QOS-AWARE ALLOCATION OF RESOURCES IN WIRELESS NETWORKS

By

MOHAMED N. MOUSTAFA

A dissertation submitted to the Graduate Faculty in Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

2001

UMI Number: 3024819

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

9/5/2001
Date

Prof. Ibrahim Habib
Chair of Examining Committee

9/5/2001
Date

Dean Mumtaz Kassir
Mumtaz Kassir
Executive Officer

Prof. Mohamed Ali (CUNY. CCNY. EE Dept.)

Prof. Leonid Roytman (CUNY. CCNY. EE Dept.)

Prof. Yi Sun (CUNY. CCNY. EE Dept.)

Prof. Nirwan Ansari (NJIT. ECE Dept.)

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

QOS-AWARE ALLOCATION OF RESOURCES IN WIRELESS NETWORKS

by

MOHAMED N. MOUSTAFA

Adviser: Professor Ibrahim Habib

The convergence of mobility, multimedia and the Internet is finally seen possible through third generation (3G) wireless systems. A key part of this evolution is Quality of Service (QoS), which is simply a set of service requirements (e.g., delay, bit error rate, guaranteed bit rate) to be met by the network while transporting traffic streams from source to destination. Essential to the notion of QoS is resource management that resolves users competition to utilize network resources according to their QoS agreements. In this thesis, we propose a novel wireless resource scheduler for controlling the transmission power and bit rate of mobile calls cooperatively while previous work has focused on handling them separately. The proposed scheme is called the Genetic Algorithm for Mobiles Equilibrium (GAME). Based on an evolutionary computational model, GAME assigns optimally both transmitting power and bit rate to every mobile station. Optimal allocation is in the sense that every user gets only enough resources necessary for meeting or exceeding its QoS requirements. Having done that, we gain further benefits as well. In addition to QoS provisioning, lower transmitting power extends a mobile station battery life and minimizes the interference seen by other users. Moreover, Improved coverage efficiency reduces infrastructure cost by requiring fewer base stations per square kilometer. In short, GAME raises the number of QoS-satisfied mobile stations in a cell. Various simulations show improvements achieved over the established scheme.

Acknowledgments

I would like to thank my adviser Prof. I. Habib for proposing the research topic, providing continuous support and directing this work.

I would like also to thank Dr. M. Naghshineh (IBM Watson Research Center) for his contribution in proposing the research topic and for his helpful comments.

I would like also to thank my examination committee for reviewing this work, and my colleagues for fruitful discussions.

Last but not least, I would like to thank my parents for their love and encouragement, my wife for her support and devotion throughout these years. I am grateful to all of you.

To my parents, my wife, and my children.

Mohamed Nabil Moustafa
September 5, 2001

Table of Contents

1. Introduction 1

1.1. System Architecture 3

1.2. Quality of Service Model 5

1.3. Radio Resource Management 6

2. Resource Control 8

2.1. Power Control 9

2.2. Combined Power and Rate Control 11

3. Genetic Algorithm for Mobiles Equilibrium 14

3.1. Radio Interface 14

3.2. Genetic Algorithms Overview 16

3.2.1. Population 17

3.2.2. Reproduction 17

3.2.3. Crossover 18

3.2.4. Mutation 18

3.2.5. Anatomy 18

3.3. Chromosome 19

3.4. Fitness Function 20

3.5. Skeleton 23

4. Multimedia Mobile Environments 25

4.1. Multimedia Services 25

4.2. Cellular Models 26

4.2.1. Outdoor General Model 26

4.2.2. Microcellular Pedestrian Model 27
5. Experiments 33
5.1. System Capacity 34
5.2. Quality of Service 36
5.3. Power Consumption 44
5.4. Spectrum Efficiency 49
5.5. Coverage Efficiency 51
5.6. Control Period Effect 53
5.7. Convergence and Scalability 58
6. Putting It Together 61
6.1. GAME Signaling on Standard Interface 61
MS-BS Signaling 63
6.2. End-To-End QoS Provisioning 64
6.2.1. UMTS QoS Architecture 64
6.2.2. Internet Protocol QoS Models 65
6.2.3. QoS Interworking and Parameter Mapping 66
7. Conclusions 68
References 70
List of Publications 73
Journal Papers 73
IETF RFCs/Internet Drafts 73
Conference Papers 73

List of Tables

Table 1. UMTS QoS Classes: Fundamental Features and attributes [2]........ 6

Table 2 Single Point Crossover Operator Effect........ 18

Table 3. Traffic Characteristics and QoS Requirements of Tested Services........ 25

Table 4 Path-Loss Formulas for High Rise Building Environment [7]........ 28

Table 5 Pedestrian Mobility Model Parameters [4]........ 31

List of Figures

Fig. 1 Simplified UMTS Architecture. 5

Fig. 2 Radio Resource Manager (RRM). 7

Fig. 3 BS-MS interface in case of IS-95 CLPC. 11

Fig. 4 *P-R* space with shaded permissible search regions. 13

Fig. 5 BS-MS interface in case of GAME-C. 16

Fig. 6 GAME Chromosome Format. 20

Fig. 7 Fitness Function Components. 22

Fig. 8 GAME Skeleton Flow Diagram. 24

Fig. 9 Manhattan-like urban model. 29

Fig. 10 Probabilities of MS movements at cross streets. 31

Fig. 11 Combinations of admissible mobile stations controlled by GAME-C 35

Fig. 12 Combinations of admissible mobile stations controlled by TPC 36

Fig. 13 Outage Probability vs. Voice Users. 40

Fig. 14 Outage Probability vs. Data Users. 40

Fig. 15 Outage Probability vs. Video Users. 41

Fig. 16 Average (Eb/N-Θ) vs. Voice Users. 41

Fig. 17 Average (Eb/N-Θ) vs. Data Users. 42

Fig. 18 Average (Eb/N-Θ) vs. Video Users. 42

Fig. 19 Average transmitter bit Rate vs. Voice Users. 43

Fig. 20 Average transmitter bit Rate vs. Data Users. 43

Fig. 21 Average transmitter bit Rate vs. Video Users. 44

Fig. 22 Average transmitter Power vs. Voice users. 46

Fig. 23 Average transmitter Power vs. Data users. 46

Fig. 24 Average transmitter Power vs. Video users. 47

Fig. 25 Average Cell Load vs. Voice users. 47

Fig. 26 Average Cell Load vs. Data users. 48

Fig. 27 Average Cell Load vs. Video users. 48

Fig. 28 Percentage of Satisfied mobile voice users versus System Load. 50

Fig. 29 Percentage of Satisfied mobile data users versus System Load. 50

Fig. 30 Voice Outage Probability versus Base Station coverage area. 52

Fig. 31 Data Outage Probability versus Base Station coverage area. 53

Fig. 32 Outage Probability vs. Control period. 56

Fig. 33 Average (Eb/N-Θ) vs. Control period. 56

Fig. 34 Average transmitter Power of Data Users vs. Control period. 57

Fig. 35 Average transmitter bit Rate of Data Users vs. Control period. 57

Fig. 36 GAME Convergence. 60

Fig. 37 GAME Scalability. 60

Fig. 38 Mobile Station-Base Station Standard Interface. 63

Fig. 39 End-to-End QoS Architecture. 65

1. Introduction

During last decade, there has been tremendous growth in two technological sectors: the Internet and wireless communications. The International Telecommunication Union (ITU) projects that by the end of 2002; there will be approximately 600 million Internet users. Running almost concurrently with this growth in the Internet has been the equally extraordinary growth in the number of mobile cellular networks and subscribers. It is expected to have over one billion mobile phone users by the end of 2002. The Internet and wireless communications have conventionally been regarded as separate technologies. This is because originally, the Internet was designed to carry mainly data traffic, whilst wireless networks were designed to carry mainly voice traffic. This boundary has become increasingly blurry in recent years, especially with the introduction of several proposals for the International Mobile Telecommunications-2000 (IMT-2000) in the International Telecommunication Union (ITU) [6, 31]. These third-generation (3G) Mobile Communications Systems are poised to be multi-service platforms supporting voice, video and data services with bit rates up to 2 Mb/s. In other words, it is the convergence of mobility, Internet, and multimedia.

An important part of this evolution is Quality of Service (QoS), which is simply a set of service requirements to be met by the network while transporting a traffic stream from source to destination. QoS attributes are usually specified in terms of Bit Error Rate (BER), guaranteed Bit Rate (R^G), delay, and so on. QoS provision is crucial for the success of a variety of packet data services, especially in a bandwidth-constrained and error prone environment such as cellular networks. Consequently, it is important to address resource management in the emerging cellular networks architectures, thus resolving users competition to utilize network resources according to their QoS agreements.

Since recent 3G partnership projects [4, 5] indicate that wideband Code Division Multiple Access (CDMA) will be the prevailing air interface in 3G Radio Access Networks (RAN), we concentrate in this study on such environment. In a CDMA network, many QoS measures, including BER, depend on the received bit energy-to-noise density ratio E_b/N_o which in turn depends as detailed in next chapter on both Mobile Station (MS) transmitter power and bit rate.

Power control is a means primarily designed to compensate for the loss caused by propagation and fading. One of the most powerful methods is the Transmitter Power Control (TPC) specified within the Interim standard 95 (IS-95) [30] that is currently operating in all CDMA based systems. In fact this scheme is included as well in all wideband CDMA 3G proposals [6, 31]. Controlling only power has some disadvantages though. First major shortcoming is that many QoS parameters are not supported since most methods were developed to handle only voice traffic. This means that the Base Station (BS) does not differentiate between users traffic during power assignments and therefore there is no bit rate guarantee. Another main problem arises when an MS reaches its maximum transmitting power level without reaching its required E_b/N_o. This situation always leads to the drop of this MS connection when this condition persists for some time [14].

In order to overcome these shortcomings we propose in this study a novel wireless resource scheduler, that assigns to every MS not only transmitter power but also transmission bit rate concurrently. Our scheme relies mainly on the Genetic Algorithm for Mobiles Equilibrium (GAME) to allot each MS the minimum resources possible for meeting or exceeding its QoS requirements. We also integrate GAME with the powerful power control of IS-95 to have the backup insurance of the current standard method, and to facilitate real time implementation. We called the integrated scheme (GAME-C). Having done that, we gain additional benefits as well. In addition to QoS provisioning, MS battery survives longer by always preferring lower transmission power, and BS coverage efficiency improves by decreasing the probability of blocking new connections or dropping current ones.

The remainder of this thesis is organized as follows. We first present a brief overview of 3G system architecture followed by a review of state-of-the art CDMA radio resources scheduling techniques in chapter 2. The proposed scheme is detailed in chapter 3. In chapter 4, we discuss the multimedia environment used in our study as well as the different applicable cellular configurations. Experiments with some numerical results are presented in chapter 5 illustrating the deployment of the proposed genetic algorithm as well as analyzing some parameters and its impact on performance. In Chapter 6 we describe different QoS classes requirements and some thoughts on QoS mapping from radio access to fixed core networks. We then close with some concluding remarks.

1.1. System Architecture

A simplified architectural model of the Universal Mobile Telecommunication System (UMTS) [6] is shown in Fig. 1. It mainly consists of three components: wireless Mobile

Station (MS)[1], Radio Access Network (RAN) and Core Network (CN). RAN, which performs all radio access specific procedures, can be viewed as a network extension providing wireless access to the Core Network (CN).

This wireless extension essentially consists of a set of interconnected Radio Network Systems (RNS). Each RNS is responsible for the resources of its set of cells and contains several Base Stations (BS) connected through a Radio Network Controller (RNC). Each BS serves a group of MSs currently residing in a cell and is responsible for intra-cell control while RNC is in charge of the inter-cell operations like the handover decisions. On the fixed network side, CN consists mainly of some edge nodes (EN) and gateway nodes (GN) that interconnect with external networks like Internet Protocol (IP), ISDN, PSTN, or old mobile networks. Thus, in addition of being able to place voice calls, an MS has the opportunity to establish communication links with any types of networks including the Internet to exchange multimedia contents, download e-mails, browse the web, etc...

Recent releases of UMTS provide user terminals with different IP-based services and extend CN features with Internet Engineering Task Force (IETF) protocols that support wide range of multimedia services [1]. Another solution for UMTS-Internet integration is to reuse the latest RAN and attaching it directly to an enhanced Mobile IP backbone through a standard interface [24], thus eliminating the usage of several CN internal nodes.

---

[1] Mobile Terminal (MT) consists of Terminal Equipment (TE) in addition to wireless Mobile Station (MS).

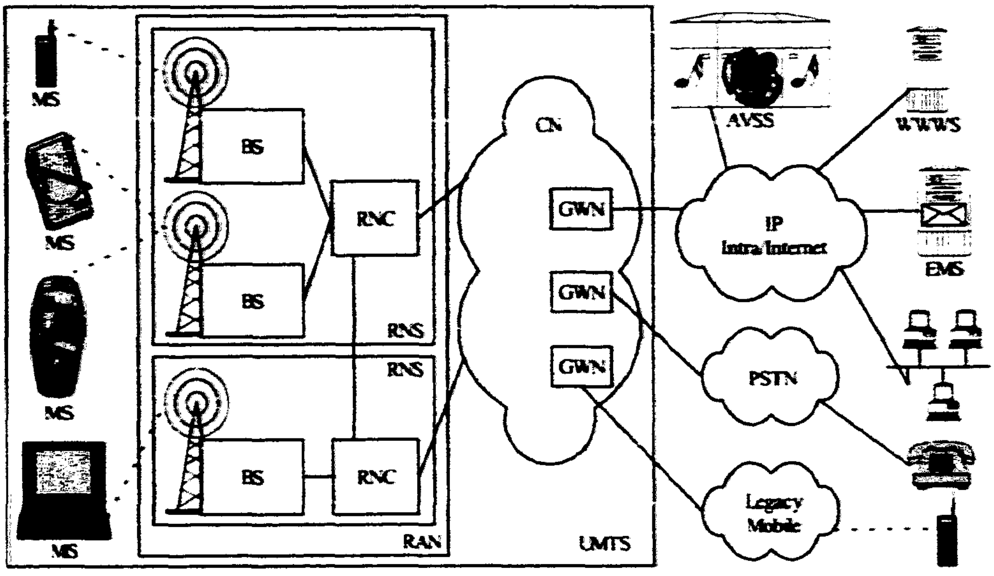

Fig. 1 Simplified UMTS Architecture.

MS: Mobile Station. BS: Base Station. RNC: Radio Network Controller. RNS: Radio Network System. RAN: Radio Access Network. CN: Core Network. GWN: Gateway Node. PSTN: Public Switch Telephone Network. AVSS: Audio/Video Streaming Server. WWWS: World Wide Web Server. EMS: Electronic Mail Server.

1.2. Quality of Service Model

Most of the services featured in UMTS can be divided into main four QoS classes primarily discriminated based on their ability to tolerate delay and bit error rate (BER). As defined in [2] and seen in Table 1, *Conversational* and *Streaming* classes preserve time relation between information entities of the stream and have a guaranteed bit rate R^G from the network. They are suitable to carry real time traffic since they define an upper limit on transfer delay within their QoS profiles. The main difference between them is the maximum transfer delay value. Conversational traffic is subject to the strict human perception in a voice over IP talk or video teleconferencing while real time Streaming traffic has slightly flexible delay requirements.

On the other hand, *Interactive* and *Background* classes are mainly intended to represent conventional Internet applications, e.g., web browsing, Telnet, FTP, and Email. They concentrate on preserving payload content by means of channel coding and retransmission, and are characterized as best effort traffic since neither R^G nor maximum delay is specified. Interactive class traffic essentially follows a Request-Response pattern, e.g., interactive web browsing or message chatting while Background class is designated for background traffic, e.g., Email or file downloading (the destination is not expecting data within a certain time). Interactive applications have higher priority than Background ones in terms of resource assignment to ensure responsiveness.

Table 1. UMTS QoS Classes: Fundamental Features and attributes [2].

		Conversational class	**Streaming class**	**Interactive class**	**Background class**
Features	***Application***	Voice over IP	Audio/Video	WWW	Email
	Delay	Strict and low	Bounded	Tolerable	Unbounded
	BER	$\leq 10^{-3}$	$\leq 10^{-5}$	$\leq 10^{-8}$	$\leq 10^{-8}$
Defined Attributes	***Maximum Bit Rate***	Yes	Yes	Yes	Yes
	Guaranteed Bit Rate (R^G)	Yes	Yes	No	No
	Maximum Delay	Yes	Yes	No	No

1.3. Radio Resource Management

As of Now, the Internet has only supported best effort service. That means the network transports traffic from source to destination without any type of QoS guarantee. This is about to change. But QoS guarantees can be attained only with appropriate resource allocation. That is why every BS should contain a Radio Resource Management (RRM) module that tries to preserve QoS across the RAN.

The RRM main role is to assign resources to users according to their QoS requirements. As Shown in Fig. 2, RRM mission starts by performing connection admission control (CAC). Since the decision is based upon resources availability, CAC consults the Radio Resource Scheduler (RRS) before accepting or rejecting the requested call. Some smart CAC can also include self-QoS functionality in addition to relying on RRS [11]. Upon call acceptance, the traffic classifier (TC), another RRM component, categorizes the incoming traffic according to its QoS specification, which is typically included in each packet header, e.g. IETF Diffserv Service Code Point [21]. Data flows are then directed to a corresponding queue according to its Qos field. Each Qos class is represented by at least one queue. Finally, the Traffic Dispatcher (TD) drains the multiple queues according to some priority logic after getting the assigned radio resources from the RRS, which relies on the channel conditions and the requested QoS in its response.

Based on the above, it is very evident that RRS bears great responsibility in having a successful RRM. That is why RRS resource control functionality is covered in details in next chapter.

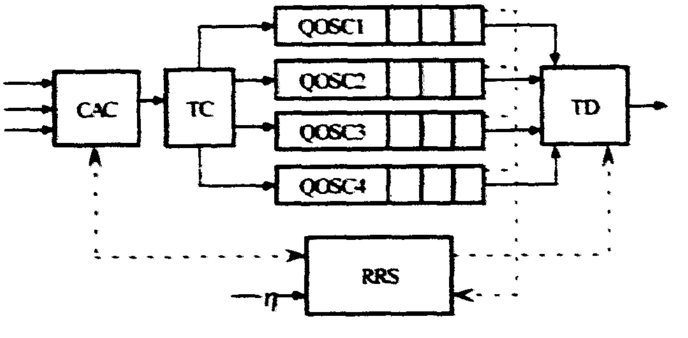

Fig. 2 Radio Resource Manager (RRM).

CAC: Connection Admission Controller, TC: Traffic Classifier, RRS: Radio Resource Scheduler, η: Interference and Noise Measurements, TD: Traffic Dispatcher, QOSCx: QoS Class x queue.

2. Resource Control

An important goal in a multiple-access system, such as the IMT-2000, is to maximize the number of simultaneous users it can accommodate. If each MS is assigned the minimum resources necessary for meeting or exceeding its QoS requirements, the capacity of the system will be maximized. That is why RRS is an essential component in QoS-aware BS, as indicated in the previous chapter. RRS has two important radio resources to control: MS transmitting power P, and bit rate R. In this chapter, we review some resource scheduling techniques highlighting their pros and cons.

In CDMA networks, BER, an important QoS measure, depends on the received bit energy-to-noise density ratio E_b/N_o given by

$$\left(\frac{E_b}{N_o}\right)_i = \frac{G_{BSi} P_i / R_i}{\left(\sum_{j \neq i}^{M} G_{BSj} P_j + \eta\right) \Big/ W} \qquad (1)$$

where W is the total spread spectrum bandwidth occupied by the CDMA signals. G_{BSi} denotes the path loss on the path between MS i and its BS. η denotes background noise due to thermal noise contained in W and M is the number of mobile users. The transmitted power of MS i is P_i which is usually limited by a maximum power level as

$$0 \leq P_i \leq P_i^{max}, \quad for \quad 1 \leq i \leq M \qquad (2)$$

R_i is the information bit rate transmitted by MS i. This rate is bounded by the peak bit rate R_i^P, designated in the traffic contract once this user has been admitted into the system.

$$0 \leq R_i \leq R_i^P, \quad for \quad 1 \leq i \leq M \qquad (3)$$

An increase in the transmission power of a user increases its E_b/N_o, but increases the interference to other users, causing a decrease in their E_b/N_os. On the other hand, an increase in the transmission rate of a user deteriorates its own E_b/N_o.

As denoted in (1), it is evident that manipulating transmission powers and bit rates of the mobile stations amounts directly to controlling QoS that is partly specified as a pre-specified target E_b/N_o value (Θ).

2.1. Power Control

Power control is a means primarily designed to compensate for the loss caused by propagation and fading. One of the most powerful methods is the Transmitter Power Control (TPC) specified within the Interim standard 95 (IS-95) [30] that is currently operating in all CDMA based systems. In fact this scheme is included as well in all wideband CDMA 3G proposals [6, 31].

TPC contains two different power control mechanisms. In the uplink (MS to BS), both open loop (OLPC) and fast closed loop power control (CLPC) are employed. In Open loop scheme, the MS determines its transmit power based on the measured received signal strength from the BS. OLPC has two main functions: it adjusts the initial access mobile transmission power and compensates large abrupt variations in the path loss attenuation. Since the uplink and downlink fading processes are not totally correlated, the OLPC cannot compensate for the uplink fading. To account for this independence CLPC is used. In this mode, as depicted in Fig. 3, the BS measures the received E_b/N_o[2] over a 1.25 ms period, and compares that to the target Θ. If the received $E_b/N_o < \Theta$, a "0" is

[2] IS-95 standard suggests that the received signal strength should be measured. However, in practice usually the SIR or E_b/N_o are used, since they have direct impact on the BER.

generated to instruct the MS to increase its power, otherwise, a "1" is generated to instruct the MS to decrease its power. These commands instruct the MS to adjust transmitter power by a predetermined amount, usually 1 dbm.

Other previous work has focused on finding adequate power levels that maximize the received bit energy-to-noise density ratio (E_b/N_o) where the transmission rate of each user is fixed [15]. Stochastic power control algorithms based on the mobile matched filter output is introduced in [25]. The authors in [28] proposed a circuit switched multirate DS-CDMA system based on a closed-form power control function. Recently, [9] presented an algorithm for controlling mobiles transmitter power levels following their time varying transmission rates.

Since MS transmitting power is constrained, as indicated in (2), receiving a "0" command from the BS is useless to an MS while already transmitting with its maximum power, P^{max}. In fact, what the MS needs under such circumstances is a command to reduce its transmitting rate instead of increasing power as can be deduced from (1) to boost its E_b/N_o. Without such rate reduction, we risk dropping this MS connection. However, the rate manipulation has to be done wisely in order not to break the other QoS requirement, which is the guaranteed bit rate R^G, as indicated in Table 1. This is simply a situation where our proposed scheme GAME excels as detailed in next chapter. It assigns power and rate values to each MS every control period such that the number of simultaneous users is maximized while preserving both QoS parameters: BER and R^G.

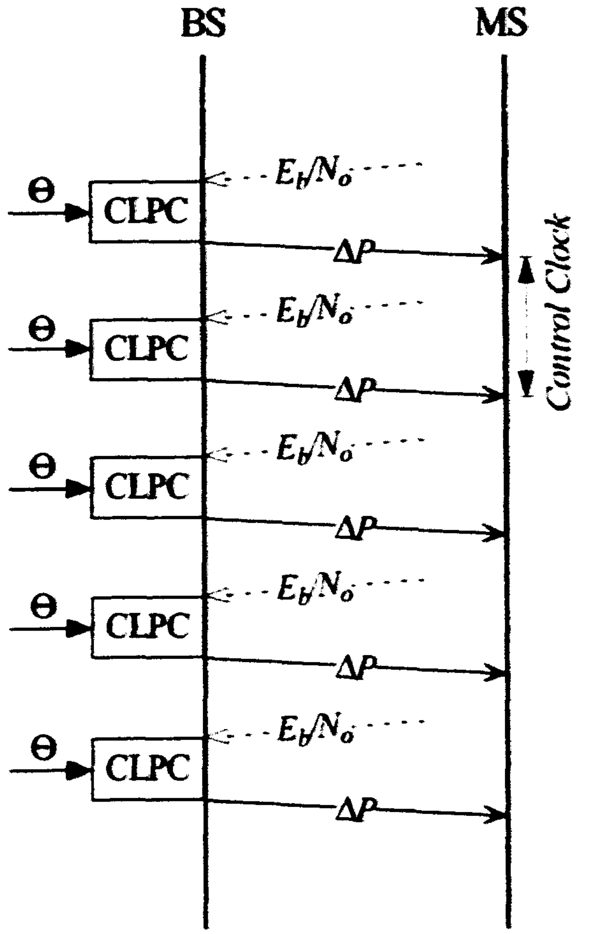

Fig. 3 BS-MS interface in case of IS-95 CLPC.

Every *control clock*, BS measures the received E_b/N_o of each MS, compares it to a threshold Θ and issues one bit feedback ΔP to MS either to increase or decrease its power by a pre-specified amount usually 1dbm.

2.2. Combined Power and Rate Control

Giving equation (1) a closer look, we can deduce that bit rate control gives an RRS additional tool to preserve QoS (BER), i.e. E_b/N_o, and hence to overcome CLPC limitation. Simply, if the BS RRS advises MS i to decrease its rate R_i at bad channel conditions instead of increasing its power P_i, then the outage probability of MS i shall decrease. The outage probability, a measure of user's satisfaction, is defined as the probability of having MS E_b/N_o falling below its threshold Θ. Several recent studies investigated this possibility. We discuss here some of them.

The authors in [22] proposed manipulating spreading gain in CDMA data networks beside power control. Spreading gain is simply the ratio of the system bandwidth (W) to the transmitter bit rate (R). Each BS appropriately balances a user's desire for a high transmission power with the amount of interference it will generate to other users. However, the optimization problem was formulated to handle only data streams as best-effort traffic. Therefore, this method does not preserve pre-specified bit rate QoS constraint (R^G).

Another scheme, proposed in [26] and indicated hereafter as (PRA), considers power and rate adaptation one at a time. The authors presented two alternatives for *P-R* adaptation. In first mode, *Mode-A*, MS transmitting bit rate is left uncontrolled while power can be increased up to some value ($P^{Limit\text{-}A} < P^{max}$) to preserve the E_b/N_o. In case of bad channel conditions, transmitting power is limited to $P^{Limit\text{-}A}$ while bit rate is reduced to meet the Θ requirement. On the other hand, *Mode-B* manipulates only bit rate. It completely suspends transmitting data when the channel gain is below a threshold. Otherwise, it reduces bit rate while having constant MS transmitting power at some level ($P^{Limit\text{-}B} < P^{max}$). Evidently, PRA provides more flexibility to the resource-scheduling problem over TPC by allowing bit rate manipulation. This flexibility can be easily noticed by visualizing the constrained search region in *P-R* space. As depicted in Fig. 4-(a), TPC set is restricted to simply horizontal lines search since R is not controlled. In the meantime, PRA has more permissible area by including vertical lines corresponding to R

adaptation when P reaches the threshold level P^{Limit}. As shown in Fig. 4-(b), PRA also embraces the space origin, which corresponds to transmission suspension of *Mode-B*. Clearly, *Mode-B* provides higher power gain but introduces time delay. Beside lack of commitment to the QoS guaranteed bit rate level R^G, the major drawback of PRA is its high dependence on the power level limit P^{Limit}, especially that the authors have not proposed a procedure for defining or optimizing it [26]. Fig. 4-(c) illustrates the flexibility of our proposed GAME by allowing the largest area in the search space without breaching the R^G level.

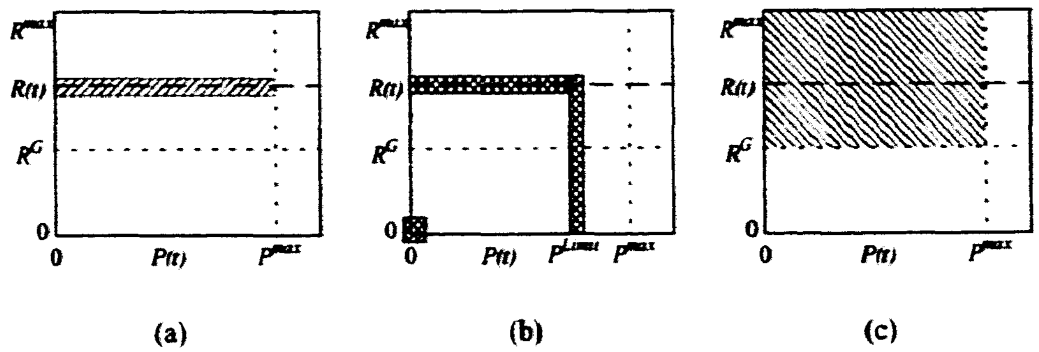

Fig. 4 *P-R* space with shaded permissible search regions.
(a) TPC [30], (b) PRA [26], and (c) GAME schemes.

3. Genetic Algorithm for Mobiles Equilibrium

In this chapter, we detail the proposed protocol. We start by describing the mobile-base station interactions followed by the genetic algorithm solution. Our analysis applies to the uplink, which we assume is orthogonal to the downlink, and can be treated independently. We concentrate on the uplink as it is generally accepted that its performance is inferior to that of the downlink [16].

3.1. Radio Interface

The main signaling messages, interchanged between MS and BS in the proposed scheme, are depicted in Fig. 5. Initially, during a call setup, MS and BS negotiate the terms of a traffic contract. This contract includes some traffic descriptors as well as some parameters representing the required QoS. P^{max}, the maximum power the MS can transmit, is acting as a constraint on the power level that can be recommended by the BS. R^P is the peak bit rate to be generated by the MS application. Connection QoS is represented by the remaining two parameters: Θ and R^G. Θ, the required received E_b/N_o, represents partly the QoS of the call since it is directly related to BER. Finally, R^G, the guaranteed bit rate represents also the QoS as seen in Table 1. It indicates the maximum bit rate that the BS ought to guarantee to the MS. Any traffic above R^G and below R^P can be accepted or rejected by GAME according to the cell load and channel conditions. In fact, R^G has also a direct relationship with the maximum tolerable delay by this type of traffic. The smaller the bearable delay is, the smaller the difference (R^P-R^G) should be.

Immediately, after the end of this call setup phase, and every *control period*, the BS triggers the GAME who tries to find the optimum power P^* and the optimum bit rate R^* for each mobile. Optimal solution is in the sense that each MS gets merely enough power, P^*, to fulfill its Θ with the maximum possible rate, R^*, where $R^* \geq R^G$. Therefore, each MS preserves its battery life and always has a guaranteed QoS. Meanwhile, the BS can admit the maximum number of calls since each one is generating the minimum interference to others.

As illustrated in Fig. 5, While GAME is solving the optimization problem, every *control clock*, the BS still has the ability to control MS power through CLPC that generates the binary ΔP based on the received E_b/N_o. If the received $E_b/N_o < \Theta$, a "ΔP=0" is generated to instruct the mobile to increase its power, otherwise, a "ΔP=1" is generated to instruct the mobile to decrease its power. Using CLPC guarantee that the BS is still in control of MS power while GAME is performing the optimization computation.

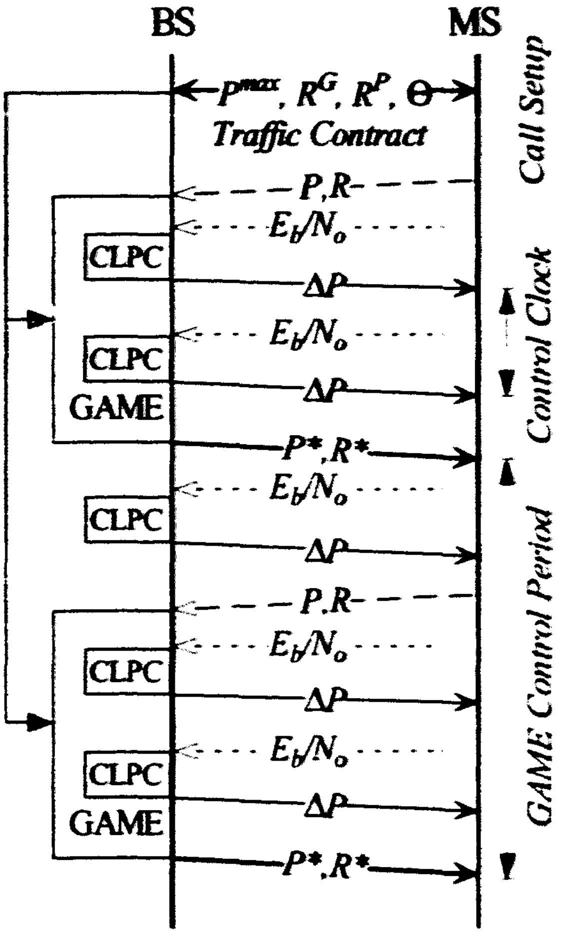

Fig. 5 BS-MS interface in case of GAME-C.

After traffic contract negotiation, GAME is launched every *control period*, measures current power *P* and rate *R*, and retrieves current traffic contract values. While optimization is being performed, the regular CLPC Δ*P* is maintained every clock. GAME returns to MS its optimal *P** and *R**.

3.2. Genetic Algorithms Overview

Genetic algorithms, as described in [7], are search algorithms based on mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search. In every generation, a new set of artificial creatures (strings) is created using bits and pieces of the fittest of the old; an occasional new part is tried for good measure. While randomized, genetic algorithms are no simple random walk. They efficiently exploit historical information to speculate on new search points with expected improved performance. Gas use random choice as a tool to guide search toward regions of the search space with likely improvement.

3.2.1. Population

GAs operate on *populations* of strings, where each string is a way of representing the variable(s) in the function to be optimized. Biologists call this objective function the *fitness* function. The string is also called *Chromosome*.

GAs work form a rich database of points simultaneously (a population of strings), climbing many peaks in parallel; thus the probability of finding a false peak is reduced over methods that go point to point. This work includes some operators like reproduction, crossover, and mutation, which are applied to successive string populations to create new strings populations.

3.2.2. Reproduction

Reproduction is a process in which individual chromosomes are copied according to their fitness function values to compose a new population. This can be implemented through a biased roulette wheel where each chromosome in the population has a roulette slot sized in proportion to its fitness.

3.2.3. Crossover

Crossover is the operator responsible of partial exchange of information between two chromosomes using cross site chosen at random as shown in Table 2.

Table 2 Single Point Crossover Operator Effect

	Before Crossover	After Crossover
Chromosome 1	☺☺☺☺☺☺☺	☺☺☺●●●●
Chromosome 2	●●●●●●●	●●●☺☺☺☺

3.2.4. Mutation

In the simple GA, mutation is the occasional (with small probability) random alteration of the value of a string position. This operator protects against losses of genetic materials that may happen in reproduction and crossover.

3.2.5. Anatomy

The mechanics of a simple genetic algorithm involve nothing more complex than copying strings and swapping partial strings as follows [17]:

1. Initialize a population of chromosomes.
2. Evaluate each chromosome in the population.
3. Create new chromosomes by mating current chromosomes; apply mutation and recombination as the parent chromosomes mate.
4. Evaluate the new chromosomes using the designated fitness function.

5. Delete old weak members of the population to make room for the new strong chromosomes.
6. If time is up, stop and return the best chromosome; otherwise go back to step 3.

3.3. Chromosome

In order to assemble a Genetic Algorithm (GA) [7], we have to define two main building blocks: *Chromosome* and *Fitness Function*, and this is actually what we discuss in this subsection and the following one.

A chromosome is the main data structure manipulated by any GA. It is simply a vector whose elements are the variables of the optimization problem, i.e., the search space components. GAME chromosome is a binary vector that encodes transmitter power, *P*, and bit rate, *R*, values of all mobile stations, *M*, in a cell. Therefore, we have initially:

$$\text{Chromosome Length} = M(N_P + N_R) \text{ bits.}$$

where N_P and N_R are the number of bits used to encode *P* and *R* respectively.

Recall that the computational complexity of solving an optimization problem increases with its search space dimensions. In other words, the time needed to reach a satisfactory solution is directly proportional to the chromosome length, which in turns depends on *M*, the total number of MSs. This is why we cluster MSs before encoding to avoid the curse of dimensionality. In order to have fair treatment, we simply group users according to their QoS class and their current traffic activity. As a consequence, we end up with only *C* clusters whose aggregate *P* and *R* are to be encoded in the chromosome. As depicted in Fig. 6, GAME chromosome length is given by:

Chromosome Length $N = C(N_P + N_R)$ bits, where $C \le M$ **(4)**

For example, if we have 300 voice users (conversational QoS class) in a cell, they can be clustered into at most two[3] groups. Talking users and silent ones. Therefore, using clustering, we can reduce the chromosome length from $300(N_P + N_R)$ to only $2(N_P + N_R)$, which clearly impacts the efficiency and responsiveness of our GAME.

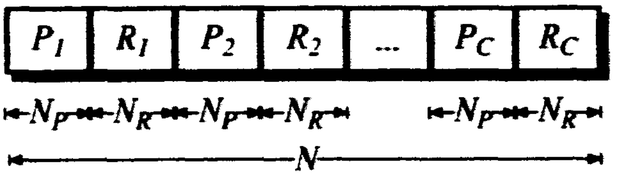

Fig. 6 GAME Chromosome Format.

N bits. Each mobile cluster occupies N_P and N_R bits to represent its power and Rate respectively. C is the total number of clusters controlled by the base station.

3.4. Fitness Function

Based on the above, we can see that the objective here is to find a nonnegative power $\underline{P}=[P_1, P_2, \ldots, P_M]$ and rate $\underline{R}=[R_1, R_2, \ldots, R_M]$ vectors within some boundaries and maximize the function F proposed initially as

$$F = \sum_{i=1}^{M} F_i^E ,$$

where F_i^E is a threshold function defined for user i as

[3] If all users are talking or all of them are silent, then we can group them into only one cluster.

$$F_i^E = \begin{cases} 1 & (E_b/N_o)_i \geq \Theta_i \\ 0 & otherwise \end{cases} \qquad (5)$$

This means to maximize the number of users that have their signal qualities above the minimum requirement, Θ. However, this objective function is incomplete since we need to give preferentiality to solutions that use less power. Hence, while limiting the *P-R* search space to solutions that maximize the number of BER satisfied mobiles through F_i^E, minimizing P is essential. Since low mobile transmitter power means little interference to others and long battery life, we proposed the power objective component, F_i^P (6), that gives credit to solutions that utilize low power and punishes others using high levels.

$$F_i^P = 1 - \frac{P_i}{P_i^{max}} \qquad (6)$$

Consequently, our main objective function F was modified to reflect this power preference as

$$F = \sum_{i=1}^{M} F_i^E + \frac{1}{M} \sum_{i=1}^{M} F_i^E F_i^P$$

The reason of multiplying F_i^P by F_i^E is to prevent those users who have failed their BER qualification from contributing to the objective score.

Another goal is to fulfill every user bandwidth request. Each call is guaranteed a specific bandwidth, R^G, according to the traffic contract. However, a user should not be prevented from getting higher rate if there is a chance. Thus, from bandwidth point of view, the R search space should avoid values below the baseline R^G while encouraging solutions to go as high as the peak bit rate R^P. This bandwidth objective is represented by

$$F_i^R = \begin{cases} (R_i - R_i^G)/(R_i^P - R_i^G) & R_i \geq R_i^G \\ 0 & otherwise \end{cases}, \qquad (7)$$

Accordingly, the final main objective function is

$$F = \sum_{i=1}^{M} F_i^E + \frac{1}{M}\sum_{i=1}^{M} F_i^E\left(F_i^P + F_i^R\right) \qquad (8)$$

Notice also that F proposed in (8) can overcome what is known as the Near-Far effect [23] that usually happens when a mobile MS1 near the BS uses power $P1$ in excess of its need and as a result prohibits another far one MS2 from reaching its Θ. Using the proposed F, this case can not happen since there exists a solution where $P1' < P1$ and hence has higher F_i^P while having the same remaining objectives scores. Therefore, GAME will prefer the solution with lower power.

In conclusion, the jointly optimal power and rate is obtained by solving the following optimization problem:

$$\max_{(P^*,R^*)\in\Omega} F(P,R) \qquad (9)$$

where F is defined in (8) and the feasible set Ω is subject to power and rate constraints

$$0 \le P_i^* \le p_i^{max},\ R_i^G \le R_i^* \le R_i^P \quad \forall i \in [1, M] \qquad (10)$$

Fig. 7 is a graphical representation of the objective function components defined in (5-7).

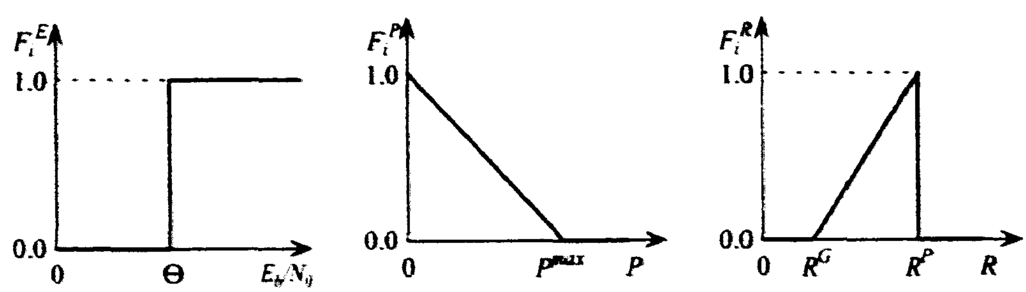

Fig. 7 Fitness Function Components.

3.5. Skeleton

The Genetic Algorithm for Mobiles Equilibrium core is a steady state GA with the ability to stop its evolution after a timeout period being expired. As illustrated in Fig. 8, the inputs are collected from the users as their current rate $\underline{R}(t)$ and power $\underline{P}(t)$ vectors. Additional information is needed as well in the input: $\underline{\Theta}$, $\underline{P}^{max}$, $\underline{R}^{G}$ and $\underline{R}^{P}$, and can be retrieved from the traffic contract established at the setup phase. The GAME starts by clustering users according to their instantaneous bandwidth and QoS requirements, and then proceeds by encoding $\underline{R}\ (t)$ and $\underline{P}\ (t)$ into chromosomes to form the initial population. The reproduction-evaluation-replacement cycle, as described in [7], evolves chromosomes into stronger ones according to fitness function F, defined in (8). Once the stopping criterion is met, the BS decodes and forwards the new $\underline{R}^*$ and $\underline{P}^*$ vectors that maximized F to the users.

There are two ways to stop the GAME progress: *Convergence* or *Time-Out*. Convergence means that the fittest chromosome, the one with highest fitness score, has the same fitness value as the average one. This indicates that GAME has saturated and no need to continue evolving. Time-Out case happens when the time slice allotted to GAME to solve the optimization problem has expired and it has to give a solution immediately. This later case usually happens when the control period is too small so GAME is triggered with high frequency.

According to the fitness function used to compare the solutions chromosomes, the fittest vectors $\underline{R}^*$ and $\underline{P}^*$ should be within the boundaries (10). In the mean time the bit rate should be as high as possible while the power level is as minimum as possible. This finest solution also should be able to make each user surmounts its required E_b/N_o value (Θ).

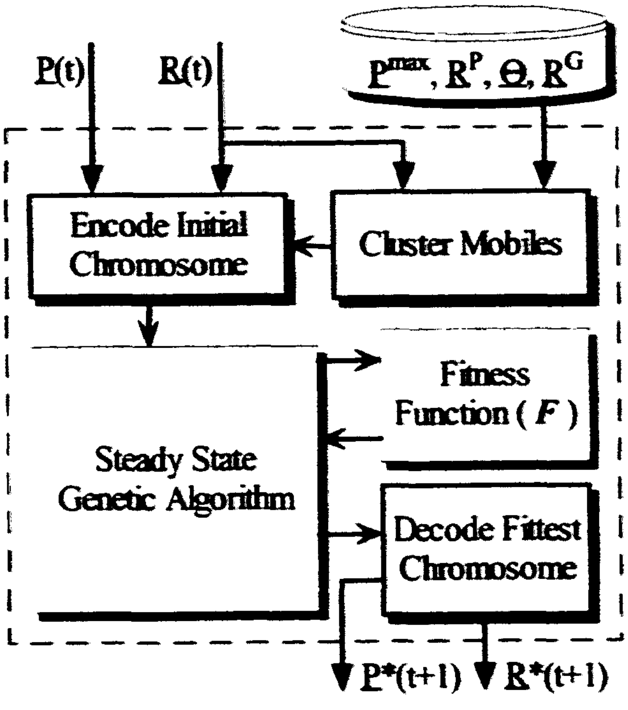

Fig. 8 GAME Skeleton Flow Diagram.

4. Multimedia Mobile Environments

Before presenting our experiments and results, we included this chapter to describe the different traffic service types as well as the cellular environments used in our simulations.

4.1. Multimedia Services

The multimedia services that have been investigated through the experiments, summarized in Table 3, covered many possible applications.

Table 3. Traffic Characteristics and QoS Requirements of Tested Services

	Voice	Video	Data
Mean Rate (Kb/s)	4.5	145	82
Peak Rate, R^P (Kb/s)	9.6	1125	144
Guaranteed Rate, R^G (Kb/s)	8.4	844	50
Required E_b/N_o, Θ (dB)	4.2	5	3.7
UMTS QoS class	Conversational	Streaming	Interactive

Voice users used in the simulations were following the On-Off model [19]. Talk-spurt and silence periods were independent and exponentially distributed with means of 1.0 and 1.35 s respectively. Talking users generated 9,600 b/s while 512 b/s were generated during silence. This traffic type fits the UMTS *conversational* QoS class. We assumed its guaranteed bit rate R^G (8.4 Kb/s) only 12.5% lower than its peak rate R^P.

MSs with video traffic represented the *streaming* QoS class. Video traffic was variable bit rate (VBR) MPEG encoded [10]. A 12 frames group of pictures (GOP) pattern (IBBPBBPBBPBB) was produced at 25 frames/s. Encoder input was 384x288 pixels with 12-bit color information. Mean bit rate was 145Kb/s while the peak rate was 1.125 Mb/s [18]. We assumed its guaranteed rate 25% away from the peak.

Data traffic generated 144 Kb/s at its peak with average of 82 Kb/s and minimum 16 Kb/s following a pareto normal distribution with cut-off [4]. This service is categorized as *interactive* QoS class, and can represent many connectionless services including web browsing. Although in [2] 3GPP has not specified a guaranteed rate for this class, we assumed a guaranteed level of 50 Kb/s (65% away from its peak) to test the efficiency of the proposed scheme in QoS provisioning.

4.2. Cellular Models

Between the antenna of the MS and the antenna of the BS, the transmitted signal undergoes a variety of effects. First, attenuation of signal strength happens due to MS-BS separation. In addition, fading occurs as a result of various signal propagation phenomena, e.g., diffraction, scattering, reflection, refraction, etc. The microcellular environment is very challenging for any RRS due to the highly dynamic path loss as described in this section.

4.2.1. Outdoor General Model

We simulated a wideband CDMA system composed of 19 hexagonal cells representing 3 inner rings in cellular coordinates. Each BS was situated at its cell center whose radius was 1 km. Mobile users were distributed uniformly over each cell space. We adopted the ITU-R distance loss model [12]. The path loss, G_{ij}, was modeled as a product of two variables

$$G_{ij} = A_{ij} \times D_{ij} \quad (11)$$

A_{ij} is the variation in the received signal due to shadow fading, and assumed to be independent and log normally distributed with a mean of 0 dB and a standard deviation of 8 dB. The variable D_{ij} is the large-scale propagation loss, which depends on the transmitter and the receiver location, and on the type of geographical environments. Let d_{ij} be the distance in km between transmitter j and receiver i, the ITU-R formula yields the following path loss equation for typical 3G CDMA system parameters [27].

$$10 \log D_{ij} = -76.82 - 43.75 \log d_{ij} \quad (12)$$

The center frequency was 1975 MHz; antenna heights of the mobile and the base station were 1.5 and 30 m respectively. We assumed that buildings occupied 20% of BS coverage area. The system spreading bandwidth W was 5 MHz and the background thermal noise density was −174 dbm/Hz. Each MS maximum transmitter power was set to 1000 mW.

4.2.2. Microcellular Pedestrian Model

Beside the general outdoor model described in the previous section, we have done experiments on a microcellular environment to test our proposed scheme over such practical urban milieu.

4.2.2.1. Path Loss Propagation Model

A microcell is a relatively small outdoor area such as a street with the BS antenna below the rooftops of the surrounding buildings. The coverage area is small and shaped by neighboring constructions. The main assumptions are relatively short radio paths (200m to 1000m), low BS antenna (3m to 10m).

In this work, we limit our study to ideal or Manhattan-like street pattern with high-rise buildings. As a consequence, there exists a clear distinction between the line-of-sight (LOS) and non-line-of-sight (NLOS) propagation. In such environment, propagation usually takes place around the sides of the buildings through reflections at the building surfaces and diffractions at building corners.

We adopt the empirical path loss model for microcells in high-rise environment developed in [7]. The model was established through measurements taken in downtown San Francisco. This urban area consists of mostly tall buildings each having more than ten stories. Table 4 summarizes the path loss formulas while Fig. 9 illustrates the considered street grid model.

Table 4 Path-Loss Formulas for High Rise Building Environment [7]

Route	Path Loss Formula	
LOS	$P_L(R_k)$=81.06 + 39.4 log f_G + 10.96 log R_k	$R_k < R_{bk}$
	$P_L(R_k)$=69.79 + 45.7 log f_G + 43.85 log (R_k / R_{bk})	$R_k > R_{bk}$
NLOS-Lateral	$P_L(R_k)$=131.19 + 12.49 log f_G + 44.86 log R_k	
NLOS-combined ST	$P_L(R_k)$= 142.37 + 29.74 log f_G + 50.37 log R_k	

Definitions of Parameters for path loss formulas, where MS and BS antenna heights (h_m and h_b) are fixed at 1.6 m and 7 m respectively:

R_k : MS-BS separation in km (0.05 < R_k < 3)

f_G : Central frequency in GHz (0.9 < f_G <2)

R_{bk} : Break point distance in km ($R_{bk} = 4\, h_b h_m/(1000\,\lambda)$)

λ : Wavelength in meters.

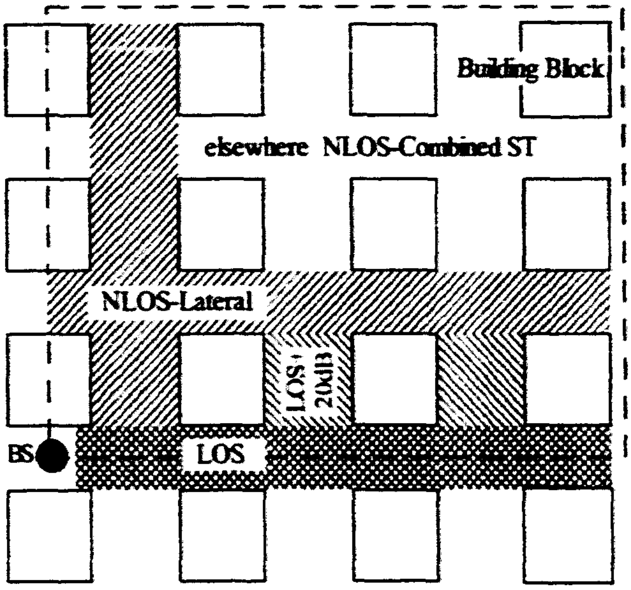

Fig. 9 Manhattan-like urban model.

Street grid illustrating deployment of different path loss formulas defined in Table 4 when BS is located between buildings [7]. Due to cell symmetry, the figure shows only the upper right quarter of BS coverage area.

4.2.2.2. Doppler Effect

In mobile radio systems, the received signal level varies with time. In part, the variation in time is due to the change in propagation loss occurring when the vehicle travels from one position to another, giving rise to relatively slow changing shadowing effects as the propagation path encounters natural and man-made obstacles. In addition, there may be a relatively fast-changing fading effect due to the interaction of multiple signal components arriving over slightly different paths that interfere with one another, sometimes constructively and sometimes destructively. The rate of variation in the fast fading effects can be related to the MS' speed, through the Doppler effect. This Doppler effect, caused by the relative motion between MS and BS antennas, is seen at the receiver as a shift in frequency

$$f_m = \frac{v}{c} f_G \tag{13}$$

where f_m is the maximum frequency shift, v is the MS speed, c is the speed of light, and f_G is the original central frequency.

Usually, the amplitudes of the multiple signal components that contribute to fast fading are subject to random variation that can be modeled by a Rayleigh distribution. The average duration of a fade below a certain level is derived in [12] for Rayleigh fading mobile radio signals and has the following value:

$$\bar{\tau} = \frac{e^{\rho^2} - 1}{f_m \rho \sqrt{2\pi}} \tag{14}$$

where ρ is the fade depth. It is defined as the ratio of the signal envelope level to the envelope rms value. f_m is the maximum frequency shift defined in (13). Moreover, the fade rate N_R has been modeled as the average number of level-crossing of the signal envelope per second [12] and calculated as:

$$N_R = f_m \rho \sqrt{2\pi} e^{\rho^2} \tag{15}$$

Giving (15) and (13) a closer look, we can deduce that the fade rate accelerates with the MS speed. For instance, at 2 GHz, an MS moving at 3 km/hr (pedestrian) suffers from 10 dB fades about 4 times each second. On the other hand, the rate increases to 133 fades/s when the MS speeds up to 100 km/hr (vehicular). In fact, these fast fades represent a great challenge to any power control algorithm. The only way to confront them is having an even faster control scheme that can adaptively adjust transmitter power accordingly.

4.2.2.3. Deployment Model

We assume an urban mobility model in a Manhattan-like structure. MSs move along streets and may turn at cross streets with a given probability. MS' position is updated every few meters and speed can be changed at each position update according to a given probability. The mobility model parameters are described in [4] and shown in Table 3. Fig. 10 illustrates the turning probability at street corners.

Table 5 Pedestrian Mobility Model Parameters [4]

Parameter	Value
Mean Speed	3 km/hr
Standard deviation for speed (normal distribution)	0.3 km/hr
Position update every	5 m
Probability to change speed at position update	0.2
Turn probability at street corners	0.5

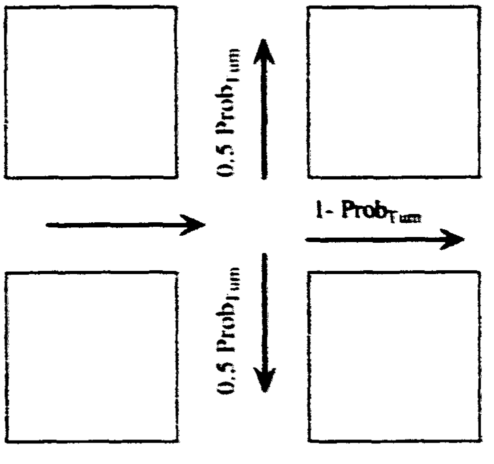

Fig. 10 Probabilities of MS movements at cross streets.

We calculated the total MS-BS path loss by combining two components. The static (propagation) component, resulting from the MS position relative to the BS, is calculated from Table 4. The dynamic (fading) component, caused by MS mobility relative to the BS, is calculated from (14) and assuming Rayleigh distributed fade depth ρ. MSs are uniformly distributed in the streets and their directions are randomly chosen at initialization.

In our study, we focus on outdoor environment. However, in order to evaluate the indoor coverage provided by an outdoor BS, an additional loss due to building penetration is to be taken into account. This penetration loss can be simulated by a normal probability distribution function having mean and standard deviation of 12 and 8 dB respectively [4].

5. Experiments

We compiled in this chapter some of the experiments done to test the proposed GAME and to compare its performance against the standard established TPC scheme. Each of the following experiments was first performed while GAME-C was in charge of resource scheduling. Then, to compare our results against the standard TPC, we repeated the experiment under the same conditions using the same seeds for random number generation but replaced GAME-C with TPC.

In the course of a ***simulation run***, 27 simulation cycles were conducted. During each cycle, the location of each mobile unit was generated randomly and its initial power was then set to a level proportional to its path loss from the nearest base station. After passing an initial stabilizing time of 300 power control clocks, statistics were collected from all MSs and averaged per service type in the following 30,000 clocks. Thus for each trial a total of 810,000 observation clocks were collected, which should be sufficient for representative simulation results. The information collected from these simulations was the average MS transmitter power, bit rate, received E_b/N_o, outage probability and the average cell loading. The outage probability of a mobile unit is defined as Prob[$E_b/N_o < \Theta$]. Cell loading is defined as the ratio of the number of active users over the maximum allowable number of users and it was approximated in [12] as

$$Cell\ Loading_b = \frac{1}{M}\sum_{i=1}^{M}\frac{G_{ib}P_i}{\sum_{j\neq i} G_{jb}P_j + \eta} \qquad \textbf{(16)}$$

This relationship illustrates the fact that the system's capacity is self-limiting, because the amount of interference is proportional to the number of users in the same or other cells. The loading is a convenient way to refer to the amount of potential capacity being used.

5.1. System Capacity

An important goal in a multiple-access system is to maximize the number of simultaneous users it can accommodate. The objective of this experiment was determining the capacity of our system, in terms of number of admissible users. We used the general outdoor model described in the previous chapter. MSs were allowed to initiate calls of all three traffic types: voice, video, and data. Notice that only the IS-95 Closed Loop Power Control (CLPC) was in effect in our comparison since we assumed that the mobiles were stationary at the time GAME is active, so there was no sudden change in link path loss, which is remedied by the open loop control (OLPC).

We started initially our first simulation run with only one voice MS and no data or video users. Gradually, we increased the number of voice users, one at a time, and repeated the simulation run. The simulation loop execution continued until an MS connection was dropped which happened when its E_b/N_o remained below its required Θ for a continuous 0.5 s. Having reached this maximum number of admissible solo voice users, M_{VOICE}, we then reset the number of voice MSs to 'zero' and repeated again the same previous simulations loop while having 'one' MS transmitting data traffic. We then increased the number of data users gradually and try all voice combinations: zero to M_{VOICE}. Finally, we increased the video users and tested all data and voice mixtures. We adjusted the GAME-C control period to 0.1 s, that means it was triggered 10 times each second.

The experiment stopped after testing all possible coexisting voice, data and video mixtures. These admissible combinations of users were determined and presented in Fig. 11 for GAME-C and in Fig. 12 for IS-95 TPC. It is evident that there are some specific limits on the maximum number of users per service that can coexist at the same time. A snapshot of these curves demonstrates that GAME-C was able to increase the maximum number of voice users (M_{VOICE}) from 320 to 375 (17% gain). The highest number of data users (M_{DATA}) has seen a boost from 30 to 47 (57% gain) while the maximum number of video users (M_{VIDEO}) inched from 13 to 14 (8% gain). As we expected, the gain in the data users was the greatest since they were the most flexible in terms of their guaranteed bit rate that reflected their high delay tolerance.

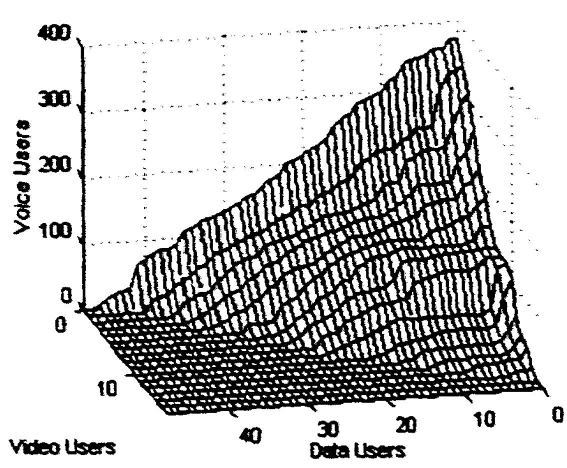

Fig. 11 Combinations of admissible mobile stations controlled by GAME-C

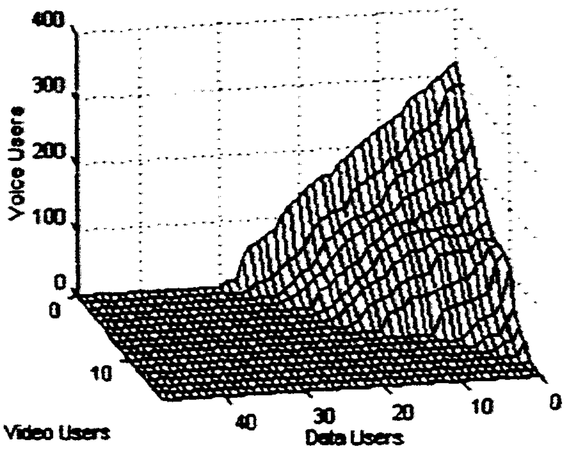

Fig. 12 Combinations of admissible mobile stations controlled by TPC

Overall, when doing the multidimensional integration to calculate the volume under each surface, CLPC attained 24707 while GAME-C achieved 41299, which is a mere profit of 67% in total admissible MSs combinations. This result realizes GAME-C objectives of increasing the system capacity and decreasing MS dropping and blocking.

5.2. Quality of Service

This experiment aimed to investigate the effect of the proposed scheme on the quality of service offered to users. We used the E_b/N_o to measure the QoS part related to BER and the average MS bit rate to assess the QoS component related to bit rate guarantee. The outage probability was also used as a user satisfaction performance measure. We tested each traffic type separately to get a clear picture of the effect on each specific class. Results on mixed traffics are reported in subsequent sections.

We started by a simulation run while having only one voice user. We then increased the number of voice users one by one and repeated the simulation run. We stopped when we reached the maximum permissible number of voice users, attained in the previous experiment. The procedure was repeated two more times separately by varying data and video users from one to the maximum tolerable number. On each simulation run, we collected some statistics including the averages plotted in Fig. 13 to Fig. 21. GAME-C control period was adjusted to 0.1 s. We used the general outdoor model described in the previous chapter. MSs were allowed to initiate calls of all three traffic types: voice, video, and data.

Unsurprisingly, outage probability increased with the growing number of MSs as shown in Fig. 13 to Fig. 15. However, each curve expressed the GAME-C superiority in users satisfaction by yielding lower outage than the standard CLPC for the same number of MSs. For instance, in Fig. 13, when the CLPC system reached its peak capacity at 320 users, the outage probability was 0.083. On the other hand, at 320 MSs, GAME-C managed to decrease this probability to 0.037 (55% gain). This enhancement persisted as well for the other two traffic types as shown in Fig. 14 and Fig. 15. When CLPC attained its maximum capacity, Data outage fell from 0.21 to 0.07 (67% gain) and video outage inched down from 0.22 to 0.20 (9% gain) by using GAME-C. Therefore, these plots verified that MSs controlled by our proposed scheme enjoyed in all cases a lower outage probability, an important measure of MS user satisfaction. Also, we confirmed the GAME-C ability to expand cell capacity and enhance BS coverage efficiency in all three types of traffic as noticed initially in the previous experiment.

E_b/N_o is an important QoS performance measure because of its relationship to BER. Fig. 16-Fig. 18 provided a proof of GAME-C QoS lead. In these figures, we plotted the average extra E_b/N_o versus the number of users. We calculated the extra E_b/N_o as the difference between the received E_b/N_o and the threshold Θ. Consistently, GAME-C supplied its users with more extra energy density than the classic CLPC scheme. It is also clear that this extra QoS faded while increasing the number of MSs since the smaller the number of users, the more the resources GAME-C had to give to MSs. It is also noteworthy to mention that on the average this extra E_b/N_o was higher than zero every time for both CLPC and GAME-C, and this demonstrates how CLPC is ably performing its job and how hard was it for GAME-C to surpass it.

Basically, there are two basic techniques to boost an MS E_b/N_o, as can be deduced from (1), increasing transmission power P, or decreasing transmitting bit rate R. Each solution has its pros and cons. The first approach, increasing power, is an easy way to enhance QoS but on the downside, it drains the mobile own battery rapidly. Besides, this action increases the interference in the face of other users, which can lead to outage and the drop of their communication links. The drawback of the second solution is that a call may be forced to cut its R to some extent. However, the main advantage is that its effect is self limited, i.e., no other users will be negatively affected but in contrast it may help them by reducing the interference. Moreover, if the reduced R is still above the guaranteed bit rate, R^G, then the QoS is still provisioned.

What GAME-C is trying to accomplish is to find a combined solution within these two extremes that get the most out of their advantages while minimizing their shortcomings. Therefore the proposed scheme increases the power marginally and decreases the bit rate slightly[4] until the resulting power is sufficient to make E_b/N_o tops Θ while introducing reasonable interference to others. We believe that this is the main cause leading to GAME-C outperforming the IS-95 CLPC since the later uses only the power as its QoS-BER provisioning tool.

We plotted in Fig. 19-Fig. 21 the average bit rate versus the number of MSs. As illustrated in the plots, using GAME-C, the rate cut increased with the number of users since more mobiles means more interference, which tweaked the solution towards the bit rate cut side. It is also clear that the amount of bit rate reduction depends on the traffic type. For instance, in case of voice traffic, GAME-C managed to keep the bit rate on the average almost the same as requested until the maximum capacity of the CPLC (320 mobiles). Afterward, GAME-C utilized its second option, which is the rate cut, aggressively until it reached its own capacity limit (375 users) where its two tools had been exhausted: the power reached a high level and the reduced rate was near the guaranteed one R^G.

In order to assess the damage in bit rate caused by GAME-C, let us compare the difference in bit rates at CLPC maximum number of possible users. At 320 users, in the case of voice traffic (Fig. 19), the average rate drop was 1.2%. Average Rate reduction was 21.2% at 30 users and 4.2% at 14 users in case of data (Fig. 20) and video (Fig. 21) services respectively. It is again clear that the rate cut in the data users case was the biggest since they had the most flexible delay constraints and the least guaranteed bit rate compared to the peak one R^P.

---

[4] GAME-C has the option to reduce bit rate as long as it is higher than the guaranteed rate R^G specified in the traffic contract.

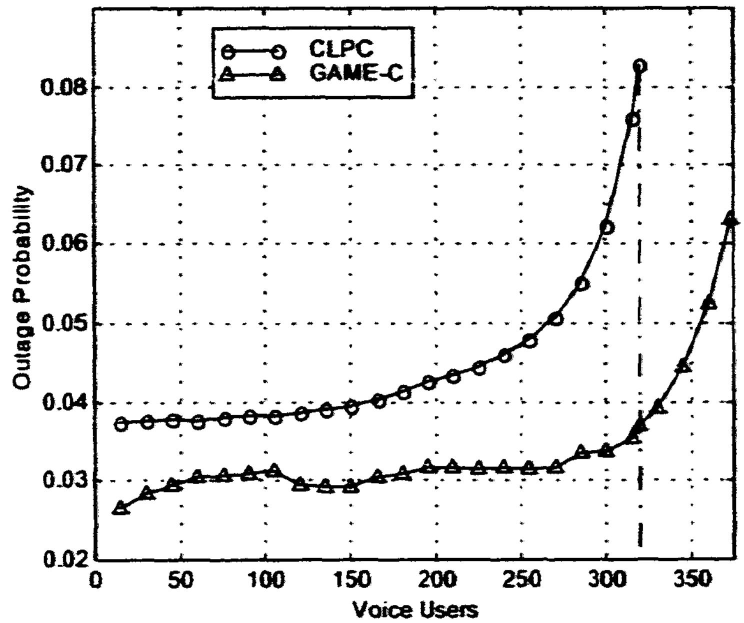

Fig. 13 Outage Probability vs. Voice Users.

'o' represents TPC. 'Δ' represents GAME-C, T =0.1 s. Video Users=0, Data Users=0.

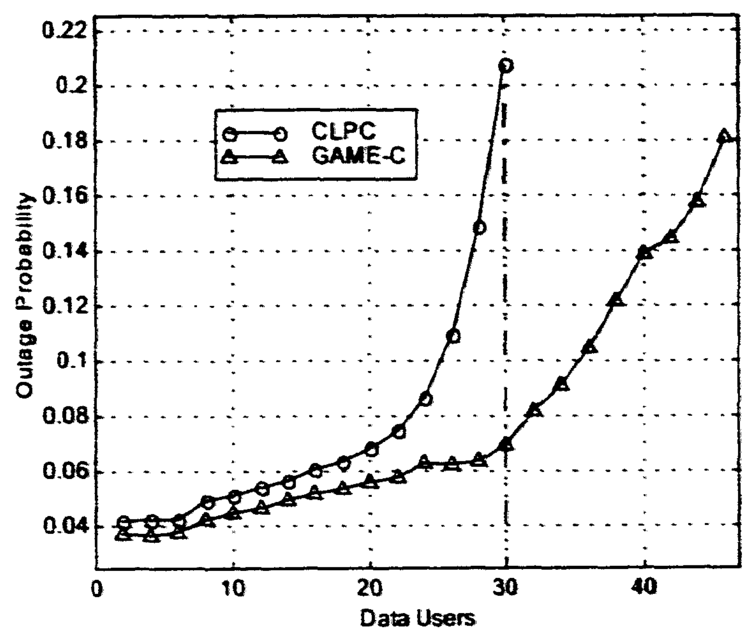

Fig. 14 Outage Probability vs. Data Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Video Users=0, Voice Users=0.

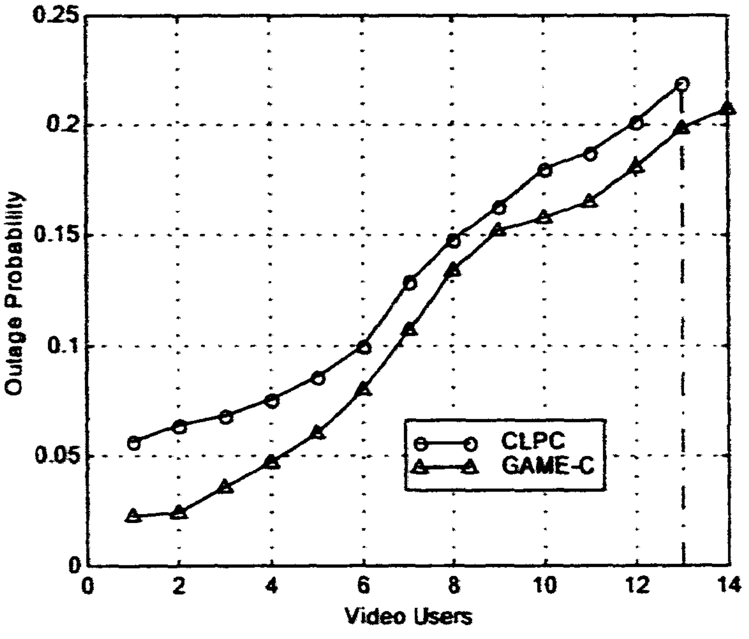

Fig. 15 Outage Probability vs. Video Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Voice Users=0, Data Users=0.

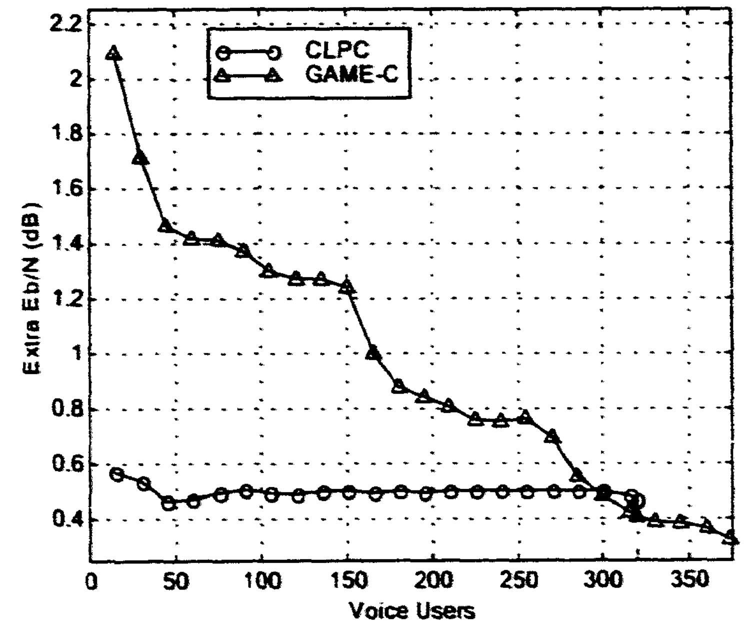

Fig. 16 Average (Eb/N-Θ) vs. Voice Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Video Users=0, Data Users=0.

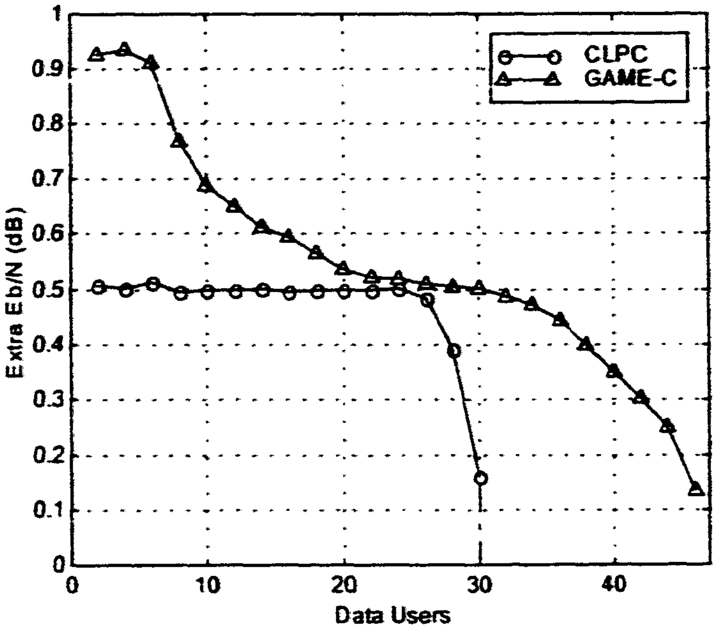

Fig. 17 Average (Eb/N-Θ) vs. Data Users.

'o' represents TPC. 'Δ' represents GAME-C, T =0.1 s. Video Users=0, Voice Users=0.

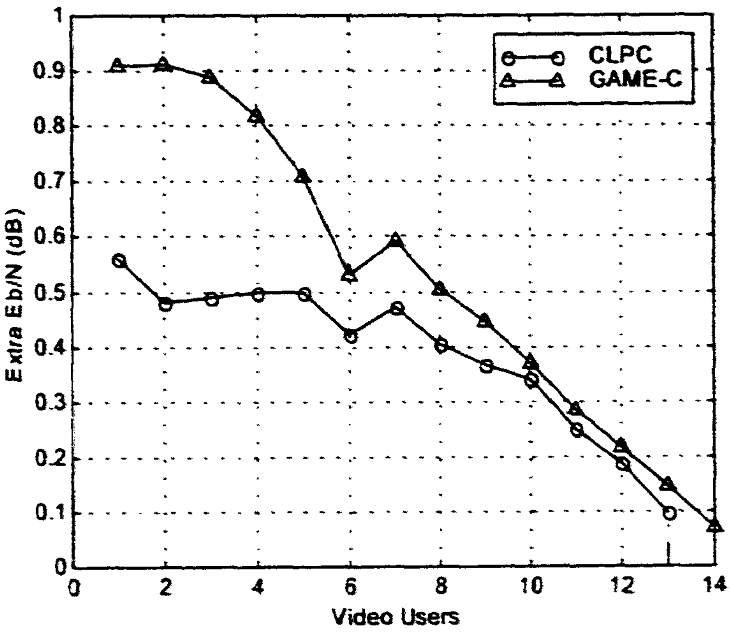

Fig. 18 Average (Eb/N-Θ) vs. Video Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Voice Users=0, Data Users=0.

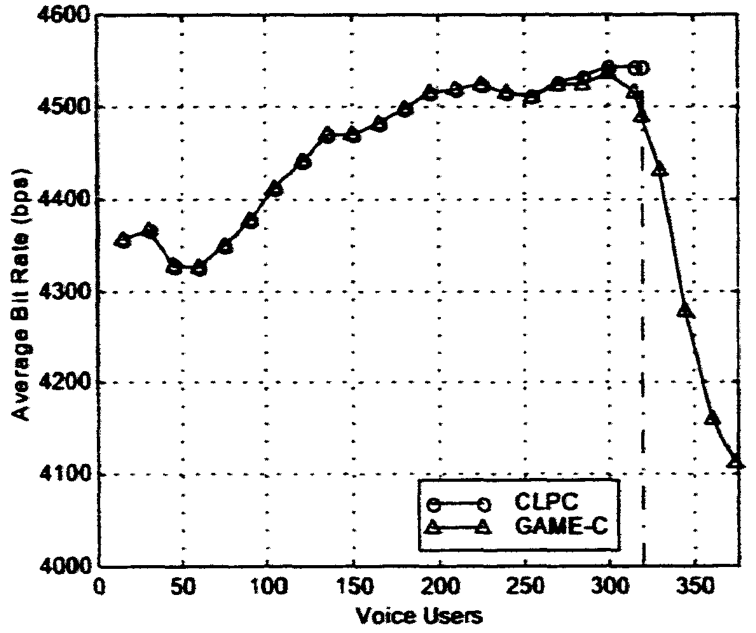

Fig. 19 Average transmitter bit Rate vs. Voice Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Video Users=0, Data Users=0.

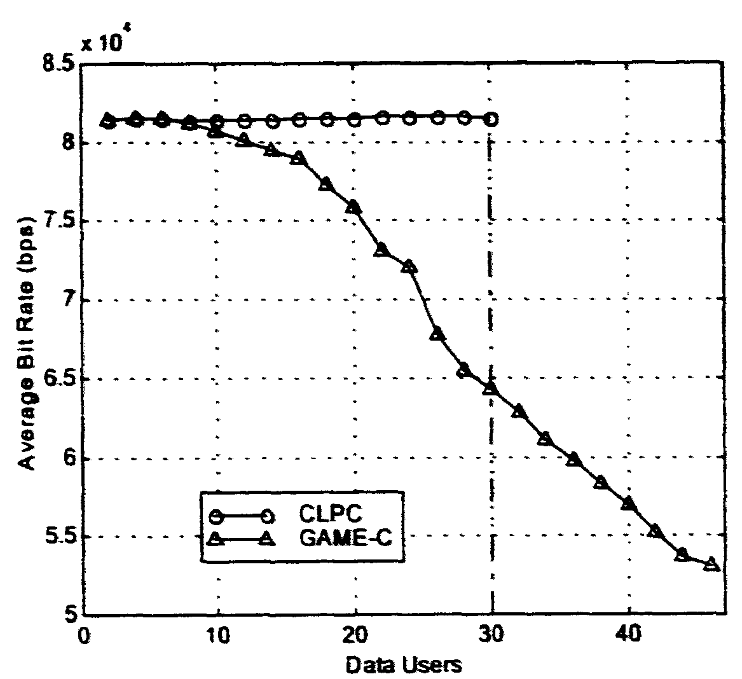

Fig. 20 Average transmitter bit Rate vs. Data Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Video Users=0, Voice Users=0.

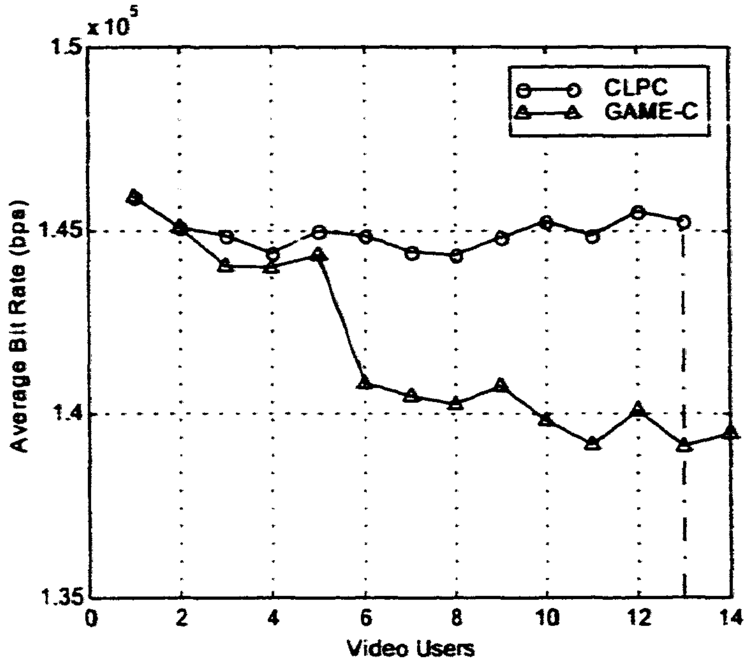

Fig. 21 Average transmitter bit Rate vs. Video Users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Voice Users=0, Data Users=0.

5.3. Power Consumption

In this experiment we aimed to study the effect of applying GAME-C on the mobile transmitter power and the cell loading. We combined different services users simultaneously to ensure the ability of the proposed algorithm to deal with the mix as good as the solo traffic type. Again, we varied the number of service users from 1 to a maximum number. This maximum number is reached once a call dropping case is attained. Each time we collected the transmitting power as well as the cell loading as defined in (16). We adjusted the GAME-C control period to 0.1 s. The experiment was repeated three times with varying mobiles number from each service type each time. The outdoor general model is used in this experiment.

As expected, GAME-C was able to reduce mobile power consumption as illustrated in Fig. 22-Fig. 24. The major reason for that savings is again the bit rate manipulation that gave the base station another degree of freedom in the restricted power allocation problem. It is obvious again, from Fig. 23, that data users were the primary beneficiaries of the power reserves. No surprise, since they were the most willing to trade their bit rate by less power and thus adding more users. If we go through the numbers at the maximum CLPC users capacity, we find the following: Power savings were 50% at 45 voice users (Fig. 22), 60% at 15 data users (Fig. 23), and 32% at 6 video users (Fig. 24). As seen, again in these plots, the average power consumed rose steadily with the growing number of users. This is natural since more users means more produced interference, which needs more power needed to overcome it.

Fig. 25-Fig. 27 depict the average cell loading versus the number of users. This is another way to confirm the power savings resulting from using GAME-C. As indicated in (16), the cell loading is a simple way to refer to the amount of potential capacity being used. That is why the average cell load increased with the number of MSs. GAME-C obviously outperformed the standard CLPC by yielding less cell loading due to power reduction. Consequently, this lower loading was translated into higher users capacity. As illustrated in Fig. 25, the maximum cell capacity in terms of users count expanded from 45 voice users, in case of CLPC, to 88 by using GAME-C. The same conclusion can be withdrawn from the other two plots (Fig. 26, Fig. 27) where data and video traffic capacities prolonged from 15 and 6 to 22 and 8 users respectively.

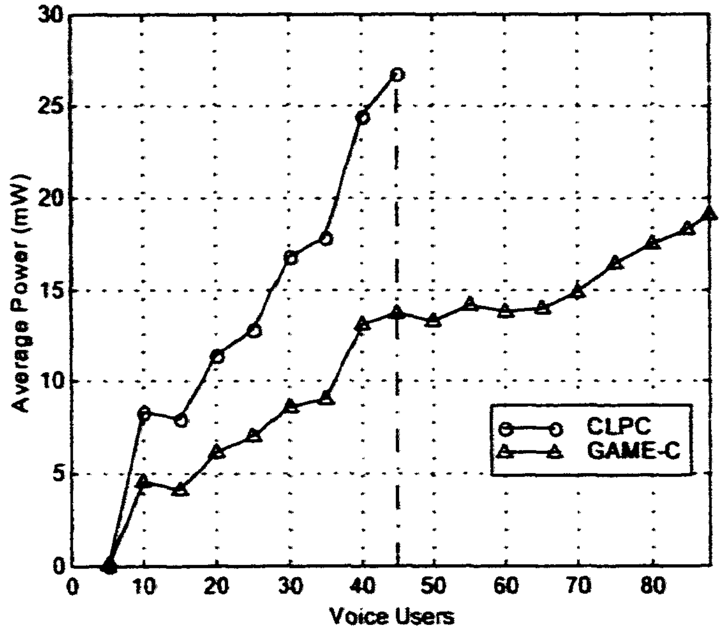

Fig. 22 Average transmitter Power vs. Voice users.

'o' represents TPC. 'Δ' represents GAME-C, *T*=0.1 s. Video Users=7, Data Users=13.

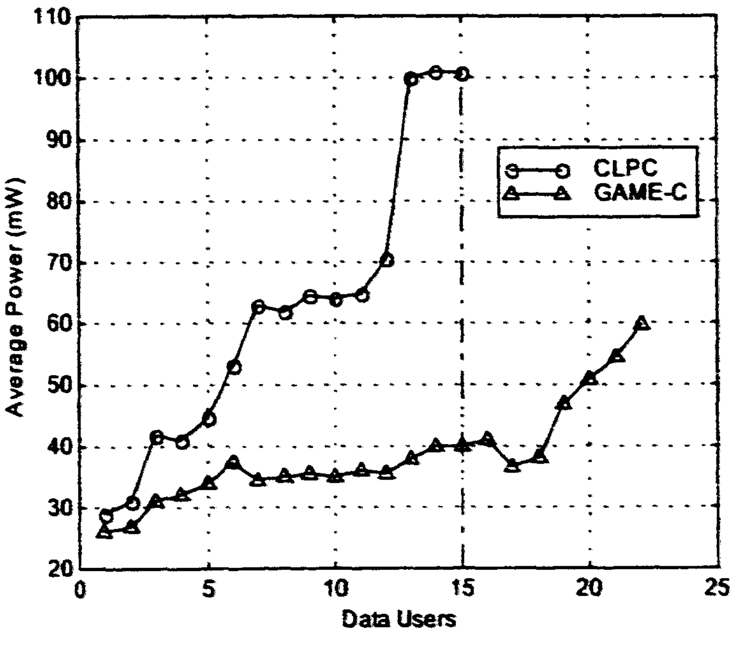

Fig. 23 Average transmitter Power vs. Data users.

'o' represents TPC. 'Δ' represents GAME-C, *T*=0.1 s. Video Users=6, Voice Users=50.

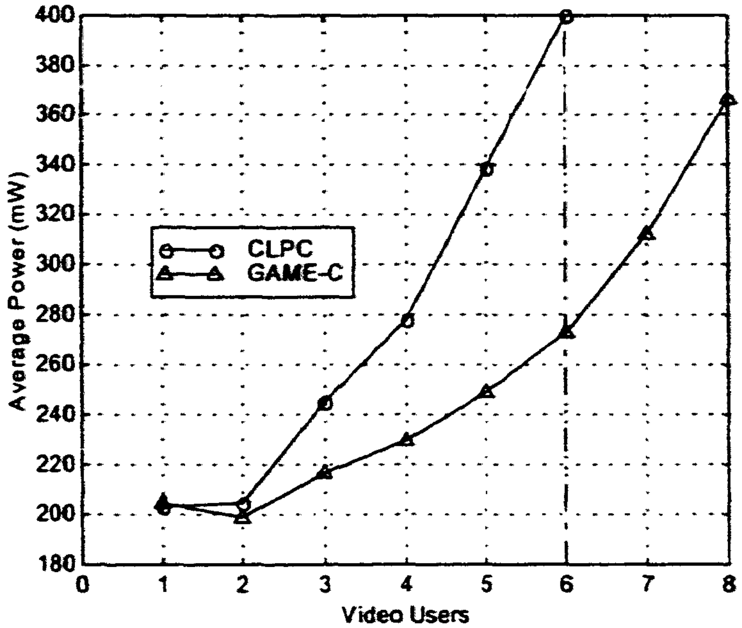

Fig. 24 Average transmitter Power vs. Video users.

'o' represents TPC. 'Δ' represents GAME-C, *T*=0.1 s. Data Users=13, Voice Users=50.

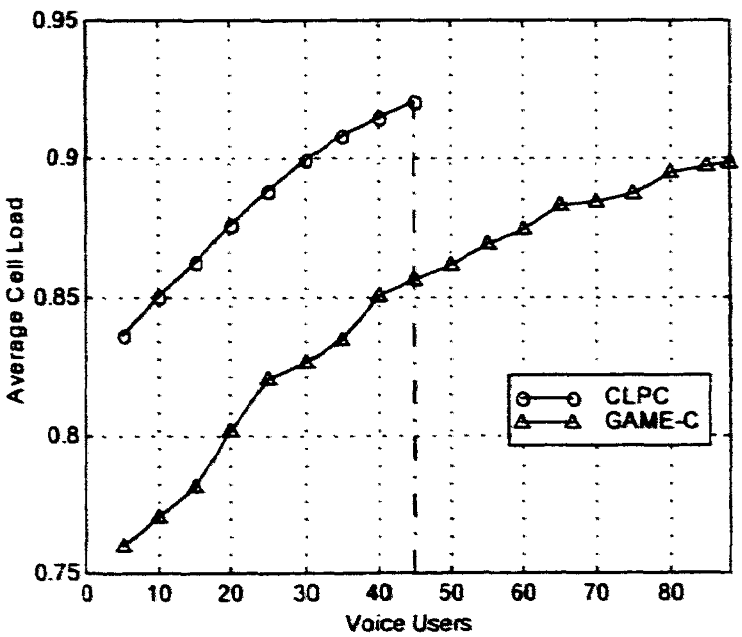

Fig. 25 Average Cell Load vs. Voice users.

'o' represents TPC. 'Δ' represents GAME-C, *T*=0.1 s. Video Users=7, Data Users=13.

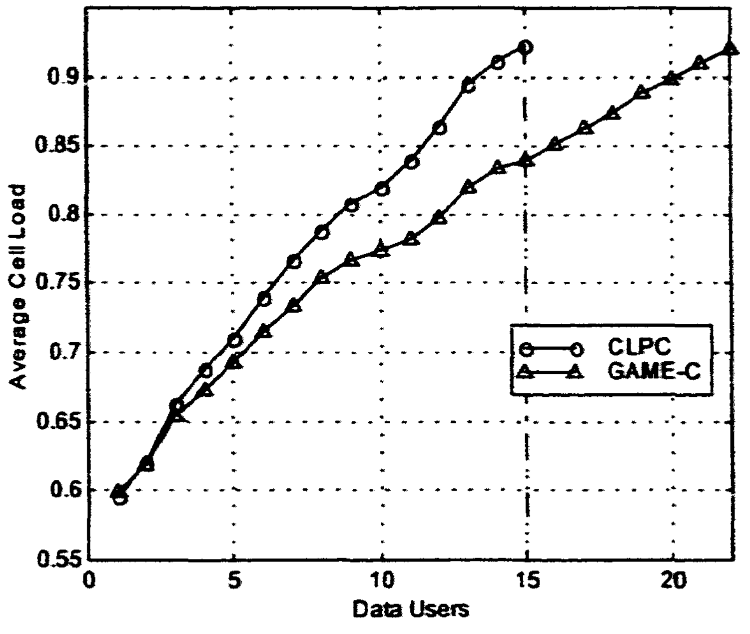

Fig. 26 Average Cell Load vs. Data users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Video Users=6, Voice Users=50.

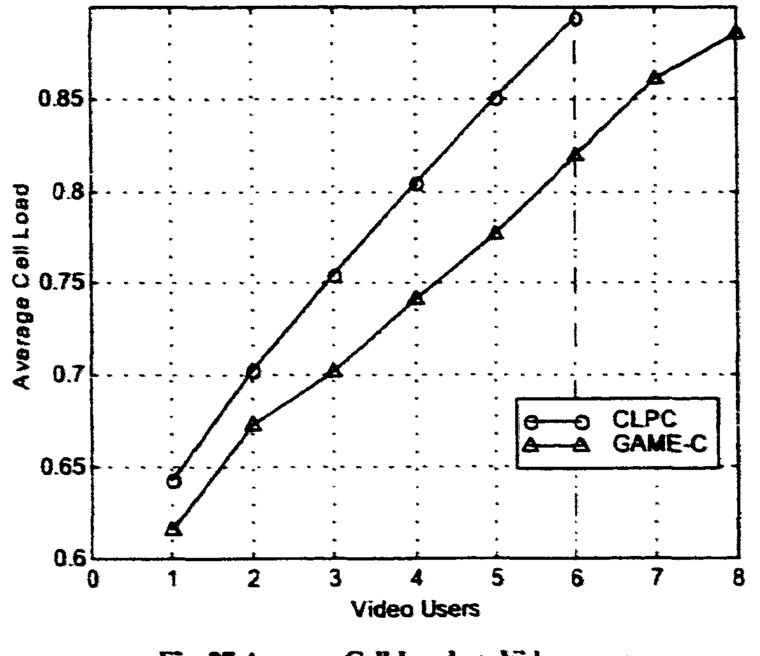

Fig. 27 Average Cell Load vs. Video users.

'o' represents TPC. 'Δ' represents GAME-C, T=0.1 s. Data Users=13, Voice Users=50.

5.4. Spectrum Efficiency

System level simulations have been performed for voice and data types of services to evaluate the Spectrum Efficiency as described in Annex B of [4]. The spectrum efficiency (ξ) is defined as the system load, measured in kb/s/cell/MHz, where there are exactly 98 % satisfied users. We consider a user as satisfied when the following conditions are fulfilled:

a) the uscr does not get blocked when arriving to the system;

b) b) the user's QoS is well preserved throughout the session, i.e., outage probability < 5% and $R > R^G$;

c) c) the user does not get dropped. A call is dropped if E_b/N_o remains below Θ for more than 5 seconds.

MSs were allowed to move in our microcellular environment according to the mobility model described in previous chapter. Periodically, we measured the system load and counted the corresponding number of satisfied users.

Fig. 28 is a plot of the experiment outcome when users were involved in voice calls, while Fig. 29 illustrates the results when MSs were holding data traffic sessions. We used up to 200 voice and 40 data users to generate wide range of system loads. As expected, GAME-C outperformed the standard TPC by yielding higher spectrum efficiency in both traffic type cases. As shown in Fig. 28 and Fig. 29, ξ increased from 117 to 145 kb/s/cell/MHz and from 224 to 387 kb/s/cell/MHz for voice and data types respectively by using GAME-C instead of TPC. These results demonstrate GAME-C competence as an RRS that utilizes the available scarce spectrum efficiently.

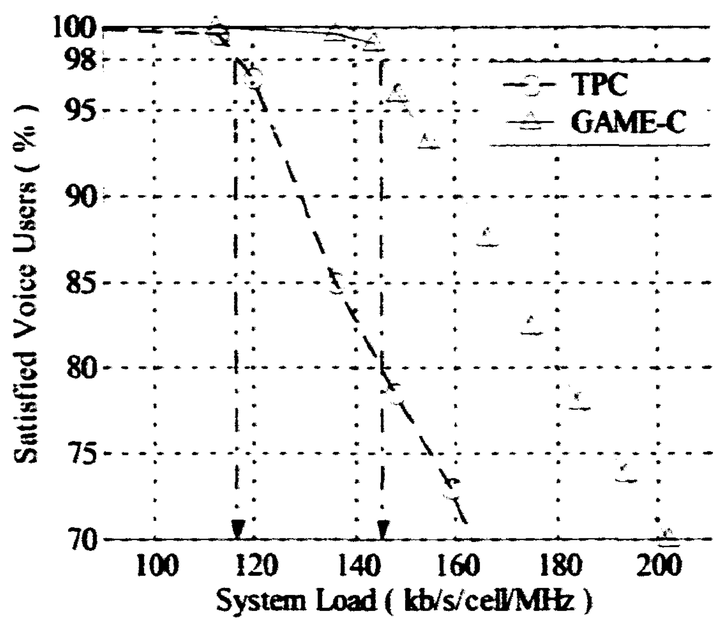

Fig. 28 Percentage of Satisfied mobile voice users versus System Load.

'Δ' represents GAME-C while 'o' represents TPC. Spectrum Efficiency is defined as the System Load at 98% satisfied users.

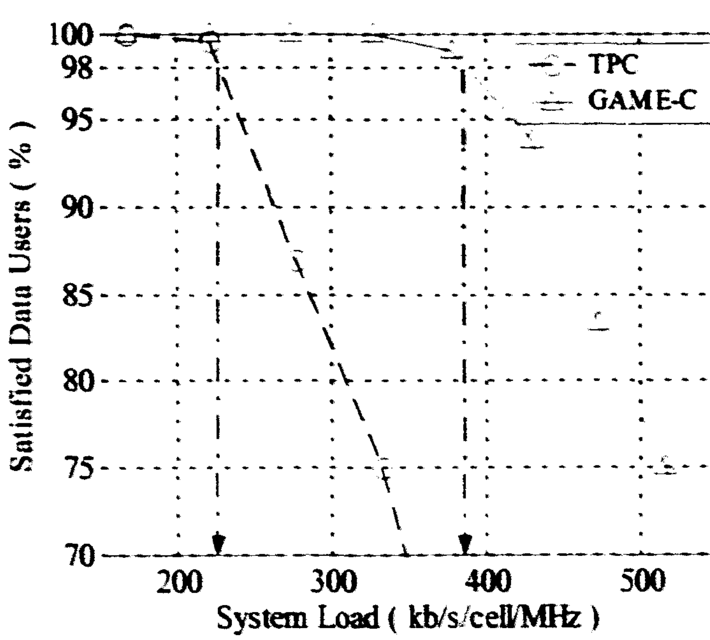

Fig. 29 Percentage of Satisfied mobile data users versus System Load.

'Δ' represents GAME-C while 'o' represents TPC. Spectrum Efficiency is defined as the System Load at 98% satisfied users.

5.5. Coverage Efficiency

To trim cost, all cellular network builders try to minimize the number of BSs required to provide adequate coverage. That is why system designers give close attention to an important performance measure, which is Coverage Efficiency (μ). Expressed in km^2/cell, μ specifies the largest cell area corresponding to the maximum tolerable outage probability by an MS. Using this figure, we can determine how many BSs per square kilometer are needed to offer MSs ample radio access services.

The objective of this experiment was determining μ of our microcellular system for both voice and data traffic types. Users were moving according to the mobility model described in previous chapter. We selected the maximum tolerable MS outage probability to be 0.05. Coverage efficiency has to be evaluated at traffic levels corresponding to low network load, since the system will be most probably interference limited at high traffic loads. We kept voice traffic density (TD) constant at 10 kb/s/km^2. Every simulation course we increase the cell area (A) in square kilometer. Next we calculate the number of MSs necessary to keep TD constant according to the following formula:

$$\text{Number of MSs} = (A \cdot TD)/\overline{R} \qquad (17)$$

where $\overline{R}$ represents each user's average transmitted bit rate in kb/s.

At the end of each simulation course, we recorded the resultant outage probability and the cell coverage area. We measured the outage probability as the ratio of the total time intervals spent by an MS having $E_b/N_o < \Theta$ over the entire simulation duration.

All MSs were moving freely according to the aforementioned mobility model. GAME-C was in charge of resource scheduling.

Fig. 30 depicts the results recorded from seven different simulation cycles using voice MSs. It is evident that outage increased with cell size due to increasing permissible BS-MS distance. A snapshot of both curves in Fig. 30 demonstrates that GAME-C was able to increase μ from 9.6 to 11.5 km²/cell in case of voice traffic.

This same experimental procedure has been repeated using data traffic MSs whose results appear in Fig. 31. Compared to standard TPC, GAME-C provided more than double coverage efficiency. μ jumped from 0.54 to 1.3 km²/cell. As we expected, the gain in the data users was the greatest since they were the most flexible in terms of their guaranteed bit rate that reflected their high delay tolerance.

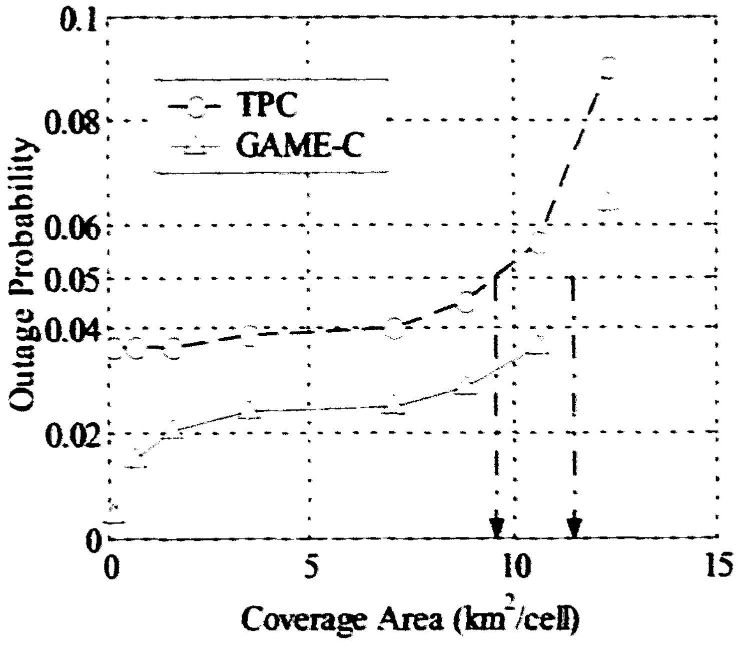

Fig. 30 Voice Outage Probability versus Base Station coverage area.
'Δ' represents GAME-C while 'o' represents TPC. Coverage Efficiency is defined as the coverage area at 5% outage.

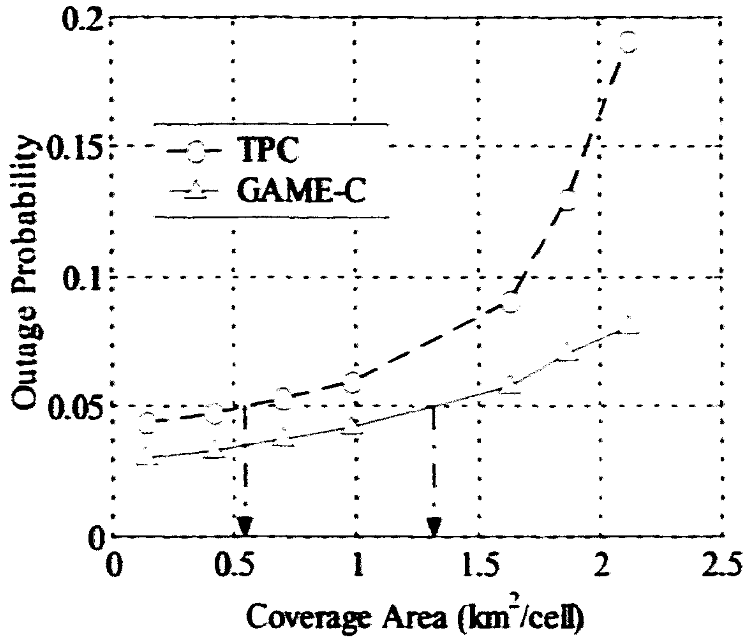

Fig. 31 Data Outage Probability versus Base Station coverage area.
'Δ' represents GAME-C while 'o' represents TPC. Coverage Efficiency is defined as the coverage area at 5% outage.

5.6. Control Period Effect

In this experiment, our objective was to study the impact of varying the control period *T* on GAME-C effectiveness in managing the wireless resources and protecting MSs QoS.

We combined several calls of different service types and kept them constant throughout the experiment. This mix was: 7 video, 12 data, and 50 voice users. We repeated the simulation while varying *T* from 0.01 s to 1 full second. At each run, we used a constant control period during the simulation. We used the outdoor general model. Again, this experiment was repeated several times to be able to catch a reasonable average statistics. The data recorded was: the outage probability, the average mobile

E_b/N_o, the average mobile transmitter power, and the average mobile transmitter bit rate. In our plots, we put the CLPC results also for comparison. Note that the CLPC values are always constant since *T* affects only GAME-C.

Fig. 32 indicates that the outage probability rose with increase in the control period. This is normal, since at small *T*, GAME was activated more frequently, and that made it able to solve any problem as soon as it appeared. On the contrary, at large *T*, a user E_b/N_o was forced to spend more time below its threshold and wait for GAME to solve it on its next launch. Fig. 33 also confirmed this conclusion, when we saw the extra E_b/N_o faded way with *T* increment. In general, the proposed scheme outperformed the standard IS-95 for all *T* values even at maximum one. At the smallest *T*, the gain of using GAME-C was at its summit. The outage probability decreased to 0.064, which is an enhancement of 51.3% and the spare E_b/N_o improved by 75.2% at 0.68dB. On the other extreme, at T=1 s, this enhancement eroded to 9.6% when the probability stood at 0.119 and the additional E_b/N_o enrichment faded to 10.3% at 0.43 dB.

Note also the behavior of GAME-C plots in Fig. 32, Fig. 33. They started with high dynamics, and then moved to a nearly saturated performance against the control period. This is normal, given the fact that as *T* tends to ∞; GAME-C tends to be pure CLPC.

We plotted also the average transmitter power and average transmitting bit rate versus the control period in Fig. 34 and Fig. 35 respectively. We reported the results of the data users only since they had the highest dynamic bit rate range and to eliminate any redundancy in our presentation.

Fig. 35 demonstrates that the bit rate recommended by GAME-C to mobiles increased with the growing control period. Consequently, the mobile transmitting power also increased with T in order to support the higher bit rate as depicted in Fig. 34. This was the reason as well for the outage probability escalation with T. At small T, GAME was constantly there to shave out excess[5] R in order to keep $E_b/N_o > \Theta$ if necessary. That is why at T=0.01 s we had the lowest bit rate, lowest outage probability, highest extra E_b/N_o and lowest power. As T moved higher, GAME launch frequency decreased. Accordingly, this setback permitted the survival of the additional rate despite QoS suffer. Therefore, at T=1 s, we had the highest bit rate, highest outage probability, lowest extra E_b/N_o and highest power.

Overall, GAME-C coped to surpass the standard CLPC for any T. At T=0.01 s, the power savings was 65.2% at 43.7 mW while the price, which is the rate cut, was the highest: 34.7% at 51.7 Kbps. On the other end, at T=1 s, the power savings was the lowest: 49.8% at 69.2 mW while the rate cut was 20.5% at 63 kbps.

[5] Excess bit rate is the rate above the guaranteed one R^G.

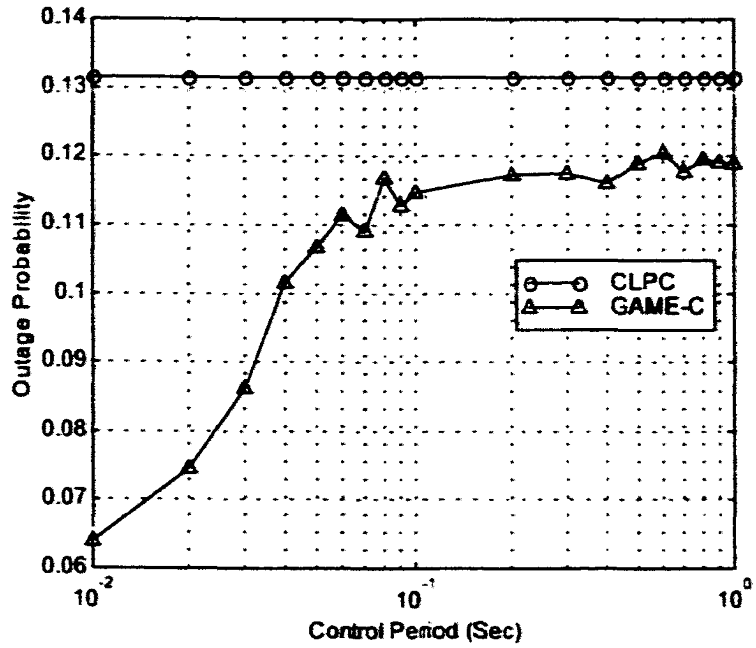

Fig. 32 Outage Probability vs. Control period.

'o' represents TPC. 'Δ' represents GAME-C. Video Users=7, Data Users=12, and Voice Users=50.

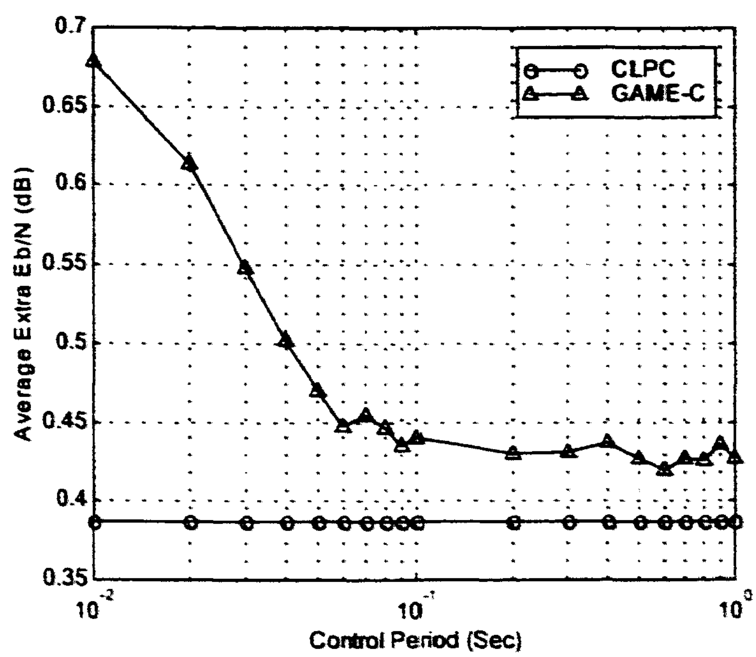

Fig. 33 Average (Eb/N-Θ) vs. Control period.

'o' represents TPC. 'Δ' represents GAME-C. Video Users=7, Data Users=12, and Voice Users=50.

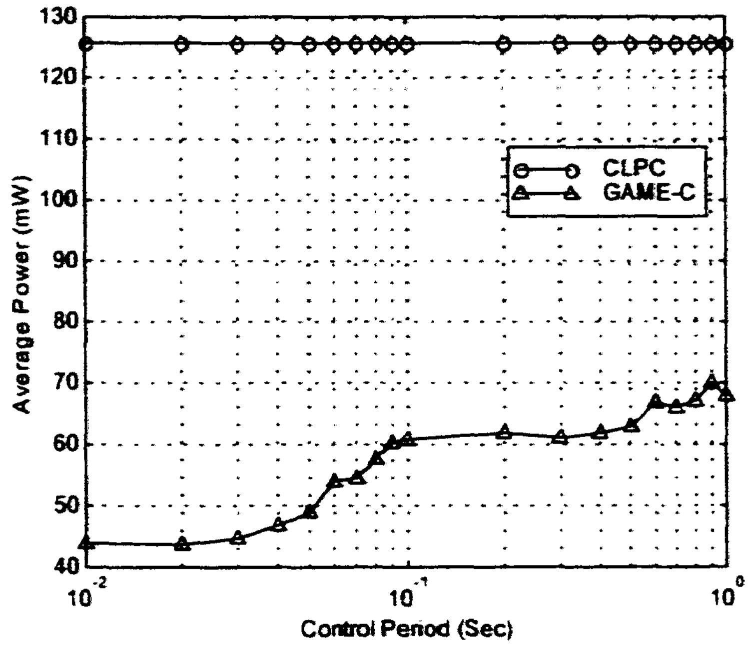

Fig. 34 Average transmitter Power of Data Users vs. Control period.

'o' represents TPC. 'Δ' represents GAME-C. Video Users=7, Data Users=12, and Voice Users=50.

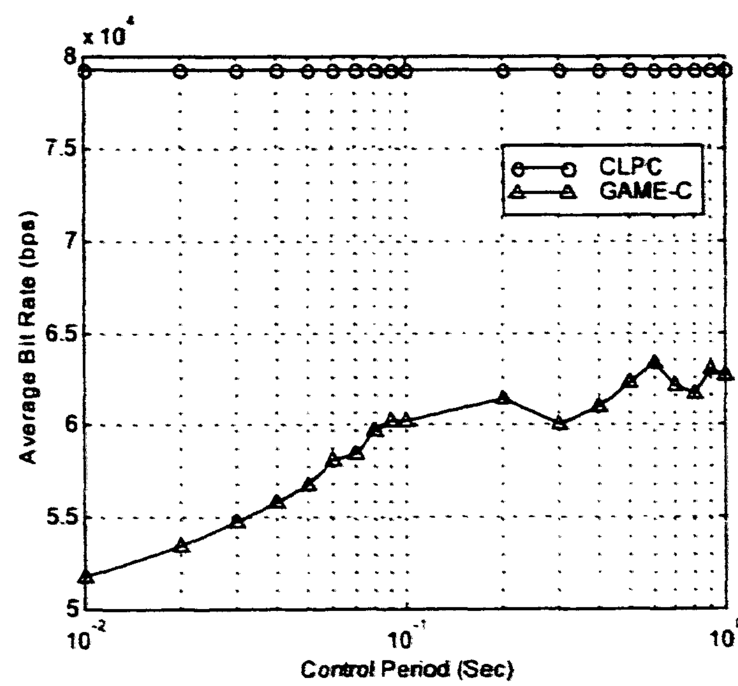

Fig. 35 Average transmitter bit Rate of Data Users vs. Control period.

'o' represents TPC. 'Δ' represents GAME-C. Video Users=7, Data Users=12, and Voice Users=50.

5.7. Convergence and Scalability

Giving our fitness function (8) a closer look, the reader can easily deduce that '0' bounds the fitness value as a lower limit (inclusive) and 'M+1' acts as an upper limit (not inclusive). This is based on the fact that all individual components: F_i^E, F_i^P, and F_i^R yield real numbers between '0.0' and '1.0'. Also, note the fact that the fitness score is a real number. The integer part is coming from the first term summation in (8) which indicates the potential number of mobile stations exceeding their E_b/N_o requirements. The mantissa part indicates the strength of this solution in meeting our different objectives and comes from the remaining terms in (8). For instance, having 30 mobile users, fitness score of '23.7' means only 23 users can meet their Θs with 70% strength. In this example, the *optimal solution* score is '30.99...9', which means all 30 users can beat the quality objective with full strength. It is noteworthy to mention that sometimes the theoretical maximum fitness value 'M+0.999...9' cannot be reached in practice since it can correspond to *P-R* values outside their constraint ranges (10).

In order to better estimate the time complexity, let us break up the genetic operations into simple computation ones. One chromosome operation (COP) includes two parent selections, one crossover, one mutation and one fitness evaluation. Obviously, the COP complexity depends on the chromosome length which in turn, as seen in Fig. 6, is a function of the number of mobile clusters C. Based on our implementation, this chromosome operation needs '17*C+21' floating point operations (FLOP). Translating this to absolute time is a function of the used hardware architecture (microprocessor, system bus, etc...). For instance, on a typical machine with 100 MFLOP/s (million FLOP per second), absolute time for one COP can be approximated in nanoseconds by

$$\text{Time needed for one COP}\,(ns) = 170 \cdot C + 210 \quad \textbf{(18)}$$

In this experiment, we investigated the time complexity imposed by GAME to reach *satisfactory* solution for the optimization problem presented in (9). Fig. 36 is a plot of the fitness score versus the average number chromosome operations required. To produce that plot, we used 99 on-off voice users. GAME control period T was constant at '0.1' s. It is evident that giving the solver more time enhances the solution quality. However, we can see that this enhancement rate degrades with time. In this experiment, GAME needed on the average 20 chromosome operations to surpass the fitness value of 99 that declares the E_b/N_o satisfaction of all MSs. Voice users were clustered as *talking* or *silent* with the same E_b/N_o requirements. Therefore, 20 COPs for C=2 needed about 11 microseconds according to (13).

Another experiment has been done to investigate GAME scalability. The objective of this experiment was to test the GAME performance against the increase in MSs number. Actually, in reporting the results, we used the number of clusters C instead of the number of MSs M since it is C that decides the chromosome length and the time for one COP. We used a mix of mobile stations in this simulation generating video, voice and data traffic to increase the range of C as much as possible. We forced GAME to stop once the fittest chromosome exceeded 95% of the fitness function maximum value. We varied the number of mobile stations from 10 to 450 and recorded the average number of COP versus the number of mobile clusters. The plot in Fig. 37 depicts the GAME scalability and how the number of users is affecting its performance. We found that GAME coped with the maximum number of mobile stations in about 2010 COP, which can be translated using (18) as 7.25 ms for 20 clusters.

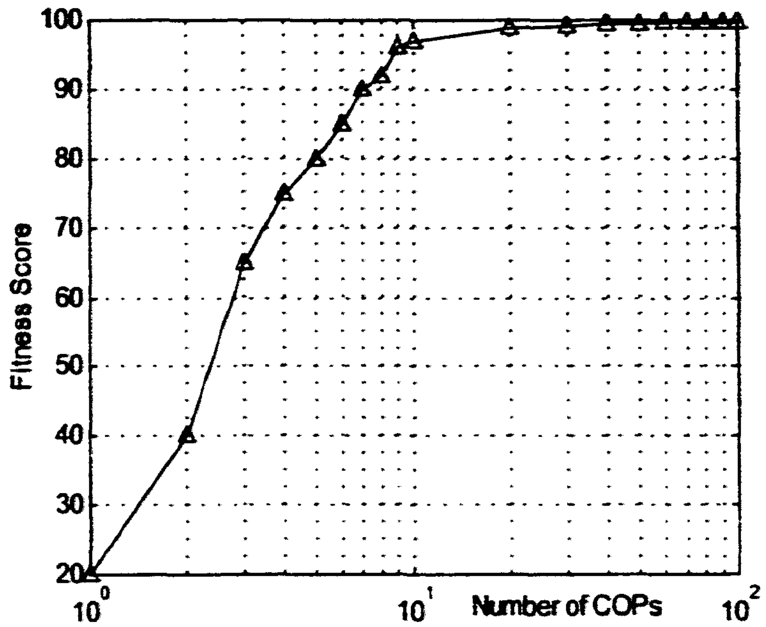

Fig. 36 GAME Convergence.

Fitness Score vs. the Average Number of Chromosome Operations (COPs).

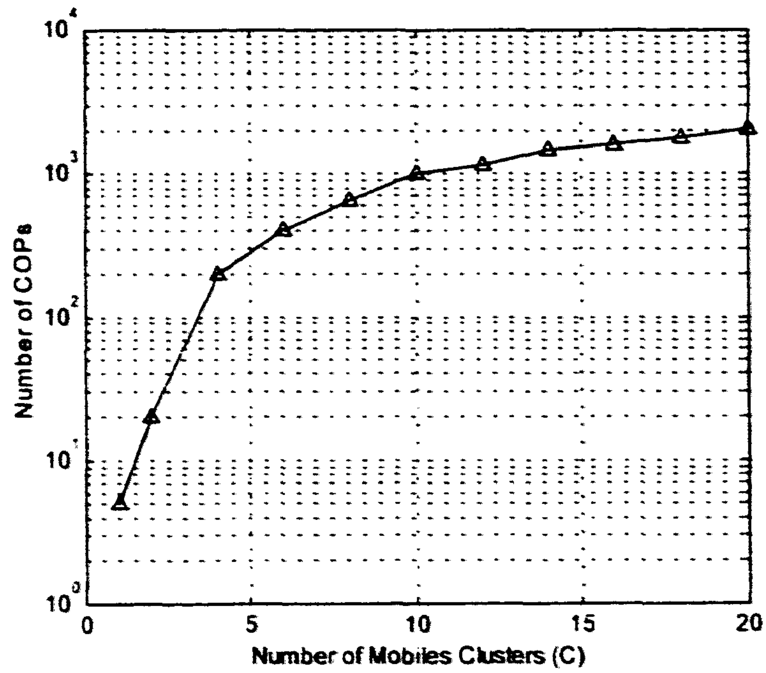

Fig. 37 GAME Scalability.

Average number of Chromosome Operations (COPs) vs. the Number of Mobile Clusters *C*. 450 users were used to generate 20 mobile clusters.

6. Putting It Together

In this chapter we try to address some integration and implementation issues. The first section describes how GAME can take advantage of current standard radio interface to exchange control signaling between MS and BS. Then, we discuss how to realize end-to-end QoS provision between wireless and wireline QoS models.

6.1. GAME Signaling on Standard Interface

According to 3GPP [3], The standard UMTS MS-RAN radio interface is layered into three protocol layers: physical (L1), data link (L2), and network (L3) layers.

The physical layer offers information transfer services to higher layers, including:

(i) Forward Error Correction (FEC) encoding/decoding and interleaving/de-interleaving of transport channels.

(ii) Modulation and spreading/demodulation and de-spreading of physical channels.

(iii) Measurements and reporting to higher layers (e.g. BER, SIR, interference power, transmit power, etc; and (iv) Closed-loop power control (CLPC).

As illustrated in Fig. 38-(a), in the Control-plane (C-plane), Layer 2 is split into two sub layers: Medium Access Control (MAC), Radio Link Control (RLC). Layer 3 is partitioned into sub-layers where the lowest sub-layer, denoted as Radio Resource Control (RRC), interfaces with layer 2 and terminates in the RAN.

Upon RRC request, MAC executes radio resource reallocation and reports measurements such as traffic volume and quality indication to RRC. MAC is responsible also for priority handling between MSs by means of dynamic scheduling.

RLC provides error correction in Protocol Data Units (PDU) by retransmission (e.g. Selective Repeat, Go Back N, or a Stop-and-Wait ARQ) and ciphering to prevent unauthorized acquisition of data. Besides, It is in charge of detection of duplicated received RLC PDUs to ensure that the resultant higher Layer PDUs are delivered only once to the upper layer and in the same order as submitted for transfer. RLC can also suspend/resume data transfer based on an RRC command.

The Radio Resource Control (RRC) layer handles the control plane signaling of Layer 3 between MSs and RAN. It performs several functions:

(i) Broadcast of system information to all MSs.

(ii) Establishment, maintenance and release of an RRC connection between MS and RAN including admission control.

(iii) Assignment of radio resources (e.g. codes, bandwidth) needed for the RRC connection while ensuring that the QoS requested for the Radio Bearers can be met.

(iv) Connection mobility functions such as handover based on measurements done by the MS.

(v) MS measurement reporting (what to measure) and control of the reporting (when to measure and how to report); and (vi) Outer loop power control by setting of the target Θ of the Closed Loop Power Control CLPC.

Naturally, RRC layer implements most of RRM techniques and specifically RRS functionalities at BS since it controls and signals the allocation of radio resources to MS. Fig. 38-(b) illustrates the signaling messages interchanged between BS-MS on the C-plane to execute the radio resources scheduling (RRS) task.

MS-BS Signaling

Initially, MS' L1 and MAC layers perform some measurements (e.g., transmitter power P, bit rate R, interference power) and report these values to MS' RRC layer. Next, MS composes a measurement report and forwards it to the BS RRC, which is in need for this information to determine which radio resources are available. Subsequently, RRS algorithm residing in BS, like GAME, determines the minimum power P^* and bit rate R^* necessary for each MS to satisfy its QoS requirements. Finally, MS' RRC layer receives those values (P^*, R^*) and in turn passes them on to appropriate layers (L1 and MAC) to be in effect. Optionally, MS' RRC can send a control message to the application coder to adjust the data encoding bit rate as well, provided that the coder has multirate capability. For example, in case of streaming video data, the coder can drop some intermediate frames or increase its quantization step when advised by RRC to reduce outgoing bit rate.

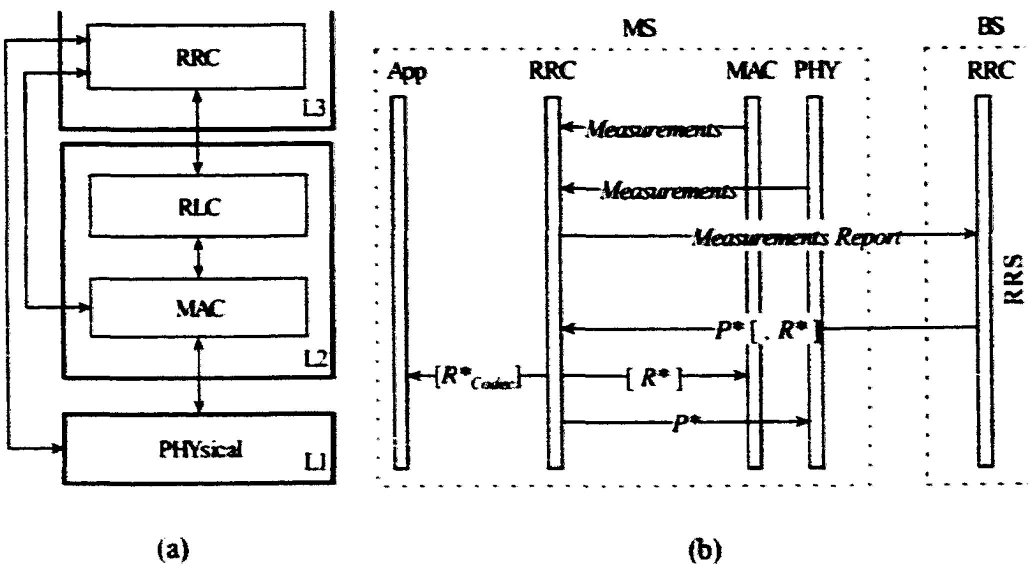

Fig. 38 Mobile Station-Base Station Standard Interface.

(a) Control-plane Layered Radio Interface Protocol Architecture. (b) Resource Control Signaling.

6.2. End-To-End QoS Provisioning

Since next generation mobile telecommunication systems are not devoted to either voice or best effort services, it is important to address Quality of Service (QoS) provisioning, which is essential for the success of packet data services, especially in a bandwidth-constrained and error prone environment such as cellular networks.

6.2.1. UMTS QoS Architecture

Naturally, Network Services are considered end-to-end, i.e., from terminal equipment (TE) to another TE. User's TE such as Personal Digital Assistant (PDA), notebook PC, or digital phone is connected to the UMTS network through MS equipment such as wireless modem. As illustrated in Fig. 39, the *Local bearer* service handles QoS between TE and MS through a QoS-capable Application Programming Interface (API). A *Bearer Service* defines all aspects between communicating end points to enable the provision of a contracted QoS. These aspects include control signaling, user plane transport and QoS management functionality. *UMTS bearer* service, offered by the UMTS operator, is in charge of QoS management inside UMTS network, i.e., from MS to CN. This is actually where the UMTS QoS model, described hereafter, is in effect. Finally, QoS support outside the UMTS network is provided by the *External bearer* service, e.g., IETF-defined Integrated Services (IntServ) and Differentiated Services (DiffServ), or simply the best effort service.

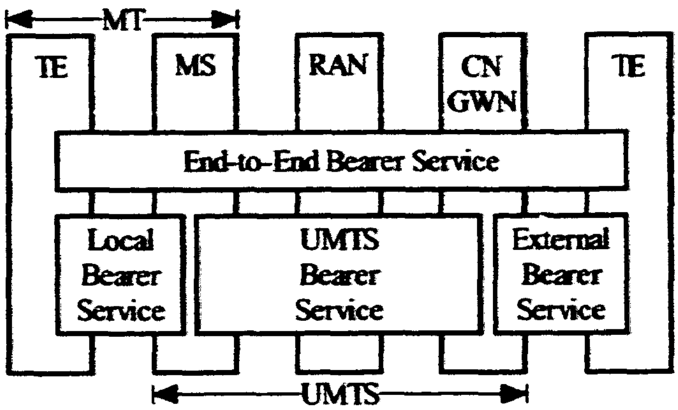

Fig. 39 End-to-End QoS Architecture.

6.2.2. Internet Protocol QoS Models

The UMTS QoS model, which includes QoS classes and attributes, is effective only within the UMTS bearer service shown in Fig. 39 and is designed to be independent of external QoS mechanisms. On the other hand, outside UMTS network, for instance, Internet applications use IntServ (signaled by RSVP) [13] or DiffServ (6-bit QoS attribute on each IP packet) [21] QoS models.

Practically, a TE application usually uses a suitable API for requesting QoS. As an example, the Generic QoS (GQoS) Winsock 2 API, defined by Microsoft, provides a QoS interface that invokes the Resource ReSerVation Protocol (RSVP) to signal resource allocation to a network.

RSVP, an IETF Layer 3 end-to-end signaling protocol standard, simply expresses application resource requirements to the network nodes on the path between a sender and its receiver(s). RSVP is receiver-driven, meaning that the receivers (not the senders) make the resource reservation. Specifically, a sender issues RSVP PATH messages to its receivers. Then, the receivers reply by RSVP RESV messages along the reverse path to request for resource reservation. Although GQOS and RSVP, are supposed to be protocol independent, the reality is that they assume the IP protocol. The relevant GQoS parameters here are:

- *Level of Service Quality*: guaranteed, controlled load, and best effort.
- Source Traffic Specifications: *Token Bucket Model* parameters (token rate in Bytes/s and token bucket size in Bytes) used to specify the rate at which permission to send traffic (or credits) accrues. The *Peak Bandwidth* (Bytes/s) is also specified to limit how fast packets may be sent back to back from the application.
- *Latency*: upper limits on the amount of delay and delay variation (both in microseconds).

6.2.3. QoS Interworking and Parameter Mapping

Obviously, to have an actual end-to-end QoS support, some mapping has to be done to ensure QoS inter-working between existing defined schemes and the UMTS model. Fortunately, this mapping overhead is needed only on boundary elements, i.e., MS and CN' GWN. Hence, the RAN and CN' other components do not have to understand external QoS mechanisms and consequently, they do not have to be upgraded frequently. Specifically, on the uplink (traffic sent by MS towards UMTS), the MS should be able to represent the external QoS requirements in a form suitable for the UMTS QoS model while GWN translates internal UMTS QoS parameters to the external network. MS and GWN mapping roles are reversed on the downlink (traffic sent to MS).

As suggested in [20], GQoS-*Level of Service Quality* maps fairly well to the UMTS-*QoS class* definition, and GQoS-*Peak Bandwidth* maps to the UMTS-*Maximum Bit Rate* attribute. UMTS-*Guaranteed Bit Rate* integrated with traffic shaping at network boundaries can support the GQoS-*Token Bucket Model*. GQoS-*Latency* maps to the UMTS-*Maximum Delay*. However, UMTS does not currently recognize delay variation at all due to rapid fluctuations in radio channel conditions.

Thus, the uplink flow starts when the TE' application uses GQoS parameters along with a QoS signaling protocol, such as RSVP through the Local bearer service. Next, the UMTS layer in the MS maps the specified parameter values in the signaling protocol to those understood within the UMTS bearer service. On the other end, GWN translates UMTS attributes, through the external bearer service, to GQoS native terminology. As a result, an end-to-end QoS support between TE and external networks is maintained.

Sometimes QoS mapping is not needed between the RAN and the Internet. This special has been investigated through the Broadband Radio Access Network for IP-Based Networks (BRAIN) project [29]. The BRAIN architecture includes a RAN, which is fully IP-based thus eliminating the need for a mediator like CN in the UMTS to connect with the Internet. In this model, RAN uses internally the same QoS techniques as the outside world. Hence, the QoS internal and external bearer services can be merged into one entity.

7. Conclusions

In this study, we introduced a novel scheme for wireless resource management in CDMA networks and applied it to wideband multimedia environment with dissimilar QoS requirements. Our proposed method, GAME-C, integrates the Genetic Algorithm for Mobiles Equilibrium (GAME) technique with the regular Closed Loop Power Control (CLPC) specified in the standards IS-95, cdma2000, and UMTS. While scheduling radio resources, GAME-C manipulates transmitting bit rate beside the traditional power control. The main algorithm is to be implemented in the base station that forwards the controlling signals to its mobile stations. Actually, GAME uses the bit rate as an additional tool to solve situations where only power control failed because of high interference. Therefore, it trades bit rate for higher users density, less connections dropping/blocking, and lower transmitting power. This rate reduction is subject also to a maximum, so the resulting allotted bandwidth is always above the guaranteed level specified in the traffic contract. Hence, QoS requirements are always met or exceeded.

The advantages of using genetic algorithms for optimization are numerous. Parallelism, GAME can be implemented as multiple synchronized threads to take advantage of the full processing power of the used hardware. Evolving nature, GAME can be stopped any moment while having the assurance that the current solution is better than all the previous ones. Scalability, mobiles can be added or removed simply by adjusting the chromosome length and leaving everything else intact.

A Small GAME control period *T* is better than a large one, since it gives GAME-C a chance to watch the situation closely. However, due to some hardware restrictions, it might be impractical for an MS to change its bit rate according to GAME recommendation every few ms. On the other hand, power can be controlled more frequently. Thus, we recommend adjusting GAME control period to a conservative T of 0.05 to 0.1 s. Therefore, GAME revises bit rate every 50 to 100 ms while CLPC updates power every 1.25ms[6]. Overall, GAME-C was able to outperform the standard CLPC with a T=0.01 s, and even with one full second control period as noticed through the experiments.

The proposed scheme performed acceptably during the experiments done to test it. The enhancements over the standard CLPC case are substantial. The volume of admissible users within a cell has seen an average expansion of 45%. The outage probability has decreased by an average of 40% with better corresponding signal quality (E_b/N_o). In the mean time, the average power consumption has been saved by 46%. Moreover, the Base station coverage area expanded on the average by 70%, demonstrating the GAME-C ability in reducing the BSs density without deteriorating coverage quality requirements. Certainly, this efficiency will be translated into lower construction cost of network infrastructure. Meanwhile, the spectrum efficiency has been stretched on the average by 48% implying the increase of QoS-satisfied users and therefore realizing the intended objectives. To pay for these enhancements, the transmitter bit rate declined on the average by 11% but without breaching the guaranteed bit rate specified in QoS requirements.

Current research work is in progress to apply the GAME on the downlink (BS to MS). We plan also to investigate some hardware implementation issues.

---

[6] IS-95 and cdma2000 use 1.25 ms for CLPC while UMTS uses 0.625 ms

References

1. 3GPP, "IP Multimedia (IM) Subsystem-Stage 2," TS 23.228 version 5.1.0, www.3gpp.org, June 2001.

2. 3GPP, "Universal Mobile Telecommunication System (UMTS); QoS Concept and Architecture," TS 23.107 version 3.5.0, www.3gpp.org, 1999.

3. 3GPP, "Universal Mobile Telecommunication System (UMTS); Radio Interface Protocol Architecture," TS 35.301 version 4.0.0, www.3gpp.org, 2001.

4. 3GPP, "Universal Mobile Telecommunication System (UMTS); Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS," ETSI TR 101 112 version 3.2.0, www.3gpp.org, 1998.

5. 3GPP2, "Introduction to cdma2000 Standards for Spread Spectrum Systems," 3GPP2 C.S0001-A, www.3gpp2.org, June 2000.

6. ETSI, "The ETSI UMTS Terrestrial Radio Access (UTRA) ITU-R RTT Candidate Submission," www.etsi.org, June 1998.

7. D. Har, H. Xia, and H. Bertoni, "Path-Loss prediction model for microcells," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1453-1462, September 1999.

8. D. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", *Addison-Wesley*, 1989.

9. D. Kim, "Rate-Regulated Power Control for Supporting Flexible Transmission in Future CDMA Mobile Networks", *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 5, pp. 968-977, 1999.

10. D. Le Gall, "MPEG: A Video Compression Standard for Multimedia Applications," *Communications of the ACM*, pp. 47-58, April 1991.

11. G. Bianchi, N. Blefari-Melazzi, "A Migration Path for the Internet: from Best-Effort to a QoS Capable Infrastructure by means of Localized Admission Control," *Proc. of International Workshop on QoS in Multiservice IP Networks QoS-IP 2001*, Lecture Notes on Computer Science, Springer-Verlag, vol. 1989, January, 2001.

12. J. Lee and L. Miller, "CDMA Systems Engineering Handbook", *Artech House Publishers*, Maryland, 1998.

13. J. Wroclawski, "The use of RSVP with IETF integrated services," *RFC 2210*, www.ietf.org/rfc/rfc2210.txt, September 1997.

14. J. Zander, "Radio Resource Management in Future Wireless Networks: Requirements and Limitations," *IEEE communications*, 1998.

15. J. Zander, "Performance of Optimum Transmitter Power Control in Cellular Radio Systems," *IEEE Trans. Veh. Technol.*, vol. 41, no. 1, pp. 57-62, 1992.

16. K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303-312, May 1991.

17. L. Davis, editor, Handbook of Genetic Algorithms, Van Nostrand Reinhold, 115 fifth avenue, New York 10003, 1991.

18. O. Rose, "Statistical Properties of MPEG Video Traffic and Their Impact on Traffic Modeling in ATM Systems", *Inst. Of Comp. Science, University of Wurzburg, Germany*. Research Report No 101.

19. P. Brady, "A Statistical Analysis of On-Off Patterns in 16 conversations," *Bell Systems Tech. Journal*, vol. 47, no.1, pp. 73-91, 1968.

20. R. Koodli, and M. Puuskari, "Supporting Packet-Data QoS in Next-Generation Cellular Networks," *IEEE communications*, February 2001.

21. S. Blake et al., "An Architecture for Differentiated Services," *RFC 2475*, www.ietf.org/rfc/rfc2475.txt, December 1998.

22. S. Oh, T. Olsen, and K. Wasserman, "Distributed Power Control and Spreading Gain Allocation in CDMA Data Networks," *Proc. of IEEE INFOCOM conf.*, 2000.

23. S. Rappaport, "Wireless Communications," *Prentice Hall*, New Jersey, 1996.

24. S. Recker et al., "Wireless Internet Architecture and Testbed for WINEGLASS," *Proc. 1st Mobile Commun. Summit*, October 2000.

25. S. Ulkus and R. Yates, "Stochastic Power Control for Cellular Radio Systems", *IEEE Trans. Commun.*, vol. 46, no. 6, pp. 784-798, 1998.

26. S. W. Kim, and Y. H. Lee, "Combined Rate and Power Adaptation in DS/CDMA Communications over Nakagami Fading Channels," *IEEE Trans. Commun.*, vol. 48, no. 1, January 2000.

27. Special Issue. "IMT-2000: Standards Efforts of the ITU", *IEEE Personal Communications*, vol. 4, August 1997.

28. T. Hu and M. Liu, "A New Power Control Function for Multirate DS-CDMA Systems", *IEEE Trans. Commun*, vol. 47, no. 6, pp. 896-904, 1999.

29. T. Robles et al., "QoS Support for an All-IP System Beyond 3G," *IEEE communications*, August 2001.

30. TIA/EIA/IS-95, "Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular Systems," *Telecommun. Industry Assoc.*, 1995.

31. TIA, "The cdma2000 ITU-R RTT Candidate Submission," www.tiaonline.org, June 1998.

List of Publications

Journal Papers

1. M. Moustafa, H. Habib, and M. Naghshineh, "Wireless Resource Management Using Genetic Algorithm for Mobiles Equilibrium," *Computer Networks, The International Journal of Computer and Telecommunications Networking*, Elsevier Science, "In Press".

2. M. Moustafa, I. Habib, M. Naghshineh, and M. Guizani "QoS-Sensitive Broadband Mobile Access to Wireline Networks," *IEEE Communications Magazine*, "Accepted 9/2001".

3. M. Moustafa, I. Habib, and M. Naghshineh, "GAME Based Dynamic Resource Scheduling in QoS Aware Radio Access Networks," *IEEE/ACM Transactions on Networking*, "Submitted 7/2001".

4. M. Moustafa, I. Habib, and M. Naghshineh, "Efficient Radio Resource Control in CDMA Microcells," *IEEE Transactions on Wireless Communications*, "Submitted 9/2001".

IETF RFCs/Internet Drafts

Moustafa, Habib, and Naghshineh, "QoS-Aware Radio Resource Control using the Genetic Algorithm for Mobiles Equilibrium," draft-moustafa-GAME-qos-resource-control-00.txt, "work in progress".

Conference Papers

1. M. Moustafa, H. Habib, and M. Naghshineh, "Genetic Algorithm for Mobiles Equilibrium," *Proc. IEEE Conference on Military Communications*, MILCOM2000, 10/2000.

2. M. Moustafa, H. Habib, and M. Naghshineh, "Genetic Algorithm for Mobiles Equilibrium Applied to Video Traffic," Proc. QoS-IP 2001, *Lectures Notes on Computer Science*, vol. 1989, Springer-Verlag, pp. 381—390, 1/2001.

3. M. Moustafa, H. Habib, and M. Naghshineh, "Wireless Resource Management using Genetic Algorithm for Mobiles Equilibrium," *Proc. sixth IEEE Symposium on Computers and Communications*, ISCC 2001, 7/2001.

4. M. Moustafa, H. Habib, and M. Naghshineh, "GAME Based QoS Provisioning in Multimedia Wideband CDMA Networks," Proc. IWQoS 2001, *Lectures Notes on Computer Science*, vol. 1989, Springer-Verlag, 6/2001.

5. M. Moustafa, H. Habib, and M. Naghshineh, "Efficient Resource Management in Microcellular CDMA Systems," *Proc. IEEE Vehicular Technology Conference*, VTC Fall'01, 10/2001.

6. M. Moustafa, H. Habib, and M. Naghshineh, "GAME Based Radio Resource Management in Wideband CDMA Networks," *Proc. IEEE Global Telecommunications Conference*, GLOBECOM, 11/2001.

7. M. Moustafa, H. Habib, and M. Naghshineh, "Efficient Radio Resource Control for Manhattan Street Environments," *IEEE International Conference on Communications*, ICC 2002, "Submitted 8/2001".