EFFICIENT CALL ADMISSION CONTROL FOR 3G NETWORKS
Chernoff-Bound based CAC

Albert Mráz, Mihály Katona
Mobile Communication Laboratory, Budapest Univ. of Technology and Economics, Magyar Tudósok krt. 2., 1117 Hungary

Dr. Sándor Imre
Mobile Communication Laboratory, Budapest Univ. of Technology and Economics, Magyar Tudósok krt. 2., 1117 Hungary

Keywords: Effective bandwidth technique, Chernoff-bound, Dynamic CAC.

Abstract: In the recent years the mobile telecommunication networks have gone through on a big development. The services of the systems have been extended very quickly, such as the number of the subscribers. The various multimedia and Internet systems have become quickly the part of our life. This phenomenon takes effect on mobile communication, too. The third-generation mobile networks could be the solution, which eliminate the defects of current systems and apply good solutions concerning both access and transport system. These 3G systems are able to meet the growing user demands and the mobile Internet requirements. Our goal is to use the scarce resources as efficiently as possible. We implemented a new, more efficient and faster call admission control than former methods. Our algorithm is able to be adapted directly for 3G mobile networks.

1 INTRODUCTION

The 3G mobile communication networks operate according to Code Division Multiple Access method. There is a major change between the methodology of the first and second generations mobile networks (FDMA and TDMA) in call admission control (CAC) policy. FDMA and TDMA networks have a hard limit for the number of subscribers operating in the network simultaneously. In the 3G CDMA systems there are no maximum limits for the number of simultaneously communicating users, because this number is determined by the user interference. Our task is to answer the question, whether the prescribed service quality (QoS) is defensible when a new call request is coming in. The main purpose is to create a call admission control method, which can adapt dynamically to an ever-changing mobile environment.

We compared our method to an algorithm (Evans and Everitt, 1999), which uses pre-calculated effective bandwidth values which on the first hand requires computationally complex (time-consuming) recalculation every time when the system parameters have been changed, and on the other hand these values are optimized for an average network scenario and not for a given one. Hence these values don't provide optimal call admission control.

We introduce a new, fast and efficient algorithm which is able to calculate the optimal effective capacity values for a given network scenario. Our solution is based on the real-time estimation of the well-known Chernoff bound (Imre, 1999).

This paper is organized as follows: In Section 2 we give an interpretation of call admission control. In the next section we describe the ON/OFF source model that we use. The introduction of the effective bandwidth technique as it was written in (Evans and Everitt, 1999) is the subject of Section 4. In section 5 we introduce the theoretical framework of CAC. In section 6 we make a transformation of CAC method so that it can be used in CDMA environment. Section 7 we present the simulation results we got and comparison between our method and the one given in (Evans and Everitt, 1999).

Mráz A., Katona M. and Imre S. (2005).
EFFICIENT CALL ADMISSION CONTROL FOR 3G NETWORKS - Chernoff-Bound based CAC.
In *Proceedings of the Second International Conference on e-Business and Telecommunication Networks*, pages 67-73
DOI: 10.5220/0001415500670073

2 GEOMETRIC INTERPRETATION OF CAC

In order to give a generalized description of call admission procedures we introduce the following scenario. Each mobile source is grouped into traffic classes. Every sources in a given traffic class have the same traffic descriptors (e.g. maximum transmission rate, average traffic rate, etc.) and number of active sources in the ith class is denoted by Ni. Therefore the state of the system in every moment can be described by a state vector

$$\boldsymbol{N} = (N_1, N_2, \ldots, N_J), \qquad (1)$$

where J refers to the number of traffic classes.

Now the call admission procedure means that we should decide whether a new call can be accepted without violating the QoS parameter guaranteed for other users or not. All state vectors can be divided into two sets (geometric interpretation at Figure 1.). Vectors that can be accepted by the network management without violating the traffic contracts belong to the first (or acceptable) set and states that must be rejected to the second (or rejected) set

$$V_{ACCEPT} := \{N: P(X > C) < \gamma\} \qquad (2)$$

$$V_{REJECT} := \{N: P(X > C) \geq \gamma\}, \qquad (3)$$

where X refers to the overall traffic of the users and C is the system capacity and γ denotes the QoS parameter. Therefore the task of CAC can be regarded as a space separation problem that is to say how to determine the surface, that separates the two regions and how to decide whether the new traffic state of the link is on the acceptable or rejected side of the separation surface.

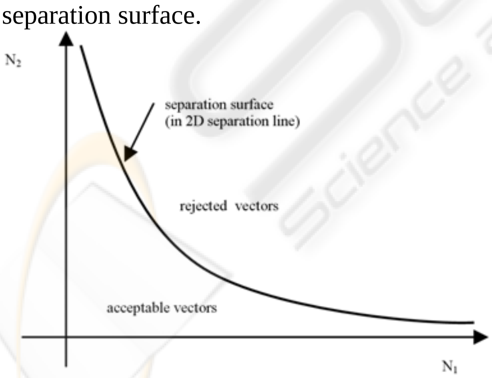

Figure 1: Geometric interpretation of CAC

Unfortunately the CAC decision cannot be carried out directly on the basis of the theoretical surface. First the calculation of exact separation surface proves to be a hard task because of the high computing complexity. Secondly the size of the space of the theoretical surface needs enormous large storage capacity.

3 APPLICATION OF ON/OFF SOURCE MODEL

In order to give a simple but widespread and close approximation of the subscribers' data traffic we use the ON/OFF model.

According to this model the traffic of a subscriber is estimated by a source which has two states, either it transmits with peak bit rate or it remains in silence, therefore the bit rate distribution of the ON/OFF source contains values different from zero only at nil and at peak bit rate on the axis of bit rate.

An ON/OFF source which describes the jth subscriber can be characterized with a peak rate (Rj or hj) and a probability (Φj) being ON. The Φj is known as the voice activity factor (VAFj). Then the mean bit rate can be calculated as

$$m_j = h_j \cdot VAF_j \qquad (4)$$

4 APPLICATION OF STATIC EFFECTIVE BANDWIDTHS

The users are grouped into traffic classes and an effective bandwidth (Kelly, 1991), (Elwalid, 1993) value is assigned to each type of user (class) which is somewhere between the mean rate and the peak rate of the variable bit rate (VBR) user. Then the actual value of the required system bandwidth is estimated by means of the effective bandwidth values and the number of the users in different classes. The effective bandwidth values are depending on the QoS parameters and on the stochastic behaviour of the users' traffic as well.

Now we describe the method of the effective bandwidth so that later we can compare it with our method. Let Z_+^J denote the set of J-tuples (J refers to the number of classes) of non-negative integers and consider the problem of finding the set of points $(N_1, N_2, \ldots, N_J) \in Z_+^J$ satisfying

$$P\left(\sum_{j=1}^{J}\sum_{i=1}^{N_j} X_{ji} > C\right) \leq e^{-\gamma} \qquad (5)$$

where X_{ji}, $i = 1,\ldots, N_j$, $j = 1,\ldots, J$ models the independent random amount of resource required by call i of class j, C is the total amount of resource available, γ is the quality of service requirement, and the N_j, $j = 1,\ldots, J$ are the number users of each class.

Inequality (5) represents the well-known tail distribution estimation problem, which requires the convolution of large number of random variables. Because calculation of convolution is rather time consuming task so the theoretical amount of resource is approximated by application of effective bandwidth without significant calculation power.

Let replace (5) by the following inequality

$$\sum_{j=1}^{J} \kappa_j N_j \le C, \qquad (6)$$

where κ_j refers to the effective bandwidth of the jth class. We remark that the effective bandwidth based solution approximates the theoretical separation hypersurface with a hyperplane, which is not the optimal solution but it realizes fast call admission decisions.

A promising way to calculate the effective bandwidth values (κ_j), is the Chernoff bound, which upper bounds the tail distribution. This means that we shall calculate the logarithmic moment generating function (LMGF) $M_j(s)$, of a class j random variable, which transforms (5) to:

$$\inf_s \left[\sum_{j=1}^{J} N_j M_j(s) - sC \right] \le e^{-\gamma}. \qquad (7)$$

The main advantage of the Chernoff bound (Kelly, 1991), (Bucklew, 1990) lies in the optimization parameter s, which allows to find the tightest upper bound for the tail.

After that a network state $(N_1, N_2, \ldots, N_J) \in Z_+^J$ is admissible if it satisfies (7).

In case of effective bandwidth concept a tangent hyperplane at the point $(N_1^*, N_2^*, ..., N_J^*)$ on the theoretical separation hypersurface can be used to approximate the theoretical hypersurface. To determine the coordinates of this point one needs to calculate the optimal value of the s parameter

$$\arg\inf_s \left[\sum_{j=1}^{J} N_j^* M_j(s) - sC \right] = s^*. \qquad (8)$$

Unfortunately to find the optimum value for parameter s according to (8) requires a large amount of computation as well, because it assumes the knowledge of separation surface; but it have to be determined once before starting the service and during the operation we shall only decide whether the inequality (6) remains valid or not.

5 DYNAMIC CHERNOFF BOUND BASED CAC

In order to provide real-time optimization which results faster and more effective CAC decision we developed a Chernoff bound based call admission control algorithm.

The following lemma has already proved in (Billingsley, 1986):

Let X be a random variable which has expectancy value. For all $C > 0$ and $s > 0$ real number it is true the following inequality holds:

$$P(X > C) < e^{\ln\left(\mathrm{E}(e^{s \cdot X})\right) - s \cdot C}. \qquad (9)$$

If we can guarantee:

$$e^{\ln\left(\mathrm{E}(e^{s \cdot X})\right) - s \cdot C} < e^{-\gamma}, \qquad (10)$$

then the call can be admissible according to (9). γ refers to the QoS parameter. We take the natural logarithm of both sides of (10) and restructure it:

$$\ln\left(\mathrm{E}(e^{s \cdot X})\right) - s \cdot C + \gamma < 0 \qquad (11)$$

On the left hand side we have the $e^{s \cdot X}$ random variable logarithmic moment generating function, in case of number of J independent source in further simplicity it will be shown as follow (Imre, 1999)

$$\Psi_{X_J}(s) := \ln\left(\mathrm{E}(e^{s \cdot X_J})\right), \qquad (12)$$

where X is indicating the distribution function of the total traffic for J piece of source.

For the use of the inequality (11) we have to solve the following problems.

- We have to determine $\Psi_{X_J}(s)$ for the overall traffic.
- A new method has to be find, which can determine the optimal value s_J^{opt} of parameter s.

5.1 Determination of the total traffic logarithmic moment generating function

This is known for independent random variables:

$$\mathrm{E}(e^{s \cdot X_J}) = \prod_{j=1}^{J} \mathrm{E}(e^{s \cdot X_j}), \qquad (13)$$

so inequality (9) can be transformed into the following form:

$$\Psi_{X_J}(s) = \ln\left(\mathrm{E}(e^{s \cdot X_J})\right) = \ln\left(\prod_{j=1}^{J} \mathrm{E}(e^{s \cdot X_j})\right) = \sum_{j=1}^{J} \ln\left(\mathrm{E}(e^{s \cdot X_j})\right), \qquad (14)$$

which is equal with use of:

$$\Psi_{X_J}(s) = \sum_{j=1}^{J} \Psi_{X_j}(s). \qquad (15)$$

This is a very important result, as the convolution is converted to a sum which is a more simple operation.

We model the users as ON/OFF sources with parameters (m_j, h_j); hence the logarithmic moment generating function is given:

$$\Psi_j(s) := K_{X_j^{ON/OFF}}(s) = \ln\left(1 - \frac{m_j}{h_j} + \frac{m_j}{h_j} \cdot e^{s \cdot h_j}\right) \qquad (16)$$

After the comparison of (16) with the result of (11) and (15), the call admission control in case of ON/OFF sources, requires the evaluation of the following inequality:

$$K_J(s) := \sum_{j=1}^{J} N_j \ln\left(1 - \frac{m_j}{h_j} + \frac{m_j}{h_j} \cdot e^{s \cdot h_j}\right) - s \cdot C + \gamma < 0. \qquad (17)$$

5.2 Determination of the optimal value for the Chernoff-parameter in case of ON/OFF sources

Gallager shown (Gallager, 1968), that the inequality (6) $\exists s > 0$, where the $e^{\ln\left(\mathrm{E}(e^{s \cdot X_J})\right) - s \cdot C} - e^{-\gamma}$ expression is minimal and the *s* value is possible to determine with derivation if $\max X_J > C > \mathrm{E}(X_J)$.

From this result for the inequality (11) $\exists\, s_J^{opt}$ if $\sum_{j=1}^{J} h_j > C > \sum_{j=1}^{J} m_j$.

To determine s_J^{opt} we have to make the derivation of expression (13)

$$K_J'(s) = \sum_{j=1}^{J} \frac{m_j \cdot e^{s \cdot h_j}}{1 - \frac{m_j}{h_j} + \frac{m_j}{h_j} \cdot e^{s \cdot h_j}} - C = 0. \qquad (18)$$

Unfortunately s_J^{opt} cannot be calculated explicitly from (18), and Gallager's lemma also does not give any kind of solution to determine s_J^{opt}. But it can be seen that $K_J'(s)$ is a strictly monotonous increasing function of $s > 0$. So s_J^{opt}, which can be found at the point of intersections $K_J'(s)$. And axis *s*, can be determined by the means of logarithmic search algorithm.

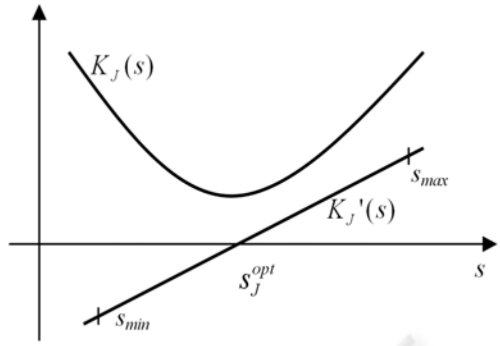

Figure 2: Determination of s_J^{opt}.

Operation of the logarithmic search algorithm:

we have to define the $[S_J^{lower}(i=0), S_J^{upper}(i=0)]$ interval which contains the s_J^{opt}, where *i* is the number of iteration steps.

$[S_J^{lower}(i=0), S_J^{upper}(i=0)]$ interval has to be mediated:

$$s_J^{median}(i) = S_J^{lower}(i) + \frac{S_J^{lower}(i) - S_J^{upper}(i)}{2}$$

we have to calculate $K_J(s_J^{median}(i))$

$$\Delta = \left| K_J(s_J^{median}(i)) - K_J(s_J^{median}(i-1)) \right|$$

if $\Delta \le d$ then $s_J^{median}(i)$ approximated s_J^{opt} satisfactorily and the algorithm stops running. *d* refers to the threshold of stop

if $\Delta > d$ then we have to calculate the value of $K_J'(s = s_J^{median}(i))$

if $K_J'(s = s_J^{median}(i))$>0, then s_J^{opt} is under the $s_J^{median}(i)$ because of $K'(s)$ being strictly monotonous; we have to choose $S_J^{upper}(i+1) = s_J^{median}(i)$, $S_J^{lower}(i+1) = S_J^{lower}(i)$

if $K_J'(s = s_J^{median}(i))$<0, then s_J^{opt} is above $s_J^{median}(i)$, we have to choose $S_J^{lower}(i+1) = s_J^{median}(i)$, $S_J^{upper}(i+1) = S_J^{upper}(i)$

we have to return to the second step.

The logarithmic search gives a fast algorithm with a small number of iterations, which can be run before every new CAC decision (Imre, 1999).

6 CAC MODEL FOR CDMA NETWORKS

In this section we show how to model CDMA cellular networks from CAC point of view.

The individual sources are grouped into traffic classes. Each user belonging to the same class has the same mean bit rate (*m*) and peak bit rate (*h*). The users are described with its required signal to interference density ratio (SIDR) as well

$$\alpha = \frac{E_b}{I_0}, \tag{19}$$

where E_b refers to the bit energy and I_o is the power-density of interference. Furthermore the sensitivity of the mobile receivers can be characterized with

$$\Gamma_k = R_k \cdot \frac{E_b}{I_0}, \tag{20}$$

which has Hz dimension and R_k refers to the peak bit rate of the kth class.

In CDMA systems the number of communicating subscribers in a moment is limited by the current SIDR.

Consider the network from the perspective of a given cell. The mobiles in this examined cell generate the interference to the receiver and the mobiles positioned in the neighbouring cells also generate it. We are taking account of three different types of cells depending on the distance from the Base Station. The first type of cell is the originated cell. The second type of cell is the neighbour cell of the originated cell and the third type of cell is the second neighbour of the originated cell (Evans and Everitt, 1999). The major part of the interference is generated at cell type 1, 2 and 3, the effect of that further ones are negligible.

6.1 Chernoff parameter transformation

The Chernoff bound calculation requires two parameters called *m* and *h*. In order to determine these parameters we use two-ways propagation model [9] for simplicity reason, as it is easy to count and implement. From (Evans and Everitt, 1999) we got the following parameters:

Type 1 (own cell):

$$h_j = \Gamma_k \cdot \left(\frac{l_1 \cdot l_2}{\left(\frac{2}{3} r_c\right)^2} \right)^2 \tag{21}$$

Type 2 (first neighbour cells):

$$h_j = \Gamma_k \cdot \left(\frac{l_1 \cdot l_2}{(r_c \sqrt{3})^2} \right)^2 \tag{22}$$

Type 3 (second neighbour cell):

$$h_j = \Gamma_k \cdot \left(\frac{l_1 \cdot l_2}{\left(\frac{(3r_c + 2r_c\sqrt{3})}{2} \right)^2} \right)^2 \tag{23}$$

Furthermore for each types of cells

$$m_j = h_j \cdot VAF_k. \tag{24}$$

We create virtual new classes, which have to be taken account by the call admission algorithm during call admission procedure

7 SIMULATION RESULTS

We implemented a simulator program (in C++) implementing the two call admission control algorithms.

At first we show the separation surface (in 3D) of a real network.

The network parameters are:

C=2.5 MHz; γ=2;

m_1=20 kbps; $m_{2=}$10 kbps; $m_{3=}$20 kbps

h_1=50 kbps; $h_{2=}$100 kbps; $h_{2=}$200 kbps

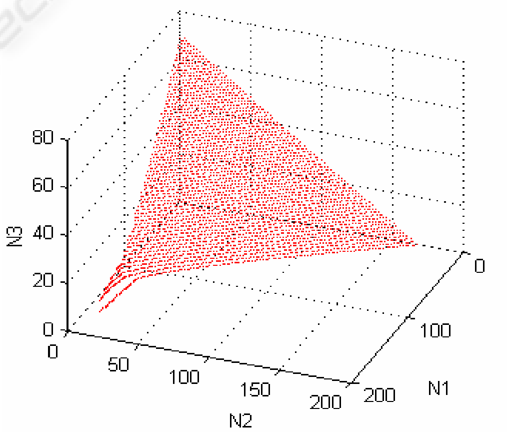

Figure 3: Separation surface in 3D.

By the second simulation (Figure 4) we watch the number of acceptable vectors by growing subscriber number in the neighbour cells:

The network parameters:

C=2.5 Mhz; γ=2;

h_1 = 50 kbps; h_2 = 100 kbps;

m_1 = 20 kbps; m_2 = 10 kbps;

E_b/I_0=7 dB; l_1=30 m; l_2=1 m; r_c=50 m.

Number of calls: 1000

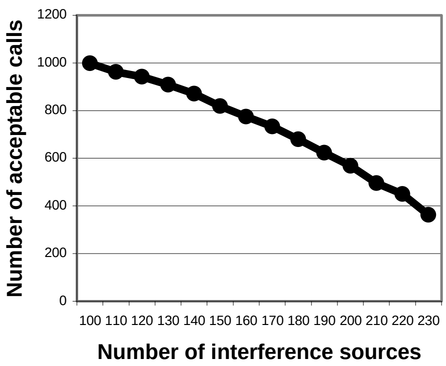

Figure 4: Acceptable calls by changing the number of interference sources.

We can see that as the number of subscribers is growing in the neighbour cells, the interference increases, that's why the number of the acceptable calls falls off.

We illustrate the differences between the two call admission control methods at the next three figures. Continuous line denotes the results of the dynamic call admission control. Dotted line signs the outcome of the effective bandwidth based method.

At the first case we can observe the formation of the acceptable calls in function of the system capacity (C).

System parameters are:

$\lambda_1 = 0.1$; $\mu_1 = 0.005$; $\lambda_2 = 0.2$; $\mu_2 = 0.002$;

$h_1 = 50$ kbps; $h_2 = 100$ kbps;

$m_1 = 20$ kbps; $m_2 = 10$ kbps; $e^{-\gamma} = 10^{-2}$

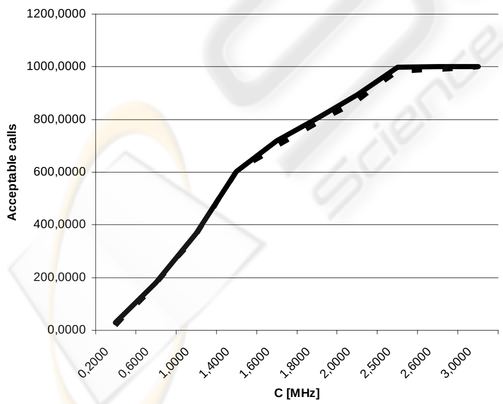

Figure 5: Number of acceptable calls by changing of C.

It's visible, that the results of the dynamic call admission control method are always higher than the values of the other method as C grows.

In the second case we changed the peak bit rate of the users in the first traffic class, and observed the formation of the number of acceptable call requests.

System parameters:

$\lambda_1 = 0.1$; $\mu_1 = 0.005$; $\lambda_2 = 0.2$; $\mu_2 = 0.002$;

C=2,5 Mhz

h_1 = **changing**; $h_2 = 100$ kbps; $m_1 = 20$ kbps; $m_2 = 10$ kbps;

$e^{-\gamma} = 10^{-2}$

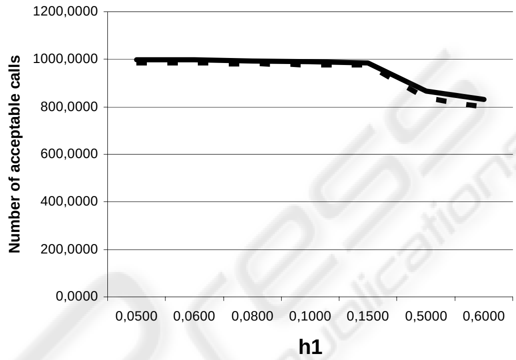

Figure 6: The number of acceptable calls by the changing of h_1.

We can observe like in the previous case, the dynamic method provides better and better results as the traffic is more and more bursty.

By the last comparison simulation we changed the time length between the call requests, which decreases as λ_2 grows.

System parameters are:

$\lambda_1 = 0.1$; $\mu_1 = 0.005$; λ_2 = **changing**; $\mu_2 = 0.002$;

C=2,5 Mhz;

$h_1 = 50$ kbps; $h_2 = 100$ kbps; $m_1 = 20$ kbps; $m_2 = 10$ kbps;

$e^{-\gamma} = 10^{-2}$.

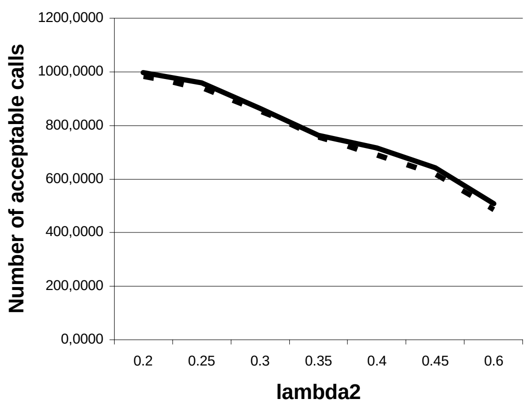

Figure 7: Number of acceptable calls by the changing of λ_2.

The increase of λ_2 means, when the average call time lengths (in the second traffic class) are decreasing, we can observe that call requests arrive more often, thus the number of the acceptable call requests is smaller. The dynamic call admission control gives more acceptable calls in this case, too.

By the running of the simulator we illustrated the number of iteration steps, which are necessary to calculate the value of s_{opt}. The stop threshold (*d)* is changing.

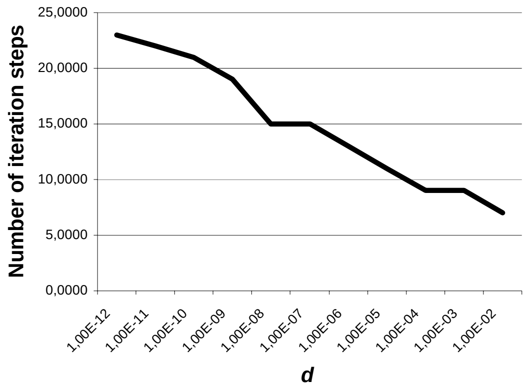

Figure 8: Number of iteration steps by different accuracy demands.

It's shown, that the value of $d = 10^{-3} - 10^{-4}$ provides a very precise approximation to the s_{opt}, in this case the number of iteration steps is only 9, what denotes we come relatively fast to the optimal value of s_{opt}.

8 CONCLUSION

In mobile communication systems we need more efficient CAC method than in the former systems, because by radio systems we can only use a very straitened and expensive resource kit (the frequency band). In the wire line networks this problem isn't important, because the capacity of the system can be extended by switching extra links. In mobile systems we can't do it.

The parameters (interference, delay etc.) of the radio channel are always changing. The effective bandwidth based algorithm (Evans and Everitt, 1999) dismisses these effects, so its efficiency isn't satisfying. We can't use the frequency band economically with this method. In addition the CAC method described in (Evans and Everitt, 1999) can't be run real time because of its large computation demand.

We made a new CAC method, which can adopt dynamic to the always changing network parameters. We had to find a faster way for calculating of the Chernoff bound. The logarithmic search algorithm gives a fast operation to the CAC method, so it becomes able to be run every time when the system parameters are changing. So we can call our new CAC method *dynamic*. In addition the simulation results show that the dynamic call admission control method gives every time more acceptable calls than the effective bandwidth based method.

The dynamic CAC method has two advantages. It can adapt to the mobile environment, and accepts more call requests than the former methods. By the acceptance of more calls we use the very expensive radio resources more economical.

In the future can be developed the dynamic CAC method with more complex network and radio channel models, because by employing of more realistic models can we get acceptable calls, growing the efficiency of the CAC method. Our plans are to implement Rayleigh- or Rice-fading, or using other traffic models than ON/OFF sources.

REFERENCES

J.S Evans and D Everitt, Effective bandwidth-based admission control for multi-service CDMA cellular networks, IEEE Transactions on vehicular technology 48 (1) (1999) 36-46

dr. Sandor Imre, Call Admission Control Methods in ATM network, 1999. pp. 100

S. Imre, K. Hankó, P. Petrás, R. Tancsics, "Optimized Effective Bandwidth Based Admission Control for Multi-Service CDMA Cellular Networks", 10th SoftCOM2002, October 08-11, 2002, Split, Dubrovnik (Croatia), Ancona, Venice (Italy), Published at FESB-Split, ISBN 953-6114-52-6, pp. 299-304

F. P. Kelly, Effective bandwidth at multi-class queues, Queuing Systems, vol. 9, pp. 968-981, Sept. 1991

A. I. Elwalid and D. Mitra, Effective bandwidth of general Markovian traffic sources and admission control of high speed networks, IEEE/ACM Trans. Netw., vol. 1, pp 329-343, June 1993.

J. A. Bucklew, Large Deviations Techniques in Decision, Simulation and Estimation, New York: Wiley 1990.

P. Billingsley, Probability and Measure, 2nd ed. New York, Wiley, 1986.

Gallager, R. G.: Information Theory and Reliable Communication, John Wiley and Sons, 1968.

R. Kohno, R. Median, and L. B. Milstein, "Spread spectrum access methods for wireless communications", IEEE Commun. Mag, Jan 1995.