A Joint Radio- IP Resource Reservation Scheme in All-IP 3rd Generation Networks

Gabor Fodor*, Gabor Malicsko+, Szabolcs Malomsoky+

*Ericsson Research, Stockholm, Sweden

Kistagangen 20-26, Kista, SE-164 80 Stockholm, Sweden,

tel.: +46 8 4043084, fax: +46 8 4047020, E-mail: Gabor.Fodor@era-t.ericsson.se

+ Ericsson Traffic Analysis and Network Performance Lab, Budapest, Hungary

Laborc 1. 1037 Budapest, Hungary,

tel.: +36 1 4377621, fax: +36 1 4377219, E-mail: Szabolcs.Malomsoky@eth.ericsson.se

Abstract - **As the All-IP architecture for W-CDMA based 3rd generation cellular networks matures within the ITU and 3GPP, there is a growing interest in devising resource reservation schemes that allocate both radio and IP resources in the access part of the network. In this paper we consider both the control- and the user plane of an end-to-end scenario with both IP and radio level QoS mechanisms and propose the addition of new parameters to the RSVP/IntServ model. Simulations indicate that these new parameters improve the efficiency of the radio resource allocation.**

I. Introduction

As the standards for 3rd generation mobile radio systems mature, there is a growing interest in the introduction of the Internet Protocol (IP) technology into the access part of the network. Ultimately, 3rd generation networks such as the UMTS/IMT-2000 are expected to implement the *all-IP architecture,* where the IP technology is used both at the user and at the network side for both voice- and data services. (This is opposed to the case where the IP stack is terminated in the base transceiver station (BTS) even if it is used as a transport technology in the access network.) This architecture provides a simplified and natural access scheme to IP based networks, such as the Internet [1].

In the radio access network, the *Packet Data Protocol* (PDP) *context* must be activated before the mobile terminal (MT) can communicate with an external IP network. The PDP context describes the characteristics of the connection to the external packet data network – type of network, network address, access point name, *quality of service (QoS)*, radio priority, and so on. [2]. Consequently, before communication between the MT and the external network can commence, the MT must request the PDP context activation and must specify the associated QoS parameters. These parameters are then used to allocate the necessary radio resources for the MT.

In an all-IP infrastructure it is desirable that a single application programming interface (API) be used to indicate the application resource requirements (including both IP and radio related resources). The most widely accepted resource reservation mechanism for IP based networks is the RSVP signaling mechanism and the associated definition of the Integrated Services (*IntServ*) classes. It follows that in order for the MT to initiate the PDP context, it must be able to derive the radio QoS parameters from the IntServ class and its associated parameters selected by the user at the API.

Unfortunately, the IntServ QoS classes specify a much coarser set of QoS parameters than what is needed by the PDP context signaling. For instance, in a PDP QoS profile a precedence class, a delay class and a reliability class must be specified [3], while in the IntServ class definitions there are no parameters for these purposes. Therefore, we propose the introduction of a new service class, referred to as the *wireless QoS class* into the RSVP/IntServ classes that would contain these parameters and allow for efficient radio resource allocation.

We organize the paper as follows. In Section II we describe the end-to-end scenario where the MT uses RSVP at the API to signal resource requirements. Section III considers the control plane aspects (i.e. RSVP-PDP signaling) of this scenario. Section IV develops a simulation model of the user plane that allows us to study the mapping of the new IntServ parameters onto the PDP context parameters. Section V presents and discusses numerical results in the case when the RSVP signaling uses the proposed wireless-QoS class. Section VI draws conclusions.

II. The End-to-End Scenario with RSVP/PDP Signaling

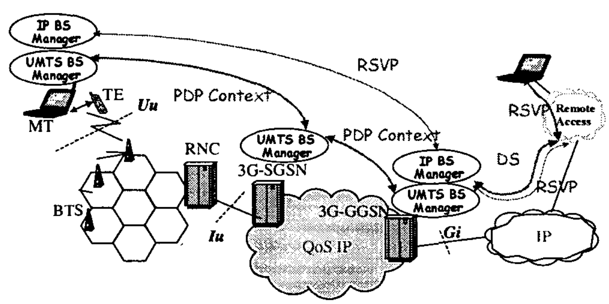

Figure 1A: *A 3G End-to-End Scenario with RSVP QoS Signaling*

We consider the end-to-end scenario depicted in Figure 1A, where the MT communicates with a remote host through the Uu, Iub, Iu, Gn and Gi interfaces, an external IP network (such as the Internet) and the remote access network. In the 3G radio access network (RAN) and in the core network (CN) the UMTS Bearer Service (BS) Manager is responsible to manage radio- and transport resources (e.g. AAL2 or IP) and to allocate the appropriate radio access bearers (RAB's) to QoS-enabled applications. For this purpose, the concept of the *UMTS BS Manager* has been introduced into the 3G architecture, which exercises the PDP context signaling.

As in Figure 1A, the scope of the UMTS BS Manager includes the RAN and CN. In order to provide for the end-to-end resource management (including the Gi interface, the external IP network and the remote access network), we introduce the concept of the *IP BS Manager,* which manages the IP resources in the external IP network and provides for inter-working in the remote access network. While the IP BS Manager in general may include the DS, RSVP or other IP QoS functionalities, in this paper we assume that it implements RSVP.

Two aspects of the Scenario in Figure 1A are noteworthy:

- Because the MT user uses the IP based API, the IP BS Manager must provide enough information to the UMTS BS Manager to allocate appropriate RAB's from the TE up to the GGSN.
- The IP BS Manager manages IP resources outside the 3G network, i.e. the IP transport resources carrying the PDP contexts within the 3G RAN and 3G CN are managed by the UMTS BS Manager.

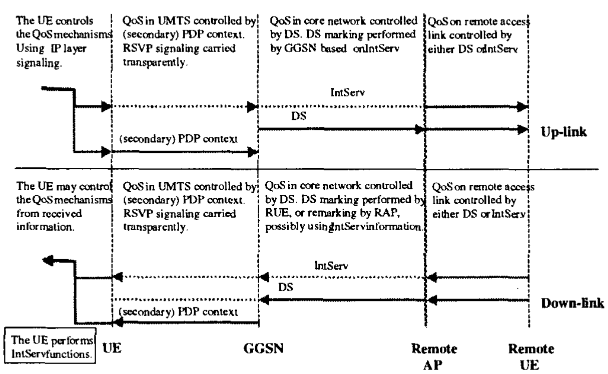

Figure 1B: *Resource Reservation Mechanisms in the End-to-end Scenario*

Figure 1B highlights the resource management mechanisms in the different parts of the end-to-end scenario of Figure 1A. From the discussion above we conclude that unlike in traditional IP networks, the RSVP signaling in 3G networks has two different roles: it helps the UMTS BS manager to allocate radio resources and it inter-works with QoS mechanisms in external IP- and remote access networks.

III. End-to-End QoS Signaling Mechanisms

It is clear that in the end-to-end scenario depicted by Figures 1A and 1B, the IP and UMTS signaling must co-operate in order to allocate resources along the end-to-end path. Figures 2A and 2B present signaling flow diagrams that are applicable to the scenario of Figure 1A.

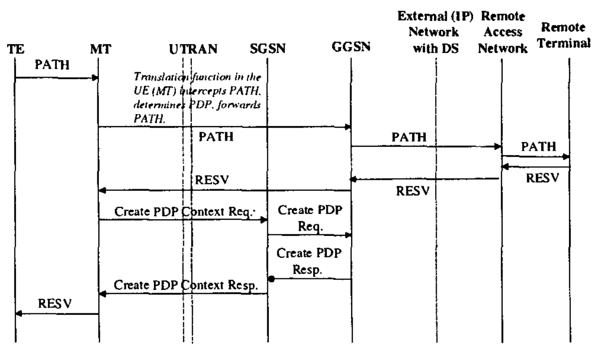

Figure 2A: *QoS Signaling Mechanisms in the End-to-end Scenario*

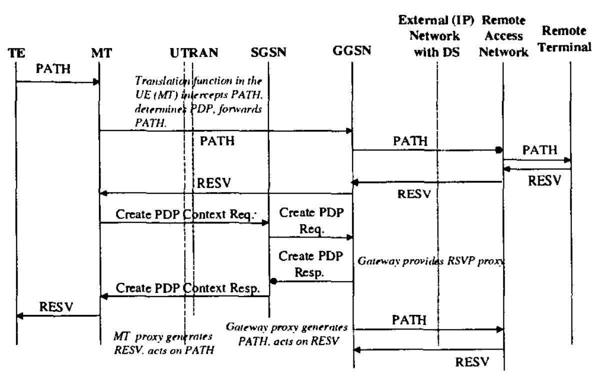

Figure 2B: *QoS Signaling Mechanisms in the End-to-end Scenario*

Since we assume that the API is implemented in the TE, in Figures 2A and 2B the TE uses an RSVP interface towards the MT. In the down-link, the PATH message is forwarded towards the GGSN by the MT. Additionally, the MT uses this message to initiate the PDP context upon the receiving the RESV message from the remote terminal. In both Figures the MT implements the UMTS BS Manager and some translation function between the IP BS Manager and the UMTS BS Manager. As the RSVP protocol is a soft-state protocol, Figure 2B introduces the RSVP proxy in the GGSN and in the MT in order to reduce the number of messages over the Uu interface.

IV. User Plane Simulation Model

In order to motivate the introduction of new IntServ parameters into the wireless QoS class and to evaluate the impact of these parameters on the efficiency of the radio resource allocation we need to develop a simulation model of the end-to-end scenario of Figures 1A and 1B. Figure 3A depicts the user plane protocol stack highlighting the layers that have significant effect on system performance and that are therefore considered in the simulation model. The simulation model is an enhanced version of the one presented in [4] and [5]. The simulator has been implemented in the PLASMA environment presented e.g. in [6] and [7].

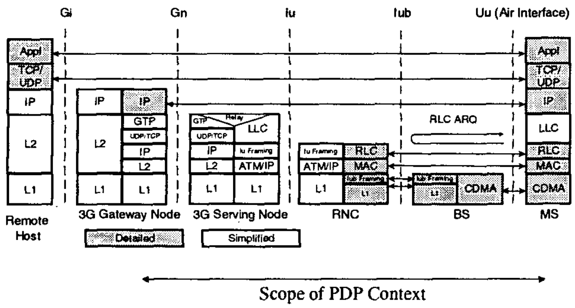

Figure 3A: *The user plane protocol stack of the end-to-end scenario of Figure 1A.*

According to Figure 3A, in the simulation model, IP packets arriving in the down-link to the UMTS network are transported to the Radio Network Controller (RNC). The IP packets are transferred to the RLC layer, where they are segmented into smaller RLC protocol data units (PDU) and the segments are buffered. For each connection, a separate RLC protocol entity exists. The RLC PDU queues of a particular IP connection are served by the MAC layer. In deterministic transmission time intervals (TTIs), the MAC layer entities ask the corresponding RLC layer entities for a certain number of RLC PDUs, which are then transferred through the radio interface in MAC frames. In our model, two RLC data transfer modes are considered: unacknowledged mode is used for UDP, and acknowledged mode for TCP.

Two traffic source types are used for performance evaluation. A UDP/IP packet generator characterized by two random variables: the size of the generated IP packets (L) and their inter-arrival time (I). This model represents traffic belonging to the conversational and streaming classes of UMTS. Traffic belonging to background and interactive classes are simply considered as best effort traffic, and are modeled as large file downloads using the TCP protocol.

In WCDMA, the loss probability of RLC PDUs over the radio interface (Block Error Rate, BLER, denoted by γ) is the function of the bit energy to noise ratio (E_b/N_0) [8].

An example BLER- E_b/N_0 curve valid in vehicular environment for 40 bytes long PDUs with MTs moving at approximately 3km/h can be seen in Figure 3B [9].

For the sake of simplicity, in our model we assume that the stochastic process of block (RLC PDU) errors is un-correlated. If unacknowledged RLC mode is used, the IP packet loss probability η for a given value of γ is:

$$\eta = 1-(1-\gamma)^{\frac{packetsize}{RLCsize}}$$

The values of η in function of the E_b/N_0 and the packet size are depicted in Figure 3C.

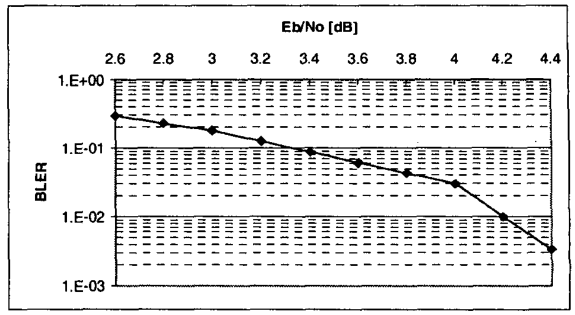

Figure 3B: *Relation between E_b/N_0 [dB] and BLER*

E_b/N_0 can be calculated from the signal to noise ratio (S/N) experienced by a specific traffic flow with the following formula [8]:

$$\frac{E_b}{N_0}_i = \frac{S}{N}_i \cdot \frac{W}{R_i} \qquad (1)$$

,where W is the total bandwidth of the carrier and R_i is the MAC server rate (defined by the number of RLC PDUs transmitted within a TTI) of the connection.

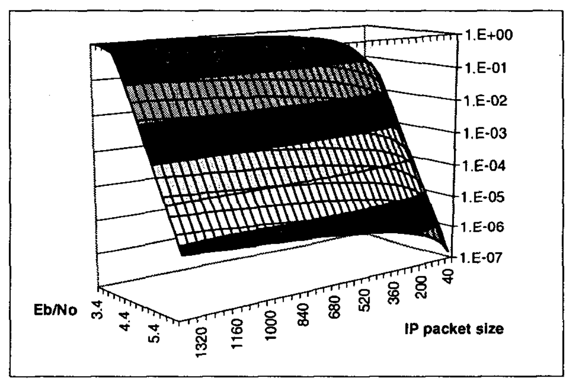

Figure 3C: *IP SDU loss ratio in Unacknowledged mode*

The signal-to-noise ratio is estimated in the following way [10]:

$$\frac{S}{N}_i = \frac{P_i^{trx} / L_i}{(a+f)\sum_{j\neq i} P_j / L_i + N} \qquad (2)$$

where:

P_i^{trx} represents the BTS transmit power of the connection,

$P_j = P_j^{trx} \cdot F_j$ is the average power level, F_j being the bearer utilization.

L_i is the path loss, $L_i = 10^{\frac{\beta + \alpha \cdot 10 \cdot \log d_i}{10}}$, where d_i specifies the distance of the MT from the BS in meters, while α and β are distance attenuation constants.

N is the background noise level.

a is a code orthogonality factor, describing the interference received from the MTs in the host UMTS cell, ranging between [0,1],

f is a factor representing the average interference experienced from neighboring cells (roughly estimated as a fraction of the own cell interference).

For efficient data transmission, every connection should keep a predefined target BLER value, maintained using fast power control [8]. In overload situations, however, the power control becomes unable to fight the growing interference; therefore the bearer rates of interactive and background class connections must be reduced by the congestion control mechanism. The number of overloads in a period is closely related to user satisfaction [11], therefore we introduce a measure U that is the percentage of time when no quality degradation is experienced by a user.

In the present model the power levels of the connections are checked once in every 10 msec interval by solving for P_i^{trx} the system of linear equations derived from (2) [10]:

$$N = \frac{P_i^{trx}}{\frac{S}{N_i} \cdot L_i} - \frac{(a+f)}{L_i} \sum_{j \neq i} P_j^{trx} \cdot F_j \qquad (3)$$

,where S/N_i's are target values, obtained from the target Eb/No's applying (1).

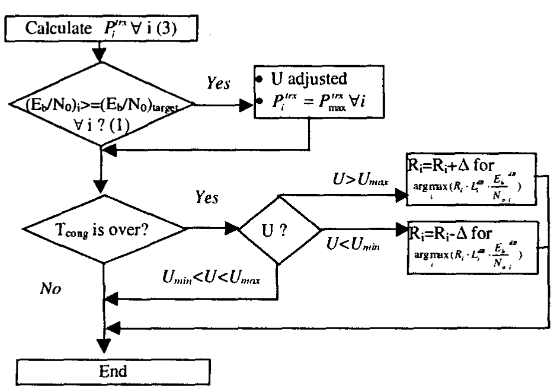

Figure 3D: *Flow chart of the congestion control*

Power levels exceeding the maximum transmission power of a connection (P_{max}^{trx}) are signs of congestion. In these cases, packet losses over the radio interface obey to the E_b/N_0 values obtained from equation (2) by substituting P_{max}^{trx}. If the U values are under a given target (U_{min}) during a longer period (T_{cong}), the service rate of the background and interactive traffic flows is reduced, according to [11]. In order to ensure an efficient control, a U_{max} value is also defined. If the interference is found to be too low ($U>U_{max}$), the service rates can be increased, as seen on Figure 3D.

V. Numerical Results and Discussion

Simulations have been carried out to analyze how the adjustment of the UMTS Radio Access Bearer Service Attributes (RABSA) [3] influence the radio performance given the same RSVP QoS descriptor. From the set of RSVP attributes, the determination of the SDU loss ratio in RABSA is not trivial, therefore in the simulations we focus on the performance effects of this parameter.

Unless specified differently, in all of the studied scenarios the parameter settings detailed in the Appendix apply.

In the first scenario, only UDP sources representing multimedia sessions are considered. The packet inter-arrival times I follow exponential distribution with a mean value of 20 msec, while the packet size L can be 80 bytes or 0 with equal probability. All sources receive 32Kbps radio bearer (TTI = 10ms, 80 bytes transmitted in each TTI), and no congestion control is applied. The SDU loss in RABSA can be adjusted in the range of $[10^{-1}...10^{-5}]$. The corresponding E_b/N_0 range based on Figure 3C and Figure 5A is approximately [3.8dB...5.6dB].

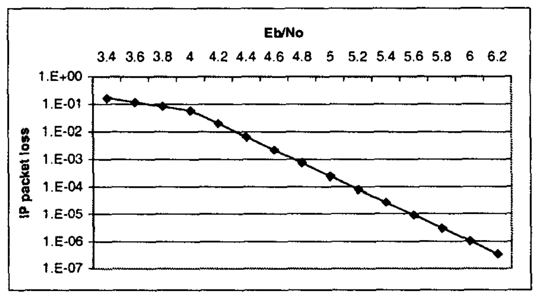

Figure 5A: *IP SDU loss ratio in Unacknowledged mode: Fig. 3C "zoomed out" for 80 byte long packets*

In the simulation we are interested in the values of U, for different E_b/N_0 levels, and increasing number of users. The results are depicted in Figure 5B.

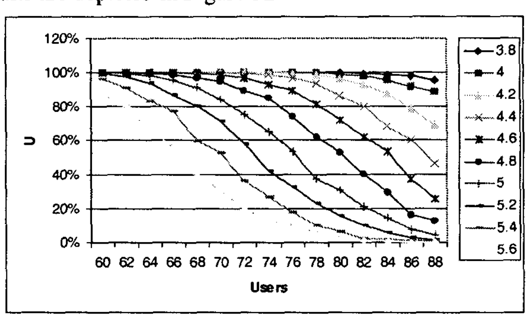

Figure 5B: *Values of U, for different E_b/N_0 levels (1. scenario)*

As it can be seen from Figure 5B, if a user satisfaction level of U = 95% is to be ensured, approximately 90 users can communicate at an E_b/N_0 level of 3.8dB corresponding to $\eta \approx 8\%$ packet loss, but less than 60 users can be admitted if the requirements are strict (5.6dB, $\eta \approx 10^{-5}$).

Apart from the number of admitted users, the throughput experienced by best effort traffic is also of substantial importance. Therefore, in the following simulation setup 4 new users are included (compared to the previous scenario) downloading large files using the TCP protocol. The bearer rates of TCP flows are subject to the congestion control described earlier. The results can be followed in Figure 5C-D.

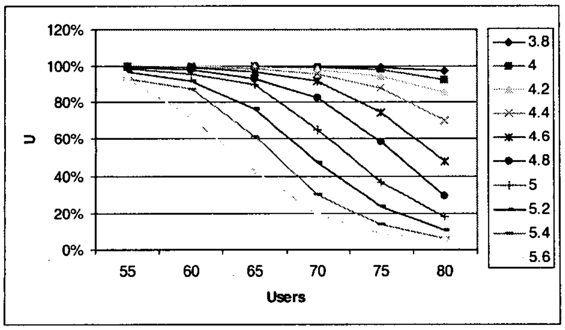

Figure 5C: *Values of U, for different E_b/N_0 levels (2. scenario)*

The phenomenon already observed in Figure 5B can be also seen in Figure 5C. That is, the number of users with UDP/IP traffic admitted at a given level of U depends largely on their required channel quality.

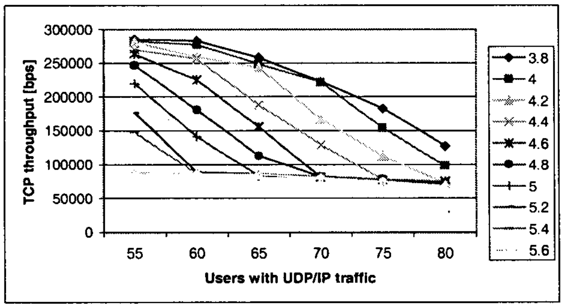

Figure 5D: *Aggregated throughputs of best effort traffic*

Figure 5D depicts the aggregated application level throughput received by the 4 additional users with best effort traffic. The different curves correspond to the different E_b/N_0 targets of the UDP flows. The target E_b/N_0 of the TCP flows is 3dB, and the RLC protocol works in acknowledged mode. The differences are straightforward when the UMTS cell is slightly loaded (55 users with UDP/IP traffic), but the advantages of the lower channel quality demands still remain notable at higher utilization levels. Moreover, it can be seen, how the bearer rates of the TCP flows are decreased due to the congestion control, as the number of users with UDP/IP traffic increases.

VI. Conclusion

In this paper, we considered the performance issues related to the use of RSVP based QoS signaling in a 3rd generation all-IP based UMTS network. Using a joint transport network and QoS aware radio interface model, we demonstrated that the translation of the RSVP attributes into UMTS Radio Bearer Service attributes is not trivial. Moreover additional information on the quality requirements of the traffic flows - like the tolerable IP packet loss probability - can significantly increase the efficiency of transport over the radio interface. For this reason, we conclude that either a new wireless RSVP QoS class should be defined, that is extended with parameters that support easy and efficient RAB establishment, or sophisticated methods need to be developed to translate between current RSVP and RAB parameters.

VII. Appendix: Simulation Parameters

We assume that IP packets are segmented into 40 bytes long RLC PDUs. The receiver MTs are uniformly distributed in a cell with a radius of 1000 meters. Each traffic source is associated with exactly one MT. Parameters required for the interference calculations are set as follows: α = 3.76dB, β = 15.3dB, $a + f = 1$, N = -100dBm and W = 4096MHz. The maximum per-connection BTS power is $P^{trx}_{\max} = 2W(33dBm)$. The user satisfaction threshold levels considered in congestion control are U_{max} = 99.5%, U_{min} = 90%. The length of the congestion control period is $T_{cong} = 1$ *sec*. The step-size of the rate control for variable rate traffic is set to Δ=32Kbps. The minimum bearer rate (MAC server rate) received by TCP flows is 32Kbps, while the maximum is 128Kbps. TTI is 10 ms for every traffic type. The target E_b/N_0 of the TCP flows is fixed to 3dB.

References

[1] **3GPP TR 23.922,** "Architecture for an All-IP Network", ver. 1.0.0, 1999

[2] **H. Granbohm and J. Wiklund,** "*GPRS - General Packet Radio Service*", Ericsson Review, No. 2, 1999.

[3] **3GPP TS 23.107,** "*Services and System Aspects; QoS Concept and Architecture*", ver. 3.2.0, 2000

[4] **A. Valko, A. Racz and G. Fodor,** "Voice QoS in 3rd Generation Mobile Networks", *IEEE Journal on Selected Areas in Communications*, Vol 17, No 1, pp 109-123, January, 1999.

[5] **G. Fodor, G. Leijonhufvud, S. Malomsoky and A. Racz,** "Comparison of ATM/AAL2 CAC Algorithms IMT-2000 Networks", *IEEE Wireless Communications and Networking Conference, WCNC* '99, New Orleans, September 1999

[6] **A. Farago, T. Henk, Z. Haraszti and I. Dahlquist,** "*PLASMA - An integrated tool for ATM network operation*", *ISS '95*, Berlin, 1995

[7] **S. Malomsoky, S. Nadas, G. Toth and P. Zarandy,** "*Simulation Based GPRS Network Dimensioning*", *ITC Specialist Seminar on Mobile Systems and Mobility*, Lillehammer, March 2000

[8] **T. Ojanperä and R. Prasad editors,** "Wideband CDMA for Third Generation Mobile Communications", Artech House, 1998

[9] **SMG2 Wideband DS-CDMA concept group (Alpha), Evaluation Document, Part 3:** "Detailed simulation results and parameters", Tdoc SMG2 970/97, 1997

[10] **K. Hiltunen and R. De Bernardi,** "WCDMA Downlink Capacity Estimation", VTC-2000-Spring, Tokyo, Japan, May 15-18. 2000.

[11] **3GPP TR 25.922,** "Radio Resource Management Strategies", ver. 3.2.0, June, 2000