Capacity Analysis and Call Admission Techniques for CDMA Packet Transmission Systems

Angela Hernández Solana[1], Antonio Valdovinos Bardají[1], Fernando Casadevall Palacio[2]

[1]Electronics Engineering and Communications Dept. University of Zaragoza. C/María de Luna, 250015. Zaragoza (SPAIN).
e-mail: anhersol@posta.unizar.es, toni@posta.unizar.es

[2]Signal Theory and Communications Dept. Polytechnic University of Catalonia.Barcelona. SPAIN.

Abstract-- **In this paper, a capacity analysis in order to set an effective threshold for connection admission control (CAC) associated to the reverse link of a multimedia wireless communication system using CDMA has been developed. Packet transmission has been considered for both real time and non real time services. A centralized demand assignment algorithm has been implemented in order to provide QoS. Performance in a multicell scenario and in time variant wireless channel conditions has been evaluated considering channel-stated scheduling algorithms. Results obtained from packet level have provided useful information in order to induce capacity bound in a multiservice system. Delay constraints associated to individual traffic classes have been included in our analysis in order to set a more accurate threshold of CAC.**

1. Introduction

Quality of Service (QoS) for an integrated CDMA system can be provided by a combination of CAC and flow control. Given that in CDMA capacity varies due to changes in the interference, the design of CAC associated two main design problems: a) to set an effective CAC threshold to guarantee the QoS for an integration of various service types, and b) to achieve the maximum efficiency in the resources utilization.

In order to support the evaluation of future CAC policies, the present study develops a capacity threshold analysis. Computer simulation results obtained from packet level have been evaluated in order to infer and support the proposed design method for CAC. The study has been associated to the Wideband Code Division Multiple Access (WCDMA) concept defined in the Universal Mobile Telecommunication System (UMTS). Both transmissions of real-time and non real-time services have performed in packetized form over the physical data channel (DPDCH). Note that, although WCDMA specification currently considers transmission of real-time services in circuit mode, it could be addressed using circuit or packet switched model, being the latter the most topical for future popular services. A centralized demand assignment protocol has been implemented to integrate services with QoS. Several scheduling strategies, have been considered and proved to perform differentiate QoS. These disciplines in coordination with a minimum transmitted power criterion (to meet the Bit Error Rate (BER) of simultaneously transmitted packets), are responsible for arranging transmission of packets (each frame=10ms) within their specified rate requirements and delay tolerances, being channel conditions of individual users in mind. This proposal differs from packet transmission modes defined for the moment in UTRA (short infrequent packet on RACH and short number of successive frames on Common Packet Channels (CPCH) [1]). We have considered the mentioned alternative in order to transmit large or frequent data packets without content for a code. A detailed description of protocol and scheduling can be found in [2].

2. Analysis of Threshold Capacity

The set of requirements associated with a MS i located in a cell k: maximum delay and delay jitter, minimum rate $r_{i,k}$ and maximum bit (BER) or block (BLER) error rates, can be mapped into an equivalent $(E_b/No)_{i.,k}$ constraint denoted by γ_i :

$$\left(\frac{E_b}{N_o}\right)_{i,k} = \frac{W}{r_{i,k}} \frac{P_{i,k} h_{i,k}}{I_{\text{int}k} + I_{ext,k} + \eta_o W} \geq \gamma_i \quad i = 1..N \tag{1}$$

with $I_{ext,k} = \sum_{\forall k' \neq k} \sum_{\forall l} P_{l,k'} h_{l,k}$, $I_{\text{int},k} = \sum_{\forall i} P_{i,k} h_{i,k}$

where E_b is the bit energy, N_o the total interference density received at the BS, η_o the thermal noise spectral density, $I_{int,k.}$ and $I_{ext,k}$ the intracell and intercell interference respectively, $P_{i,k}$ the transmitted power associated to the MS i located in cell k, $h_{i,k}$ the path loss between MS i and the BS k, N the number of users in the cell k and W the available bandwidth in the cell (chip rate). It has been shown that, considering that power output constraints are applied to MS, $0 < P_{i,k} < p_{i,k,\max}$, the power control problem in a cell is feasible if and only if [3]:

$$\sum_{j=1}^{N} C_j = \sum_{j=1}^{N} 1 / \left(\frac{W}{r_{j,k}\gamma_j} + 1 \right) \leq \frac{1 - \eta(t)}{(1 + f(t))} = C_{\max}(t) \tag{2}$$

$$\eta(t) = \frac{\eta_o W}{min_j \left[p_{j,k,max} h_{j,k}(t) \left(\frac{W}{r_{j,k}\gamma_j} + 1 \right) \right]} \quad , \quad f(t) = \frac{I_{ext,k}(t)}{I_{\text{int},k}(t)}$$

We consider $C_{max}(t)$ is the maximum available capacity in the cell, C_i the consumed capacity by a MS and $\eta(t)$ is the reserved capacity needed to guaranty the power constraints to MS's in a frame. A minimum value of parameter η equal to 0.1 is

0-7803-7605-6/02/$17.00 ©2002 IEEE.

considered in order to limit the maximum total power received at the BS $I_{total,k}$. If condition (2) is achieved for a set of rates and E_b/N_o values, then the power can be obtained using (3).

$$P_i(t) = \frac{\eta_o W}{\hat{h}_{i,k}(t)\left(\frac{W}{r_{i,k}\gamma_i}+1\right)C_{res}(t)} \quad \text{with} \quad C_{res}(t) = 1-(1+f(t))\sum_{j=1}^{N} C_j \qquad (3)$$

Note that in a certain frame, where N_{sch} is the number of scheduled users, the relation (4) is satisfied.

$$C_{max}(t) - \sum_{i=1}^{Nsch} C_i = \frac{C_{res}(t)}{1+f(t)} - \frac{\eta(t)}{1+f(t)} \geq 0 \qquad (4)$$

The ratio between inter and intracell interference (f(t)), and the channel condition should be known in order to achieve the minimum transmitted power criterion with accuracy. Ratio f is estimated from the ratio measured in the previous frame.

Taking into account condition (2), scheduling at the BS arranges transmissions of MS's according to algorithms related with delay and rate requirements. MS's which can not met Eb/No requirements are delayed although could be in contact with the base through the Dedicated Control Channel (DPCCH) (to perform closed power control). Unused resources (power and instant of transmission) are assigned to the rest of the users. Note that waste of capacity associated to DPCCH must be added in (2). In addition to the scheduling strategies based on delay and rate requirements, we have considered a channel–state dependent scheduling algorithm based on the required value of parameter η [2]. MS's which require high value of η_i could be delayed in order to improve permits to more users.

Once we have described theoretical expressions for capacity, we will develop an effective CAC threshold in order to guarantee the QoS. This problem has been addressed before in quite a lot of works [4][5][6]. [4] presents useful expressions in evaluation of system capacity. However, since the erlang capacity is calculated based on only blocking rate, the communication quality is not guaranteed and its relation with erlang capacity is left unclear. Distribution function of interference is determined in order to evaluate blocking probability. [4] shows that the mean and variance of other-cell interference can be approximated by the intracell interference multiplied by a constant coefficient f. In [5] expressions of the QoS and GoS (grade of service) as functions of traffic intensity and CAC thresholds have been derived although only a type of service is considered. In [6], an effective capacity (rate), taking into account the stochastic nature of traffic and delay requirements in a multimedia scenario, is used to characterize the resources required by multimedia services. The proposed method, though useful, has been shown to be conservative.

In any of previous references, no specific MAC protocol has been developed in order to provide differentiated QoS. We have introduced this specific factor. In our case, unlike [4][5][6], the scheduler limits the number of MS's transmitting in a frame, thus, a priori, every MS having permit to transmit must reach its Eb/No constraints (condition (2) is satisfied). So, only imperfections in the estimation of ratio f could prevent it, given that almost perfect channel estimation could be assumed thanks to DPCCH. Thus, users can only not satisfy their QoS constraints related with delay constraints. Real time services drop packets that exceed delay constraints, so a specific drop probability constraint is imposed as QoS criterion. Simulation results show that initial capacity $C_{max}(t)$ maintains the same statistical distribution, which has been seen to correspond to a gamma distribution, with the same mean an variance independently of the number of active users in the system, being only dependent of the propagation parameters, closed power control inaccuracy and the spatial distribution of MS's within the cell. However, in order to perform CAC, we use an alternative measure of the capacity (5), more dependent of current active users, which provide a more accurate bound.

$$C_{ini}(t) = 1-\eta(t) - f(t)\sum_{i=1}^{Nsch} C_i(t) \qquad (5)$$

Simulations show that $C_{ini}(t)$ can be characterized by a F-distribution, given that the product of f(t) and $\sum_{i=1}^{Nsch} C_i(t)$ can be modeled as a product of Gamma distributions. This distribution is maintained considering either one type of traffic or integration of real time and non real time services. However, in order to simplify the analysis, we will only consider real time services, and only one type of traffic.

Note that if there are N_i MSs admitted in a cell, the probability of k_i active users may be expressed as a binomial distribution (9), where ρ_i is the activity factor associated to class i MS's. Distribution of active users is not the same as distribution of scheduled users. However, while N_i is below the CAC threshold (characterized by the maximum number of MS allowed, $N_{max,i}$) they could be considered equal, which simplifies parameters of the F-distribution associated to C_{ini}.

$$P_{u,i}(k_i) = \binom{N_i}{k_i}\rho_i^{k_i}(1-\rho_i)^{N_i-k_i} \qquad (6)$$

Dropping probability can be calculated as the fraction of all active users that can not transmit due to the absence of available capacity in a specific frame (7). (p_{Cini} is the pdf of C_{ini}.)

$$P_{drop} = \frac{\sum_{k_i=0}^{N_i} P_{u,i}(k_i)\left(\sum_{j=0}^{k_i} j\left(\int_0^{jC_i} p_{Cini}(m)dm\right)\right)}{\sum_{k_i=0}^{N_i} k_i P_{u,i}(k_i)} \qquad (7)$$

Tolerance to delay is not included in (7), so the maximum number of users allowed, considering a threshold to the P_{drop}, represents a conservative bound of the admission region. In fact, P_{drop} is the probability that a packet has been delayed.

Instead of C_i, an effective capacity, $C_{i,ef}$ (introducing the delay influence) must be derived in order to provide a more accurate result. Probability (ε_i) is the probability that delay associated to packets of any connection k belonging to class i exceed delay requirements. So, if we are able to obtain the delay distribution of the individual sources, the capacity of the system in terms of number of MS's can be calculated as (8) and, in addition, the effective capacity associated to MS of class i as (9):

$$N_{max} = max(Ni \,|\, \forall_{k=0:N_i} \ \Pr[delay_k > Dmax_k] \le \varepsilon_i) \qquad (8)$$

$$C_{i,ef} = E[C_{max}] / N_{max} \qquad (9)$$

We need compute the delay distribution for individual sources, considering their BLER and service rate requirements, the variable wireless capacity and the error control scheme (selective repeat algorithm is implemented). Considering all this factors together in a single analytical model seems inaccessible, so we have made the analysis in two steps. Probability that delay does not exceed delay requirements is computed as the probability that capacity assigned to the individual connections, in a time interval equal to delay tolerance $Dmax_k$ counted from the instant of the packet arrival to the queue (t_a), is higher than the capacity required to transmit all packets waiting in the queue when packet arrives, $q_k(t_a)$ (10).

$$\begin{aligned} &\Pr[delay_{k,i} > Dmax_{k,i}] = \\ &= 1 - \Pr(C_{asig,k\in\{N_i\}}(t_a, t_a + Dmax_k) \ge q_k(t_a) * C_i) \qquad (10) \\ &= 1 - \sum_{n_q} \Pr[C_{asig,k\in\{N_i\}}(t_a, t_a + Dmax) \ge n_q * C_i / q_k(t_a) = n_q] \Pr(q_k(t_a) = n_q) \end{aligned}$$

The probability (11)

$$\Pr[C_{asig,k\in\{N_i\}}(t_a, t_a + Dmax) \ge n_q * C_i / q_k(t_a) = n_q] \qquad (11)$$

is calculated using a model that includes the selective repeat scheme, the capacity assigned to a connection calculated in function of the delay probability calculated in (7), P_d, and the error probability related to the Eb/No requirement, P_e. Calculus of P_e includes the effects of the little deviations in the Eb/No due to errors in the estimation of ratio f and channel conditions (they have been statistically modeled). Model considers that all packets waiting to transmit must receive service within their specific delay requirement. Calculus of (11) involves two cases: the queue has packets waiting to transmit (12) and the queue is empty when a new packet is generated (13).

$$\Pr\left(C_{asig,k\in\{N_i\}}(t_a, t_a + Dmax) \ge n_q * C_i / q_k(t_a) = n_q\right) =$$

$$= \begin{cases} \left(\sum_{D=1}^{Dmax-nq+1} W^*\right) \prod_{n=1}^{n_q} \left(\sum_{D=1}^{Dmax-nq+n} \left(\sum_{w=n-1}^{D-1} W \right)\right) & \text{if } n_q > 1 \qquad (12) \\ \sum_{D=1}^{Dmax} W^* & \text{if } n_q = 1 \qquad (13) \end{cases}$$

where

$$W = \left(W_{q,1}(w) W_{s,1}(w_s) + W_{q,0}(w) W_{s,0}(w_s)\right)(1-P) \qquad (14)$$

$$W^* = \left(\sum_{w=0}^{\mathrm{int}((D-1)/2)} \binom{D-1-w}{w} P_e^{\,w} P_d^{\,D-1-2w} \right)(1-P) \qquad (15)$$

with $w_s = D - w - 1$ and P the probability that a packet can not be transmitted or can be transmitted with error. D_{max} is equal to the delay requirement in frames minus one. Calculus of W involves two parts, the delay until the first attempt to transmit the packet generated in t_a (W_q) and the delay until that packet and the packet waiting for retransmission (with smaller lifetime) are transmitted without error (W_s). W_q is calculated using:

$$W_{q,1}(w) = W_{q,-1}(w) * (1-P) \qquad (16)$$

$$W_{q,0}(w) = W_{q,-1}(w) * P_e \qquad (17)$$

where

$$W_{q,-1}(w) = \begin{cases} \sum_{w_q=n-1}^{w} \binom{w_q - 1}{n-2} P^{w_q - n + 1} (1-P)^{n-1} P_d^{\,w - w_q} & \text{if } n > 1 \\ P_d^{\,w} & \text{if } n = 1 \end{cases}$$

and W_s with expressions (18) and (19)

$$W_{s,1}(w_s) = \sum_{i=0}^{\mathrm{int}(w_s/2)} \binom{w_s - i}{i} P_e^{\,i} P_d^{\,w_s - 2i} \qquad (18)$$

$$W_{s,0}(w_s) = \begin{cases} P W_{s,0}^{\,1}(w_s, w_s > 2) P_e + W_{s,0}^{\,2}(w_s) & \text{if } w_s > 2 \\ P P_e + W_{s,0}^{\,2}(w_s) & \text{if } w_s = 2 \\ 1 & \text{if } w_s = 0 \end{cases} \qquad (19)$$

where

$$\begin{aligned} &W_{s,0}^{\,1}(w_s, w_s > 2) = \\ &= \sum_{i=1}^{\mathrm{int}((w_s-2)/2)} \binom{w_s - 2 - i}{i} \left(P_e^{\,2i} + (P_e P_d)^i\right) P_d^{\,w_s - 2 - 2i} + \qquad (20) \\ &+ \sum_{k=2}^{\mathrm{int}((w_s-2)/2)} \binom{w_s - 2 - k}{k} P_d^{\,w_s - 2 - 2k} \sum_{j=1}^{k-1} \binom{k}{j} P_e^{\,2j} (P_e P_d)^{k-j} + P_d^{\,w_s - 2} \end{aligned}$$

$$\begin{aligned} &W_{s,0}^{\,2}(w_s) = \sum_{i=1}^{w_s - 1} P \left(\sum_{j=0}^{\mathrm{int}((i-1)/2)} \binom{i-1-j}{j} (P_e P_d)^j P_d^{\,i-1-2j} \right)(1-P) * \qquad (21) \\ &* \left(\sum_{k=0}^{\mathrm{int}((w_s-i-1)/2)} \binom{w_s - i - 1 - k}{k} P_e^{\,k} P_d^{\,w_s - i - 1 - 2k} \right) \end{aligned}$$

On the other hand, the probability of queue occupation when a packet arrives is calculated from an analytical model that considers the evolution in the queue occupation when the source doesn't generate traffic an when the source is in ative state (Remember that traffic is assumed to be bursty). The interaction between sources and the corresponding changes in the assigned capacity due to changes in traffic distributions are included in the model. In this case, delay tolerance of users is only represented by the maximum size allowed to the queue and no additional considerations are taken into account.

In a first approach to the problem, the queue occupation could be calculated resolving the discrete time system represented by the state transition diagram of figure 1.

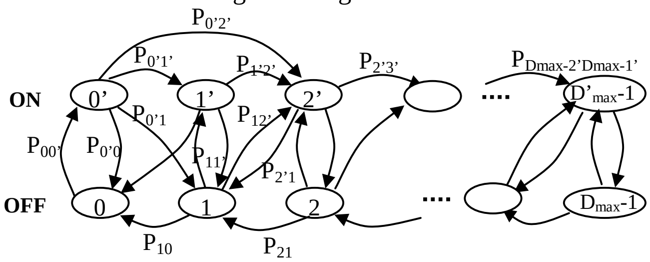

Figure 1

Transition probabilities are denoted by:

$$
\begin{aligned}
&P_{0'1'} = (1-P_{onoff})(1-P) \\
&P_{0'2'} = (1-P_{onoff})P \\
&P_{u'(u+1)'} = (1-P_{onoff})P && \text{with } u \geq 1 \\
&P_{u'u} = P_{onoff}\,P && \forall\, u \qquad (22) \\
&P_{u'(u-1)} = P_{onoff}(1-P) && \text{with } u \geq 1 \\
&P_{uu'} = P_{offon}(1-P) && \text{with } 0 \leq u < D_{max}-1 \\
&P_{(D_{max}-1)(D_{max}-1)'} = P_{offon} \\
&P_{u(u+1)'} = P_{offon}\,P && \text{with } 1 \leq u \leq D_{max}-1
\end{aligned}
$$

where probability P includes the error probability and delay probability. In this case, D_{max} is the delay tolerance. Note that in state 0' there are no packets waiting in the queue and the generated packet is not able to transmit given that at least on frame of delay is needed in order to send the corresponding access request on the DPDCH.

To incorporate a more realistic model, we have extended the model considering a separate delay and error probability function of the number of active users (P(u) with u the number of active users). In this case, the system to be resolved is :

$$
\bar{p} = \bar{p}\begin{bmatrix} A_{on} & A_{on-off} \\ A_{off-on} & A_{off} \end{bmatrix} \quad \text{and} \quad \bar{p}.\bar{e}^{T} = 1 \quad \text{with } \bar{e} = [1,1....1]
$$

$$
\bar{p} = (\bar{p}_{0'}\ \bar{p}_{1'}\ ....\bar{p}_{D_{max}-1'}\ \bar{p}_{0}\ \bar{p}_{1}\ ....\ \bar{p}_{D_{max}-1}) \qquad (23)
$$

where components $\bar{p}_i$ of the vector $\bar{p}$ represents the occupation probability of buffer for different number of user actives (from 1 to N_{max}). Note that $\bar{p}_i = (p^{i}{}_{1} p^{i}{}_{2} ....p^{i}{}_{N_{max}})$ and A_{on}, A_{off}, A_{onoff} and A_{offon} are $D_{max} \times D_{max}$ matrix of $N_{max} \times N_{max}$ components. The specific components of matrix A are no included here but could be derived easily as an extension of the model showed in figure 1.

Although inherent approaches, the model developed has been proved to be very accurate. Analysis presented here has only considered one type of traffic. However, the extension for an integration of several real time users with different BLER and rate requirement is immediate considering the effective capacity computed for each type of traffic.

In a first instance we can consider a conventional CAC that allows a new user into the radio access network if condition (24) is satisfied. Extensions of this work must include algorithms that consider the difference between new and handover users and adaptive reserve of resources depending on the types of traffic and their mobility.

$$
\sum_{i \in cell} C_{i,ef} + C_{new,ef} < E[C_{max}] \qquad (24)
$$

3. System Model And Simulation Results

We propose a cellular system model composed by 19 hexagonal cells. Only interference from the first-tier of adjacent cells is considered. Wrap-around technique is used to avoid border effect. Macro cell propagation model proposed in [7] is adopted for path loss. Log-normally distributed shadowing with standard deviation of $\sigma = 8dB$ is added according with the model proposed in [8]. Additionally, a multi-path fading environment proposed in [9] is considered. 11dB antenna gain and thermal noise power of –103dBm are assumed [7]. MSs have a maximum output power of 27 dBm according with class 2 defined in [9], and move with a speed between 25 to 50Km/h. Initial localization is selected randomly (uniform distribution).

Two kind of traffic sources: Real Time Services (Class I-Data services with delay constraints of 300 ms) and Non-real Time Services (Class II - Data services with non-delay constraints) are considered. Convolutional coding rate ½ together with a retransmission scheme (ARQ) are used to achieve BLER=10^{-2} in both cases. Real-time sessions are based on Packet Calls with a number of packets exponentially distributed with mean 35 packets, while a service of 36kps (transport block of 360bits and real transmission rate of 120kbps) is assumed. Average inter-packet arrival time is 10ms, while packet call inter arrival time is exponentially distributed with mean 1s. Non-real time traffic sources are based on the model presented in [10]. Two types are considered: 8kbps and 32Kbps data rate services.

We present some results considering shadow and fast fading. Figure 2 shows C_{max} when only class I MS's are present in the cell.Variations in statistics are only due to changes in the distributions for MS's in the cell. Gamma distribution with the same mean and variance as C_{max} provides a very good approach.

Figure 3 shows C_{ini} for 56 and 76 MS's in the cell. F-Distribution is very closed to simulation results. Results obtained with a gamma distribution with the same mean and variance seem to be much less suitable.

Figure 4 shows C_{ini}, when there are 30 class I MS's and a variable number class II.1 and class II.2 MS's. Gamma and F-distribution models look also to be good approach.

Figure 5 compares the dropping probabilities computed with models (7) (Pr_delay_model) and (10) (Pr_drop_model) and simulations results when only class I MS's are considered. Delay probability is far from dropping probability obtained in the simulation. However, the dropping probability calculated when delay tolerance is included in the model is very closed to those obtained in the simulation. Pdrop_model has been calculated considering the distribution of capacity modeled from the specific propagation conditions and spatial distribution of MS's obtained for each point of the graphic.

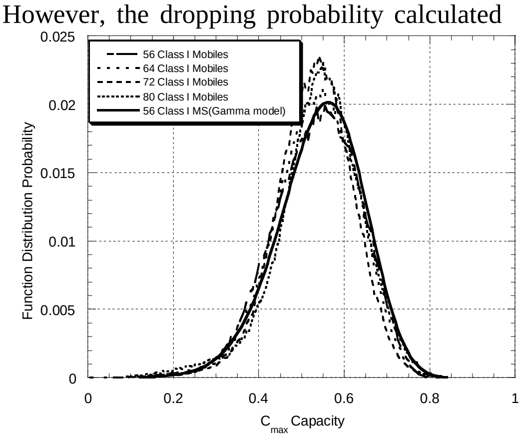

Figure 2. (Steps of pdf distribution are 0.005)

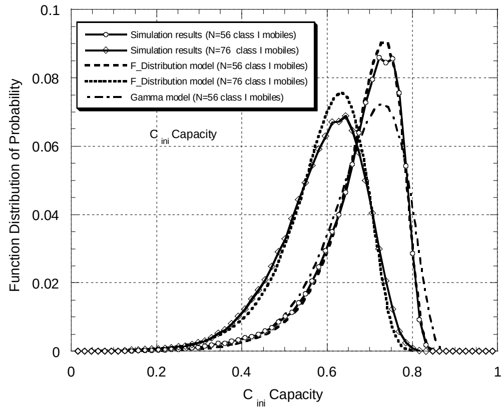

Figure 3. (Steps of pdf distribution are 1/63.74).

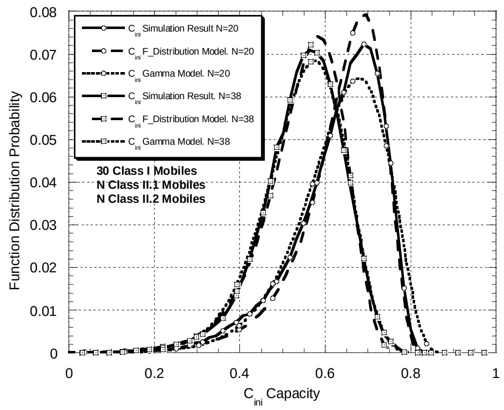

Figure 4

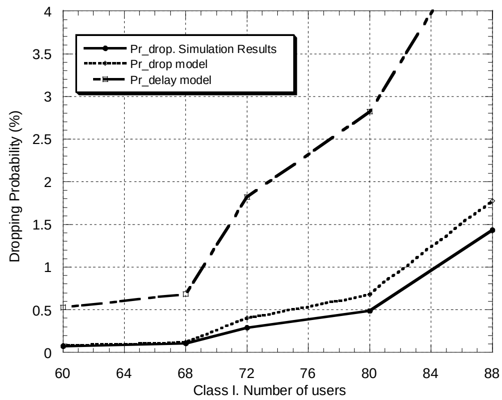

Figure 5

4. Conclusions

This paper presents a capacity analysis in order to set thresholds for connection admission control. Analysis has been associated to reverse link of the WCDMA, where packet transmission has been considered for both real time and non real time services using a centralized demand assignment protocol. Results obtained from packet level provide useful information in order to induce capacity bound. Delay constraints associated to individual traffic classes have been included in the analysis in order to set a more accurate threshold of CAC.

Acknowledgments

This work has been supported by the grants CICYT TIC2001-2481 and TIC2001-2222 from the Ministry of Science and Technology of Spanish Government and FEDER.

References

[1] 3GPP TS 25.321 " MAC Protocol Specification", v3.6, December 2000.

[2] Angela Hernández Solana, Antonio Valdovinos Bardají, Fernando Casadevall Palacio, "Performance Analysis of Packet Scheduling Strategies for Multimedia Traffic in WCDMA", Proceedings of VTC'2002 Spring.

[3] Ashwin Sampath, Sarath Kumar and Jack M. Holtzman, "Power Control and Resource Management for a Multimedia CDMA Wireless System", PIMRC'95.

[4] A. M. Viterbi and A. J. Viterbi, "Erlang Capacity of a power controlled CDMA system", IEEE Journal on Selected Areas, vol. 11, Aug. 1993.

[5] Yoshihiro Ishikawa, Narumi Umeda, "Capacity Design and Performance of Call Admission Control in Cellular CDMA Systems", IEEE J. Select. Areas Comm, Vol. 15, No 8. October 1997.

[6] Jean Q. J. Chak, Weihua Zhuang, "Capacity Analysis for Connection Admission Control in Indoor Multimedia CDMA Wireless Communications". Wireless Personal Communications 12, 2000.

[7] 3GPP TS 25.942 "RF System Scenarios". V3.1.0., June 2001.

[8] Viterbi A.J., Viterbi A. M., "Soft handoff extends CDMA cell coverage and increases reverse link capacity", IEEE Journal on Selected Areas, Vol. 12, Nº 8, October 1994.

[9] 3GPP TS 25.101 " UTRA (UE) FDD; Radio Transmission and Reception", v3.5.0, March 2001.

[10] "Universal Mobile Telecommunications System (UMTS). Selection procedures for the choice of radio transmission technologies of the UMTS", UMTS Technical Report 30.03, version 3.2.0. April 1998.