Handover Provisioning in WCDMA Systems

Saied M. Abd El-atty, Angelos N. Rouskas and Dimitrios N. Skoutas,

Dept. of Information and Communication Systems Engineering
University of the Aegean
Samos, GR-83200, GREECE
e-mail:{ssoliman, d.skoutas, arouskas}@aegean.gr

ABSTRACT

In this paper, we propose a prioritization of handover (HO) calls over new calls in WCDMA systems employing orthogonal variable spreading factor (OVSF) codes as channelization codes. The code occupancy of the system is modeled by a Markov chain and the differentiation between HO and new call is performed at the code level by introducing a "guard code" scheme. The scheme belongs to the well known family of guard channel schemes and reserves some code capacity to favor the continuation of HO calls over the new calls. As the management of the general case is intractable, we solve certain numerical instances of the problem and manage to calculate several popular performance metrics like new call blocking and HO failure probabilities and code utilization.

Keywords

OVSF, Markov Chain, Handover, WCDMA, Code reservation.

1. INTRODUCTION

In third generation mobile communication networks, wideband-code division multiple access (WCDMA) is adopted as the air interface. In WCDMA systems, orthogonal variable spreading factor (OVSF) codes are chosen as the channelization codes to separate streams of data among the mobile users. OVSF codes allow WCDMA systems to dynamically allocate a variable transmission rate to each user or connection by assigning one (single-code operation) or multiple codes (multi-code operation) with different spreading factors. In such networks both soft and hard handovers are possible depending on the type of channel used by the application. Conversational applications use dedicated channels where soft handover is possible, while streaming interactive and background services use the shared channel where only hard handover is possible. A hard handover failure will lead to a temporary connection outage and additional delay induced in the packets waiting to be transmitted in the new cell, while a soft handover failure is also undesirable due to the signaling overhead induced during the re-establishment of the connection. As both cases lead to poor network performance and should be avoided, it is important to favor HO requests over new call request during the OVSF code allocation process.

To the best of our knowledge, the existing literature in OVSF code management focuses only on the code blocking and code assignment and reassignment issues, without distinguishing between new and HO connection requests. In this paper, we propose a "guard code" scheme for HO traffic management in WCDMA system employing OVSF codes, which prioritizes HO calls over new ones and thus improves HO failure rate due to capacity shortage compared to the new call blocking rate. The scheme introduces a code reservation threshold, which allows HO calls to use the total capacity of the OVSF code tree, but restricts the new calls to use only a part of the tree capacity. We model the code occupancy of the OVSF code tree by a Markov chain, where the arrival rates at each state depend on the existing code reservation threshold, and numerically solve the flow balance equations and compute the steady state distribution of certain problem instances. We study the behavior of the system in terms of new call blocking, HO failure probability and code utilization under different traffic loads and different code reservation thresholds.

The remainder of the paper is organized as follows: In Section 2 we make a brief presentation of the OVSF channelization codes as they are utilized by 3G WCDMA systems. The parameters of the employed system model are presented at Section 3. The proposed guard code reservation scheme as well as the respective performance analysis is shown at Section 4. Finally, Section V presents some numerical results and the paper is concluded in Section VI.

2. BASIC CONCEPTS

In DS-CDMA systems, such as 3G systems, Orthogonal Variable Spreading Factor (OVSF) codes are employed as channelization codes in order to preserve the orthogonality among different physical channels [1]. At the transmitter, each bit of a traffic flow is multiplied with an OVSF code. Thus, the spectrum of each signal is on purpose widened in order to occupy all the available bandwidth. As a result, the transmitted signal has lower power density and a noise-like spectrum. At the receiver the despreading is performed by multiplying with the same OVSF code sequence initially used at the spreading procedure.

MOBIMEDIA 2007, August 27-29, Nafpaktos, Greece

DOI 10.4108/ICST.MOBIMEDIA2007.1591

2.1 OVSF Channelization Codes

The method that is used for the generation of the OVSF codes is described in detail in [2], [3]. OVSF codes are generated by a binary tree with L layers. Each node represents a channelization code $C_{SF,k}$, where SF is the spreading factor of the code and k is the code number, $1 \le k \le$ SF.

The higher the spreading factor the lower the transmission rate supported by a code. The spreading factors are doubled between consecutive layers, as we move from the higher (root) layer to the lower (leaf) layer. Leaf codes have the maximum spreading factor (SF_{max}) and therefore the minimum data rate, which is denoted by R. The transmission rate supported by an OVSF code with spreading factor SF is always a multiple of a power of two of the lowest available rate.

A difficulty in the assignment of an OVSF code is the orthogonality constraint. According to this constraint the assignment of a code is possible, if only if none of its ancestor codes and none of its descendant are already occupied. Once a code is assigned, all of its ancestors, as well as all of its descendants are blocked and can be used after the code is released. Hence, at the 4-layer OVSF code tree shown at Figure 1, the assignment of code $C_{4,1}$ immediately blocks its ancestor codes $C_{2,1}$, $C_{1,1}$, and its descendant codes $C_{8,1}$, and $C_{8,2}$.

2.2 Code Blocking and Capacity Blocking

The blocking of a new incoming call can be characterized either as capacity blocking or as code blocking. If the remaining OVSF code tree capacity is lower than the rate requirement of an incoming call then the latter will be blocked and we will refer to this case as capacity blocking. For example at the OVSF code tree of Figure 1 the available capacity is 5R since the free codes $C_{8,6}$, $C_{4,2}$ and $C_{4,4}$ support transmission rates of 1R ,2R and 2R respectively. Therefore, if an incoming call requests a rate of 8R then the call will be blocked due to the lack of capacity.

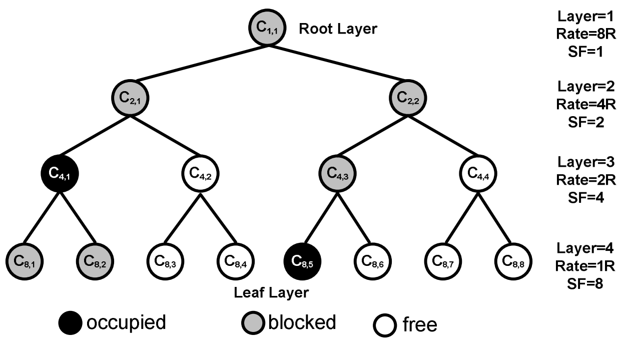

Figure 1 The OVSF code tree.

However, capacity shortage is not the only cause of call blocking. As a result of the statistical nature of the arrival and departure processes the occupied OVSF codes are randomly spread across the code tree during system operation. Consequently, the capacity of the code tree becomes fragmented and that in turn causes code blocking. Code blocking is defined as the condition that a new call cannot be supported although the system has enough capacity to support the rate requirement of the call [4]. For example, as we previously mentioned, the available capacity of the OVSF code tree at Figure 1 is 5R. However, if an incoming call requests a rate of 4R will be blocked as the codes $C_{2,1}$ and $C_{2,2}$ which support rate 4R are both blocked by lower layer occupied codes.

A first possible countermeasure for the alleviation of code blocking is the clever selection among possible candidate codes during the assignment procedure. Nevertheless, as a consequence of the statistical nature of the departure process, the complete elimination of code blocking is accomplished only if a code reassignment procedure is employed [5], [6]. A reassignment procedure reallocates ongoing calls to other codes so that a new call can always be supported if the system has adequate free capacity to support the requested rate.

3. SYSTEM MODEL

We consider a WCDMA system which supports multi-rate transmission rate by employing an OVSF code tree with L layers. The codes at the leaf layer have the maximum SF (SF_{max}) and consequently the lowest possible transmission rate R, while codes at some layer with spreading factor SF have transmission rate $2^{(SF_{\max}/SF)} \cdot R$.

We denote the code *capacity occupancy state* (COS) by c, $0 \le c \le C_T$, where C_T is the total capacity of the OVSF code tree and is equal to the number of leaf layer codes. We assume that two types of calls, namely new and HO calls, share the total system capacity C_T and that code blocking is completely eliminated by some code reassignment procedure so that an incoming request, either new or HO, is blocked only due to capacity shortage. A call is assigned a code of rate kR, where $k = 2^{(SF_{\max}/SF)}$ is a power of two as above and SF is the requested spreading factor.

Finally, without loss of generality, we assume that the only possible requested rates are 8R, 4R, 2R and 1R. The HO load is assumed to be known in advance by the measurements collected in the NodeB station. For simplicity of the model analysis, we assume that the distribution of the arrival traffic is Poisson and denote by $\lambda_{n,k}$ and $\lambda_{h,k}$ the arrival rates of new and HO calls, respectively, requesting a rate of kR.

We also assume that the mean service time and the cell residence time of a call follow negative exponential distributions and therefore the average code holding time for a call with transmission rate kR is also exponentially distributed with mean of $1/\mu_k$.

4. RELATED WORK

Many reservation schemes have been proposed in the literature in order to alleviate the HO dropping rate and favor HO calls over new ones. At [7], the authors propose a Dynamic Multiple Threshold Bandwidth Reservation (DMTBR) scheme, which employs multiple bandwidth reservation thresholds in order to provide QoS provisioning in wireless multimedia networks. At [8], a virtual guard channel (VGC) scheme for handoff calls is evaluated in comparison with the conventional guard channel (GC) scheme. At [9] the performance of a handoff scheme with guard channels in ATM-based mobile networks is considered. The handoff guard channels model together with DOVE (Delay Of Voice End-user) algorithm is proposed at [10]. The aim of this scheme is to preserve the QoS of handoff calls while at the same time to guarantee the QoS of new calls. A multi-level dynamic guard channels (MLDGC) scheme which aims to efficiently enhance the admission priority for handoff calls is proposed at [11]. Finally, at [12] a scheme based on the fractional guard channel policy is proposed and its performance is evaluated via simulation.

Concluding, the main idea behind all the guard channel schemes is the reservation of additional bandwidth for the HO calls from the total available capacity. This is achieved by prohibiting the acceptance of new calls if the channel utilization has superseded some threshold value. In 3G WCDMA systems the bandwidth allocation to the connections is achieved through the use of the OVSF codes. Hence, any guard scheme proposed for 3G systems should be aware of the OVSF code tree characteristics. Therefore, although the scheme which we propose belongs to the well known family of guard channel schemes it differs from other previously proposed schemes since it is adapted to the features of the OVSF code tree.

5. THE GUARD CODE RESERVATION SCHEME

As we already mentioned, the failure of an incoming HO call yields poor network performance in the case of soft HO, and deterioration of the QoS perceived by the user, in the case of hard HO. Thus, the failure of a HO attempt due to code capacity shortage is less acceptable compared to the blocking of a new call due to the same reason. As we previously mentioned the most well known methods which are employed in order to alleviate the HO dropping rate are the guard channel reservation schemes which are easy to implement and manage.

In our analysis, we consider OVSF codes as the precious resource at the downlink of WCDMA systems and apply a *guard code* reservation scheme to favor incoming HO calls over new calls. As long as the available code capacity is not lower than a threshold value C_H, the BS allocates codes to either new or HO calls. If, however, the available capacity falls below C_H, only HO calls are accepted and new calls are blocked. Assuming that the HO load is known by the measurements collected at the base station, the base station can dynamically adapt the threshold value in the guard code reservation scheme.

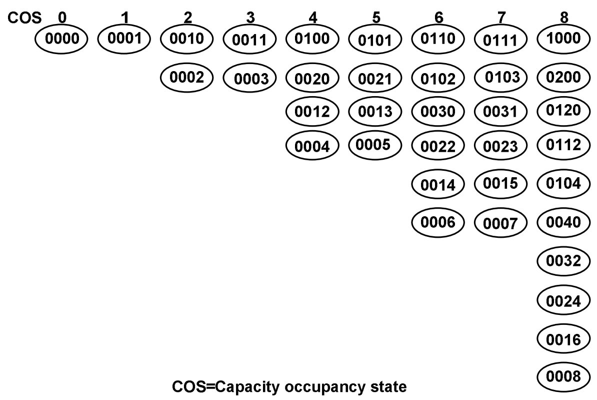

Figure 2 Code occupancy state space when C_T=8.

5.1 System State Space

We denote the system code occupancy state by $(x_8\, x_4\, x_2\, x_1)$, where x_8, x_4, x_2, x_1 are integers denoting the number of occupied codes of rate 8*R*, 4*R*, 2*R* and 1*R*, respectively, such that $8x_8+4x_4+2x_2+x_1=c$. For example the system state (0012) corresponds to code occupancy *c*=4 with no codes of rate 8*R* and 4*R*, one code of rate 2*R* and two codes of rate 1*R* occupied. Depending on the possible transmission rates in the system, we can obtain all the feasible occupancy states in the OVSF code tree by using a generation function as in [5]. For example, all the possible states in an OVSF code tree with C_T=8 are 36 as shown in Figure 2. When the OVSF code tree size increases the total number of states will increase dramatically. For example, at C_T =16, total number of states=201 states, while at C_T =32, total number of states=1625 states [5].

5.2 Performance Analysis

The code occupancy of the OVSF code tree can be modeled by a Markov chain with multiple transitions among the feasible states. Figure 3 shows only a small part of the Markov chain, illustrating all the feasible transitions into and out of state (0112), when C_T=16 and C_H=6. As the used code capacity is *c*=6, the available capacity is C_T - *c*=10 and the threshold is C_H=6, no new calls of rate 8*R* and 4*R* are allowed in the system and the corresponding transitions from this state with arrival rates $\lambda_{n,8}$ and $\lambda_{n,4}$ are absent from the transition diagram. The flow balance equation for the same state is given in the following:

$$
\begin{aligned}
&P_{0112}\cdot\left(2\mu_1+\mu_2+\mu_4+\lambda_{n,1}+\lambda_{h,1}+\lambda_{n,2}+\lambda_{h,2}+\lambda_{h,4}+\lambda_{h,8}\right)=\\
&P_{0113}\cdot\left(3\mu_1\right)+P_{0122}\cdot\left(2\mu_2\right)+P_{0212}\cdot\left(2\mu_4\right)+P_{1112}\cdot\left(\mu_8\right)+\\
&P_{0111}\cdot\left(\lambda_{n,1}+\lambda_{h,1}\right)+P_{0102}\cdot\left(\lambda_{n,2}+\lambda_{h,2}\right)+P_{0012}\cdot\left(\lambda_{n,4}+\lambda_{h,4}\right).
\end{aligned}
\quad (1)
$$

For constructing the general form of the flow balance equations the following notations are necessary:

- P_j: the steady state probability of the code capacity state $j=(x_8\, x_4\, x_2\, x_1)$.
- C_j: code capacity occupancy in state *j*.
- $n_k(j)$: the number of calls with transmission rate *kR* in state *j*.
- $j\oplus k$: the new state after accepting a call of rate *kR* into the system while in state *j*.
- $j\ominus k$: the new state after completion of a call of rate *kR*, while in state *j*.
- $I_n(j,k)$ and $I_h(j,k)$: functions that indicate the possibility of accepting a new or HO call of rate *kR* while in state *j*, such that:

$$
I_n(j,k)=\begin{cases}1 & \text{if } C_{j\oplus k}\le C_T-C_H,\\ 0 & \text{otherwise.}\end{cases}\quad \text{and}
$$

$$
I_h(j,k)=\begin{cases}1 & \text{if } C_{j\oplus k}\le C_T,\\ 0 & \text{otherwise.}\end{cases}
$$

Now we can derive the general form of the flow balance equations for any state j, as follows:

$$P_j \cdot \sum_{k=1,2,4,8} \left(n_k(j)\cdot \mu_k\right) + P_j \cdot \sum_{k=1,2,4,8} \left(I_n(j,k)\cdot \lambda_{n,k} + I_h(j,k)\cdot \lambda_{h,k}\right) = \sum_{k=1,2,4,8} \left(P_{j\oplus k} \cdot I_h(j,k)\cdot n_k(j\oplus k)\cdot \mu_k\right) + \sum_{\substack{k=1,2,4,8 \\ C_{j\ominus k}\le C_T - C_H,\, n_k(j)>0}} \left(P_{j\ominus k} \cdot \left(I_n(j\ominus k,k)\cdot \lambda_{n,k} + I_h(j\ominus k,k)\cdot \lambda_{h,k}\right)\right). \quad (2)$$

The complete set of equations includes the sum of steady-state probabilities of all code occupancy states which is equal to unity, $\sum_j P_j = 1$. The solution of the above set of equations yields the steady state probability P_j for each state j.

The probability $P_{COS}(c)$ of having a capacity occupancy c is defined as:

$$P_{COS}(c) = \sum_{\forall j: C_j = c} P_j \qquad 0 \le c \le C_T. \quad (3)$$

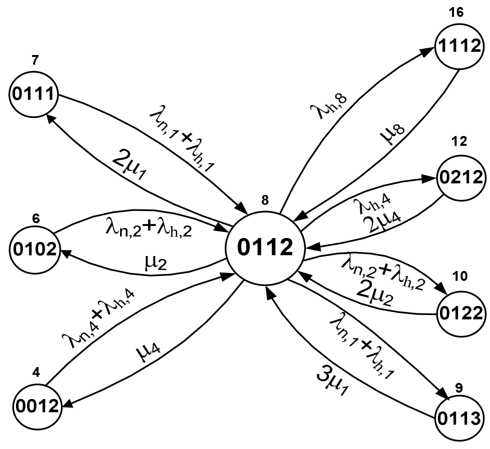

Figure 3 Markov state transitions to and out of state (0112) when C_T=16 and C_H=6.

The new call blocking probability $P_{nb}(kR)$ of the calls requesting code rate kR (k=1, 2, 4 and 8) is given by:

$$P_{nb}(kR) = \sum_{c=C_T - C_H - (k-1)}^{C_T} P_{COS}(c). \quad (4)$$

The HO failure probability $P_{hf}(kR)$ of the calls requesting rate kR is given by:

$$P_{hf}(kR) = \sum_{c=C_T-(k-1)}^{C_T} P_{COS}(c). \quad (5)$$

Another important performance measure is code utilization U_T, which is defined as the percentage of the total code capacity being utilized and can be expressed as:

$$U_T = \frac{\sum_{c=1}^{C_T} c \times P_{cos}(c)}{C_T}. \quad (6)$$

6. NUMERICAL RESULTS

The performance of the proposed guard code reservation scheme is evaluated in an OVSF code tree with total capacity C_T =16. All the possible code occupancy states of the system, as well as the flow balance equations were generated by a C++ program. Different input parameters, such as call arrival and call service rates, as well as capacity threshold values produced different instances of the problem. The equations were then fed to Mathematica to calculate the steady states probabilities equations and metrics of interest, i.e. the new call blocking and handover failure probabilities as well as code utilization. We used several thresholds values C_H=4, 6, and 8, different traffic intensities and different traffic mix of new and HO calls. Furthermore, in our study we assume two different distributions for the traffic patterns of both new and handover calls. The first distribution is (8R:4R:2R:1R) = (10:20:30:40), where the lower rate calls dominate the overall traffic volume in the system while the second distribution is uniform (8*R*:4*R*:2*R*:1*R*) = (25:25:25:25). We will refer to these two distributions as A and B respectively.

6.1 New Call Blocking and Handover Failure Probabilities

Figures 4 to 7 show the new call blocking and handover failure probabilities P_{nb} and P_{hf} as the offered traffic load increases following the traffic patterns A and B respectively. In this particular problem instance, we assumed that the HO traffic load is only a low percentage (20%) of the total offered load while the code reservation threshold is C_H=4. As illustrated in Figures 4 and 6 for both distributions the higher rate calls suffer higher blocking and dropping than the lower rate calls. However, due to the guard code scheme the HO failure rate of every call rate (Figs. 5 and 7) is significantly lower than the new call blocking probability of the corresponding call rate.

Furthermore, we can also observe that with traffic pattern B the blocking probability of new calls as well as the failure probability of handover calls is higher compared to traffic pattern A. This phenomenon is caused by the increased percentage of higher rate calls at traffic pattern B. The higher rate calls experience capacity blocking more often than lower rate calls.

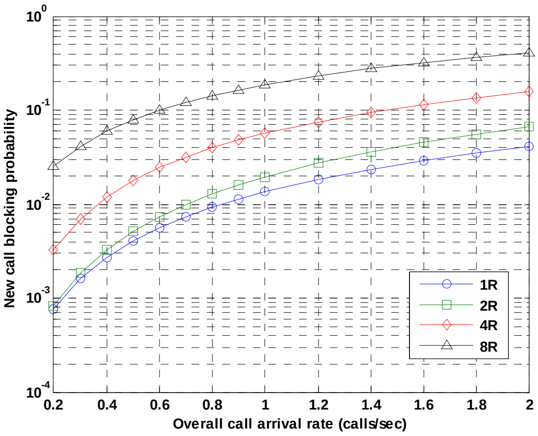

Figure 4 New call blocking probability versus overall call arrival rate (λ) at traffic pattern A.

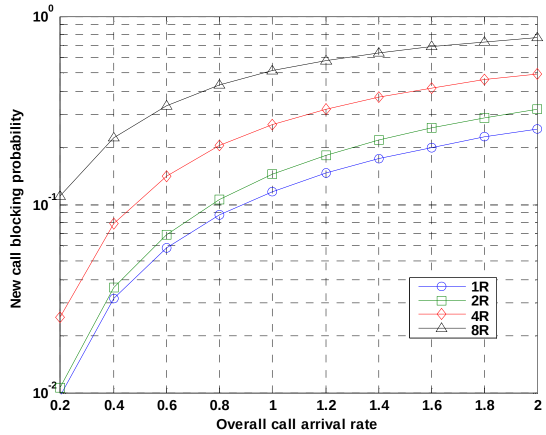

Figure 6 New call blocking probability versus overall call arrival rate (λ) at traffic pattern B.

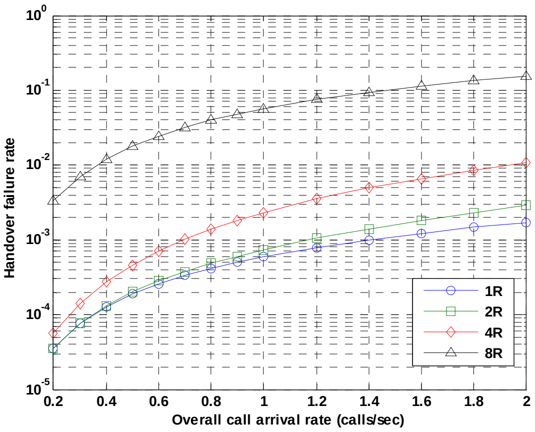

Figure 5 Handover failure rate versus overall call arrival rate (λ) at traffic pattern A.

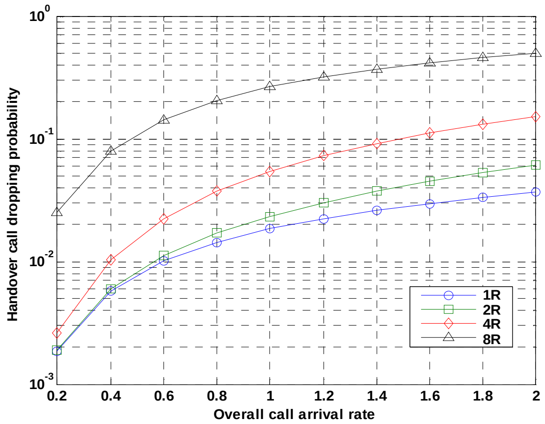

Figure 7 Handover failure rate versus overall call arrival rate (λ) at traffic pattern B.

6.2 The Capacity Threshold

The value of the code reservation threshold plays a significant role at the utilization of the available bandwidth as well as to the overall system performance. In this scenario we employ traffic pattern A and we compare in Figure 8 the code bandwidth utilization U_T at different code reservation thresholds (C_H=0, 4, 6, and 8). The case C_H =0 corresponds to the complete sharing scheme where no code capacity reservation for HO calls exists. As shown in the figure, U_T increases nearly linearly as the offered load increases and it achieves the highest values when there is no capacity reservation, as expected. Of course, this benefit comes with the cost of increased HO failure rates.

Figure 9 shows the performance of P_{hf} and P_{nb} probabilities of calls requesting rates 1*R* and 4*R* at different code capacity thresholds while the overall arrival rate λ=1call/sec.

It seems that for this certain problem instance the increase of threshold above C_H =4, does not yield any significant improvement at the HO failure rate, while cause unnecessary deterioration of the new call blocking.

7. CONCLUSION

A guard code scheme has been introduced to favor HO calls over new calls in WCDMA systems employing OVSF codes as channelization codes. The reservation process takes place at the code management level, and new calls are admitted to the system if the available capacity is above a certain threshold. The code occupancy state of the system and the transitions between the different states are modeled by a Markov chain. Due to the large number of states even for moderate total capacity values, certain numerical instances of the problem are solved and the performance metrics confirm the effectiveness of the proposed scheme.

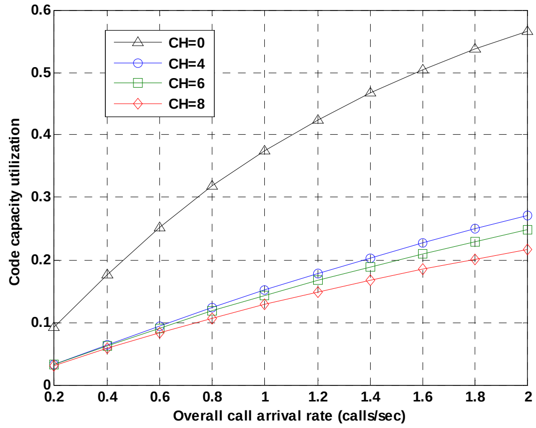

Figure 8 Code capacity utilization versus overall call arrival rate (λ).

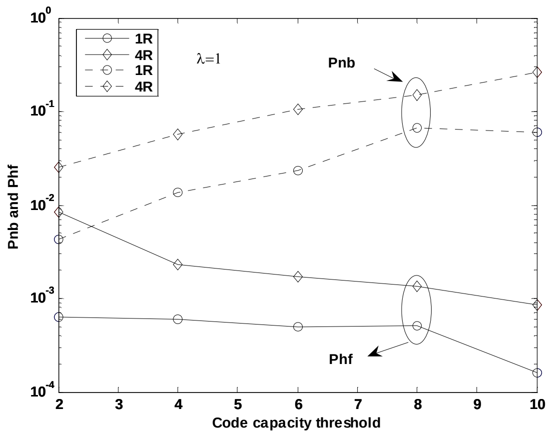

Figure 9 New call blocking versus HO failure rate at different code reservation thresholds.

8. ACKNOWLEDGMENTS

This work was supported through an ARCHIMEDES grant by the Greek Ministry of Education and Religious Affairs.

9. REFERENCES

[1] Harri Holma and Antti Toskala, WCDMA for UMTS Radio Access for Third Generation Mobile Communications, John Wiley and Sons, 2004.

[2] 3GPP, Technical Specification TS V5.2.0, Spreading and Modulation (FDD), 2002-2009.

[3] F.Adashi, M.Sawahashi, and K.Okawa, “Tree structured generation of orthogonal spreading codes with different lengths for forward link of DS-CDMA mobile radio,” Electronic Letters, vol. 33, Jan 1997, pp. 27–28.

[4] T. Minn, K.Y. Siu “Dynamic Assignment of Orthogonal Variable Spreading Factor Codes in W-CDMA”, IEEE Journal on Selected Areas in Communications, Vol. 18, No. 8, August 2000, pp. 1429-1440.

[5] Y.-C. Tseng and C.-M. Chao, “Code Placement and Replacement Strategies for Wideband CDMA OVSF Code Tree Management”, IEEE Trans. on Mobile Computing, vol. 1, no. 4, Oct.-Dec. 2002, pp. 293-302.

[6] A. Rouskas and D. Skoutas, “OVSF code assignment and reassignment at the forward link of W-CDMA 3G systems”, Proc. IEEE PIMRC, vol. 5, pp. 2404-2408, 2002

[7] X. Cheng, B. Li and Y. Fang, “A Dynamic Multiple-Threshold Bandwidth Reservation (DMTBR) Scheme for QoS Provisioning in Multimedia Wireless Networks”, IEEE Trans. Wireless Comm., vol. 4, no. 2, March 2005, pp.583-592.

[8] J. Zheng, Y. Zhang and E. Regentova, “Virtual Guard Channel for Handoff Calls in Integrated Voice/Data Wireless Networks” IEEE Comm. letter vol. 10, no.10, April 2006, pp. 263-266.

[9] Lee, S. and Park, S, “Handoff with Dynamic Guard Channels in ATM-Based Mobile Networks”, Proc 13th international Conference on information Networking. IEEE ICOIN, January 1998, pp.21 – 23.

[10] Okuda, T. Ideguchi, T. and Tian, X. "Improvement of the dropping probabilities in mobile communication networks by using guard channels protocol with delay of new user", Proc IEEE Vehicular Technology Conference, vol. 6, Sept. 2004, pp. 4558- 4562.

[11] Tung Chong Wong, Jon W. Mark and Kee Chaing Chua, "Multi-level Dynamic Guard Channels for Priority Access in Cellular Systems" NETWORKING 2005, Springer LNCS 3462, pp. 598 – 609.

[12] Ioannis C. Panoutsopoulos, Stavros Kotsopoulos and Vasilios Tountopoulos “Handover and New Call Admission Policy Optimization for G3G Systems”, Springer Wireless Networks, vol. 8, no.4, July 2002, pp. 381-389.