OVSF code slots sharing and reduction in call blocking for 3G and beyond WCDMA networks

DAVINDER S SAINI[1] , VIPIN BALYAN[2]
Department of Electronics and Communication Engineering,
Jaypee University of Information Technology,Waknaghat, Distt. Solan, Himachal Pradesh – 173215[1]
Jaypee Institute of Information Technology, Sector-128, Distt. Noida, Uttar Pradesh – 201307[2]
INDIA
davinder.saini@juit.ac.in[1]
vipin.balyan@rediffmail.com[2]

Abstract: OVSF (Orthogonal Variable Spreading Factor) codes are used in WCDMA uplink and downlink transmission for multirate traffic. The orthogonal property of these codes leads to code blocking and new call blocking. Two code (or slot) sharing assignment schemes are proposed to reduce the effect of code blocking using OVSF and NOVSF (non blocking OVSF) codes. The schemes favor the real time calls as they are given higher priority in all 3G and beyond networks. The benefit of the proposed scheme is the better handling of non quantized rates compared to other novel single code and multi code assignment schemes. Both single code and multi code options are analyzed for OVSF as well as NOVSF codes. Simulation results are discussed to show the benefits of the proposed scheme.

Key- Words: OVSF codes, slots, time sharing, multiplexing, single code assignment, multi code assignment, real time, code blocking.

1 Introduction

Two important issues under research in WCDMA systems are spreading codes [1,2] and modulation[2]. Spreading is the fundamental operation of WCDMA radio interface. The spreading codes in WCDMA are of two types namely channelization code and scrambling code. The channelization codes in WCDMA are OVSF codes. The channels in the forward link and reverse link use theses codes for transmission. OVSF codes are shorter in length and are made from orthogonal function. The orthogonality property of OVSF codes makes it suitable for WCDMA. The signals from two or more UEs in the reverse link are transmitted to same BS in the cell from separate locations. This change in distance gives rise to change in time for the signals to reach at the BS. The orthogonal property of the OVSF codes is disturbed due to different arrival times. Hence the OVSF codes are not used for calls separation in the reverse link. The facility to handle variable call rate is also incorporated in OVSF codes. In contrast to OVSF codes, scrambling codes are quite long (with the exception of the reverse link of the short scrambling code). Scrambling codes are generated from the stream called pseudo noise (PN) sequences [3]. In WCDMA the requirement is have codes with high value of autocorrelation and low value of crosscorrelation. High autocorrelation properties are desired to recover the intended signal and to reduce the effect of multipath signals. The low crosscorrelation properties are required to minimize the effect of interfering signals. The PN codes and orthogonal codes individually do not have both good autocorrelation and good crosscorelation properties. Orthogonal codes are sets of binary

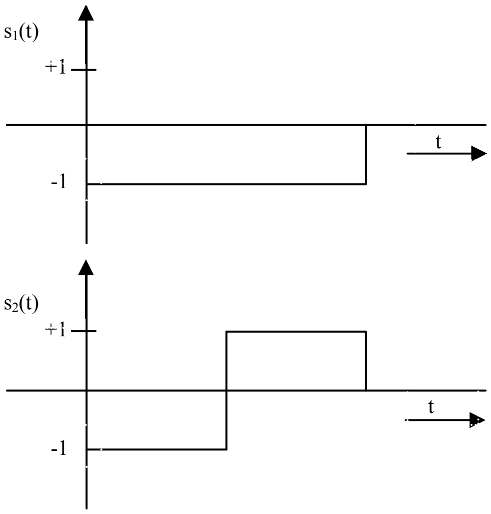


Fig.1. Orthogonal signals

sequences that have a cross-correlation coefficient equal to zero. A set of periodic signals $s_i(t)$ and $s_j(t)$ is orthogonal if

$$\frac{1}{E}\int_0^T s_i(t)s_j(t)dt = \begin{cases}0, & \mathrm{i} \neq \mathrm{j}\\ 1, & \mathrm{i} = \mathrm{j}\end{cases} \qquad (1)$$

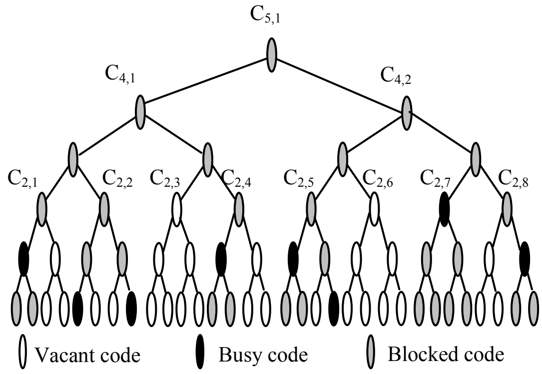
where, E is the energy of the signal given by

$$E = \int_0^T s_i^{\,2}(t)dt \qquad (2)$$

In Equation (1), T is the period of the signals $s_i(t)$ and $s_j(t)$, and E is the signal energy as defined in Equation (2). The set of signals representing the orthogonal binary sequences 00 and 01 is shown in Figure 1.

The OVSF codes designed to facilitate multirate traffic. This is because multiple spreading factor options. The OVSF codes are generated [4] using to Walsh functions. The codes with different spreading factors handle calls with different rates. One of the important properties of OVSF codes is that, when a code is used, all of its ancestors and descendants are blocked from assignment. This is required to assign orthogonal codes to each new call. The SF of a code decides the rate of the call that can be supported by an OVSF code. Lower is the position of a code in tree, higher is SF and vice versa. Once a code is assigned, all its ancestors and descendants are blocked. It limits the number of OVSF codes. So, OVSF codes should be allocated efficiently. Further, the fair allocation of codes become difficult as the scattered lower rate codes block high rate codes [5-6]. In OVSF based networks, the treatment of voice calls and data calls is different. The voice calls require single code for full call duration at fixed rate, while the data calls have the flexibility of variable data rate. The treatment of real time and on real time calls in most of the systems including GSM, CDMA etc. is different. In [7], the voice calls and real time calls are handled with either complete partitioning (CP) or partial partitioning (PP) methods. The voice traffic, the traffic model used is Erlang Law for both cases CP and PP. On the other hand for data traffic, the models depend upon the scheme used for resource allocation and affects code blocking which reduces the spectral efficiency of WCDMA system up to 25%. The code blocking appears due to two reasons namely internal fragmentation [8]and external fragmentation [8]. The internal fragmentation is due to quantized rate handling capacity of OVSF codes. The networks with non quantized calls, always gives wastage of capacity of the code it occupies. The capacity wastage is high as codes near the root are used where it can reach close to 100%. The external fragmentation is due to scattered vacant codes such that the high rate call is rejected despite the system has enough capacity to handle it. This paper describes time slot sharing method which avoids the limitation of code blocking to avoid both the limitations discussed above. The slot usage benefits are explained for tradition OVSF codes and non blocking OVSF codes where the codes in different layers are orthogonal to each other. The preference is given to handle real time (priority) calls. The remainder of the paper is organized as follows.

Section 2 discusses review of OVSF and non blocking OVSF based WCDMA systems. Section 3 gives explanation of the proposed OVSF slot sharing assignment scheme for single code and multiple code option. Section 4 describe NOVSF slot sharing scheme using single code and multiple codes. Section 5 gives simulation parameters and results to show the code blocking reduction benefits in the proposed schemes. The paper is concluded in section 6.



Layer	Bandwidth	SF
5	32*R*	1
4	16*R*	2
3	8*R*	4
2	4*R*	8
1	2*R*	16
0	1*R*	32

Fig.2. OVSF code tree with six layers

2 OVSF and non blocking OVSF Codes

2.1 OVSF Code Tree and code blocking

Consider an L layer (L=8) WCDMA OVSF code tree with layer 7 for root and layer 0 for leaves. The call data rates and spreading factor (SF) and for layer $l, l \in [0, L-1]$ are $2^l R$ (R=15 *kbps*) and 2^{L-l} respectively. Figure 2 shows OVSF code tree with six layers. There are three possible statuses of the codes as explained below,

Busy code: Code assigned to a call (marked black in Figure 2).

Blocked code: Code blocked (marked gray in Figure 2) due to busy ancestor or descendant.

Vacant code: Unoccupied code (marked white in Figure 2) can be assigned to future calls.

A code in a layer is orthogonal to codes in the same layer and all the codes in different branches, but not to the codes in the same branch. As discussed, if a code is assigned to the incoming call, its ancestors and descendants are blocked from assignment. In Figure 2, code $C_{2,7}$ is being occupied by the call. Therefore all its ancestors and descendants are blocked from assignment. The data rate handled varies from R, $2R$,…, $32R$ with corresponding code SF =32,16,…,1. The overall bandwidth transmitted is always $32R$. To make transmitted bandwidth same irrespective of input calls with rate R, $2R$, …, $32R$ require SF= 32,16,….,1. The maximum capacity of the tree is $32R$ out of which $15R$ (sum of bandwidth of black nodes) is currently used. So the maximum available capacity is $17R$. If a call with rate $8R$ wants vacant code, it will not be handled by the system because there is no vacant code with capacity $8R$. We come to the situation in which a call is not handled even the system have enough capacity to handle it. This is called code blocking which leads to call blocking. There is large number of code assignment schemes proposed in literature to reduce code blocking.

2.2 Non blocking OVSF (NOVSF) codes

The non blocking OVSF (NOVSF) codes [9] provide zero code blocking. There are two categories of NOVSF codes. The first category of codes employ time multiplexing giving different incoming call rates the flexibility of using time slots from a single layer only. The layer used and number of slots per code depends on the type of call and majority. In the second category, the structure of the non blocking OVSF code tree is exactly same as that of OVSF code tree. The codes in different layers are orthogonal to each other (unlike traditional OVSF codes). The SF of the root can be chosen according to the type of wireless network. The paper aims to provide time slot usage description for two scenarios: (i) when all calls are treated similarly (ii) some calls are given higher priority.

2.3 Existing code assignment schemes

The code assignment schemes which do not incorporate code sharing can be categorized into single code and multi code assignment schemes. The single code assignment scheme used single code to handle incoming calls [10]. The leftmost code assignment (LCA) [11], crowded first assignment (CFA) [11], fixed set partitioning (FSP) [12] and dynamic code assignment (DCA) [13] are single code assignment schemes. In the leftmost code assignment scheme, the code assignment is done from the left side of the code tree. In crowded first assignment, the code is assigned to a new call such that the availability of vacant higher rate codes in future is more. In the fixed set partitioning, the code tree is divided into a number of sub trees according to the input traffic distribution. In dynamic code assignment scheme, the blocking probability is reduced using reassignments based on the cost function. The DCA scheme requires extra information to be transmitted to inform the receiver about code reassignments. The incoming calls are converted into quantized rates (if not so). A fast OVSF code assignment design is given in [14] which aims to reduce number of codes searched with optimal/suboptimal code blocking. The code assignment scheme uses those vacant codes whose parents are already blocked. This leads to occurrence of more vacant codes in groups, which ultimately leads to less code blocking for higher rate calls. The integration of calls is done in [15] for allocation of OVSF codes when a quantized or non-quantized call arrives, and further, the voice calls and data calls are treated differently as former are delay sensitive and later can be stored in buffer.The single code assignment schemes are simpler, cost effective and require single rake combiner at the BS/UE. The multi code assignment schemes use multiple codes in the OVSF code tree and hence multiple rake combiners to handle single call. It reduces code blocking compared to single code assignment schemes but the cost and complexity is more[16,17,18]. The multi-code assignment schemes provide additional benefit of handling non-quantized (not in the form of $2^l R$, $l, l \in [0, L-1]$ data rates. The DCA scheme with different QoS requirements is given in [19]. The performance of fixed and dynamic code assignment schemes with blocking probability constraint is given in [20]. The throughput performance is proved to be better in this paper. The multi rate multi code compact assignment (MMCA) [21] scheme uses the concept of compact index to accommodate QoS differentiated mobile terminals. The maximally flexible assignment scheme [22] discusses two code assignments namely rearrangeable code assignment and non rearrangeable code assignment schemes. It define flexibility index to measure the capability of assignable code set. Both schemes provide the maximal flexibility for the code tree after each code assignments.

The code/slot sharing in all above schemes further reduces the code blocking. The code assignment scheme in [23] uses the higher layer code sharing on the periodic basis. This is significant requirement for calls with large peak to average ratio of data rates. The code utilization is increased due to code sharing. Using idle/unused code slots for bursty traffic, the fairness and per connection data rate guarantee is achieved. The code assignment scheme discussed in [24] describes three scheduling algorithms to support multimedia transmissions. The credit management and compensation mechanism provides rate guarantee and fair access to mobile terminals. Each call is allowed to use multiple codes in time sharing manner. The channel sensitive algorithm is given to use multiple codes when there are channel errors leading to bad signals. For bad quality transmission, codes with higher spreading factors are used. It defines shared count of ancestors to use the code from partially shared area or fully shared area. Zero wastage designs in [25] are defined in which the rake combiner's usage is made dynamic and the amount of rake combiners used depend upon the rate type, with more combiners given to the rate which deviates significantly from the quantized one. The average number of rakes per call is arbitrarily assumed, and if the rakes used for a particular call are less than the average (which happens for quantized or near quantized calls), the unused rakes can be used by future calls. The performance is significantly improved compared to the fixed rate systems. The amount of codes used on

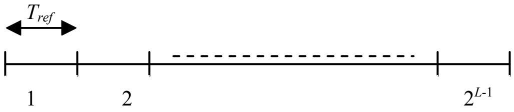


Fig.3 Total time slots in any one layer

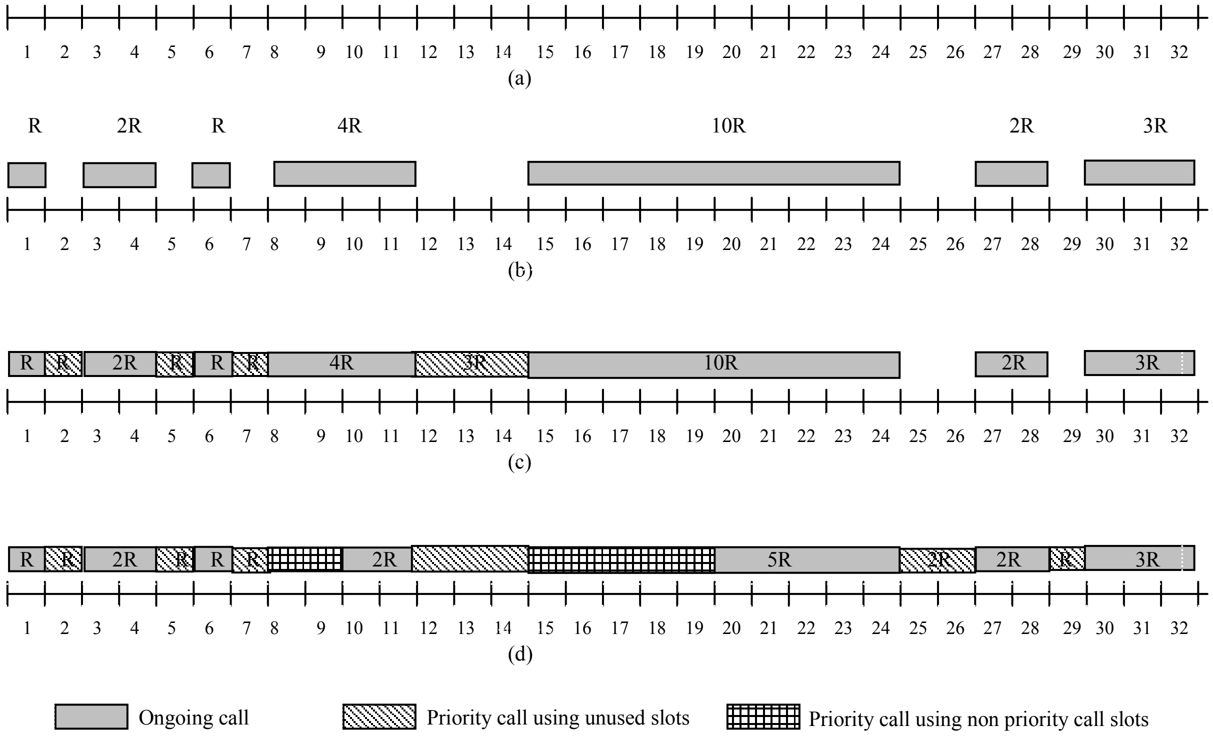


Fig.4. Illustration of slot usage in a 32 OVSF CDMA system. All calls except 4R and 6R are non priority calls. (*a*) total vacant codes available (*b*) Status of the slots at the arrival of 4*R* priority call (*c*) Status of the slots after handling 4R call when the system has one rake (*d*) Status of the slots after handling 4R call when the system has two rakes (*e*) Status of the slots after handling 6R call when the system has two rakes

average is less than the codes required for existing multi code designs.

3 OVSF based time slots sharing

For a layer *l*, the number of codes in OVSF code tree is 2^{L-l} and number of slots per code is 2^{l}. It is sufficient to use any one layer from the code tree for time slot sharing with total 2^{L} vacant slots. For example, if layer 4 is used, there are 8 orthogonal codes in the OVSF code tree with 16 time slots per code making total slots equal to 2^{L-1} (as shown in Figure 3). The slots in a layer are denoted by S_x, where x is the slot number. The identifier S_x does not depend on layer number *l* as any layer can be opted.

3.1 Slots sharing with single code

Two call scenarios are considered, (*i*) general case: all the calls are treated similarly (ii) some layer calls are given higher priority.

3.1.1 General case: All calls are treated similarly

For a new call *kR*, if *k* consecutive vacant slots are available, the call is handled. Otherwise the call is rejected. The call rejection is only due to insufficient slots.

3.1.2 Priority and non priority calls coexist

The calls are divided into two categories, namely, priority calls (*P*-calls) and non priority calls (*NP*-calls). When a *NP*-call arrives, the slots are assigned according to the procedure discussed earlier. On the other hand, if *P*-call with rate *kR* arrives, the availability of *k* vacant slots is checked, and if the vacant codes are available call is handled as discussed earlier. If *k* vacant slots are not available, the call can still be handled by allocating some slots currently assigned to *NP*-calls. Thus, the rate of the *NP*-call is reduced at the cost of handling *P*-call. If at the arrival of a new *P*-call of rate *kR*, there are N_p ongoing *P*-calls and N_n ongoing *NP*-calls with corresponding used slots $S_p^i, 1 \le i \le N_p$ and $S_n^j, 1 \le j \le N_n$ respectively, the algorithm identifies the optimum *NP*-call whose some slots can be assigned to the new call. In our design we use single level sharing where only one *P*-call can share the slots of a *NP*-call . It may happen that none of existing *NP*-calls has enough slots and therefore, the slots from two or more ongoing calls are required. For a *P*-call of *kR* rate, if the maximum vacant slots available are $k', k' < k$ and there are some *NP*-ongoing calls with at least $2(k\text{-}k')$ slots, the *P*-call call can be handled with level-1 sharing. However, it will increase the time duration of ongoing *NP*-call. If rates of *NP*-calls are denoted by vector $R^n = [R_1^n, R_2^n, ..., R_{N_n}^n]$ from left to right of the code tree, optimum *NP*-call is identified by $R_{i-opt}^n = \max(R_i^n), i \in [1, N_n]$ and R_{i-opt}^n is use to handle *P*-call. The requirement of 2(*k*-*k'*) slots of an optimum *NP*-call guarantee that the existing *NP* calls has at least half the slots available for its own after giving $(k\text{-}k')$ slots to *P*-call. This maintains fairness of capacity distribution between *P*-call and *NP*-call. For a i^{th} *NP* call with rate R_i, if there are total *S* slots out of which S' slots are given to a *P*-call, the new rate of *NP* call becomes $R_i \times (S - S')/S$.

3.2 Multi code assignment

3.2.1 General case: All calls are treated similarly

Same as discussed in section 3.1.1.

3.2.2 One or more calls are given higher priority

If a priority call with *kR* rate arrives, find the slot availability as explained in section. If *r* groups with sufficient vacant slots are not available, the slots from non priority calls are used. Let fraction $k'R \mid k' < k$ is handled by *q*|*q*<*r* groups of vacant slots, the rate portion $(k - k')R$ has to be handled by non priority busy call slots. The selection of busy code depends on number of available rakes and number of rakes used will increase level sharing. One rake is used to handle vacant slots group and one rake for each *NP* call whose slots are used.

Level-1 sharing: number of rakes used -2.
Level-2 sharing: number of rakes used -3 and so on.

For remaining $(k - k')$ slots, arrange *NP*-calls in descending order of slots assign to them. If a *NP*-call of maximum slots (say) k_1 with $k_1 \ge 2 \times (k - k')$ exists, new call of rate *kR* is handled using k' vacant slots and $(k - k')$ slots of *NP*-call. This is level-1 sharing utilizing two rakes. If a call rate is not handled and available rakes are more than used rakes, increase level sharing with using more *NP*-calls till $\sum_{i=1}^{m} k_i = (k - k')$, where *m* denotes *NP*-calls whose slots are used, utilizing *m*+1 rakes.

Consider a 32 slots OVSF CDMA system as shown in Figure 4(*a*), where all the users except 4*R* are non priority users. For a particular time, let the status of the code slots is shown in Figure 4(*b*). If a new 6*R* priority call arrives, availability of four vacant slots is checked. As shown in (*c*) there are 3*R* rate vacant slots between 12-14 position and three *R* rate slots available at 2, 5 and 7 positions respectively. This call will be handled without sharing. If a new call of rate 10*R* arrives, vacant slots are not enough to support it. The assigned slots to *NP*-calls are arranged in descending order and half of the slots (required slots) of these calls are assigned to new call. As shown in (*d*), 3*R* available vacant slots and remaining 7*R* is provided by ongoing *NP*-calls slots arranged in descending order $10R, 4R, 3R, 2R, 2R, R$ and R. The call is handled using 3 vacant slots and 5*R*, 2*R* and *R* slots of 10*R,* 4*R* and *R* rate call respectively.

4 NOVSF based time slots sharing

4.1 Single code assignment

Consider a NOVSF CDMA system with *L* (layer numbers 0 to *L*-1) layers in the code tree. The code in layer *l*, $l \in [0, L-1]$ is represented by $C_{l,x}$, $1 \le x \le 2^{L-l}$. The time frame for code $C_{l,x}$ can be represented as sum of 2^l fixed slots of width T_{ref}, where T_{ref} represents the duration of code time frame in layer 0 known as reference time slot (RTS). For a layer *l*, the code time frame can be represented as sum of 2^l RTSs. For a code $C_{l,x}$ the slots are $S_{l,2^l(x-1)+1}$ to $S_{l,2^l x}$ with width of each slot T_{ref}. For every layer, the total numbers of slots are 2^{L-1}. For a call of rate $2^l R$, any of the 2^l unused slots (consecutive or scattered) in layer *l* can be used for the new call. This is different from the code allocation without time slot usage where a call can be handled only if there is at least one vacant code (equivalent to 2^l consecutive vacant slots).

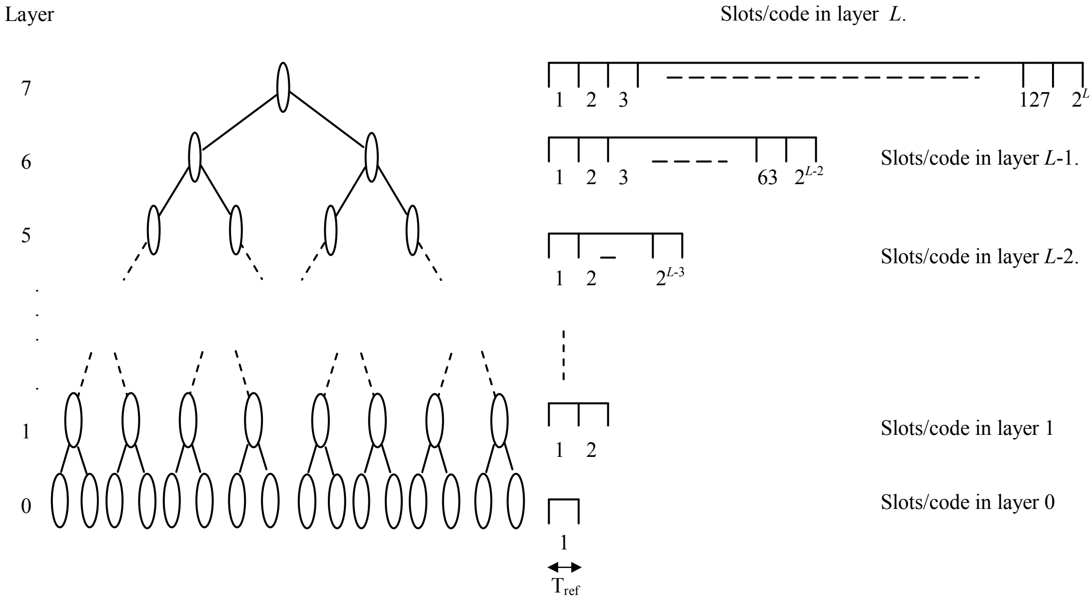


Fig.5. (a) 8 layer N-OVSF code tree (b) Number of slots/code for each layer in part (a).

The division of code time frame into slots is illustrated in Figure 5 for $l, (l \le L-1)$ layer OVSF code tree. To illustrate code slots sharing benefits, two scenarios of WCDMA systems are considered (i) general case where all the calls are treated similarly (ii) one (or more) layer calls is (are) given higher priority as compared to others.

4.1.1 General case: All calls are treated similarly

Consider full vacant code tree. For a new call of rate *kR*, find $\min(l) | k \le 2^l$. The call *kR* uses *k* vacant slots of a vacant code $C_{l',x}$, where $1 \le x \le 2^{L-l'}$ and $l' = \min(l, L-1)$ with $C_{l',x}$ has at least *k* vacant slots. If a code $C_{l',x}$, is used and it has $j, j > k$ vacant slots initially, the vacant slots of code $C_{l',x}$, which can be used by future calls are $j - k$. If layer l' do not have *k* consecutive vacant slots, find the vacant slots in codes of layer l'+1 to *L*-1. If we define slot utilization as the ratio of slots used by ongoing calls to the total slots available in the OVSF CDMA system, the code slot sharing increases slot utilization and reduce code blocking significantly.

4.1.2 One or more users are given higher priority

When one or more users (layers) are given higher priority, some of the busy slots of the ongoing non priority calls can be utilized for priority calls. The completion time for non priority calls increases as a consequence. Let layer vectors $\overline{l^p} = [l_1^p, l_2^p, ...l_P^p]$ and $\overline{l^{np}} = [l_1^{np}, l_2^{np}, ...l_{L-P}^{np}]$ represents *P* priority and *L-P* non priority layers. For non priority calls a threshold vector $\overline{H^{np}} = [H_1^{np}, H_2^{np}, ...H_{L-P}^{np}]$ (signifying maximum number of time slots which can be used by priority calls) is used to decide whether or not the incoming priority call can be handled by a non priority layer l_i^{np}. If the number of slots occupied by priority class users in layer l_i^{np} is less than H_i^{np}, some of the busy slots in layer l_i^{np} can be used by the priority layer call. The completion time of the of the non priority call increases (compared to the completion time when there is no slot sharing). For a code $C_{l_i, x_{l_i^{np}}}, 1 \le x_{l_i^{np}} \le 2^{L-l_i^{np}}$, the slot utilization $UC_{l_i, x_{l_i^{np}}}$ is defined as

$$UC_{l_i, x_{l_i^{np}}} = \frac{2^{l_i} - \sum SC_{l_i, x_{l_i^{np}}}}{2^{l_i}} \tag{5}$$

In Equation (5), $\sum SC_{l_i, x_{l_i^{np}}}$ represents the sum of slots of busy codes $C_{l_i, x_{l_i^{np}}}$, used by the previous priority calls (if any). Slot utilization represents the fraction of the code capacity utilized by ongoing non priority call. The code $C_{l_i, x_{l_i^{np}}}$ with the least value of slot utilization is the candidate for handling of new call. The aim is not to reduce the number of slots significantly for the ongoing busy call in non priority layer. This maintains the fairness for amount of slots used by ongoing non priority calls. If a tie occurs for slot utilization of two or more busy codes, any one code can be used for priority call handling.

If the priority users can be of non quantized rate nature, consider P' priority classes represented by vector $\overline{R^p} = [R_1^p, R_1^p, ...R_p^{P'}]$. All P' priority users fall in *P* different layers represented by layer vector $\overline{l^p} = [l_1^p, l_2^p, ...l_p^P]$, where the priority users with rates R_j^p will require time slots from the priority layer l_k^p if $R_j^p \le 2^{l_k^p} \mid l_k^p = \min(l_k^p)$ for all *j*. Then for a new call of rate $R_j^p = kR$, the procedure will be same as for quantized rate priority users. This may lead to code capacity wastage and may be avoided by the use of multiple codes as will be discussed in section 4.2.

To illustrate the delay induced in completion of non priority call by assigning its slots to priority call, assume t_0 is the initial completion time of the call handled by the code $C_{l_i^{np}, x_{l_i^{np}}}, 1 \le x_{l_i^{np}} \le 2^{7-l_i^{np}}$ in non priority layer l_i^{np} with rate $2^{l_i^{np}} R$ whose busy code threshold is not exceeded. The busy code is identified using the slot utilization Equation (5). If a new priority call seeking a code in layer $l_{i_1}^p$ arrives and there is no vacant code in layer $l_{i_1}^p$, the initial time t_0 of the code $C_{l_i^{np}, x_{l_i^{np}}}$ is divided into two components $t_0^{'}$ and $\left(t_0 - t_0^{'}\right)$. The time $t_0^{'}$ signifies the elapsed time of the code $C_{l_i^{np}, x_{l_i^{np}}}$ when the call of $2^{l_{i_1}^p} R$ arrives. The proposed design increases remaining time $\left(t_0 - t_0^{'}\right)$ of the call by assigning $2^{l_i^p} R$ code slots to $2^{l_i^p} R$ priority call. Subsequently, the new remaining time for ongoing non priority call after the arrival of $2^{l_i^p} R$ call becomes

$$t_1 = \left(t_0 - t_0^{'}\right) \times \frac{2^{l_i^{np}}}{2^{l_i^{np}} - 2^{l_{i_1}^p}} \tag{6}$$

Also, the total duration of non priority call becomes $t_0^{'} + t_1$. If some future priority call $l_{i_2}^p$ requires vacant slots and threshold of this non priority call is not exceeded, the $l_{i_2}^p$ number of slots from this call can be utilized by priority call $2^{l_{i_2}^p} R$. The time t_1 of the code $C_{l_i^{np}, x_{l_i^{np}}}$ is divided into two components $t_1^{'}$ and $\left(t_1 - t_1^{'}\right)$. The new remaining time of the non priority call becomes $t_0^{'} + t_1^{'} + t_2$ where t_2 is given as

$$t_2 = \left(t_1 - t_1^{'}\right) \times \frac{2^{l_i^{np}} - 2^{l_{i_1}^p}}{2^{l_i^{np}} - 2^{l_{i_1}^p} - 2^{l_{i_2}^p}} \tag{7}$$

The result can be generalized for m^{th} level sharing of non priority call $2^{l_i^{np}} R$ to handle priority call $2^{l_{i_m}^p} R$. Therefore the remaining time becomes $\sum_{j=1}^{m-1} t_j^{'} + t_m$, where tm is defined as

$$t_m = \left(t_{m-1} - t_{m-1}^{'}\right) \times \frac{2^{l_i^{np}} - \sum_{j=1}^{m-1} 2^{l_{ij}^p}}{2^{l_i^{np}} - \sum_{j=1}^{m} 2^{l_{ij}^p}} \tag{8}$$

The discussion in the section deals with users with rates in the form of 2^l. If a new priority call of rate *kR*, $k \neq 2^l$ requires code, the *k* consecutive vacant slots are searched for codes in layer $l | k \le 2^l$. If the vacant slots are not available, the codes in layer $l_i^p | l_i^p > l$ are searched and *k* vacant slots in layer l_i^p are utilized to handle the new call. If vacant slots are not present, *k* number of busy slots used by non priority calls are used as discussed above.

4.2 Multi code assignment

Consider an OVSF system equipped with *r* rake combiners. We illustrate multi code assignment for two cases as for single code assignment.

4.2.1 General case

Consider full vacant code tree. For a new call of rate *kR*, find $\min(l) | k \le 2^l$. If there are $q_1 | q_1 \le r$ codes in layer *l* with some vacant slots each, providing total k_1 vacant slots the fraction of *kR* rate handled by layer *l* codes is k_1R. The balance rate $k_2R=(k-k_1)R$ is to be handled by codes in the lower layers (*l*-1,…1). Starting with layer *l*-1, check the availability of maximum $q_2=r-q_1$ codes for finding k_2 vacant slots. The procedure can be repeated till layer 1 to find the k_2 vacant slots. Therefore multi code scheme uses the vacant slots of maximum *r* different codes to handle new call.

4.2.2 One or more users are given higher priority

There are two possibilities (i) the call requires code from the non priority layer (ii) the call requires code from the priority layer. If a non priority call *kR* requires codes (slots) from the non priority layer $l_{i_1}^{np}$ (say) and if there are $q_1 | q_1 \le r$ codes (consecutive or scattered) with some vacant slots and all q_1 codes provides total *k* (or $k_1 | k_1 < k$) vacant slots, the fraction of *kR* rate handled by layer

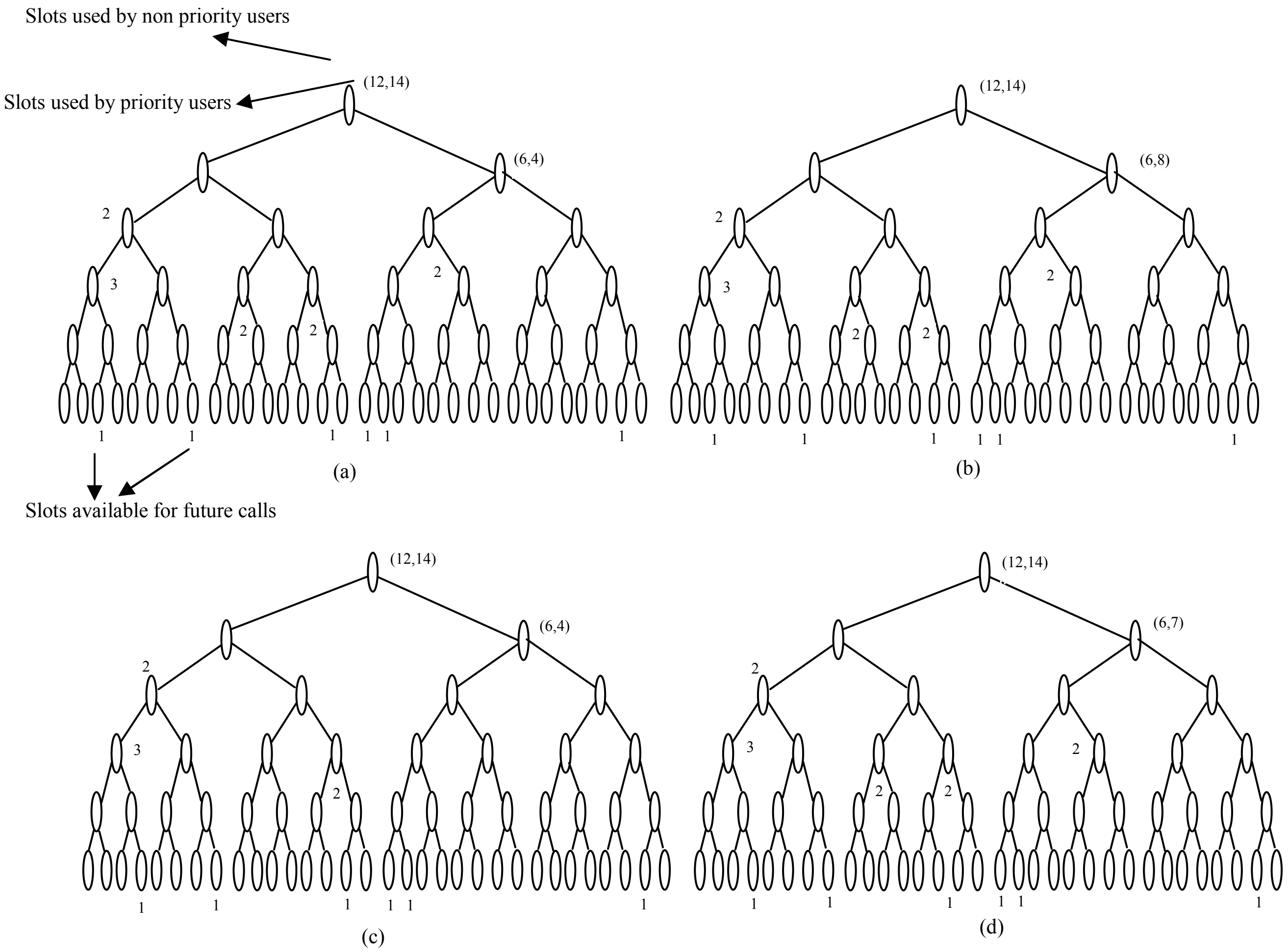


Fig.6. Illustration of slot usage in 6 layer NOVSF CDMA systems. All calls except 4*R* and 6*R* are non priority calls. (a) Status of the slots at the arrival of 4*R* priority call (b) Status of the slots after handling 4*R* call when the system has one rake (c) Status of the slots after handling 4*R* call when the system has two rakes (d) Status of the slots after handling 6*R* call when the system has two rakes

$l_{i_1}^{np}$ codes is *kR* (or k_1R) as discussed in section 3.2.1. However, if a priority layer call requires slots in priority layer and slots are available then assign new call to them. If vacant slots are insufficient to handle call of rate *kR*, then call will be handled using slots in both priority and non priority layers. If a code in priority layer $C_{l_{i_1}^p, m_1}, 1 \leq m_1 \leq 2^{L-1-l_{i_1}^p}$ has $MC_{l_{i_1}^p, m_1}$ number of vacant slots available, the total number of vacant slots of codes in priority layer $l_{i_1}^p$ used to handle call of rate *kR* is $\sum_{j=1}^{q_1} MC_{l_{i_1}^p, x_j}, 1 \leq x_j \leq 2^{L-1-l_{i_1}^p}$. The balance rate $k_2R=(k\text{-}k_1)R$ is to be handled by $q_2 \mid q_2 \leq (r - q_1)$ codes in the non priority layer $l_{i_2}^{np}$ where $l_{i_2}^{np} = \min(l_{i_z}^{np}) \mid l_{i_z}^{np} < l_{i_1}^{np}$. In lower most non priority layer (say) $l_{i_{L-P}}^{np}$, number of rakes available is $r - \sum_{l=1}^{L-P-1} q_l$. The total number of slots still required (k_{L-P}) can be formulated as

$$k_t = k - \sum_{l=1}^{L-P-1} k_l \qquad (9)$$

The total numbers of codes which can be used in layer $l_{i_{L-P}}^{np}$ are given by

$$q_{L-P} = r - \sum_{l=1}^{L-P-1} q_l \qquad (10)$$

If all the non priority lower layers are unable to handle the rate *kR*, the vacant slots are searched in higher layers *i.e.* $\sum_{j=1}^{P} \sum_{m=1}^{q_j} MC_{l_{i_j}^p, x_m^j} \leq k$, the vacant slots are searched in higher non priority layers $l_{i_{L-P+1}}^{np} = \min(l_{i_{L-P}}^{np}) \mid l_{i_{L-P+1}}^{np} > l_{i_1}^p$ in ascending order of layer number. If layer $l_{i_{L-P+z}}^{np}$ denotes the uppermost non priority layer, the number of rakes available in layer $l_{i_{t+z}}^{np}$ is $r - \sum_{l=1}^{L-P+z-1} q_l$. The number of slots available at layer $l_{i_{L-P+z}}^{np}$ is $k_{L-P+z} = k - \sum_{l=1}^{L-P+z-1} k_l$. If the number of slots required is less than k_{L-P+z}, the call is handled. If there are insufficient codes (vacant slots) available even in uppermost non priority layer, the layers $l_{i_{L-P+1}}^{np}$ to $l_{i_{L-P+z}}^{np}$ be checked for the threshold of number of busy codes. The procedure given in section can be used for using few vacant slots for incoming priority call.

Consider a six layer *NOVSF* code tree as with the status of the codes as shown in Figure 6(*a*). In layers 4 and 5, the variable (*a*,*b*) represents that *a* number of code slots are used by non priority users and *b* number of code slots are used by priority users. In other layers the variable *c* on a particular code represents *c* number of vacant slots for that code which can be used for future calls. Let the users with rates 4*R* and 6*R* are priority users. If a new user with rate 4*R* arrives and the system is equipped with one rake, the call require sharing of slots as there is no group of 4 consecutive vacant slots. If threshold for non priority layer *l* is 2^{l-1}, 4 vacant slots of code $C_{4,2}$ layer are used and the status of the code tree is shown in Figure 6(*b*). If the system is equipped with two rakes, codes $C_{2,6}$ and $C_{1,5}$ are used as shown in Figure 6(*c*). If a new user with rate 6*R* arrives codes $C_{2,1}$ and $C_{4,2}$ are used as shown in Figure 6(*d*).

5 Simulation results

Consider *L*=8 layer OVSF code tree as in the downlink of OVSF System. For simulation, following classes of users are considered with rates *R*, 2*R*, .. 8*R* respectively. The arrival rate λ is assumed to be Poisson's distributed with mean value varying from 0-4 calls per unit of time. Call duration is exponentially distributed with mean value of 3 units of time. The maximum capacity of the tree is 128*R* (*R* is 7.5*kbps*). Simulation is done for 5000 users and result is average of ten simulations.

There are 8 possibilities for arrival calls rates 2^iR, $, i \in [0,7]$. We assume that there are two sub cases (*a*) layer 1,2,3 users are given higher priority and these users make 20% of the arrival distribution and 80% of the calls are non priority calls (*b*) layer 1,2,3 users are priority users and these make 80% of the arrival distribution along with 50% non priority calls. The lowermost layers (1,2,3 *etc.*) are chosen for priority as real time calls are generally are priority calls. It is assumed that the users with two or more different priority layers are not given priority within themselves (though the provision can be made for the same also). The threshold for number of busy codes in non priority layer *i*, is 2^{7-i} (threshold for layer 7 is 1 as there is only one code). For quantized users system, let $\lambda_{q_i}, i \in [0,7]$ is the average arrival rate of i^{th} class calls. The average service time (denoted by, $1/\mu$ where μ is average service rate) is assumed to be 1 for all arrival classes. The traffic load for the i^{th} class of users is given by $\rho_{q_i} = \lambda_{q_i} / \mu$. Consider that there are G_i, *i*=0,1,2,…7 servers in the i^{th} layer corresponding to G_i number of vacant codes. The total codes (servers) in the system assuming eight set of classes are the given by vector $G=\{G_0,G_1,G_2,G_3,G_4,G_5,G_6,G_7\}$. The maximum number of servers used to handle new call is equal to the number of rake combiners. The code blocking for the i^{th} class is defined by

$$P_{B_i} = \frac{(\rho_{q_i})^{G_i} / G_i!}{\sum_{n=1}^{G_i} (\rho_{q_i})^{n} / n!} \qquad (11)$$

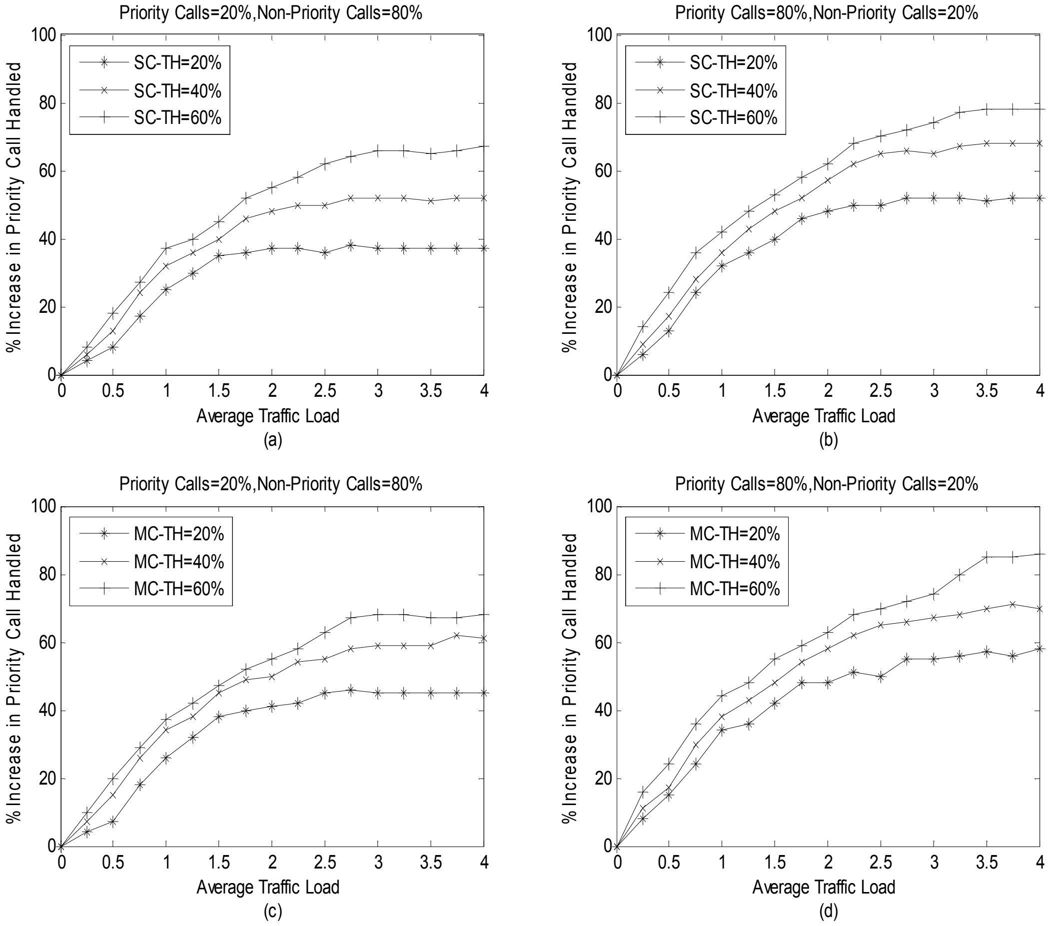
Define λ_q as the average arrival rate of all calls in the system and is given by $\lambda_q = \sum_{i=0}^{7} \lambda_{q_i}$. The average code blocking for quantized users is

$$P_B = \sum_{i=0}^{7} (\lambda_{q_i} / \lambda_{nq}) P_{B_i} \qquad (12)$$

Three threshold levels are considered for single code and multi codes

- Threshold=20%
- Threshold=40%
- Threshold=60%

As the threshold increases slot utilization increases for single code and multi codes assignment as shown in Figure. 7. The NOVSF provides zero code blocking; therefore results are not shown for this scheme.



(a) (b) (c) (d)

Fig. 7 Comparison of percentage of priority call handled in single code (a),(b) and multi codes (c),(d) designs for arrival rates distribution: Priority Calls=20%, Non-Priority Calls=80%, and Priority Calls=80%, Non-Priority Calls=20%

6 Conclusion

Real time calls are given higher priority in all 3G and beyond systems. The code blocking is the major limitation of OVSF based 3G systems which lead to call blocking in both real time and non real time calls. The code slots sharing schemes are proposed to reduce the code blocking for real time calls which gives reduction in call blocking. The use of multiple rakes along with the code sharing facility can be used to make real time call blocking close to zero. The code sharing (time sharing) is the complicated task and may require lot of effort. Work can be done to optimize the assignment of code slots for different order of priority within priority users.

References:

[1] E.H. Dinan and B. Jabbari , Spreading codes for direct sequence CDMA and wideband CDMA cellular networks, *IEEE Communications Magazine*, 1998.

[2] Spreading and Modulation (FDD), *3GPP TS 25.213 (V6.0.0). Technical Specification (Release 6), Technical Specification Group Radio Access Network,* 3GPP, 2003.

[3] R.L. Pickholtz, D.L. Schilling and L.B. Milstein, Theory of spread spectrum communications-A tutorial, *IEEE Transaction on Communications*, Vol. 30, 1982, pp. 855–884.

[4] F. Adachi, M. Sawahashi, K. Okawa, Tree structured generation of orthogonal spreading codes with different lengths for forward link of DS-CDMA mobile radio, *IEEE Electronic Letters*, Vol. 33, 1987, pp. 27-28.

[5] J. S. Chen, W. C. Chiang, N. C. Wang, and Y. F. Huang , Adaptive load balance and handoff management strategy for adaptive antenna array wireless networks, *in Proceedings of 12th WSEAS International Conference of Communication*, 2008, pp. 213-219.

[6] J. S. Chen, N. C. Wang, Z. W. Hong and Y.W. Chang, An adaptive load balance strategy for small antenna based wireless networks, *WSEAS Transaction on Communication,* Vol.8, No.7, 2009, pp. 588-597.

[7] G. Budura, C. Balint, A. Budura, and E. Marza, Traffic models and associated parameters in GSM/(E)GPRS networks, *WSEAS Transaction on Communication,* Vol. 8, No.8, 2009, pp. 833-842.

[8] C. M. Chao, Y. C. Tseng and L. C. Wang, Reducing internal and external fragmentation of OVSF codes in WCDMA systems with Multiple Codes, *IEEE Transaction on Wireless Communication*, Vol.4, pp. 1516-1526.

[9] H. Çam, Nonblocking OVSF Codes and enhancing network capacity for 3G wireless and beyond systems, *Special Issue of Computer Communications on 3G Wireless and Beyond for Computer Communications*, Vol. 26, No.17, 2003, pp. 1907-1917.

[10] E. Dahlman and K. Jamal, Wide-band services in a DS-CDMA based FPLMTS system, *in Proceedings of IEEE Vehicular Technology Conference*, No. 3, 1996, pp. 1656-1660.

[11] Y.C. Tseng, C.M. Chao and S.L. Wu, Code placement and replacement strategies for wideband CDMA OVSF code tree management, *in Proceedings of IEEE GLOBECOM*, Vol. 1, 2001, pp.562-566.

[12] J.S. Park and D.C. Lee, Enhanced fixed and dynamic code assignment policies for OVSF-CDMA systems*, in Proceedings of IEEE ICWN*, 2003, pp. 620-625.

[13] T. Minn and K.Y. Siu, Dynamic Assignment of orthogonal variable spreading factor codes in W-CDMA, *IEEE Journal of Selected Areas in Communication*, Vol. 18, No.8, 1998, pp. 1429-1440.

[14] V. Balyan and D. S. Saini, Vacant codes grouping and fast OVSF code assignment scheme for WCDMA networks, *Springer Journal of Telecommunication Systems*, 2011, DOI 10.1007/s11235-011-9469-5.

[15] V. Balyan and D. S. Saini, Integrating new calls and performance improvement in OVSF based CDMA Networks, *International Journal of Computers & Communication*, Vol.5, No. 2, ,2011, pp.35-42.

[16] W.T. Chen, H.C. Hsiao and Y.P. Wu, A novel code assignment scheme for W-CDMA systems, *in Proceedings of IEEE Vehicular Technology Conference*, 2001, Vol. 2, pp.1182-1186.

[17] J. S. Park, L. Huang, D. C. Lee, and C.C. Jay Kuo, Optimal code assignment and call admission control for OVSF-CDMA systems constrained by blocking probabilities, *in Proceedings of IEEE GLOBECOM*, Vol.5, 2004, pp. 3290-3294.

[18] Y. Yang and T. S. P. Yum, Multicode multirate compact assignment of OVSF Codes for QoS differentiated terminals, *IEEE Transaction on Vehicular Technology*, Vol. 54, No. 6, 2005, pp. 2114-2124.

[19] B. J. Chang and P. S. Chang, Multicode-based WCDMA for reducing waste rate and reassignments in mobile cellular communications, *Computer Communication*, Vol. 29, No. 11, 2006, pp. 948–1958.

[20] H. W. Ferng, H. L. C. Shiung and D. Y. C. Tsung, An OVSF code tree partition policy for WCDMA systems based on the Multi-Code approach, *in Proceedings of IEEE Vehicular Technology Conference,* 2005, No. 2, pp. 1212-1216.

[21] F. A. P. Cruz, J. L. A. Vazquez, A. J. Seguin and L. O Guerrero, Call admission and code allocation strategies for WCDMA Systems with multirate traffic, *IEEE Journal of Selected*

Areas in Communications, Vol.24, 2006, pp. 26-35.

[22] Y. Yang and T. S. P. Yum, Maximally flexible assignment of orthogonal variable spreading factor codes for multirate traffic*, IEEE Transaction on Wireless Communication*, Vol. 3, No. 3, 2004, pp. 781–792.

[23] A.C. Kam, T. Minn, and K.Y. Siu, Supporting rate guarantee and fair access for bursty data traffic in WCDMA, *IEEE Journal of Selected Areas in Communication*, Vol. 19, No. 11, 2001, pp. 2121–2130.

[24] L. Xu, X. Shen, and J.W. Mark, Dynamic bandwidth allocation with fair scheduling for WCDMA systems, *IEEE Transaction on Wireless Communications*,Vol.9, No.2, 2002, pp. 26-32.

[25] D .S .Saini, Reducing wastage capacity in OVSF based CDMA networks using dynamic rake combiners, *WSEAS Transaction on Communication,* Vol. 10, No.6, 2009, pp. 163-174.

Davinder S Saini was born in Nalagarh, India in January 1976. He received B.E degree in electronics and telecommunication engineering from College of Engineering Osmanabad, India in 1998. He received M.Tech degree in communication systems from Indian Institute of Technology (IIT) Roorkee, India in 2001. He received PhD degree in electronics and communication from Jaypee University of Information Technology Waknaghat, India in 2008. He is with Jaypee University of Information Technology Wakhnaghat since june 2002. Currently, he is Associate Professor in electronics and communication department. His research areas include Channelization (OVSF) codes and optimization in WCDMA, routing algorithms and security issues in MANETs.

Vipin Balyan received B.E degree in Electronics & Communication with honors India in 2003 and M.Tech degree in Electronics & Networking from LaTrobe University, Bundoora, Melbourne, Australia in 2006. He has been with RKGIT, Ghaziabad affiliated to U.P.Technical University, Lucknow, Uttar Pradesh, India as a Lecturer for 2 years. He was associated with Jaypee University, Wakhnaghat, India as a Senior Lecturer. He is currently working as a Senior Lecturer in Jaypee Institute of Information Technology, Noida, India and pursuing his Ph.D on *Efficient Single Code Assignment in OVSF based WCDMA Wireless Networks***.**