for each class are likely to be parameterized differently for different user profiles accounting for variations in access rates, burstiness, etc. We show some examples in Table 3-1.

Table 3-1. Example of Multiple User Classes and Performance Requirements

User Profile		1	2	3	4	5
Service Class/ Traffic Model	Service Name	Video Conference	Voice	Streaming Video	Web Browsing	E-mail
	Service Class Group	Conversational	Conversational	Streaming	Interactive	Background
	Type of switching	Circuit	Circuit	Packet	Packet	Packet
	Max. Bit Rate Down (kbps)	64	13	144	1000	16
	Max. Bit Rate Up (kbps)	64	13	-	500	16
	Session Arrival Distribution	Exponential	Exponential	Exponential	Exponential	Exponential
	Session Length Distribution	Exponential	Exponential	Exponential	Exponential	Exponential
	Burst Arrival Distribution	Exponential	Exponential	Exponential	Pareto	Pareto
	Burst Length Distribution	Pareto	Exponential	Pareto	Pareto	Pareto
	Maximum BER - Down	1.0E-06	1.00E-03	1.00E-06	1.00E-08	1.00E-06
	Maximum BER - Up	1.0E-06	1.00E-03	-	1.00E-08	1.00E-06
	Required Eb/No - Dw (dB)	8	7	12	1	3
	Required Eb/No - Up (dB)	8	7	-	1	3
	Tolerated Delay (ms)	100	100	200	-	-
Propagation Environment	Propagation Scenario	Indoor Slow	Outdoor Fast	Outdoor Slow	Indoor Slow	Outdoor Slow
	Terminal speed (km)/h	0	120	120	3	60
Terminal Type	User Terminal	Laptop	Mobile Phone	Palmtop	Laptop	Mobile phone

Accurate demand characterization should include proper description of the different geographical distributions for the users classified in each profile. The methodology presented here suggests the use of demographic database layers [18] as input to this process and discusses possibilities of mapping these data into *user layers*. We discuss some of the main data inputs in the next subsections.

3.1. Residential Database (Census Data)

One important layer to be considered in forecasting traffic demand is the residential geographic user distribution. In the U.S., this type of information usually comes from census data; the database is a collection of polygons that correspond to census tracts or census lots (a smaller subdivision of census tracts). Relevant information is available about the geographical area corresponding to each of these polygons, including:

- residential population;
- number of households; and
- statistics such as ethnicity, gender and age distribution.

Figure 3-1 illustrates this type of database, including information about population distribution in the San Francisco Bay area.

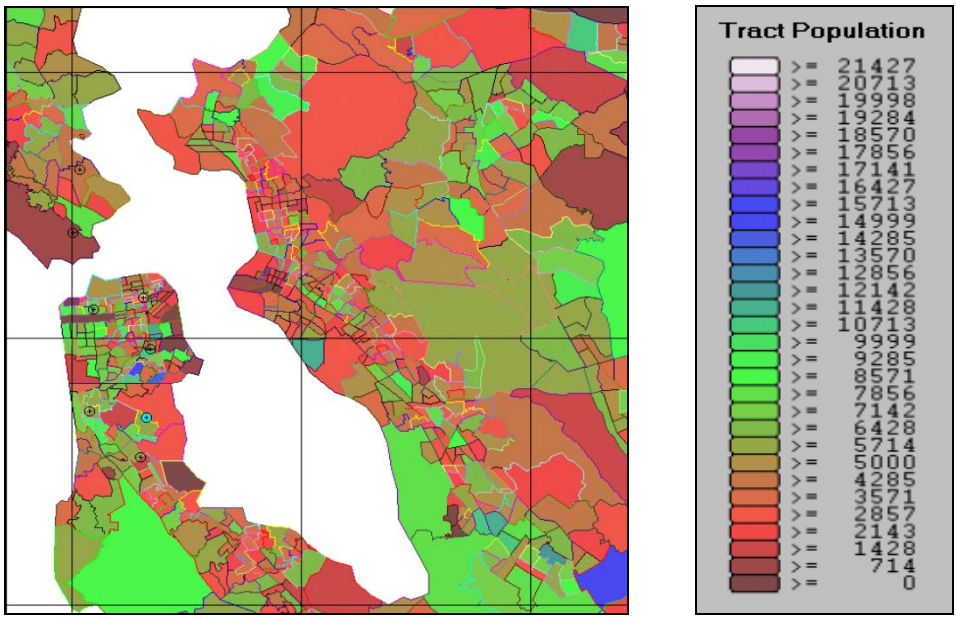


Figure 3-1. Example of residential database (census tracts) for San Francisco Bay area.

This type of database constitutes one of the layers of information to be used in the demand forecast. Considerations need to be made to forecast user profiles and correlate them with some proportionality factors to the attributes available for the geographical areas. For instance, typical mobile forecasts assume that the market penetration will be a factor applied directly to the population

in each polygon, while fixed wireless applications apply the proportionality over the household attribute.

Besides the census data, other commercially available databases classify the areas with other attributes that may be of interest to the operator such as average income, average number of years in school, level of familiarity with "high-tech" consumer items, etc. This allows another level of refinement where sub-layers could be used. For instance, the forecast could assign different values of market penetration and average bandwidth demand depending on the household income in each area.

From the residential geographical distribution, using different market forecasting factors, one traffic layer distribution will be created for each applicable class related to the residential type of user.

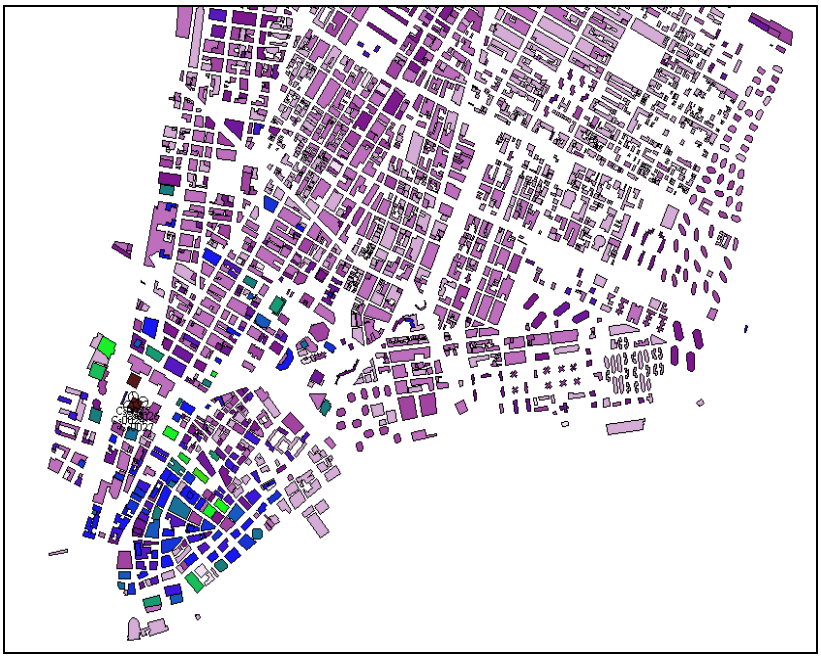
3.2. Business Database

From the business layer, using different market forecast factors, one traffic layer distribution will be created for each applicable class related to the business type of user.

Another layer in building a geographic traffic distribution profile is a business database. This type of data may come from different sources in each market, such as chambers of commerce databases, Geographical Information Systems (GIS) government databases, etc. Due to the diversity of sources they do not follow any standard format, but are typically either a collection of polygons or a collection of points, geographically distributed, with attributes including:

- number of employees at the location;
- average business revenue;
- average monthly expenditures on telephone/data services;
- available office area in square feet; and
- type of business.

Figure 3-2 illustrates this type of database for the Manhattan area.



D	StreetAddress		ZIP	Area	Height	Owner	LotClassification
390001	00082	WALL STREET	10005	53253	150	THE OTTO GERDAV CO	03
260021	00045	WALL STREET	10005	493187	312	45 WALL ST L L C	09
270001	00053	WALL STREET	10005	208831	130	55 WALL COMPANY INC	01
270009	00059	WALL STREET	10005	390531	399	63 WALL INC	03
270017	00067	WALL STREET	10005	303175	292	GESELLSCHAFT FUER IMM	03
310011	00085	WALL STREET	10005	680213	370	EIGHTY-FIVE WALL ETC.	04
330011	00095	WALL STREET	10005	459000	312	CHEMICAL BANK	04
330022	00099	WALL STREET	10005	91368	335	P A BUILDING CO	03
350010	00107	WALL STREET	10005	990205	296	CITICO RP	04
350014	00037	WALL STREET	10005	377214	0	WHW ASSOCIATES	03
380001	00100	WALL STREET	10005	463664	365	100 WALL LLC C/O THE	04
450009	00014	WALL STREET	10005	838749	0	14 WALL T REALTY CO.	04
390004	00078	WALL STREET	10005	50044	148	ANADAE RLTY CP	03
400001	00072	WALL STREET	10005	324705	186	AI REALTY CORP	03
400003	00060	WALL STREET	10005	1617206	0	MORGAN GTY TRUST CO	04
400014	00048	WALL STREET	10005	287766	412	THE BANK O F NEW YORK	04
430001	00044	WALL STREET	10005	285049	0	NEW YO RK LIF E INSURAN	03
430002	00040	WALL STREET	10005	1061266	760	40 WALL DEVEL ASSOCS	04
430005	00030	WALL STREET	10005	108963	162	30 WALL ASSOCIATES LL	03
430006	00026	WALL STREET	10005	23199	50	US GOVERNMENT	29
460001	00002	WALL STREET	10005	173159	290	FIELDSTONE CAPITAL, I	03
370008	00105	WALL STREET	10005	295376	320	110 WALL COMPANY	03

Figure 3-2. Example of business database in Manhattan area.

To forecast traffic distribution based on this type of database, one typically relies on market penetration data per type of business and estimates the range of the required data rate depending on the number of employees. For instance, a high-technology business will tend to have more employees who require wireless Internet access than a grocery store.

3.3. Road Traffic

The road traffic database accounts for vehicular and pedestrian users. It locates geographically, usually as vectors, the main roads, highways and streets in a certain area. Typical information available in this database is the number of vehicles that transit on given sections of the road per day. Forecast of data traffic must also take into consideration mobile traffic at the busiest hour.

Again, using different market forecast methods, one traffic layer distribution must be created for each applicable class related to vehicular and pedestrian users.

3.4. Morphology Weighting

As discussed above, traffic demand from residential users is typically available as a set of polygons. In performing resource planning, the designer is confronted with the question of how to geographically distribute the traffic associated with each polygon. The simplest approach would be to assume traffic to be uniformly distributed over the entire area covered by the polygon. However, further refinement is possible if information regarding land usage (morphology) is available. Morphology databases classify the terrain as to what type of land-cover exists at each location. Typical terrain classifications include water, areas of low/medium/dense vegetation, urban and suburban areas, as exemplified in Figure 3-3.

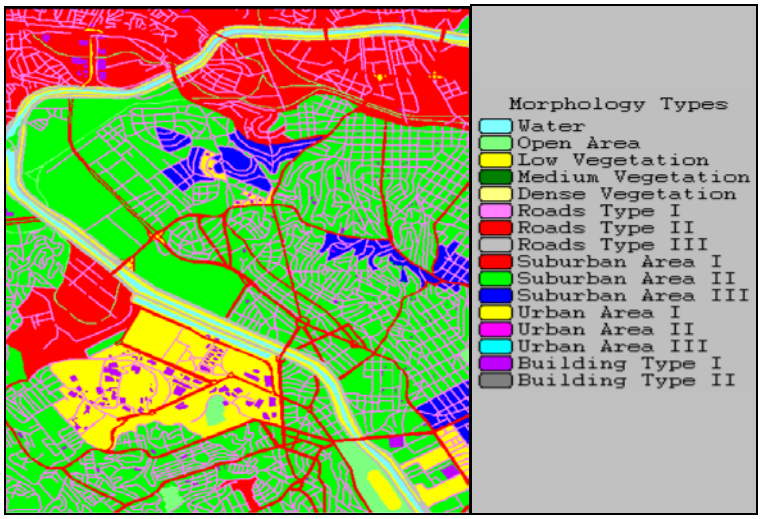


Figure 3-3. Example of morphology database.

When distributing traffic geographically based on a residential database, one may consider different weights to account for the relative probability that a user in a certain polygon will be located in an area of dense vegetation versus a suburban area, for instance. This is illustrated in Table 3-2.

Table 3-2. Example of Morphology Weights for Traffic Distribution.

Morphology Types	Type 1 City Urban Market	Type 2 City Suburban Market	Type 1 City Rural Market
Water	0	0	0
OpenArea	5	5	5
Low Vegetation	2	2	2
Dense Vegetation	1	1	1
Roads - Type I	10	10	10
Roads - Type II	12	12	12
Roads - Type III	15	15	15
Suburban Area I	40	40	40
Suburban Area II	60	60	60
Suburban Area III	80	80	80
Urban Area I	100	100	100
Urban Area II	120	120	120
Urban Area III	150	150	150
Building Block - Type I	200	200	200
Building Block - Type II	250	250	250

3.5. Market Penetration Forecast

As discussed in the previous sections, multiple layers of traffic sources (residential, business, road traffic, etc.) must be considered individually as to how they correlate with expected load for each traffic type. Other factors that could also have an impact are event-related locations, which may be permanent (stadiums, theatres, and show arenas) or temporary (festivals, athletic games, etc.). Such considerations present further refinement of load forecast to be used for network planning and design.

Network design must take into consideration not only current requirements but also how these requirements may change over the near and medium terms. Market forecast data must be used to modify the information discussed above. Important metrics include market penetration, market share factor and population growth from year to year.

Clearly, the network operator has a direct influence in some of these factors, as the growth of its customer base will depend on the provided service quality as well as the aggressiveness of the marketing and pricing strategies. However, service quality and return on investment to allow aggressive pricing will depend on a good dimensioning system that uses appropriate combinations of

all factors mentioned above. Therefore, at the same time that the operator can "forecast" market penetration, it can also influence the market and make this forecast become a reality.

3.6. Combining Multiple Layers

Once all the information described in the previous section comes together, it must be combined, as illustrated in Figure 3-4. One should note that from a modeling point of view, this combination does not consist of a simple sum. First, each traffic class must be taken into account separately; besides, traffic models for each class are likely to be parameterized differently for different user profiles (accounting for variations in access rates, burstiness, etc.). Finally, there is the question of time dependencies, discussed next.

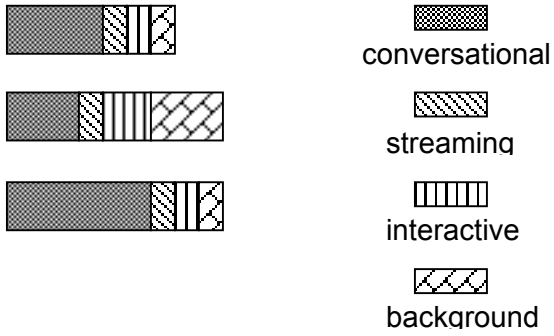


Figure 3-4. Profiles for residential, business and road traffic for a given region, taking into account census data, morphology weighting, market forecast, etc. The total load for purposes of network dimensioning is the superposition of all these factors.

Suppose that, for each layer, the method used helped us obtain the traffic distribution during the busiest hour. However, the busiest hour for each layer does not necessarily correspond to that of the other layers, as illustrated in Figure 3-5.

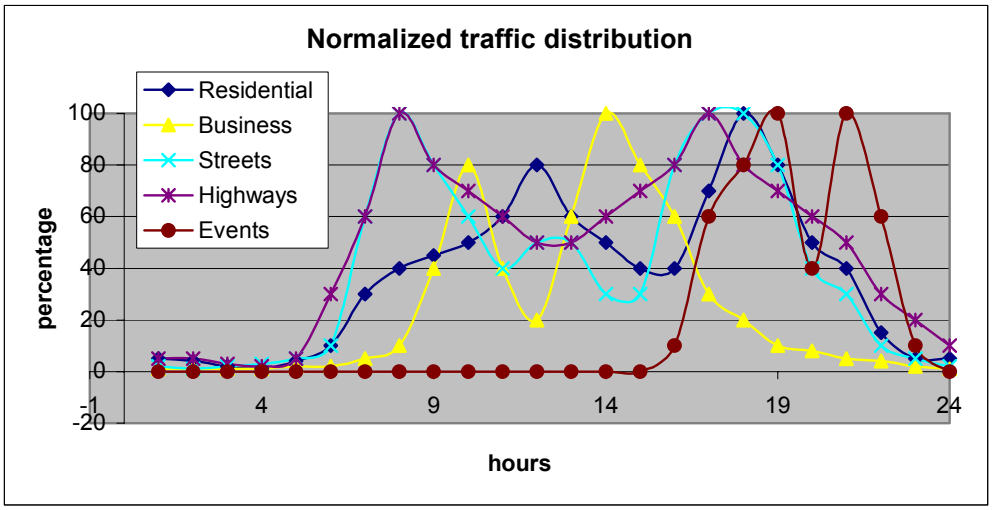


Figure 3-5. Example of traffic distribution dynamics with time for a given class.

It is not reasonable, therefore, to combine all layers at their peak hours, since it would lead to over-dimensioning. Time dependency weighting must be applied to correctly assess demand at any given time.

For instance, suppose we are evaluating system behavior at 2:00 PM. The business layer should be considered with 100% weight, but the residential layer would be scaled to under 50% of its peak value. If we look at the combined traffic grid at this time, we would most likely find the highest bandwidth demand on those cells located in business areas, and few channels needed on the residential ones. However, if we look at the 7:00 PM time frame, heaviest demand for bandwidth is now more likely to come from residential areas.

Since we are looking for ways to dimension channels on a per sector basis, not only on a system basis, one cannot ignore the spatial distribution of the busiest hour over different layers and the resulting required bandwidth per sector will be the maximum demand from looking at all time periods. To allow for all possibilities one would need to create combined class grids for all hours of the day. For simplicity, however, usually only two or three hours are selected. A typical system analysis looks into hours where most relevant layers have their peaks, such as residential and business in the example above.

Figure 3-6 illustrates one traffic grid created as a combination of multiple input layers (residential, business, etc.) for a certain application type and time of day. This grid is represented by small boxes of fixed size (grid resolution), each one with a particular set of traffic attribute values. Attributes could be expressed in terms of user density, throughput demand density, etc.

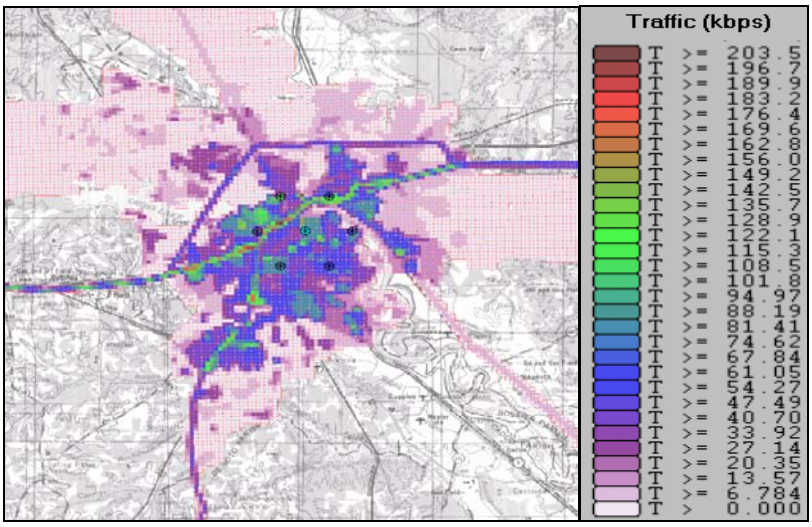


Figure 3-6. Example of one-layer traffic grid for a specific service class and time of the day.

The previous illustration displays the effective traffic demand load offered for a given application type. This information must still be augmented by the profile of QoS requirements with each region on the grid. Multiple traffic grids are combined together to characterize a demand scenario for a certain time of the day, as illustrated in Figure 3-7.

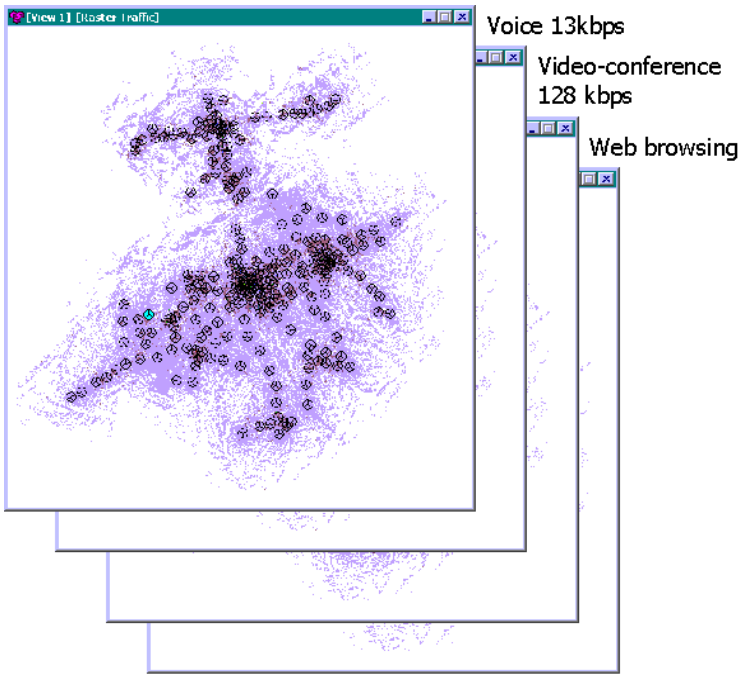


Figure 3-7. Example of one demand scenario for a specific time of the day , with multiple traffic grids (one grid for each service class).

3.7. Additional Input Variables for Existing Systems

Besides the methodology described in the previous section for demand characterization, other valuable information may be used when the system in question is an existing system. This situation is expected to happen in most of the cases for the third-generation systems where existing PCS and cellular operators will upgrade their networks to satisfy the user demand for data communications. The most important pieces of information to take into account from existing systems are switch traffic and handoff statistics per cell.

Switch traffic data (Erlangs/cell) statistics describe the existing voice traffic distribution per cell. This information is expected to have a strong correlation with the data traffic distribution for the same system when services are available. In an alternative to the traffic grid creation method described in the previous section, one option would be to create traffic grids based on the information of traffic per sector, spreading the traffic associated to each sector under its best-server coverage area. Factors could be applied to map Erlangs into number of users and profile extrapolations to factor those voice-user distributions in to data-user distributions for different profiles. This process is illustrated in Figure 3-8.

The switch handoff statistics (intercell dynamics) are the most valuable information to allow the appropriate dimensioning of the demand overhead caused by the handoff situations, where traffic may be served by sectors that are note the best server at the location, or may be taking resources from more than one sector at a certain time (soft handoff).

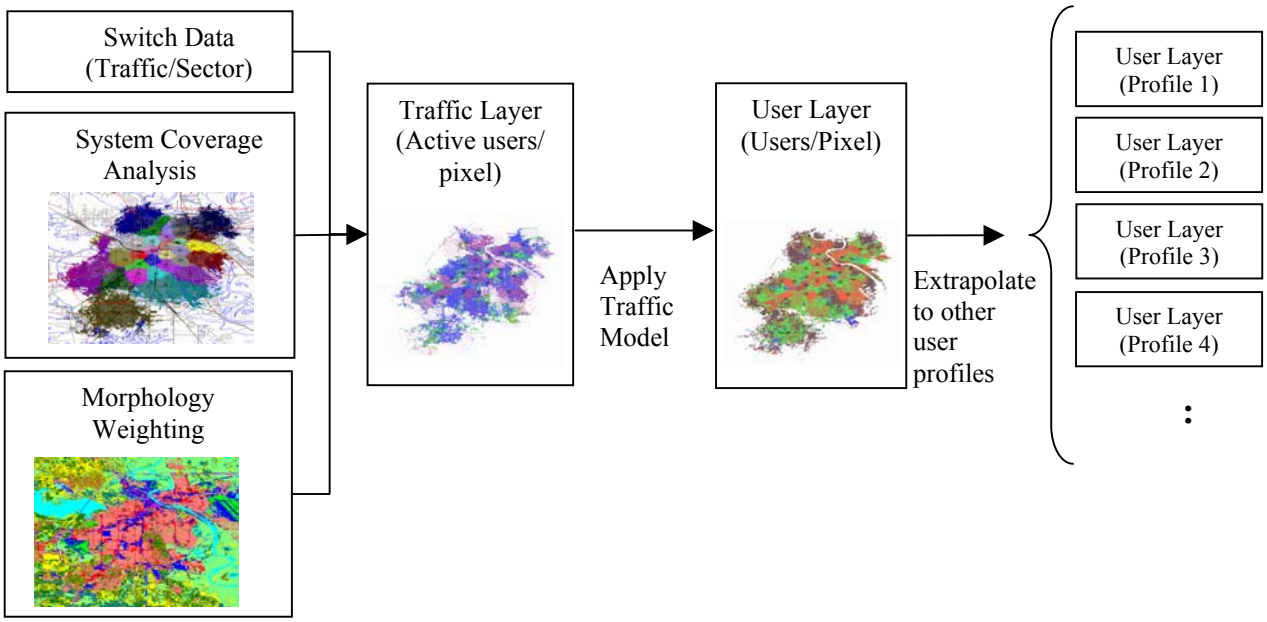


Figure 3-8. Using existing switch data to create traffic distributions.

3.8. Traffic Simulation Input

Once a traffic grid is created for each class, as a combination of multiple layers, the next step is to cross this information with the cellular system design, considering cell locations and their coverage areas. Resource requirements assessment must, of course, consider the type of technology selected for the system, access methods, modulation schemes, etc. System-level simulation is the most likely means to perform this assessment, due to the difficulty of conducting mathematical analysis that realistically models such a complex system.

Figure 3-9 summarizes the process described in this section, where multiple user distribution layers were first processed in order to estimate user profiles and dismember them into class layers, that combined together with a certain distribution weighting define one simulation scenario for a 3G system to be dimensioned.

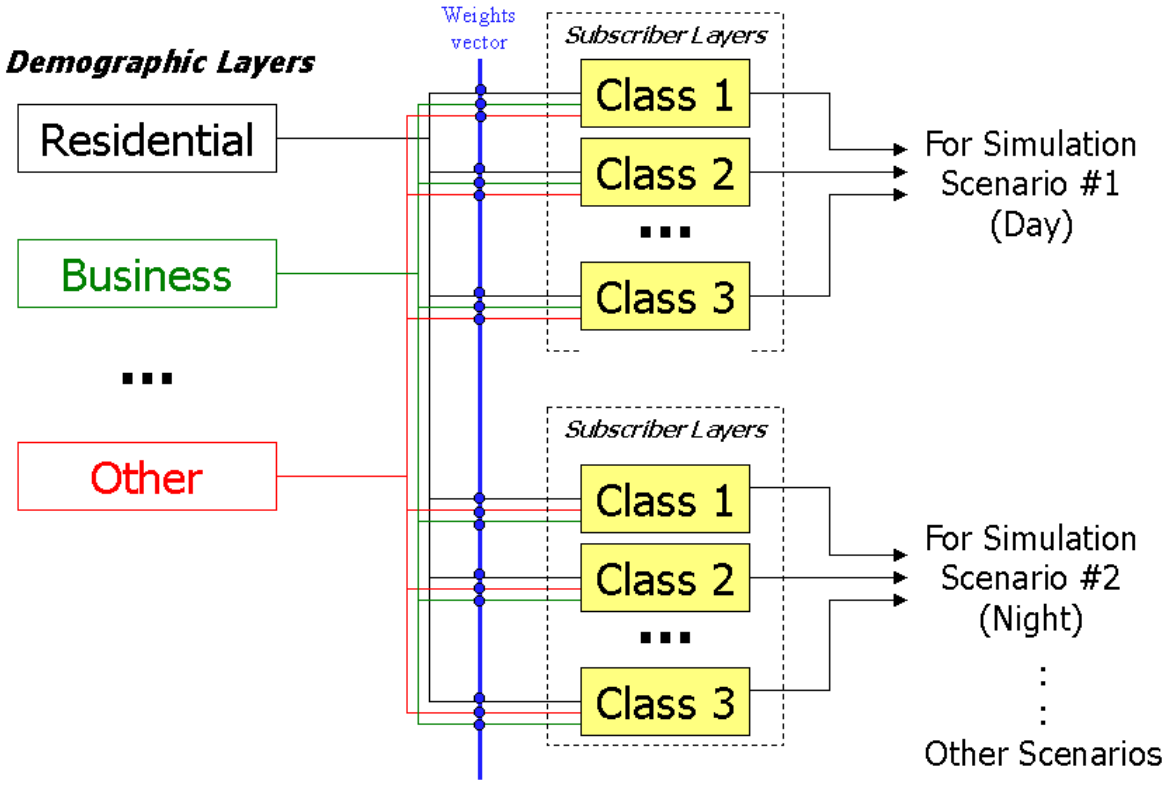


Figure 3-9. Block diagram of demand database estimation.

3.9. Summary

In this chapter, we proposed a practical methodology for demand characterization. We presented the concept of "user profile," which ties together a given combination of service class, propagation environment and terminal type. Next, we developed a practical approach to explore the dynamics of user geographical distribution in creating multi-service, multi-class traffic layers that will serve as input for network traffic simulation algorithms.

In the next chapter we discuss the concept of Quality of Service, with a focus on the physical layer for third-generation systems, and how these QoS requirements map into higher layers, including mechanisms of radio resource management, such as power control and congestion control. The concepts of system capacity and traffic mix dependence are also explored in that chapter.

Chapter 4. Providing QoS in 3G Systems

4.1. UMTS Architecture

The Universal Mobile Telecommunication System (UMTS) architecture has been proposed as the third-generation evolution for GSM systems, which are the most widely adopted standard for 2G wireless systems. For this reason, UMTS is expected to be widely implemented. Most European and Asian markets have been using GSM for many years now and in the US many operators have decided to migrate their 2G IS-136 (TDMA) systems to GSM as part of a migration path to 3G so they can also follow the worldwide trend and benefit in the long term from large scale production.

The standardization body for UMTS is the 3G Partnership Project (3GPP). In the US, a parallel trend for 3G is the evolution of IS-95 systems to CDMA2000, which is being standardized by the 3GPP2. The long-term goal of the standardization bodies is to achieve a global cellular system with wideband capacity to support multimedia mobile communications. Although most of the discussion of multimedia traffic modeling is common to both 3GPP and 3GPP2 paths, whenever this document deals with specific technology details, we will be focusing on the UMTS standard.

The vision of UMTS in its final implementation is to achieve an all IP core network where the type of radio access is of little importance, as much as it does not make a difference today for wired IP systems whether the lower layers use Ethernet, FDDI, etc. As an evolution path in that direction, a packet switched domain has been added to the core network of GSM. Basically a UMTS network consists of the radio access network, called UMTS Terrestrial Radio Access Network (UTRAN), which uses Wideband CDMA (WCDMA) as its access technology, the circuit-switched (CS) domain and the packet-switched (PS) domain. These two latter elements define the core-network (CN). These elements are illustrated in Figure 4-1 [19].

The User Equipment (UE) Domain includes the User Services Identity Module (USIM) Domain and the Mobile Equipment (ME) Domain. The USIM is typically a smart card, much like the SIM unit in GSM, that contains subscriber identity and security information that will be used for authentication and authorization procedures. The ME can be further subdivided into two entities: the Mobile Termination (MT) that performs the radio transmission and the Terminal Equipment (TE) that contains the end-to-end application. An example of a ME could be a laptop (TE) connected to a mobile phone (MT).

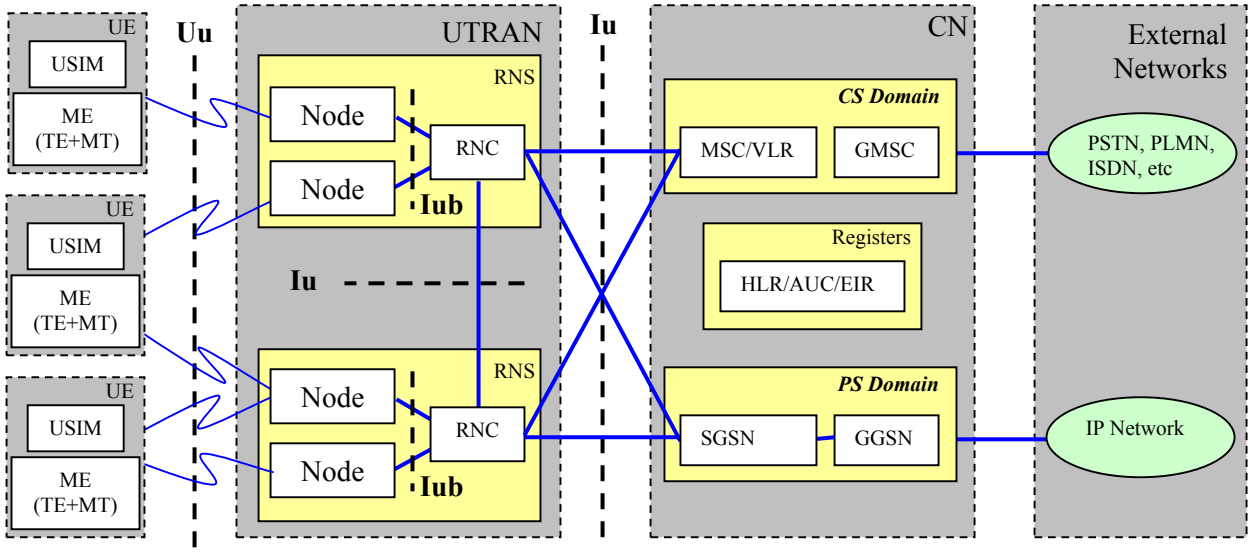


Figure 4-1. 3GPP UMTS Architecture (according to Release R3 [19]).

The UTRAN defines the Access Network Domain and its main elements are the Base Stations (BS) or Node Bs and the Radio Network Controllers (RNC). The combination of one RNC and its supported BSs is referred to as the Radio Network Subsystem (RNS).

The CN domain supports communications among UTRAN elements and also works as a bridge between the local and external networks. It supports mobility management and user and control data switching, among other important functions. The CN domain includes the circuit-switched elements, which are similar to those of GSM, such as Mobile Switching Center (MSC), Visitor Location Register (VLR), Gateway MSC (GMSC), as well as the packet switched elements such as the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). Additional elements of the CN are the Home Location Register (HLR), Authentication Center (AuC) and Equipment Identity Register (EIR), which are databases used for subscriber identification and security purposes.

From the point of view of one user, the Core Network Domain can be further subdivided in "Serving Network Domain" which is its access point to the core network, the "Transit Network Domain," as the core network path between the serving network domain and the remote party, and lastly the "Home Network Domain," to which the USIM is related by subscription.

Together, the Access Network and the Core Network define the Infrastructure Domain.

The external networks can be grouped into circuit switched (PSTN, other PLMNS, etc.) and packet switched, referred to in Figure 4-1 as the "IP Network."

The future of UMTS specs (Release 5 and up) predicts an all-IP network. All traffic is supposed to become packet switched (IP based) and a new CN subsystem called the IP Multimedia Subsystem

(IMS) should be added, which would deal with the VoIP calls and interact with the PSTN. In that scenario, the CS domain shown in Figure 4-1 would disappear.

4.2. UMTS End-to-End QoS Provisioning

4.2.1. Service Bearers

The UMTS architecture supports the concept of end-to-end quality of service, i.e., from one Terminal Equipment (TE) to another TE [20]. The bearer service layer architecture is shown in Figure 4-2.

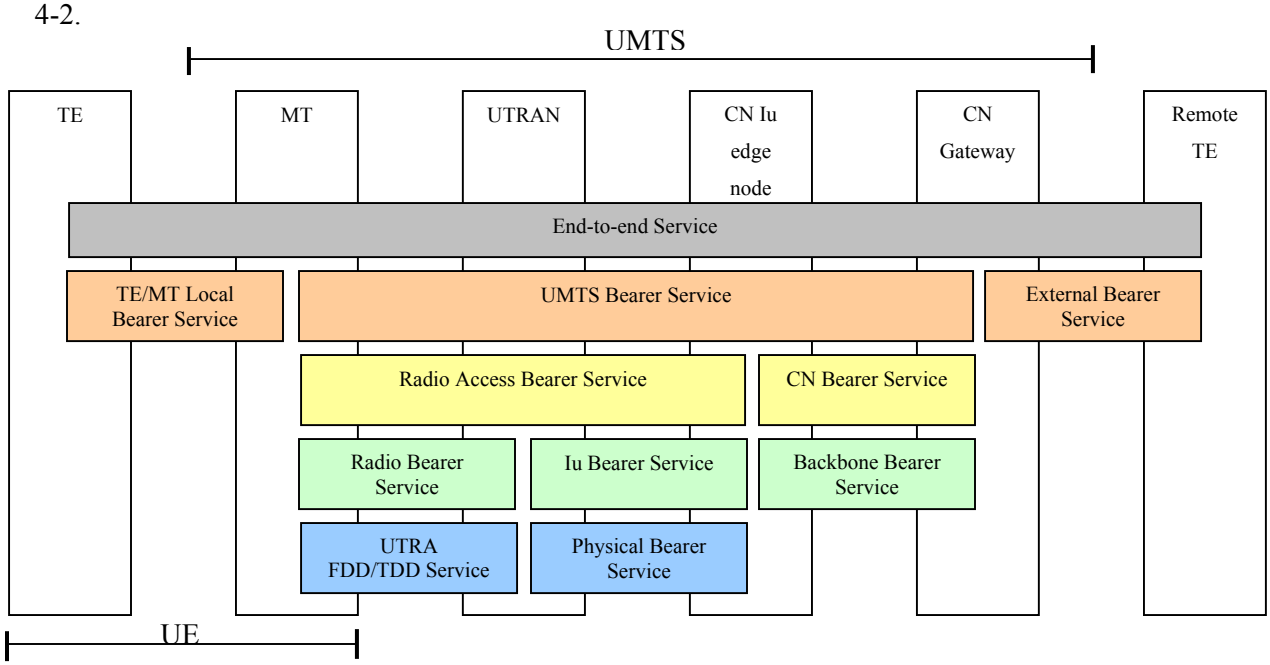


Figure 4-2. UMTS QoS Architecture [20].

Figure 4-2 illustrates how end-to-end services are supported through the use of bearer services of the underlying network. Each higher layer bearer service consists of bearer services of the lower layers. For instance, to achieve end-to-end service between two TEs, the TE from one side communicates with its MT through the *TE/MT Local Bearer Service* and its MT communicates with the infrastructure domain through the *UMTS Bearer Service*. If the other user is in another network, the *External Bearer Service* is used to achieve communication between the gateway of the local UMTS network and the external network. We are more concerned here with the UMTS network QoS provisioning, which is described in [21] and provided by the UMTS Bearer Service. For further reading, end-to-end QoS provisioning including the External Bearer Service is described in [22].

Inside the UMTS network, the *UMTS Bearer Service* consists of the *Radio Access Bearer* (RAB) service and the *Core Network Bearer* (CNB) service. Both of these services are specific to UMTS cellular technology and deal with aspects particular to the wireless mobile environment, such as user mobility and propagation profile.

The RAB service is responsible for security of signaling and user data between MT and the CN Iu edge node, while providing the negotiated QoS for the *UMTS Bearer Service* or default QoS for signaling. A RAB service is kept while the MT moves. The RAB consists of two bearer services: the *Radio Bearer Service* and the *Iu Bearer Service*. The *Radio Bearer Service* deals with the radio interface aspects (UTRAN), while the *Iu Bearer Service* provides transport between the UTRAN and CN.

The CNB service is responsible for the efficient use of the backbone network to satisfy the contracted UMTS. It uses a generic backbone network service according to the operator's choice.

4.2.2. Layer Structure

The User Equipment (TE+MT), being at one end of the connection, is responsible for controlling the end-to-end IP bearer of a connection. It is also responsible for the UMTS bearer. Figure 4-3 illustrates the protocol stack as seen at the UE, showing its main protocols and functional entities. It was created based on information described in [23],[24] and [25].

As can be seen from Figure 4-3, UMTS provides protocols and services corresponding to Layers 1 and 2 and the lowest part of Layer 3 in the OSI reference model. Other higher layer signaling functions related to UTRAN are mobility management and call control, which belong to the CN.

The PHY provides access to the radio medium. It handles tasks related to the physical transmission of information over the mobile radio medium. These tasks include error protection/detection/correction, multiplexing of transport channels, rate matching, closed loop power control and synchronization.

The MAC layer controls the radio resources by selecting the appropriate transport format, which includes reservation and priority schemes. It handles priority between different services of a mobile and between different mobiles (by means of dynamic scheduling). It also provides traffic indicators to the RRC layer to support resource management. MAC is responsible for service multiplexing on the physical layer and provides an unacknowledged transfer mode to the upper layers.

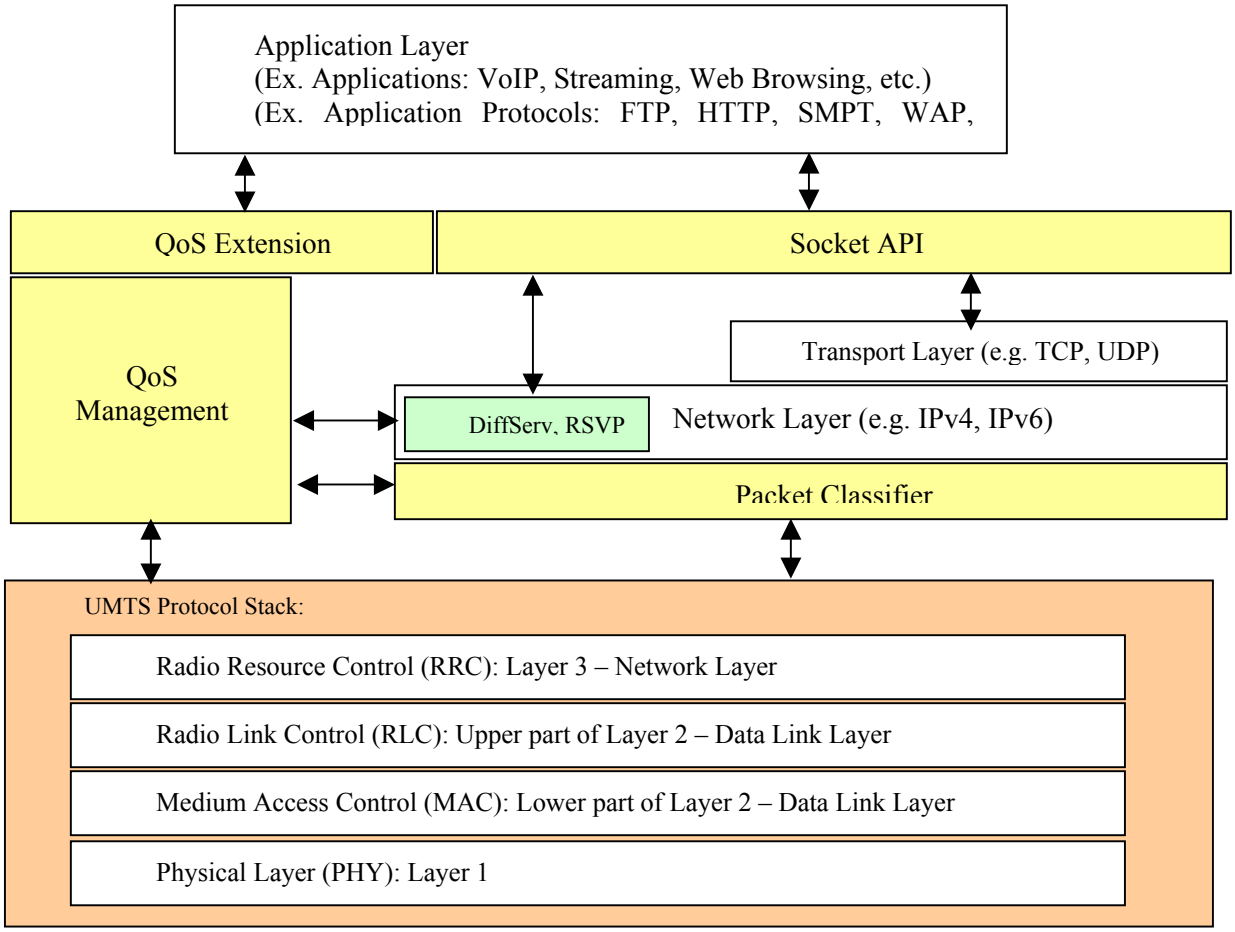


Figure 4-3. Protocol stack as seen at the UE, showing its main protocols and functional entities. Adapted from [23],[24] and [25].

The RLC layer implements the UMTS retransmission algorithms, which are optimized with different parameters (e.g. retransmission unit size) for different radio environments. It includes tasks such as segmentation/reassembly, RLC header addition/removal and transmitter/receiver buffering. Those tasks will be executed or not depending on the mode in which it is operating (transparent, unacknowledged or acknowledged modes).

RRC is the layer responsible for establishment, maintenance and release of connections and RABs, including admission control, mobility control (handoffs), radio resource allocation between cells, control of requested QoS and outer-loop power control, among other functions [25].

Within the UMTS protocol stack, peer-to-peer communication is provided by means of Service Access Points (SAP), such as the Physical Channels, Transport Channels and Logical Channels [26]. Logical Channels implement the SAPs between RLC and MAC and can be grouped in Control and Traffic channels. The transport channels are distinguished by how the information is transferred over

the radio interface (e.g. shared, dedicated, random access and broadcast). They constitute the SAPs between MAC and PHY. The physical channels are distinguished by the transmission characteristics over the air (e.g. carrier frequency, scrambling code and channelization code)

Layers 3 and part of Layer 2 in the UMTS protocol stack are "vertically" partitioned into the User Plane (UP) and Control Plane (CP). The User Plane refers to traffic generated by any traffic source above the UMTS protocol stack, including traffic used for signaling in the application layer. The Control Plane refers to the signaling traffic exchanged inside the UMTS network [23].

4.2.3. The UMTS Bearer Service

UMTS Bearer Service implements packet switched services through a packet data protocol (PDP) context. The QoS profile describing the quality of the UMTS bearer service is negotiated during the PDP context activation procedure and the QoS attributes are associated to the resulting PDP context. The QoS profile is the set of QoS attributes that are essential to describe the application's needs in terms of target transfer delay, reliability, bit rate, priority, etc.

Multiple PDP contexts with different QoS requirements are supported for the same IP address. One UMTS bearer corresponds to one PDP context.

The process of QoS activation between the UE and the GGSN (CN Gateway for PS domain) is described in [23] as follows.

1. Radio Access Bearer Establishment:
    - From the application needs, the UE maps QoS needs to relevant PDP context QoS parameters. Next the UE requests a PDP context activation from the SGSN.
    - The SGSN confirms whether the subscriber is entitled to use the requested QoS profile (authorization). Next, it will verify its resource availability and in case it is positive, the SGSN will signal the RAN for a Radio Access Bearer (RAB) to be established according to the requested QoS profile.
    - The RAN performs internal admission control and resource reservation and sets up a RAB in case those procedures are successful.
2. CN Bearer Service Establishment:
    - Once the RAB is successfully established, the SGSN makes a PDP-context activation request to the GGSN.
    - In case the GGSN has adequate resources for the requested QoS, it opens as transport connection (e.g. a tunnel) between the SGSN and GGSN. At this point a successful PDP context is established between the UE and the GGSN.

When an external network is involved, which is expected to happen in most connections, the UE and the Gateway nodes will be responsible for the mapping of the QoS parameters used in the IP network with the internal QoS parameters. These functions are described in [22].

Still according to [22], interworking with the external network may be realized in a number of ways, including:

- signaling along the flow path (e.g., RSVP, LDP);
- packet marking or labeling along the flow path (e.g., DiffServ, MPLS);
- interaction between policy control and/or resource management elements.
- Service Level Agreements enforced by the border routers between networks.

The IP BS Manager may either communicate end-to-end through relevant signaling protocols or it can translate its QoS parameters (which use standard IP mechanisms) to map them to UMTS parameters. The mapping functions at the UE will translate QoS requirements from the application layer into either PDP context parameters or IP layer parameters (e.g. RSVP). At the SGSN side, the IP QoS parameters are mapped into UMTS QoS parameters and vice-versa.

4.3. Mapping QoS Parameters to the Air Interface

Many QoS schemes have been discussed to support different applications requirements in the WCDMA air interface. Examples of such mechanisms are discussed in [27], where hard QoS guarantees are provided for real-time applications, such as voice and video, using connection-oriented transmission, while non-real-time applications are supported on a best-effort basis through transmission-rate request access schemes that use the bandwidth left unused by other traffic. This type of approach is also discussed in [14].

In providing QoS for WCDMA systems, there are important mechanisms in place that will define admission control, load control and interference reduction to maximize system capacity. Those mechanisms need to be considered in dimensioning the system.

Higher layer QoS parameters, such as required data rates and delay guarantees are mapped onto the lower layers as different requirements of bandwidth, link quality (signal-to-interference ratio), priority modes, etc. In the same way, lower layer metrics such as the link load are used as decision criteria for admission control and interference reduction on those systems. Therefore, it makes sense

to concentrate the discussion of dimensioning the radio network on the variables that are most closely related to the air interface.

In this section, we start by discussing mechanisms available for Radio Resource Management. Then we concentrate on the main metric used by such mechanisms, which is the link load, to justify the importance of correctly characterizing the traffic offered to the system.

4.3.1. Radio Resource Management

In WCDMA systems, variable data rates are supported through the use of different spreading factors. Dedicated channels are usually allocated for real-time applications such as conversational and streaming classes, although full IP solutions for real-time applications should be supported in the future. Non-real-time applications are handled as scheduled packet-data over either common, shared or dedicated channels.

To maximize capacity and at the same time to provide QoS guarantees for the allocated services, Radio Resource Management (RRM) needs to be in place. The following are the main tasks related to RRM.

- *Power control:* Used in the downlink and in the uplink to maintain transmitting powers at the minimum requirements that will satisfy the targets expected for each class (eventually updated by an outer loop). Implemented on both subscriber and base-station ends of the air interface.
- *Handover:* To manage mobility, the system needs to support soft-handoff between neighbor cells.
- *Load control:* Based on received power or on throughput, load control determines how the system monitors the levels of interference to keep them under acceptable planned ranges and how it manages to bring the load back to acceptable levels when those ranges are exceeded.
- *Admission control:* Based on estimates of the increase in total interference power on the uplink or on total throughput being supported, the network estimates the noise rise (or throughput rise) that would result from accepting the requested RABs and decides whether to grant access and with which QoS parameters. Admission control is performed both on downlink and uplink.
- *Packet scheduling functions:* support non-real-time transmissions over packet data channels. It decides on bit rates and schedules transmissions to achieve optimum resource utilization while satisfying the required QoS parameters.

RRM functions are mostly implemented at the RNC, with the exception of load control, which is also monitored at the base station, and power control, which is performed on the base station and subscriber ends.

RRM functions can be implemented based on hard blocking or soft blocking. In either case, the system will stop accepting new connections at some point to not sacrifice existing connections and planned coverage. Typically, soft blocking mechanisms allow more efficient system utilization.

Although the subject of Radio Resource Management is a fascinating one that has been widely explored in recent research ([27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38] and [39]), it is not in the scope of this work to compare different methods or strategies of radio resource management, such as whether it is better to use call admission control versus bandwidth reservation. Instead, the focus of this research is on the importance of presenting the correct inputs to a system simulator so that, given a certain RRM strategy, the correct system capacity can be estimated and appropriate dimensioning of the system is possible.

Based on many published works on the subject ([28], [30], [31], [32], [33], [34], [36], [40] and [41]), it is observed that link load is widely used as the metric of choice to indicate a system's ability to accept more users and, therefore, it is used as the main criteria in algorithms for load control, admission control and packet scheduling. Moreover, it is the most widely used system metric of capacity, and therefore its target operating point is typically a design parameter.

The importance of this work comes from the fact that for a given resource management strategy and a given system target load, the total system capacity or throughput still depends on factors such as the user geographical distribution and the service mix. These latter variables are typically overlooked in many studies, where it is common to find important conclusions being drawn from homogeneous traffic distributions and arbitrarily determined service mixes.

In particular the importance of the service mix is illustrated in the next section.

4.3.2. Link Load and Noise Rise

One important step for verification of the system capacity in the system-level simulation, such as those performed later in this dissertation, is to include the same approach used in the real system for admission and congestion control. This includes the check on whether the system is exceeding the "target loading factors." In this sense, the loading factor in CDMA system is one of the most important design parameters and maximizing capacity achievable for a given loading factor is a design objective.

The load or loading factor expresses the ratio of received interference divided by the sum of received interference plus thermal noise [42].

$$\ell = \frac{I}{I+N} = \frac{I_0}{I_0+N_0}$$ **Equation 4-1**

The load factor is always smaller than 1 and tends to 1 as interference grows, i.e., as the system gets more loaded.

The **noise rise** indicates how much the interference raises the noise floor. It is therefore always greater than 1, and is typically expressed in dB. Therefore, it is equivalent to:

$$N_{rise} = \frac{I+N}{N} = \frac{I_0+N_0}{N_0} \quad \text{and} \quad N_{rise}dB = 10*\log(N_{rise})$$ **Equation 4-2**

Therefore, the noise rise is related to the loading factor by [1]:

$$N_{rise} = \frac{1}{1-\ell} \quad \text{and} \quad N_{rise}dB = -10*\log(1-\ell)$$ **Equation 4-3**

or

$$\ell = \frac{N_{rise}-1}{N_{rise}}$$ **Equation 4-4**

4.3.2.1. Downstream Load

The load of a sector on the downstream is calculated taking into account the load contribution of each user, which varies with the user location. It is given by:

$$\ell_{Dw} = \sum_{j=1}^{N}\left\{\nu_j \frac{\left(\frac{E_b}{N_0}\right)_j}{\frac{W}{R_j}}\left[\left(1-\alpha_j\right)+\left(\frac{o}{s}\right)_j\right]\right\}$$ **Equation 4-5**

In the equation, W is the chip rate, R_j is the data rate for user j and Eb/No_j is the required signal energy per bit divided by noise plus interference spectral density to achieve a certain quality of service for user j when connected at data rate R_j., ν_j is the activity factor for user j at the physical layer, α_j is the orthogonality factor in the downlink and $(o/s)_j$ represents other cell-to-same cell interference ratio as seen by user j at each specific location. The sum is calculated over all N users connected to the sector.

Equation 4-5 is a theoretical formula that expresses the overall interference received at the mobile location by evaluating the overall intra-cell interference and using a proportionality factor to estimate the additional fraction coming from other cells. As can be seen from that expression, the total received interference at one location depends on the combination of all services and data rates supported at a given moment.

4.3.2.2. Estimating Downstream Load (for Practical Purposes)

Equation 4-5 cannot be used to estimate the load when the intention is to use the load factor as a criteria for admission control for real systems or simulations. To do so we would have to assume or estimate the orthogonality factor at each location or use an average value for that variable. For the same reason, the ratio $(o/s)_j$, which also varies per location, would have to be measured by the mobile or estimated through simulations. Therefore, one way to estimate the downstream load is:

Equation 4-6

$$\ell_{Dw} \cong \sum_{j=1}^{N} \left\{ \nu_j \frac{\left(\frac{E_b}{N_0}\right)_j}{\frac{W}{R_j}} \left[\left(1 - \bar{\alpha}\right) + \bar{i} \right] \right\}$$

Here, $\underline{\alpha}$ is the mean (estimated or assumed) orthogonality factor for all locations served by the sector, and $\underline{i}$ is the average of $(o/s)_j$, i.e., represents average other-cell to same-cell interference ratio.

Another (more practical) way to estimate the downlink load factor in real systems or simulations is based on the total throughput of the sector, and is given by:

Equation 4-7

$$\ell_{Dw} = \frac{\sum_{j=1}^{N} R_j}{R_{\max}}$$

Here R_j is the bit rate of the j^{th} user and R_{max} is the maximum achievable throughput for the cell.

One additional option of load estimation uses the total sector transmit power, as shown below.

$$\ell_{Dw} = \frac{P_{Total}}{P_{\max}}$$ **Equation 4-8**

Where P_{total} is the total base station power being transmitted and P_{max} is the maximum base station transmission power allowed.

4.3.2.3. Upstream Load

The formula used to estimate the upstream loading factor for a multi-user multi-rate WCDMA system can be written as [42]:

$$\ell_{up} = \left(1+\frac{o}{s}\right)\sum_{j=1}^{N}\left[\frac{1}{1+\frac{W}{\left(\frac{E_b}{N_0}\right)_j R_j \nu_j}}\right]$$ **Equation 4-9**

In the equation, W is the chip rate, R_j is the data rate for user j and Eb/No_j is the required signal energy per bit divided by noise plus interference spectral density in order to achieve a certain quality of service for user j when connected at data rate R_j. ν_j is the activity factor for user j at the physical layer and o/s represents other cell to same cell interference ratio as seen by the sector receiver.

4.3.2.4. Estimating Upstream Load (for Practical Purposes)

When the intention is to use load factor as a criteria for admission control for real systems or simulations, there are different approaches to estimating the downlink load factor.

Because all parameters relevant to the uplink load calculation may be directly estimated or measured by the system (real or simulated), one could use Equation 4-9. Another way to estimate load for the upstream is through direct application of the definition of loading, given by Equation 4-1.

That means that the uplink load may be calculated directly by measuring the total power received at a sector, together with an assumed value for the noise floor. That relationship is rearranged below.

$$\ell_{up} = 1 - \frac{N_{floor}}{P_{Rx\_Total}}$$ **Equation 4-10**

Here, P_{Rx_Total} is the total measured power (noise plus interference) at the reception of a sector, and N_{floor} is an estimate of the thermal noise floor, typically calculated as:

$$N_{floor} = KTBN_f$$ **Equation 4-11**

K is the Boltzman constant (1.38×10^{-23}), T is the ambient temperature in Kelvin, B is the receiver bandwidth in Hertz and N_f is the receiver's noise figure. Instead of using Equation 4-11, N_{floor} may also be measured in real systems as the received power at the unloaded sector (no users active in the system).

Finally, for simulation purposes, the uplink load is typically estimated by computing the sum of all received powers at a sector from all users in the system (not only the users connected to that sector). This option is shown below.

$$\ell_{up} = 1 - \frac{N_{floor}}{N_{floor} + \sum_{k}^{all} P_{Rx_k}}$$ **Equation 4-12**

In the equation above, P_{Rxk} is the power received from the k^{th} user and is summed over all users in the system.

4.3.3. Capacity Dependence on User Mix

From inspection of Equation 4-6 and Equation 4-9, we notice that different scenarios of user mix (proportion of users in different classes) would result in approximately constant total throughput values for a given loading factor if the required $(Eb/No)_j$ thresholds are all set to the same value.

To demonstrate that observation, let us define the total data rate supported by one channel at one sector in a multi-user multi-rate scenario as:

$$R_{Total} = \sum_{All\_cell\_users} R_j \nu_j$$ **Equation 4-13**

Assuming Eb/No_j is the same for all users (j=1..N), let us replace Eb/No_j in Equation 4-6 and Equation 4-9 as a constant ρ where:

$$\rho = \left(\frac{E_b}{N_0} \right)_{required}$$ **Equation 4-14**

where *Eb/No* is assumed constant for all classes

Therefore, we can rewrite the **downlink load** given in Equation 4-6 as:

$$\ell_{Dw} \cong \left[\left(1 - \bar{\alpha}\right) + \bar{i} \right] \frac{\rho}{W} \sum_{All\_cell\_users} R_j \nu_j$$ **Equation 4-15**

$$\ell_{Dw} \cong \left[\left(1 - \bar{\alpha}\right) + \bar{i} \right] \frac{\rho}{W} R_{Total}$$

For the **uplink load**, given in Equation 4-9, let us also consider that in practice the right side of the denominator is much greater than 1, which is a reasonable assumption [42]. Therefore Equation 4-9 could be rewritten as:

$$\ell_{Up} \cong \left(1 + i\right) \frac{\rho}{W} \sum_{All\_cell\_users} R_j \nu_j$$ **Equation 4-16**

$$\ell_{Up} \cong \left(1 + i\right) \frac{\rho}{W} R_{Total}$$

Considering that the total loading factor is used as a criterion for admission control, it determines the capacity of the cell as a design parameter. Therefore, it is valid to say that if the Eb/No_j requirement for all classes were homogeneous, the total throughput capacity of a cell to achieve a desired loading factor would not depend on the particular mix of service classes.

However, in the more realistic scenario where different classes have different thresholds set for operation in terms of *Eb/No* required, the total load will be given by a weighted sum of the

different data rates times the required thresholds. This means that different total throughput per cell can be achieved for the same loading factor and network layout depending on the user mix that defines the demand offered to the system. This observation emphasizes once more the importance of appropriately describing the traffic demand offered to the system.

4.4. Summary

In this chapter, we discussed the concept of quality of service, with a focus on the physical layer for third-generation systems and how these QoS requirements map into higher layers. Mechanisms of radio resource management, such as power control and congestion control were also discussed. Next, we explored how system capacity (expressed as total throughput) depends on the user mix, i.e., the proportions of different services that are combined to achieve a certain total traffic demand, demonstrating the importance of an accurate characterization of different user mixes with their appropriate model parameters.

In the next chapter, we discuss simulation techniques that may be used to dimension multimedia systems, starting by a discussion of system-level simulation methods, then presenting some solutions to the problem of modeling traffic inputs for those simulations. We end the chapter with a discussion of the simulation process applied to multimedia wireless systems.

Chapter 5. Simulation of a Given User Mix

In this chapter, we discuss simulation techniques that may be used to dimension multimedia systems. We start the chapter by presenting the multiple types of system-level simulations, then we select one method (static simulations) and discuss its main limitations, namely the difficulty in correctly estimating the snapshot sizes for simulations. Next, we propose an approach to overcome that limitation, in a method described here as "multiplexing simulation." Results of multiplexing simulations performed for this research are presented and discussed. At the end of the chapter, a broader view of the whole system simulation process for multimedia wireless systems is presented and discussed.

5.1. Simulation Techniques for Mobile Networks

In this section, we revise the main concepts involved in system-level simulations and the methods used for simulating mobile networks today.

5.1.1. Static Simulation

The main characteristic of static simulations is that they do not model the time evolution of the simulated system, i.e., there is no correlation between consecutive simulation times.

A summary of the main characteristics of static simulation is given below [65].

- Outage is defined as signal-to-interference ratio falling below a given threshold.
- Handoff levels and signal and interference levels impact the analysis.
- Fading is simulated for each call in each instant of time (snapshot) independently, but its impact is evaluated instantaneously through probability (given that fading was randomly found for this call to be below the acceptable threshold, then this call is in outage).

As additional points, we highlight the following limitations of the static simulation method.

- Handoff thresholds impact QoS, however handoff timers and hysteresis cannot be perfectly modeled.
- Static simulations do not capture the time correlations of signal due to fading and mobility.
- These simulations cannot capture burst admission probabilities and performance during bursts.

Static simulations may widely vary in their degree of complexity. They may range from analytical modeling to more statistically representative approaches such as Monte-Carlo simulations. In the case of analytical modeling, coverage and capacity are determined through analysis assuming that mobile densities and fading models are analytically (although statistically) tractable. This option is not discussed in this chapter, since it is not considered to be a "true" simulation method.

For further discussion of the static simulation methods, we divide them in two main categories, statistical pixel simulation and Monte Carlo Simulation.

5.1.1.1. Statistical Pixel Simulation

This approach is referred to in [66] as "Statistical Snap-Shot" analysis. In that paper, it is proposed as a solution to evaluate performance of second-generation CDMA systems (single-class voice users) based on the traffic (mean number of active users) assumed per pixel on a traffic layer.

Instead of simulating calls from different combinations of users' locations, this method assumes that calls are generated from all pixels and then it scales the results according to the traffic values.

For each pixel, it obtains the BS-MS and the MS powers (more than one BS if in soft handoff) that would be transmitted on the downlink and uplink, respectively, for a call at that pixel and then it scales those powers according to the traffic declared for that pixel. For instance, if at a given pixel the traffic is 0.2 active users (Erlangs), the powers computed for the pixel would be 20% of that needed for an actual call.

The method is still iterative in the sense that the results have to converge, after power control adjustments are applied to powers on each iteration, based on total received interference at downlink and uplink. Although it is not mentioned in [66], the most natural metric to test convergence would be the sum of all powers transmitted and received at each sector.

5.1.1.2. Monte-Carlo Simulation (or Monte-Carlo Snapshots)

In this case, multiple instants of the system operation (snapshots) are simulated in the system, each of them with a different set of mobile locations. The number of users active in the snapshots should reflect an expected distribution of active users on the network area.

At each snapshot, multiple calls are generated and offered to the system, at locations randomly distributed and with density proportional to the expected geographical traffic density.

For each call, the relevant BS and MS transmitted powers are obtained, and power control algorithms are emulated so those powers are adjusted to achieve the target signal-to-interference or carrier-to-interference (C/I) ratios.

Using the powers calculated on the step above, new C/I ratios are obtained, and an iterative process may be necessary between this step and the step above until powers converge, similar to the method described in the statistical pixel simulation method.

One fundamental difference between the statistical pixel simulation and the Monte-Carlo simulation methods is that in the pixel simulations, outages tend always to happen on the last pixels tested for each sector, when load is already close to capacity limits, while on the latter, calls in outage are statistically averaged over many iterations. One way to overcome this limitation in the statistical pixel simulation and avoid biased outage maps is to evaluate pixels in random order, but in this case we are gradually transforming the pixel simulation into a Monte Carlo simulation, as multiple random scenarios would have to be evaluated for statistical significance.

One of the most severe limitations of static simulations is the difference in the concept of outage. While in static simulations outage is defined as the event of the C/I ratio falling below a certain threshold, in the case of dynamic simulations (discussed next) the concept of outage accounts for the amount of time a certain signal-to-interference ratio stays below that threshold. One possibility to overcome that particular limitation is to introduce this broader outage concept in static simulations as well, through the use of minimum duration outage probabilities, as discussed next.

For static simulations, fading effects are typically considered through the use of random variables whose values are drawn from the cumulative distribution function (CDF) curve of a given distribution. Typically those curves express the probability of a given signal level or signal-to-interference ratio falling below a specified threshold, given the variable's mean value and other distribution parameters. Typical distributions used in this context are the log-normal distribution (for shadow fading) and the Rayleigh distribution for non-line of sight multipath fading.

Through the use of the level crossings theory [67], it is possible to obtain different types of curves that express the probability of the signal level being above (or below) a certain threshold level for a period longer than a specified limit, given the signal's mean level and other distribution parameters relevant to that signal. If those types of curves are used instead of the typical signal strength distributions, they could enhance the results achieved with Monte Carlo static simulations when compared to typical static simulation, getting results closer to those of dynamic simulations. In case this type of approach is to be considered, different distributions should be used for different environments and mobile speeds, where each distribution represents the "probability that a signal

stays below a certain value for a time longer than a certain threshold T_{acc} for a signal at that environment and mobile speed" [67].

5.1.2. Dynamic Simulation

In dynamic simulations, there is a time correlation between consecutive simulation times, i.e., events may trigger other events. The most important consequence of this approach is that it is possible to model feedback loops that are present in many radio link control and RRM algorithms. Additionally, but not as importantly, dynamic simulations allow modeling user mobility along pre-defined user trajectories.

According to [62], dynamic simulators may be grouped in two categories.

- *Event-driven:* Events simulated in the system are the triggers of new events or tasks. The simulation time jumps from event to event.
- *Time-Driven:* Simulation time steps are fixed. Time steps are typically selected according to the network frame structure.

A summary highlighting the main advantages of dynamic simulations is given below [65].

- Outage is defined as signal-to-interference ratio falling below a given threshold for a given duration of time. This is the concept of "minimum duration outage," also presented in [67].
- Fading and outage durations are functions of fading environment and mobile speed.
- Handoff timers and handoff delays impact QoS.
- Capture performance at call level.

However, one main disadvantage of that type of simulation is that they are more complex to implement. More importantly, dynamic simulations may require impractically long execution times before statistically representative results are obtained.

5.2. Proposed Approach

In selecting a method to be used for simulating 3G systems, one has to be aware of the increased complexity of those systems when compared to original voice-only 2G systems. This is one strong argument for us to explore here the possibility of extrapolating the concepts of static simulation presented before to make them applicable for multimedia mobile systems.

One of the main challenges of this approach is the correct representation of the traffic offered to the system at each snapshot, when considering that different classes present different traffic behaviors.

After demand characterization is done, as described in Chapter 3, we are left with a set of user layers, with each layer consisting of a geographical distribution of users described by a given user profile. Since the network simultaneously supports multiple user types, the simulation of a certain scenario must take into account active users of all profiles together.

There are numerous papeers in the literature that describe static simulation processes and provide simulation results for the performance of WCDMA systems in response to a certain geographical distribution of simultaneously active users of a certain service class (single class simulations), such as in [44], [45] and [46], or of multiple service classes as in [47],[48], [49] and [50]. However, one important step to come before this process, which is typically not discussed in those papers, is how to estimate the number of simultaneously active users in a statistically representative manner so that the simulation results reflect the response of the system to a certain traffic forecast. This is one of the main contributions of the present work, specifically discussed in this section. Figure 5-1 and Figure 5-2 illustrate this point.

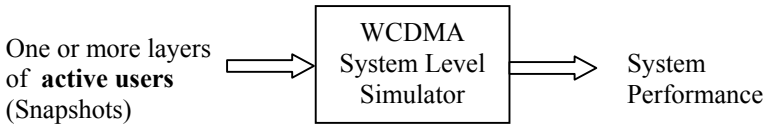


Figure 5-1. Typical approach to simulation for WCDMA.

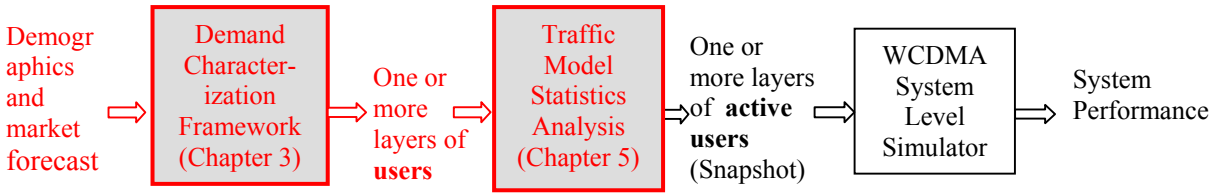


Figure 5-2. Contribution of the current discussion (shaded).

We emphasize here the importance of this contribution in that it allows proper dimensioning and, therefore, a more reliable business case analysis for operators planning to migrate to multimedia technologies, by properly reflecting the impact of the expected demand on the system. For the same reason, the system design will benefit from a proper rollout plan for deployment, as well as efficient resource allocation based on expected demand and growth. Furthermore, with reliable means of

characterizing the expected traffic offered to the system, the design may be optimized by iterations in the network configuration (e.g. base-station locations, antenna models, and tilts), as well as fine adjustment of many configurable parameters of operating networks, such as handoff thresholds, soft-handoff parameters, neighbor lists, target loads, and target power limits.

Of the benefits mentioned above, resource allocation is an important and fundamental phase for all interference-limited systems, since proper performance analysis is needed in define how much spectrum (how many channels) is needed at each sector and allow topology iterations in case the allocated spectrum is not sufficient. This performance analysis is completely dependent on traffic intensity and location. In particular, for those systems that require frequency planning, such as GPRS and EDGE, one more benefit at this phase is to use this traffic-sensitive performance analysis in the decision-making phase of selecting the best frequency plan.

The appropriate reproduction of the traffic load in terms of statistically representative snapshots depends both on an appropriate traffic **demand characterization,** which defines how users of each service are geographically distributed (discussed in chapter 3) and on the appropriate **statistical modeling** of the traffic distributions in order to capture the correct distributions of simultaneously active users at a given time. That latter step would not be needed in case dynamic simulations were to be used, as those distributions would be captured through the simulation process itself.

The simulation of multiple snapshots for a given set of user layers and the statistical combination of results obtained from those snapshots allows the estimation of performance and quality of service achieved in the system for a certain offered traffic load. The approach to obtain the snapshot figures in a way that is consistent with the traffic profile described is one of the key challenges of this phase.

The proposed framework consists of de-coupling the treatment of temporal randomness (generation of traffic bursts) and spatial randomness (geographical distribution and movement) for the purposes of system simulation.

The main objective of this decoupling is to reduce complexity and consequently processing effort in simulating third-generation communication systems, which are already complex enough.

To illustrate the challenges of the problem described and demonstrate the value of the solution being proposed here, we first illustrate the discussion by presenting below an example of how the equivalent problem is typically dealt with in voice networks.

5.2.1. Voice Example

Suppose we want to estimate how many channels are needed per sector in a typical 2G voice system, given a geographical distribution of users and the description of their behavior (Erlangs

per user). Furthermore, consider that the distribution is not absolute, in the sense that it does not "glue" the users to geographic locations. Instead, it provides the proportion of users distribution for different locations. A full simulation of such a system would include at least the following steps.

Strategy A:

1. For each user, at its geographical location, reproduce randomly (following a Poisson process) the calls generated for that user.
2. For each time slice (assuming a given time unit), estimate how many users would be connected to each sector. Use propagation analysis to estimate for each active user which site would be the "best-server."
3. Advance time, generate new events (end some bursts, begin new bursts), and proceed with sector load estimation.
4. For more detailed treatment, in advancing time user mobility could be reproduced as well.
5. At the end, process statistics of load offered to each sector.

Although the procedure above may sound like the "right" thing to do, this process could become extremely complex and slow for a large number of users and sectors. Moreover, if the decision criteria is not as simple as a best-server analysis, the simulation may become unreasonably complex. Also, as will be demonstrated in the next section, long characterization periods are needed to achieve statistically representative results in such systems. All of these problems together could render the simulation process infeasible or at least impractical.

In contrast, a Monte Carlo static simulation for the same voice system works as if we pretended to be observers in the sky taking stroboscopic pictures of the system. For each "picture," you need to simply distinguish the users that are active from those that are not. If all users on the ground can be assumed to have similar profiles (i.e., Erlangs per user), it is possible to reproduce the distribution of total active users over many pictures without having to reproduce the real burst distributions for each user. This could be done using the Poisson distribution, which is analytically described in a simple format. Therefore, a much simpler procedure that would provide similar results as those described above would be as follows.

Strategy B:

1. For each iteration, simulate a Poisson random variable for a traffic volume equivalent to the total traffic in the system. This random variable will represent the number of active users, K, in the "picture" corresponding to that iteration (snapshot). This step deals with the temporal randomness of the users.
2. For the snapshot calculated in that iteration, randomly activate K users in the traffic grid (which has a total of N>>K users). This random activation should follow the probability density of users in the geographical grid. This step corresponds to the spatial randomness of the users.
3. For those active users, identify the serving sectors.
4. Proceed to the next iteration.
5. At the end, process statistics of load offered to each sector.

The procedure explained in Strategy B in the example is not a new methodology proposed in this work. It is actually a procedure typically used for simulation of cellular voice systems due to its simplicity.

However, for multi-class systems (such as 3G), it has been a trend in the literature to deal with traffic in a rather different way that either follows the principles of Strategy A with a full dynamic simulation (e.g. OPNET Modeler™ simulations) or simply bypasses the challenge of estimating the size of the snapshot by assuming that all classes follow Poisson processes, whose distributions are analytically described, as in [23] and [42].

The objective of this work is to propose a methodology that allows static simulations to be performed for 3G systems while still appropriately modeling the traffic behavior of the multiple types of service classes served by the system.

5.3. Snapshot Generation

The proposed framework suggests a solution for multi-class systems in which the equivalent of Strategy B could be used without the dangerous assumption of Poisson processes for all data classes. It consists of initially performing the "multiplexing simulation" that will provide practical distributions of the number of active users for non-Poisson classes. These distributions will be used in the same way as those analytically available for Poisson classes in the process of randomly estimating snapshots sizes. Figure 5-3 illustrates the procedure proposed in this work.

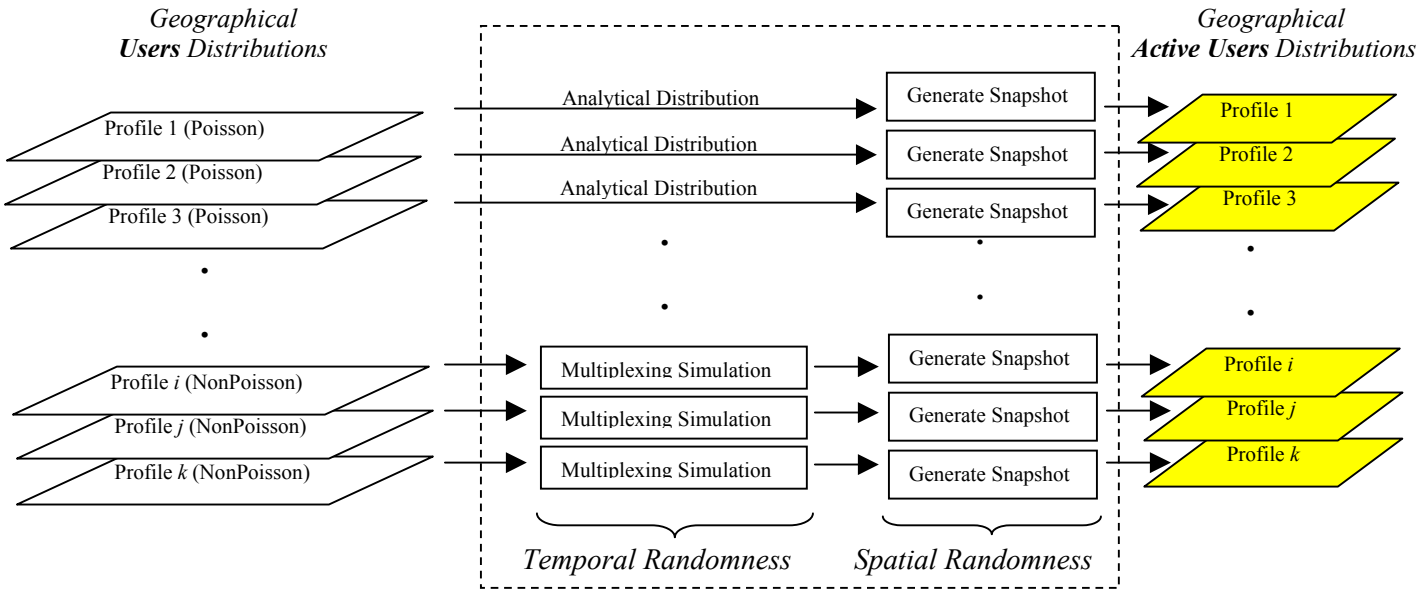


Figure 5-3. Snapshot solution for multi-class traffic layers including Poisson and Non-Poisson profiles.

The procedure illustrated constitutes the snapshot generation block (dashed box), which will deal with both temporal and spatial user randomness.

The snapshot generation consists of determining the number of active users of each user profile based on the geographical layer distribution and on the traffic parameters associated with that profile. The main traffic parameters associated with users of a certain profile that will allow such calculation are:

- session inter-arrival time distribution,
- session length distribution,
- burst inter-arrival time distribution, and
- burst length distribution.

The session level expresses the rate at which connections arrive in a connection-oriented service or, in a more informal definition, the rate of "awakeness" of the users of a certain profile. A burst here is considered a sequence of packets small enough such that packet lengths and packet inter-arrival times within the burst are constant. Therefore, the statistical nature of the received data is taken into account at the burst level. Although, in reality, there is also a non-uniform inter-arrival distribution at the packet level, we assume here that its randomness may be modeled as an aggregate burst length distribution. To illustrate the distinction of session and burst levels in this discussion, consider for instance a service that stays "always ON" for hours, such as a wireless IP connection for

web applications. Even though the connection is established, the user is not actually using resources in the network while it is idle. Each time the user actually uses the connection, for instance by starting to browse a web page, a session has started. The exchange of data during that connection (download data, reading time, sending info) constitutes the bursts. If the user closes the browser or stays idle for a long time, the system will consider it idle again and that is the end of the session.

With regards to the aforementioned variables, several studies have presented specific distributions that best describe different types of multimedia applications [15]. Some comments on typical distributions are presented here.

- *Voice:* Typically, for voice applications, inter-arrivals times and service times are exponentially distributed. Also, for voice applications, the concept of session and burst are coincident, as each call is at the same time considered a "burst of voice" and a "session" (in a circuit-switched connection).
- *Data:* For data applications, session distributions are typically Poisson, with characterization parameters (mean inter-arrival rate and mean session length) that differ depending on the application. Burst inter-arrival distributions also depend on the type of application.
    - For FTP and Telnet connections, burst inter-arrivals are said to be exponential (chapter 15 of [15]).
    - For TCP, connections are described as having heavy-tailed interarrival distributions. Typical distributions suggested for TCP are Weybull and Pareto (Chapter 15 of [15]).

Summarizing what has been discussed above, the proposed framework for the task of snapshot generation consists of the following steps.

1. Simulate the burst and session arrival processes for the users of a certain traffic profile. This preliminary simulation step, which is a separate task from the actual WCDMA network simulation block, does not take into account the geographical distribution and it aims only to determine the statistical multiplexing effect obtained for that traffic profile. For clarity in further discussions, these simulations will be referred to here as "multiplexing simulations." This multiplexing simulation should be performed for those profiles that cannot be analytically evaluated, which is the case for non-Poisson processes.
2. After simulation of a certain profile is concluded, the outcome is the distribution of active users for that profile as a function of the total number of users.

3. One set of simulations should be executed for each traffic profile. The profiles for which the desired distributions are known do not need to be evaluated.
4. After all multiplexing distributions are obtained, the snapshots for each profile are generated from those distributions, i.e., the number of active users of a certain profile is a random variable that follows its multiplexing distribution.
5. Once the number of active users is known, they are randomly distributed on the spatial plane following the proportions of the geographical user distribution.

5.3.1. Multiplexing Simulator

As part of this research, a "multiplexing simulator" has been created. This simulator accepts different types of distributions and their parameters as inputs, and provides graphical and file outputs of the "multiplexing distributions." An illustration of the interface created for this simulator is shown in Figure 5-4.

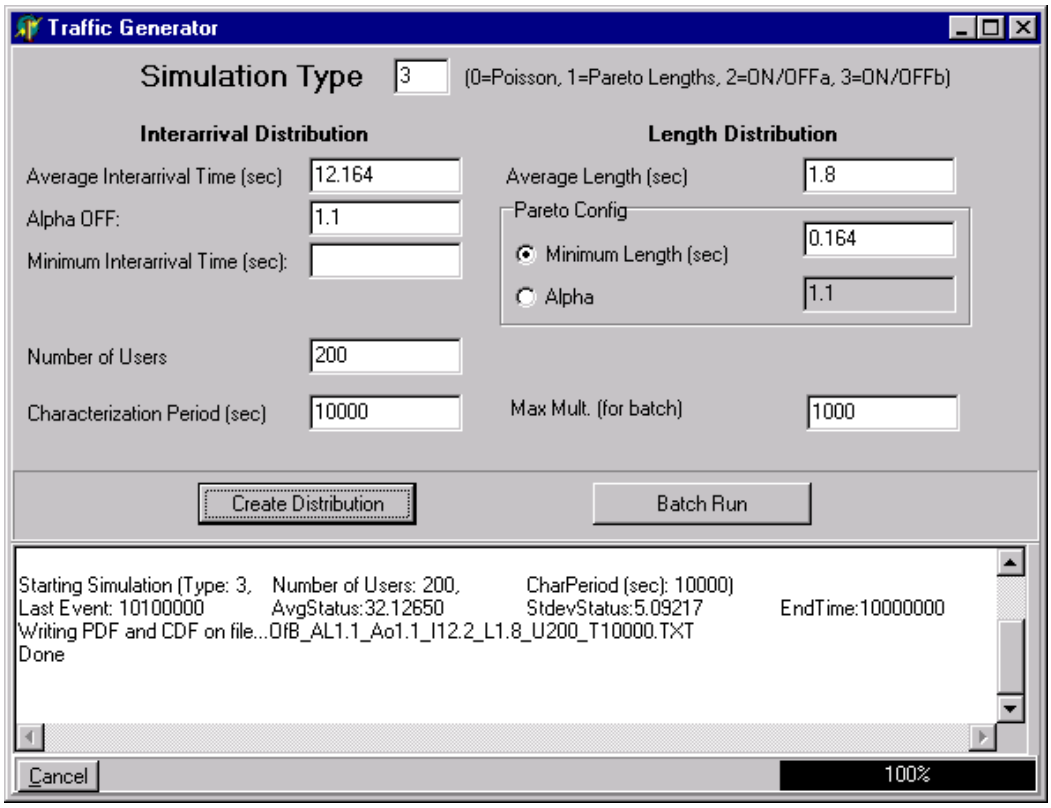


Figure 5-4. Interface of multiplexing simulator developed for this study.

The simulator generates as output both text and graphical files, which include the probability mass function (PMF) and cumulative distribution function (CDF) curves for the number of simultaneously active users. Those are illustrated in Figure 5-6 and Figure 5-5.

Traffic Type(0=Exp,1=Par,2=ON/OFFa,3=ON/OFFb)	3		
Average Interarrival Time (sec)	12.164		
Alpha OFF (if ON/OFF):	1.1		
Average Length (sec)	1.8		
Minimum Lenght=K (if Pareto)(sec)	0.164		
Alpha Lenght (if Pareto):	1.10024		
Characterization Period (sec)	10000		
Number of Users	200		
Average Status	32.1265		
StDev Status	5.09217		
NSim	Accum	PDF	CDF
0	0	0	0
1	0	0	0
2	121	1E-05	1E-05
3	15	1E-06	1E-05
4	59	6E-06	2E-05
5	22	2E-06	2E-05
6	6	6E-07	2E-05
7	15	1E-06	2E-05
8	47	5E-06	3E-05
9	7	7E-07	3E-05
10	3	3E-07	3E-05
11	16	2E-06	3E-05
12	4	4E-07	3E-05
13	14	1E-06	3E-05
14	150	1E-05	5E-05
15	798	8E-05	1E-04
16	2292	0.0002	4E-04
17	5282	0.0005	9E-04
18	9770	0.001	0.002
19	18550	0.0018	0.004
20	36493	0.0036	0.007
21	62890	0.0062	0.014
22	103774	0.0103	0.024
23	153917	0.0152	0.039
24	219913	0.0218	0.061
25	308001	0.0305	0.091
26	405058	0.0401	0.131
27	516667	0.0512	0.183
28	610131	0.0604	0.243
29	685995	0.0679	0.311
30	756680	0.0749	0.386
31	787559	0.078	0.464
32	786783	0.0779	0.542
33	768384	0.0761	0.618
34	716154	0.0709	0.689
35	638533	0.0632	0.752
36	556844	0.0551	0.807
37	468386	0.0464	0.853
38	383144	0.0379	0.891
39	306160	0.0303	0.922
40	234946	0.0233	0.945
41	172626	0.0171	0.962
42	127054	0.0126	0.975
43	89038	0.0088	0.983
44	64018	0.0063	0.99
45	41069	0.0041	0.994
46	25279	0.0025	0.996
47	16574	0.0016	0.998
48	9455	0.0009	0.999
49	5196	0.0005	0.999
50	2372	0.0002	1
51	1414	0.0001	1
52	696	7E-05	1
53	295	3E-05	1
54	408	4E-05	1
55	232	2E-05	1
56	140	1E-05	1
57	127	1E-05	1
58	157	2E-05	1
59	67	7E-06	1
60	25	2E-06	1
61	15	1E-06	1

Figure 5-5. Text output sample.

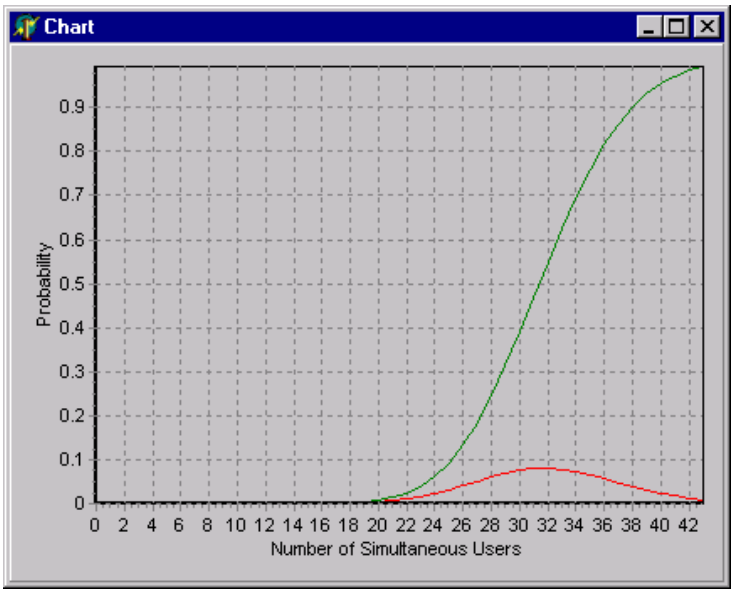


Figure 5-6. Graphical output sample.

The following assumptions are used in all simulation scenarios:

- Simulations are performed only at the burst level.
- Distributions are applied to each user independently, i.e., from each user, new bursts arrive at the server at random times, following the specified "inter-arrival time distribution." Each burst has a randomly generated burst length following a specified "length distribution."
- The following parameters are defined for "inter-arrival time distribution."
  - Option for exponential or Pareto distribution
  - Average inter-arrival time (always required)
  - For Pareto case, additional choice of one of the following:
    - Alpha or α (parameter for Pareto distribution)
    - Minimum possible value assumed by the variable
- The following parameters are defined for "length distribution."
  - Burst length distribution type (choice of exponential or Pareto)
  - Average burst length (always required)
  - For Pareto case, additional choice of one of the following:
    - Alpha or α (parameter for Pareto distribution)
    - Minimum possible value assumed by the variable

The simulations performed aimed to verify the impact of using different distributions while keeping the same mean values for inter-arrival rates and burst length. It is assumed that by equalizing

the means, the simulations are producing the same average volume of traffic (in bits) over a certain simulation period, i.e., equivalent average load is imposed on the system. We use the simulation results then to verify how many simultaneous channels would be active in the system when responding to the traffic offered.

The distributions simulated in the tool as well as results are presented in the next subsections. Additional distribution models could easily be added to the current implementation.

5.3.2. Multiplexing Simulation: Exponential Interarrivals – Exponential Lengths

For user behavior descriptions, the M/M/∞ process is used to describe users that have both inter-arrival and service times following exponential distributions (Poisson process). For those processes, the distribution of active users has its PMF given analytically by [3]:

$$P[X = k] = \frac{1}{k!}\left(\frac{N\lambda}{\mu}\right)^{k} e^{-\frac{N\lambda}{\mu}}$$ **Equation 5-1**

X is the random variable for the number of simultaneously active users, N is the total number of users, λ is the mean inter-arrival rate and $1/\mu$ is the mean service time. Notice that for Poisson processes, the distribution of active users is completely defined by the mean parameters given for inter-arrival rate and service time. Moreover, the expected value of active users in the system in a M/M/∞ process is given by:

$$E[X] = \frac{N\lambda}{\mu}$$ **Equation 5-2**

And the standard deviation is of the variable X is given by:

$$STD[X] = \sqrt{\frac{N\lambda}{\mu}}$$ **Equation 5-3**

The Poisson option was created in the tool for consistency checking and validation purposes only, as its results can be determined analytically.

We are particularly interested in observing the expected value of the number of active users for other distributions that are not Poisson, and comparing those to the results for the Poisson process. This is important because most of the work currently done in dimensioning wireless 3G systems assumes Poisson processes even for data traffic.

5.3.3. Multiplexing Simulation: Exponential Interarrivals – Pareto Lengths

For the more generic case where, for instance, the service time is heavy tailed (M/G/∞), the expected number of active users still converges to the same value as in the M/M/∞ process (Chapter 15 of [15]).

Although we do not formally prove it in this work, we also expect that the distribution itself coincides for both M/M/∞ and M/G/∞ processes, based on the following reasoning.

i. The Erlang formula, which was developed for the M/M/c/c process, applies also for M/G/c/c [3], i.e., it does not depend on the service time distribution and holds true as long as the inter-arrivals follow an exponential distribution.
ii. Therefore, for calculating the blocking probability, the Erlang formula could be used to estimate the multiplexing effects of aggregated users.
iii. Following the same reasoning and considering that the M/M/∞ process corresponds to the M/M/c/c process when $c \to \infty$ and that the M/G/∞ process corresponds to the M/G/c/c process when $c \to \infty$, we conclude that the distributions for the number of simultaneously active users is the same for both processes.

Therefore, we should expect that simulations based on exponential inter-arrival times and Pareto burst lengths should converge to the same distributions as the Poisson simulations, both converging to the formulas given in the previous section for the M/M/∞ process.

Still, to verify the conclusions discussed above, multiplexing simulations were performed for the exponential interarrivals and Pareto lengths scenario. Those results are presented in Figure 5-7 and Figure 5-8. The following parameters were used in the simulations:

- Mean burst Inter-arrival time: 12.164 seconds
- Mean burst Length: 1.8 seconds

- Distributions simulated:
  - Exponential-arrivals and exponential lengths ("Exp" legend)
  - Exponential-arrivals and Pareto lengths
    - $\alpha = 1.1$ (strong self-similarity) ("Lpr_AL1.1" legend)
    - $\alpha = 1.2$. ("Lpr_AL1.2" legend)
    - $\alpha = 1.3$ (weak self-similarity) ("Lpr_AL1.3" legend)
- Characterization period: Multiple simulations with increasing characterization periods ranging from 3600 to 36 million seconds.

Scenario 1:

- Number of users: 200

Scenario 2:

- Number of users: 1000

We can draw the following observations from the results presented in Figure 5-7 and Figure 5-8.

1. The exponential-arrivals and exponential-length (Exp-Exp) simulation provided results that match those expected from theory ($M/M/\infty$) even for short characterization periods.
2. Although it was already mentioned that the results for the exponential-arrivals and Pareto-Length (Exp-Par) should follow the same distribution obtained for the Exp-Exp simulations, it was observed that results yielded a smaller mean number of simultaneous active users for short characterization periods.
3. As the characterization period grows the results for the Exp-Par simulation get closer to those of the Exp-Exp simulation.
4. As α grows from 1.1 to 1.3 (lengths become less self-similar), the behavior also gets closer to that of the Exp-Exp simulation.
5. Results for both 200 and 1000 users simulations (Scenarios 1 and 2) followed the same patterns mentioned above. The results for 1000 users typically evolved more smoothly than those for 200 users as the characterization periods grow.

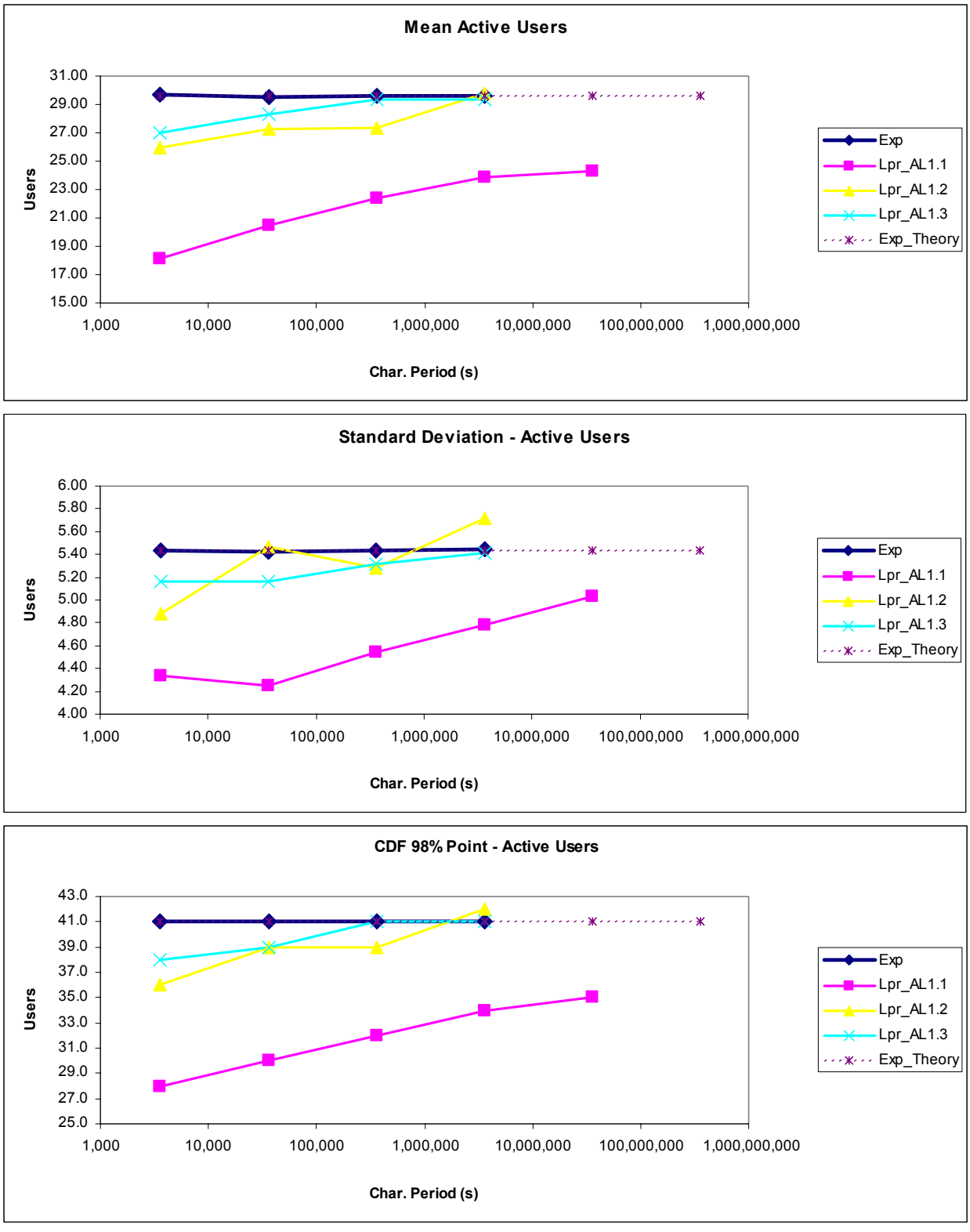


Figure 5-7. Simulation results for exponential arrivals, Pareto lengths with 200 users (Scenario 1). Mean inter-arrival time = 12.164 seconds, mean length = 1.8 seconds.

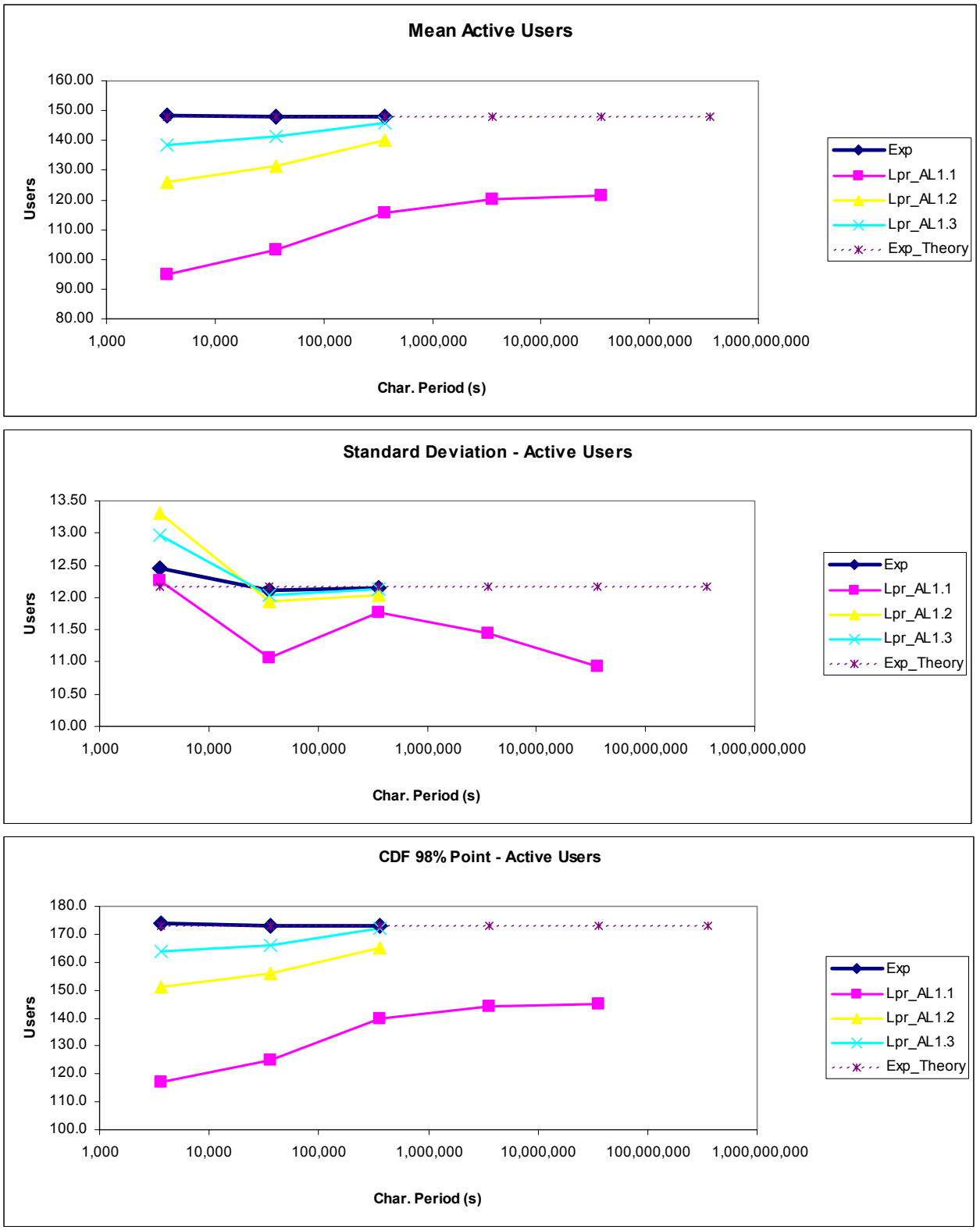


Figure 5-8. Simulation results for exponential arrivals, Pareto lengths with 1000 users (Scenario 2). Mean inter-arrival time = 12.164 seconds, mean length = 1.8 seconds.

If we choose the longer characterization periods as the most representative result for each scenario (long-term results), it is even more clear that mean and 98% results get lower as the length distributions become more self-similar (smaller α). It is also interesting to notice that, as expected, these distributions all tend to the analytical results obtained from the analysis of an M/M/∞ system as the number of points simulated grows. Moreover, they consistently approach that distribution from lower values, i.e., the mean and standard deviation values grow slowly tending to the M/M/∞ distribution. This effect is expected because a few large bursts that may not have already happened at this point of the simulation will be responsible for raising the statistics at some point. This type of behavior is expected on simulations of heavy-tailed random variables, as described in Chapter 3 of [15].

The long-term simulation results for each distribution type are presented in Table 5-1 and Table 5-2, and in Figure 5-9.

Table 5-1. Results for Scenario 1 with 200 Users

Scenario 1 200 Users	Mean	StDev	CDF 98%
Poisson (Theory)	29.6	5.4	41
Poisson (Simulated)	29.6	5.4	41
Exp-Par 1.3	29.4	5.4	41
Exp-Par 1.2	29.8	5.7	42
Exp-Par 1.1	24.3	5.0	35

Table 5-2. Results for Scenario 1 with 1000 Users

Scenario 2 1000 Users	Mean	StDev	CDF 98%
Poisson (Theory)	148.0	12.2	173
Poisson (Simulated)	147.9	12.2	173
Exp-Par 1.3	145.9	12.1	172
Exp-Par 1.2	139.9	12.0	165
Exp-Par 1.1	121.6	10.9	145

Long-Term results for 200 Users **Long-Term results for 1000 Users**

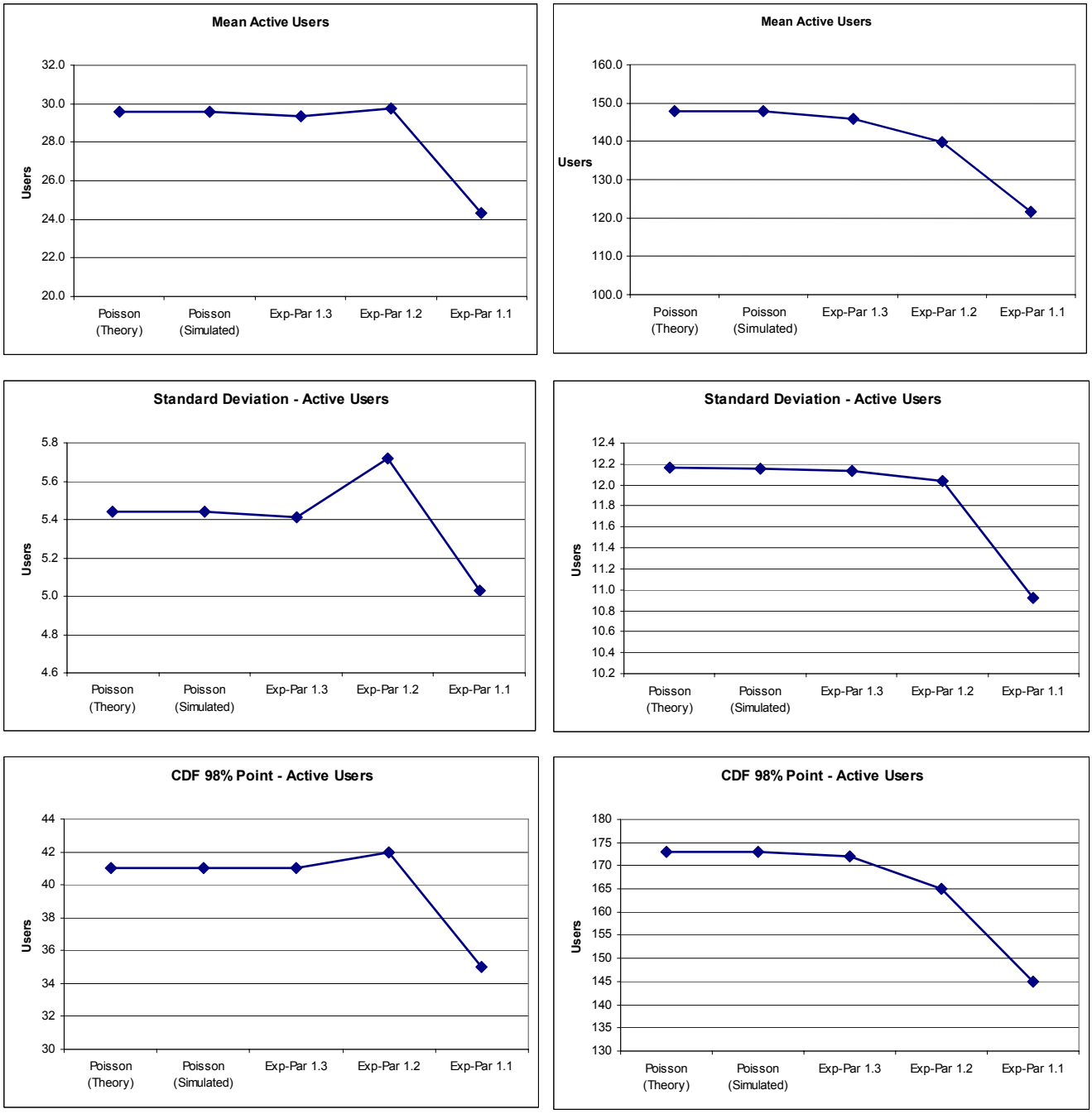


Figure 5-9. Long-term results for Exp-Par simulations (Scenarios 1 and 2).

5.3.4. Multiplexing Simulation: ON/OFF Model

In [13] web traffic is described as a superposition of ON/OFF processes. Both ON and OFF times distributions are described as heavy tailed, with the ON distribution ($\alpha \cong 1.2$) being much heavier-tailed than the OFF distribution ($\alpha \cong 1.5$). The same reference also mentions that the self-similar characteristic of a source is determined by the distribution that is heavier-tailed (ON distribution in this case).

We have simulated the following scenarios, using an ON/OFF traffic model.

- Mean burst inter-arrival time: 12.164 seconds
- Mean burst length: 1.8 seconds
  - Mean "ON" length =1.8 seconds
  - Mean "OFF" length = 12.164 - 1.8 = 10.364 seconds
- Distributions simulated:
  - Exponential-arrivals and exponential lengths ("Exp" legend)
  - Exponential Arrivals and Pareto lengths:
    - $\alpha_{ON} = \alpha_{OFF} = 1.1$ (strong self-similarity) ("OfB_AL1.1" legend)
    - $\alpha_{ON} = \alpha_{OFF} = 1.2$ ("OfB_AL1.2" legend)
    - $\alpha_{ON} = \alpha_{OFF} = 1.3$ (weak self-similarity) ("OfB_AL1.3" legend)
- Characterization period: Multiple simulations with increasing characterization periods ranging from 3600 to 36 million seconds.

Scenario 3:

- Number of Users: 200

Scenario 4:

- Number of Users: 1000

The results obtained for both scenarios are presented next in Figure 5-10 and Figure 5-11.

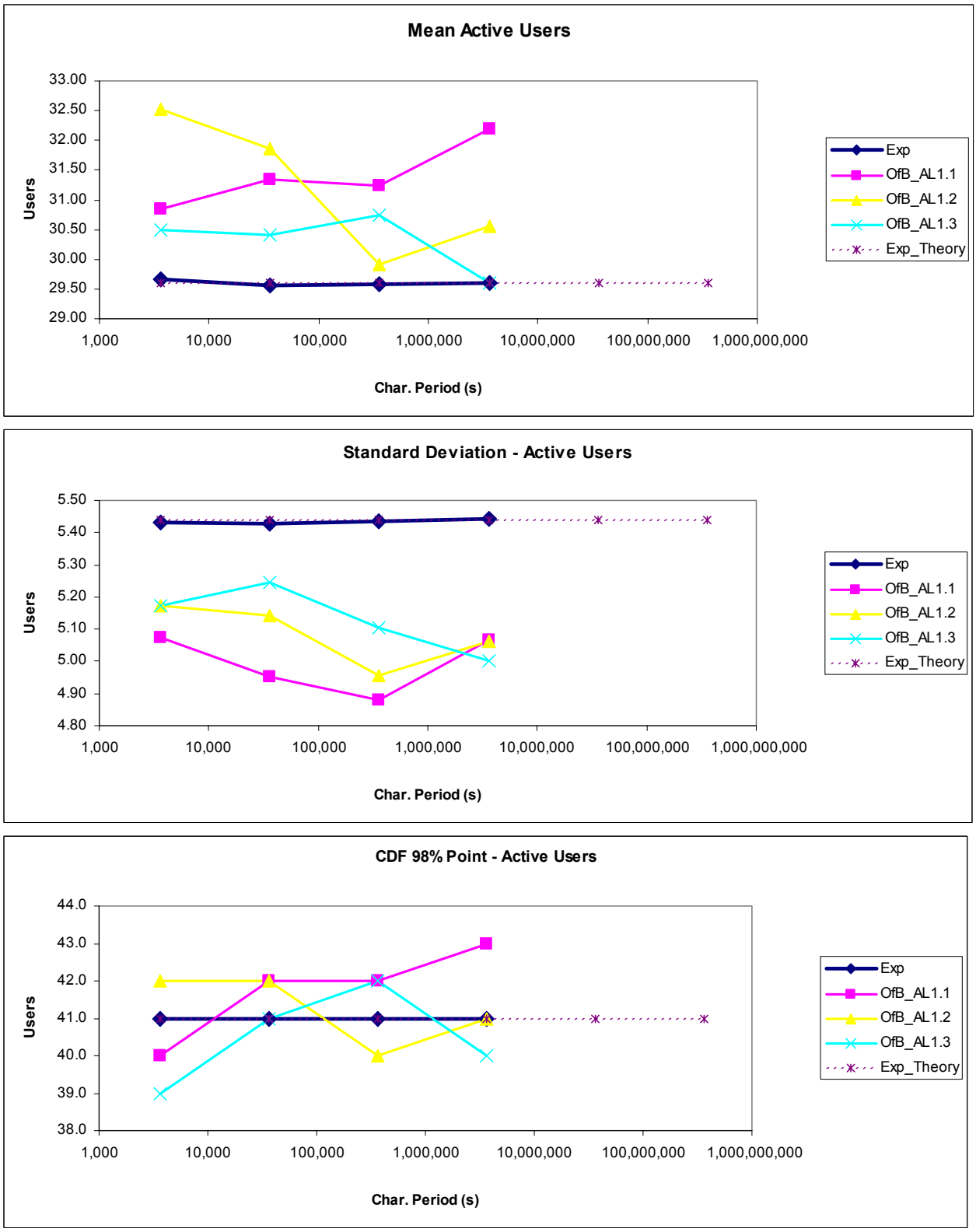


Figure 5-10. Simulation results for ON/OFF with 200 users (Scenario 3). Mean inter-arrival time = 12.164 seconds, mean length = 1.8 seconds.

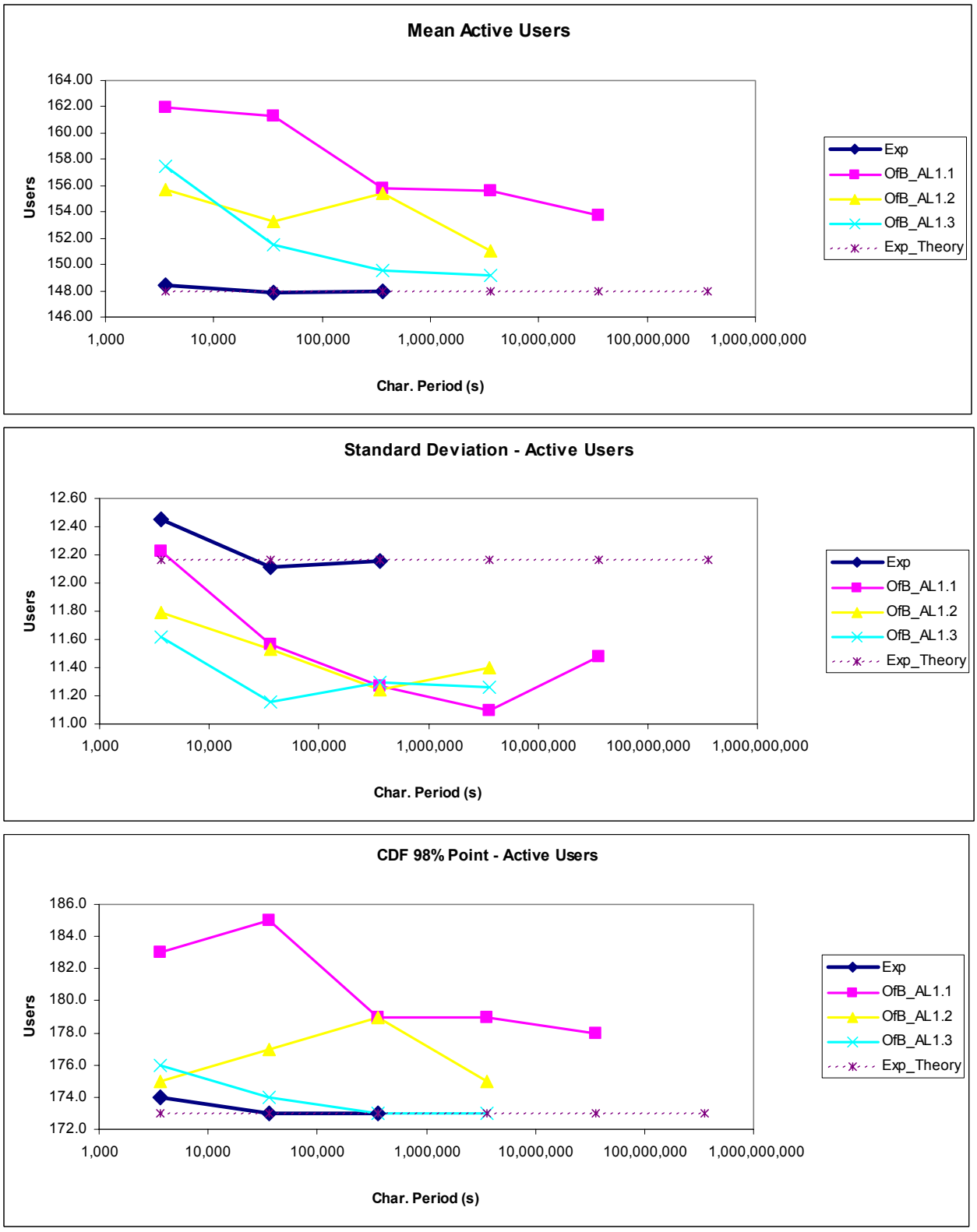


Figure 5-11. Simulation results for ON/OFF with 1000 users (Scenario 4). Mean inter-arrival time = 12.164 seconds, mean length = 1.8 seconds.

We can draw the following observations from the results presented above.

1. Results from the ON/OFF simulations have a consistently larger mean number of simultaneous active users compared to the results of the Exp-Exp simulations.
2. Results from the ON/OFF simulations have a consistently smaller standard deviations compared to the results of the Exp-Exp simulations.
3. As the characterization period grows, the results get closer to those of the Exp-Exp simulation, but still do not seem to tend to those results.
4. As α gets smaller (distributions become more self-similar), the mean number of active users gets larger, as well as the 98% CDF point (i.e., the number of users at which the CDF curve reaches 98% probability).

One important conclusion from the observations above is that as the user behavior becomes more self-similar, even if the traffic volume offered to the system is kept the same, more capacity is required to handle that traffic for the same QoS. In other words, while the mean inter-arrival times and lengths are kept the same, i.e., the equivalent throughput requested from the system is the same, more resources have to be available at the system if congestion is to be avoided.

The long-term simulation results for each distribution type are presented in Table 5-3 and Table 5-4, as well as in Figure 5-13.

Table 5-3. Results for Scenario 3

Scenario 3 200 Users	**Mean**	**StDev**	**CDF 98%**
Poisson (Theory)	29.6	5.4	41
Poisson (Simulated)	29.6	5.4	41
ON/OFF 1.3	29.6	5.0	40
ON/OFF 1.2	30.5	5.1	41
ON/OFF 1.1	32.2	5.1	43

Table 5-4. Results for Scenario 4

Scenario 4 1000 Users	Mean	StDev	CDF 98%
Poisson (Theory)	148.0	12.2	173
Poisson (Simulated)	147.9	12.2	173
ON/OFF 1.3	149.1	11.3	173
ON/OFF 1.2	151.0	11.4	175
ON/OFF 1.1	153.7	11.5	178

Suppose, for instance, that we want to dimension the system to handle 98% of all offered traffic. If the mean offered traffic were constant, i.e., if users would offer traffic to the system in an organized (predictable) manner, it would suffice to dimension the system to handle exactly the mean number of active users at a time (30 channels in Scenario 3, 148 channels in Scenario 4). However, as user behavior becomes random, more capacity needs to be offered by the system to satisfy the 98% success point. In case this random behavior follows a Poisson process, that "extra capacity" needed to accommodate the randomness would require 41-30 =11 additional channels for 200 users or 173-148 = 25 additional channels for 1000 users. Furthermore, for users that follow self-similar behavior, that capacity should be even larger to accommodate extra unpredictability (due to less multiplexing compression of self-similar traffic). That extra capacity would mean 43-41 = 2 more channels in Scenario 3 (200 users) or 178-173 = 5 more channels in Scenario 4 (1000 users). This discussion is summarized below (Figure 5-12).

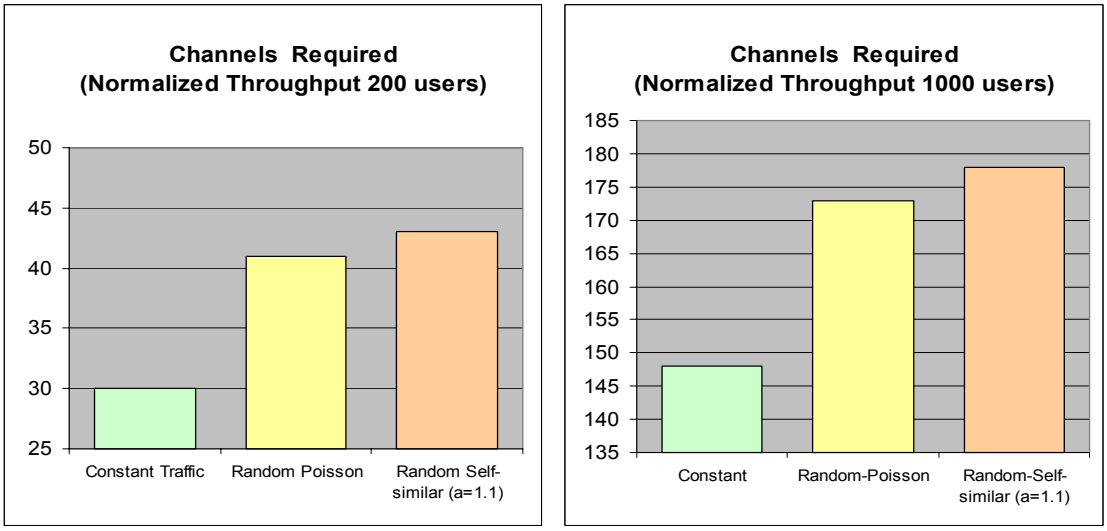


Figure 5-12. Increase in capacity needed as user unpredictability grows.

Long-Term results for 200 Users **Long-Term results for 1000 Users**

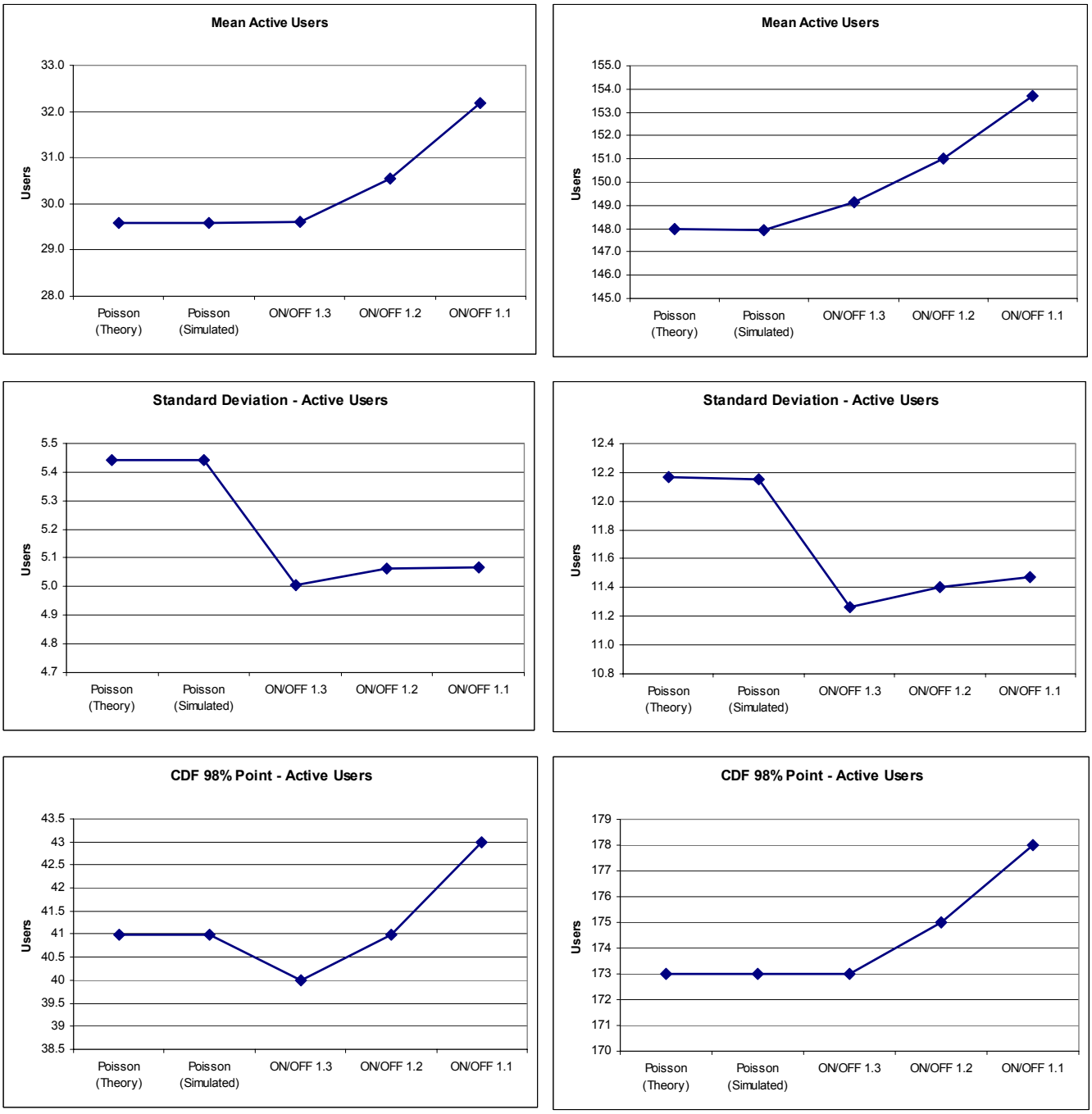


Figure 5-13. Long-term results for Exp-Par simulations (Scenarios 3 and 4).

5.3.5. Sensitivity and Reusability of Multiplexing-Simulations

In this section, we discuss the sensitivity of the multiplexing simulations presented before, with regards to sample size (number of users) and simulation time. We also discuss the possibility of reusability for such simulations, in a way that results may be tabulated or fitted to curves so lookup is possible during system-level simulations.

5.3.5.1. Sensitivity of Multiplexing-Simulations

As presented in the previous sections, multiplexing simulations have been performed for multiple scenarios, which included multiple dimensions. The sensitivity of those simulations with respect to the traffic model, to the number of users (sample size) and to the characterization period are evaluated in this section.

Sensitivity to Traffic Model:

For the analysis of sensitivity to the traffic model, the following simulation options were grouped and compared:

- exponential arrivals-exponential lengths (Exp-Exp),
- exponential arrivals-Pareto lengths (Exp-Par) and
- ON-OFF with Pareto lengths for both ON and OFF length distributions (Par-Par).

Based on the results presented, it is seen that the arrival process model assumptions have a greater impact in the results than the length process. In other words, if the arrival process is kept the same, as was the case between simulations on groups Exp-Exp and Exp-Par, the results of active users distributions tend to be similar. This observation from the simulations results has also been justified based on traffic theory, as presented in Section 5.3.3. As can be observed in that section as well, the impact of using different length distributions is mainly reflected in the time it takes for the results to converge to expected values, i.e., as α gets smaller in Figure 5-7 and Figure 5-8, the characterization period needed for the simulations to converge grows significantly.

If the arrival process is changed, as is the case between groups Exp-Par and Par-Par, the results suffer a much greater impact, as the total number of active users in the system effectively grows with the increase in self-similarity (as α gets smaller). This can be clearly observed in the simulation results presented in Section 5.3.4, more specifically in Figure 5-13.

Sensitivity to Number of Users:

For the analysis of sensitivity to the number of users (sample size), the following simulation options were grouped and compared:

- Sample sizes = 200 users
- Sample size = 1000 users

Throughout all simulations performed, the same trend of results has been observed for both sets of simulations, i.e., if we consider the ratio between the results obtained from simulation over the expected value from theory for Poisson arrivals, the ratios of long-term results of each simulation type are similar for both sample sizes evaluated, in each simulated scenario.

This observation is better illustrated in Figure 5-14, where we compare the ratio simulation result to the expected Poisson results in terms of average value, standard deviation and 98% CDF point for the number of simultaneously active users.

Sensitivity to Number of Users:

For the analysis of sensitivity to the characterization period, we compared results obtained for different ranges. In the scenarios performed, characterization periods (simulation times) ranged from 1hr (3,600 s) to 10,000 hrs (36 million seconds), with each period being 10 times longer than the previous one.

The main observation in this topic is the fact that the more predictable a model is, such as a for Poisson traffic model, the smaller the characterization period needed for results to converge. Conversely, the more self-similar the model is, i.e., the smaller the α value, the longer it takes for simulations to converge.

This fact was observed in all sets of simulations performed and may be verified by looking at Figure 5-7, Figure 5-8, Figure 5-10 and Figure 5-11.

5.3.5.2. Reusability and Parameterization

Since multiplexing simulations are location independent, the topic of reusability of those simulations naturally comes to mind. It would save the designer time if in many situations the phase of multiplexing simulation could be bypassed through the use of an “estimated distribution” of active users, based on the lookup of tabulated results from pre-processed simulations.

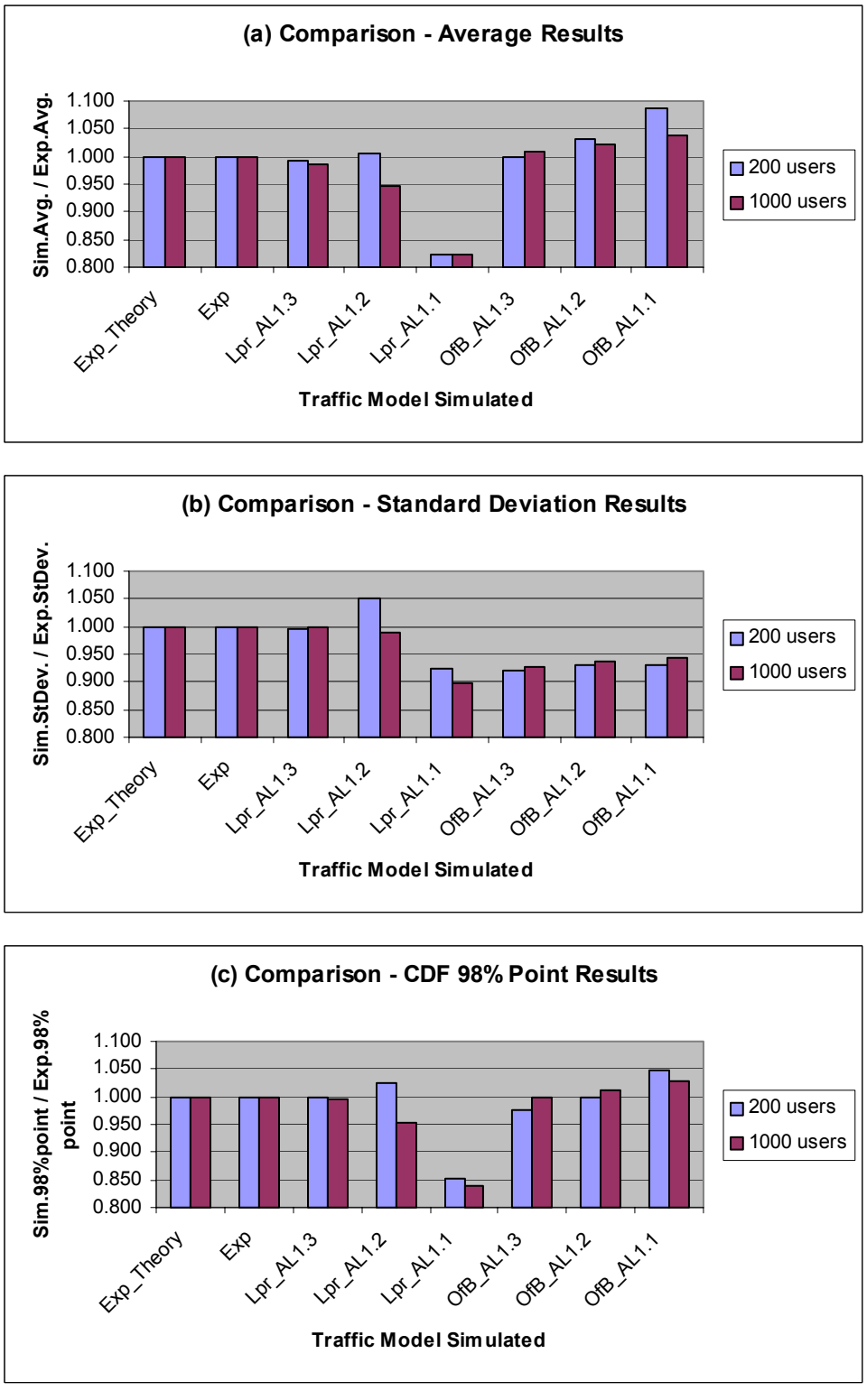


Figure 5-14. Sensitivity to sample size. Comparison between long-term simulation results (simulation result/expected from theory for Poisson model). (a) Average number of simultaneous users. (b) Standard deviation. (c) 98% CDF point.

For this purpose, we look back at the results presented in this section, with a focus on the trends observed in each case.

It has been observed that a good starting point to parameterize results is to calculate the expected distribution of and equivalent Poisson process to use as a reference, based on the equations provided in Section 5.3.2. The "equivalency" mentioned here refers to the amount of traffic to be simulated, which is given by:

$$A = \frac{N\lambda}{\mu} = N\frac{E[L]}{E[I_a]}$$ **Equation 5-4**

In the equation above, N is the total number of users (sample size), λ is the mean inter-arrival rate (i.e., the inverse of the expected interarrival time *E[Ia]*) and μ is the inverse of the mean service time *E[L]*. Notice that for Poisson processes, A is also the expected value of active users in the system.

One possibility of parameterization, for instance, could be given for the ON-OFF scenarios, where, as demonstrated in Figure 5-14, it makes sense to estimate the distribution of active users in the system by comparison with the Poisson distribution.

Next we present a very simple example using linear regression only of how this could be done for the ON/OFF scenarios. The graphics shown in Figure 5-15 provide the following curves for the distribution of X, the number of active users in the system.

$$E_{ON/OFF}[X] = (-0.2951*\alpha + 1.3853)*E_{Poisson}[X]$$ **Equation 5-5**

$$STD_{ON/OFF}[X] = (-0.0726*\alpha + 1.0185)*STD_{Poisson}[X]$$ **Equation 5-6**

$$CDF^{98\%}_{ON/OFF}[X] = (-0.2552*\alpha + 1.317)*CFD^{98\%}_{Poisson}[X]$$ **Equation 5-7**

The estimations above would be valid only for values of α within the range of 1.1 to 1.3.

Please notice that the example presented here intends only to illustrate the possibility of parameterization of results.

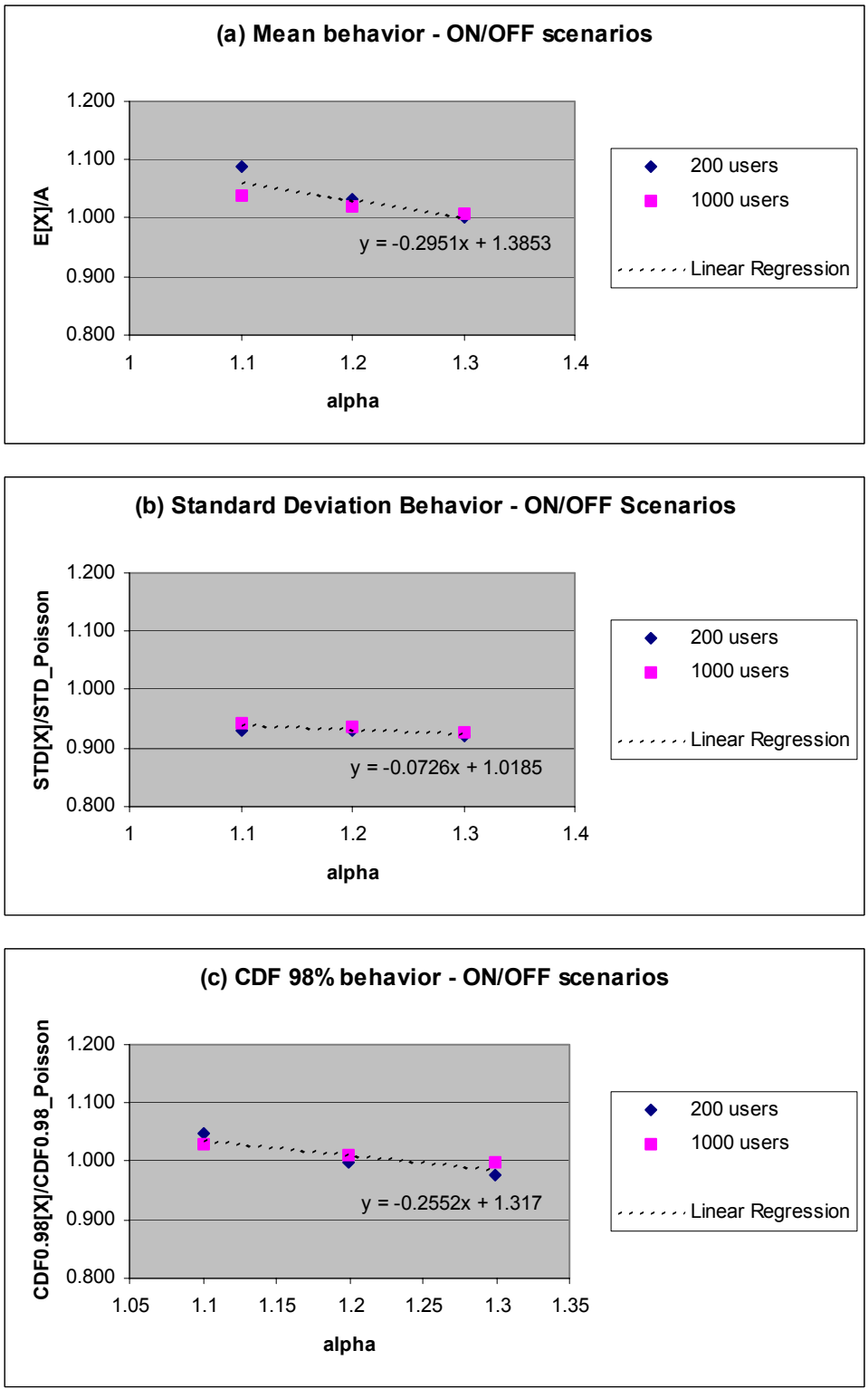


Figure 5-15. Parameterizing distribution of active users for ON/OFF model based on reference values for Poisson distribution. (a) Estimating the mean value. (b) Estimating standard deviation. (c) Estimating 98% C.D.F. point.

5.4. WCDMA Simulation

The procedure described in the previous section constitutes the snapshot generation block, which deals with both temporal and spatial user randomness. The simulation of the 3G systems will use as input the geographical distributions of active users (snapshot output) for all blocks simultaneously. Additional information needed to simulate how the system responds to the offered traffic include the user profile, which has already been used in generating the snapshot, and the radio network configuration, site locations, sector antennas, etc. Those additional blocks are described next.

5.4.1. User Profile Configuration

As already explained in Chapter 3, each traffic layer is associated with the description of a user profile. In this section, we formalize the concept of user Pprofile as follows. A user profile is defined by its set of service configuration, user terminal and environment parameters. These elements are shown in Figure 5-16.

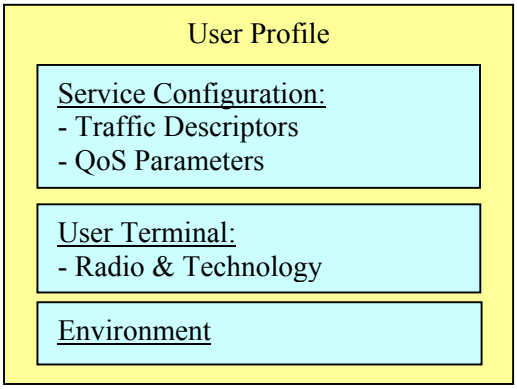


Figure 5-16. User profile configuration elements.

Typical parameters in each group are illustrated in Figure 5-17.

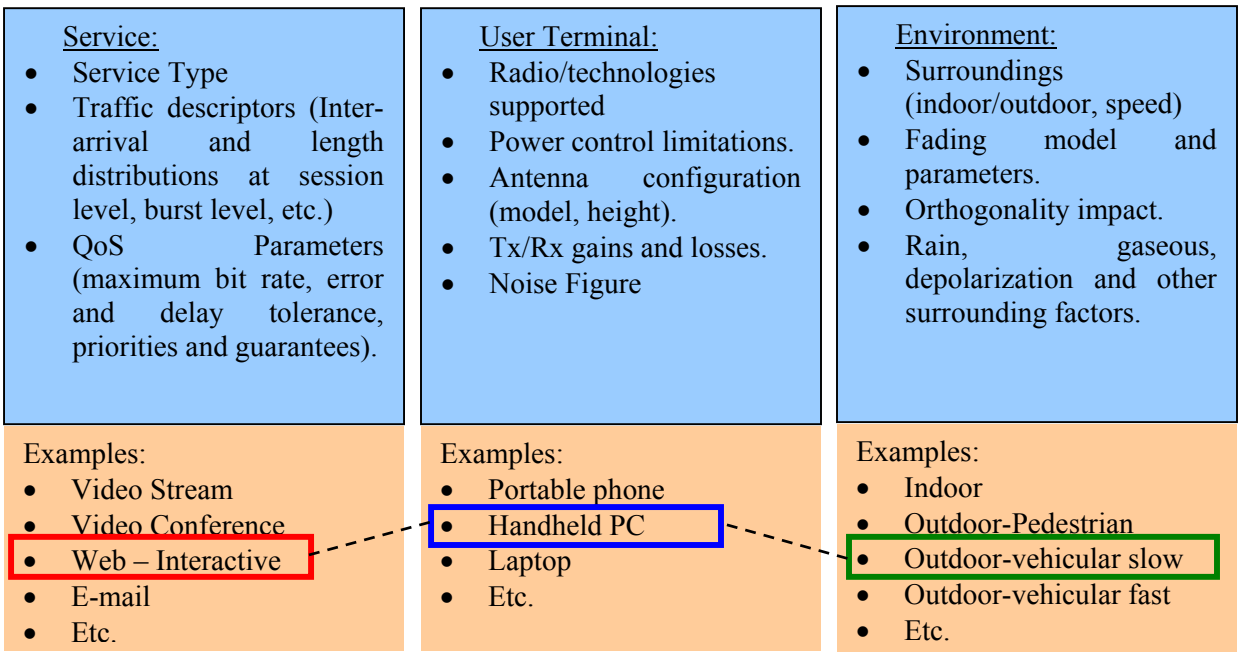


Figure 5-17. Concept of user profile as a combination of service, user terminal and environment configurations. User profile shown in example (connected boxes) is "Web on handheld PC at slow vehicle".

5.4.2. Radio Network Configuration

The Radio Network Configuration will describe the Radio Access Network and, more specifically, the base stations (Node Bs). For the simulation to correctly capture the way the network will respond to the traffic offered, it has to take into account the sector locations (latitude and longitude), transmission parameters such as power configuration for the control channels, power limitations for traffic channels that are subject to power control, link budget parameters such as gains and losses at transmission and reception, noise figure, diversity gains; antenna configuration such as height above ground, antenna model, radiation patterns, tilt and azimuth, and the technology specific parameters such as handoff thresholds, installed carriers, etc. This is illustrated in Figure 5-18.

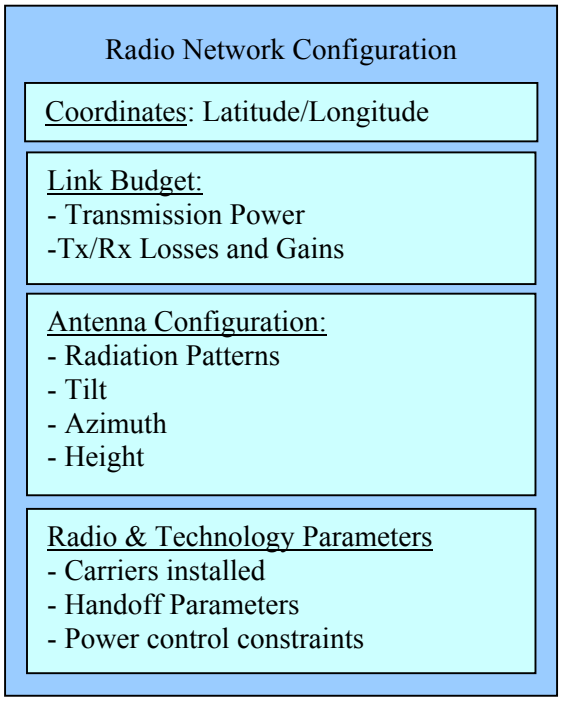


Figure 5-18. Radio network configuration elements.

5.4.3. Terrain and Propagation Data

Once sector locations are defined through the sector database and user locations are defined through the snapshot generation, the propagation analysis has to take place to evaluate path loss between transmission and reception between those elements. The path loss analysis allows appropriate calculation of transmission powers needed on power controlled channels, as well as calculation of noise rise at each point. A block diagram of the terrain and propagation data components is illustrated in Figure 5-19.

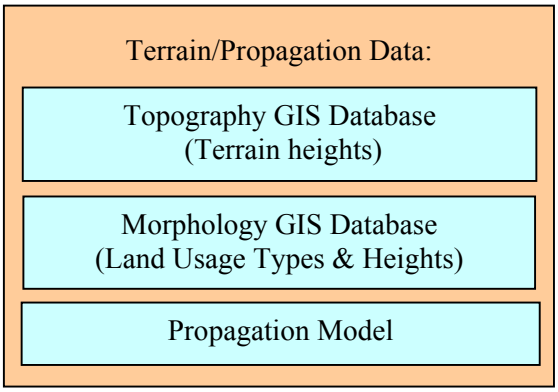


Figure 5-19. Terrain and propagation data.

5.4.4. Simulating the Air Interface

Once all blocks have been defined, the system simulation will be performed. Using the framework described in the previous sections, it is possible to verify for each active user, what would be the system resources (physical channels, power, etc) to allocate each service. For that purpose, the simulator will use all blocks together and perform iterative calculations for system load and power control algorithm convergence. A block diagram summarizing the information flow for the system simulation is shown in Figure 5-20.

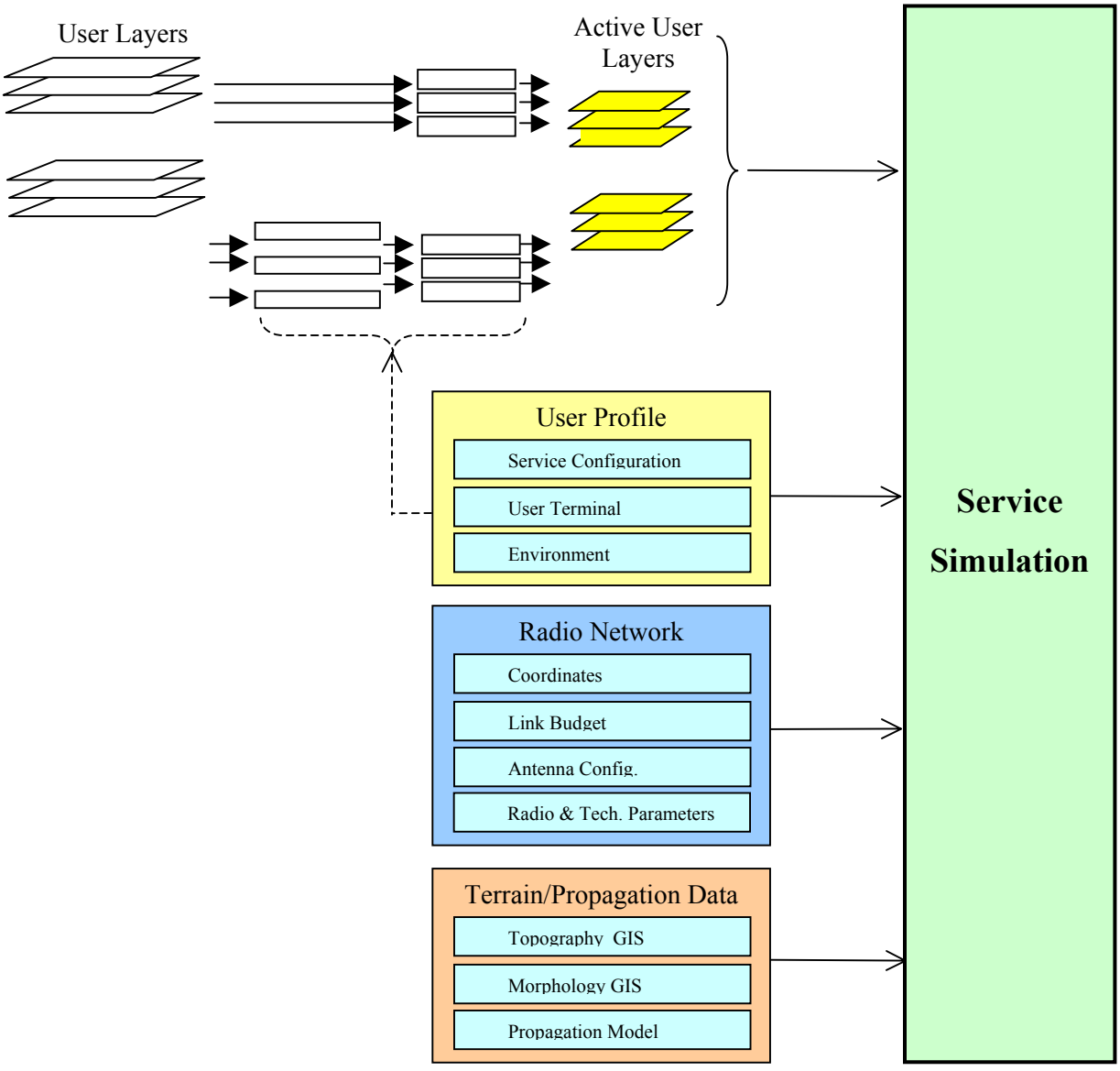


Figure 5-20. Information flow for 3G system simulation.

As a consequence of allocating the appropriate resources for each active user (within the limits of availability of the Radio Access Network), it is possible to then calculate the overall interference that is being generated by the sectors at each user location (downlink noise rise) and the overall interference generated by the users at each sector location (uplink noise rise). The updated noise rise calculation creates new conditions for the power control algorithms, which will create new figures of noise rise. This iterative process goes on alternately on downlink and uplink until convergence is reached in both directions.

In addition to convergence for the offered load, some other Radio Resource Management functions performed by the simulation algorithm are described below.

- *Load control and admission control:* Sectors will monitor load and refuse service in case a specified load limit has been achieved. Service may also be refused due to limits on number of codes, total throughput, etc., depending on the criteria defined.
- *Handover:* based on pilot *Ec/Io* levels and user class capabilities to support soft handoff, the same service bearer will be served by more than one sector. Handover gains are calculated based on the difference of signal levels available at each pixel.

Figure 5-21 illustrates the main functions performed by the simulation module, as well as the main iterative functions performed in that algorithm.

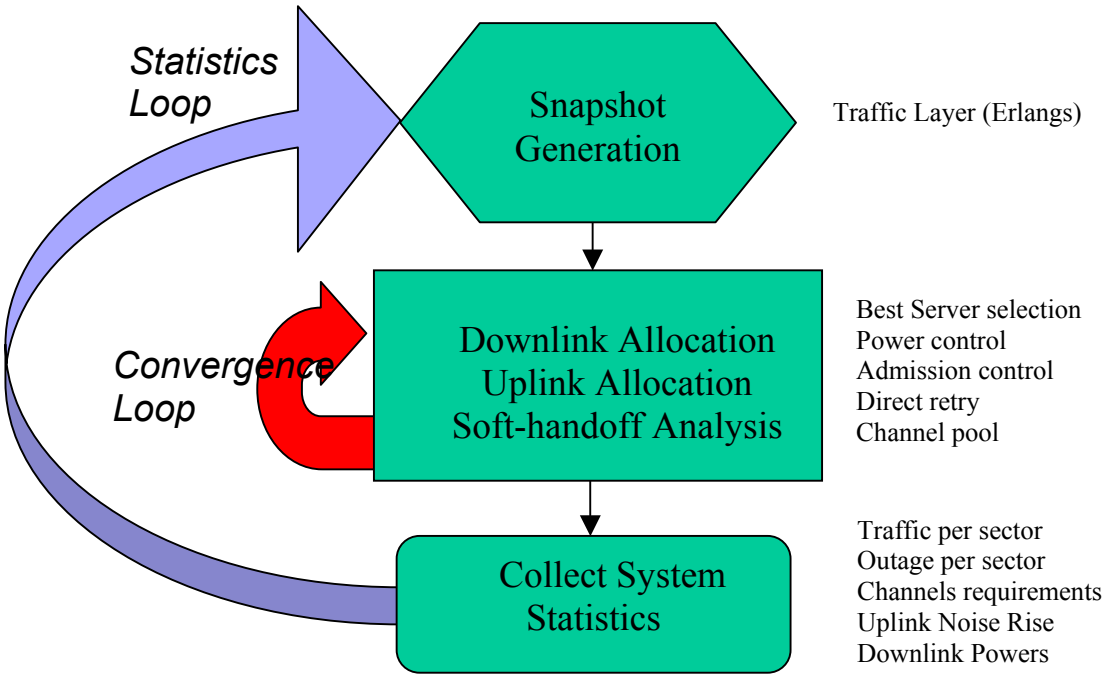


Figure 5-21. Iterative algorithm in 3G system simulation

The main outputs of the simulation process are the sector parameters that depend on the traffic load, namely the noise rise per sector and the power transmitted by the traffic channels. Besides, statistics are collected with regards to many variables including throughput, number of channels/codes used, main/handoff traffic allocation, outage, etc.

5.4.5. Collecting Statistics from the System Simulation

The traffic and QoS reports are generated by collecting statistics for the performance variables for each user profile during the simulation snapshots. At the end of the process a report containing all the statistical variables is presented allowing the user to generate any type of reports form these statistical variables. The following statistical variables are captured during the traffic simulation.

- Average and standard deviation of number of users supported by each sector as the main server, separated by user class and totalized for all user classes.
- Average and standard deviation of number of users supported by each sector as a soft-handoff server, separated by user class and totalized for all user classes
- Combined average and standard deviation of number of users.
- Average and standard deviation of the number of connections that could not be supported due to lack of downlink and/or uplink traffic channel coverage (when pilot coverage is available).
- Average and standard deviation for each sector's traffic channels power levels, separated by user class and totalized for all user classes.
- Average and standard deviation for noise rise level perceived by each sector.

5.4.6. The Dimensioning Loop

The appropriate simulation of the system for a given demand characterization will allow appropriate performance analysis of the network, i.e., will allow correct identification of the percentage of outage expected at the system due to the Radio Network limitations.

As soon as these limitations are identified on a per sector basis, the next step in the design is to increase resources in the bottleneck sectors, based on the feedback provided by the simulation algorithm. After that, new simulations are required to make sure that the new performance figures are within acceptable levels. This procedure of course is also an iterative process, where the feedback

from the simulation algorithm, namely achieved QoS results, serve as input for modifications in the Radio Network. This process is illustrated in Figure 5-22.

In dimensioning traffic for wireless systems, the main resource for allocation is the spectrum, i.e., the operator needs to estimate and maximize the capacity that can be handled per installed carrier for a given network design. Assuming a system layout with given site locations and sector configurations and using the methodology previously described to estimate the demand on that network, the most important dimensioning problem is to determine the amount of resources that need to be allocated at each sector to satisfy the demand while providing the required quality of service. In the event that the number of resources needed exceeds the spectrum availability (dictated by license agreements), a re-design of the network with the addition of new cells and cell splitting is considered if demand and QoS are to be satisfied. If the system design is for a new network layout, this process allows more flexibility and many iterations may be executed to optimize site location and sector configuration.

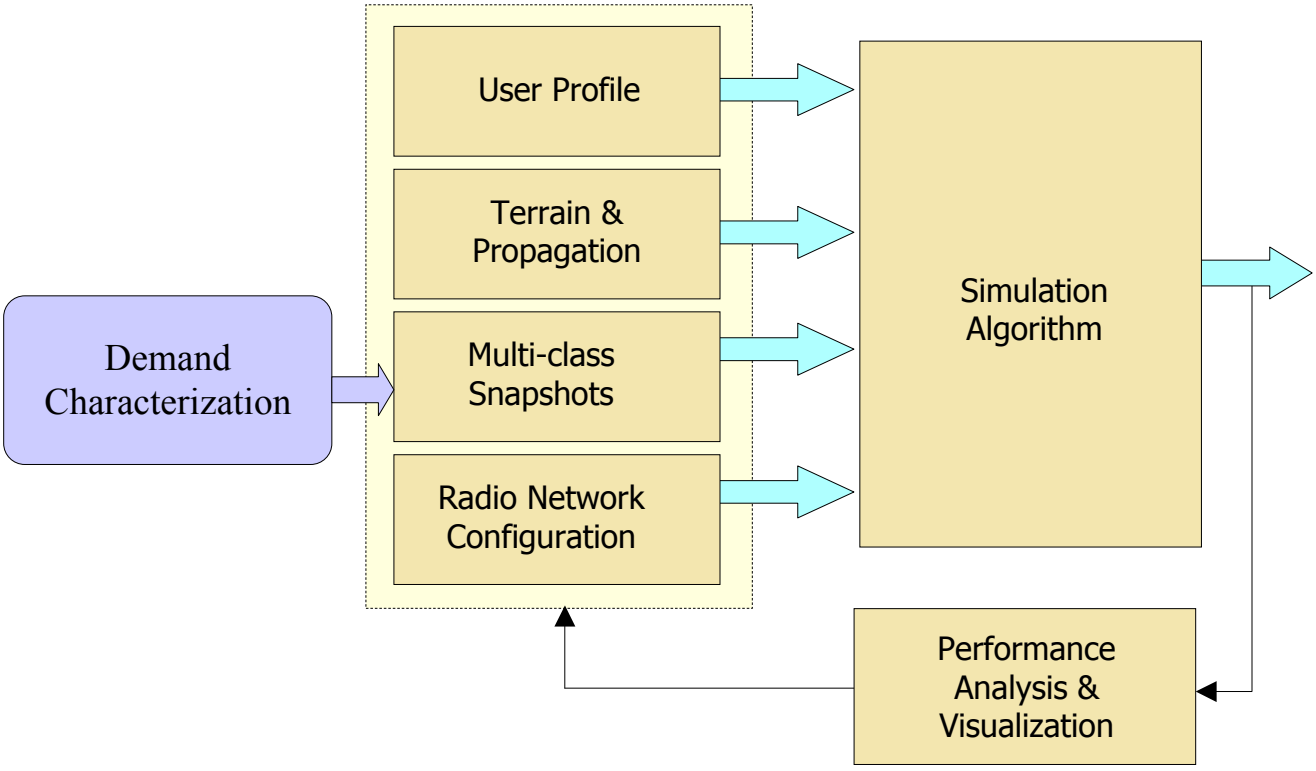


Figure 5-22. Overall block diagram of dimensioning process proposed in this framework.

Notice that although QoS must be achieved from end-to-end in the 3G networks, we are particularly concerned in this section with the bottleneck of the air interface, which is strongly limited by spectrum availability and hardware implementation, among other things. Once the Radio Network is appropriately dimensioned, e.g., the appropriate number of required carriers and sector locations

have been obtained, the final simulation of the system will provide the relevant figures of throughput per sector and its statistical distribution, and those figures will be used for the core network dimensioning, which is not in the scope of this discussion.

5.5. Summary

We started this chapter by discussing system simulation techniques that could be used to dimension multimedia systems, and selecting static Monte-Carlo simulations from among the multiple options presented. Next, we proposed the concept of "multiplexing simulation" as a solution to model traffic as statistically significant inputs for generation of snapshots to be used in system-level simulations. Results of some multiplexing simulations were presented and discussed and sensitivity and reusability issues are addressed. Finally, we presented a generic view of the whole system simulation process for multimedia wireless systems.

In the next chapter, we perform UMTS case study simulations using the proposed framework of demand characterization and system-level simulation. Multiple simulation scenarios are performed and results are used to discuss sensitivity and validation of the proposed methodology.

Chapter 6. Case Studies: UMTS Simulations

6.1. Objectives

As the next natural step in this dissertation, in this chapter we perform UMTS simulations using the proposed framework. For this analysis, we selected the city of Shreveport, LA, for which the author has previously performed 2G system dimensioning for PCS networks and for which important GIS input information is available, including demographics, terrain elevation and morphology data, as well as field-tuned propagation models.

Once simulation is performed to reflect the most relevant aspects of the system, different user mixes may be studied, yielding a better understanding of expected system performance.

The main objectives we accomplish in this phase of the research are as follows.

- Select a suitable WCDMA simulator and, for validation purposes, compare it to at least one other simulator available in the public domain.
- Analyze sensitivity of dimensioning results to main assumptions in traffic modeling and demand characterization.
- Validate the proposed methodology against baseline results of simplified models or, when applicable, existing data.
- Identify the main metrics of system capacity such as total throughput, throughput per sector, power transmitted per sector, and uplink noise rise per sector.
- Verify the impact of different demand scenarios on those capacity metrics.

To achieve the objectives listed above, multiple simulation scenarios are conducted in this phase, as described in the following sections.

6.2. Selecting a Simulation Tool

Since the main objective here is not to design and implement an optimal simulator, but to evaluate the proposed system dimensioning methodology, we have decided to use for this phase an "off-the-shelf" UMTS simulation tool. Options considered included both commercial tools as well as tools available in the public domain for that purpose.

6.2.1. Main Tool – CelPlanner Suite

In selecting a simulation tool, the author has considered many factors, including the following:

- Accuracy and reliability of the simulation tool.
- Cost and availability of the tool to the author.
- Author's knowledge and understanding of the tool's capabilities and limitations.
- Learning curve that is required to start obtaining productive results using the tool, which is directly related to the author's familiarity with it.
- Author's influence in having new features/modifications (mainly those related with additional reports and output variables) implemented in the tool when deemed necessary.

Based on the criteria above, the author's choice was to use the UMTS module of the CelPlanner Suite software package from CelPlan Technologies, Inc., in whose specification the author has had direct participation, relying strongly on knowledge acquired during this research. That software package has been implemented at this time and its commercially available version has been released.

Other options that were considered in choosing the simulation tool included tools available in the public domain for research purposes, such as WCDMAsim^{TM}, developed by the Mobile and Portable Radio Research Group (MPRG) at Virginia Tech. Our limitation to use that tool is due to the fact that it is a link-level simulator, which does not directly support network simulation features. Another option of a tool available in the public domain for research purposes is the NPSW tool, developed by Nokia, which runs over a MATLAB® platform by The MathWorks, Inc.

Another alternative commercially available tool considered was the OPNET Modeler® UMTS module, which supports detailed dynamic simulations over multiple layers. OPNET has kindly provided a temporary research license for this purpose. However, we decided to keep this option as a backup, due to the effort it would require to learn how to use the tool well enough to develop the complex model required for the author to be able to accomplish satisfactory and reliable results.

6.2.2. Validation Tool – NPSW

For validation analysis, we have decided to select a second simulation tool, to be used in a comparative study with equivalent inputs to those used in a given scenario. That procedure allows us

to validate the proposed methodology against baseline results of simplified models. This selection was done following the same criteria used for the first tool selection. For its simplicity and due to the fact that it is an open implementation, the tool selected for that purpose is the NPSW tool, developed by Nokia, which runs over a MATLAB® platform (by The MathWorks, Inc.). The author had access to this tool as an annex to the book *Radio Network Planning and Optimization for UMTS,* edited by J. Laiho, A. Wacker and T. Novosad [23], all engineers from Nokia.

6.3. System Description

As mentioned before, the analysis area used for all simulation scenarios is the city of Shreveport, LA, for which an extensive amount of GIS and propagation data was available to the author from a previous 2G PCS networks design experience.

In the following section, we illustrate the main data elements available for this market, and make specific comments about some of the details related to each item.

6.3.1. GIS Database

6.3.1.1. Image

The image database consists of scanned/digitized raster images of maps of the area. Those images are all geo-reference, i.e. their locations are correlated to specific coordinates, in the format of a raster file or bitmap. The function of the image database is to work as a geo-reference for the market area, where main geographical references can be correctly mapped and identified, to the level of streets and other functional references typically found in geographical maps.

For the area of Shreveport, the following map scales were available as geo-referenced images:

- USGS maps 1:250k scale
- USGS maps 1:100k scale
- USGS maps 1:24k scale

Illustrations of those images are shown next in Figure 6-1 to Figure 6-3.

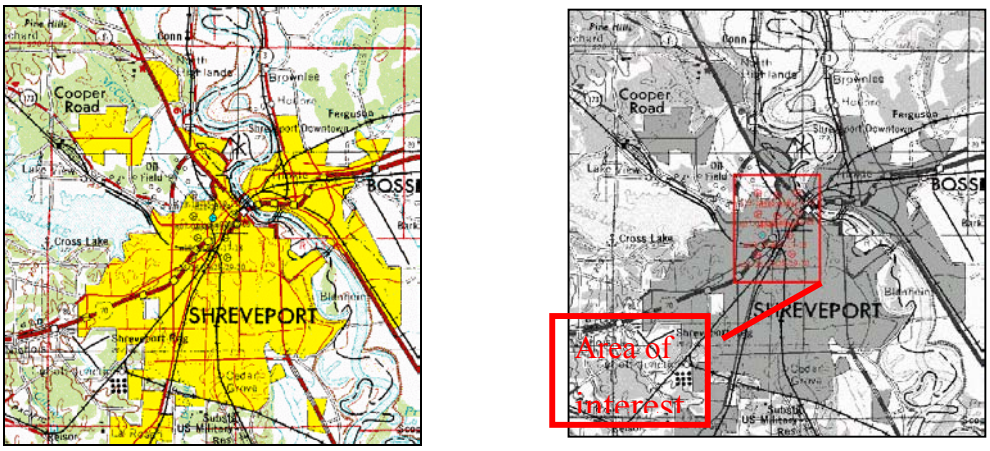


Figure 6-1. Raster image scanned from 1:250K maps.

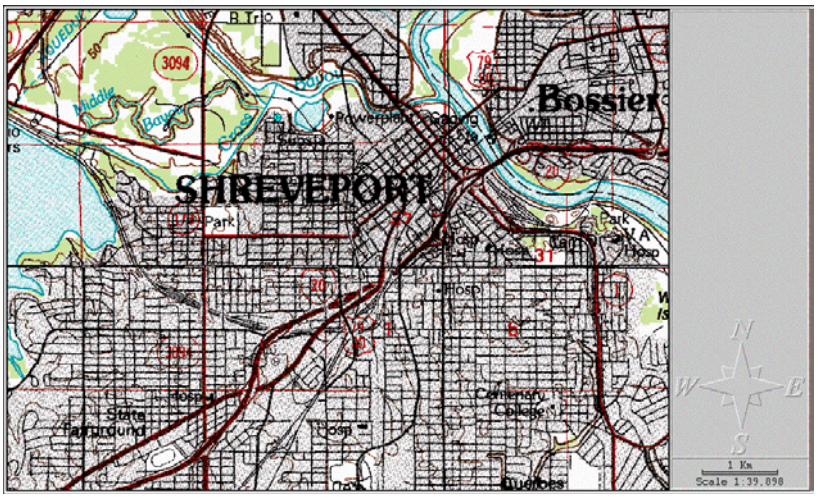


Figure 6-2. Raster image scanned from 1:100K maps.

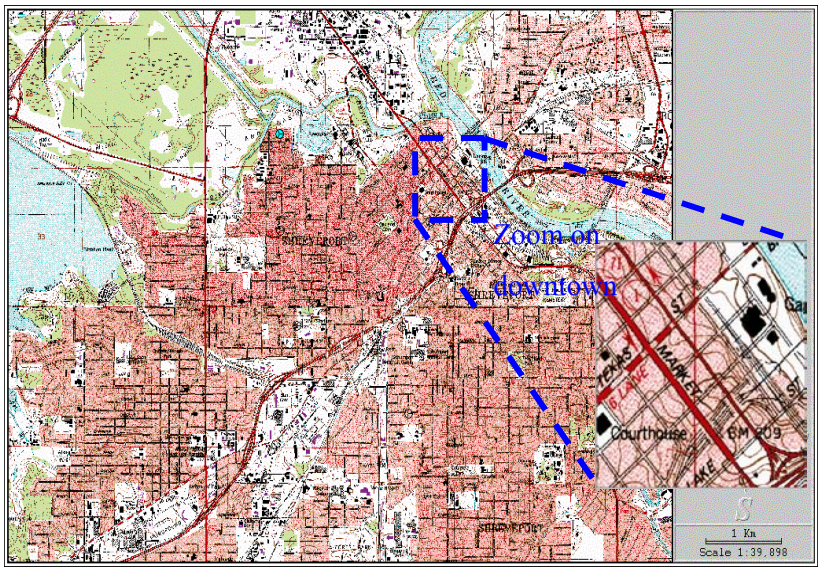


Figure 6-3. Raster image scanned from 1:25K maps.

6.3.1.2. Landmarks Database

The landmarks database consists of vector elements representing the main types of landmarks. The data available for this specific market was produced from USGS (year 2000) data, and provides vector information in multiple layers for landmarks such as water bodies, primary and secondary roads, streets, railroads, etc. Figure 6-4 graphically represents this information for a portion of the area of interest. This information is typically used as geographical reference during the process of tuning the propagation model.

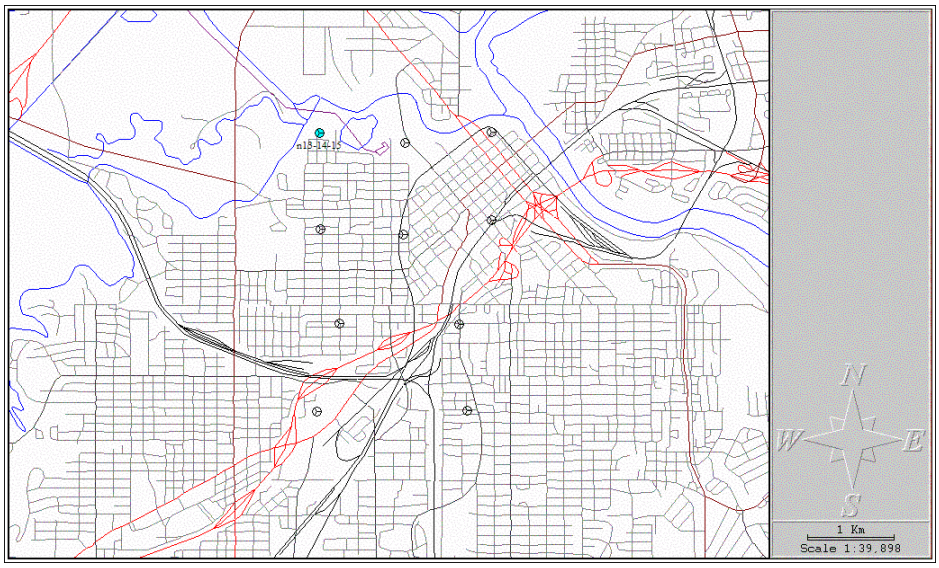


Figure 6-4. Landmarks (vector) database for Shreveport area

6.3.1.3. Topography Database

The topography database provides information regarding terrain height, typically represented as the height "above mean sea level" (AMSL). The data available for this market is in raster format (one height value per pixel) in the resolution of 1 arc sec (each pixel represents approximately a 30x30 meter area). This database has been converted from USGS DEM raster files. This information is illustrated in Figure 6-5.

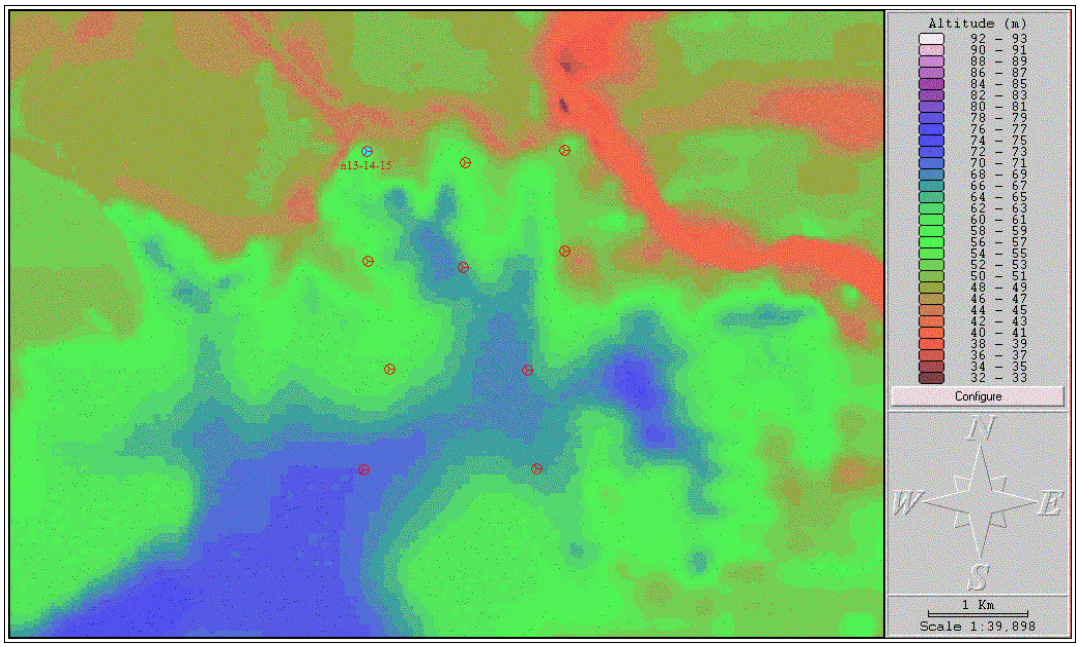


Figure 6-5. Topography database (raster) in 1arc sec resolution for Shreveport area.

6.3.1.4. Morphology Database

The morphology database used in this analysis consists of raster files in which each pixel has as attribute the type of usage or cover existing in that area, as well as mean clutter type for that cover. Examples of morphology types include: water, wetlands, shrubs, forest, residential areas, roads, commercial, etc. The morphology database available for this market is in 1 arc sec (30 meter) resolution and was derived from USGS LULC (Land Use and Land Cover) data with morphology heights added per category. The original 32 classes available from the USGS data were converted to thirteen types, to group types that presented similarities with respect to their propagation characteristics. Figure 6-6 illustrates the information contained in the morphology database.

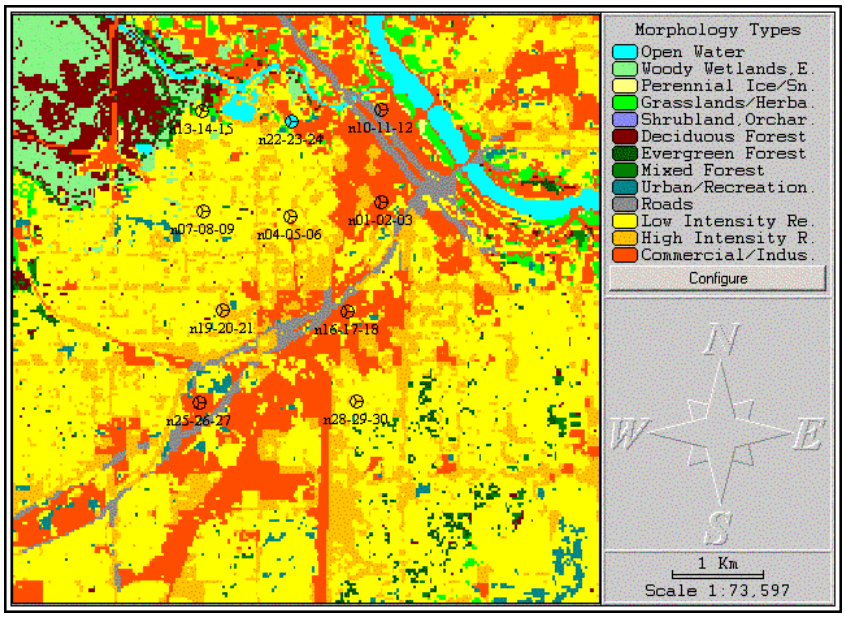


Figure 6-6. Morphology database (raster) in 1 arc sec resolution for Shreveport area.

Topography and morphology databases are considered together when analyzing link propagation. Therefore, both terrain obstacles (such as hills, valleys, etc) and clutter obstacles (such as building blocks, trees, etc.) are taken into account together when evaluating the propagation path and estimating path loss based on a given propagation model. Figure 6-7 illustrates this point by showing a terrain profile.

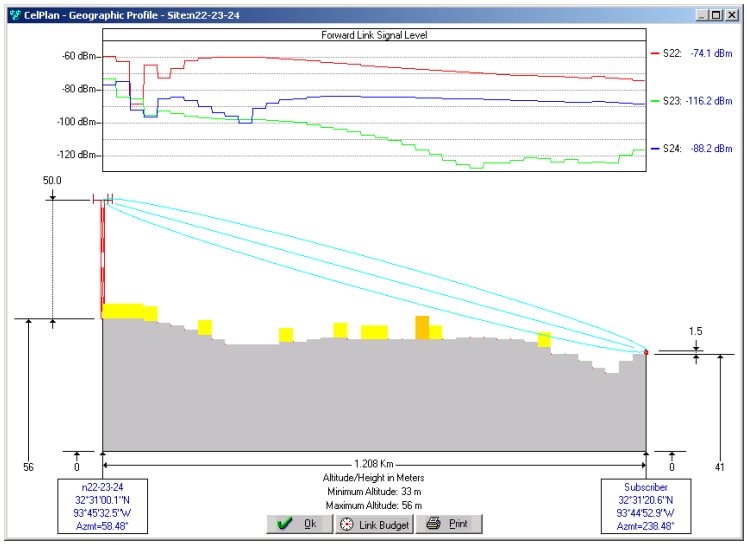


Figure 6-7. Example topography (gray) and morphology (colors) information along a propagation path.

6.4. Simulation Scenarios

We have chosen four scenarios for simulation that will be evaluated for multiple purposes. A brief description of the scenarios evaluated is presented in this section, as well as the main objectives of each one.

Scenario A: One class only, homogeneous traffic distribution

This scenario is meant to serve as a baseline, as it simulates a simple and reproducible situation that allows comparison of results with those expected from theory and from other simulation scenarios that include more sophisticated configurations.

Scenario B: One class only, heterogeneous traffic distribution

We use this scenario to compare results with previous simulations and discuss sensitivity of results to demand characterization with respect to geographical distribution. This scenario also serves as a baseline for comparison with the next scenario (Scenario C), in which the same type of configuration is evaluated using a different simulation tool.

Scenario C: One class, heterogeneous traffic, public domain simulator

The purpose of this scenario is to compare results with the previous simulation (Scenario B), for validation purposes. We focus on ascertaining whether results using our selected simulation tool (CelPlanner Suite) prove to be satisfactorily close to those achieved with NPSW.

Scenario D: Multiple classes, heterogeneous traffic distribution

Traffic layers are created using the same user distribution (snapshot), but now considering multiple user profiles, i.e., the user distribution will actually correspond to three separate snapshots, one for each user profile. The total throughput requested from the system here is much larger than in the previous scenarios, and we use the results to discuss sensitivity to demand characterization with respect to multiple classes. We also use this scenario for further validation against the public domain simulator, by executing this analysis in both simulation tools.

6.5. Methodology

For each scenario, the following information is presented.

- Description of the scenario's inputs and considerations.
- Summary of results such as overall achieved throughput, throughputs achieved per sector, transmitted and received powers per sector and other information that may be relevant to that particular scenario.
- Discussion of results addressing the main topics the scenario was intended to evaluate.

With regards to the methodology used in the simulations, much of the process is common to all scenarios. Some of those common tasks will be presented in this section.

6.5.1. Configuring Scenario Inputs

Next we present the main inputs and assumptions used in the configuration of all scenarios. Whenever there is additional information relevant only to a particular scenario, it will be described in the section corresponding to that specific scenario.

6.5.1.1. Users Geographical Distribution

For the evaluation of Scenario A, we have created a homogeneous traffic grid that has the same density of users in every pixel. A total of 1140 users have been distributed over the rectangular area indicated as the area of interest. The geographical area has approximately 3.6 km in the East-West direction by 4.6 km in the North-South direction (total area = 16.6 km^2), so the traffic density is about 71.5 users/km^2.

We follow the procedures already described in Chapter 3, creating one traffic layer for Scenarios B, C, and D that included the same number of users, but with a heterogeneous geographical distribution.

The parameters used in creating both geographical distributions (uniform and weighted by morphology) are shown in Table 6-1, while Figure 6-8 shows examples of the resulting grids, and Figure 6-9 shows a detailed view (zoom) of those grids, where the numbers in each square (bin) represent the mean number of active users in that bin, multiplied by 1000.

Considering that a bin (small square) in the grid has been chosen to have 30 by 30 meters, and considering the user density per square kilometer previously mentioned, each bin has a traffic intensity of about 0.062 users/bin in the homogeneous case, while that attribute varies per bin location in the heterogeneous traffic grid.

Table 6-1. Configuration Used when Creating Traffic Grids and Morphology Weights Used for Heterogeneous Distribution.

Traffic Grid Creation Parameters	
Resolution (bin size):	30 m
Active users:	1140
Samples per traffic bin:	16
Traffic Distribution Criterion:	by Morphology weight
Morphology weight factors	
Open Water	0
Woody Wetlands	1
Perenial Ice, Snow, Bare Rocks	1
Grassloands, Herbaceous, Pasture	10
Shrubland, Orchards, Vineyard	3
Deciduous Forest	5
Evergreen Forest	5
Mixed Forest	5
Urban/Recreational Grasses	40
Roads	80
Low Intensity Residential	110
High Intensity Residential	150
Commercial/Industrial/Transportation	200

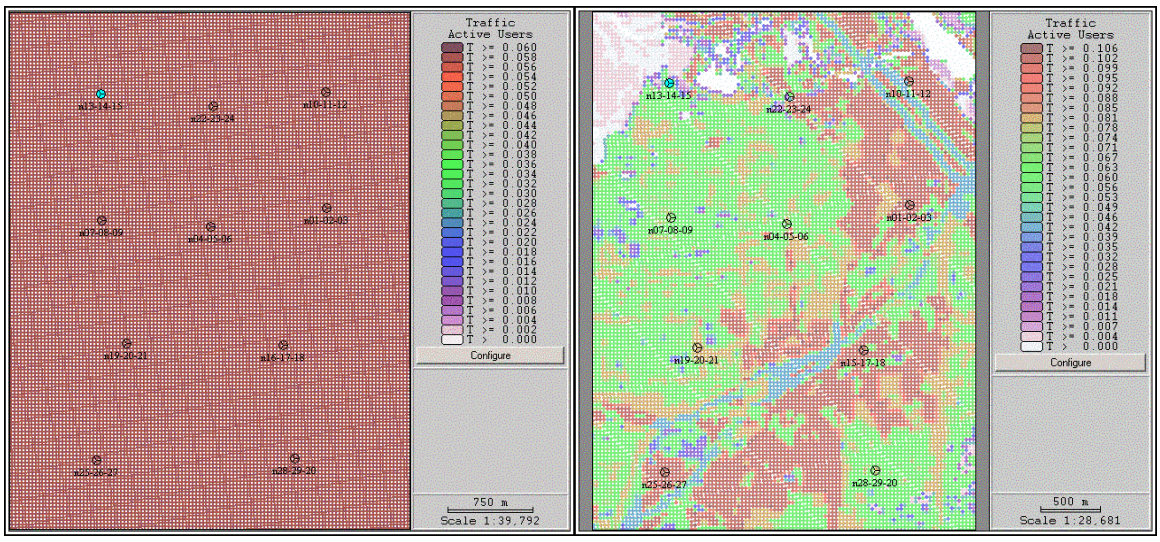


Figure 6-8. Resulting traffic grid. Left: homogeneous grid, right: Heterogeneous grid (created based on morphology weights).

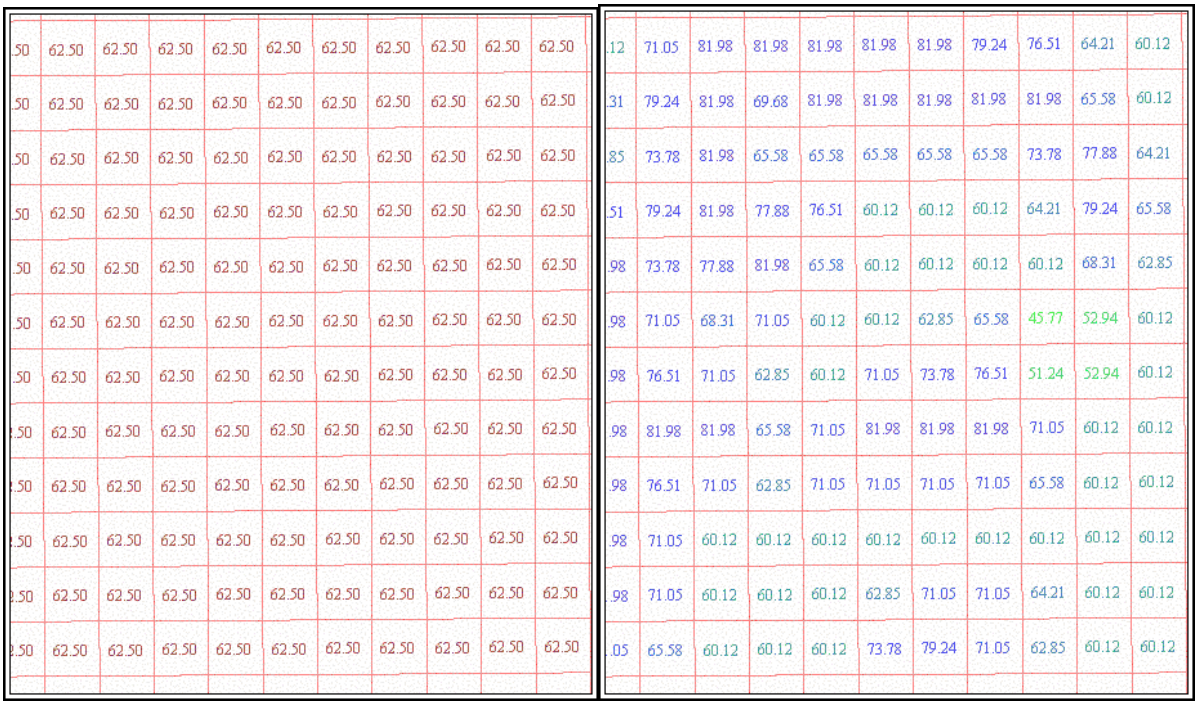
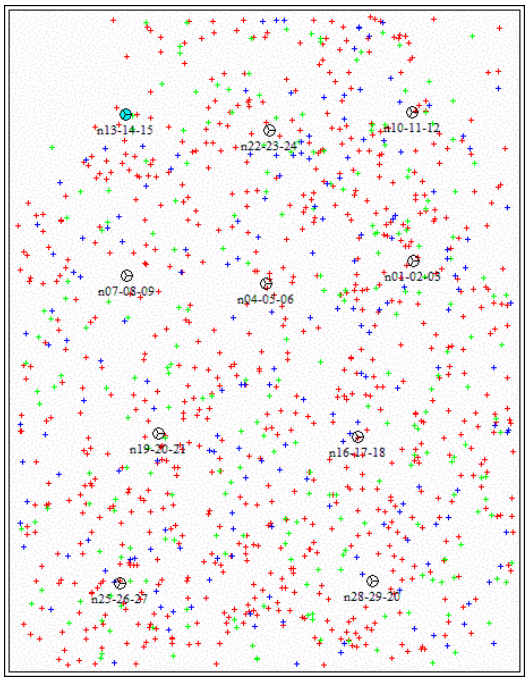


Figure 6-9. Resulting traffic grid. Left: homogeneous grid, right: heterogeneous grid (created based on morphology weights).

As described in Chapter 5, a statistically representative analysis will deal with the randomness of user locations, by performing multiple (hundreds or thousands) of snapshots where users are randomly distributed proportionally to the traffic intensity of a given traffic grid (Monte Carlo simulation). In this section, however, to allow the validation questions to be further evaluated, we concentrate on one snapshot at a time. This allows, for instance, a fair comparison between tools (e.g., comparing scenarios B and C), as for equivalency purposes exactly the same snapshot will be used for both scenarios. That snapshot is illustrated in Figure 6-10, where different colors indicate different user profiles (used in Scenario D). For the single-class scenarios (Scenarios A, B and C), the same snapshot was used, but with all users belonging to the same profile.



Example of one snapshot:

(Active users in each class)

Class 1 - Data 8kbps Portable: 720 users

Class 2 - Data 64 kbps Portable: 240 users

Class 3 - Data 144 kbps Portable: 180 users

Figure 6-10. Snapshot used as focus for comparison between scenarios B and C.

6.5.1.2. User Profiles

To distinguish the different user profiles to be simulated in the system, each profile (referred in the tool as user class) is described to the simulation tool with regards to its service characteristics, terminal configuration and environment conditions. Illustrations of such configurations are shown in Table 6-2 to Table 6-5.

For the scenarios simulated in this chapter, three user profiles have been considered. For scenarios A, B and C, only profile 1 is used, while scenario D simulates the three profiles simultaneously. When performing the multi-class scenario analysis, for simplicity, the three class distributions will be obtained from the same traffic grid, applying scaling factors to differentiate among them (Table 6-5).

Table 6-2. Service Configuration Parameters for the Three Service Classes Configured

Service Configuration				
Identification		Data 8 kbps	Data 64 kbps	Data 144 kbps
Service Type		Data	Data	Data
Service Data Rate	Forward	8 kbps	64 kbps	144 kbps
	Reverse	8 kbps	64 kbps	144 kbps
Activity Factor	Forward	1	1	1
	Reverse	1	1	1
Data Overhead Factor		0	0	0
Maximum Number of Simultaneous Servers (allowed in soft-handoff)		3	3	3
Required Traffic Eb/Io	Forward	9	7.1	6.1
	Reverse	6.5	4	3.5
Traffic Channel Maximum Power	Forward	0.2	0.84	1.5
Power Control Target Margin above Threshold (dB)	Forward	1	1	1
	Reverse	1	1	1

Table 6-3. Terminal Configurations (same configuration used for all classes)

User Terminal Configuration			
Identification	Portable Phone 1.5m	Palmtop 1.5m	Laptop 1.5 m
Maximum Output Power (W)	0.5	0.5	0.5
Receiver Noise Figure (dB)	8	8	8
Antenna Pattern	Omni-directional	Omni-directional	Omni-directional
Antenna Height	1.5 m	1.5 m	1.5 m
Antenna Nominal Gain (dBi)	-0.64 dBd (1.5 dBi)	-0.64 dBd (1.5 dBi)	-0.64 dBd (1.5 dBi)

Table 6-4. Environment Configuration (common for all classes)

Environment Configuration	
Identification	Outdoor
Human Body Attenuation (dB)	1.5
Penetration Attenuation (dB)	0

A common outdoor propagation environment model was used for all service classes, this accounts for 1.5 dB.

Table 6-5. User Profile Definitions (each user profile is a combination of service, terminal and environment)

User Profiles			
Identification	Data 8k Portable	Data 64k Palmtop	Data 144k Laptop
Service Class	Data 8 kbps	Data 64 kbps	Data 144 kbps
Terminal	Portable Phone 1.5m	Palmtop 1.5m	Laptop 1.5m
Environment	Outdoor	Outdoor	Outdoor
Allocation Priority	0	0	0
Traffic Factor	0.632	0.211	0.158
Color	red	green	blue

6.5.1.3. Base-Station Parameters

The description of the service classes refers mostly to the radio-access bearers (RAB) configurations, and specifies parameters that will be used when negotiating network access for each bearer establishment, such as allocated resources, QoS expectations, etc. Besides that, we need to configure the network itself and the resource availability for that network. In order to do so, we need to describe the Radio Access Network (RAN), and its base-station configurations.

The radio network chosen for this project consists of a total of 10 base stations, all of them with identical configurations, namely 3-sector 50m cells, with 7 degrees of tilt in all sectors, and azimuths 90, 210 and 330 degrees. The network does not follow a regular hexagonal grid, as it has been distributed with a slightly higher concentration (smaller radius) on the six top sites, and larger radiuses for the lower four base station. This topology is illustrated in Figure 6-11. For further reference in this chapter, each sector in the system has been given a unique identification number from 1 to 30, and each base-station (consisting of a 3-sector site) has been named based on the sector numbers that comprise it, in the sequence of ascending azimuth. Therefore, if for instance, we refer to sector 15 of site “n13-14-15,” we are uniquely identifying it as the 330-degree azimuth sector located at that site.

Figure 6-11. Network topology: sectors are numbered within a site based on ascending azimuth (90, 210 and 330 degrees).

In describing the base stations, transmission and reception parameters must be specified for each sector, such as physical installations (location, antenna heights, etc.), RF parameters (antenna patterns, tilt, transmitting powers, transmission and reception losses, etc.) and Radio Resource Management parameters that are specific to the sectors, such as thresholds for handoff algorithms, minimum and maximum power setups for traffic channels, common channels power allocations, etc.

The configurations of a base-station setup and link-budget parameters for a specific sector in the base-station are illustrated in Table 6-6 and Table 6-7, respectively.

Table 6-6. Base Station Configuration Parameters

Common Cell Parameter (valid for all cells)	
Number of sectors	3
Prediction resolution	1 arc.sec.
Prediction radius	9.5 km
Propagation Model	Single slope (40 dB/dec)
Sector Parameters	
Azimuth	90, 210 and 330 degrees
Tilt	7 degrees
Antenna Model	kt65deg15
Operation Frequency	2000
Pilot Minimum Ec/Io (Tdrop)	-18 dB
Pilot SHO Add threshold	-14 dB

Table 6-7. Base Station Link-Budget Parameters

Common Link Budget Parameters at Base Station	
Forward Link	
Pilot Channel Power (W)	1
Paging Channel Power (W)	0.75
Sync Channel Power (W)	0.25
Traffic Channels Total Power (W)	0.1
Traffic Channel Min. Power (W)	0.001
Traffic Channel Max. Power (W)	1.5
Traffic Power Standard Deviation	0
Number of Paging Channels	1
RF Power Scaling	0
Orthogonality Factor	0.5
Transmission Losses (dB)	14
Transmission Antenna Nominal Gain (dBd)	15
Reverse Link	
Reception Antenna Nominal Gain (dBd)	15
Diversity Gain (dB)	2
Reception Losses (dB)	14
Receiver Noise Figure (dB)	5

6.5.1.4. UMTS Configurable Parameters

In addition to the parameters that are configurable on the user profiles and on the base-stations, some parameters are globally defined for the whole system, and refer to the technology as well as some of the techniques being used for Radio Resource Management, such as load and admission control parameters. An illustration of such parameters is shown in Table 6-8.

Table 6-8. Global UMTS System Parameters

UMTS System Parameters	
Bandwidth per carrier (MHz)	3.84
Maximum Throughput per Carrier	no limit specified
Maximum Number of Vocoders per carrier	no limit specified
Forward Maximum Load Factor	1
Reverse Maximum Load Factor	1
Maximum Allocation Failures per Call	100
BS (sector) Minimum Output Power (W)	0.001
BS (sector) Maximum Output Power (W)	10

6.5.2. Running Simulations

After inputs have been defined for each scenario, simulations were executed and the results saved in different formats for post-processing.

For a simulation to be executed, one or more traffic grids are open, and a set of simulation parameters is configured, such as the total number of snapshots, convergence criteria, random seed, etc. Those parameters are illustrated in Figure 6-12.

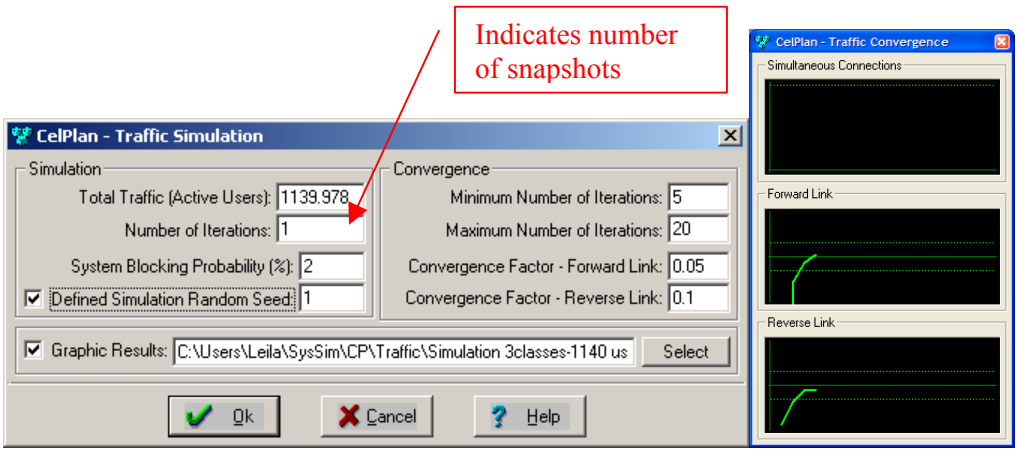


Figure 6-12. Simulation setup parameters and simulation status dialog while running.

As mentioned before, to perform the simulations of this section in a controlled manner that allows comparisons with results obtained from other simulation tools, we chose to run these simulations for one snapshot only, and save the set of generated active users in a file that can be later converted and imported into other tools.

During the simulation phase, as described in detail in Section 5.4.4, important functions of the operation of the network are reproduced, including the procedures for power control, load/admission control and handover management. In this process, overall interferences on both downlink and uplink for each connection vary depending on transmitted powers and accepted/rejected connections. For that reason, an iterative process goes on for both downlink and uplink, until the interference and, therefore, load converges in both directions. This whole convergence process is performed for each snapshot evaluated, and statistics are collected along the process.

6.5.3. Processing Results Statistics

Once a simulation is performed for a certain scenario, a large volume of information is available for post processing. That information may be presented in the format of a summary (Figure 6-13), or a full report that may be exported as a text file and opened as a spreadsheet for post-processing (Figure 6-14).

On those reports, important information is available that allows the assessment of the system capacity in response to the offered traffic, as well as system diagnosis that identifies the main reasons why traffic could not be handled in each scenario.

Out of the many output variables available in the report, we highlight here the ones that will be more relevant to this research. Unless otherwise indicated, the variables described below are presented on a per sector, per user profile basis, as well as overall results for the whole system. When computed over multiple snapshots, all variables are presented in terms of their mean and standard deviation. The following metrics are considered.

- Number of accepted connections (service bearers).
- Number of overhead connections (in soft handoff).
- Overall throughput.
- Total power transmitted per sector on the traffic channels (in addition to the common channels, which have fixed and pre-defined powers, that are input variables for the simulation).
- Uplink load factor and noise rise.
- Number of rejected connections, grouped by rejection reasons, which include the following:
    - Rejection due to lack of coverage. Since this means that no sector was able to provide coverage for this service request, this statistic is not grouped on a per-sector basis, but as a global result for the system.
    - Rejection due to limit on downlink power allowed per RAB.
    - Rejection due to limit on total downlink power available at sector.
    - Rejection due to limit on uplink power available at mobile station.
    - Rejection due to load control on downlink.
    - Rejection due to load control on uplink.
    - Rejection due to hard-blocking based on maximum channel count or data-rate allowed per sector.

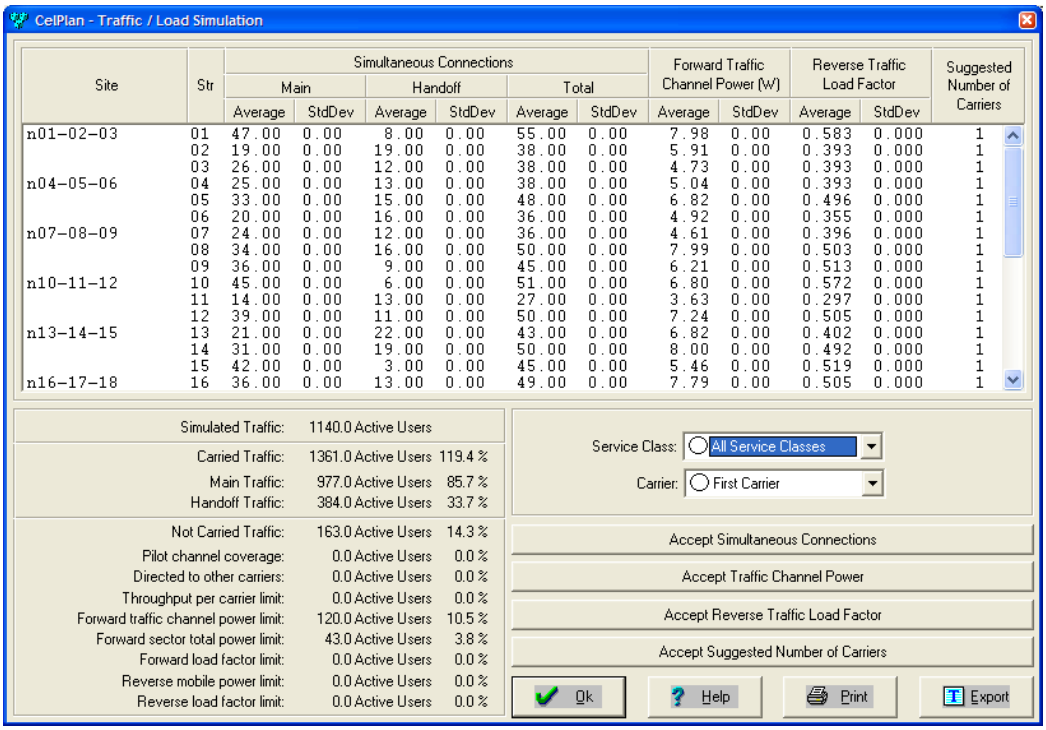


Figure 6-13. Example of simulation results summary.

Traffic Simulation Report

Simulated Traffic 1139.978 Active Users
Iterations 1
System Blocking 2 %

Site	Sector	First Carrier Carried Traffic															First Carrier Not Carried				
		FCSCMain (Conn)		FCSCHff (Conn)		SCTotal (Conn)		FWDRTot (Kbps)		RVDRTot (Kbps)		FWPower (W)		RVLoad (Factor)		NoiseRise (dB)	OtherCarriers (Conn)		DataRate (Conn)		FWPower (Conn)
		Avrg	StDv	Avrg	StDv	Avrg	StDv	Avrg	StDv	Avrg	StDv	Avrg	StDv	Avrg	StDv		Avrg	StDv	Avrg	StDv	Avrg
Class Data 8k Portable																					
n01-02-03	1	10	0	4	0	14	0	112	0	112	0	2.051	0				0	0	0	0	5
...																					
n28-29-20	30	17	0	4	0	21	0	168	0	168	0	3.351	0				0	0	0	0	12
Sum		469	0	87	0	556	0	4448	0	4448	0	83.52	0				0	0	0	0.00	201.00
Class Data 64k Palmtop																					
n01-02-03	1	3	0	0	0	3	0	192	0	192	0	2.186	0				0	0	0	0	4
...																					
n28-29-20	30	5	0	0	0	5	0	320	0	320	0	3.468	0				0	0	0	0	4
Sum		92	0	13	0	105	0	6720	0	6720	0	71.49	0				0	0	0	0.00	90.00
Class Data 144k Laptop																					
n01-02-03	1	3	0	0	0	3	0	432	0	432	0	3.763	0				0	0	0	0	3
...																					
n28-29-20	30	1	0	0	0	1	0	144	0	144	0	1.156	0				0	0	0	0	4
Sum		56	0	3	0	59	0	8496	0	8496	0	72.58	0				0	0	0	0.00	74.00
All Service Classes																					
n01-02-03	1	16	0	4	0	20	0	736	0	736	0	8	0	0.58	0	3.770995	0	0	0	0	12
...																					
n28-29-20	30	23	0	4	0	27	0	632	0	632	0	7.975	0	0.586	0	3.834159	0	0	0	0	20
Sum		617	0	103	0	720	0	19664	0	19664	0	227.6	0				0	0	0	0.00	365.00

First Carrier

Carried traffic - Main traffic	617 Active Users	54.1 %
Carried traffic - Handoff traffic	103 Active Users	9 %
Not carried traffic - Pilot channel coverage	0 Active Users	0 %
Not carried traffic - Directed to other carriers	0 Active Users	0 %
Not carried traffic - Throughput per carrier limit	0 Active Users	0 %
Not carried traffic - Forward traffic channel power li	365 Active Users	32 %
Not carried traffic - Forward sector total power limit	158 Active Users	13.9 %
Not carried traffic - Forward load factor limit	0 Active Users	0 %
Not carried traffic - Reverse mobile power limit	0 Active Users	0 %
Not carried traffic - Reverse load factor limit	0 Active Users	0 %

All Suggested Carriers

Carried traffic - Main traffic	1140 Active Users	100 %
Carried traffic - Handoff traffic	951 Active Users	83.4 %
Not carried traffic - Pilot channel coverage	0 Active Users	0 %

Figure 6-14. Example of simulation results report.

After simulation is executed for a given scenario, a large set of additional analyses may be performed based on the key results obtained from simulation. Those analyses include the geographical display (prediction) of performance parameters, such as the achievable signal-to-interference ratios, for the pilot and traffic channels, soft handoff overheads, etc. Those predictions have to be calculated after the traffic simulation is done, because they depend on results such as the total powers transmitted and received per sector, which vary depending on the traffic load. Examples of the most relevant predictions will be presented in the next sections of this chapter, for different scenarios, and therefore are not illustrated here.

In the next sections we will detail the inputs and outputs obtained for each scenario, as well as highlight any other important aspect relevant specifically to that scenario. We will also discuss the results obtained, focusing on the main objectives of each scenario.

6.6. Scenario A: One Class, Homogeneous Traffic Distribution

6.6.1. Inputs and Assumptions

A summary of the main input parameters used for Scenario A is shown in Table 6-9.

Table 6-9. Summary of Input Parameters for Scenario A

Input Parameters	Scenario A
Traffic Simulation Parameters	
Traffic Distribution (homogeneous/heterogeneous)	Homogeneous
Number of 8 kbps users	1140
Number of 64 kbps users	0
Number of 144 kbps users	0
Throughput Requested (kbps)	9120
Number of snapshots	1
Maximum number of iterations	20
Random seed for snapshot	1
System Parameters	
Maximum allowed rate per sector (for hard-blocking) [kbps]	2500
Maximum number of channel elements per sector	999 (no limit)
Maximum Load Factor on Downlink (load control)	1 (no limit)
Maximum Load Factor on Uplink (load control)	1 (no limit)
Maximum number of failures per connection	100
Maximum BS total power (W)	10
User Profile Parameters (8kbps users)	
Required Eb/No downlink	9 dB
Required Eb/No uplink	6.5 dB
Maximum power allowed per RAB (W) downlink	0.2
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0

Sector Parameters	
Pilot Power (W)	1
Other common channels power (W)	1
Orthogonality factor	0.5
Maximum physical power allowed per RAB (W) downlink	1.5
Net Tx Gains (Gains - Losses) (dB)	1
Net Rx Gains (Gains - Losses) (dB)	3
Receiver Noise Figure (dB)	5

6.6.2. Results

Once simulation was executed, the results were saved in the format of screen shots and reports. The most relevant results of this analysis are shown next.

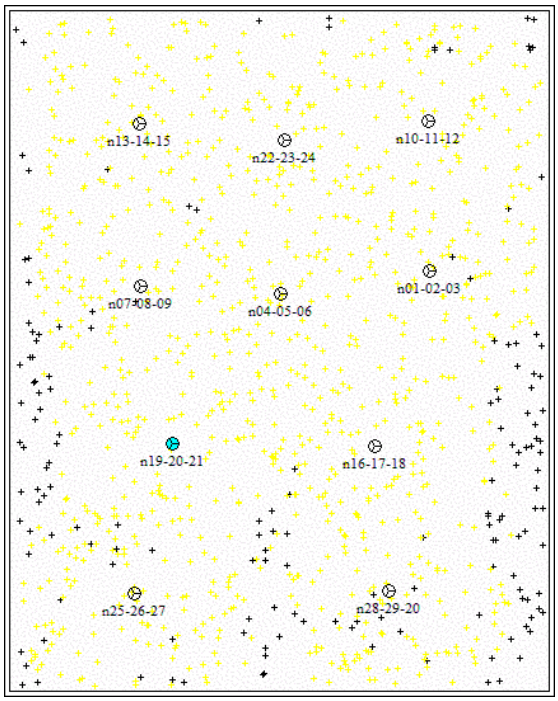


Figure 6-15. Graphical result of snapshot for Scenario A. Points in yellow show served connections. Points in black are unsuccessful connections.

From the simulated scenario, multiple statistics were collected for all sectors, in a format previously illustrated (Figure 6-14). A brief summary of the main statistics resulting from this scenario simulation is shown next.

Table 6-10. Summary of Results for Scenario A

	Avg. Sector	StDev	Sum
Requested Connections	38		1140
Served Connections	30.5	10.0	914
SHO Connections	9.3	4.0	278
Total Connections	39.7	8.2	1192
Throughput (kbps)	243.7		7312
Total Rate (kbps)	317.9		9536
SHO overhead	30.42%		30.42%
Not carried traffic - Pilot channel coverage	0		0
Not carried traffic - Directed to other carriers	0		0
Not carried traffic - Throughput per carrier limit	0		0
Not carried traffic - Forward traffic channel power limit	7.4		222
Not carried traffic - Forward sector total power limit	0.1		4
Not carried traffic - Forward load factor limit	0		0
Not carried traffic - Reverse mobile power limit	0		0
Not carried traffic - Reverse load factor limit	0		0
Total Dw TCHs Power (W)	5.6	1.4	168.85
Uplink Load Factor	0.42	0.09	
Uplink Noise Rise	2.4	0.7	

After saving the results into the project configurations, mainly those referring to total transmitted and received powers per sector, a full set of predictions was performed, and geographical illustrations of those results are shown next (Figure 6-16 to Figure 6-30).

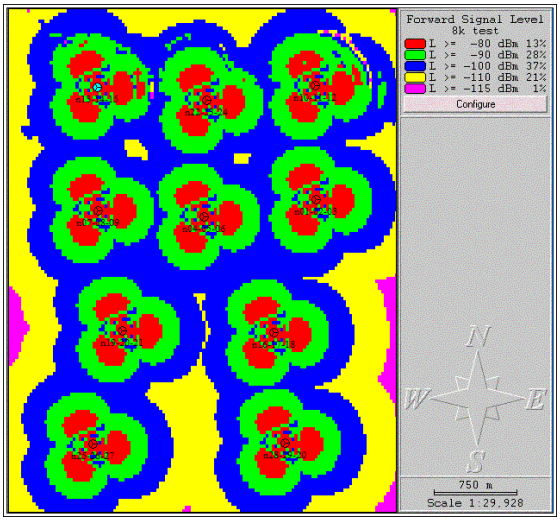


Figure 6-16. Pilot signal strength (dBm.)

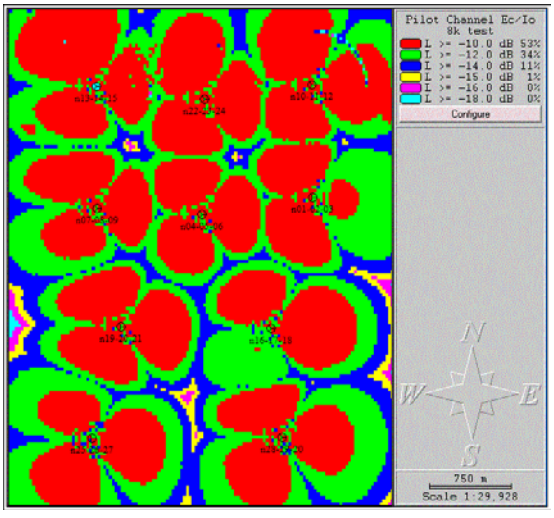


Figure 6-17. Pilot channel Ec/Io (dB).

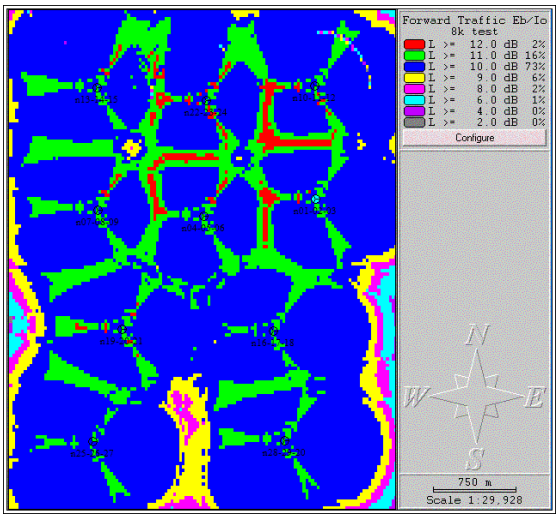


Figure 6-18. Downlink traffic channel Eb/Io (dB)

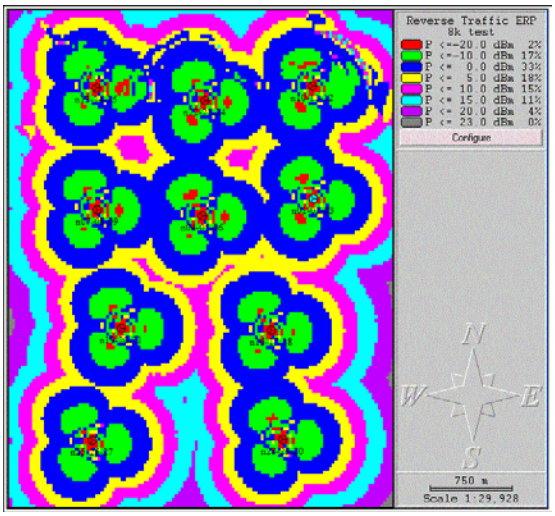


Figure 6-19. Uplink required MS power (dBm).

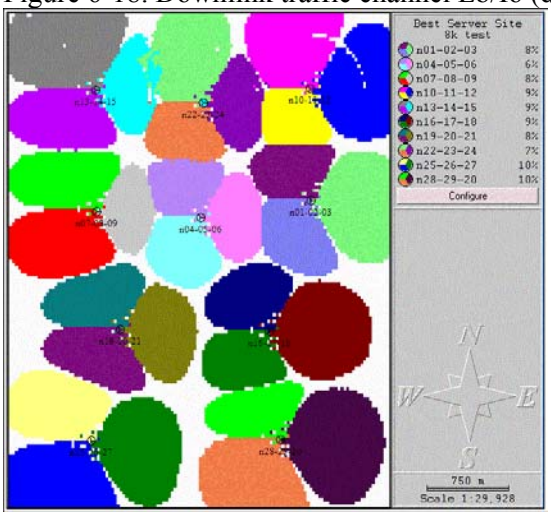


Figure 6-20. Pilot best server.

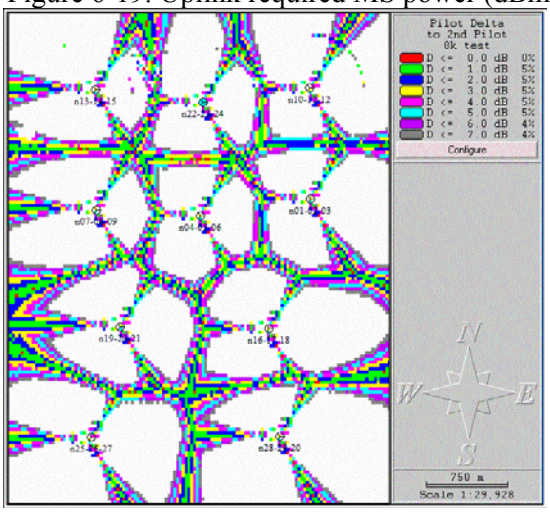


Figure 6-21. Pilot delta to 2nd pilot.

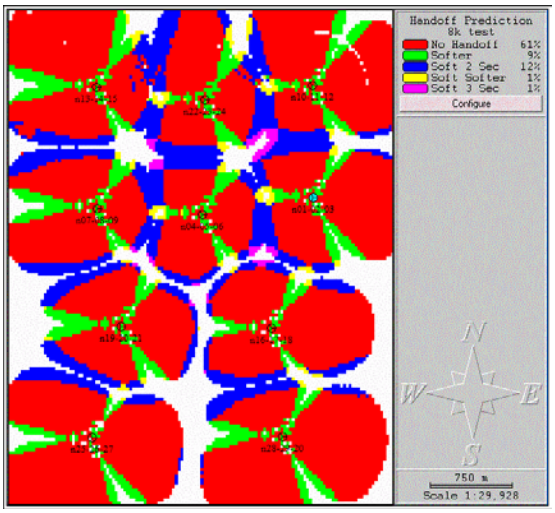


Figure 6-22. Handoff Areas.

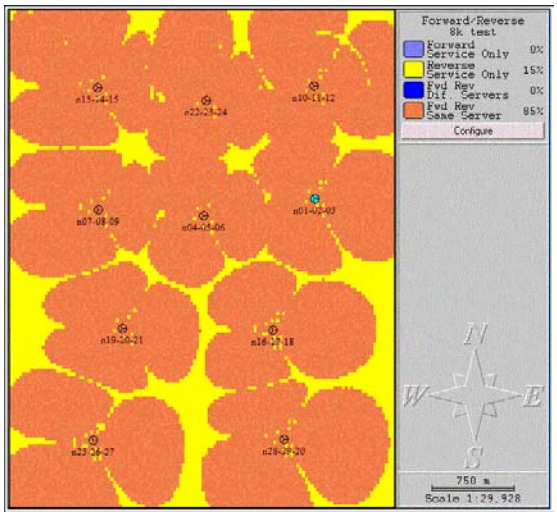


Figure 6-23. Down/uplink coverage comparison

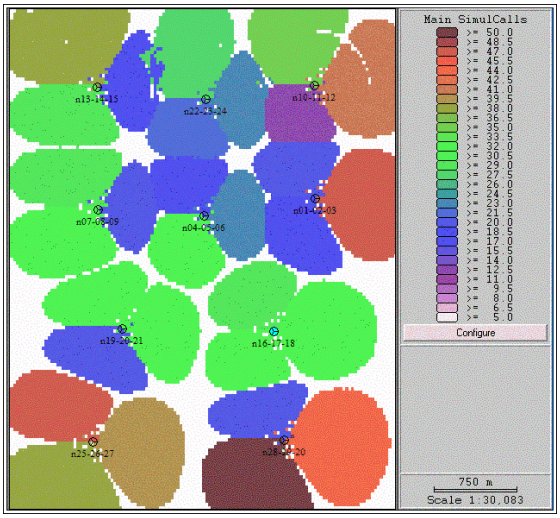


Figure 6-24. Number of users served.

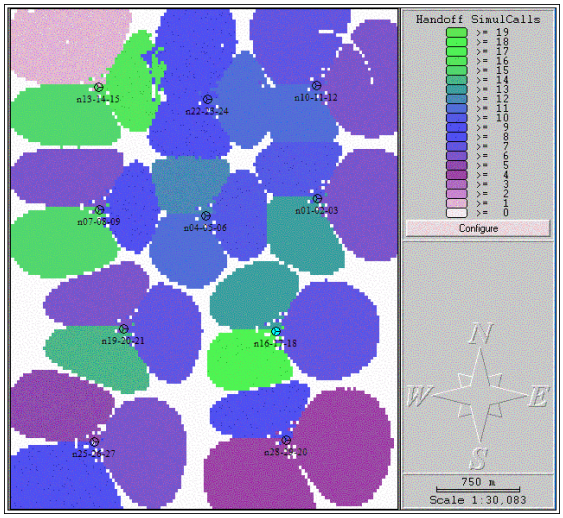


Figure 6-25. Number of SHO connections.

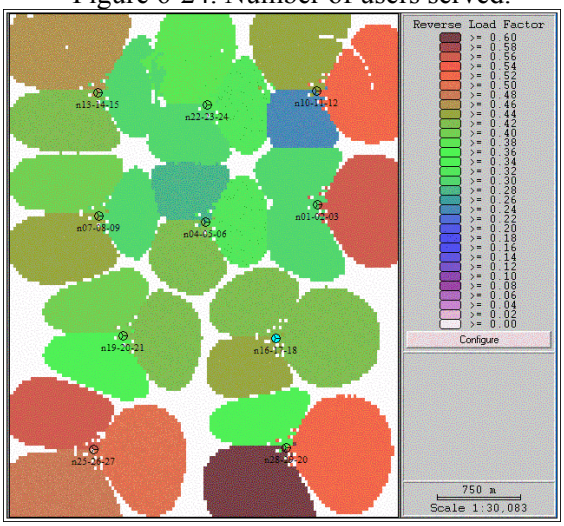


Figure 6-26. Uplink load factor.

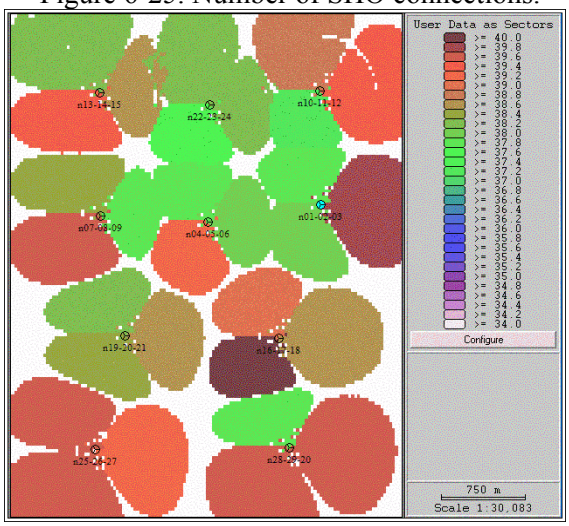


Figure 6-27. Total BS Tx power (dBm).

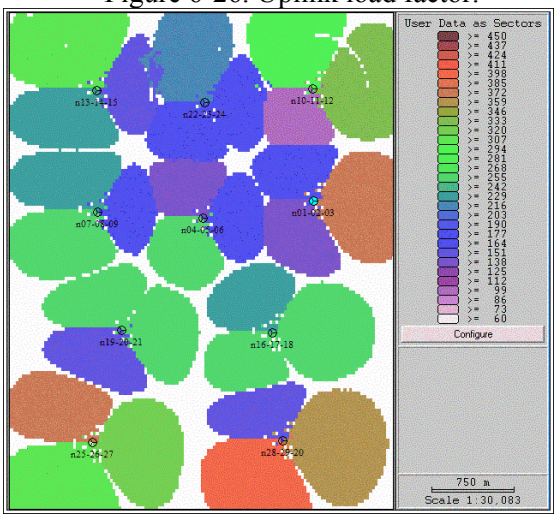


Figure 6-28. Sector throughput (kbps).

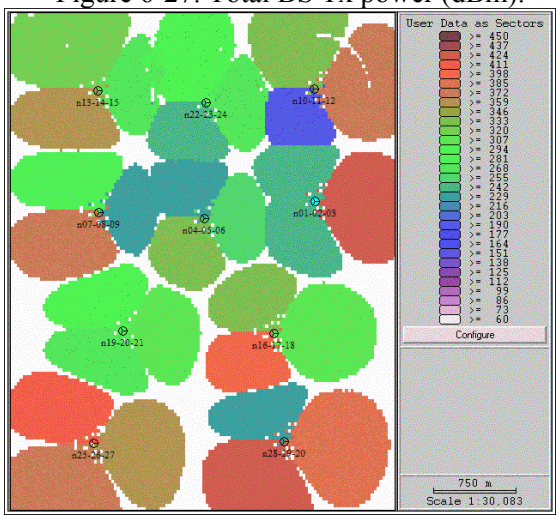


Figure 6-29. Sector total rate (kbps).

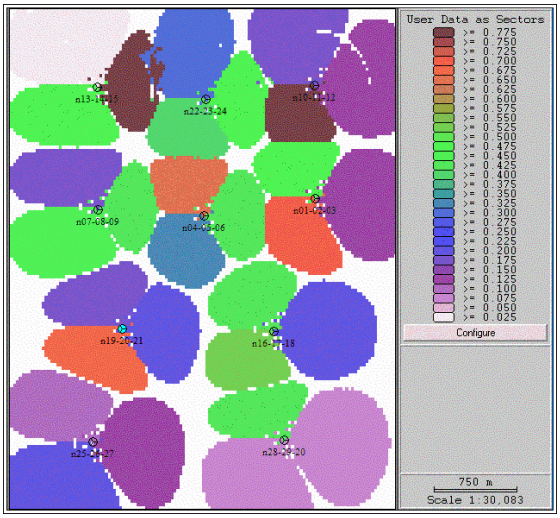


Figure 6-30. Sector SHO overhead.

6.6.3. Discussion

Although traffic was homogeneously distributed for this scenario, the network topology (i.e., the base station locations) was not regular in terms of following a homogeneous hexagonal grid. A total of ten sites have been distributed with a slightly higher concentration (smaller radius) on the six top sites, and larger radiuses for the lower four base station. For that reason, as expected the results in terms of throughput, load and coverage are not regular for the sectors in this system.

From the results presented in Figure 6-16 and Figure 6-17, it is clear that pilot coverage is not an issue in this particular scenario, as all pixels within the service area are above the required thresholds for pilot Ec/Io, set as –18 dB.

In Figure 6-18 we observe that the Eb/Io requirement (threshold of 9 dB) for the downlink traffic channel is also satisfied for most pixels on the top part of the system (six cells of smaller radius), while holes in the coverage are expected to appear at the edges of the lower part of the system (corresponding to the four sites of larger radius).

From Figure 6-19 we observe that uplink coverage is not the limiting factor in this scenario, as the required mobile transmission powers are well within the maximum allowed of 0.5 watts (27 dBm), for the entire service area.

The observations made above are confirmed by the results on Table 6-10, where we observe that all users not served in this system were rejected due to excess power requirements on the downlink, either on a per-channel basis (222 rejected users) or on the total downlink power allowed per sector (4 users). This observation reflects the fact that neither pilot channel nor uplink channels are limiting

service here, or, put in a different way, the system is not necessarily optimized in order to have its coverage in different layers equalized. This subject will be discussed in more detail in Chapter 7.

In Figure 6-20 we show the best-server area definition of each sector, using different colors to identify sectors. The gaps between sectors seen in this figure are due to the fact that in this prediction, the definition of a sector best-server area is given only to those points where coverage is available for the pilot, downlink and uplink traffic channels without the help of soft-handoff gain. Therefore, the gaps between sectors in that prediction do not mean that service is not available at those areas, but that it would not be available if soft-handoff (SHO) were not to be considered.

In Figure 6-21 we illustrate one useful prediction in the phase of system optimization, which is the estimation of potential pilot conflicts based on the difference in pilot strength between sectors. This type of prediction is used in order to allow appropriate pilot offset planning, which combined with search window size planning will prevent pilot offset conflicts in areas where handoff is desired.

In Figure 6-22 we identify the areas where soft-handoff is expected to happen in this system, based on pilot Ec/Io ratios available from the surrounding sectors. In Figure 6-23, we show an interesting comparison of coverage achieved for both uplink and downlink directions, when soft-handoff gains are not accounted for.

Figure 6-24 and Figure 6-25 show results that come directly from the simulation (no post-processing), and represent the total number of users served by each sector as main server or as secondary server (soft-handoff), respectively. It is interesting to observe here that the total number of users served is smaller in the areas of smaller cell radius such as sectors 04-05-06 and those surrounding them. On the other hand, sectors on the edges of the system, pointing outwards, are the ones that tend to carry the most traffic, since they suffer less interference from other cells, for the same signal strengths. For that reason, those are also the sectors that present largest coverage radiuses.

As a consequence of more traffic being carried by some sectors (mainly those at the system edges), the uplink load factor, illustrated in Figure 6-26, also follows the same trend described above, i.e., higher load factors are observed for those sectors at the edge of the system. It is interesting to remember here that the uplink load factor is a function of the total received power at a sector and, therefore, it is expected to grow as traffic grows within a sector.

In a system that is limited by the downlink, as is the case in this particular configuration, we expect to find all or most sectors transmitting close to its maximum allowed. That is the reason why in Figure 6-27 we observe that most sectors are transmitting very close to the maximum allowed power, which is 10 Watts (40 dB). We also observe that the sectors where the power is below maximum (not much below) are the ones expected to experiment the highest SHO areas.

Figure 6-28 and Figure 6-29 show the total throughput and the total rate carried by each sector respectively. Here, the throughput is defined as the data that is carried by a sector as the main server of its users, while the total rate includes also the SHO overhead from additional connections that the sector is supporting. As illustrated in Table 6-10, the sum of the throughput of all sectors in this system is 7.3 Mbps, which gives an average per sector of 243.7 kbps. However, due to additional soft-handoff overhead, each sector is carrying in average 317.9 kbps, which gives a mean SHO overhead of about 30% for this scenario. Figure 6-30 shows the SHO overhead observed in simulation for each particular sector, and demonstrates once again, that the internal sectors are those that carry the largest soft-handoff fraction.

6.6.3.1. Comparison with Expected Ranges from Theory

In this section we compare the results obtained from Scenario A with expected values from theory. This is possible because many simplifications have been applied to this scenario, such as the configuration of only one class, and homogenous spreading of traffic. Yet, as mentioned before, not everything is ideal in this scenario, since the base-stations locations do not follow a perfect hexagonal grid, therefore creating geographical relationships that differ from those typically expected in a perfectly ideal scenario. Nevertheless, we will use formulations and values that are typical for similar scenarios and evaluate here whether the results found in this simulation agree with the range of expected values from theory.

In order to do so, we will use some of the formulation presented in section 4.3.2, where we discuss load factors and noise rise estimations. Some common inputs to be used in that analysis, based on the inputs used to simulate Scenario A, are described in Table 6-11:

Table 6-11. Input Parameters Relevant to Theoretic Comparison

Input parameters (uplink)		
Variable	**Symbol**	**Value**
Bandwitdh (KHz)	W	3840
Eb/No required (dB)		6.5
Eb/No required (linear)	Eb/No	4.47
Activity Factor	ν	1
User Data Rate (kbps)	R	8
Number of sectors		30
Other-cell/same-cell ratio	*o/s*	Discussed next

We initially set the other-cell to same-cell ratio (o/s) based on literature. Different authors typically show different ranges for o/s, which vary with the number of sectors, antenna tilt and antenna configurations. In [7], theoretical estimations based on regular hexagonal grids show that for the downlink the o/s varies with the proximity of the user to the base-station, while for the uplink one o/s is considered for each sector. In both cases, the estimation of o/s depends on the path loss ratio (attenuation exponent over distance), γ, which typically varies between 2 and 5 (in our scenarios, γ=4). In [7], the o/s factor is shown to range from 0.49 to 0.72 on the uplink for omni-directional cells, with γ=4. Sectorized cells are expected to present lower values for that variable, since interference is more isolated. Typical suggested values in literature [42], [43], use 0.55 as a rule of thumb for non-tilted three-sector cells. In [23], simulation results are available for the specific case of 3-sector sites using 65 degrees antennas. In those simulations, o/s values are in the range of 0.38 (for 14 degrees tilt) to 0.88 (no tilt). However, even though results are tabulated according to the tilt value, it has been observed that the effect of tilt on the o/s value is directly related to the mean radius of a cell, and therefore those values cannot be directly applied. In the particular case of our simulated scenario, where the network topology does not follow a regular hexagonal grid, it is clear that any rule to estimate o/s beforehand for this system will be simply an approximation. Yet, based on all that has been discussed, we expect that this value falls somewhere between 0.38 and 0.55. Therefore, in order to evaluate the capacity formulas, we will initially simply use the median value in this range (0.465) as a reasonable o/s assumption. Later on in this section, we will estimate the actual ratio from simulation results.

Table 6-12. Initial Assumption for o/s Ratio

Variable assumption	Symbol	Value
Other-cell/same-cell ratio	*o/s*	0.465

Using the data given above, and based on the fact that all users have the same profile, i.e., request the same data rate *R*, have the same voice activity factor *v* and same *Eb/No* requirements, we can rewrite the equation given on Equation 4-9, to express it as a function of *N* (Equation 6-1), where *N* is the number of users served by a sector.

$$\ell_{up} = \left(1+\frac{o}{s}\right)\frac{N}{1+\frac{W}{\left(\frac{E_b}{N_0}\right)R\nu}} \qquad \textbf{Equation 6-1}$$

The equation above has two unknowns left, even after we assumed a "typical" value for o/s. These two unknowns provide a function between the capacity of a sector (N) and its load factor (ℓ). Therefore, by looking at one of the unknowns from the simulation results, it is possible to estimate the capacity expected for the system and verify consistency with results obtained for the other unknown. In order to do that, we may also reorder Equation 6-1 in order to estimate N as a function of the load factor, as in Equation 6-2. By multiplying N (the expected capacity of a sector) by the number of sectors, we can then estimate the expected capacity of the system.

$$N = \frac{\ell_{up}}{\left(1+\frac{o}{s}\right)}\left[1+\frac{W}{\left(\frac{E_b}{N_0}\right)R\nu}\right]$$

Equation 6-2

A summary of results obtained from simulation which will be used for comparison with the theory expectations is shown in Table 6-13.

Table 6-13. Summary of Results Relevant to Theory Comparison

Users served		914
Users connected on SHO		278
Users served total		1192
SHO Overhead		0.30
Average Users served per sector	N	30.47
Average Load Factor per sector	ℓ	0.415

Therefore, using N from the results table, we can estimate below the load factor expected from theory (0.412) using Equation 6-1, which is in good agreement with the results obtained from simulation (0.415), as illustrated in Table 6-14.

Table 6-14. Expected Load Factor per Sector

Average Users served per sector	N	30.47
Expected load factor per sector	ℓ	**0.412**

Conversely, using ℓ from the simulation results, we can estimate the expected system capacity from theory as 922, by using Equation 6-2, which again is in good agreement with the results obtained from simulation (914 users), as illustrated in Table 6-15.

Table 6-15. Expected Users per Sector and Total Users Served

Average Load Factor per sector	ℓ	0.415
Expected users served per sector	N	30.74
Expected users served total		**922**

Finally, instead of assuming a value for o/s we may calculate the observed value for this particular scenario, by rearranging Equation 6-1 as shown below.

$$\frac{o}{s} = \frac{\ell_{up}}{N}\left[1 + \frac{W}{\left(\frac{E_b}{N_0}\right)R\nu}\right] - 1$$

Equation 6-3

In this case, we use both N and ℓ from simulation results and calculate o/s based on Equation 6-3. That calculation shows that the achieve o/s ratio is 0.478, which is in good agreement with the value of 0.465 initially assumed for this scenario based on results presented in the literature.

Table 6-16. Estimated o/s Ratio for Scenario A

Average Users served per sector	N	30.47
Average Load Factor per sector	ℓ	0.415
Estimated o/s	o/s	**0.478**

6.7. Scenario B: One Class, Heterogeneous Traffic Distribution:

6.7.1. Inputs and Assumptions

A summary of the main input parameters used for Scenario A is shown in Table 6-17.

Table 6-17. Summary of Input Parameters for Scenario B

Input Parameters	Scenario B
Traffic Simulation Parameters	
Traffic Distribution (homogeneous/heterogeneous)	Heterogenous
Number of 8 kbps users	1140
Number of 64 kbps users	0
Number of 144 kbps users	0
Throughput Requested (kbps)	9120
Number of snapshots	1
Maximum number of iterations	20
Random seed for snapshot	1
System Parameters	
Maximum allowed rate per sector (for hard-blocking) [kbps]	2500
Maximum number of channel elements per sector	999 (no limit)
Maximum Load Factor on Downlink (load control)	1 (no limit)
Maximum Load Factor on Uplink (load control)	1 (no limit)
Maximum number of failures per connection	100
Maximum BS total power (W)	10
User Profile Parameters (8kbps users)	
Required Eb/No downlink	10 dB
Required Eb/No uplink	6.5 dB
Maximum power allowed per RAB (W) downlink	0.2
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0
Sector Parameters	
Pilot Power (W)	1
Other common channels power (W)	1
Orthogonality factor	0.5
Maximum physical power allowed per RAB (W) downlink	1.5
Net Tx Gains (Gains - Losses) (dB)	1
Net Rx Gains (Gains - Losses) (dB)	3
Receiver Noise Figure (dB)	5

6.7.2. Results

Once the simulation executed, the results were saved in the format of screen shots and reports. The most relevant results of this analysis are shown next.

In Figure 6-31, we show the graphical results of the simulation for this scenario.

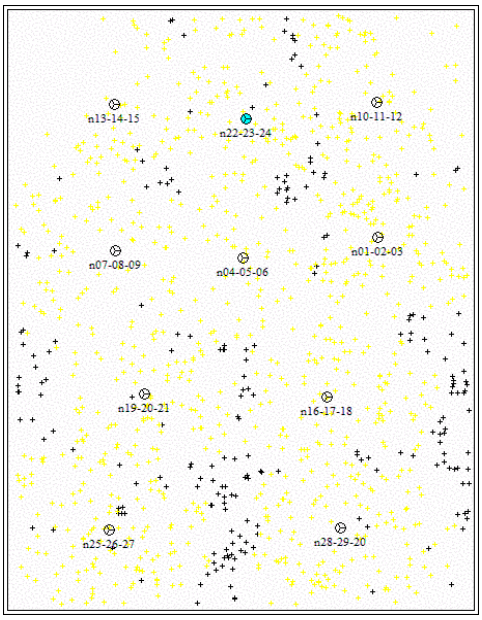


Figure 6-31. Graphic result of Scenario B snapshot. Points in yellow show served connections. Points in black are unsuccessful connections.

From the simulated scenario, multiple statistics were collected for all sectors, in a format previously illustrated (Figure 6-14). A brief summary of the main statistics resulting from this scenario simulation is shown in Table 6-18.

Table 6-18. Summary of Results for Scenario B

	Avg. Sector	StDev	Sum
Requested Connections	38		1139.978
Served Connections	30.8	9.8	925
SHO Connections	9.3	3.3	280
Total Connections	40.2	9.2	1205
Throughput (kbps)	246.7		7400
Total Rate (kbps)	321.3		9640
SHO overhead	30.27%		30.27%
Not carried traffic - Pilot channel coverage	0		0
Not carried traffic - Directed to other carriers	0		0
Not carried traffic - Throughput per carrier limit	0		0
Not carried traffic - Forward traffic channel power limit	7.0		211
Not carried traffic - Forward sector total power limit	0.1		4
Not carried traffic - Forward load factor limit	0		0
Not carried traffic - Reverse mobile power limit	0		0
Not carried traffic - Reverse load factor limit	0		0
Total Dw TCHs Power (W)	5.7	1.5	169.84
Uplink Load Factor	0.419	0.11	
Uplink Noise Rise	2.4	0.8	

After saving the results into the project configurations, mainly those referring to total transmitted and received powers per sector, a full set of predictions was performed, and geographical illustrations of those results are shown next (Figure 6-32 to Figure 6-46).

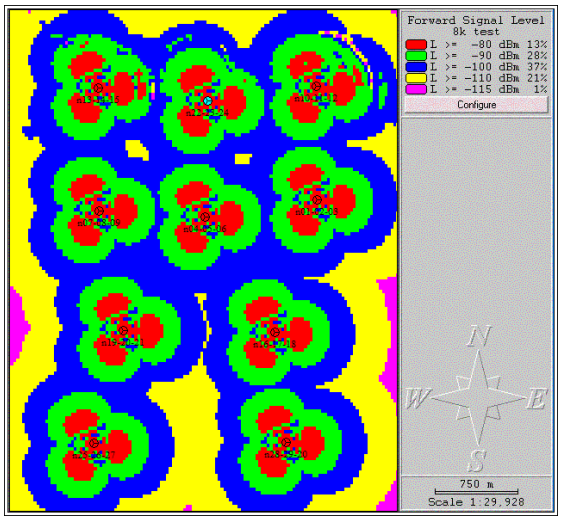


Figure 6-32. Pilot signal strength (dBm).

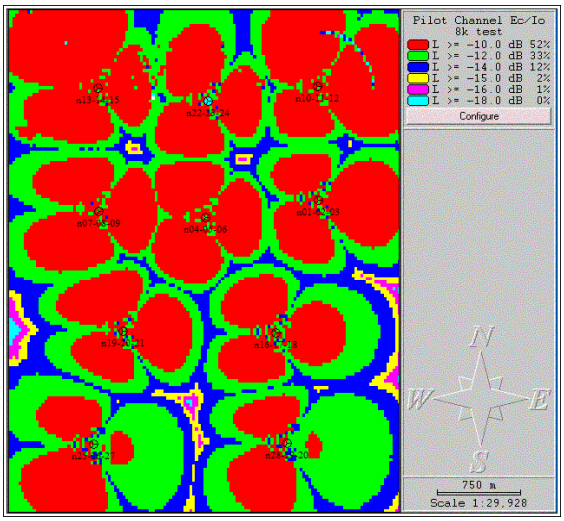


Figure 6-33. Pilot channel Ec/Io (dB).

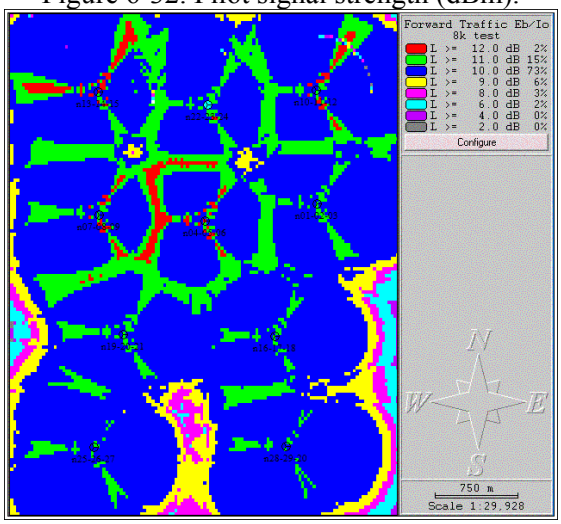


Figure 6-34. Downlink traffic channel Eb/Io(dB)

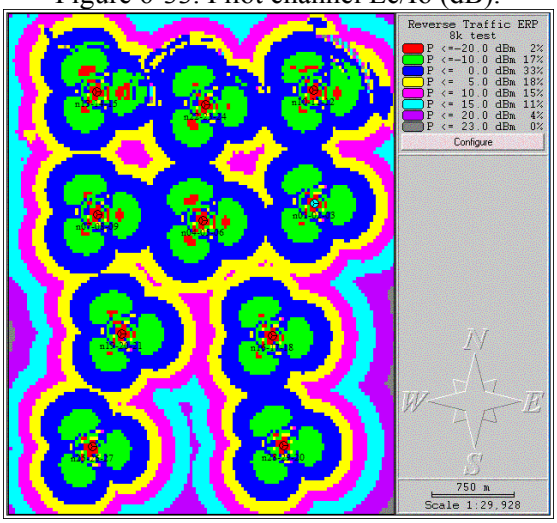


Figure 6-35. Uplink required MS power (dBm).

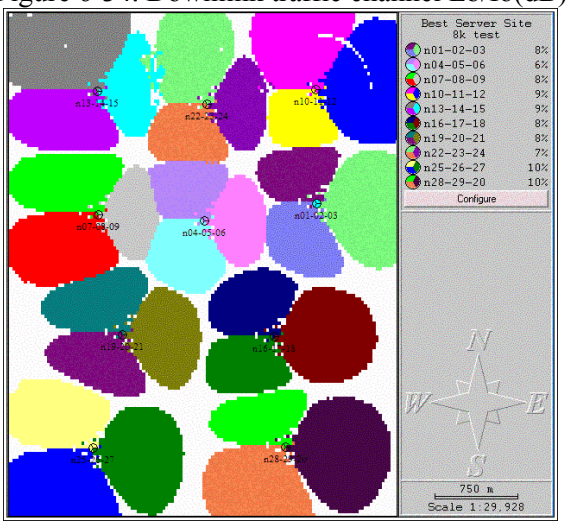


Figure 6-36. Pilot best server.

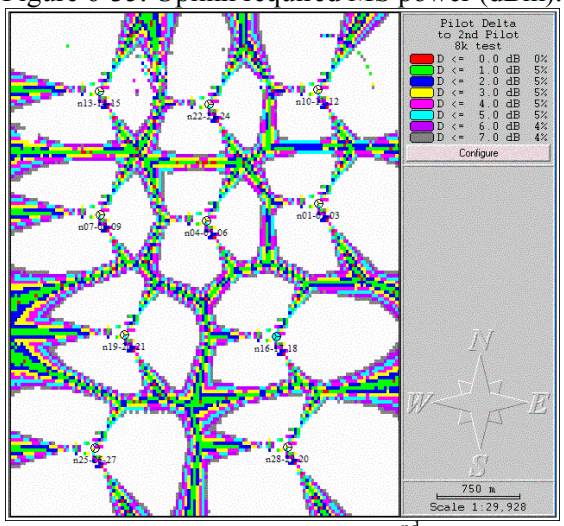


Figure 6-37. Pilot delta to 2nd pilot.

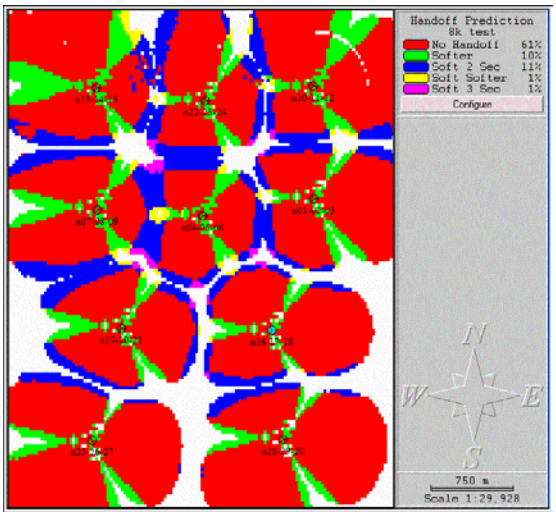


Figure 6-38. Handoff areas.

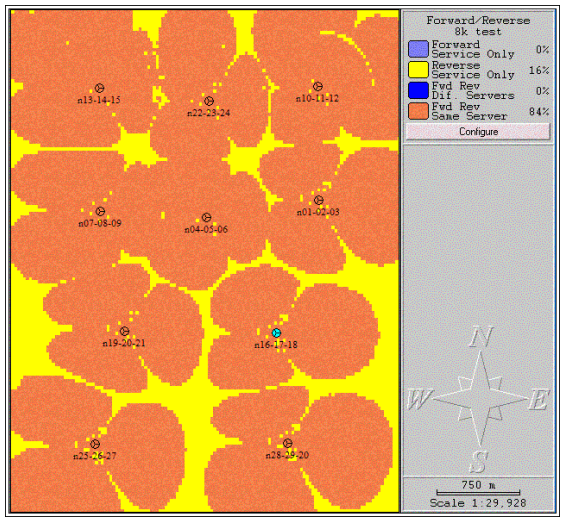


Figure 6-39. Down/uplink coverage comparison

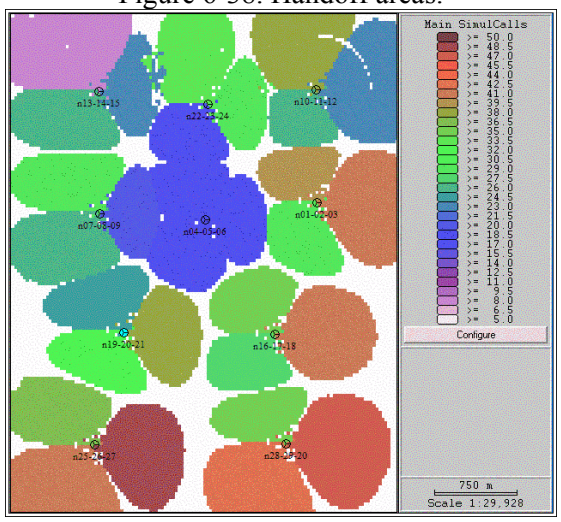


Figure 6-40. Number of users served.

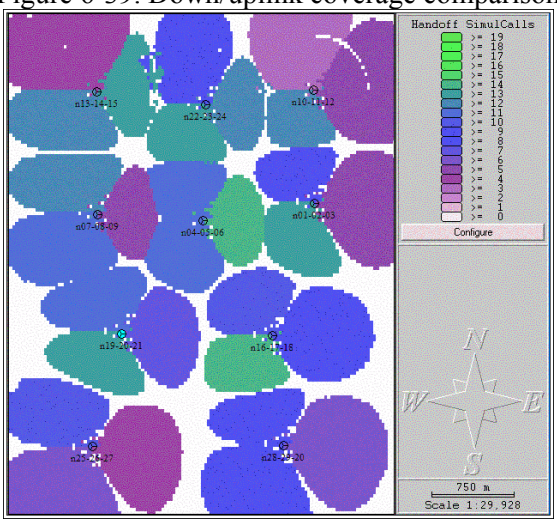


Figure 6-41. Number of SHO connections.

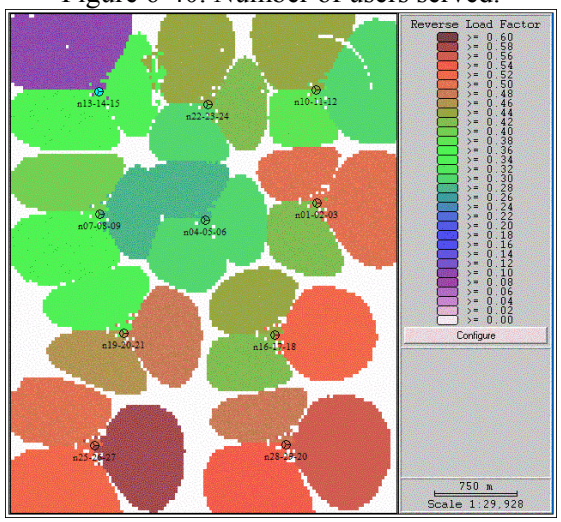


Figure 6-42. Uplink load factor.

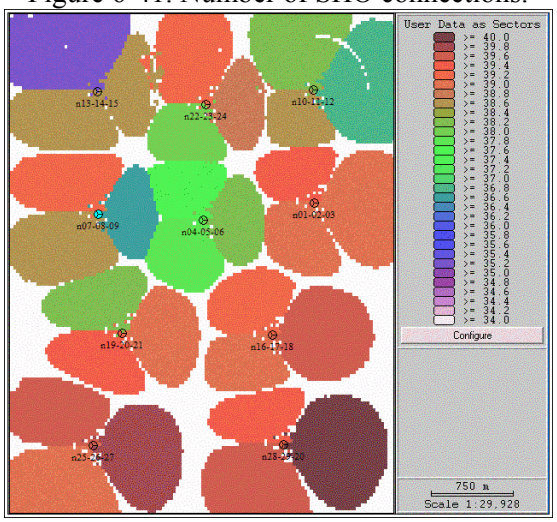


Figure 6-43. Total BS Tx power (dBm).

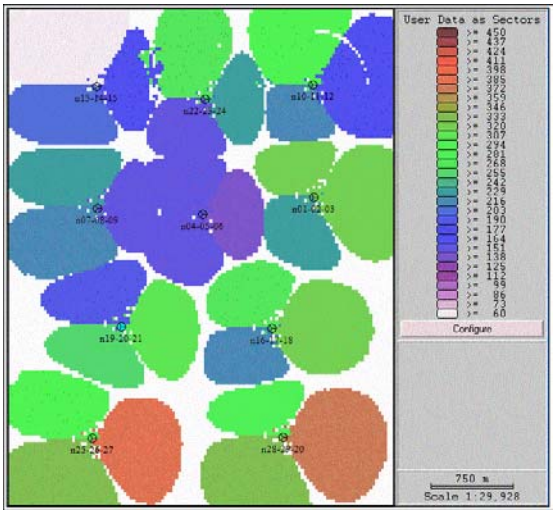


Figure 6-44. Sector throughput (kbps).

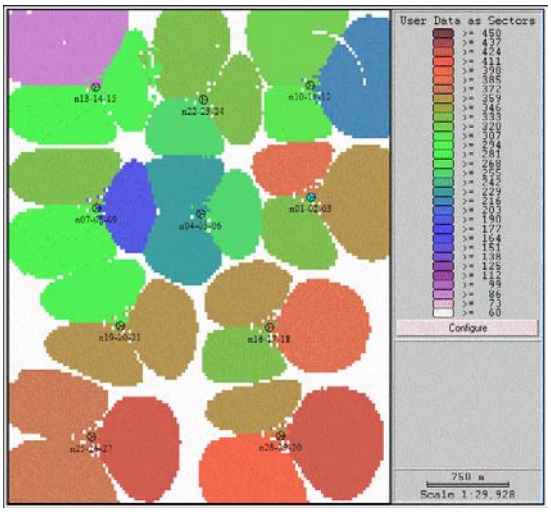


Figure 6-45. Sector total rate (kbps).

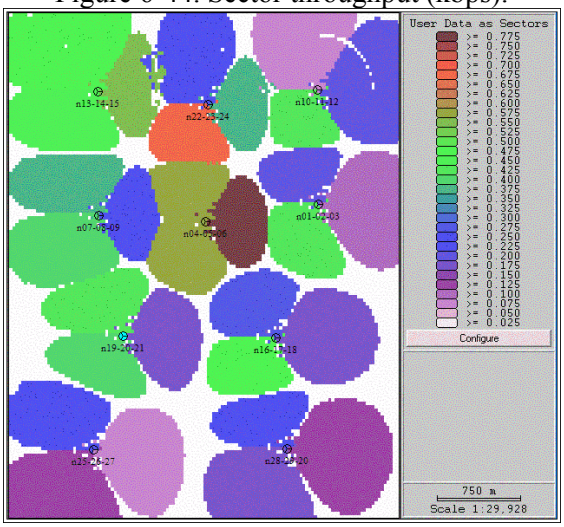


Figure 6-46. Sector SHO overhead.

6.7.3. Discussion

The main focus for discussion in this section is to compare the results observed for Scenario B, which used heterogeneous traffic distribution with those observed on Scenario A, which used homogenous traffic. To do that, we should keep in mind the different traffic grids used in each scenario, which were illustrated in Figure 6-8. In the heterogeneous distribution, we observe that the highest traffic concentrations are in the North-East corner and in the diagonal that connects it to the southwest corner. This high traffic stripe corresponds roughly to the base-stations numbers: n10-11-12, n01-02-03, n16-17-18 and n25-26-27. On the other hand, the area with lowest traffic density corresponds to the northwest corner of the system, particularly concentrated in sector 15 (third sector

of site n13-14-15). Other sectors are in intermediate density areas. The sector concentration areas and their identifications are illustrated in Figure 6-47.

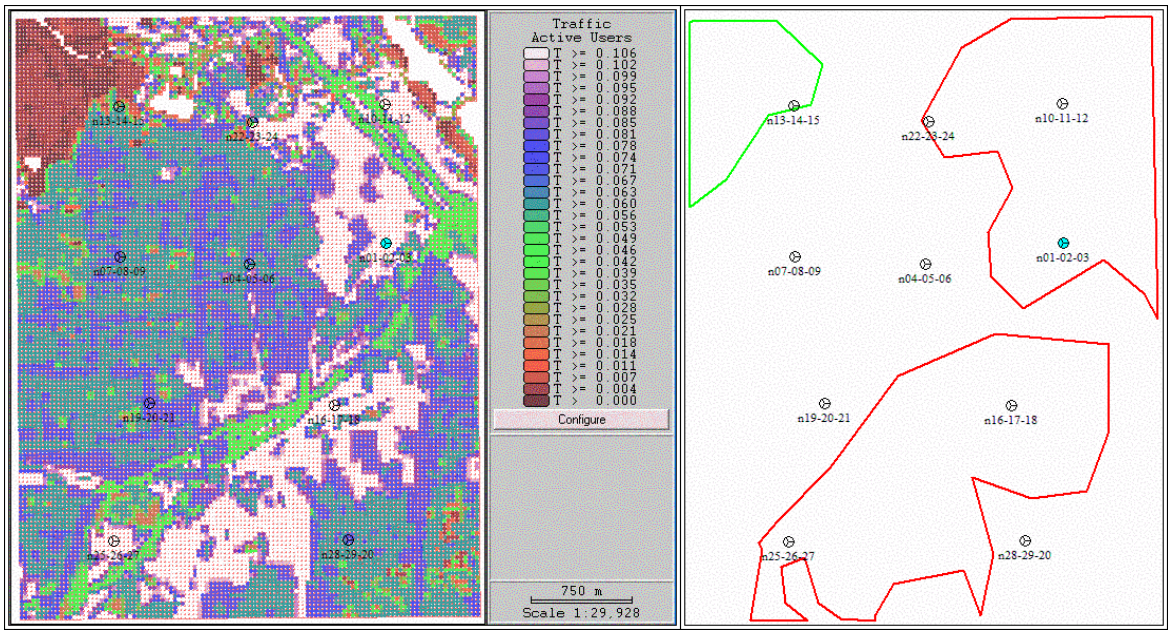


Figure 6-47. Sector with highest (red) and lowest (green) traffic densities.

By comparing the results obtained in this scenario with those from a regular traffic grid (Scenario A), we observed that mean values for traffic carried, overall throughput, load factors and transmitting powers have not varied much, as the same overall traffic has been offered to the system in both cases.

However, as expected, traffic now concentrates more heavily in a few sectors, reducing their footprints (as in sector 11). As expected, some gaps between sectors were found to have increased while coverage for other sectors has actually also increased (as in sector 15). In other words, for this system to have a better traffic balance between sectors, the RF configuration should take into account those irregularities in traffic distribution, where some tilts could be reduced in order to allow more overlap between sectors, and some pilot powers could also be individually adjusted for the same reason. While the topic of changing the system configuration in order to improve performance will be further discussed in Chapter 7, the points observed here strongly illustrate the fact that in a system where traffic cannot be assumed to be homogeneous, network configuration and optimization have to take this characteristic into account.

Overall, we observed that the main indicators of traffic load in the system, namely the total transmit power in the downlink, and the load factor in the uplink, have achieved wider ranges of variation in the heterogeneous scenario than in the homogeneous scenario. This observation is clear

from comparison of Table 6-10 and Table 6-18, where it is seen that both variables present higher standard deviations in Scenario B. In the same direction, we observe from comparison between Figure 6-27 and Figure 6-43 (total transmit BS transmit power for Scenarios A and B, respectively), that while powers range from 37.5 to 40 dBm in the homogenous case, it ranges from about 35 to 40 dBm in the heterogeneous case. Likewise by comparing Figure 6-26 with Figure 6-42 (uplink load factor for Scenarios A and B, respectively), we observe that the uplink load factor ranges from 0.25 to 0.6 in the homogenous case, it ranges from about 0.1 to 0.6 in the heterogeneous case.

6.8. Scenario C: One Class, Heterogeneous Traffic, Alternative Simulator:

6.8.1. Inputs and Assumptions

A summary of the main input parameters used for Scenario A is shown in Table 6-19.

Table 6-19. Summary of Input Parameters for Scenario C

Input Parameters (NPSW)	Scenario C
Traffic Simulation Parameters	
Traffic Distribution (homogeneous/heterogeneous)	Heterogenous
Number of 8 kbps users	1140
Number of 64 kbps users	0
Number of 144 kbps users	0
Throughput Requested (kbps)	9120
Number of snapshots	1
Maximum number of iterations	20
Random seed for snapshot	1
System Parameters	
Maximum allowed rate per sector (for hard-blocking) [kbps]	no limit
Maximum number of channel elements per sector	no limit
Maximum Load Factor on Downlink (load control)	1 (no limit)
Maximum Load Factor on Uplink (load control)	1 (no limit)
Maximum number of failures per connection	no limit
Maximum BS total power (W)	10
User Profile Parameters (8kbps users)	
Required Eb/No downlink	10 dB
Required Eb/No uplink	6.5 dB
Maximum power allowed per RAB (W) downlink	0.2
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0

Sector Parameters	
Pilot Power (W)	1
Other common channels power (W)	1
Orthogonality factor	0.5
Maximum physical power allowed per RAB (W) downlink	1.5
Net Tx Gains (Gains - Losses) (dB)	1
Net Rx Gains (Gains - Losses) (dB)	3
Receiver Noise Figure (dB)	5

6.8.2. Equivalency Concerns

As mentioned before, the scenario presented here uses a different simulation tool compared to all other scenarios evaluated in this chapter. The objective of this analysis is to compare results with those obtained for Scenario B, where inputs are roughly equivalent. However, since no two simulation tools are identical in terms of input parameters or algorithms, we do not expect to find results that are identical, but rather to be able to observe results that are consistent with each other.

In the process of configuring equivalent scenarios in both tools, many topics have been of concern, of which the most important are highlighted here.

6.8.2.1. Database and Coordinate System Equivalency

While typical commercial network planning applications (such as CelPlanner) use a geodetic coordinate system (based on latitude-longitude), the coordinates in NPSW refer to a generic metric reference grid. In order to deal with issues such as projections and coordinate translations, we configured CelPlanner to a user-defined coordinate system, with origin coordinates centered in the region we wanted to simulate, i.e., using Shreveport downtown coordinates. This way, all base-stations and user x and y coordinates could refer to the coordinate system in both tools.

6.8.2.2. Conversion of Antenna Patterns

The antenna pattern file corresponding to the 65-degree antenna used for all sectors in this system had to be converted to CelPlanner format in order to allow full equivalency with the pattern available in the NPSW database. Although both tools use text files as the interface format to import/export antenna data, the contents of those files, units and headers differ considerably, and therefore extra

care had to be taken in order to understand the NPSW format before conversion was performed. The converted pattern is illustrated in Figure 6-48.

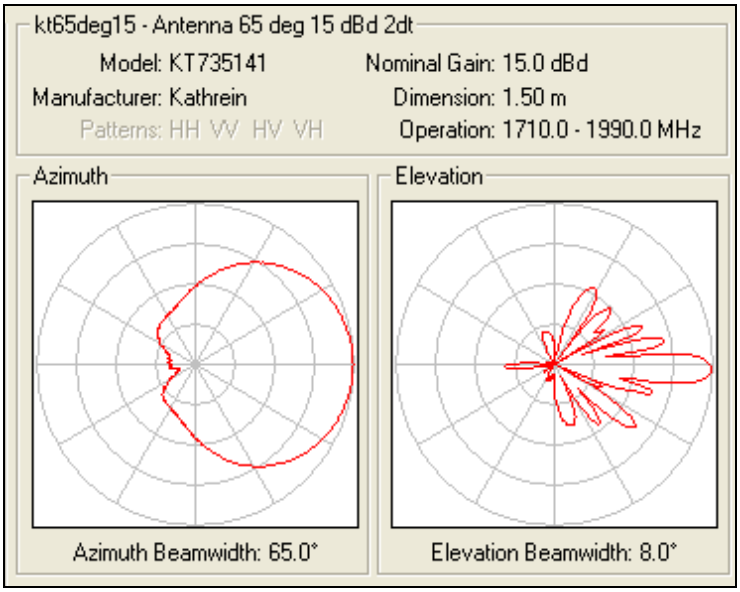


Figure 6-48. Antenna pattern (65 degrees) used on both cases.

6.8.2.3. Propagation Model Equivalency

Both tools used support multiple propagation models, some more sophisticated than others. In CelPlanner, propagation models are configurable to use high-resolution topography and morphology database, at multiple resolutions. Multiple models are available and most of them are user-adjustable through the use of model parameters, or self-adjustable through automatic calibration algorithms based on field-collected signal strength measurements. In NPSW, the range of propagation models and input databases available is more restricted, and resolution for prediction analysis is not easily adjustable. For that reason, we chose to configure in both tools the least complex model available, which correspondeds, in this case to a single-slope (40 dB/decade) model insensitive to terrain and morphology databases. Using those models, we created propagation predictions for all sites in the system, and compared their coverage radius for the same signal strength thresholds. Despite differences in the predictions granularity (resolution), cell radiuses were found to be very similar, as illustrated in Figure 6-49. In that picture, presentation areas, legend colors and thresholds were configured to provide the easiest visualization and comparison possible.

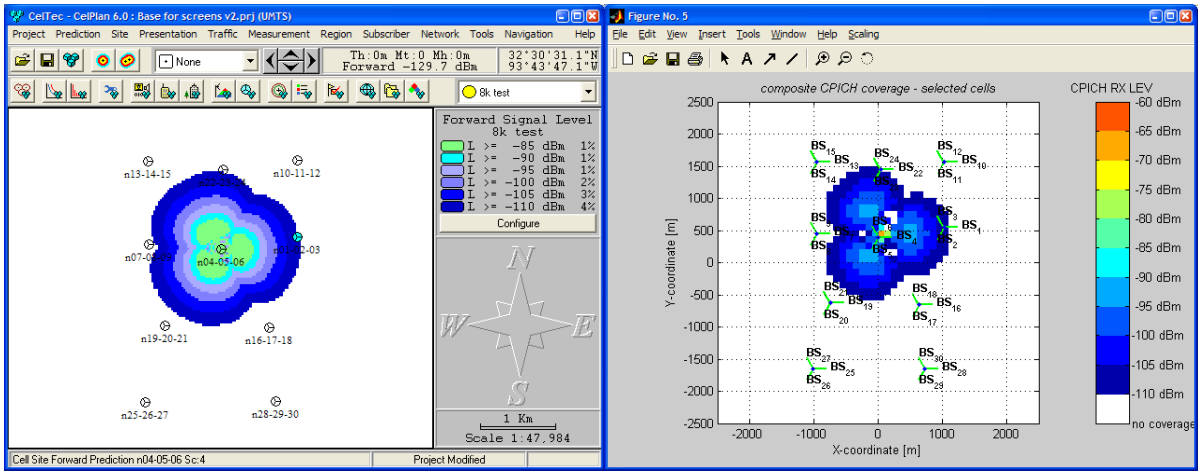


Figure 6-49. Configuring propagation model equivalency. Left: CelPlanner prediction, right: NPSW prediction.

6.8.2.4. Snapshot Conversion to MS Database

Once the coordinate system had been dealt with, the conversion of the user database to be used in NPSW was generated from the log file of one single snapshot generated in CelPlanner. The equivalent user distribution load in each tool is shown next (Figure 6-50).

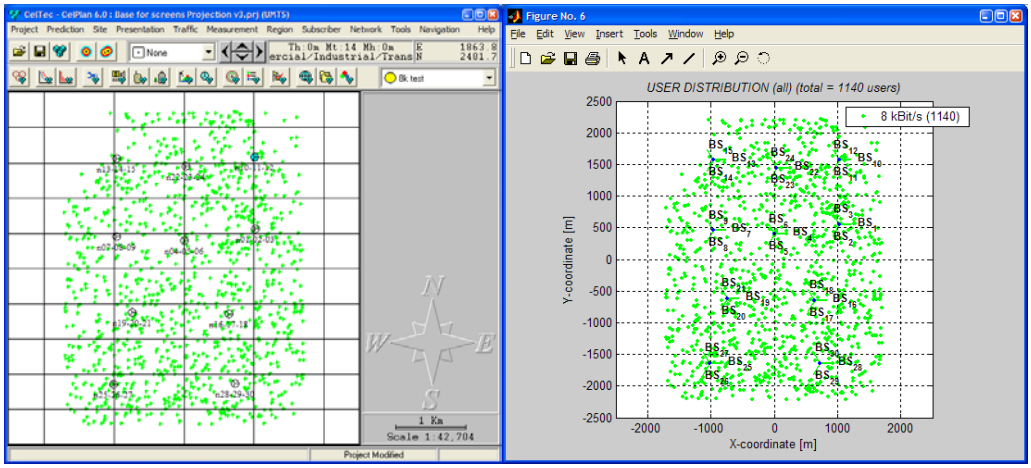


Figure 6-50. Single snapshot, left: used in CelPlanner for Scenario B, right: used in NPSW for Scenario C.

Besides converting the user locations from one snapshot, many parameters related to the user configuration (such as data rate, terminal power limits, antenna configuration, etc.) were mapped for

equivalency between tools. One example of the input format for the MS database in NPWS is shown next (Figure 6-51).

xPos	yPos	Ground Height	Ant. Height	Max TX Power	Min TX Power	Ant Gain	Body losses	Bit rate UL	Bit rate DL	Used	Speed
m	m	m	m	dBm	dBm	dBi	dB	b/s	b/s	Carrier	km/h
1570.888	933.721	0	1.50E+00	27	-43	0	2	144000	144000	1	3
150.238	697.281	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-1167.992	-1678.779	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-1055.632	1166.771	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1342.458	-649.549	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1218.548	-1509.179	0	1.50E+00	27	-43	0	2	144000	144000	1	3
653.658	700.251	0	1.50E+00	27	-43	0	2	144000	144000	1	3
421.258	-1768.859	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-815.382	665.381	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-954.332	-1028.679	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1283.758	628.861	0	1.50E+00	27	-43	0	2	144000	144000	1	3
491.648	-385.709	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-863.422	-1711.799	0	1.50E+00	27	-43	0	2	144000	144000	1	3
470.808	2192.291	0	1.50E+00	27	-43	0	2	144000	144000	1	3
487.768	926.581	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-814.482	-1009.519	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-1083.042	1245.571	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1181.458	2195.601	0	1.50E+00	27	-43	0	2	144000	144000	1	3
825.918	722.871	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-570.512	-528.249	0	1.50E+00	27	-43	0	2	144000	144000	1	3
-1505.542	-1104.739	0	1.50E+00	27	-43	0	2	144000	144000	1	3
846.178	-1404.359	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1416.988	-740.869	0	1.50E+00	27	-43	0	2	144000	144000	1	3
756.368	105.641	0	1.50E+00	27	-43	0	2	144000	144000	1	3
128.578	-1689.859	0	1.50E+00	27	-43	0	2	144000	144000	1	3
1328.968	417.241	0	1.50E+00	27	-43	0	2	144000	144000	1	3
374.618	1289.571	0	1.50E+00	27	-43	0	2	144000	144000	1	3

Figure 6-51. Mobile terminals database, as used in NPSW input.

6.8.2.5. Network Data Cconversion to BS database

The mapping of the network database included the equivalency in parameters related to cell location, sector configurations (powers, azimuth, tilt, etc.), RF parameters (cable losses, amplifier gains, etc.) and power configurations (pilot, other common channels, limits for the RAB traffic channels). One difficulty found in this phase was that not all parameters could be directly mapped between tools, due to differences in modeling and algorithms. For that reason, many parameters that are located in the BS database in CelPlanner were found in the system parameters setup (or sometimes hard-coded in the Mmatlab routines) in NPSW. Since documentation was not fully available for that implementation, mapping in this phase was done with some of the parameters being assumed on a reasonable basis. Figure 6-52 and Figure 6-53 show base-station parameters are configured in each tool.

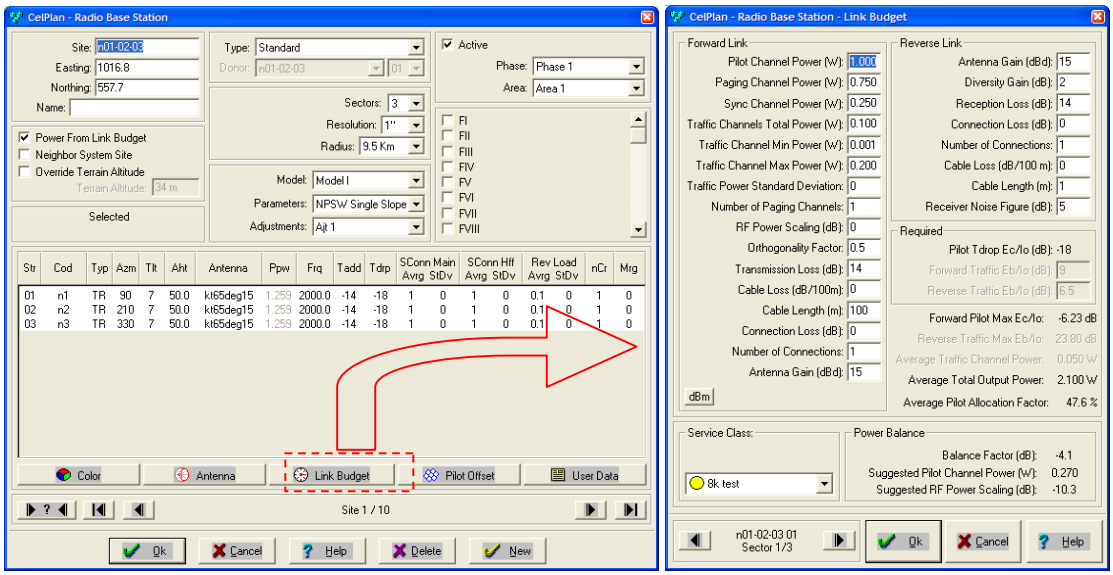


Figure 6-52. Base-station data configuration in CelPlanner.

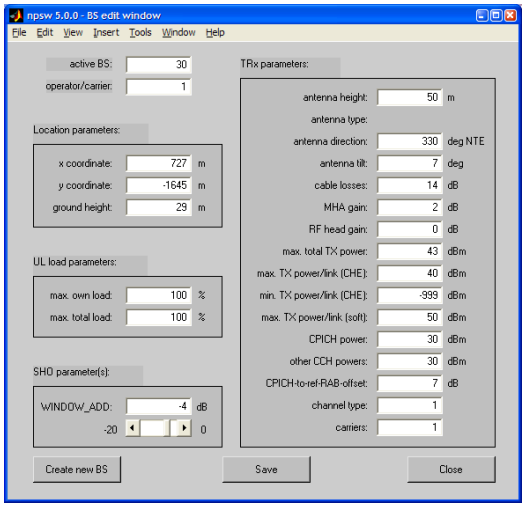


Figure 6-53. Base-station data configuration in NPSW.

6.8.2.6. System Parameters Equivalency

NPWS is a Matlab based tool, and many other parameters had to be changed/configured directly in the corresponding NPSW models, since dialogs were not available for all configurable parameters that needed to be mapped for equivalency with CelPlanner simulations. The main modules that had to be edited for equivalency in NPSW included the system parameters module (*npswsys.m*), which initializes parameters global to the system, the global parameters module (*npswini.m*), which includes

parameters common base stations parameters, common mobile station parameters, convergence and setup parameters for the uplink and downlink iterations, propagation model parameters, etc; and the link performance tables (*[channel type]LinkPerfTables.m*), which contains information from link-level simulation results such as Eb/No requirements, orthogonality factors, SHO gains, etc. An illustration of the format of the modules listed is shown (Figure 6-54).

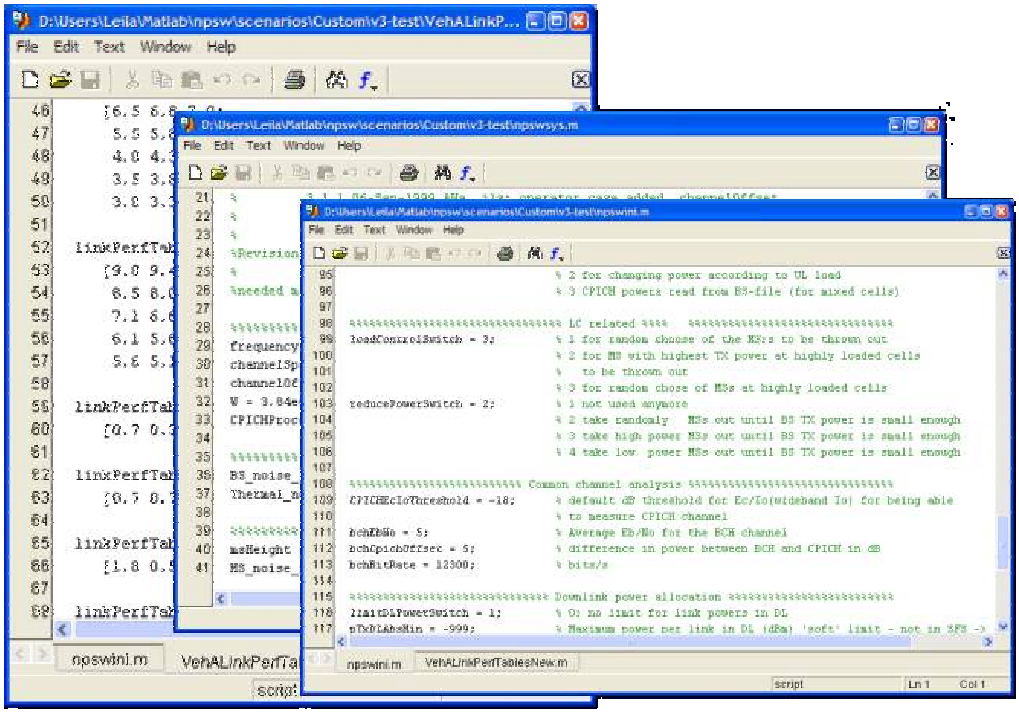


Figure 6-54. System parameters configuration in NPSW.

6.8.2.7. Other Topics of Concern Regarding Equivalency

Besides all topics covered in mapping equivalency between the tools, there were additional topics that could not be perfectly mapped due to differences in the nature of both tools with regards to how the system is modeling and the simulation algorithms themselves. Besides, even though the internal algorithms of CelPlan are well know by the author, only high-level documentation [64] was available to explain the algorithms implemented in the NPSW tool. As mentioned before, some reasonable assumptions were made whenever possible, but some particular topics for concern with regards to tool equivalency are listed below.

- **Iteration algorithms:**

Both tools are static simulators that estimate overall interference in each direction, accept or reject connections based on some specific criteria and estimate interference generated again in the next iteration. However simple that may sound, there are many ways to implement the convergence loops, and apparently the implementation in NPSW has separated, for simplicity, the downlink and uplink modules, that run alternately until convergence is achieved in both. Although this simplifies calculations and speeds convergence, there are some intricacies between both links that are missed in this solution, causing the system to actually diverge in some instances, where whole sectors could "escape" the loop and reject all connections. In fact, we had to reject some original scenarios being evaluated with NPSW because some of the sectors would fall into this situation.

Besides, it has been noticed that some "convergence accelerators" are used in some instances in the NPSW algorithm, mainly when excess load causes the iterative process to oscillate between high and low load. These accelerators are typically implemented in the format of failure tests, where if a user is repeatedly oscillating between the state of served and not-served, it is put in permanent outage after a few iterations. Although the "maximum number of failures" is a parameter in the tool, it had to be configured to a low value in order for simulations to converge in NPSW. In CelPlanner, the convergence loop is integrated, and pilot, downlink and uplink coverages have to converge together, which apparently makes the convergence loop more robust. For that reason, convergence for the equivalent scenario was achieved without using a "maximum number of failures" limit, i.e., the algorithm achieves convergence in fewer iterations than the configured limit.

- **Convergence criteria:**

Typical convergence criteria for systems using CDMA technology are the transmit powers in the downlink and the receive powers in the uplink, both from the point of view of the base-stations. However, there are again different ways to implement those tests, and they could be based on local convergences per sector, per base-station or for the entire system. It was not clear based on available documentation how equivalent both algorithms are.

- **Nomenclature and units equivalency:**

To the best of the author's knowledge, parameters have been mapped based on conventional nomenclature used in the UMTS industry. However, there is still the possibility that some of the units and parameter definitions were not completely equivalent between both tools.

6.8.3. Results

Once simulation was executed, results are automatically saved in NPSW in the format of internal Matlab variables. For that reason, not all variables available for the previous scenarios will be available for comparison here.

Table 6-20 shows a brief summary of the main statistics resulting from this scenario simulation, which were available for display or analysis at the end of the simulation process.

Table 6-20. Summary of Results for Scenario C

	Avg. Sector	Sum
Requested Connections	38.0	1140
Served Connections	37.0	1110
SHO Connections	14.3	429
Total Connections	51.3	1539
Throughput (kbps)	296.0	8880
Total Rate (kbps)	410.4	12312
SHO overhead	38.6%	38.6%
Not served due to Downlink Pwr limit	1.0	30
Not served due to Uplink Pwr limit	0.0	0
Total Dw TCHs Power (W)	4.07	122.10
Uplink Load Factor	0.34	
Uplink Noise Rise	1.81	

Similarly to what is done in CelPlanner, after computing the results related to total transmitted and received powers per sector, a set of predictions is performed, and geographical illustrations of those results are shown next (Figure 6-55 through Figure 6-65).

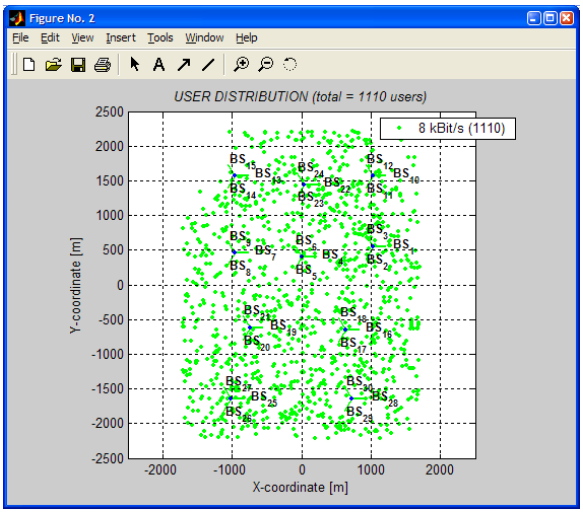


Figure 6-55. Served users.

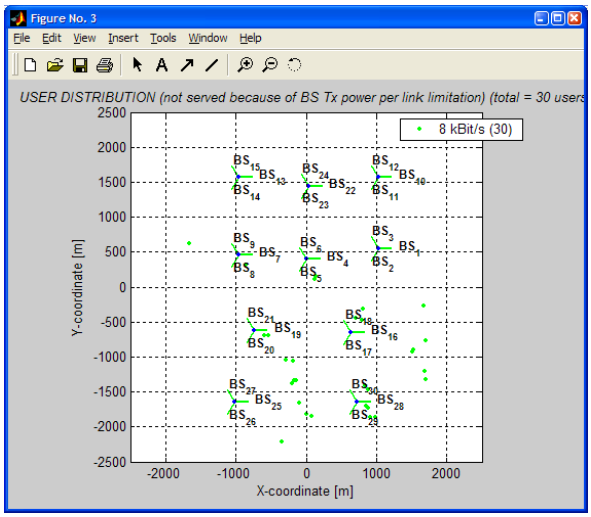


Figure 6-56. Rejected (downlink power limit).

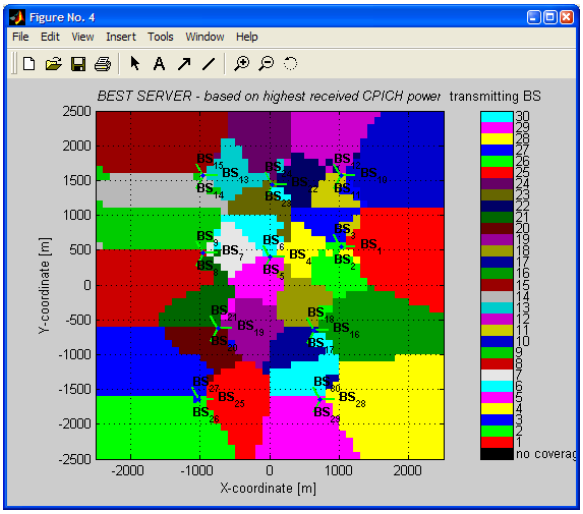


Figure 6-57. Best server (pilot power).

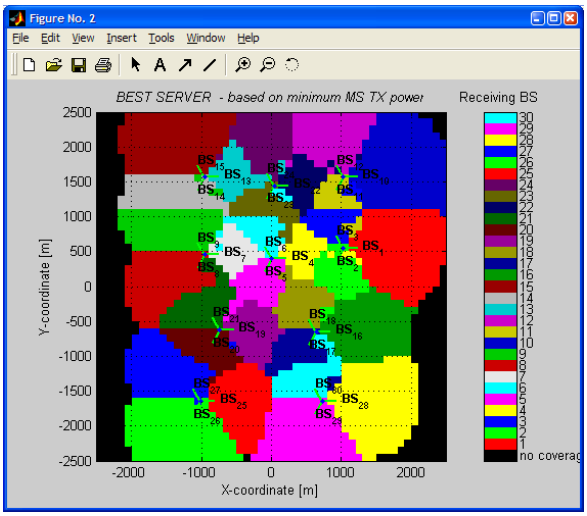


Figure 6-58. Best server (uplink).

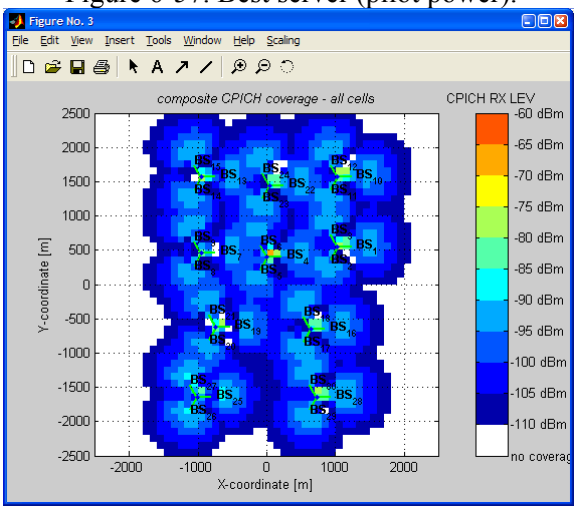


Figure 6-59. Pilot signal strength (dBm).

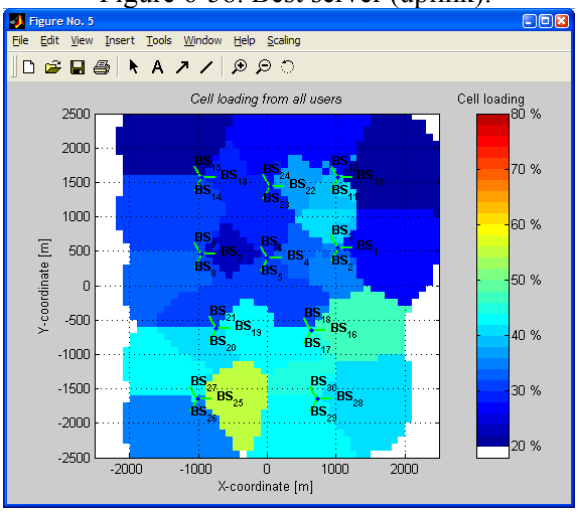


Figure 6-60. Uplink load factor per sector.

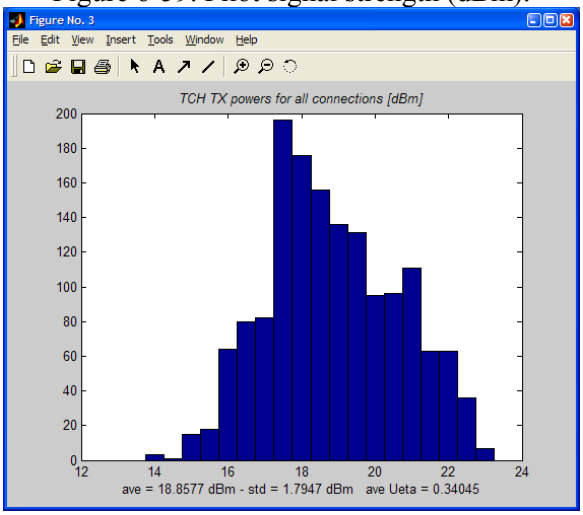


Figure 6-61. Histogram:Tx powers for all connections.

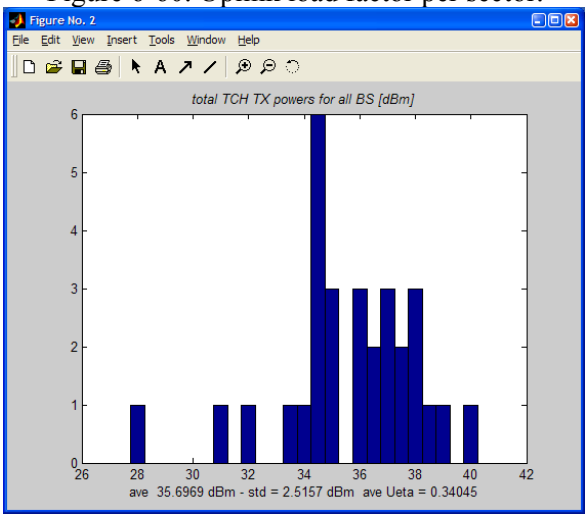


Figure 6-62. Histogram:Total BS Tx powers all BSs.

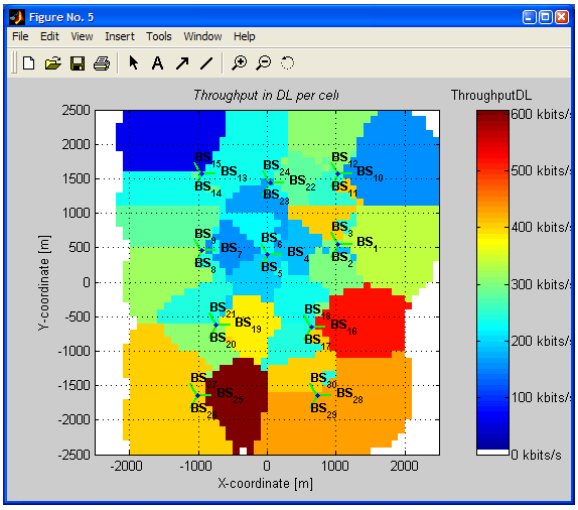


Figure 6-63. Sector throughput

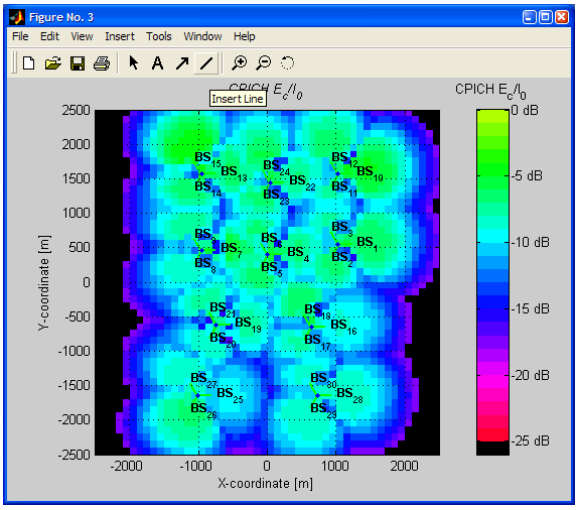


Figure 6-64. Pilot Ec/Io (dB).

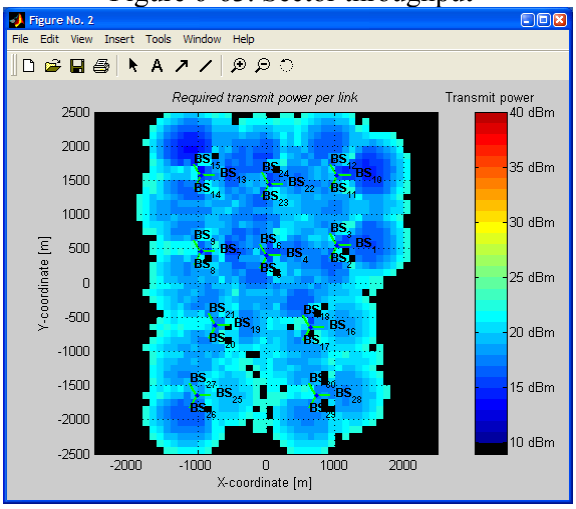


Figure 6-65. Required MS Tx power (dBm).

6.8.4. Discussion

Although fully equivalent results were not expected, results obtained for Scenario C, simulated in NPSW, are consistent in many aspects to those obtained for Scenario B, simulated using CelPlanner. The main observations from the comparison between both scenarios are highlighted below.

- Both simulations achieved a large number of served calls, namely 925 in CelPlanner against 1110 in NPSW, out of 1140 calls requested, a 16% difference in service rate.
- In both tools, the bulk reason for calls being rejected was the limit in downlink available power per traffic channel (per RAB). In NPSW, that reason accounted for all call rejections, while in CelPlanner, 98% of call rejections were due to the same reason, while 2% were due to limit in the total sector power.

- As a consequence of the topic above, it is observed that pilot coverage and uplink coverage were not limiting factors in this scenario.
- Uplink load factor per sector in NPSW was found to range from a minimum of 0.06 (on sector 15), to a maximum of 0.54 (on sector 25). In CelPlanner, the minimum uplink load was found to be 0.11 (also for sector 15), and the maximum was 0.59 (for sector 25 as well). This shows a good agreement in terms of the uplink load results.
- Soft handoff overhead values per sector in NPSW were found to range from 3% to 100%. The five sectors with lowest SHO overhead (ranging from 3% to 20%) were, in ascending order, sectors 1, 25, 26, 16 and 12. The highest SHO overhead (at around 100%) were found in sectors 4 and 23. In CelPlanner, the lowest SHO overhead factors (ranging from 8% to 14%, were found in sectors 12, 25, 1, 28 and 26, while the highest SHO overhead factors (68% to 78%) were found in sectors 23 and 4. Therefore, although absolute values for SHO overheads were found to be spread over a wider range in NPSW than in CelPlanner, both tools were very consistent in terms of which sectors carried more or less overhead.

Overall, it was observed that results were very consistent qualitatively and to a lesser degree in terms of absolute values. We conclude that the consistency found between results in this particular scenario is enough to validate the use of either tool from this point down, as both tools have shown to be sensitive to the same system aspects when subjected to the same system limitations.

Later in this chapter we also show that NPSW results are in good agreement with the more complex configuration simulated in Scenario D, which we ran using both tools.

6.9. Scenario D: Multiple Classes, Heterogeneous Traffic Distribution:

6.9.1. Inputs and Assumptions

This is a more complex scenario where three different classes are simultaneously active in the system. The same network configuration (base stations and user distribution) has been kept from the previous scenarios. However, the original 1140 users were split into three groups, with 720 users on the lowest rate user profile (8 kbps), 240 users in the mid-rate profile (64 kbps) and 180 users on a high-rate profile (144 kbps). Additionally, even though the maximum allowed power per sector remained the same (10 Watts) compared to previous scenarios, here there is a distinction in the maximum power allowed per Radio Access Bearer, with more power being allowed for the higher

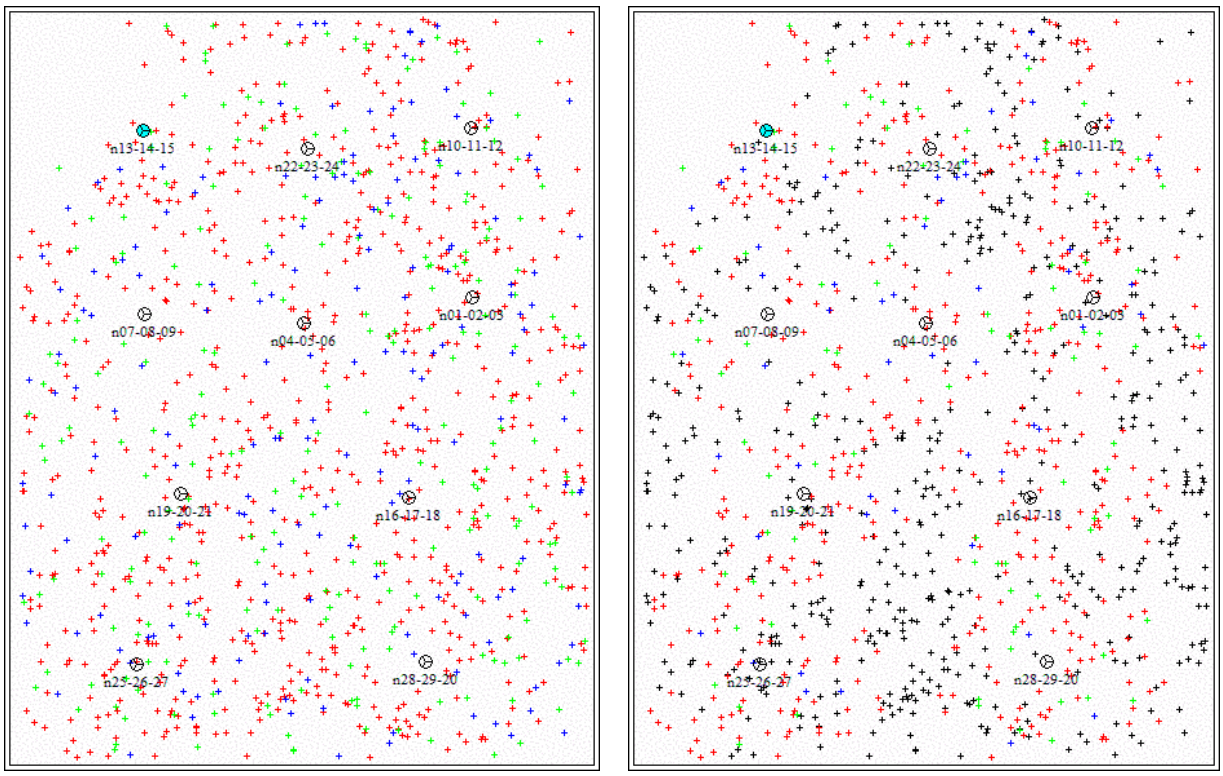
rate classes (0.2, 0.84 and 1.5 Watts, respectively). However, we have assumed that the maximum uplink power per user (from mobile terminal) is still common for all profiles (0.5 Watts). A summary of the main input parameters used for Scenario D is shown in Table 6-21.

Table 6-21. Summary of Input Parameters for Scenario D

Input Parameters	Scenario D
Traffic Simulation Parameters	
Traffic Distribution (homogeneous/heterogeneous)	Heterogenous
Number of 8 kbps users	720
Number of 64 kbps users	240
Number of 144 kbps users	180
Throughput Requested (kbps)	47040
Number of snapshots	1
Maximum number of iterations	20
Random seed for snapshot	1
System Parameters	
Maximum allowed rate per sector (for hard-blocking) [kbps]	2500
Maximum number of channel elements per sector	999 (no limit)
Maximum Load Factor on Downlink (load control)	1 (no limit)
Maximum Load Factor on Uplink (load control)	1 (no limit)
Maximum number of failures per connection	100
Maximum BS total power (W)	10
User Profile Parameters (8kbps users)	
Required Eb/No downlink [dB]	9
Required Eb/No uplink [dB]	6.5
Maximum power allowed per RAB (W) downlink	0.2
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0
User Profile Parameters (64kbps users)	
Required Eb/No downlink [dB]	7.1
Required Eb/No uplink [dB]	4
Maximum power allowed per RAB (W) downlink	0.84
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0
User Profile Parameters (144kbps users)	
Required Eb/No downlink [dB]	6.1
Required Eb/No uplink [dB]	3.5
Maximum power allowed per RAB (W) downlink	1.5
Maximum power allowed per RAB (W) uplink	0.5
Receiver Noise Figure (dB)	8
Terminal Antenna gain (dBi)	0
Sector Parameters	
Pilot Power (W)	1
Other common channels power (W)	1
Orthogonality factor	0.5
Maximum physical power allowed per RAB (W) downlink	1.5
Net Tx Gains (Gains - Losses) (dB)	1
Net Rx Gains (Gains - Losses) (dB)	3
Receiver Noise Figure (dB)	5

6.9.2. Results

Once simulation was executed, the results have been saved in the format of screens and reports. The most relevant results of this analysis are shown next.



Traffic Offered:

Class 1 - Data 8kbps Portable:	720 users
Class 2 - Data 64 kbps Portable:	240 users
Class 3 - Data 144 kbps Portable:	180 users

Traffic Served:

Class 1 - Data 8kbps Portable:	469 users
Class 2 - Data 64 kbps Portable:	92 users
Class 3 - Data 144 kbps Portable:	56 users
Users not served:	523 users

Figure 6-66. Graphical result of snapshot for Scenario D.

From the simulated scenario, multiple statistics were collected for all sectors, in a format previously illustrated (Figure 6-14). A brief summary of the main statistics resulting from this scenario simulation is shown in Table 6-22.

Table 6-22. Summary of Results for Scenario D

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	15.6	5.7	469	3.1	1.6	92	1.9	1.1	56	20.6	5.7	617
SHO Connections	2.9	2.3	87	0.4	0.6	13	0.1	0.4	3	3.4	2.4	103
Total Connections	18.5	5.7	556	3.5	1.6	105	2.0	1.2	59	24.0	5.7	720
Throughput (kbps)	125.1	45.9	3752.0	196.3	100.7	5888.0	268.8	159.2	8064.0	590.1		17704
Total Rate (kbps)	148.3	45.9	4448	224.0	104.6	6720	283.2	171.2	8496	655.5	107.5	19664
SHO overhead	18.55%		18.55%	14.13%		14.13%	5.36%		5.36%	11.07%		11.07%
Not carried - Pilot channel coverage										0		0
Not carried - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried - Forward traffic channel power limit	6.7	5.0	201	3.0	2.2	90	2.5	2.1	74	12.2	8.6	365
Not carried - Forward sector total power limit	1.7	2.9	50	1.9	1.7	58	1.7	1.5	50	5.3	4.7	158
Not carried - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.8	0.9	83.5	2.4	1.2	71.5	2.4	1.5	72.6	7.6	1.2	227.6
Uplink Load Factor										0.55	0.1	
Uplink Noise Rise										3.55	0.7	

After saving the results into the project configurations, mainly those referring to total transmitted and received powers per sector, a full set of predictions was performed, and geographical illustrations of those results are shown in Figure 6-67 through Figure 6-93.

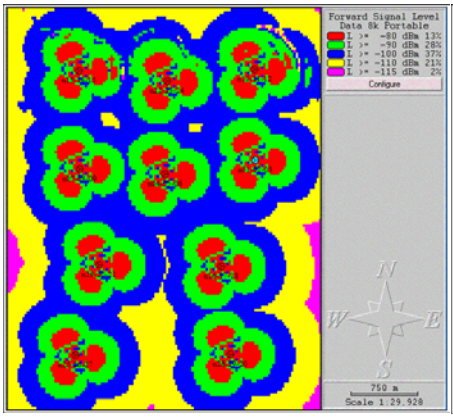


Figure 6-67. Pilot signal strength (dBm).

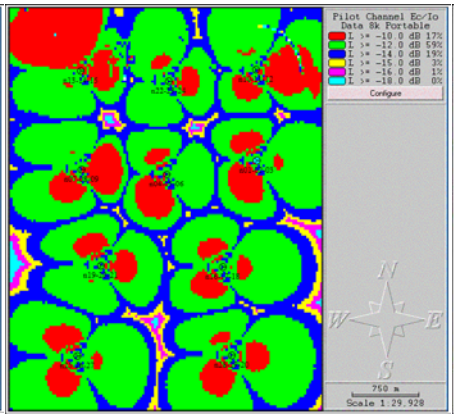


Figure 6-68. Pilot channel Ec/Io (dB).

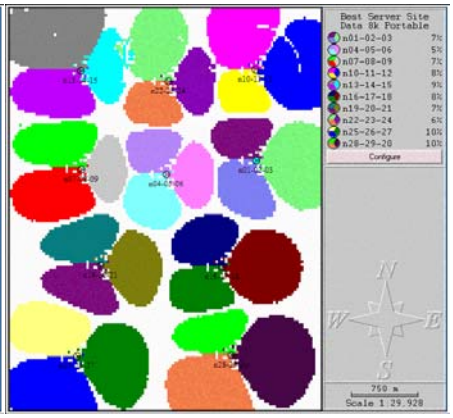


Figure 6-69. Pilot best server.

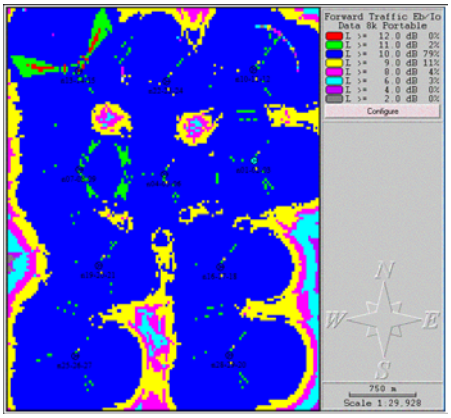


Figure 6-70. Downlink TCH Eb/Io (dB) – 8 kbps.

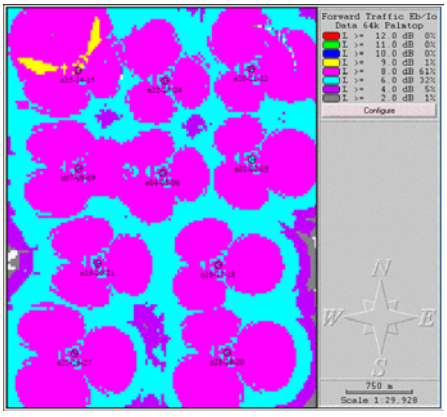


Figure 6-71. Downlink TCH Eb/Io (dB) – 64 kbps.

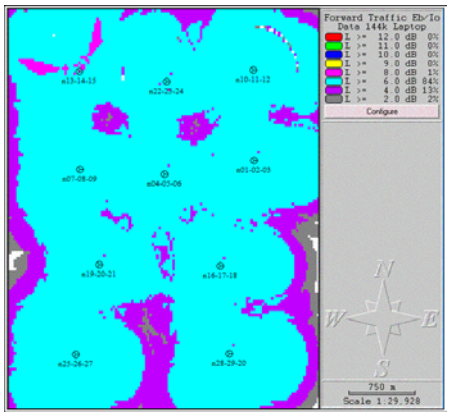


Figure 6-72. Downlink TCH Eb/Io (dB) – 144 kbps.

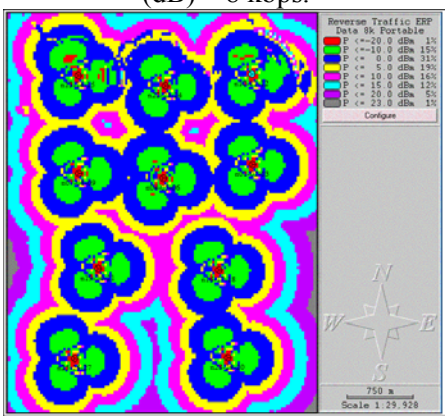


Figure 6-73. Uplink required MS power (dBm) – 8 kbps.

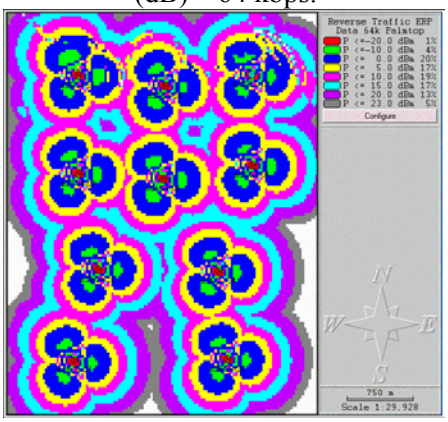


Figure 6-74. Uplink required MS power (dBm) – 64 kbps.

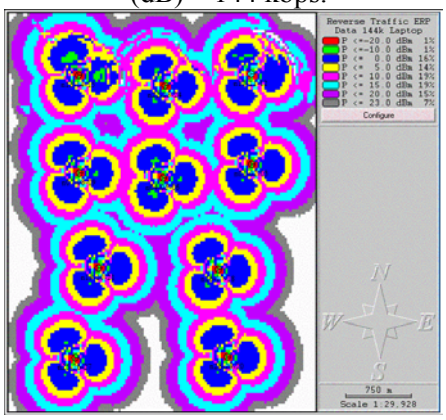


Figure 6-75. Uplink required MS power (dBm) – 144 kbps.

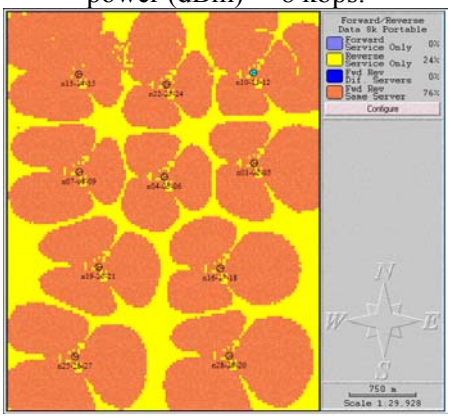


Figure 6-76. Downlink/uplink coverage comparison – 8kbps.

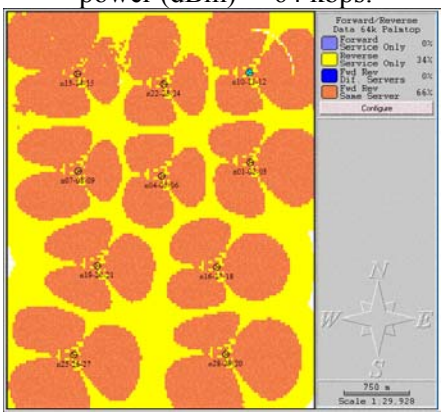


Figure 6-77. Downlink/uplink coverage comparison – 64kbps.

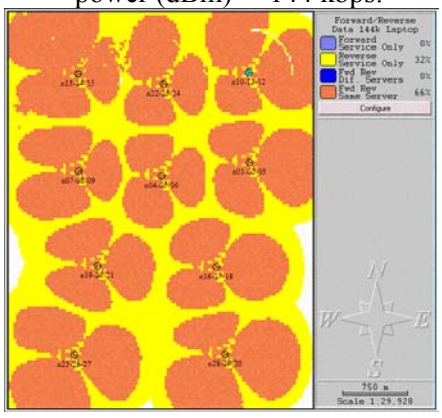


Figure 6-78. Downlink/uplink coverage comparison – 64kbps.

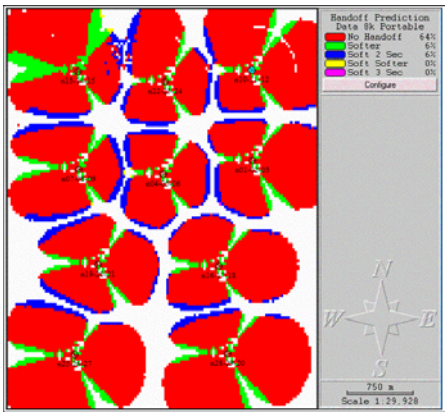


Figure 6-79. Handoff areas – 8 kbps.

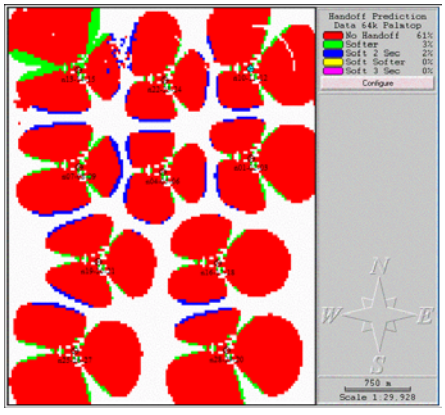


Figure 6-80. Handoff areas – 64 kbps.

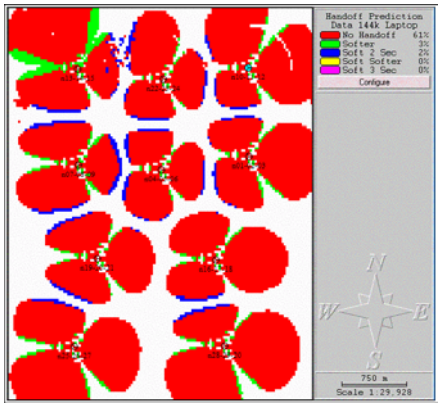


Figure 6-81. Handoff areas – 144 kbps.

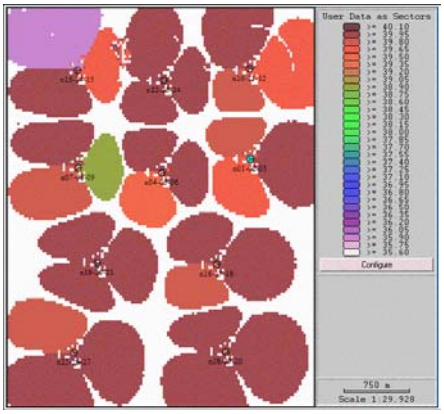


Figure 6-82. Total BS Tx power (dBm).

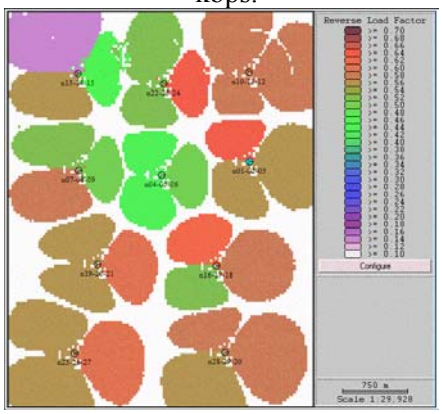


Figure 6-83. Uplink load factor.

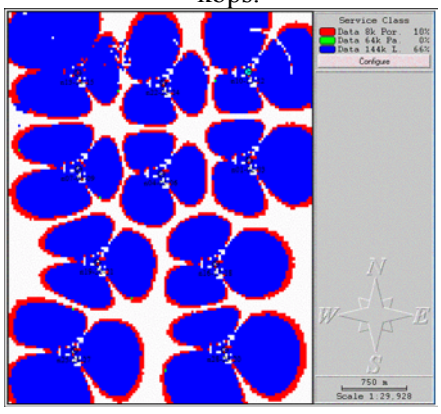


Figure 6-84. Service class coverage.

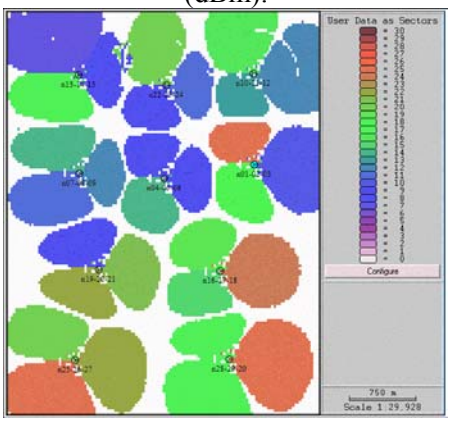


Figure 6-85. Number of users served - 8 kbps.

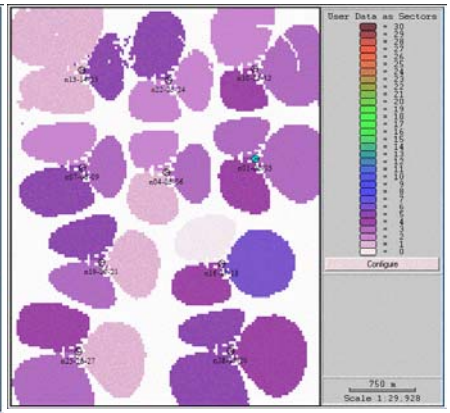


Figure 6-86. Number of users served - 64 kbps.

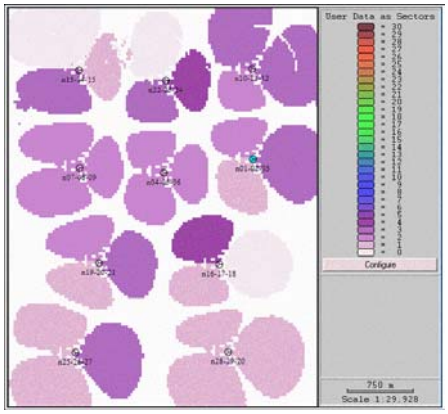


Figure 6-87. Number of users served - 144 kbps.

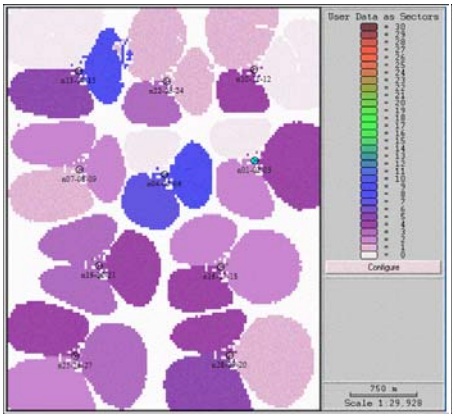


Figure 6-88. Number of SHO connections - 8 kbps.

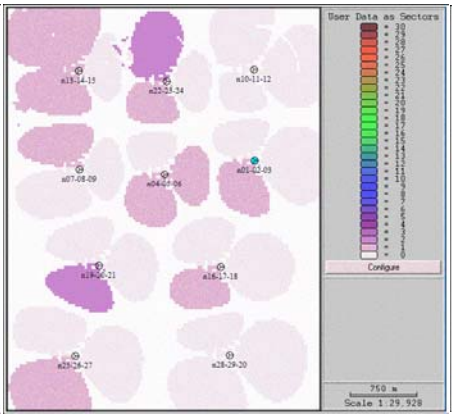


Figure 6-89. Number of SHO connections - 64 kbps.

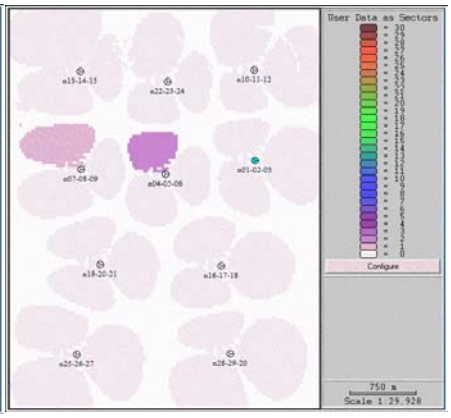


Figure 6-90. Number of SHO connections - 144 kbps.

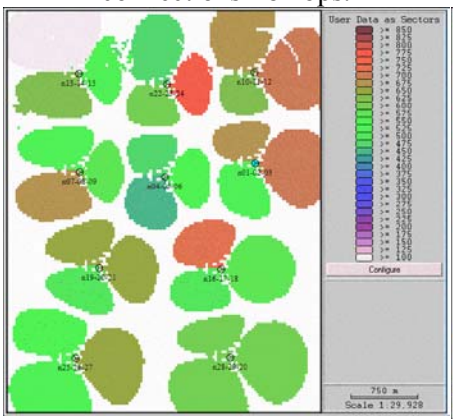


Figure 6-91. Sector throughput (kbps).

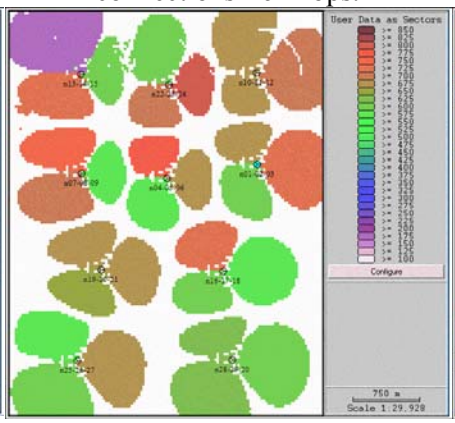


Figure 6-92. Sector total rate (kbps).

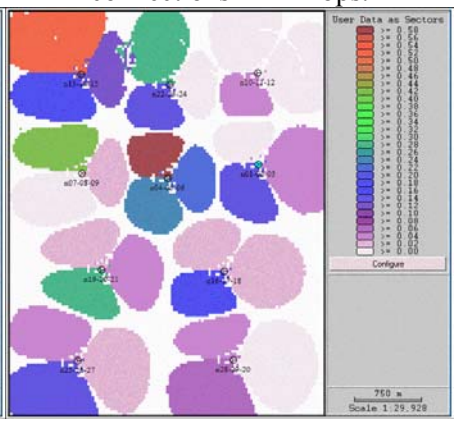


Figure 6-93. Sector SHO overhead.

6.9.3. Discussion

In this scenario, much higher aggregate throughput was obtained by the same Radio Access Network, which supported in this simulation a total of 17.7 Mbps, compared to 7.4Mbps in Scenario B. This is an interesting observation, considering that even in the single class scenario, not all offered traffic was carried. The increased capacity here is a consequence, mainly, of more power being allowed on a per channel basis (even though maximum sector powers have been kept at 10W), when in both cases we observed that the downlink was the limiting factor.

As a consequence of the increased load in the system, most sectors worked much closer to their maximum allowed power, raising the total power transmitted in the traffic channels per sector from 5.7 W (Scenario B) to 7.6 W, out of a maximum of 8W allowed (since all sectors were configured with fixed 2W for common channels).

Another consequence of increasing the traffic offered to the system is the rise in uplink load factors (which ultimately express the total received power at a sector). Mean uplink load factors have risen from a sector average of 0.42 to an average of 0.55.

Based on the predictions presented in the previous section, (Figure 6-67 through Figure 6-93) other interesting observations can be made in comparison with those results presented for scenario B. While Figure 6-67 shows that pilot signal strength (dBm) has been kept, Figure 6-68 indicates that the effective pilot coverage (proportional to pilot's Ec/Io has been reduced due the increase in overall downlink interference. In the plots shown in Figure 6-70 to Figure 6-72, we see that the maximum achievable Eb/Io in the uplink has also been reduced even for the 8kbps class prediction (Figure 6-70), due to the increase in downlink interference. Furthermore, the higher-rate classes (64 and 144kbps) are much more limited in terms of maximum achievable Eb/Io rates, even though they are allowed to transmit at higher powers (0.84 and 1.5 W, respectively). For that reason, it is noticeable that some holes will be found in the downlink coverage for all classes, with holes varying in size for each class. Those holes can be observed if we consider that the required thresholds of Eb/Io are 9dB, 7.1 dB and 6.1 dB respectively for the classes of 8 kbps, 64 kbps and 144 kbps.

In Figure 6-73 to Figure 6-75, we observe that the required terminal power also increases for the higher rate classes. However, here we do not notice holes in the uplink coverage, as in all cases the required power is well within the maximum allowed limits for this simulation (27 dBm).

Another interesting observation comes from comparing the plots in Figure 6-91 and Figure 6-93, where we can notice that the SHO overhead is inversely proportional to the throughput carried by a sector. Take sector 15, for instance (upper left corner). While it is the sector with the least number of served connections as primary access server, it is left with enough power margin on the downlink to accept a larger fraction of secondary connections (soft handoff), and for that reason it is one of the sectors with most SHO overhead.

Finally, as an additional validation analysis for this simulation, we have performed the equivalent scenario in NPSW, and observed that results are extremely consistent between both simulations. A summary of results observed, compared between both tools is shown below.

As a conclusion for this scenario, we highlight again the importance of the characterization of the traffic offered in the determination of the total capacity of a system. As illustrated and discussed already in Chapter 4 and Chapter 5, the capacity of a system is directly tied to the nature of the traffic offered to it. Furthermore, for the same physical system (base station locations, sector configurations, etc), capacity is also strongly tied to the Radio Resource Management techniques and parameters configured at that system. This particular fact will be further explored in Chapter 7, where this current

scenario or demand characterization will be used as a baseline and modifications in the system configuration will be tested for capacity improvement.

Table 6-23. Summary of results comparison between simulation tools for Scenario D

		CelPlanner	NPSW
Served Users	8k	469	421
	64k	92	96
	144k	56	57
	Total	617	574
Throughput (kbps)		17704	17720
Not served users:			
Not carried traffic - Forward traffic channel power limit	8k	210	263
	64k	90	137
	144k	74	105
Not carried traffic - Forward sector total power limit	8k	50	36
	64k	58	7
	144k	50	8

6.10. Summary

We have executed in this chapter four simulation scenarios that served different purposes.

In scenario A we used the simplest possible traffic distribution (geographically homogeneous) and used that scenario to compare results with those expected from theory. The results observed were demonstrated to be consistent with those expected, allowing us to begin the process of validation for the simulation model chosen, which used the CelPlanner simulation tool.

In scenario B, we modified the traffic distribution in order to make it geographically heterogeneous, based on the methodology discussed in Chapter 3 of this work. The differences observed between this scenario and the previous one allowed us to demonstrate the impact of demand characterization with respect to geographical distribution of traffic.

In scenario C, the objective was to model a second simulation tool (NPSW) to simulate a scenario that would be as equivalent as possible to the configuration used in scenario B. That exercise proved to require a large effort of mapping equivalency between tools, both in terms of input parameters as well as algorithms and capabilities supported. Still, we have achieved satisfactory results in that comparison, which has also been extended for validation of results in scenario D. Based on the consistency found, we achieved a comfortable level of confidence in proceeding with one tool of our choice for further simulations in this dissertation.

In scenario D, the same user distribution of scenarios A, B and C has been split into three layers, for three different user profiles that included higher data rates. This posed a much larger demand on

the system, for the same system configuration in terms of base-stations configuration and maximum powers allowed per sector. It was observed that the system was able to carry much more throughput at this time, which demonstrates its sensitivity to demand characterization with respect to multiple classes.

In the next chapter, for a given demand characterization and network topology (baseline used in scenario D), we will explore modifications that could be made in the system that would allow it to increase its capacity. While the topic of capacity improvement for a given input is a whole discipline related to optimization, the intention of that chapter is much more modest, as we intend to explore the subject only as a demonstration of the importance of correctly modeling those inputs (here, the demand characterization) to achieve system improvement.

Chapter 7. System Optimization and Capacity Improvement

In this chapter, we apply the elements of demand characterization in optimizing a given network configuration. This analysis complements the framework proposed in this dissertation, demonstrating the close dependence between the processes of system dimensioning and optimization with the accurate modeling of traffic demand offered or forecasted for the system.

We start this discussion by reminding the reader that the term "optimization" has different meanings for different people even when looking at the specific subject of a wireless network. According to [68], the set of meanings is large and includes some of the points highlighted below.

1. System design improvements, such as cell locations, site configurations, etc., targeting network performance improvement and efficiency, where more or less freedom is allowed for modifications in existing configurations depending on the degree of maturity of a system (e.g.: Is the system completely built? Can new sites be added or physically modified?).
2. System implementation auditing, in order to ensure that the proposed design and parameters have been correctly configured and are working properly.
3. Fine-tuning of adjustable parameters in order to maximize system performance.
4. Appropriate diagnosis of problems observed in the network (typically based customer complaints) and proper identification of their causes.
5. Anticipation of system growth and potential problems expected in the system, and implementation of design and fine-tuning solutions that will cope with the expected system and demand evolution.

As one can observe, all tasks above involve some type of analysis, simulation and optimization algorithms that rely on a good description of the traffic offered to the system. However, some of those concepts of "optimization" involve also system diagnosis tools and equipment, as well as feedback from the customer base. The focus of our discussion in this chapter will refer mainly to those definitions of optimization that concentrate on the design and fine-tuning of the system. Therefore, the optimization of a system in this discussion refers mainly to two groups: RF design modifications and Radio Resource Management modifications. In each group, and depending on the degrees of freedom allowed for a given system, one or more of the parameters may be modified in the system configuration.

In the RF design category, typical optimization measures deal with modifications in the network topology (site locations) and structure, power setup and antenna design. Typical variables explored in this type of optimization are:

- RF network topology (number of sites, locations, number of sectors, etc);
- Site configuration selection and setup (model, height, tilt, azimuth);
- Common channels power adjustment (mainly the pilot powers).

The second category refers to adjusting Radio Resource Management parameters, of which the most typically adjustable variables are:

- Power control limits;
- Target parameters for admission and load control purposes;
- Soft-handoff related parameters (such as search window sizes, neighbor lists, reporting ranges, addition hysteresis, removal hysteresis and replacement hysteresis).

Throughout this chapter, whenever relevant, we use examples of optimization procedures based on the original network configuration presented in our case study (Chapter 6). All examples given will be compared to a baseline case, which we have chosen to be the one that was evaluated in Scenario D in Chapter 6 (section 6.9), where a 10-cell, 30-sector network is subject to a multiple class, geographically heterogeneous traffic distribution.

In the previous chapter, for validation and comparison purposes, all simulation results presented for the baseline project were created based on a single snapshot. However, as previously discussed in Chapter 5, statistically significant results of system performance simulations are achieved when a large number of snapshots is taken into account.

While each snapshot is a random set of active users that follows a given user density distribution, multiple snapshots will reflect the variability of those user positions with regards to their geographical density (Monte Carlo method). Therefore, a more statistically significant approach to evaluate performance of a system is to use results that were obtained from averaging the system performance over multiple snapshots. For this reason, all simulations performed in this chapter, were executed over a large number of snapshots. In deciding the amount to use, we initially performed several simulations over the baseline scenario, with variable number of snapshot (up to a few hundred) and verified that results with more than 20 snapshots presented small variability, relative to their absolute values. For that reason, for this particular system, 20 snapshots were considered a sufficiently small number to allow fast simulation executions while still providing some statistical significance. For a clearer illustration of the effects of averaging results over multiple snapshots, we show below the

simulation results achieved for the same scenario (Scenario D, presented in section 6.9), executed over 1 single snapshot and over 20 random snapshots (Table 7-1 and Table 7-2, respectively).

Table 7-1. Summary of Simulation Results for Baseline Project, for One Snapshot (same as in section 6.9)

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	15.6	5.7	469	3.1	1.6	92	1.9	1.1	56	20.6	5.7	617
SHO Connections	2.9	2.3	87	0.4	0.6	13	0.1	0.4	3	3.4	2.4	103
Total Connections	18.5	5.7	556	3.5	1.6	105	2.0	1.2	59	24.0	5.7	720
Throughput (kbps)	125.1	45.9	3752.0	196.3	100.7	5888.0	268.8	159.2	8064.0	590.1		**17704**
Total Rate (kbps)	148.3	45.9	4448	224.0	104.6	6720	283.2	171.2	8496	655.5	107.5	19664
SHO overhead	18.55%		18.55%	14.13%		14.13%	5.36%		5.36%	11.07%		11.07%
Not carried traffic - Pilot channel coverage										0		0
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	6.7	5.0	201	3.0	2.2	90	2.5	2.1	74	12.2	8.6	365
Not carried traffic - Forward sector total power limit	1.7	2.9	50	1.9	1.7	58	1.7	1.5	50	5.3	4.7	158
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.8	0.9	83.5	2.4	1.2	71.5	2.4	1.5	72.6	7.6	1.2	227.6
Uplink Load Factor										0.55	0.1	
Uplink Noise Rise										3.55	0.7	

Table 7-2. Summary of Simulation Results for Baseline Project, for 20 Snapshots

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	16.3	4.5	488.25	2.9	0.5	87.85	1.6	0.3	49.4	20.9	4.8	625.5
SHO Connections	2.9	1.2	88.3	0.3	0.2	8.55	0.2	0.1	4.8	3.4	1.4	101.65
Total Connections	19.2	4.2	576.55	3.2	0.4	96.4	1.8	0.3	54.2	24.2	4.4	727.15
Throughput (kbps)	130.2	35.9	3906.0	187.4	30.6	5622.4	237.1	46.7	7113.6	554.7		**16642**
Total Rate (kbps)	153.7	33.7	4612.4	205.7	28.5	6169.6	260.2	49.7	7804.8	619.6	69.3	18586.8
SHO overhead	18.08%		18.08%	9.73%		9.73%	9.72%		9.72%	11.69%		11.69%
Not carried traffic - Pilot channel coverage										0		0
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	6.5	4.4	193.75	3.5	2.2	103.85	2.6	1.5	77.8	12.5	7.9	375.4
Not carried traffic - Forward sector total power limit	1.3	1.5	38	1.6	1.2	48.3	1.8	1.1	52.8	4.6	3.7	139.1
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.9	0.7	86.4	2.2	0.4	66.4	2.2	0.4	67.1	7.3	1.0	219.9
Uplink Load Factor										0.54	0.1	
Uplink Noise Rise										3.41	0.6	

To further illustrate the effect of performing multiple snapshots, we show in Figure 7-1 the graphical results of both simulation cases. On both sides, the colored dots represent successful connections (different colors for different classes), while black dots are unsuccessful connections. On the left side, the dots represent the users in one snapshot, while the picture at the right represents the users from all snapshots (superimposed). As it can be observed from comparing the pictures, displaying the results of a simulation for multiple snapshots superimposed allows a quick visualization of the geographical location of the problem spot areas (regions where dots are mostly black). This type of display will be presented for all scenarios evaluated in this chapter.

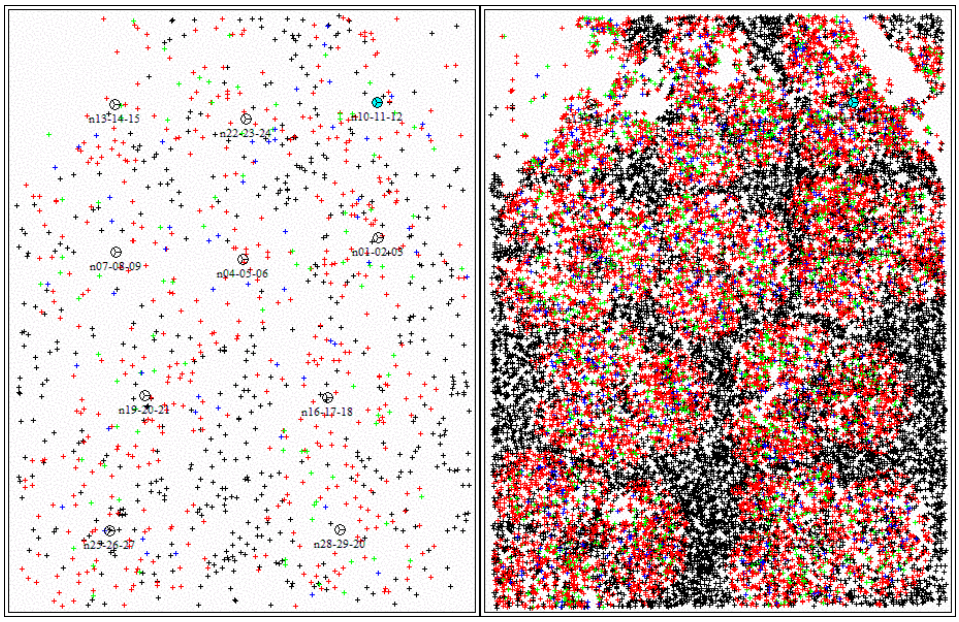


Figure 7-1. Graphical results of simulated users over 1 snapshot (left) and over 20 superimposed snapshots (right).

It is important to emphasize here that comprehensive network optimization is outside the scope of this dissertation. Our aim in this chapter is to present a framework that employs traffic demand characterization for optimization purpose, and to demonstrate the importance of such input in the process. In other words, the optimization of a system to a certain set of inputs is only as good as the input modeling accuracy. Therefore, appropriate system dimensioning and optimization has to rely on sensitive models of demand characterization, which are the main focus of this work.

In the following sections, we will discuss in more detail some of the variables that can be used for system optimization, and present solutions and techniques where demand characterization can be used as input in order to achieve network performance improvement.

7.1. RF Design Optimization

As mentioned before, the process of improving (or optimizing) a network performance may include modifications in the RF network topology, in the antenna parameters configuration and in the power setup for coverage adjustment. Each of those topics is discussed next.

7.1.1. RF Network Topology

In any RF system, the network topology (i.e., the geographical distribution of the RF network, with its transmission and reception points) is a key element in assuring proper network performance. When dealing with such a design, the main trade-offs refer to achieving the most solid (free of holes) coverage over the target area, with the least resources possible.

In CDMA based systems, this trade-off is made even more difficult, as the concept of coverage is directly affected by the amount of traffic (and therefore interference) that the system will handle when loaded. In other words, a CDMA system that has solid and continuous pilot coverage when unloaded may present large gaps in its coverage when loaded by heavy traffic. For that reason, typical RF design for CDMA systems pre-assume a certain operating load level for the cells, and evaluates coverage thresholds based on link-budgets that take that load into account.

The design and optimization problem becomes even more complex when, as it is typically the case, a pre-assumption of load level per site is not possible, and loads actually vary by location, due to heterogeneously distributed traffic density.

Therefore, if unloaded systems would provide a certain coverage radius based on link budget analysis, it follows that loaded systems make that radius smaller (some times much smaller), and heterogeneously distributed traffic makes those radiuses variable per location. This point is illustrated below (Figure 7-2):

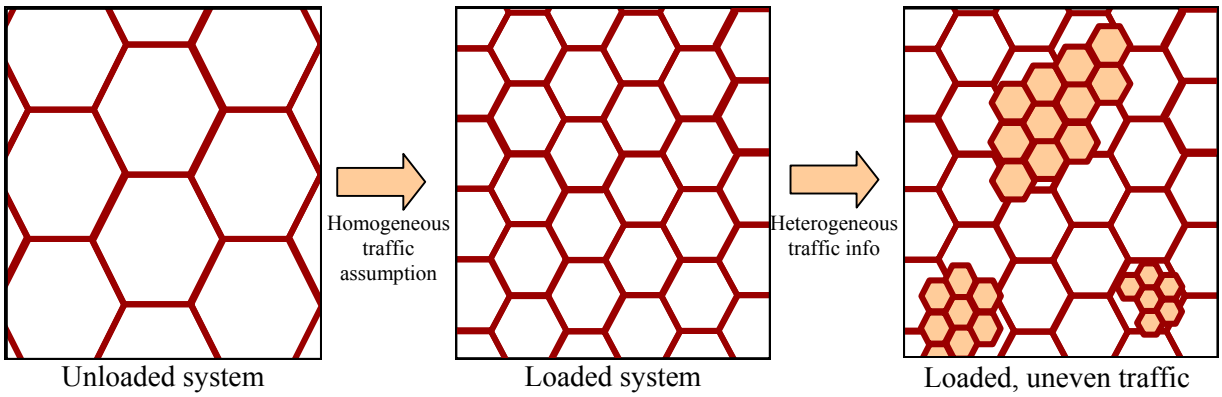


Figure 7-2. Effects of taking traffic data into account during RF network design and optimization.

For the reasons discussed above, the processes of RF network design and optimization should include as input a complete and accurate demand characterization, which will add complexity to those

processes, and in most cases require that the system be evaluated through system simulation, as discussed in this work.

7.1.2. Site Configuration and Antenna Parameters Optimization

When a system is already deployed and creating new sites is not a favored option, optimization becomes a bigger challenge, as it has to be achieved through "softer" methods. Among the measures available to influence network performance, physical changes in the cells' RF configuration are explored and desirable due to the strong impact they may have in some cases, at relatively small implementation costs. This is true for any wireless network; however, the main difference when it refers to CDMA based systems is that the decision making process on what needs to be changed is not as straightforward as in other technologies.

In traditional non-CDMA wireless systems, where interference issues are dealt with through frequency planning, optimizing RF parameters typically means targeting a few main objectives:

(i) **Maximizing the coverage** of each sector so that no holes exist in the footprint.

(ii) **Adjusting cell sizes** in an inverse proportion to traffic density, in order to keep total traffic per cell at an approximate constant level.

(iii) **Reducing the coverage overlap** between cells to a minimum that is just enough for hard-handoffs to be achieved with success, i.e., to a minimum that does not create holes in the system even in fading conditions.

Those objectives are typically achieved through cell shaping, i.e., adjustments of coverage shape by means of antenna pattern and down-tilt modifications, and size adjustment through variations in transmitted power.

In CDMA based systems, however, many other dynamics have to be taken into account when deciding how to shape cells. Among the most important differences we have:

(i) Cell coverage should still be maximized so that no holes exist in the system, but in CDMA systems, due to cell breathing, holes sizes and positions depend on the amount of traffic being offered to the system, which varies with time and geographical location.

(ii) Adjusting cell sizes depends not only on the traffic covered by a given cell, but also on the load of its neighbor cells and geometry.

(iii) **Cell overlap should not be minimized, but rather adjusted to an optimum point**, where the benefits of soft-handoff balance the disadvantages of additional

inter-cell interference and transmission overhead. This is a very subtle trade-off, as soft-handoff, for instance, may have simultaneously (on the same system configuration) advantages and disadvantages depending on the network loading.

In the uplink, soft-handoff will always cause the overall inter-cell interference to decrease, as mobiles that can be heard by a cell will interfere less if they are also being power controlled by that cell.

However, in the downlink, the gain provided by the Eb/Io reception improvement at the mobiles may come at the expense of additional power being transmitted by the serving sectors. In those cases, the total downlink power (transmitted by all sectors in the system) may fall below (handoff gain) or above (loss) what it would be without handoff.

In addition to the conflicting effects of soft-handoff on downlink and uplink power and interference balances, another overall disadvantage of soft-handoff is its larger consumption of system resources such as channel elements. These restrictions have also to be taken into account when resources are scarce.

For all those reasons, the fine-tuning of how much overlap area is desirable will depend significantly on the geographical traffic distribution, as well as on questions such as which is the limiting link (downlink or uplink) in the system and in each individual cell.

Optimization of CDMA systems is a complex problem, and simulation capabilities are typically required in order to explore all the aspects and trade-offs of different solutions.

Despite the abundance of literature on 3G WCDMA and 2G CDMA systems, there are few contributions that address the cell design and optimization problem from the Radio Network point of view. While capacity analysis has been a very frequent subject, it typically is dealt with based on simulations over uniform traffic distributions.

On the particular subject of design and optimization, there are few works available in the literature and most of those sources deal with single-class, second-generation (IS-95) CDMA systems only. For instance, in [51], the impact of non-homogenous traffic distribution is evaluated over an IS-95 system, but it starts from the assumption that most of those systems are uplink limited, and therefore it restricts its analysis and simulation capabilities only to the uplink dynamics. While it is typically (but not generally) an accepted concept for IS-95 networks that reverse link limits the capacity, in third-generation systems that pattern is not confirmed anymore, mainly due to technological an system improvements in the uplink implementations for 3G, and the expected asymmetry in terms of

traffic behavior. Still, some interesting results can be observed from that research [50], mainly with respect to the clear impact that traffic-sensitive adjustments may have on system capacity. For instance, it has been demonstrated in that work that for an **uplink-limited, single-class CDMA IS-95 network, decreasing the ratio of inter-cell to intra-cell interference for the highest traffic density cells has an effect in increasing network capacity.** This conclusion will be used in some exploratory examples presented next.

7.1.2.1. Tilt Modifications

We expect that by increasing mechanical downtilt in a cell we should reduce inter-cell interference and, therefore, reduce the o/s ratio for that cell. In order to determine which cells need more o/s reduction and by how much, we observe from inspection of the CDMA capacity formulas (Equation 7-1) that in order for capacity (N) to go up, we should either reduce o/s or increase the load factor (ℓ).

$$N = \frac{\ell_{up}}{\left(1+\frac{o}{s}\right)}\left[1+\frac{W}{\left(\frac{E_b}{N_0}\right)R\nu}\right]$$

Equation 7-1

Therefore, considering that the load factor is a parameter directly available from simulation results, we suggest a rule that increases downtilt (reduces o/s) for the cells with lowest load factors. That criteria was applied in the following example.

Example: Tilt modifications

In order to explore the impact of modifying some aspects of the cell configurations on network capacity, we created a scenario where the antenna's downtilt was modified according to a selected criterion. This criterion followed conclusions provided in [51], namely that the o/s ratio (inter-cell to intra-cell interference) should be reduced for those sites that present the highest load.

This scenario is compared against the baseline scenario, where all cells used mechanical tilt equal to 7 degrees.

The criterion previously suggested, of increasing tilt for sectors with lowest uplink load factors has been applied to all sectors, with tilt values allowed to vary within a limited range, from a minimum of

7 degrees downtilt to a maximum of 14 degrees. Only one-degree steps were allowed, therefore suggested values were rounded to the closest integer. The new tilt configuration used in that scenario is shown in Table 7-3.

Table 7-3. Parameters Used for Tilt Scenario (based on uplink load factor)

Cell Name	Sector ID	Up Load Factor	Antenna Tilt (degrees)	
			Baseline	Suggested
n01-02-03	1	0.58	7	8
n01-02-03	2	0.56	7	8
n01-02-03	3	0.64	7	7
n04-05-06	4	0.50	7	9
n04-05-06	5	0.45	7	10
n04-05-06	6	0.46	7	10
n07-08-09	7	0.51	7	9
n07-08-09	8	0.59	7	8
n07-08-09	9	0.53	7	9
n10-11-12	10	0.60	7	8
n10-11-12	11	0.59	7	8
n10-11-12	12	0.59	7	8
n13-14-15	13	0.48	7	10
n13-14-15	14	0.58	7	8
n13-14-15	15	0.16	7	14
n16-17-18	16	0.60	7	8
n16-17-18	17	0.53	7	9
n16-17-18	18	0.62	7	7
n19-20-21	19	0.60	7	8
n19-20-21	20	0.56	7	8
n19-20-21	21	0.58	7	8
n22-23-24	22	0.66	7	7
n22-23-24	23	0.53	7	9
n22-23-24	24	0.53	7	9
n25-26-27	25	0.61	7	8
n25-26-27	26	0.56	7	8
n25-26-27	27	0.57	7	8
n28-29-30	28	0.60	7	8
n28-29-30	29	0.58	7	8
n28-29-30	30	0.59	7	8

In order to illustrate the impact on the pilot channel coverage, Figure 7-3 shows the sector overlap available for both the baseline (left) and modified (right) designs, evaluated over an unloaded network.

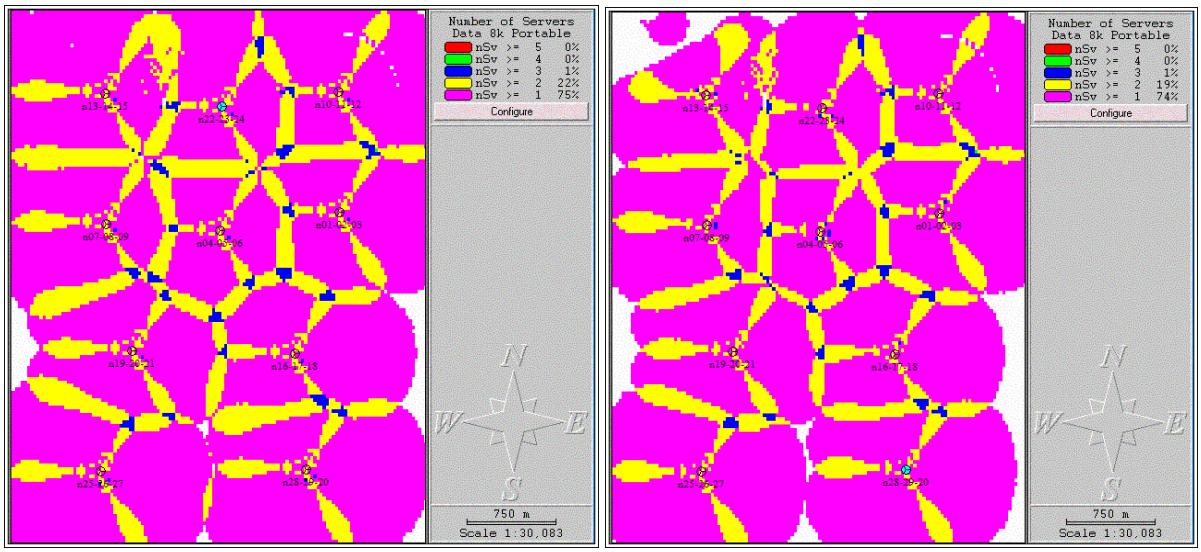


Figure 7-3. Pilot overlap comparison over unloaded system. Left: baseline, right: modified tilt.

Using the modified tilts, we performed the traffic simulation for the same traffic demand, over 20 snapshots. A summary of the simulation results is shown in Table 7-4. Graphical results for the same simulation are illustrated in Figure 7-4.

Table 7-4. Summary of Simulation Results for Tilt Scenario

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	15.8	4.6	473.85	3.0	0.6	88.75	1.8	0.3	52.55	20.5	5.1	615.15
SHO Connections	2.7	0.8	80.75	0.3	0.2	8.45	0.2	0.1	5	3.1	0.9	94.2
Total Connections	18.5	4.5	554.6	3.2	0.6	97.2	1.9	0.3	57.55	23.6	4.7	709.35
Throughput (kbps)	126.4	37.0	3790.8	189.3	38.7	5680.0	252.2	48.9	7567.2	567.9		**17038**
Total Rate (kbps)	147.9	35.7	4436.8	207.4	35.5	6220.8	276.2	47.4	8287.2	631.5	82.2	18944.8
SHO overhead	17.04%		17.04%	9.52%		9.52%	9.51%		9.51%	11.19%		11.19%
Not carried traffic - Pilot channel coverage										0		0
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	6.8	4.5	202.6	3.5	2.0	103.85	2.5	1.5	75.65	12.7	7.9	382.1
Not carried traffic - Forward sector total power limit	1.4	1.4	43.45	1.6	1.2	47.35	1.7	1.2	51.8	4.8	3.7	142.6
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.7	0.7	82.3	2.2	0.4	66.6	2.3	0.4	70.0	7.3	1.0	218.9
Uplink Load Factor										0.54	0.1	
Uplink Noise Rise										3.45	0.6	

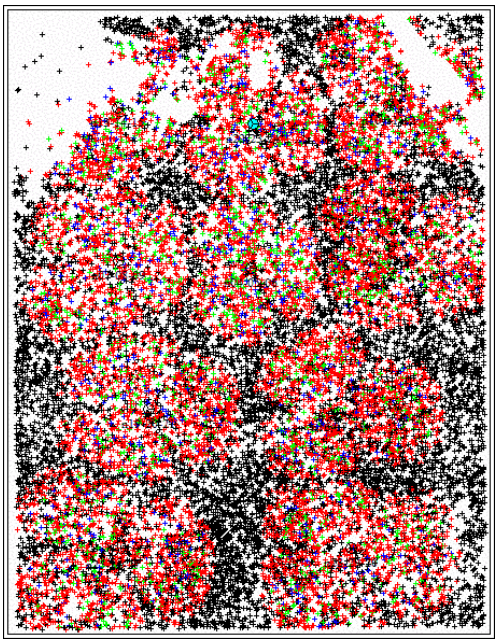


Figure 7-4. Graphical simulation results for tilt scenario.

By comparing the results obtained above with those in the baseline project, we observe that the total system capacity has grown from 16.6 to 17.0 Mbps, a mere 2.4% improvement. Soft-handoff overhead has been kept at similar levels (11.2%, compared to 11.7% in the baseline project).

One of the reasons that may explain the small impact of changing tilts in this case is the fact that the most restricting limitations here were and still are related to the maximum downlink power allowed per traffic channel and the total maximum allowed per sector. For those reasons, users further away from the cells were already being put in outage on the baseline, and this condition has not changed with the tilt adjustment.

In addition to the configuration tested in this example, we have tested other configurations where tilts were also modified proportionally to the uplink-estimated o/s, based on post-processed results from the baseline simulations. The range of variation was still kept between 7 and 14 degrees. In that solution, higher o/s uplink ratios got higher tilts (closer to 14), while lower o/s ratios got tilts closer to 7. Results have shown to be very similar to those presented above, i.e., no significant improvement was achieved by varying tilt alone.

7.1.2.2. Pilot Power Modifications

A powerful resource in adjusting the size of a cell relative to its neighbors in CDMA systems is to vary the pilot power.

Here, we suggest that pilot powers be modified according to a technique typically used in optimizing 2G CDMA systems, which is also evaluated in [51], in which sectors in areas of higher traffic density are configured with higher pilot powers, in order to decrease their inter-cell/intra-cell interference ratio, and therefore to be able to have more control over the interference they suffer (since intra-cell interference is power controlled by the sector itself). Although this technique has been proven to work for single-class, uplink limited second-generation systems, it has not been evaluated or extensively explored in multi-class third-generation systems that may be either uplink or downlink limited.

In order to explore the setting of pilot power values, we propose that the original (non-optimized system) be evaluated with respect to its original coverage footprint, in terms of its pilot signal strength. This type of analysis allows us to identify how much traffic would be covered by each sector in the hypothetical situation of homogeneous interference at all points. In other words, instead of considering the traffic actually captured in simulation in the baseline project, one can look at the original traffic density and traffic amounts offered to each sector against the sector's received pilot strength. Figure 7-5 illustrates this type of analysis for our baseline project, showing the best-server area of each sector based on received pilot power.

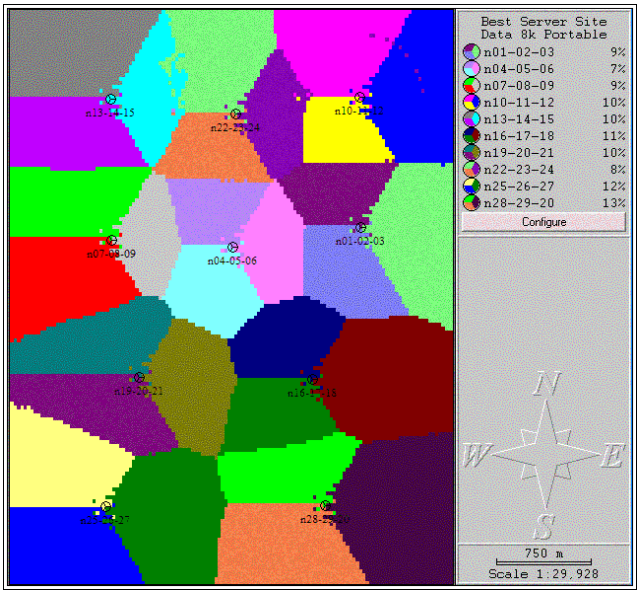


Figure 7-5. Pilot signal strength (RSSI) best server prediction.

Using the area definitions shown in Figure 7-5, it is possible then to cross that information with the demand offered to the system, in terms of users covered per sector, as well as estimate the geographical area of each sector, and identify the traffic density in terms of users per square kilometer.

Next, pilot powers should be adjusted based directly on the user density per sector, i.e., sectors with higher traffic density get higher pilot powers and vice-versa. The example shown below illustrates this procedure for our baseline project.

Example: Pilot Power Modifications

In the original geographic area considered, the traffic density varies from a minimum of 14.94 users/km^2 (sector 15) to a maximum of 111.51 users/km^2 (sector 3). That strong heterogeneity (a factor of about 7:1) has proven to reduce system capacity when cells are equally designed, i.e., have all the same RF configuration and power setup ([53] and [54]). The user density per sector in the system is presented in Table 7-5.

Table 7-5. Users Density per Pilot Best-Server Area

Cell Name	Sector ID	Traffic (users)	Area (sq.km)	Density (users/sq.km)
n01-02-03	1	50.19	0.63	79.55
n01-02-03	2	38.05	0.43	88.75
n01-02-03	3	39.61	0.36	111.51
n04-05-06	4	28.87	0.37	78.56
n04-05-06	5	32.15	0.42	77.19
n04-05-06	6	23.94	0.33	72.06
n07-08-09	7	22.78	0.34	66.72
n07-08-09	8	38.60	0.59	65.31
n07-08-09	9	33.61	0.52	64.95
n10-11-12	10	33.94	0.59	57.27
n10-11-12	11	28.05	0.30	92.53
n10-11-12	12	35.21	0.48	73.00
n13-14-15	13	20.61	0.38	54.71
n13-14-15	14	20.21	0.52	39.04
n13-14-15	15	9.45	0.63	14.94
n16-17-18	16	80.38	0.89	90.04
n16-17-18	17	41.64	0.43	97.82
n16-17-18	18	41.01	0.45	91.10
n19-20-21	19	47.89	0.57	84.54
n19-20-21	20	36.57	0.51	71.08
n19-20-21	21	38.28	0.54	70.63
n22-23-24	22	30.82	0.38	82.15
n22-23-24	23	25.77	0.35	73.18
n22-23-24	24	34.76	0.53	66.19
n25-26-27	25	74.80	0.76	98.29
n25-26-27	26	39.06	0.52	74.59
n25-26-27	27	40.96	0.55	74.51
n28-29-30	28	60.93	0.90	67.33
n28-29-30	29	47.63	0.57	83.85
n28-29-30	30	44.21	0.47	93.44

Pilot powers adjustments were based directly on the density column of Table 7-5. Another decision explored in this analysis was with regards to which range to spread those powers. While the original system had all pilots fixed at 1W, it was observed that pilot coverage exceeded the footprint of traffic channel coverage on the forward link, i.e., there was excess power being dedicated to the pilot channel that could have been used more efficiently in the traffic channel pool. For that reason, we experimented with a varying range of powers, and found the best range to be the configuration of all pilot powers around a mean value of 0.25W.

As a consequence, pilot powers used in this analysis were in the range of 0.06W to 0.42W. While these values have provided the best results in the simulations, it is important to notice that in real systems sector 15 should probably be treated as an exception, i.e., despite optimization suggestions, a floor in the pilot power should be used in order not to practically "kill" the coverage of sectors with low traffic density. Table 7-6 illustrates the values used for pilot power in each sector, compared to the baseline project.

Table 7-6. Parameters used for Pilot Power Scenario (based on geographical user density)

Cell Name	Sector ID	Pilot Power (W)	
		Baseline	Suggested
n01-02-03	1	1	0.30
n01-02-03	2	1	0.33
n01-02-03	3	1	0.42
n04-05-06	4	1	0.29
n04-05-06	5	1	0.29
n04-05-06	6	1	0.27
n07-08-09	7	1	0.25
n07-08-09	8	1	0.24
n07-08-09	9	1	0.24
n10-11-12	10	1	0.21
n10-11-12	11	1	0.34
n10-11-12	12	1	0.27
n13-14-15	13	1	0.20
n13-14-15	14	1	0.15
n13-14-15	15	1	0.06
n16-17-18	16	1	0.34
n16-17-18	17	1	0.36
n16-17-18	18	1	0.34
n19-20-21	19	1	0.31
n19-20-21	20	1	0.26
n19-20-21	21	1	0.26
n22-23-24	22	1	0.31
n22-23-24	23	1	0.27
n22-23-24	24	1	0.25
n25-26-27	25	1	0.37
n25-26-27	26	1	0.28
n25-26-27	27	1	0.28
n28-29-30	28	1	0.25
n28-29-30	29	1	0.31
n28-29-30	30	1	0.35

Using the modified pilot powers, we performed the traffic simulation for the estimated traffic demand, over 20 snapshots. A summary of the simulation results is shown in Table 7-7. Graphical results for the same simulation are illustrated in Figure 7-6.

Table 7-7. Summary of Simulation Results for Pilot Power Scenario

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	16.3	4.7	489.1	3.8	0.7	114.05	2.4	0.5	72.85	22.5	5.3	676
SHO Connections	0.1	0.2	4.3	0.0	0.0	0.3	0.0	0.0	0.25	0.2	0.2	4.85
Total Connections	16.4	4.7	493.4	3.8	0.7	114.35	2.4	0.5	73.1	22.7	5.2	680.85
Throughput (kbps)	130.4	37.8	3912.8	243.3	43.5	7299.2	349.7	66.5	10490.4	723.4		**21702**
Total Rate (kbps)	131.6	37.3	3947.2	243.9	43.6	7318.4	350.9	66.5	10526.4	726.4	112.3	21792
SHO overhead	0.88%		0.88%	0.26%		0.26%	0.34%		0.34%	0.41%		0.41%
Not carried traffic - Pilot channel coverage										4		132
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	4.2	3.5	124.75	2.3	1.6	69.7	1.7	1.1	52.4	8.2	6.1	246.85
Not carried traffic - Forward sector total power limit	0.7	1.2	22.3	1.0	1.2	29.6	1.1	1.3	33.45	2.8	3.6	85.35
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.1	0.8	64.1	2.4	0.6	70.8	2.7	0.6	80.7	7.2	1.6	215.5
Uplink Load Factor										0.66	0.1	
Uplink Noise Rise										4.83	0.8	

By comparing the results obtained above with those from the baseline case, we observe that the total system capacity has grown from 16.6 to 21.7 Mbps, which means a significant improvement in capacity of about 30%. This improvement is explained by two main effects that have been caused in the system:

(i) The fact that pilot powers were reduced in average over the whole system made more power to be available per sector to be dedicated to traffic channels, as a limit of 10W had been imposed for the total power transmitted by any sector (to simulate a hardware limitation). This effect is clearly noticed in the significant reduction of rejected calls due to "Forward sector total power limit" on the table above, which was reduced from 158 rejected connections to 85.

(ii) The second effect has to do with the uneven pilot power distribution among sectors, which allowed a better equalization of the traffic carried per sector. While in the baseline case each sector carried in average 20.5 users with a 5.3 standard deviation, in the optimized pilot project the mean number of users per sector was raised to 22.5, with standard deviation reduced to 5.2 users.

Another interesting observation from this scenario is that soft-handoff overhead has been drastically reduced to 0.4%, compared to 11.7% in the baseline project. While this is a risky margin of handoff overhead (because too little overlap is allowed for user mobility), those results are consistency with the conditions found in this system, since soft-handoff has been proven to

compromise even more the system capacity and power allocations on the downlink, which is where the original system was already limited.

For that reason, if we compare Figure 7-6 with the picture at the right on Figure 7-1, we observe that the black shadows in the system are more concentrated now in the regions between sites, and between sectors, giving more definition to the areas of service and without service.

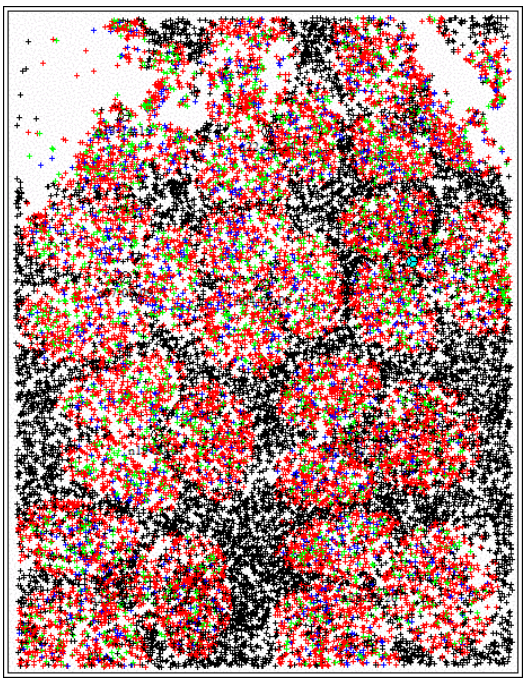


Figure 7-6. Graphical simulation results for pilot power scenario.

Finally, using post-processed results from simulations, we created the prediction of expected chip energy over interference spectral density (Ec/Io) for the pilot coverage in both cases (original baseline project and modified tilt version). Those results are illustrated in Figure 7-7. They demonstrate that pilot channel coverage was over dimensioned in the original scenario, and has been reduced significantly in order to provide the best balance possible between pilot and traffic-channel coverage, allowing more power to be dedicated to the traffic channels.

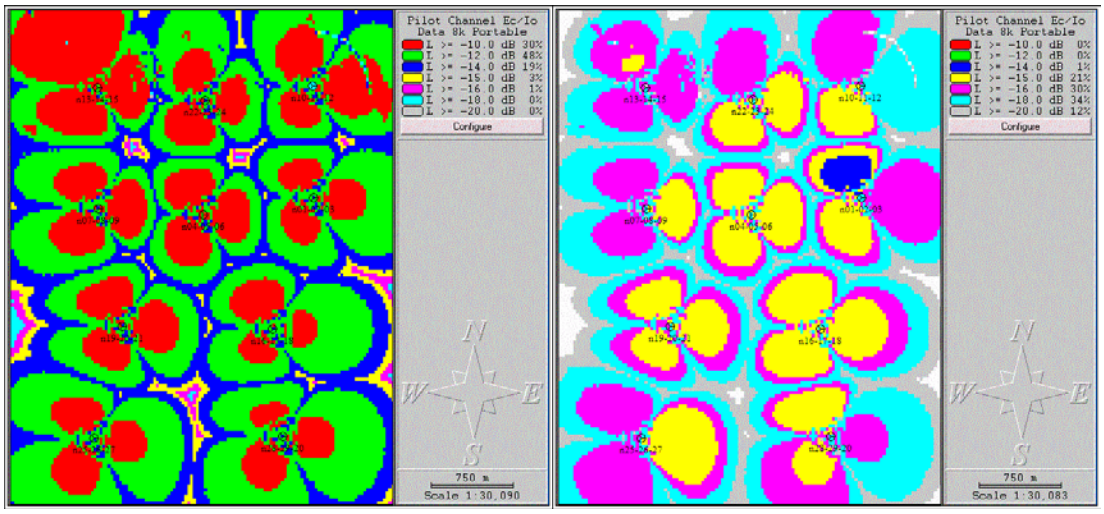


Figure 7-7. Pilot Ec/Io (db) prediction baseline project (left) and modified pilot scenario (right).

7.2. Resource Management Parameters Scenario:

7.2.1. Power Control Management

Parameters and variations of the power control algorithms in place may also be exploited in optimizing capacity for wireless networks.

While it is not within the scope of this discussion to look for better power control algorithms, we do wish to explore the use of the proposed framework and the information on demand characterization in order to find the best parameters to use as setup in typical power control algorithms established for 3G networks.

7.2.1.1. Target Power Proportion per Service

For different QoS service classes allowed in the 3G systems, particularly in the WCDMA standards, power control algorithms are allowed to use different values of maximum power per class. Those values can be adjusted in order to provide a balance in coverage extension among all classes, therefore avoiding the effect of smaller footprints being available for those classes demanding more capacity. In other words, by reducing the maximum power allowed for low data-rate classes, we reduce the maximum distance we cover those users, and in return we are allowed to cover higher-rate users at higher distances.

As one can imagine, a good balance here would mean to adjust maximum powers per class in a way that all footprints fall in the same range, not favoring any specific class.

In order to achieve that, we based our analysis in the formula for the sensitivity of a link, which determines the minimum signal strength required in the link for a given service class, based on that class data rate R_{class} and on its Eb/Io requirements ($Eb/Io_{Reqclass}$). This equation is given in Equation 7-2, where N+I refer to the total noise plus interference received at a certain location (user location for downlink, cell location for uplink), and W refers to the channel bandwidth.

$$S_{Class} = \left(\frac{E_b}{I_o}\right)_{\mathrm{Re}qClass} * R_{Class} * \frac{(N+I)}{W} \qquad \textbf{Equation 7-2}$$

Additionally, the transmitted power required at the other end of the link Ptx_{class} is directly proportional to the sensitivity of the user and the path loss L_{path} between both ends, as shown in Equation 7-3.

$$Ptx_{Class} = S_{Class} * L_{Path} \qquad \textbf{Equation 7-3}$$

If we want all classes to achieve the same cell radius, it means that all of them should be limited to the same maximum path loss. For that reason, classes with lower sensitivities should be configured to have lower limits for their maximum transmitted powers, as shown in Equation 7-3, where $Ptx_{MaxClass}$ is the maximum power allowed for a class and $(L_{path})_{MaxClass}$ is maximum path loss allowed for that same class.

$$(L_{Path})_{MaxClass} = \frac{Ptx_{MaxClass}}{\left(\frac{E_b}{I_o}\right)_{\mathrm{Re}qClass} * R_{Class} * \frac{(N+I)}{W}} \qquad \textbf{Equation 7-4}$$

Therefore, we suggest that one configures the best power proportion between classes by equalizing the maximum path loss for all classes $(L_{path})_{MaxClass}$, which considering that $(N+I)/W$ is common for all classes, means equalizing the term given in Equation 7-5

$$Equalize: \quad \frac{Ptx_{MaxClass}}{\left(\frac{E_b}{I_o}\right)_{\mathrm{Re}qClass} * R_{Class}}$$ **Equation 7-5**

Example: Adjusting target power proportion per service - Uplink

In this example, we use the formulation presented above in order to estimate the ideal limits in transmitted power for the downlink. First we obtain the product of $(Eb/Io_{Reqclass})*R_{class}$ for each class. Next, we suggest power limits that are directly proportional those products. Table 7-8 below shows the results obtained for the uplink parameters, normalized to class 1 (sixth row). We limit the maximum transmit powers of all classes to 0.5 Watts, which is the value that was used for all classes in the original simulation (last row).

Table 7-8. Estimating Proportionality of Power Control Limits for Different Classes (uplink)

Uplink	Symbol	Class 1	Class 2	Class 3
Rates	R	8	64	144
Required Eb/Io (dB)	E	6.5	4	3.5
Required Eb/Io (linear)		4.47	2.51	2.24
E*R		35.7	160.8	322.4
E*R normalized		1.00	4.50	9.02
Suggested Ptx (norm. to maximum = 0.5)		**0.06**	**0.25**	**0.50**

A new simulation has been performed using the above suggested power proportions per class. However, since the system evaluated was not limited at any moment on the uplink, changes in the proportion had no impact in capacity.

Example: Adjusting target power proportion per service – Downlink

In this example, the same procedure was applied to the downlink direction. Table 7-9 shows the results obtained for the downlink power limits, normalized to class 1 (sixth row). We limit the maximum of all classes to 1.5 Watts, which is the maximum value that was used for any class in the original simulation (last row).

Table 7-9. Estimating Proportionality of Power Control Limits for Different Classes (downlink)

Downlink	Symbol	Class 1	Class 2	Class 3
Rates	R	8	64	144
Required Eb/Io (dB)	E	9	7.1	6.1
Required Eb/Io (linear)		7.943	5.129	4.074
E/R		63.5	328.2	586.6
E/R normalized		1.00	5.17	9.23
Suggested Ptx (norm. to maximum =1.5)		**0.16**	**0.84**	**1.50**

A new simulation has been performed using the above suggested power proportions per class. The simulation was performed over 20 snapshots. A summary of the simulation results is shown in Table 7-10. Graphical results for the same simulation are illustrated in Figure 7-8.

Table 7-10. Summary of Simulation Results for Downlink Power Proportion per Class Scenario

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	12.9	3.7	387.2	3.3	0.6	99.95	2.1	0.3	61.75	18.3	4.2	548.9
SHO Connections	1.2	0.7	36.45	0.3	0.2	10.2	0.2	0.1	6.05	1.8	0.9	52.7
Total Connections	14.1	3.4	423.65	3.7	0.6	110.15	2.3	0.3	67.8	20.1	3.8	601.6
Throughput (kbps)	103.3	29.3	3097.6	213.2	38.4	6396.8	296.4	44.8	8892.0	612.9		18386
Total Rate (kbps)	113.0	27.3	3389.2	235.0	36.6	7049.6	325.4	46.9	9763.2	673.4	75.0	20202
SHO overhead	9.41%		9.41%	10.21%		10.21%	9.80%		9.80%	9.87%		9.87%
Not carried traffic - Pilot channel coverage										0		0
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	10.4	6.1	311.5	3.4	2.0	102.1	2.5	1.5	73.9	16.3	9.5	487.5
Not carried traffic - Forward sector total power limit	0.7	1.0	21.3	1.3	1.0	37.95	1.5	1.0	44.35	3.5	2.8	103.6
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	1.9	0.5	56.7	2.5	0.4	75.9	2.8	0.4	82.9	7.2	1.0	215.5
Uplink Load Factor										0.55	0.1	
Uplink Noise Rise										3.53	0.6	

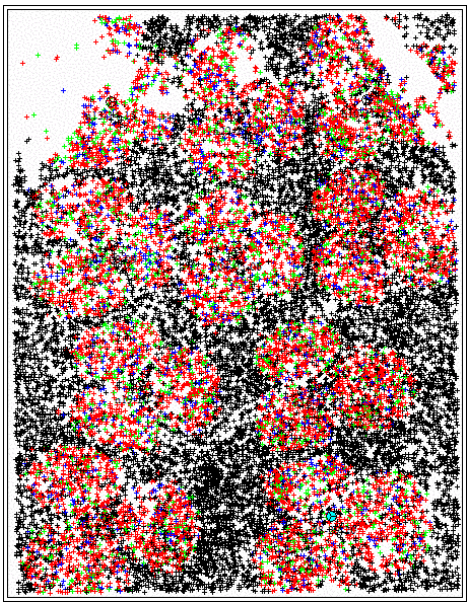


Figure 7-8. Graphical simulation results for downlink power proportion per class scenario.

In this case, some improvement has been observed when compared to the original configuration, which has actually used very similar parameters, i.e., power limits of 0.2, 0.84 and 1.5 Watts for classes 1, 2 and 3 respectively.

In other words, power limits per class on the downlink were already close to the ideal proportion, where only class 1 had been a little overestimated. Nevertheless, this small change has demonstrated to have an impact in total system capacity, as the 8kbps users were now limited by smaller maximum downlink powers, and therefore, were not allowed to be served as far as they were before. This limitation in their coverage area allows service to go only to the point where users of the other classes are also being served, and therefore creating more capacity for the other classes.

The total system capacity has been improved from 16.6 to 18.4 Mbps, a significant 10% improvement, while the soft-handoff overhead has been only slightly reduced from 11.7% to 9.9%. These results demonstrate the impact and importance of such adjustments on the limiting links of the system.

The radius equalization provided by this proportion adjustment has also helped equalize the proportion of served users per class, as illustrated in Table 7-11

Table 7-11. Proportion of Users Ooffered and Served per Class in the Original and Adjusted Scenarios

	Offered Demand		Served Original		Served Adjusted	
User Class	**Users**	**%**	**Users**	**%**	**Users**	**%**
Class 1	720	63%	488.2	78%	387.2	71%
Class 2	240	21%	87.8	14%	99.5	18%
Class 3	180	16%	49.4	8%	61.75	11%
Total	**1140**		**625.4**		**548.45**	

7.2.1.2. Power Factor per Sector

For systems where there is some flexibility in increasing the maximum power allowed per sector, an additional resource is available to improve system capacity, which is similar in principle to adjusting the pilot powers. It consists of applying a multiplying factor to either the transmitted (downlink) or received (uplink) traffic powers of a sector. These factors, as in the pilot power adjustment, are applied as a function of traffic density in the system.

In the downlink, the power factor (also referred by some equipment vendors as scaling factor) is a multiplying factor applied uniformly to all channels on the downlink of a sector. Those factors are sector parameters and therefore may vary on a per-sector basis. In the uplink the power factor is also a

multiplying factor applied uniformly over all received traffic powers of a sector. In the same way, this parameter is adjustable on a per-sector basis.

In [51], the authors demonstrate the improvements that result from applying uplink power factors that are proportional to traffic density on an uplink limited single-class IS-95 CDMA system.

In this work, we extend the technique and suggest that it may be applied in either direction, or in both, as a means to improve capacity over the most limiting link. We also provide, in this framework, the contribution of a solution that is applicable for multi-class systems such as 3G networks, using the methodology of demand characterization provided in this work.

In order to evaluate the impact of such technique, we performed an example scenario applied to the downlink powers. Since this system is downlink limited, and the powers in the uplink are already controlled, we did not expect any increase in capacity by changing the maximum powers allowed per mobile station.

Example: Modifying downlink power factor per sector

In this scenario, similarly to the example presented for pilot power adjustments, we used the information of geographical user density (provided in Table 7-5). One difference is that while in the pilot power adjustments we scaled the powers to satisfy a desired mean value (0.25), here our constraint is to keep traffic power allocations under a controlled limit. For that reason, the traffic density proportion was scaled to fit all power factors within the range of 1 to 2 (0 to 3 dB).

The resulting downlink power factors suggested for each sector are shown in Table 7-12. Notice that in order not to "kill" coverage in the sector of smallest traffic density (sector 15), we have treated that sector as an exception, and applied the power factor to the other 29 sectors of the system only.

A simulation has been performed using the suggested power factors per sector. The simulation was performed over 20 snapshots. A summary of the simulation results is shown in Table 7-13. Graphical results for the same simulation are illustrated in Figure 7-9.

Table 7-12. Parameters Used for Downlink Power Factor Scenario (based on geographical user density)

Cell Name	Sector ID	Downlink Traffic Power Factor			
		Before		After	
		linear	dB	linear	dB
n01-02-03	1	1	0	1.56	2
n01-02-03	2	1	0	1.69	2
n01-02-03	3	1	0	2	3
n04-05-06	4	1	0	1.55	2
n04-05-06	5	1	0	1.53	2
n04-05-06	6	1	0	1.46	2
n07-08-09	7	1	0	1.38	1
n07-08-09	8	1	0	1.36	1
n07-08-09	9	1	0	1.36	1
n10-11-12	10	1	0	1.25	1
n10-11-12	11	1	0	1.74	2
n10-11-12	12	1	0	1.47	2
n13-14-15	13	1	0	1.22	1
n13-14-15	14	1	0	1	0
n13-14-15	15	1	0	1	0
n16-17-18	16	1	0	1.70	2
n16-17-18	17	1	0	1.81	3
n16-17-18	18	1	0	1.72	2
n19-20-21	19	1	0	1.63	2
n19-20-21	20	1	0	1.44	2
n19-20-21	21	1	0	1.44	2
n22-23-24	22	1	0	1.59	2
n22-23-24	23	1	0	1.47	2
n22-23-24	24	1	0	1.37	1
n25-26-27	25	1	0	1.82	3
n25-26-27	26	1	0	1.49	2
n25-26-27	27	1	0	1.49	2
n28-29-30	28	1	0	1.39	1
n28-29-30	29	1	0	1.62	2
n28-29-30	30	1	0	1.75	2

Table 7-13. Summary of Simulation Results for Downlink Power Factor Scenario

	Class 8 kbps			Class 64 kbps			Class 144 kbps			All Classes Combined		
	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum	Avg.	StDev	Sum
Requested Connections	24		720	8		240	6		180	38		1140
Served Connections	17.1	4.6	511.5	3.4	0.7	102.7	2.2	0.4	64.55	22.6	5.4	678.75
SHO Connections	2.0	1.0	58.55	0.2	0.1	5.4	0.1	0.1	4.1	2.3	1.2	68.05
Total Connections	19.0	4.1	570.05	3.6	0.6	108.1	2.3	0.4	68.65	24.9	4.7	746.8
Throughput (kbps)	136.4	37.1	4092.0	219.1	44.4	6572.8	309.8	57.2	9295.2	665.3		**19960**
Total Rate (kbps)	152.0	32.7	4560.4	230.6	41.5	6918.4	329.5	52.6	9885.6	712.1	101.7	21364.4
SHO overhead	11.45%		11.45%	5.26%		5.26%	6.35%		6.35%	7.04%		7.04%
Not carried traffic - Pilot channel coverage										0		1
Not carried traffic - Directed to other carriers	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Throughput per carrier limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Forward traffic channel power limit	5.9	3.7	176	3.2	1.7	95.1	2.4	1.5	71.85	11.4	6.8	342.95
Not carried traffic - Forward sector total power limit	1.1	1.7	32.15	1.4	1.3	41.85	1.4	1.2	43.45	3.9	4.2	117.45
Not carried traffic - Forward load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse mobile power limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Not carried traffic - Reverse load factor limit	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0	0
Total Dw TCHs Power (W)	2.7	0.6	80.4	2.4	0.4	70.8	2.7	0.4	81.0	7.7	1.1	232.2
Uplink Load Factor										0.64	0.1	
Uplink Noise Rise										4.51	1.0	

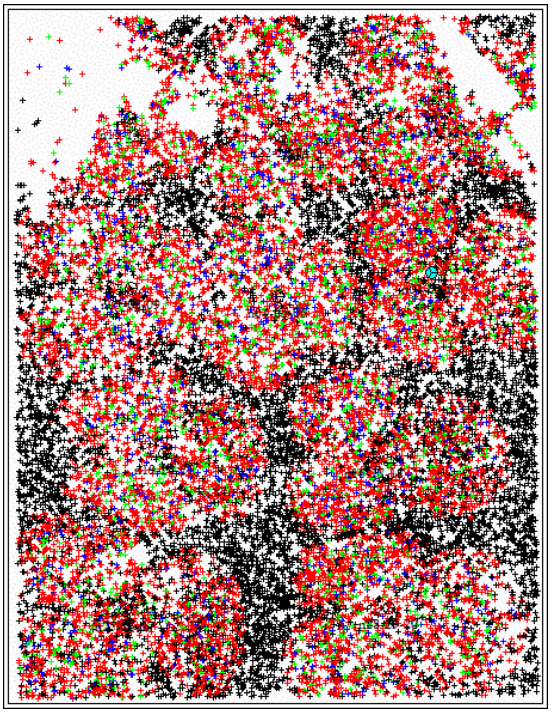


Figure 7-9. Graphical simulation results for downlink power factor scenario.

In this example, reasonable improvement has been observed when compared to the original configuration. The total system capacity improved from 16.6 to 19.9 Mbps, a significant 15% improvement. Soft-handoff overhead has been reduced from 11.7% to 7%.

Most importantly, we observed that the system has extended now its overall service area, getting closer (but not yet there) to a balance between traffic channel and pilot channel coverage limits. This effect can be clearly observed in Figure 7-9, where the service areas are closer to a continuous footprint now, compared to the original design. That observation is also confirmed by the fact that in this adjusted system we reached the point of having a user rejected by pilot limits, instead of traffic channel limitation that had been the pattern for all simulations in the original system so far.

It is important to notice that the system did not reach ideal coverage yet, mainly because we did not allow the powers to follow completely the existing proportions of traffic density in this experiment. In additional simulations performed, we observed that when the downlink power factors are allowed to grow by a factor of up to 8 (9 dB), there is an inversion in the nature of this system's limitations, where most of connections refused are due to limits in pilot power and some even in uplink power.

7.2.2. Congestion Control

Congestion control is the method used to limit the amount of traffic accepted in a system in response to overload conditions in the network. Particularly in 3G systems, where the quality of service of active users must be guaranteed within certain thresholds, the system cannot afford to accept more connections in detriment of the existing ones, or has to take some actions if existing load grows beyond certain limits. For those reasons, sector loads are kept under restrict control, through the use of admission control and load control techniques. Algorithms used for that purpose are not completely standard, and different congestion control algorithms may be adopted by different vendors and implementations.

In admission control techniques, the system typically tries to estimate the amount of load increase that would be caused to it in case a requested service (RAB) is accepted. It then accepts or rejects the RABs depending on specified acceptance thresholds.

In load control, the system keeps monitoring (or estimating) the current load in each sector, using methods already discussed in section 4.3.2 of this work. Typical methods of estimating load in WCDMA systems are through the measurement of total receive power in the uplink and total transmitted power on the downlink.

When a problem is flagged, i.e., when load has reached certain specified limits (plus margins allowed for stability), the system takes some action in order to prevent or solve overload problems. Those actions may include the following.

- Reduction of traffic volume in the non-real-time traffic classes, through the use of packet scheduling algorithms. These reductions will take into account the classes and priorities negotiated for those bearers, and respecting the minimum guaranteed rates.
- Reduction of data rates of real-time traffic classes such as voice or data, within the limits of acceptable quality of service.
- Temporary blocking or reduction in power control commands, which ultimately also reduces quality for existing RABs.
- Forced handoffs to sectors with lower loads.

While it is not part of the network planning and optimization to decide on which algorithms to use, typical parameters to be optimized with regards to congestion control are the target values for maximum load, which may be different for downlink and uplink, and may also vary for admission control and load control functions. Besides, target margins must be configured, in order to give the system some room to exceed by small amounts the load values that it is trying to keep on average.

In optimizing the parameters for load control targets, the effects of non-uniform traffic density have to be taken into account once again, as the choice of target loads, which may be configurable per sector or for the whole system, will depend on the traffic distribution and ratios of inter-cell to intra-cell interference that vary according to location and sector geometry.

While load control algorithms are vastly explored in the literature, there is little information available on the subject of techniques to optimize those targets. In this work, we ran some experimental simulations, in which only global values for target load in the system have been tested, and compared to the baseline case.

Since no load control was used in the baseline project, any limit imposed here (lower than 1) should yield a reduction in capacity, instead of an increase. This expectation has been confirmed using simulations and configuring targets for values slightly below 1.

However, another objective has been aimed in this example, and it refers to the fact observed in the previous examples that power restrictions were causing the system to refuse connections at the edges of the cells, causing discontinuity between sectors that would ultimately disturb mobility. If we intended to make this system work in reality, even with the original tight restriction of maximum 10W per sector, one option would be to sacrifice capacity in return for continuous coverage. This can be achieved through the use of load control, which will ultimately prevent the sector from reaching high levels of interference, and therefore, extend the maximum allowed path loss and by consequence cell radiuses.

The results obtained with those experiments are discussed next.

Example: Adjusting global target load factors

In this example, we performed multiple simulations where the only parameter modified in each case was the target load factor on the uplink. While we know the downlink is the limiting direction, we assume that "load control" is already imposed in that direction through the established limit of 10W per sector (as in the original case studied). Therefore, we explore here the effects that limiting service acceptance per sector has on providing a more continuous coverage to this system. For this reason, results for the multiple scenarios simulated are presented in terms of graphical displays of service area, as shown in Figure 7-10.

Target Load on Uplink: 0.5	Target Load on Uplink: 0.4	Target Load on Uplink: 0.3
Capacity achieved: 14 Mbps	Capacity achieved: 10.5 Mbps	Capacity achieved: 7.2 Mbps

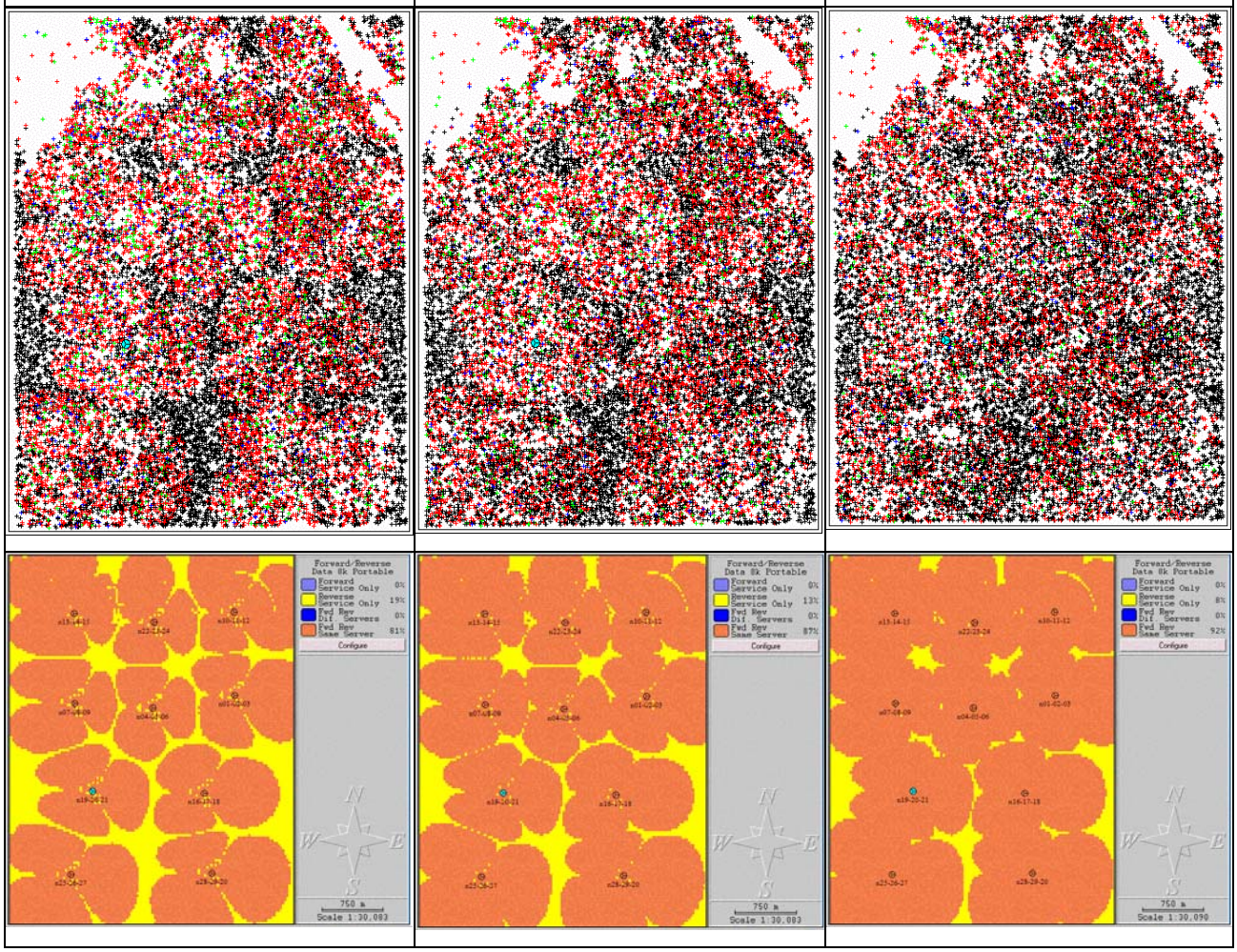


Figure 7-10. From left to right: Results for load targets of 0.5, 0.4 and 0.3. Top: graphical results for 20 snapshots simulations. Bottom: service area prediction after post-processing.

From the results presented above, it is clear that even though capacity has been sacrificed significantly, the effects of a restrict load control has provided footprint continuity for this system, an important and desirable feature of any real network.

7.3. Summary

In this chapter, we have presented techniques that allow the use of the demand characterization methodology in the process of optimizing a given network configuration.

The framework proposed in this dissertation has been complemented with the suggestion of techniques and methodologies to achieve system optimization based on accurate modeling of traffic demand, and has at the same time illustrated the importance of demand characterization in this process.

The techniques for achieving optimal design were divided into radio network topology, sector configuration parameters and radio resource management parameters. We have performed example simulations and experiments in each of those groups.

While in this discussion each technique has been applied separately and individually compared to a baseline original project, it is natural to conclude that the optimal system will result from a combination of all those techniques applied together. That exercise, however, is beyond the scope of this work, since, as mentioned before, it is not the intention to propose here a comprehensive optimization framework for 3G networks, which would be, in itself subject for another dissertation.

This concludes the development of this work. In the following chapter we summarize results and conclusions observed along this research, and indicate potential areas of future research related to this work.

Chapter 8. Conclusions and Related Areas of Research

8.1. Summary and Contributions

The main contribution of this work is the proposal of a framework for more realistic characterization of the type of traffic to be offered to next generation multimedia wireless systems and the impact of such traffic on the dimensioning of these systems. An important component of this work is the development of demand-sensitive techniques for system dimensioning and performance optimization.

Although many simulations have been described in literature for multimedia systems, their typical starting point assumes that the active user layers are known inputs. This work focuses on techniques that will allow the definition of those inputs both quantitative and qualitatively based on existing demographic data as well as on previous knowledge of voice traffic distribution on the network.

It addresses the problem of obtaining the distributions of active users based on user distributions and on the traffic models associated to each user profile. The input-modeling problem is then separated into the temporal and the spatial domains, using the concept of snapshot, which is extrapolated from current voice network dimensioning techniques. In order to deal with the temporal domain, we propose a method described as "multiplexing simulation," in which self-similar traffic models are handled through pre-simulations that allow the estimation of distributions of active users in the system. Results of multiplexing simulations performed for this research were presented and discussed.

We have next performed UMTS simulation case studies using the proposed framework in order to discuss validation and sensitivity of the methodology. Issues such as validation against simplified models, across different simulation tools and to different demand characterization inputs are discussed and multiple simulation scenarios are used for that purpose.

To complete the proposed framework, we investigated techniques that could be used to improve capacity, by optimizing system variables such as network topology, sector RF configurations and radio resource management parameters, presenting additional case-study simulations to exemplify and illustrate those topics. The contribution in this part is the proposal of a methodology that relies on demand characterization to achieve demand-sensitive system optimization.

Many systems today have already started General Packet Radio Systems (GPRS) operation as their first step on the evolution from GSM systems towards 3G. To date, as presented by many

European operators in a recent conference [63], the current "dimensioning technique" for these systems has focused only on the voice traffic through traditional methods and completely neglected the data traffic when planning the system. Although it is a valid argument that data traffic is still negligible on those networks today, it is also true that large investments are being made to migrate those networks towards offering a variety of multimedia services, and the only means to get a return on that investment includes proper dimensioning of the systems to support data traffic with appropriate quality of service without sacrifice of the existing voice subscribers. In that sense, the main benefits expected from the work presented here are to:

- Allow proper dimensioning and therefore a more reliable business case analysis (Capital and operational expenditures) for operators planning to migrate to 3G technologies or other multimedia wireless systems such as MMDS.
- Allow operator to design a proper rollout plan for the system deployment.
- Allow efficient resource allocation based on expected demand and growth.
- Allow demand-sensitive system optimization.

8.2. Related Areas of Research

Other areas of research that were **not** covered in this dissertation but could be explored in future works and benefit from the methodology proposed here are described next.

8.2.1. Customization or Selection of Resource Management Algorithms

From the system design point of view (i.e., from the operator's and design engineer's angle), there is not much room for modifications in the resource management algorithms as a method to improve capacity once the equipment vendor (with its specific implementation choices) has been selected. In this work, we chose in the simulation phase specific techniques that are typical for resource management (e.g. admission control based on target load, handoff management, etc.) and maintained those techniques throughout the studies.

Although we did not explore modifications in the algorithms themselves during this work, we believe that the area of Resource Management algorithms, which is largely explored in the literature,

could employ the methodology proposed here as a means to evaluate and compare different techniques, allowing the selection among multiple techniques to be done on a demand-sensitive way. In order to do that, modifications and customizations would have to be implemented in the simulation tools, to allow the reproduction of the techniques in the simulation phase.

In the same way, we believe that operators could also benefit from the proposed framework in the process of selecting equipment vendors and their implementations, since there is much flexibility in the standards to allow different implementations by each vendor, and a comparison methodology that focuses on the operator's specific target market would be desirable.

8.2.2. Traffic Shaping for Capacity Maximization

As part of the business plan analysis, the operator has to define the products and services to be supported by the system. The definition of services and service classes to be supported, with different QoS guarantees and treatments, has to be appropriately made based on expected market penetration, return on investment (ROI) and demand analysis.

We believe that the operator could benefit from this methodology in order to model different services to be offered by the system and to evaluate methods to maximize ROI depending on the service classes definitions of choice.

In the same way, from the knowledge of traffic mixes that could maximize system capacity and return on investment, it is possible to develop applications that stimulate the growth in the demand of certain service classes, in what could be called "traffic aware" application development.

Finally, as another means to shape traffic, the proposed methodology could be used to help in the definition of pricing policies, in order to allow a fair estimation of resource usage proportions among different classes, and also to allow more informed decisions on the stimuli for demand growth for certain service classes over others, influencing market penetration of the system.

8.2.3. Applicability to a Wireline Scenario

Since a significant part of this framework refers to the demand characterization of traffic based on geographical and demographical information sources, as well as appropriate traffic modeling, another contribution of this work is that part of it could be directly applied to the wireline scenario, such as

Internet Service Providers (ISP) based on DSL, Cable modem and ISDN technologies. This is true in what refers to methods presented to map demographic variables into users distributions of different classes, and model expected traffic behaviors, even if with different parameters. However, the models proposed here for system-level simulations would not apply directly to the wireline scenario, where system dynamics, restrictions and technology capabilities are different in many aspects. Therefore, the application of the first part of this framework (demand characterization) for the wireline scenario and the development of equivalent dimensioning and optimization techniques that are adapted and particular to those scenarios is another topic suggested as a related area of research.

Bibliography

[1] Z. Sahinoglu and S. Tekinay, "On Multimedia Networks: Self-Similar Traffic and Network Performance," IEEE Communications Magazine, vol. 37, no. 1, Jan. 1999, pp. 48-52.

[2] L. Kleinrock, *Queueing systems Volume 1: Theory*, John Wiley & Sons, New York, 1975.

[3] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed., Addison-Wesley, Reading, Mass., 1994.

[4] D.L. Spohn, *Data Network Design*, McGraw-Hill, New York, N.Y., 1997.

[5] W. Willinger and V. Paxson, "Where Mathematics Meets the Internet", Notices of the AMS, Sept. 1998.

[6] M. Sexton, A. Reid, *Broadband Networking*, Artech House, Norwood, MA, 1997.

[7] J.S. Lee, L.E. Miller, *CDMA Systems Engineering Handbook*, Artech House, Boston, MA, 1998.

[8] B. Jabbari, "Teletraffic Aspects of Evolving and Next-Generation Wireless Communications Networks," IEEE Personal Communications, vol. 3, no. 6, Dec. 1996, pp. 4-9.

[9] M.W. Oliphant, "The Mobile Phone Meets the Internet," IEEE Spectrum, vol. 36, no. 8, Aug. 1999, pp. 20-28.

[10] L.L. Peterson and B.S. Davie, *Computer Networks: a Systems Approach*, 2nd ed., Morgan Kaufmann, San Francisco, 2000.

[11] H. Heffes and D.M. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexor Performance," IEEE Journal on Selected Areas in Communications, vol. 4, no.6, Sept. 1986, pp.856-868.

[12] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, "On the Self-similar Nature of Ethernet Traffic (Extended Version)," IEEE/ACM Transactions on Networking, vol. 2, no. 1, Feb. 1994, pp. 1-15.

[13] M.E. Crovella and A. Bestravros, "Self-Similarity in World Wide Web Traffic – Evidence and Possible Causes," IEEE/ACM Transactions on Networking, vol. 5, no.6, Dec. 1997, pp. 835-845.

[14] J.W. Roberts, "Traffic Theory and the Internet," IEEE Communications Magazine, vol.39, no.1, Jan. 2001, pp. 94-99.

[15] K. Park and W. Willinger, ed., *Self Similar Network Traffic and Performance Evaluation*, John Wiley & Sons, New York, 2000.

[16] W. Willinger et al., "Self-Similarity through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level," Proceedings ACM SIGCOMM '95, 1995, pp. 100-113.

[17] J. Roberts, U. Mocci and J. Virtamo, ed., *Broadband Network Teletraffic: Final Report of Action COST 242,* Springer, Berlin, 1996.

[18] L.Z. Ribeiro and L.A. DaSilva, "A Framework for the Dimensioning of Broadband Mobile Networks Supporting Wireless Internet Services," IEEE Wireless Communications, vol. 9, no.3, June 2002, pp. 6-13.

[19] 3GPP TS 23.101 version 3.1.0 Release 1999, *General UMTS Architecture*, 3rd Generation Partnership Project, Technical Specification, 2001-12.

[20] 3GPP TS 23.107 version 3.7.0 Release 1999, *QoS Concept and Architecture*, 3rd Generation Partnership Project, Technical Specification, 2002-01.

[21] 3GPP TS 23.107 version 4.3.0 Release 4, *QoS Concept and Architecture*, 3rd Generation Partnership Project, Technical Specification, 2002-01.

[22] 3GPP TS 23.107 version 5.3.0 Release 5, *QoS Concept and Architecture*, 3rd Generation Partnership Project Technical Specification, 2002-01.

[23] J. Laiho, A. Wacker and T. Novosad, ed., *Radio Network Planning and Optimization for UMTS*, John Wiley & Sons, West Sussex, England, 2002.

[24] V.K. Garg and O.T.W. Yu, "Integrated QoS Support in 3G UMTS Networks," IEEE Wireless Communications and Networking Conference, WCNC'2000, vol.3, Sept. 2000, pp. 1187 –1192.

[25] V.K. Garg and M. Garg, "Role of MAC/RLC in achieving higher data rates and QoS in third generation (3G) UMTS system" IEEE Int. Conf. Personal Wireless Communications, ICPWC'2000, Dec. 2000, pp. 459 –463.

[26] M.D. Yacoub, *Wireless Technology: Protocols, Standards, and Techniques*, CRC Press, Boca Raton, CA, 2001.

[27] S. Dixit, Y. Guo and Z. Antoniou, "Resource Management and Quality of Service in Third-Generation Wireless Networks," IEEE Communications Magazine, vol. 39, no. 2, Feb. 2001, pp. 125-133.

[28] L. Jorguseski, J. Farserotu and R. Prasad, "Radio Resource Allocation in Third-Generation Mobile Communication Systems," IEEE Communications Magazine, vol. 39, no. 2, Feb. 2001, pp. 117-123.

[29] N. Dimitriou, R. Tafazolli and G. Sfikas, "Quality of Service for Multimedia CDMA," IEEE Communications Magazine, vol. 38, no. 7, July 2000, pp.88-94.

[30] K. Kim, Y. Han, C.H. Yim and K.S. Jeong, "A Call Admission Algorithm with Optimal Power Allocation for Multiple Class Traffic in CDMA Systems," IEEE Vehicular Technology Conference, VTC'2000 52nd , vol. 6, pp. 2666-2671.

[31] Y. Guo and H. Chaskar, "A Framework for Quality of Service Differentiation on 3G CDMA Air Interface," IEEE Wireless Communications and Networking Conference, WCNC'2000, vol.3, Sept. 2000, pp. 975- 979.

[32] Y. Guo and B. Aazhang, "Resource Allocation and Capacity in Wireless CDMA Networks Using Adaptive Power Control and Antenna Array Multiuser Receiver," IEEE Symposium on Computers and Communications, ISCC'2000 5th , July 2000, pp. 723 –730.

[33] Y. Guo and B. Aazhang, "Call Admission Control in Multi-Class Traffic CDMA Cellular System using Multiuser Antenna Array Receiver," IEEE Vehicular Technology Conference, VTC'2000-Spring, Tokyo, vol. 1, May 2000, pp. 365-369.

[34] Y. Guo and B. Aazhang, "Capacity of Multi-Class Traffic CDMA System with Multiuser Receiver," IEEE Wireless Communications and Networking Conference, WCNC'1999, vol.3, Sept. 1999, pp. 975- 979.

[35] R. Koodli and M. Puuskari, "Supporting Packet-Data QoS in Next-Generation Cellular Networks," IEEE Communications Magazine, vol. 39, no. 2, Feb. 2001, pp. 180-187.

[36] D. Imbeni and M. Karlsson, "Quality of Service Management for Mixed Services in WCDMA," IEEE Vehicular Technology Conference, VTC'2000-Fall, Boston, vol. 2, Sept. 2000, pp. 565-572.

[37] J. Jiang and T.H. Lai, "An Efficient Approach to Support QoS and Bandwidth Efficiency in High-Speed Mobile Networks," IEEE Int'l Conf. Comm., ICC 2000, New Orleans, vol. 2, June 2000, pp. 980-984.

[38] J. Jiang and T.H. Lai, "Call Admission Control vs. Bandwidth Reservation: Reducing Handoff Call Dropping Rate and Providing Bandwidth Efficiency in Mobile Networks," Int'l Conference on Parallel Processing, Toronto, Aug. 2000, pp. 581-588.

[39] J. Jiang and T.H. Lai, "Bandwidth Management Providing Guaranteed Call Dropping Rates for Multimedia Mobile Networks," IEEE Int'l Conference on Multimedia and Expo, ICME 2000, New York, vol. 2, Aug. 2000, pp. 963-966.

[40] K. Sipilä et al., "Estimation of Capacity and Required Transmission Power of WCDMA Downlink Based on a Downlink Pole Equation," IEEE Vehicular Technology Conference, VTC'2000-Spring, Tokyo, vol. 2, May 2000, pp. 1002-1005.

[41] J. Laiho-Steffens, A. Wacker and P. Aikio, "The Impact of the Radio Network Planning and Site Configuration on the WCDMA Network Capacity and Quality of Service," IEEE Vehicular Technology Conf., VTC'2000-Spring, Tokyo, vol. 2, May 2000, pp. 1006-1010.

[42] H. Holma and A. Toskala, ed., *WCDMA for UMTS – Radio Access for Third Generation Mobile Communications*, John Wiley & Sons, 2000.

[43] V.K. Garg, *IS-95 CDMA and cdma 2000: Cellular/PCS Systems Implementation*, Prentice Hall, Upper Saddle River, 1999.

[44] S. Hämäläinen, H. Holma and A. Toskala, "Capacity Evaluation of a Cellular CDMA Uplink with Multiuser Detection," IEEE Int'l Symp. Spread Spectrum Techniques and Applications, Mainz, Germany, vol.1, Sept. 1996, pp. 339-343.

[45] K. Sipilä et al., "Modeling the Impact of the Fast Power Control on the WCDMA Uplink," IEEE Vehicular Technology Conf., VTC'1999-Spring, Houston, vol. 2, May 1999, pp. 1266-1270.

[46] K. Sipilä et al., "Soft Handover Gains in a Fast Power Controlled WCDMA Uplink," IEEE Vehicular Technology Conf., VTC'1999-Spring, Houston, vol. 2, May 1999, pp. 1594-1598.

[47] A. Wacker et al., "The Impact of the Base Station Sectorisation on WCDMA Radio Network Performance," IEEE Vehicular Technology Conf., VTC'1999-Fall, Amsterdam, vol. 5, Sept. 1999, pp. 2611-2615.

[48] A. Wacker et al., "Static Simulator for Studying WCDMA Radio Network Planning Issues," IEEE Vehicular Technology Conf., VTC'1999-Spring, Houston, vol. 3, May 1999, pp. 2436-2440.

[49] J. Laiho-Steffens et al., "The Impact of the Subscriber Profile on WCDMA Radio Network Performance," IEEE Vehicular Technology Conf., VTC'1999-Fall, Amsterdam, vol. 5, Sept. 1999, pp. 2490-2494.

[50] J. Laiho-Steffens, A. Wacker and K. Sipilä, "Verification of 3G Radio Network Dimensioning Rules with Static Network Simulations," IEEE Vehicular Technology Conf., VTC'2000-Spring, Tokyo, vol. 1, May 2000, pp. 478-482.

[51] R.G. Akl et al., "Multicell CDMA Network Design," IEEE Trans. Vehicular Technology, vol. 50, no. 3, May 2001, pp. 711-722.

[52] J. Yang and W.C. Lee, "Design Aspects and System Evaluation of IS-95 based CDMA Systems," 1997 IEEE 6th IEEE Int. Conf. Universal Personal Communications, vol. 2, October 1997, pp. 381-385.

[53] M. Soleimanipour and G.H. Freeman, "A Realistic Approach to the Capacity of Cellular CDMA systems," IEEE Vehicular Technology Conf. 1996, vol. 2, May 1996, pp. 1125-1129.

[54] T. Suzuki, K.Takeo, M. Nishino and Y. Amezawa, "Microcell Quality Control Scheme for PCS CDMA Systems considering Non-Uniform Traffic Distribution," IEEE Int. Conference on Universal Personal Communications, vol. 1, Oct. 1993, pp. 239-243.

[55] R.T. Love et al., "A Pilot Optimization Technique for CDMA Cellular Systems," IEEE VTS 50th Vehicular Technology Conf., VTC 1999-Fall, 1999, vol. 4, pp. 2238-2242.

[56] D. Grillo, R.A. Skoog, S. Chia and K.K. Leung, "Teletraffic Engineering for Mobile Personal Communications in ITU-T Work: The Need to Match Practice and Theory," IEEE Personal Communications, vol. 5, no. 6, Dec. 1998, pp. 38-58.

[57] P. Wirth, "The Role of Teletraffic Modeling in the New Communications Paradigms," IEEE Comm., Aug. 1997, pp. 86-92.

[58] J.N. Daigle, *Queueing Theory for Telecommunications*, Addison-Wesley, Reading, MA, 1992.

[59] M.D. Yacoub, *Foundations of Mobile Radio Engineering*, CRC Press, Boca Raton, FL, 1993.

[60] J.S. Lee, L.E. Miller, *CDMA Systems Engineering Handbook*, Artech House, Boston, MA, 1998.

[61] T. G. Robertazzi, *Computer Networks and Systems: Queueing Theory and Performance Evaluation*, 2nd ed., Springer-Verlag, 1994.

[62] T. Halonen, J. Romero and J. Melero, ed., *GRM, GPRS and EDGE Performance*, John Wiley & Sons, West Sussex, England, 2002.

[63] IIR Conference on GPRS and UMTS Cell Planning, Vienna, Austria, June 2002.

[64] A. Wacker, J. Laiho, K. Sipilä, K. Heiska, K. Heikkinen, *NPSW Matlab Implementation of a Static Radio Network Planning Tool for Wideband CDMA*, Software Manual, v5.0.0, (accompanying documentation for ref. [23]), July 2001.

[65] S. Kumar and S. Nanda, "High Data-Rate Packet Communications for Cellular Networks Using CDMA: Algorithms and Performance," IEEE Journal on Selected Areas in Communications, vol.17, no. 3, March 1999, pp. 472-492.

[66] Z. Dziong, M. Krishnan, S. Kumar and S. Nanda, "Statistical 'Snap-Shot' for Multic-Cell CDMA System Capacity Analysis," IEEE VTC 1999, pp. 1234-1237.

[67] N. Mandayam, P. Chen and J. Holtzman, "Minimum Duration Outage for Cellular Systems: A Level Crossing Analysis," IEEE Vehicular Technology Conference, VTC 1996 46th, vol. 2, 28 Apr-May 1996, pp. 879-883.

[68] S. Baxter, *Wireless CDMA RF Performance Optimization* (Course Notes – RF200 v3.46), Scott Baxter & Associates, May 2002.

Vita

Leila Zurba Ribeiro received her M.S. degree in Electrical Engineering from the State University of Campinas, Brazil, in July 1995 and her B.S. degree in Electrical Engineering from the Federal University of Santa Catarina, Brazil in December 1992, where she was awarded for highest final grades. For her Ph.D. research, she received an IGERT (Integrated Graduate Education and Research Training) fellowship, funded by the National Science Foundation (NSF). Her MS research work was funded by scholarships from the CNPq and CPqD (both are Brazilian government research institutions). She is also a member of the IEEE.

Leila joined CelPlan Technologies, Inc. in December 1993, in Campinas, Brazil, as a design engineer, and has since then worked in the design and optimization of wireless communications systems. In 1996 she moved to US branch of the same company and in 1999 she became Director of Systems Engineering. She currently manages and oversees optimization and design projects for wireless networks and is also involved in the process of research and specifications for the development of design and optimization tools. In parallel with working at CelPlan, Leila has also taught as an adjunct professor at the MS in Telecommunications program at George Mason University in the Spring 2001.

Recent Publications:

L.Z. Ribeiro and L.A. DaSilva, “A Framework for the Dimensioning of Broadband Mobile Networks Supporting Wireless Internet Services,” IEEE Wireless Communications, vol. 9, no.3, June 2002, pp. 6-13.

L. Z. Ribeiro and L. A. DaSilva, "Traffic Demand Characterization for Multimedia Mobile Networks," IEEE Vehicular Technologies Conference, pp. 620-624, October 2001.

Contact Information:

Email: leilazr@attglobal.net