QOS ANALYSIS OF A UMTS CELL WITH DIFFERENT SERVICE CLASSES

Paolo Lollini
Dipartimento di Sistemi e Informatica
Università degli Studi di Firenze
viale Morgagni 65,
50134, Firenze, Italy
email: lollini@dsi.unifi.it

Andrea Bondavalli
Dipartimento di Sistemi e Informatica
Università degli Studi di Firenze
viale Morgagni 65,
50134, Firenze, Italy
email: bondavalli@unifi.it

Felicita Di Giandomenico
ISTI Institute
Italian National Research Council
via Moruzzi 1,
56124, Pisa, Italy
email: digiandomenico@isti.cnr.it

ABSTRACT
The objective of this work is to evaluate the Quality of Service (QoS) of a Universal Mobile Telecommunication System (UMTS) cell. Among UMTS characteristics, we focus the attention on two aspects strictly related with the identified QoS measures: the random-access procedure and the admission control strategy. Using a model-based approach, we model and evaluate a UMTS cell under several service classes, thus providing useful insights on the perception of the QoS both from the users'perspective and from the mobile telephone operators'one.

KEY WORDS
UMTS - Random Access Procedure - Admission Control - Stochastic Activity Network - Simulation - QoS Analysis.

1 Introduction

The telecommunication industry is continuously investing on innovative and competitive wireless services. High QoS is an always increasing requirement which absorbs significant effort by system suppliers. Quantitative assessment of the degree of QoS offered by a network is therefore of primarily importance. Recent works ([1, 2]) analyze some QoS measures (mainly related to service availability) in a General Packet Radio Service (GPRS) environment. In this paper, we concentrate on the Universal Mobile Telecommunications Systems (UMTS), which is the third generation mobile communication system standardized by 3GPP.

Though the evaluation of the QoS of a UMTS network in a single service scenario is a widely discussed topic in the literature, only few works aim at analyzing the UMTS system in presence of different traffic scenarios. [3] evaluates the performance of Code-Division Multiple-Access (CDMA) schemes with integrated services, namely real-time voice services and non real-time data services. [4] focuses the attention on multimedia services (real-time and non real-time) and it proposes an adaptive resource allocation system to provide appropriate QoS according to service requests from end users. In [5] the performance of a mixed traffic scenario on the downlink of a CDMA is evaluated, taking into account speech traffic and web-browsing traffic sharing the same frequency carrier.

This paper contributes to the analysis on the service accomplishment level perceived by UMTS users. The focus is on those UMTS characteristics which mainly impact on the analyzed QoS measures, namely the random-access procedure and the admission control strategy, and it accounts for different classes of users behavior, mainly distinguished by their throughput, the workload they add on the cell, the duration of the requested service, the duration of the idle period and the set of sub-channels available for preamble transmission. Following a modular approach, we build the UMTS cell model as composition of more simple submodels representing different system characteristics, and we solve it using a simulation method.

The rest of this paper is organized as follows. Section 2 presents the system under analysis and the identified QoS indicators. Section 3 contains an overview of the modeling technique, the list of assumptions used in the modeling definition and the description of the submodel representing the random-access procedure. In Section 4 the numerical results of the simulation studies are presented and discussed, while conclusions are drawn in Section 5.

2 The investigated system and the QoS indicators

In this paper we focus the attention on two UMTS characteristics having important effects on the QoS, both from the users'perspective and from the mobile telephone operators'one: the random-access procedure and the admission control strategy. These mechanisms are needed to decide whether a new service request can start based on the available network "capacity". Actually they mainly influence the so called "connection-level" QoS ([4]), that are the quality indicators related to the connectivity properties of the network, like the call blocking or dropping probability.

When a user needs a service from the network, its User Equipment (UE) sends a channel request to the network through the Physical Random Access CHannel (PRACH), a specific channel dedicated to the uplink transmission of channel request. Since the network does not control the PRACH usage, the access method, based on a random-access procedure, may cause collisions among requests by different UEs, thus worsening the expected QoS. Once the network receives the channel request, it performs

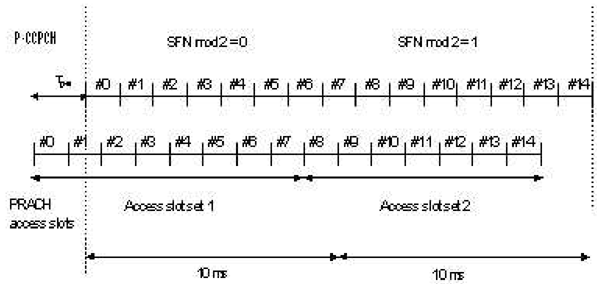


Figure 1. PRACH structure

	Sub Channel Number											
SFN Mod 8	0	1	2	3	4	5	6	7	8	9	10	11
0	0	1	2	3	4	5	6	7				
1	12	13	14						8	9	10	11
2				0	1	2	3	4	5	6	7	
3	9	10	11	12	13	14						8
4	6	7					0	1	2	3	4	5
5			8	9	10	11	12	13	14			
6	3	4	5	6	7					0	1	2
7						8	9	10	11	12	13	14

Table 1. Available uplink access slots for different sub-channels

the admission control procedure to decide if a traffic channel can be allocated to this new request. The goal is, in general, to ensure that the interference created after adding a new call does not exceed a pre-specified threshold, thus preventing the QoS to degrade below a certain level. In this section we give an overview of these two mechanisms detailing only those aspects that will be used in the modeling phase.

The random-access transmission in the 3GPP specifications [6, 7] is based on a slotted ALOHA approach. It consists of several consecutive preamble transmissions with increasing power at well-defined time-offsets, called access slots. The preambles are sent through the PRACH depicted in Figure 1. It consists of 15 access slots divided in two frames: the first frame is synchronized with the even frames of the Primary Common Control Physical CHannel (SFN mod 2 = 0, where SFN is the System Frame Number) and it consists of the slots from 0 to 7; the second frame is synchronized with the odd frames of the PCCPCH (SFN mod 2 = 1) and it consists of the slots from 8 to 14. As shown in Table 1, the access slots are further divided into 12 sub-channels, and each sub-channel defines a sub-set of the total set of uplink access slots that can be used for preamble transmission, on the basis of the SFN.

Before starting the random-access procedure, the UE receives some information from the Broadcast CHannel (BCH) like available signatures, available slots in the available sub-channels, maximum allowed preamble retransmissions (M_{max}), ramping factor, persistency value and initial preamble power. Then the UE randomly selects one of the available sub-channels for transmission ($SelSubCh$), an access slot in that sub-channel, a preamble spreading code from the set of available codes and it starts a preamble transmission using the selected uplink slot, preamble signature and preamble transmission power. When the network receives the preamble, it replies with a positive acquisition indicator (Ack) if the admission control procedure has been passed, otherwise with a negative acquisition indicator (Nack). The network does not send any reply (NoAck) if it does not receive the preamble (e.g. in case of not enough preamble transmission power) or in case of preamble collisions (more preamble transmissions over the same time slot). In case of Nack or NoAck signals, the UE selects the next available access slot in the set of available PRACH sub-channels, it randomly selects a new signature, it increases the preamble transmission power by the ramping factor obtained from BCH and, if the number of transmission attempts is still less than the maximum allowed, it starts a new preamble transmission. In case no more transmission attempts are available, the random-access procedure is terminated (*random-access procedure failure*). If the UE receives an Ack signal, it can finally start the transmission of the message using the allocated traffic channel.

Call Admission Control (CAC) is one of the important means to guarantee Quality of Service (QoS) in the telecommunication systems. When a UE asks for a service, the CAC algorithm decides whether to accept or reject this new request. There are several types of admission control algorithms studied in the literature, each one having different properties and aiming at optimizing different network parameters (e.g. [8, 9]). [10] presents a multi-cell CAC algorithm based on soft handover: the algorithm works on balancing the load over the network by transferring some connections from the overloaded cells to the neighboring cell with lighter load, thus achieving a more balanced resource utilization over the whole network. In [11] a CAC strategy founded on fuzzy logic is proposed, where the rules for the admission criterion are based on cell parameters like congestion state, available load and total interference.

Here we consider an admission control algorithm based on the workload of the UMTS cell: a new call is accepted if the workload level reached after adding the call does not exceed a pre-specified threshold, both in uplink and in downlink. Equivalently:

$$\eta_{ul} + \Delta L_{ul} < \eta_{ul\_threshold} \quad , \tag{1}$$

$$\eta_{dl} + \Delta L_{dl} < \eta_{dl\_threshold} \quad , \tag{2}$$

where η_{ul}, ΔL_{ul} and $\eta_{ul\_threshold}$ (or η_{dl}, ΔL_{dl} and $\eta_{dl\_threshold}$) are, respectively, the cell workload before the admission of the new call, the workload increment due to the admission of the new call and the pre-specified threshold level in uplink (or downlink).

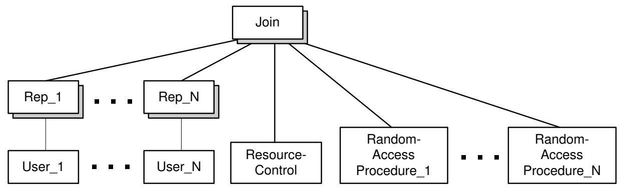


Figure 2. Composed model

The QoS indicators that we analyze are the following:

1. The call blocking probability (P_{block}), that is the probability that a call is blocked due to a random-access procedure failure.

2. The mean access time (T_{access}), that is the mean duration of the random-access procedure from the time a user sends a channel request to the time it receives a traffic channel or the random-access procedure fails.

3. The cell workload in uplink (η_{ul}) and downlink (η_{dl}).

4. The number of traffic channels allocated, that corresponds to the average number of served users.

The first two measures concern the QoS from the users' perspective, while the others are more mobile operators' oriented.

3 The UMTS cell model

All the developed models are derived using the Stochastic Activity Networks [12] (SAN) formalism. The UMTS cell model is built through the composition of three types of submodels, each one capturing a different aspect of the system behavior (see Figure 2). Suppose to have N different service classes (e.g. conversational, interactive, background, etc.) distinguished by the throughput, the workload they add to the cell (ΔL_{ul} and ΔL_{dl}), the duration of the service requested ($ServTime$), the duration of the idle period ($IdleTime$), the duration of the idle period in case of random-access procedure failure ($BlockTime$) and the set of sub-channels available for preamble transmission ($SelSubCh$). For each service class *i* (with *i*=1,...,N), we build one model capturing the user's behavior belonging to this class ("User_*i*" submodel) and one model representing the random-access procedure associated to the same class ("Random-Access Procedure_*i*" submodel). On the contrary, the traffic channel allocation, the admission control strategy and the collisions control on preamble transmissions are shared by all the service classes and are modeled in the "Resource-Control" submodel.

As depicted in Figure 2, the models are composed together using the *Join*[1] and *Rep*[2] operators [13] made available by the SAN formalism. The composed model is defined as a tree in which the leafs are the submodels and each non-leaf node is a Join or a Rep node. Since each "User_*i*" submodel (with *i*=1,...,N) represents the behavior of a single user belonging to the service class *i*, then it has to be replicated (through the Rep_*i* operator) x_i times if x_i is the total number of users within the service class *i*. The root of the tree represents the complete composed model, that is the UMTS cell model.

For the sake of brevity, in this paper we only detail the random-access procedure model; the descriptions of the other models can be found in [14, 15].

The UMTS cell model is based on these assumptions:

- The number of users camped in the cell is constant, and no user can move from a service class to another.

- The admission control strategy is based on the workload, according to the previous relations (1) and (2).

- If the cell accepts the service request, a Dedicated CHannel (DCH) is assigned to the user until the service is completed (an ongoing call can not be dropped).

- When the service is completed (or the random-access procedure fails) a user sends a new service request to the network after a period defined by the $IdleTime$ (or $BlockTime$) parameter.

- The cell uses two Frequency Division Duplex (FDD) carriers.

3.1 The random-access procedure_*i* model

The random-access procedure model for a generic service class *i* (with *i*=1,...,N) is depicted in Figure 3 and it represents the PRACH structure sketched in Section 2. In this subsection we outline the meaning of its main elements.

Place `ready_`*i* is shared by all the replicas of the "User_*i*" submodel and its marking represents those users belonging to the service class *i* that are requiring a new service. The input gate `sync_prach` is used to synchronize the service requests with the beginning of one of the two frames of the PRACH. The number of tokens in `num_slot` (initially set to 8) identifies the current slot number in the current frame, while the current frame is identified by the number of tokens in place `count_sfn` (initially set to 0): it is the first frame (with 8 access slots) if the mark of `count_sfn` is odd, otherwise it is the second frame (with 7 access slots). Places `num_slot` and `count_sfn` are shared with the "Resource-Control" submodel that modifies their marks following the PRACH structure specification. In particular, a deterministic transition removes from place `num_slot` one token every $\frac{20}{15}$ ms (slot length), as

[1] A Join is a general state-sharing composition node used to compose two or more submodels (the children of the Join node).

[2] A Rep is a special case of the Join node used to construct a model consisting of a number of identical copies of a submodel (the child of the Rep node).

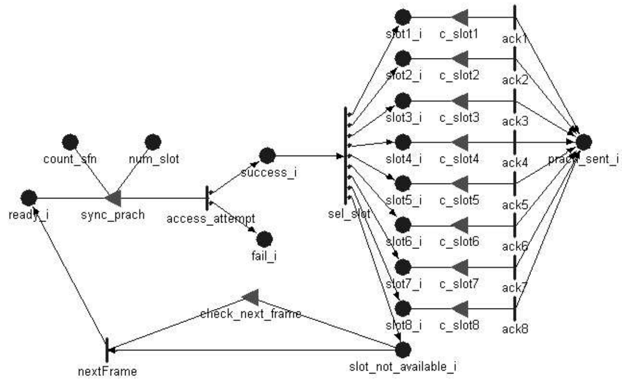


Figure 3. Random-access procedure_*i* model

each frame has a length of 10 ms. When all the tokens have been removed, its marking is reset to 7, a token is added to count_sfn (now it is odd), and so on.

When one of the two frames begins, each UE decides whether to send the preamble on the PRACH or not, based on the persistency value received through the BCH. This choice is modeled using the instantaneous transition access_attempt that moves a token from place ready_*i* to place success_*i* (with a probability Asc) if the preamble transmission is allowed, or to place fail_*i* (with a probability $1 - Asc$) if the preamble transmission is not allowed. Place fail_*i* is shared by all the replicas of the "User_*i*" submodel. If a UE has further preamble transmission attempts, inside the corresponding "User_*i*" submodel one token will be moved from place fail_*i* to place ready_*i* in the next transmission time interval, otherwise a random-access procedure failure occurs.

Tokens in place success_*i* are uniformly distributed through the firing of the instantaneous transition sel_slot into a selection of places slot1_*i*,...,slot8_*i*, representing the PRACH access slots. The selection is based on the current frame number (mark of count_sfn) and on the PRACH sub-channels available for the service class ($SelSubCh$), following the relationships defined in Table 1. Places slot1_*i*,...,slot8_*i* represent the slots belonging to the first PRACH frame if the mark of count_sfn is even, otherwise they represent the slots belonging to the second PRACH frame (with place slot8_*i* never selectable). Tokens in place slot_not_available_*i* represent those UE that could send the preamble (they passed the persistency check) but they have to wait for the next frame because of slots unavailable in the current frame. The input gate check_next_frame and the instantaneous transition nextFrame synchronize the new channel requests with the next frame. The input gates c_slot1,...,c_slot8 synchronize the preamble transmissions with the selected access slots, as each preamble transmission actually happens when the selected access slot is the current one. When j is the current slot (mark of num_slot = j, with j=1,...,8), the input gate c_slot$j$ enables the corresponding instantaneous transition ack$j$ and all the tokens in slot$j$_*i* are moved to place prach_sent_*i*, whose marking represents the number of users belonging to service class i that have sent the preamble to the network.

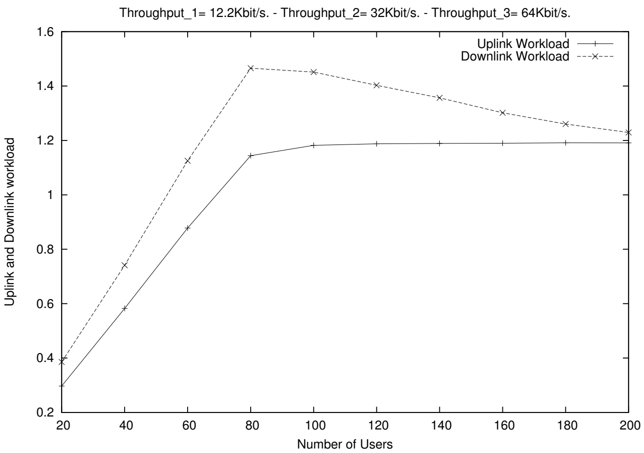
4 QoS evaluation

We evaluated the identified QoS measures in a representative scenario with three service classes. A steady-state simulation has been performed, using the simulator provided by the Möbius tool [16, 17], at varying of the number of users camped in the cell. Table 2 summarizes the settings for the main system parameters, distinguished by service class. The cell workload in uplink ($\eta_{ul\_threshold}$) and downlink ($\eta_{dl\_threshold}$) are doubled with respect to their typical values (about 65% in uplink and 80% in downlink) as we are assuming that the cell uses two FDD carriers, while the other parameters' settings are characteristic of a macrocellular environment. In all the simulations we are assuming that each service class has the same number of users ($\frac{1}{3}$ of the total number of users camped in the cell).

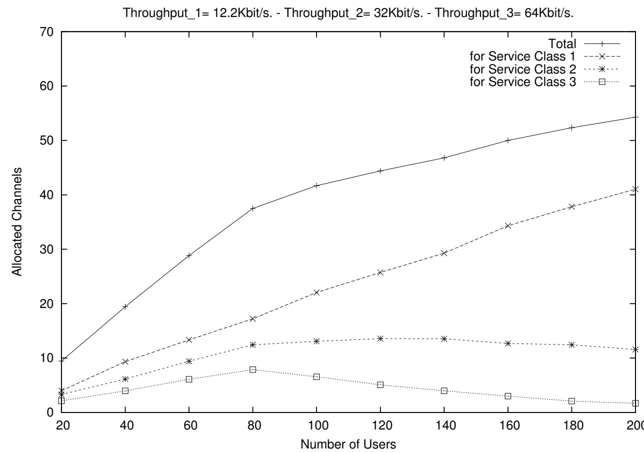
Parameters	Serv.Class 1	Serv.Class 2	Serv.Class 3
$Throughput$	12.2 kbps	32 kbps	64 kbps
ΔL_{ul}	0.016	0.037	0.050
ΔL_{dl}	0.013	0.046	0.083
$\eta_{ul\_threshold}$	1.3		
$\eta_{dl\_threshold}$	1.6		
Asc	60 %	50 %	
$SelSubCh$	0,1,2,3	4,5,6,7	8,9,10,11
M_{max}	8		
$ServTime$	300 sec.	266 sec.	131 sec.
$IdleTime$	150 sec.	300 sec.	
$BlockTime$	30 sec.		

Table 2. Main parameters'values

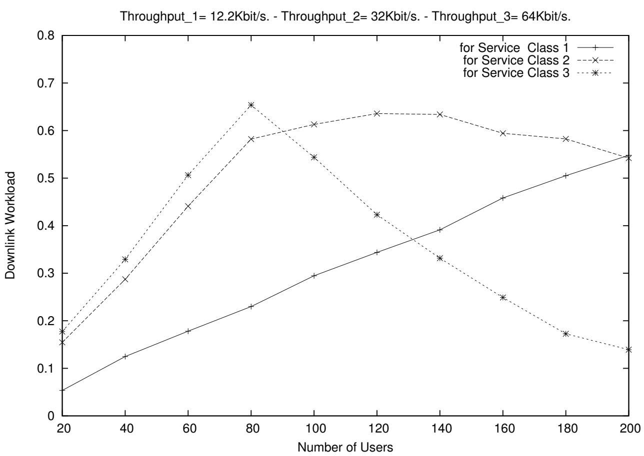
Figure 4(a) shows the uplink (η_{ul}) and downlink (η_{dl}) workload at varying of the number of users camped in the cell. It can be noted that both of them increase linearly until the maximum workload level in uplink is reached (with about 80 users in the cell). From this point, η_{ul} remains constant while η_{dl} decreases. This trend can be explained by analyzing the plots in Figure 4(b) that describes the number of allocated channels for each service class. It can be noted that, once the system reaches the maximum uplink workload level, the number of channels allocated to the first service class increases, while decreases for the other two service classes. This is due to the chosen admission control strategy, that accepts a new call if the workload level reached after adding the call does not exceed a pre-specified threshold, both in uplink and in downlink. Since the workload increment caused by the service class 1 is lower than the others, then it is more likely that the system accepts these types of calls when the workload of the



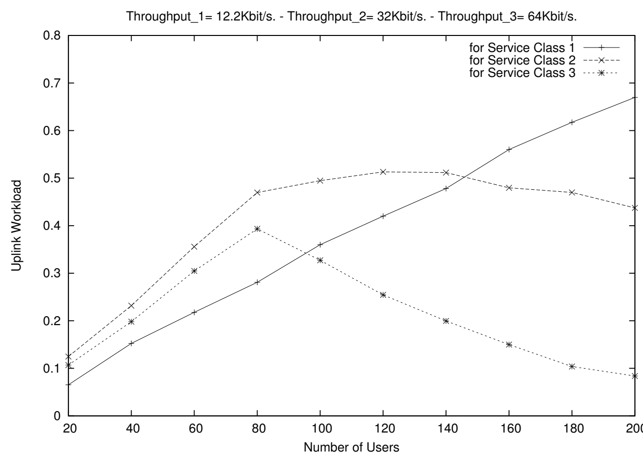
(a)



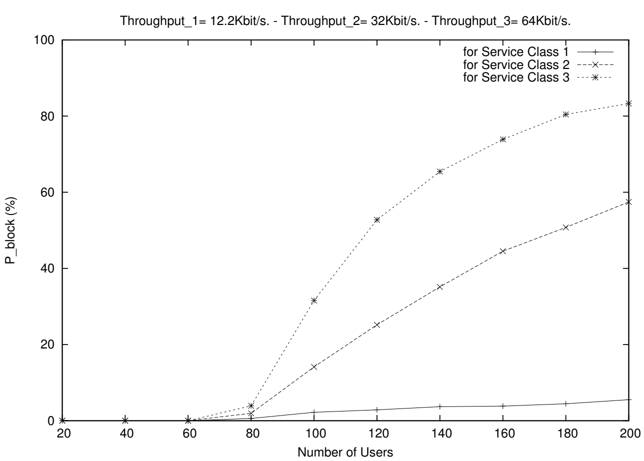
(b)



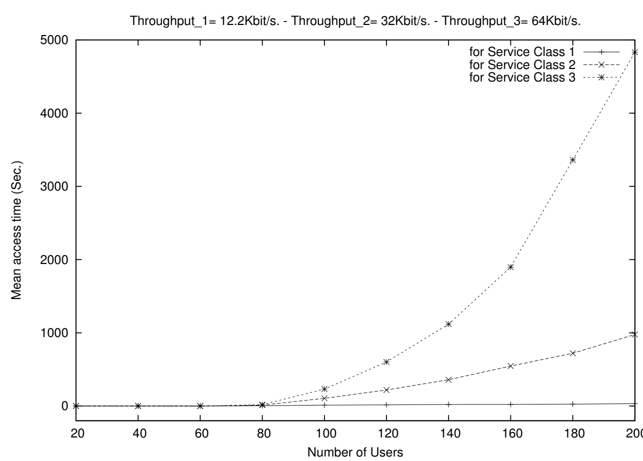
(c)



(d)



(e)



(f)

Figure 4. (a) Total workload, (b) Allocated channels, (c) Downlink workload, (d) Uplink workload, (e) Call blocking probability, and (f) Mean access time

cell is near its maximum level. Figures 4(c) and 4(d) show, respectively, the downlink (η_{dl}) and uplink (η_{ul}) workload variations for each of the three service classes. Again, the workload contribution of service class 3 rapidly becomes lower if the number of users in the cell increases beyond the 80 units. Finally, Figures 4(e) and 4(f) show, respectively, the call blocking probability and the mean access time for each service class. These measures depend on the number of users that try to get a channel at the same time; at steady-state, this number is only a fraction of the total population in each service class, since some users are not requiring any service and some others are currently being served. From the Figures we observe that, as expected, the users in the first service class have a good QoS perception as both measures are really low. On the contrary, when the cell reaches its maximum workload level the two measures rapidly increase for the users in service class 2 and 3, thus worsening the perceived QoS.

5 Conclusions

This paper has presented a study on modeling and analysis of a UMTS cell, accounting for different classes of services. QoS measures, relevant from both a user's and an operator's perspective, have been evaluated, to better understand the underlying processes and get useful insights on proper configurations of UMTS cells. Actually, this analysis constitutes a first, necessary step towards the study of a more complex scenario, involving multiple cells and accounting for dependability critical conditions, such as cell outages. These constitute the main directions for future developments.

6 Acknowledgements

This work has been partially supported by the European Community through the IST-2001-38229 CAUTION++ project. The authors want also to gratefully acknowledge the contribution given by Gabriele Petruzzi to the early phases of this work.

References

[1] S. Porcarelli, F. Di Giandomenico, A. Bondavalli, M. Barbera, & I. Mura, Service-level availability estimation of GPRS, *IEEE Transactions on Mobile Computing*, *2*(3), 2003, 233-247.

[2] P. Lollini, A. Bondavalli, & F. Di Giandomenico, Evaluation of the impact of congestion on service availability in GPRS infrastructures, To appear in *Lecture Notes in Computer Science*, Springer-Verlag, 2005.

[3] S. Z. Ozer, & S. Papavassiliou, Performance analysis of CDMA systems with integrated services, *IEEE Transactions On Vehicular Technology*, *52*(4), 2003.

[4] L. Huang, S. Kumar, & C.-C. Jay Kuo, Adaptive resource allocation for multimedia QoS management in wireless networks, *IEEE Transactions On Vehicular Technology*, *53*(2), 2004.

[5] M. Cesana, & A. Capone, Impact of mixed voice and data traffic on the UMTS-FDD performance, *Proc. IEEE Global Telecommunications Conference*, *1*, 2002, 758-762.

[6] ETSI TS 25.214 *Physical layer procedure (FDD)*, Release 1999, http://www.3GPP.org

[7] ETSI TS 25.321 *MAC protocol specification*, Release 2002, http://www.3GPP.org

[8] M. Andersin, Z. Rosberg, & J. Zander, Soft and safe admission control in cellular networks, *IEEE/ACM Transaction on Networking*, *5*(2), 1997, 255-265.

[9] Y. Ishikawa, & N. Umeda, Capacity design and performance of call admission control in cellular CDMA systems, *IEEE Journal on Selected Area in Communications*, *15*(8), 1997, 1627-1635.

[10] A. I. Zreikat, & K. Al-Begain, Soft handover-based CAC in UMTS systems, *Proc. 10th International Conference on Telecommunications*, *2*, 2003, 1307-1312.

[11] P. Dini, & S. Guglielmucci, A call admission control strategy based on fuzzy logic for WCDMA systems, *Proc. IEEE International Conference on Communications*, *4*, 2004, 2332-2336.

[12] J. F. Meyer, A. Movaghar, & W. H. Sanders, Stochastic activity networks: structure, behavior, and application, *Proc. International Workshop on Timed Petri Nets*, Torino, Italy, 1985, 106-115.

[13] W. H. Sanders, & J. F. Meyer, Reduced base model construction methods for stochastic activity networks, *IEEE Journal on Selected Areas in Communications*, *9*(1), 1991, 25-36.

[14] P. Lollini, A. Bondavalli, & F. Di Giandomenico, *QoS evaluation in a UMTS cell*, CNR-ISTI Technical Report, Doc.Id: 2005-TR-15, http://dienst.iei.pi.cnr.it/

[15] G. Petruzzi, *Modellizzazione ed Analisi della QoS offerta da una cella UMTS*, MS thesis (in Italian), Univ. of Florence - D.S.I., 2004.

[16] A. Williamson, *Discrete event simulation in the Möbius modeling framework*, MS thesis, Univ. of Illinois at Urbana-Champaign, 1998.

[17] D. D. Deavours, G. Clark, T. Courtney, D. Daly, S. Derisavi, J. M. Doyle, W. H. Sanders, & P. G. Webster, The Mobius framework and its implementation, *IEEE Transactions on Software Engineering*, *28*(10), 2002, 956-969.