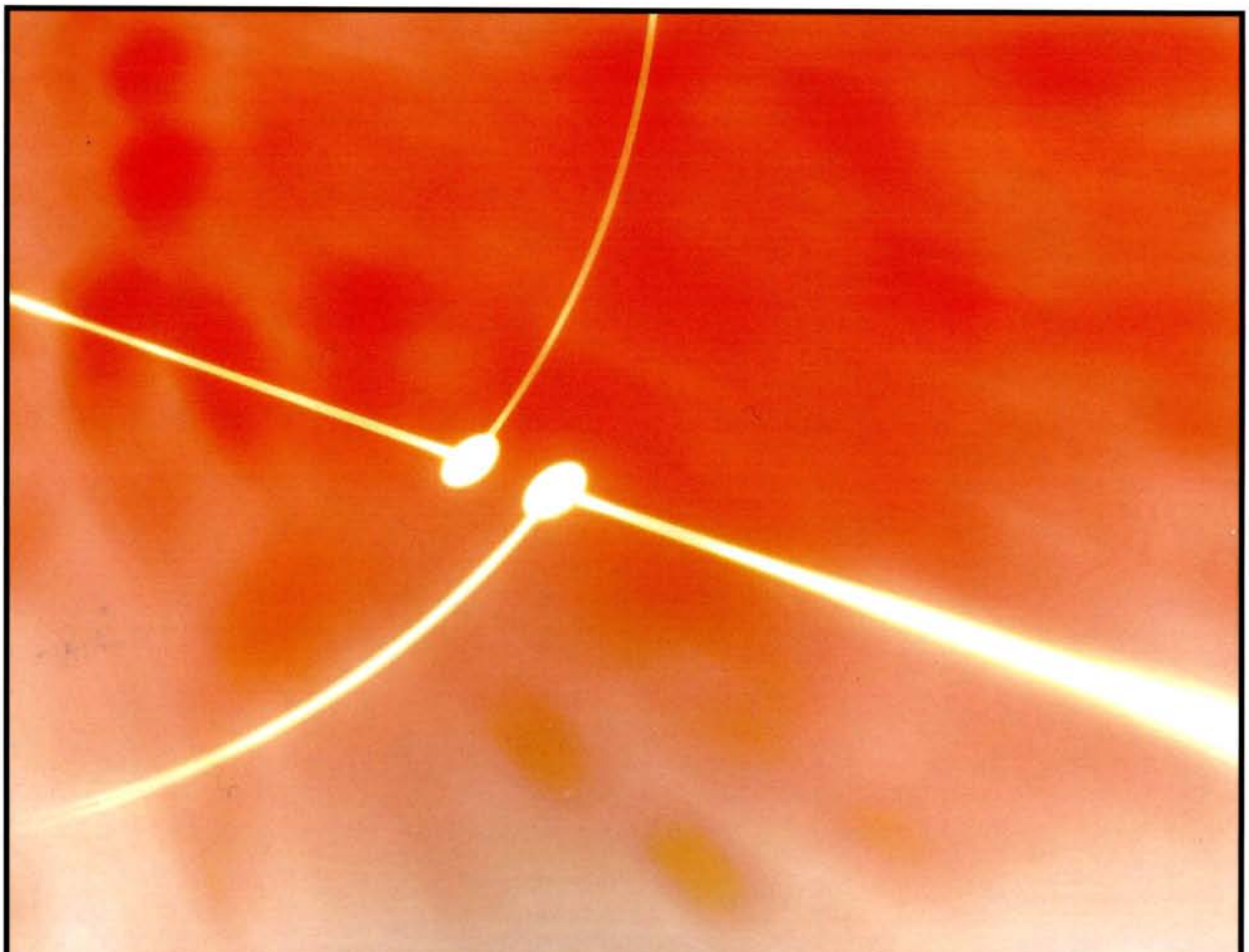
PREMIER REFERENCE SOURCE

Quality of Service Architectures for Wireless Networks

Performance Metrics and Management



Sasan Adibi, Raj Jain, Shyam Parekh, & Mostafa Tofighbakhsh

Quality of Service Architectures for Wireless Networks:
Performance Metrics and Management

Sasan Adibi
Research In Motion (RIM), Canada

Raj Jain
Washington University in St. Louis, USA

Shyam Parekh
Bell Labs, Alcatel-Lucent, USA

Mostafa Tofighbakhsh
AT&T Labs, USA

Information Science REFERENCE

INFORMATION SCIENCE REFERENCE

Hershey · New York

Table of Contents

Chapter 8
User Based Call Admission Control Algorithms for Cellular Mobile Systems

Hamid Beigy
Sharif University of Technology, Iran

M. R. Meybodi
Amirkabir University of Technology, Iran

ABSTRACT

Call admission control in mobile cellular networks has become a high priority in network design research due to the rapid growth of popularity of wireless networks. Dozens of various call admission policies have been proposed for mobile cellular networks. This chapter proposes a classification of user based call admission policies in mobile cellular networks. The proposed classification not only provides a coherent framework for comparative studies of existing approaches, but also helps future researches and developments of new call admission policies.

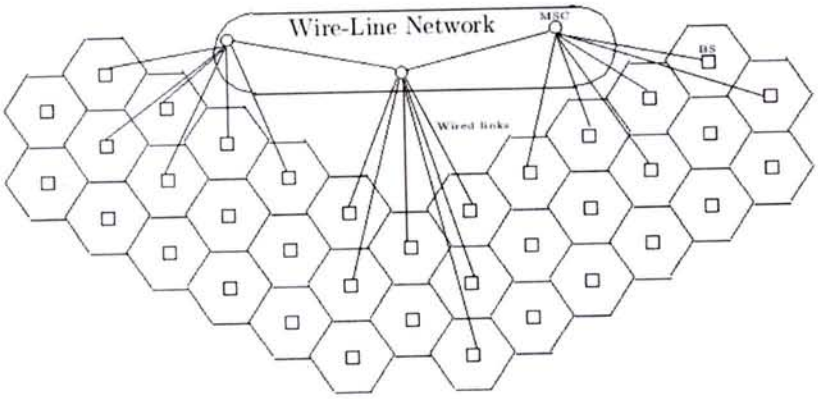
1. INTRODUCTION

The frequency spectrum allocated to the mobile communication networks is very limited. This means that the frequency channels have to be reused as much as possible in order to support the many thousands of simultaneous calls that may arise in any typical mobile communication network (Katzela & Naghshineh, 1996). Thus, the efficient management and sharing of channels among numerous users become an important issue. In cellular networks the geographical area covered by the network is divided into smaller regions called cells. Each cell is serviced by a base station, located at its center. The base station is used to service the users located at that cell. A number of base stations are again linked to a central server called mobile switching center, which also acts as a gateway of the mobile communication network to the existing wire-line networks such as PSTN, or internet. A base station communicates with users (mobile stations) through wireless links and with mobile switching centers through dedicated links. The model of such a network referred to as *cellular network* is shown in figure 1 (Das & Sen & Jayaram, 1998).

We assume that the network uses a fixed channel assignment algorithm, which means that each base station has a fixed number of channels (capacity).

DOI: 10.4018/978-1-61520-680-3.ch008

Figure 1. System model of cellular networks

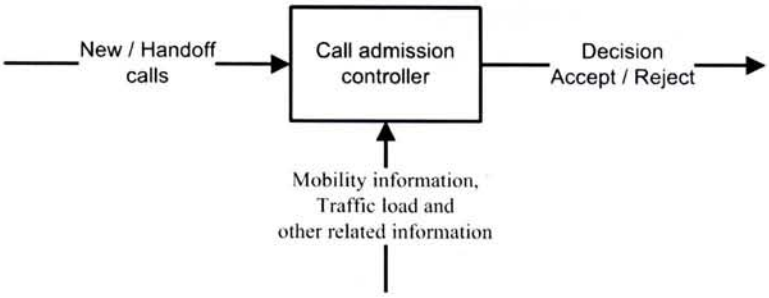


This capacity is interpreted in terms of bandwidth and is independent of used multiple access technology such as FDMA, TDMA, or CDMA. In order for a mobile user to be able to communicate with other user(s), a connection usually must be established between the users. The establishment and maintenance of a connection in cellular networks is the responsibility of the base stations. In order to establish a connection, a mobile user must first specify its traffic characteristics and quality of service (QoS) requirements. This traffic specification may be either implicit or explicit depending on the type of services provided by the network. For example, in a cellular phone network, the traffic characteristics and QoS requirements of voice connections are known a priori to the base station, and therefore, they are usually specified implicitly in a connection request. The next generation wireless networks are expected to eventually carry multi-media traffic such as voice, mixed voice and data, image transmission, email and etc. The traffic characteristics and the QoS requirements of connections for these services may not be known a priori to the base station. In these networks, mobile users must specify explicitly the traffic characteristics and the QoS requirements as a part of the connection request. Then, the base station determines whether it can meet the requested QoS requirements and, if possible, establish a connection.

When a call is originated and attempted in a cell, one channel allocated to the base station is used for the communication between the mobile station and the base station as long as channel is available. When all channels in a cell are in use while a call is attempted, then it will be blocked and cleared from the system. When a call gets a channel, it will keep the channel until its completion, or until it moves out of the cell, in which case the used channel will be released. When the mobile station moves into a new cell while its call is ongoing, a new channel needs to be acquired in the new call for further communication. This process is called *handoff* and must be transparent to the mobile user. During the handoff, if there is no channel is available in the new call for the ongoing call, it is forced to terminate before its completion.

When a user moves from one cell to another, the base station in the new cell must be responsible for all the previously established connections. A significant responsibility involves allocating sufficient resources in the cell for maintaining the QoS requirements of the established connections. If sufficient resources are not allocated to the handoff calls, the QoS requirements may not be met, which in turn may result in forced termination of the connection. Since the forced termination of established connections is usually more objectionable than rejection of a new con-

Figure 2. The call admission control algorithm



nection request, it widely believed that a cellular network must give a higher priority to the handoff connection requests as compared to new connections requests. Handoff problems are expected to become more and more important since the size of cells in emerging cellular networks tends to be smaller, which implies that handoff would occur more frequently, to attain a higher capacity.

In order to satisfy the QoS requirements, call admission control algorithms are needed, which determine whether a call should be either accepted or rejected at the base station and assign the required channel(s) to the accepted call. This results in a distributed call admission control strategy, which can be applied to every base station. Whenever a new call arrives, the call admission policy takes the call as input and based upon the current traffic conditions of network, decides whether or not to accept the user, as illustrated in figure 2. Call admission control in mobile cellular networks became a high priority in network design and research due to the rapid growth of popularity of wireless networks. A large number of call admission policies have been proposed for mobile cellular networks. However, despite years of research efforts, the call admission problem remains a critical issue and a high priority, especially given the perspectives of continually growing speed and size of future wireless networks. It is often difficult to characterize and compare various features among different policies. A good and detailed classification helps the researches and engineers to understand the similarities and differences among various schemes and decide which techniques are best suited for particular use (Beigy & Meybodi, 2003d; Ghaderi & Boutaba, 2006; Cruz-Perez & Ortigoza-Guerrero, 2007). These classifications not only provides a coherent framework for comparative studies of existing approaches, but also helps in future researches and developments of new call admission policies. This chapter is based on the classification given in (Beigy & Meybodi, 2003d).

The rest of this chapter is organized as follows: Section 2 describes the call admission problem and presents the proposed classification. Section 3 gives the non-prioritized call admission policies and the prioritized call admission policies are given in section 4. Optimal policies are given in section 5 and section 6 concludes the chapter.

2. CALL ADMISSION CONTROL

The challenges in the wireless networks are to guarantee the QoS requirements while taking into account the limited number of channels and interference between them. The study of the different schemes to accept calls in communication networks is known as the call admission control problem. Call admission control for high-speed wire-line networks have been intensively studied in the last few years. There are two major differences between wireless and wire-line networks due to the link characteristics and user mobility. The transmission links for the broadband wire-line networks are characterized by high transmission rates and very low error rates. In contrast, wireless links have a much smaller transmission rates and a much high error rates. The second major difference between the two networks is the user mobility. In wire-line networks, the user-network interface remains fixed throughout the duration of a connection whereas the user-network interface in a wireless environment may change throughout the connection. Due to the user mobility, call admission control becomes much more complicated in the wireless networks than wire-line networks. An accepted call that has not completed in the current cell may have to be handed off to another cell. During the handoff, the call may not be gain a channel in the new cell to continue its service due to the forced call termination. Thus, the new calls and handoff calls to be treated differently in terms of resource allocation. Since users tend to be much more sensitive to forced call termination (call dropping) than to the call blocking, handoff calls are normally assigned higher priority over the new calls.

Call admission control is one method to manage radio resources in order to adapt to the traffic variations. Call admission control denotes the process to make a decision for new admission according to the amount of the available resources versus users QoS requirements, and the effect upon the QoS of the existing calls imposed by new calls. Call admission control plays a very important role in cellular networks because it directly controls the number of users in the network and must be designed to guarantee the QoS requirements. The usual network performance indicators are the blocking probability of new calls, the dropping probability of handoff calls, the computation and communication overheads, and the total carried load. Good call admission control policies have to balance the dropping probability of handoff calls and the blocking probability of new call in order to provide the desired QoS requirements.

There has been much research into call admission control policies for cellular networks. A good call admission control algorithm must have the following features in order of importance.

- Maximize channel utilization in a fair manner to all calls
- Minimize the dropping probability of connected calls
- Minimize the reduction of the QoS for the connected calls
- Minimize the blocking probability of new calls

Call admission control policies can be divided into a number of different categories depending on the comparison basis. For example, when call admission control policies are compared based on decision policies, they can be divided into user (number)-based CAC (NCAC) and interference-based CAC (ICAC) policies (Ishikawa & Umeda, 1997). NCAC policies accept/reject calls based on the number of users in the cell. Using ICAC, a base station, by monitoring the interference on a call-by-call basis, determines whether or not a new call is acceptable. The new call is blocked if the observed interference level exceeds a CAC interference threshold. Each base station should measure the total power of received signals in the spreading bandwidth before dispreading them. ICAC therefore requires overheads for base station hardware and complicates its architecture, while NCAC can be implemented by means of base station software.

Before we start presenting NCAC schemes, we give a general framework for call admission control, which will be used throughout this chapter, is developed. We consider network cells with N classes of calls $W = \{w_1, \ldots, w_N\}$, and C full duplex channels. Class $w_i\,(for\ 1 \le i \le N)$ consists of a stream of statistically identical calls with Poisson arrival at rate λ_i and independent identical exponentially distributed call holding times with the same mean $1/\mu$. Assume that all classes need only one channel for each call. Let c denotes the number of busy channels in the cell. The state space S of a cell is given by $S = \{c \mid c \leq C\}$. We define the admission policy $u : S \times W \rightarrow \{0,1\}$, where $u(x,w)$ specifies the probability of acceptance of calls of class w when the cell is in state x. At any time t, the decision to accept or reject calls of class w depends only on the current state of the cell or its neighboring cells. From the point of the call admission controller, the process can be modeled as a Markov process, where the transition rates between the states $x, y \in S$ for a call of class w, are given by

$$q(x,y) = \begin{cases} \sum_{w=1}^{N} u(x,w)\lambda_{w} & if\ y = x+1 \\ x\mu & if\ y = x-1 \\ 0 & otherwise \end{cases} \quad (1)$$

Function $u(x,w)$ may be deterministic or stochastic (probabilistic), static or dynamic. Based on function $u(x,w)$, the call admission control policies can be divided into non-prioritized, prioritized, and optimal policies, as shown in figure 3. In non-prioritized policies (Hong & Rappaport, 1986), all calls are accepted when the requested channels are free, while in prioritized policies, one group of calls have a higher priority than other groups, for example, the handoff calls have the higher priority than new calls. In prioritized policies, when the requested channels are not available, the call may be queued or rejected. Optimal policies accept/reject calls to maximize throughput of the network.

3. NON-PRIORITIZED CALL ADMISSION CONTROL POLICIES

In these call admission control policies (Hong & Rappaport, 1986), no single class is treated differently than any other classes. This is the simplest scheme and involving checking to guarantee that the requested bandwidth is available for the calls. If the bandwidth requirements can be met, then the call is accepted and the bandwidth is allocated; otherwise the call is blocked. This policy always accepts calls as long as doing so leads to a state

Figure 3. Classification of user based call admission control algorithms

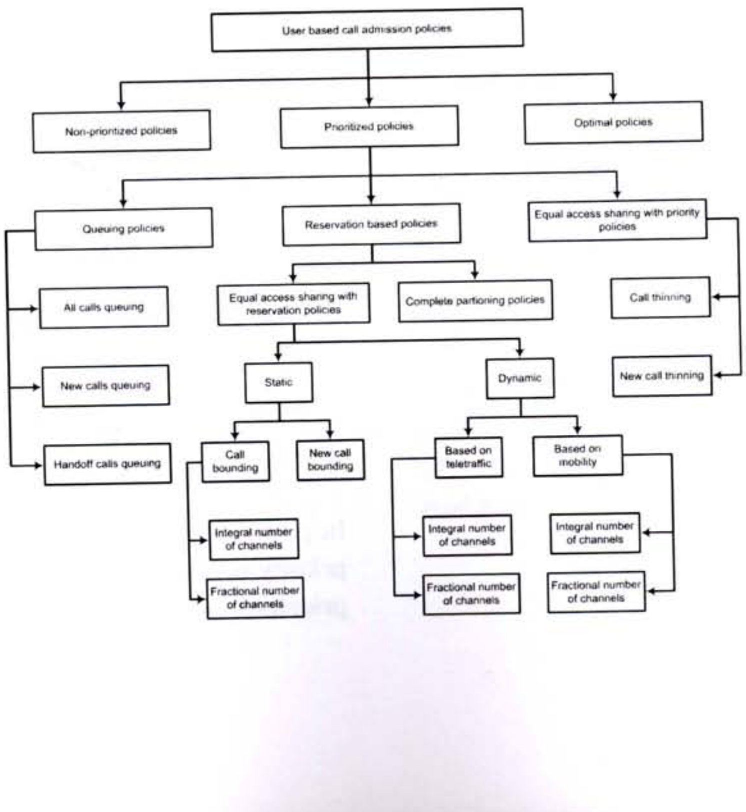
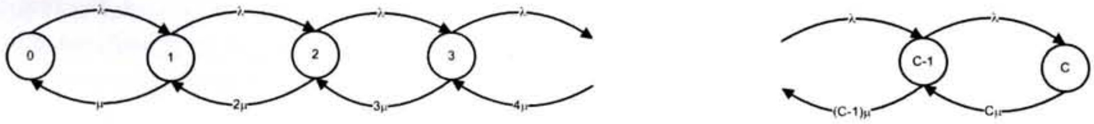


Figure 4. State transition diagram for non-prioritized scheme



in the state space S, that is,

$$u(x,w)=\begin{cases}1 & if\ x+1\in S\\ 0 & otherwise\end{cases} \tag{2}$$

In order to study the performance of this scheme, we consider a homogenous cellular network where all cells have the same number of channels, C, and experience the same arrival rates for all classes of calls. Without loss of generality, we consider two classes of calls: new and handoff calls. We assume that the arrival of new and handoff calls are Poisson distributed with rates λ_n and λ_h, respectively and the call holding time of calls are exponentially distributed with the mean $1/\mu$. Note that the same service rate for both types of calls implies that the base station of a cell does not need to discriminate between new and handoff calls, once they are connected. These assumptions have been found reasonable as long as the number of mobile users in a cell is much greater than the number of channels allocated to that cell. Define the state of a cell at time t by the total number of occupied channels, $c(t)$. Thus, the channel occupancy can be modeled by a continuous time Markov chain with states $0, 1, \ldots, C$. Figure 4 shows the state transition diagram of a system with C channels and non-prioritized call admission scheme.

Define the steady state probability $P_n = \lim_{t\to\infty}$ Prob $[c(t)=n]$ as the probability of n channels being occupied. Given this, it is straight forward to derive probability P_n (for $n=0, 1, \ldots, C$). The steady state probability P_n that n channels are busy is given by the following expression.

$$P_n=\left(\frac{\rho^n}{n!}\right)P_0, \tag{3}$$

where

$$P_0=\left[\sum_{k=0}^{C}\left(\frac{\rho^k}{k!}\right)\right]^{-1}, \tag{4}$$

and $\rho=(\lambda_n+\lambda_h)/\mu$. Two commonly used performance measures for cellular networks are: *dropping probability of handoff calls* (B_h) and *blocking probability of new calls* (B_n). The dropping probability of handoff calls represents the probability that a handoff call being dropped during handovers. This probability is defined as the ratio between the number of calls dropped by the system and the total number of admitted calls. The blocking probability of new calls represents the probability that a new call being denied access to the network. This probability is defined as the percentage of calls that are denied access to the network. Given the state probabilities, we can drive the blocking probability of new calls and the dropping probability of handoff calls.

$$B_n=B_h=\frac{\rho^C}{C!}P_0 \tag{5}$$

4. PRIORITIZED CALL ADMISSION CONTROL POLICIES

In prioritized call admission control policies, a priority is assigned to each class of calls. These priorities are implemented through function *u(x,w)*. For example, from the point of view of a mobile user, dropping of an ongoing call is less desirable than blocking of a new call. Therefore, to reduce the chances of unsuccessful handoff calls, the system assigns a higher priority to the handoff calls. Thus the function *u(x,.)* has a higher value for handoff calls than the new calls. The prioritized call admission policies can be divided into three groups: *equal access sharing with priority, reservation based and queuing based policies*, and *queuing priority policies*, which are described in the rest of this section.

4.1 Equal Access Sharing with Priority Policies (EASWP)

In these call admission control policies, all classes of calls have access to all channels but some classes have a higher priority than others. This priority is implemented through the use of function *u(x,w) p(x,w)*, where *p(x,w)* is the probability of accepting calls of class *w* when the cell is in state *x*. The reported EASWP policies can be classified as call thinning and new call thinning schemes, which are briefly described below.

4.1.1 Call Thinning Schemes

In these schemes, the state of system, *x*, is the number of busy channels in the cell. Call thinning schemes, in turn can be divided into two subclasses: static and dynamic schemes: In what follows, we explain these schemes for two classes of calls. In static call thinning schemes, *u(x, w)* is determined based on a priori information and remain fixed during the operation of the network. Ho & Lea (1999) proposed a static call thinning scheme and linear programming was used to determine the optimal values of *p(x, w)*. In this scheme, we have

$$u(x,w) = \begin{cases} p(x,w) & if\ x < C \\ 0 & if\ x = C \end{cases} \qquad (6)$$

A restricted version of this scheme, which is called *fractional guard channel* (FGC) scheme, was proposed by Ramjee & Towsley & Nagarajan (1997). In this scheme, the handoff calls have higher priority over the new calls. This scheme accepts new calls with certain probability that depends on the channel occupancy of the cell and accepts the handoff calls when the cell has free channels. In this scheme, we have

$$u(x,w) = \begin{cases} 1 & if\ x < C\ and\ w = handoff\ calls \\ p(x) & if\ x < C\ and\ w = new\ calls \\ 0 & if\ x = C \end{cases} \qquad (7)$$

Since *p(x)* only appears when new calls arrives, *p(x)*'s are called new call admission probabilities. The idea behind this scheme is to smoothly throttle the new call stream as the network traffic is building up. Thus, when the network is approaching the congestion, the accepted new calls become thinner. Due to the flexible choice of new call admission probabilities, this scheme can be made very general. Figure 5 shows the state transition diagram of a homogeneous network with *C* channels and FGC scheme.

Define the steady state probability $P_n = \lim_{t\to\infty}$ Prob $[c(t)=n]$ as the probability of *n* channels being occupied. The steady state probability P_n that *n* channels are busy is given by the following expression (Ramjee & Towsley & Nagarajan, 1997).

Figure 5. State transition diagram for FGC scheme

Figure 6. State transition diagram for UFC scheme

$$P_n = \left(\frac{\rho^n}{n!} \prod_{k=0}^{n} \gamma_k \right) P_0, \tag{8}$$

where

$$P_0 = \left[\sum_{k=0}^{C} \left(\frac{\rho^k}{k!} \prod_{k=0}^{n} \gamma_k \right) \right]^{-1}, \tag{9}$$

$\gamma_k = [\alpha + (1-\alpha)p(k)]$, $\alpha = \dfrac{\lambda_h}{\lambda_n + \lambda_h}$, and $\rho = (\lambda_n + \lambda_h) / \mu$. Given these state probabilities, we can drive the blocking probability of new calls and the dropping probability of handoff calls.

$$B_n(C,p) = P_0 \left(1-\alpha\right) \sum_{m=0}^{C} \frac{\rho^m}{m!} \prod_{k=0}^{m-1} \gamma_k$$

$$B_h(C,p) = P_0 \frac{\rho^C}{C!} \prod_{k=0}^{C-1} \gamma_k \tag{11}$$

The most disadvantage of this scheme is that no algorithm is given to find p(x)s. In order to find p(x), a restricted version of this scheme called uniform fractional guard channel scheme (UFC) is introduced by Beigy & Meybodi (2004a). In this scheme, the new call admission probabilities are independent of channel occupancy. Thus, in this scheme, we have

$$u(x,w) = \begin{cases} 1 & x < C \; and \; w = handoff \; calls \\ p & x < C \; and \; w = new \; calls \\ 0 & x = C \end{cases} \tag{11}$$

Figure 6 shows the state transition diagram of a homogeneous network with C channels and UFC scheme.

The steady state probability P_n that n channels are busy is given by the following expression:

$$P_n = \left(\frac{(\rho\gamma)^n}{n!} \right) P_0, \tag{12}$$

where

$$P_0 = \left[\sum_{k=0}^{C} \left(\frac{(\rho\gamma)^k}{k!} \right) \right]^{-1}, \tag{13}$$

$\gamma = [\alpha + (1-\alpha)p]$, $\alpha = \dfrac{\lambda_h}{\lambda_n + \lambda_h}$, and $\rho = (\lambda_n + \lambda_h) / \mu$. Given these state probabilities, we can drive the blocking probability of new calls and the dropping probability of handoff calls.

$$B_n(C,p) = 1 - \alpha \left[1 - \frac{(\rho\gamma)^C}{C!} P_0 \right]$$

$$B_h(C,p) = P_0 \frac{(\rho\gamma)^C}{C!} \tag{14}$$

$B_n(C,p)$ and $B_h(C,p)$ have interesting properties, which enable us to design an algorithm for finding the optimal value of parameter p. It was shown that $B_n(C,p)$ and $B_h(C,p)$ are monotonically decreasing and increasing function of p, respectively (Beigy & Meybodi, 2004a). The algorithm 1 is given for finding the optimal value of p and can be described as follows. At first, the algorithm considers the case when all channels are shared between handoff and new calls. If the complete sharing does not satisfy the level of QoS, then the algorithm considers the case when all channels are exclusively used for handoff calls. If the exclusive use of channels for handoff calls does not satisfy the level of QoS, then the number of allocated channels to the cell is not sufficient and the algorithm terminates; otherwise the algorithm searches for the optimal value of p. The search method used in this algorithm is binary search.

Algorithm 1: The algorithm for finding the optimal value of p

```
Algorithm FindUFCParameter
set upper ← 1; lower ← 0
if(B_h (C,1)≤ P_h)thenreturn 1
end if
if(B_h (C,0) ≥ P_h) then
return 0
end if
while ((upper -lower) < 0.0001)
doset p ← ⌊(upper + lower) /2⌋
if(B_h (C,1)> P_h)thenset upper ←
p
else
set lower ← p
end if
end while
return p
end Algorithm
```

In dynamic call thinning schemes, *u(x,w)* is adapted based on information gathered during the operation of the network. Some dynamic call thinning algorithms are reported in (Ayyagari & Empremides, 1999; Wu, & Wong & Li, 2002). A dynamic call thinning scheme for multi-media cellular network is presented Ayyagari & Empremides (1999). In this scheme, calls are classified on the basis of channel requirement and a propriety level is associated with each class of calls. This scheme collects calls in a time period and then accepts calls with the higher priorities. In (Wu, & Wong & Li, 2002), a call admission scheme called stable dynamic call admission control scheme is suggested. The aim of this scheme is to maximize the channel utilization (minimize the new call blocking probability) subject to a hard constraint on the dropping probability of handoff calls. In this scheme, status information is exchanged periodically among neighboring cells, and even next neighboring cells if necessary. The exchanged information includes the channel occupancies and the new call arrival rates. Each cell updates its acceptance ratio (the maximum fraction of new calls to be accepted in the cell) in the next control period at the beginning of that period. The control action is obtained by solving system of equations which specifies the average dropping probability of handoff calls must be equal to the QoS of the system. Beigy & Meybodi (2004b) proposed a learning automaton based algorithm to adjust the value of p, in which a learning automaton is associated to each cell. In this algorithm as shown in Algorithm 2, when a handoff call arrives, it is accepted as long as there is a free channel. If there is no free channel, the handoff call is blocked. When a new call arrives to a particular cell, the learning automaton associated to that cell chooses one of its actions. If action ACCEPT is selected by the automaton and the cell has a free channel, then action ACCEPT is rewarded. If there is no free channel to be allocated to the arrived new call, the call is blocked and the action ACCEPT is penalized. When the automaton selects action REJECT, the algorithm computes an estimation of the dropping probability of handoff calls ($\hat{B}_h$) and uses it to decide whether or not accept new calls. If the current estimate of dropping probability of handoff calls is less than the given threshold p_h and there is a free channel, then the new call is accepted and action REJECT is penalized; otherwise, the new call is rejected and action REJECT is rewarded.

Algorithm 2: The learning automata based algorithm for finding the optimal value of p

```
Algorithm AdaptiveUFC-I
if (NEW CALL) thenif (action of
learning automaton is ACCEPT)
thenif (c(t) < C) then
accept call and reward action
ACCEPT
 else
reject call and penalize action
ACCEPT
end if
else
if (c(t) < C and B̂_h < p_h) then
accept call
 else
reject call
end if
Compute B̂_h if (new call is ac-
cepted and B̂_h < p_h) then
penalize action REJECT
else
reward action REJECT
end if
end if
end if
end Algorithm
```

The simulation results reported in (Beigy and Meybodi, 2004b) shows that this algorithm cannot maintain the specific level of QoS for the dropping probability of handoff calls. This problem may be due to the existence of delay in the cellular network, because the selected action of learning automaton is immediately rewarded / penalized. Since the effect of the estimated new call admission probability is specified after a time period, then the reward/ punishment of learning automaton must be given in the end of that period. In order to overcome this problem, another algorithm is given in (Beigy & Meybodi, 2004b), in which the action probability vector of learning automaton is adjusted upon the arrival the next new call. This algorithm, as shown in Algorithm 3, uses a learning automaton to accept/reject new calls and a pre-specified level of dropping probability of handoff calls is used to penalize/reward the action selected by the learning automaton. This algorithm can be described as follows. When a handoff call arrives, it is accepted as long as there is a free channel. If there is no free channel, the handoff call is dropped. When a new call arrives to a particular cell, the learning automaton associated to that cell chooses one of its actions. If action ACCEPT is selected by automaton and the cell has at least one free channel, the incoming call is accepted and the selected action is rewarded. If there is no free channel to be allocated to the arrived new call, the call is blocked and action ACCEPT is penalized. When the automaton selects action REJECT, then the new call is rejected and the base station computes an estimation of the dropping probability of handoff calls ($\hat{B}_h$) and uses it to reward or punish action REJECT. If the current estimate of dropping probability of handoff calls is less than the given threshold p_h, then action REJECT is penalized; otherwise, action REJECT is rewarded. Beigy & Meybodi (2003a) showed that this algorithm finds the optimal value of the of UFC's Parameter.

Algorithm 3: The learning automata based algorithm for finding the optimal value of p

```
Algorithm AdaptiveUFC-II
if (NEW CALL) thenif (action of
learning automaton is ACCEPT)
thenif (c(t) < C) then
accept call and reward action
ACCEPT
 else
reject call and penalize action
ACCEPT
end if
else
reject call & compute B̂_h if (B̂_h
< p_h) then
penalize action REJECT
else
```

```
reward action REJECT
end if
end if
end if
end Algorithm
```

Figure 7 shows the performance of this under different handoff traffic when the other parameters of the cell are fixed. Note that the level of QoS is maintained by this algorithm for various handoff traffic conditions. Figure 8 shows the B_n and B_h for this algorithm for two typical different handoff traffic loads, which shows that that the admission probability converges to its optimal value.

4.1.2 New Call Thinning Schemes

Below we explain the one new call thinning scheme (Fang & Zhang, 2002). In this scheme, the state of system, x, is the number of ongoing new calls in the cell. For the sake of simplicity assume that we have two classes of calls: new calls and handoff calls. This scheme, which limits the new calls in the system, gives a higher priority to the handoff calls over the new calls. This scheme accepts new calls with certain probability that depends on the number of ongoing new calls in the cell and accepts the handoff calls when the cell has free channels. In this scheme, we have

$$u(x,w) = \begin{cases} 1 & if\ x < C\ and\ w = handoff\ calls \\ p(x) & if\ x < C\ and\ w = new\ calls \\ 0 & if\ x = C \end{cases} \tag{15}$$

4.2 Reservation Based Call Admission Control Policies

In reservation based call admission control policies, some of the channels allocated to the cell are reserved for the higher priority calls. In these policies, we have u(x,w)=0 for some x and w. In these call admission control policies, all classes of calls are accepted equally within a specified bandwidth of the maximum channel capacity that depends on the given class. Once the available channel capacity has been used, only calls that are of a high priority will be accepted to use the remaining (reserved) channels (bandwidth). This has the effect of prioritizing a traffic class above the other traffic classes. In the reservation based policies, classes of calls can be grouped and fix a threshold for each group. When restricted to simple form, these policies dedicate a certain number of channels for each group and the remaining channels are shared among all groups. To define a simple form for these policies, we form W groups, $G_1, \ldots, G_W$, such that each w belongs to only one group G_W. The reservation based policies can be stated as:

Figure 7. Performance of the adaptive UFC algorithm for different handoff traffic

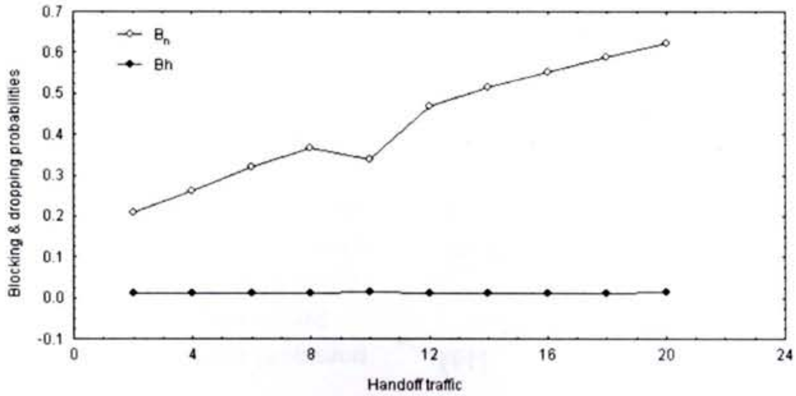
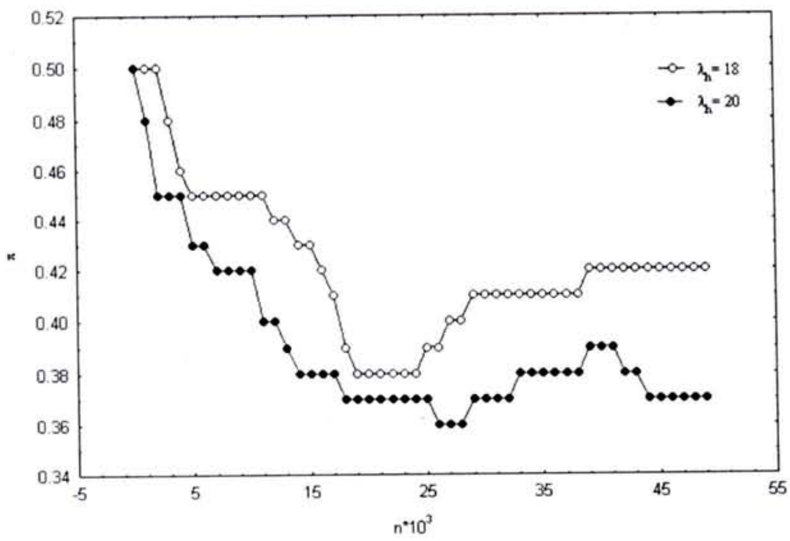


Figure 8. Convergence of the proposed algorithm for different handoff traffic



$$u(x,w) = I\{x \leq T_w\} \wedge I\{x+1 \in S_{G_W}\} \quad (16)$$

where T_W is the maximum channel capacity for calls of class w in group G_W and S_{G_W} is the set of channels associated to group G_W. These policies can be divided into two main groups: *equal access sharing with priority* (EASWR) and *complete partitioning schemes*, which are explained in the following subsections.

4.2.1 Equal Access Sharing with Reservation (EASWR)

In these call admission control policies, we have $S = S_{G_1} = S_{G_2} = ... = S_{G_W}$. Thus, all classes of calls can use any channel and calls are accepted with equal probabilities within a specified bandwidth of the maximum channel capacity. Once the available channel capacity has been used, only calls that are of high priority will be accepted to use the remaining (reserved) channels. This has the effect of prioritizing one class above the other classes, that is,

$$u(x,w) = I\{x \leq T_w\} \wedge I\{x+1 \in S\}, \quad (17)$$

where T_W is the maximum channel capacity for calls of class *w*. Based on the manner used for determination of the values of $T_w s$, the EASWR policies can be divided in two main groups: *static* and *dynamic EASWR* schemes. In static EASWR schemes, values of $T_w s$ are determined based on the a priori information about the network and remain unchanged during the operation of the network while in dynamic EASWR schemes, $T_w s$ are adapted during the operation of the network. The static EASWR schemes can be divided into two main groups: *call bounding* and *new call bounding* schemes. In the call bounding schemes, the call admission is based on the number of ongoing calls (number of busy channels) in the cell while in the new call bounding schemes; the call admission is based on the number of ongoing new calls in the cell. In dynamic EASWR schemes, function *u(x,w)* is adapted according to the some available information. In dynamic EASWR schemes, the number of channels is allocated and reserved dynamically using *traffic analysis* and *prediction of mobile terminal movement*.

Static call bounding schemes: In the call bounding schemes, admission of a new call is based on the number of ongoing calls in the cell, independent of type of calls. In other words, the state x of a cell is defined as the number of busy channels in the cell. Based on the values of $T_w s$, the call bounding schemes can be divided into two schemes: *reserving integral number of channels* and *reserving fractional number of channels*. In the reserving integral number of channels, all $T_w s$ are integer values while in reserving fractional number of channels, at least one of $T_w\,'s$ are fractional numbers. In the reserving integral number of channel schemes, range of function *u(x,w)* is the set of $\{0,1\}$. When only two groups G_1 and G_2 (one for new calls and the other for handoff calls) are considered, this scheme is referred to as *guard channel policy*, or *cutoff priority policy* in which a fixed number of channels is reserved in each cell exclusively for handoff calls (Hong & Rappaport, 1986). Under such policy, new calls and handoff calls are treated equally on a first-come first-served basis for channel allocation until a predetermined channel utilization threshold is reached. Let T be this threshold. At this point, new calls are simply blocked and only handoff call requests are accepted. In other words, a new call is accepted if $c < C - T$, where $T \geq 0$ is the number of channels reserved specifically for handoff (guard channels), that is,

$$u(x,w) = \begin{cases} 1 & if\ x < C\ and\ w = handoff\ calls \\ 1 & if\ x < T\ and\ w = new\ calls \\ 0 & if\ x = C \end{cases} \tag{18}$$

Figure 9 shows the state transition diagram of a homogeneous network with C channels and guard channel scheme. The system is modeled by a typical M/M/C/C queuing model.

The steady state probability P_n that n channels are busy is given by the following expression.

$$P_n = \begin{cases} \dfrac{\rho^n}{n!} P_0 & if \quad n \leq T \\ \alpha^{-T} \dfrac{(\rho\alpha)^n}{n!} P_0 & if \quad T < n \leq C \end{cases} \tag{19}$$

where

$$P_0 = \left[\sum_{k=0}^{T} \left(\frac{\rho^k}{k!} \right) + \alpha^{-T} \sum_{k=T+1}^{C} \frac{(\rho\alpha)^k}{k!} \right]^{-1}, \tag{20}$$

$\alpha = \dfrac{\lambda_h}{\lambda_n + \lambda_h}$ **and** $\rho = (\lambda_n + \lambda_h) / \mu$. Given these state probabilities, we can drive the blocking probability of new calls and the dropping probability of handoff calls.

$$B_n(C,T) = P_0 \alpha^{-T} \sum_{m=T+1}^{C} \frac{(\rho\alpha)^m}{m!}$$

$$B_h(C,T) = P_0 \alpha^{-T} \frac{(\rho\alpha)^C}{C!} \tag{21}$$

It has been shown that B_n (B_h) is a monotonically decreasing (increasing) function of T and there is an optimal threshold T^* in which the blocking probability of new calls is minimized subject to the hard constraint on the dropping probability of handoff calls. Algorithm 4 can be

Figure 9. State transition diagram for guard channel scheme

used to find the optimal value of threshold T^* (Ramjee & Towsley & Nagarajan, 1997; Haring & Marie & Puigjaner & Trivedi, 2001).

Algorithm 4: The algorithm for finding the T^*

```
Algorithm FindGCParameter
set upper ← 1; lower  ←  0
if(B_h (C,C)≤ P_h)thenreturn C
end if
if(B_h (C,0) ≥ P_h) then
return 0
end if
while ((upper -lower) < 0.0001)
doset p ← ⌊(upper + lower) /2⌋
if(B_h (C, ⌊p⌋)> P_h)thenset upper
← p
else
set lower ← p
end if
end while
return ⌊p⌋
end Algorithm
```

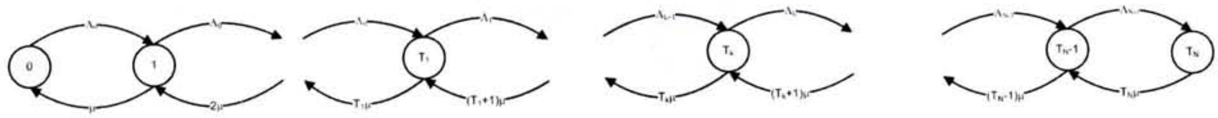
Chang & Kim (2001) proposed an algorithm to find the optimal number of guard channels in a general multi-cell networks, which minimizes the weighted average of dropping probability of handoff calls in a cluster while satisfying the pre-specified QoS for new calls and co-channel interference constraints. Approximate analyis of guard channel scheme supporting two classes of calls (new and handoff calls) with different average channel holding times were done by Fang & Zhang (200) and Yavuz & Leung (2006). Chen & Lee (2001) considered two traffic classes of voice and transactions and proposed a static guard channel scheme to maintain the upper bound of dropping probability of handoff transaction calls. In this approach, $(C-T)$ guard channels are reserved for handoff transaction calls, but new calls and handoff voice calls have the same priority. Thus, this scheme fails to maintain the upper bound for dropping probability of handoff voice calls. In order to maintain the upper bound for dropping/blocking probability for different classes of calls, call admission schemes with multi-thresholds are introduced.

In (Yin & Li & Zhang & Lin, 2000), dual-threshold reservation (DTR) scheme is given for integrated voice/data wireless networks. In DTR scheme, three classes of calls, data calls (both new and handoff calls), new voice calls and handoff voice calls in increasing order of level of QoS are considered. The basic idea behind the DTR scheme is to use two thresholds, one for reserving channels for handoff voice calls, while the other is used to block data calls into the network in order to preserve the blocking performance of voice calls in terms of the dropping probability of handoff calls and the blocking probability of new calls, that is,

$$u(x,w)=\begin{cases}1 & if\ x<C\ and\ w=handoff\ voice\ calls\\ 1 & if\ x<T_2\ and\ w=new\ voice\ calls\\ 1 & if\ x<T_1\ and\ w=new\ data\ calls\ or\ \ handoff\ data\ calls\\ 0 & if\ x=C\end{cases} \tag{22}$$

DTR assumes that the bandwidth requirement of voice and data are the same. The equations for blocking probabilities of DTR are derived using a two-dimensional Markov chain and the effect of different values for number of guard channels on dropping and blocking probabilities are studied, but no algorithm for finding the optimal number of guard channels is given. Beigy & Meybodi (2003b) and Beigy & Meybodi (2003c) proposed two algorithms to find the optimal values of T_1 and T_2 for a single cell and multi-cells system, respectively when the average channel holding times for new and handoff calls are the same. Tzeng & Lu (Tzeng & Lu, 2008) designed a call admission control scheme that uses two thresholds; one threshold is used to determine whether or not to accept a new call arrival into a cell, and the other threshold is used to limit the total

Figure 10. State transition diagram for multi-threshold guard channel scheme



number of calls in a cell. The objective of this scheme aims to satisfy the total completion time requirement of mobile users while maximizes channel utilization.

Beigy & Meybodi (2005a) generalized the idea of two-threshold guard channel scheme to multi-classes and *multi-threshold guard channel scheme* for N classes of calls was introduced. In this scheme, a homogenous cellular network was considered where all cells have the same number channels C and experience the same call arrival rates for all types of calls. In each cell, the arrival of calls of class k ($k=1,...,N$) is Poisson distributed with arrival rate λ_k and the channel holding time of calls of class k is exponentially distributed with the same mean $1/\mu$. Thus, the total call arrival rate is $\Lambda_0 = \lambda_1 + \lambda_2 + ... + \lambda_N$. Assume that the calls of class k has a certain level of QoS such that its blocking probability must be less than q_k. Without loss of generality, it is assumed that $q_1 \geq q_2 \geq ... \geq q_N$. This implies that calls for class k require fewer resources than calls of class $k+1$, i.e. calls for class $k+1$ have a higher priority than calls of class k. To provide the specific level of QoS for calls, the allocated channels of each cell are partitioned into N subsets. In order to partition the channel sets, ($N-1$) thresholds, $T_1, T_2, ..., T_{N-1}$ ($0 \leq T_1 \leq T_2 \leq ... \leq T_{N-1} \leq C$ are used. For the sake of simplicity, two additional fixed thresholds $T_0 = -1$ and $T_N = C$. The procedure for accepting calls in multi-threshold guard channel scheme is given in equation (23) can be described as follows. A call from class k is accepted when the number of busy channels is smaller than T_k; otherwise the call is blocked.

$$u(x,w) = \begin{cases} 1 & if\ x < T_w \\ 0 & otherwise \end{cases} \tag{23}$$

Let $c(t)$ denote the number of occupied channels in the given cell and $\Lambda_k = \sum_{j=k+1}^{N} \lambda_j$, $\alpha_k = \frac{\Lambda_k}{\Lambda_0}$, and $\rho = \frac{\Lambda_0}{\mu}$. In the multi-threshold guard channel scheme, $c(t)$ is a continuous-time Markov chain (birth-death process) with states $0, 1, ..., C$. Figure 10 shows the state transition diagram of a system with C channels and multi-threshold guard channel scheme. The system is modeled by a typical M/M/C/C queuing model.

The steady state probability P_n that n (for $T_k \leq n \leq T_{k+1}$) channels are busy is given by the following expression:

$$P_n = P_0 \frac{(\rho\alpha_k)^n}{n!} \prod_{j=1}^{k} \left(\frac{\alpha_{j-1}}{\alpha_j}\right)^{T_j}, \tag{24}$$

where

$$P_0 = \left[\sum_{k=0}^{N-1} \prod_{j=1}^{k} \left(\frac{\alpha_{j-1}}{\alpha_j}\right)^{T_j} \sum_{n=T_k+1}^{T_{k+1}} \frac{(\rho\alpha_k)^n}{n!} \right]^{-1}. \tag{25}$$

Given these state probabilities, the blocking probability of calls of class k is calculated using the following equation.

$$B_k(T_1, ..., T_N) = P_0 \sum_{n=T_k+1}^{T_N} P_n. \tag{26}$$

Properties of $B_k(T_1, ..., T_N)$ have been studied in (Beigy & Meybodi, 2005a). It was shown that $B_k(T_1,...,T_N)$ is a monotonically increasing function of T_k and a monotonically decreasing function of T_j ($j \neq k$). Algorithm 5 can be used for finding the optimal values of thresholds $T_1,...,T_{N-1}$ for the following problem: given C channels allocated to a cell, the objective is to find the optimal values of $T_1,...,T_{N-1}$ in a such a way that it minimizes $B_1(T_1,...,T_N)$ subject to the constraints $B_k(T_1,...,T_N) \leq q_k$ (for k=2,...,N).

Algorithm 5: The algorithm for finding the optimal values of $T_1,...,T_{N-1}$

```
Algorithm FindMTGCParameters
set T0← -1; T1← T2← T3← ....← TN←
C;
if BN(T1,T2,...,TN) ≤ qNthenreturn
(T1,T2,...,TN)
end if
for k ← N down to  2 dowhile Tk-1
> 0 and Bk(T1,T2,...,TN) > qkdoif
not MinBlockCheck (T1,T2,...,TN,
k+1) thenfor m ← 1 to k-1 doset
Tm ← Tm-1
end for
end if
end while
end for
if there is at least one class
that QoS is not satisfied then
return 'the number of assigned
channel to this cell is small'
end if
return (T1,T2,...,TN)
end Algorithm
function MinBlockCheck (T1,...,TN
, k)
if k = N and Tk-1 < Tkand
Bk(T1,T2,...,Tk-1+1,...,TN) ≤ qk-
thenset Tk-1← Tk-1+1
return trueelse if k < N
and not MinBlockCheck (
T1,...,TN, k+1) and Tk-1 < Tkand
Bk(T1,T2,...,Tk-1+1,...,TN) ≤ qk-
thenset Tk-1← Tk-1+1
return true
end if
return false
end function
```

Beigy & Meybodi (2005a) also considered the problem of finding a call admission control scheme that minimizes the number of required channels while preserving the QoS level for all priority levels (all classes of calls) and Algorithm 6 is given to find such optimal number of channels and guard channels.

Algorithm 6: The algorithm for finding the optimal values of $T_1,...,T_{N-1}$

```
Algorithm FindMTGCMinChannels
set T0← -1; T1← T2← T3← ....← TN←
0;
while at least one constraint is
not satisfied do
MinChannelCheck (T1,...,TN, 1)
end if
for k ← N down to  1 dowhile Tk-1
> 0 and Tk-1 < Tkand all con-
straints when set Tk is set to
Tk ← Tk+1 are satisfied doset Tk←
Tk+1
end while
end for
return (T1,T2,...,TN)
end Algorithm
function MinChannelCheck (T1,...,TN
, k)
if k = N and BN(T1,T2,...,TN) >
qkthenset TN← TN+1
return trueelse if k < N
and not MinChannelCheck (
T1,...,TN, k+1) and Tk < Tk+1and
Bk(T1,T2,...,Tk+1,...,TN) > qk-
thenset Tk← Tk+1
return true
end if
```

```
return false
end function
```

In reserving integral number of channels, a number of channels are exclusively reserved for highest priority calls which results in less channels available to lowest priority calls and hence the total carried traffic suffers. In these schemes, if only the blocking probability of highest priority calls is considered, these schemes give very good performance, but the blocking probability of lowest priority calls is degraded to a great extent. This effect can be degraded by reserving fractional number of channels.

In schemes that reserve fractional number of channels, the call admission controller has more control on both the dropping probability of handoff calls and the blocking probability of new calls. When only two groups G_1 and G_2 (one for new calls and the other for handoff calls) are considered this policy is referred to as *limited fractional guard channel scheme* (LFG) in which a fractional number of channels is reserved in each cell exclusively for handoff calls (Ramjee & Towsley & Nagarajan, 1997). The LFG scheme uses an additional parameter p and operates the same as the guard channel policy except when T channels are occupied in the cell, in which case new calls are accepted with probability p, that is,

$$u(x,w) = \begin{cases} 1 & if\ x < C\ and\ w = handoff\ calls \\ 1 & if\ x < T\ \ and\ w = new\ calls \\ p & if\ x = T\ \ and\ w = new\ \ calls \\ 0 & if\ x = C \end{cases} \tag{27}$$

Figure 11 shows the state transition diagram of a homogeneous network with C channels and LFG scheme.

The steady state probability P_n that n channels are busy is given by the following expression:

$$P_n = \begin{cases} \dfrac{\rho^n}{n!} P_0 & if \quad n \le T \\[2ex] \gamma\alpha^{-(T+1)} \dfrac{(\rho\alpha)^n}{n!} P_0 & if \quad T < n \le C \end{cases} \tag{28}$$

where

$$P_0 = \left[\sum_{k=0}^{T} \left(\frac{\rho^k}{k!} \right) + \gamma\alpha^{-(T+1)} \sum_{k=T+1}^{C} \frac{(\rho\alpha)^k}{k!} \right]^{-1}, \tag{29}$$

$\gamma = [\alpha + (1-\alpha)p]$, $\alpha = \lambda_h / (\lambda_n + \lambda_h)$ and, and $\rho = (\lambda_n + \lambda_h) / \mu$. Given these state probabilities, we can drive the blocking probability of new calls and the dropping probability of handoff calls.

$$B_n(C,T,p) = (1-p)\frac{\rho^T}{T!} + \gamma\alpha^{-(T+1)} \sum_{m=T+1}^{C} \frac{(\rho\alpha)^m}{m!}$$

$$B_h(C,T,p) = P_0 \gamma\alpha^{-(T+1)} \frac{(\rho\alpha)^C}{C!} \tag{30}$$

It has been shown that B_n (B_h) is a monotonically increasing (decreasing) function of $T + p$ and therefore there is an optimal pair (T^*, p^*), which minimizes the blocking probability of new calls subject to the hard constraint on the dropping probability of handoff calls. The following algorithm (Algorithm 7) can be used to obtain the optimal pair (T^*, p^*) (Ramjee & Towsley & Nagarajan, 1997).

Algorithm 7: The algorithm for finding the optimal pair (T^*, p^*),

```
Algorithm FindLFGParameter
set upper ← 1; lower ← 0
if(B_h (C,C,0)≤ P_h) thenreturn
(C,0)
end if
```

```
if(B_h (C,0,0) ≥ P_h) then
return (0,0)
end if
while ((upper -lower) < 0.0001)
doset p ← ⌊(upper + lower) /2⌋
if(B_h (C, ⌊p⌋, p-⌊p⌋)> P_h) thenset
upper ← p
else
set lower ← p
end if
end while
return (⌊p⌋, p-⌊p⌋)
end Algorithm
```

Vazquez-Avila & Cruz-Perez & Ortigoza-Guerrero (2006) compared uniform fractional channel scheme, limited fractional channel, scheme, guard channel scheme from different performance criteria.

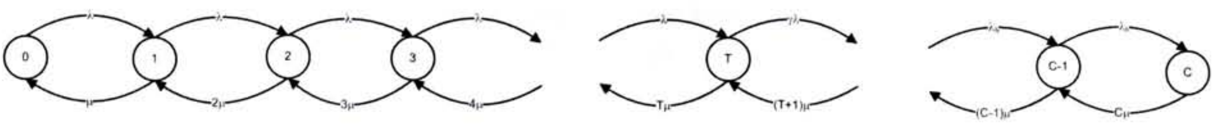
New call bounding schemes: In new call bounding schemes, new calls are accepted if the number of channels used by new calls is less than a threshold (bound for new call) provided that the cell has enough channels for allocating to the incoming new calls. In other words, the state, x, of a cell is defined as the number of ongoing new calls in the cell. Fang & Zhang (2002) proposed a new call bounding scheme in which $T_w s$ are integers. In this scheme, when a new call arrives, if the number of new calls in a cell exceeds a threshold then the new call is blocked; otherwise it will be accepted and the handoff call is rejected only when all channels in the cell are occupied. The idea behind this scheme is that we would rather accept fewer new calls than dropping the ongoing calls in the future, because customers are more sensitive to the call dropping than the call blocking. In (Chung & Chiu, 2002), a new call bounding scheme is given for integrated voice/data wireless networks. In this scheme, it is assumed that the number of ongoing data calls always is constant. This scheme accepts the incoming voice request if the number of voice connections is less than the voice threshold T_1. Since the number of data connections is fixed, there is no call admission control for data connections. In (Chung & Chiu, 2002), no algorithm is given to determine the optimal value of T_1.

Fang (2003) proposed a call admission scheme, which is a generalization of fractional guard channel and new call bounding schemes for multiple classes of calls. This scheme accepts calls with a certain probability, which is determined by the number of busy channels belonging to the priority level of the arriving call in the cell. Fang (2003) analyzed the blocking probabilities of calls when all classes of calls have the same average channel holding time. Wang & Fang & Pan (2008) are studied two variants of the call admission scheme given in (Fang, 2003) for the case that different classes of calls have arbitrary channel requirements and different average channel holding times and their blocking performance analysis are carried out using multi-dimensional Markov process. The first variant uses the information about the total amount of busy channels (bandwidth units) and the second variant utilizes the number of users belonging to the same priority level.

Dynamic EASWR schemes based on teletraffic analysis: In these schemes, function *u(x,w)* is adapted based on the estimated traffic. Since all ongoing calls in the neighboring cells are potential handoff calls to the test cell, these schemes estimate the handoff arrival rate as a function of the number of ongoing calls in the neighboring cells. In these schemes, the number of reserved channels can be an integral number or a fractional number.

The *linear weighting scheme* is given in (Acampora & Naghshineh, 1994a; Acampora & Naghshineh, 1994b) uses the mean number of ongoing calls in the neighboring cells, *I*, within a maximum cell distance *d* from the test cell in determining of the call admission. Let S_d denotes the set of cells in a maximum cell distance *d* from the test cell and c_i denotes the number of ongo-

Figure 11. State transition diagram for limited fractional guard channel scheme



ing calls in the neighboring cell *i*. In this scheme, the state of the system at each time instant is defined as

$$x = \left[\frac{1}{|S_d|} \sum_{i \in S_d} c_i \right] \tag{31}$$

In linear weighting scheme, the new calls are only accepted to the originating cell if

$$u(x,w) = \begin{cases} 1 & \text{if } x < C \text{ and } w = \text{handoff calls} \\ 1 & \text{if } x < T_1 \text{and } w = \text{new calls} \\ 0 & \text{if } x = C \end{cases} \tag{32}$$

Note that the guard channel scheme is a special case of this algorithm where S_d=*i*. Peha & Sutivong (2001) proposed a call admission scheme called *weighted sum scheme,* which uses the weighted sum of the number of ongoing calls in the test cell and in the neighboring cells in determining the admission. Let c_i be the mean number of ongoing calls in the neighboring cells with distance I and p_i be the weight of these cells such that $\sum_{i=1}^{\infty} p_i = 1$ and $p \geq 0 (for\ i \geq 0)$. The state of system in weighted sum scheme at each time instant is defined as $x = \left[\sum_{i=0}^{\infty} p_i c_i \right]$. In this scheme, the new calls are only accepted to the originating cell if

$$u(x,w) = \begin{cases} 1 & \text{if } x < C \text{ and } w = \text{handoff calls} \\ 1 & \text{if } x < T_1 \text{ and } w = \text{new calls} \\ 0 & \text{if } x = C \end{cases} \tag{33}$$

The optimal value of weights p_i can be determined experimentally. The *distributed call admission scheme, proposed by*Naghshineh & Schwartz (1996), does not need the exchange of status information upon the arrival of calls (new and handoff calls). Rather, it only requires the exchange of such information periodically. The admission control algorithm calculates the maximum number of calls that can be accepted in the test cell without violating the QoS of the existing calls in that cell as well as calls in its neighboring cells. One of the main features of this scheme is its simplicity in that the admission decision can be made in real time and does not require much computational effort but this scheme cannot always guarantee the target call dropping probability.

Yu & Leung (1997) introduced a dynamic guard channel scheme in which each base station dynamically adapts the number of channel to be reserved based on the current estimates of the rate at which mobiles in the neighboring cells are likely to incur a handoff into this cell. The objective of the adaptation algorithm is to maintain a specified level of QoS for handoff calls despite of the temporal fluctuations in the traffic into the cell. The determination of the number of channels to be reserved is based on an analytical model which relates number of reserved channels to the dropping probability of handoff calls and the blocking probability of new calls.

In (Oliveria & Kim & Suda, 1998), the number of channels that must be reserved is estimated according to the requested bandwidth of all ongoing connections. Each base station keeps monitoring the dropping probability of handoff calls and the utilization of channels in its cell. Then base station according to this information adjusts the number of guard channels. Lee & Park (1998) proposed a call admission algorithm in which when a new or a handoff call arrives at the test cell, a number of channels in the neighboring cells is reserved. The number of channels to be reserved varies dynamically depending on the number current ongoing calls in the test cell and its neighboring cells.

Choi & Shin (1998) have proposed a scheme based on prediction of the probability that a call will be handed off to a certain neighboring cell from aggregate history of handovers in each cell and determines the number of reserved channels. In this scheme, each base station records the number of handoff failures and adjusts the reservation by changing the estimation window size. Boumerdassi & Beylot (1999) proposed a call admission algorithm for multi-rate personal communication networks in which the number of channels that must be reserved is determined periodically based on the estimated parameters, such as handoff rate. In the beginning of each period, the traffic parameters are estimated and it is assumed that for a given period, traffic parameters are fixed. In this scheme, when the number of occupied channels reaches the threshold T_1, the cell reserves a resource in the neighbors for which the probability of transition is high. If they have free channels, the reservation takes place immediately; otherwise, the algorithm waits for a free channel.

In (Ramanathan & Sivalingam & Agrawal & Kishore, 1999), two dynamic EASWR algorithms are given for wireless networks that support several types of traffic such as voice, data, and video applications, each with different channel requirements. The objective of these algorithms is to accept all handoff calls. Then the base station accepts new calls if and only if the additional channels need to accept all incoming handoff calls (the number of channels to be reserved) and this new call is available. The number of reserved channel is determined according to the estimation of the exact arrival time and channel requirements of future handoff calls. An extension of guard channel scheme is given in (Bozinovski & Popovski & Gavrilovska, 2000; Bozinovski & Popovski & Gavrilovska, 2000). This scheme operates same as the guard channel scheme when a new call arrives and $x < T_1$ or $x = C$; when $T_1 \leq< C$, the algorithm estimates the dropping probability of handoff calls during a period. Then the algorithm accepts new call if the estimated dropping probability of handoff calls is less than the predetermined QoS; otherwise reject the new call. A dynamic channel reservation algorithm, which is presented in (Rappaport & Purzynski 1996), the number of channel to be reserved in each cell is determined dynamically based on the number of ongoing calls in the neighboring cells. This scheme ensures that QoS is maintained in all cells.

Beigy and Meybodi (in press) proposed two learning automata based algorithms to determine the near optimal number of the guard channels when the parameters traffic parameters are unknown and possibly time varying. In these algorithms, learning automata are used to adapt the number of guard channels as the network operates. Let $g(t)$ be the number of guard channels at time instant t which takes values in interval $[g_{min}, g_{max}]$, (for $0 \leq g_{min} < g_{max} \leq C$). In these algorithms, each base station uses one learning automaton with action set $\underline{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ alpha}, where $r = g_{\max} - g_{\min} + 1$. Selection of action α_i by learning automaton means that the base station uses $g(t) = g_{\min} + \alpha_i - 1$ guard channels.

The operation of these algorithms can be described as follows. These algorithms accept handoff calls as long as the cell has free channels. When a new call arrives at a given cell, the learning automaton associated to this cell chooses one of its actions, say α_i. If the cell has at least $g_{\min} + \alpha_i - 1$ free channels, then the call will be accepted; otherwise it will be blocked. Then the base station computes the current estimate of the dropping probability of handoff calls $\hat{B}_h$ and based on the result of comparison of this quantity with the specified level of QoS (p_h), the reinforcement signal will be produced and the action probability vector of the learning automaton will be updated using a learning algorithm. The differences between these algorithms are the way that they produce reinforcement signal for the learning automata and learning algorithm used to update the action probability vector.

The first algorithm uses a SL_{R-I} learning automaton in each cell and the reinforcement signal at time instant n is equal to $\psi\left(\left|\hat{B}_h - p_h\right|\right)$, where $\psi : R \longrightarrow [0,1]$ is a projection function. The projection function is considered to be a continuous, nondecreasing and nonnegative function that maps the set of real numbers into [0,1], for example $\psi(x) = x$ can be a projection function, which maps [0,1] into [0,1]. The continuity of the projection function is needed because the response produced by the environment is a real number in interval [0,1], its nonnegativity is needed in order to maintain the reward and penalty nature of updating, and the nondecreasing property is needed for preserving the relative strength of the reinforcement signal. It is obvious that when $\hat{B}_h$ is far from p_h, and then the reinforcement signal will be large, which causes the selected action of the learning automaton to be penalized. When $\hat{B}_h$ is near to p_h, the reinforcement signal will be small and near to zero which causes the selected action of the learning automaton to be rewarded. In other words, when $\hat{B}_h$ is greater than p_h, the chosen number of guard channels is too small and when $\hat{B}_h$ is smaller than p_h, the number of guard channels chosen by learning automaton is large. In other words, the reinforcement signal is an indicator of the relative distance of the dropping probability of handoff calls to the predefined level of QoS.

Simulation results showed that the blocking probability of new calls for the first algorithm is lower than the blocking probability of the guard channel algorithm, but it can not maintain the predefined level of QoS, as evidenced by the results of simulation. The second algorithm tries to minimize the blocking probability of new calls and at the same time to maintain the specified level of QoS. This algorithm uses a L_{R-I} learning automaton in each cell for determination of the number of guard channels. The selected action of learning automaton in a cell will be rewarded if the incoming new call is accepted and the current estimate of dropping probability of handoff calls

Figure 12. Blocking probabilities of new calls for learning automata based dynamic guard channel algorithms

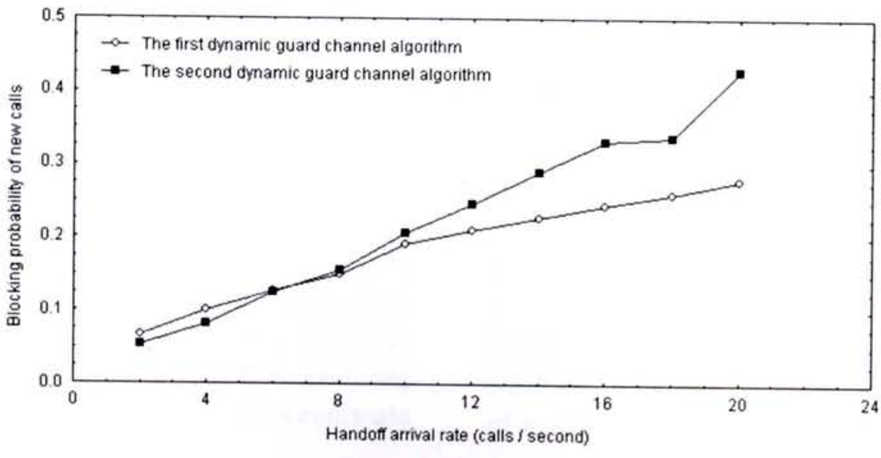
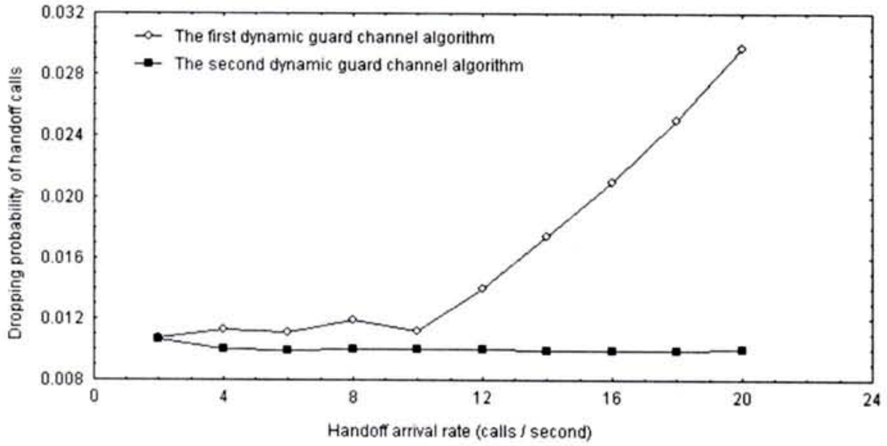


Figure 13. Dropping probabilities of handoff calls for learning automata based dynamic guard channel algorithms



$\hat{B}_h$ is less than the specific level of QoS (p_h) or the incoming new call is rejected and the current estimate of dropping probability of handoff calls is greater than the specific level of QoS; the selected action neither rewarded nor penalized otherwise. Figures 12 and 13 show the performance of these algorithms.

Yu & Leung (1996) proposed a call admission algorithm is given, in which when a new or handoff call arrives at a neighboring cell, number of channels that must be reserved in the test cell is increased by a fraction amount and when a call is completed at or moved out of the neighboring cells, the number of reserved channels is decreased by the same fractional amount. Han & Nilsson (2000) proposed a *population-based channel reservation scheme*. This scheme dynamically adjusts the number of channels that must be reserved for handoff calls according to the amount of cellular traffic in its neighboring cells. Assume that cell i have n_i neighboring cells. Whenever a call which consumes b channels is accepted into cell j as either a newly call or a handoff call, the base station of the cell requests a fractional channel reservation for the amount of b / n_j to each of its n_j neighboring cells. Whenever this call is leaving the cell either by call completion or by handoff into one of its neighboring cells, the base station requests a fractional channel release for the same amount as requested for the reservation to each of its n_j neighboring cells, even to the cell into which this call is handed over. This step is to inform the neighborhood of appearance and disappearance of a potential handoff. Each base station network maintains a counter that records transactions for fractional channel reservation or release requests from its neighboring cells. Every time it receives a fractional channel reservation request or a release request, it increments or decrements the counter by the requested amounts, respectively.

Beigy and Meybodi (2005b) proposed an adaptive limited fractional guard channel algorithm for two classes of calls: new and handoff calls. The objective of this algorithm is to adapt parameter $T + p$ in such a way that minimizes the blocking probability of new calls subject to the constraint that the dropping probability of handoff calls be at most p_h. Since $T + p$ is a continuous parameter, the algorithm uses a continuous action-set learning automaton for adaptation of the value of parameter $T + p$. Let $x(n) = T(n) + p(n)$ be the parameter of the limited fractional guard channel algorithm at instant *n*, and *x(n)* takes values in the interval $\left[x_{\min}, x_{\max}\right]$, where $0 \leq x_{\min} < x_{\max} < C$. The action-set for learning automaton is the real line and it uses the Gaussian distribution, $N(\mu, \sigma)$, to choose its actions. This Gaussian distribution is updated using the reinforcement signal, which is emitted from the environment. Initially, the learning automaton chooses one of its actions with equal probability using a Gaussian distribution with a large variance. Since *x(n)* and *μ(n)* must be in the interval $\left[x_{\min}, x_{\max}\right]$, the above mentioned learning automaton cannot be used directly to adapt the value of $T+p$, and hence a projected version of learning automaton will be used. In the projected version of learning automaton, a constraint set $H = \left\{y \middle| x_{\min} \leq y \leq x_{\max}\right\}$ is used for updating μ as well as choosing actions of learning automaton. In the projected version, when the updated value of μ goes outside of the constraint set *H*, then μ is pushed into *H* and when the action chosen by the learning automaton does not belong to *H*, then the action is pushed into *H*. This algorithm can be described as follows. Each base station is equipped with a learning automaton for adapting $T+p$. When a new call arrives at a given cell, the learning automaton associated to that cell chooses one of its actions, say *x(n)*. Let $T(n) = \lfloor x(n) \rfloor$ and $p(n) = x(n) - \lfloor x(n) \rfloor$. If the number of busy channels of a cell is less than *T(n)*, then the incoming call will be accepted; when the cell has *T(n)* busy channels, then a call will be accepted with probability $p(n)$; otherwise the incoming call will be blocked. On the arrival of a new call the base station computes the current estimate of the dropping probability of the handoff calls and based on the result of the comparison of this quantity with the specified level of QoS, then it computes the reinforcement signal as $\left|\hat{B}_h - p_h\right|$ and the learning automaton updates its action probability distribution. Beigy and Meybodi (2005b) showed that this algorithm finds the optimal number of channels that must be reserved.

In (Beigy & Meybodi, 2002a; Beigy & Meybodi, 2002b), two adaptive limited fractional guard channel algorithm based upon continuous action-set learning automata are reported. These algorithms adjust the number of channels to be reserved in the cell according the traffic of the cell and the predefined QoS. The differences between these algorithms are the learning algorithm used for learning automata and the ways that the reinforcement signal will be produced. Salamah & Lababidi (Salamah & Lababidi, 2005) proposed an adaptive channel reservation scheme for cellular networks. In this algorithm, the base station measures the signal strength to predict the handoff. When there is no handoff in the new future, some of the reserved channels can be used for new calls.

Dynamic EASWR schemes based on mobility: The most salient feature of the mobile wireless network is the mobility, which can be used for adjusting the $T_w s$. Since the handoff occurs when the mobile users are moving during the call connection, thus good call admission control algorithms should consider the mobility pattern. Hence, in order to make a reservation schemes effectively adapt to the network traffic situations, the user mobility information must be deployed. In these schemes, each base station adjusts the reservation by employing the mobility information. The mobility pattern is influenced by many factors such as destinations of mobile users, the layout of the network, and the traffic condition in the network. Since it is not easy to specify the mobility pattern of each mobile user in detail, therefore the statistical mobility patterns of users are more useful. Based on the values of thresholds, $T_w s$, these schemes can reserve an integral or fractional number of channels.

Concept of *shadow cluster*, which is introduced by Levine & Akylidiz & Naghsineh (1997), estimates the future resource requirements based on the current movement pattern of the mobile users. The fundamental idea of the shadow cluster concept is that as an active user travels to other cell, the region of influence also moves. The base stations currently being influenced are said to form a shadow cluster, because the region of influence follows the movement of the active mobile terminal like a shadow. However, the strength of this scheme depends on the accuracy of the knowledge of users' movement patterns, such as trajectory of a mobile user, which is difficult to predict in real time systems. Hou & Fang (2001) proposed an *integral mobility based channel reservation scheme* in which mobile users are classified in two classes according to their velocities: high and low speed users. Thus the average cell dwell time of high speed users are shorter than that of the low speed users. Based on the velocity of each mobile user, the handoff probability of each class is predicted and the number of channels that must be reserved is determined. It is also noted that the better performance will be achieved if this scheme and a new call bounding scheme are combined. Hu & Sharma (2003) proposed a dynamic reservation scheme for multimedia cellular networks in which the handoff calls have a higher priority than the new calls. The prerequisite of this scheme is that base stations can estimate future trajectory of mobile computers with high degree of accuracy, which is possible in today's increasing improved position location techniques. This scheme uses the Kalman filter to predict the next cell for every mobile computer. Huang and et. Al. (Huang & Chuang & Yang, 2008) proposed a reservation based adaptive call admission algorithm in which a fuzzy logic system is used to estimate the number of channels to be reserved for handoff calls and particle swarm optimization (PSO) technique used to adjust the parameters of the membership functions in the fuzzy logic system. In (Martinez-Bauset & Gimenenz-Guzman & Pla, 2008) the problem of optimizing admission control policies in mobile multimedia cellular networks when predictive information regarding the movement of mobile terminals is available was studied. For the optimization process a reinforcement learning approach was used.

When a fractional number of channels are reserved, $T_w s$ are real numbers and the call admission controller have more control on both the dropping probability of handoff calls and the blocking probability of new calls because the rounding of T_w 's lost some information. *Fractional mobility based channel reservation scheme* is given in (Hou & Fang, 2001) in which mobile users are classified in two classes according to their velocities: high and low speed users. Thus the average cell dwell time of high speed users are shorter than that of the low speed users. Based on the velocity of each mobile user, the handoff probability of each class is predicted and values of T_w 's are determined.

4.2.2 Complete Partitioning Policies

Complete partitioning policies are subsets of reservation based call admission policies. In these policies, we have $S = S_{G_1} \cup S_{G_2} \cup \ldots \cup S_{G_W}$ Complete partitioning policies partition the channels among the different classes of calls by dedicating a certain number of channels to each class. This policy takes place when the threshold point for traffic class w is inside the state space, i.e. $T_w \in S$. These policies isolate each class of calls and the resulting process is simply the aggregation of N independent $M / M / T_w / T_w$ processes. Leong & Zhuang (2002) considered a cellular network that supports two traffic types of voice (constant-rate) and data (variable-rate). In this scheme, voice calls have a higher priority than the new calls. The channels in each cell are partitioned into two subsets, one for voice calls and the other for data calls. Each partition uses the standard LFG policy to accept/reject new calls in that class. Ahn & Kim (2003) proposed a dynamic channel allocation for multimedia cellular networks that uses the guard channel scheme for maintaining the level of QoS and works the same as the guard channel scheme when a new or handoff call arrives, but when a call is terminated or completed, it differs from the guard channel policy. If a call that uses a guard channel is terminated or completed, then that channel is reserved for future incoming handoff calls. On the other hand, if a call that uses an ordinary channel is terminated or completed, then the bandwidth adaptation, which is the allocation of freed bandwidth to the ongoing calls, is applied. In order to allocate the freed bandwidth to the ongoing calls, a Lagrangean relaxation procedure is used that leads to a sub-optimal solution. Kulavaratharash & Aghvami (1999) divided channels of a cell into two groups: ordinary and guard channel groups. The new calls are accepted if the ordinary channel group has free channel; otherwise the call will be blocked. For handoff calls, three different strategies are used: 1) first guard channel group and then the ordinary channel group is selected, 2) first ordinary channel group and then the guard channel group are selected, and 3) randomly one of the preceding strategies is selected. In order to improve the blocking probability of new calls without trading off the dropping probability of handoff calls, an algorithm is given in (Kulavaratharash & Aghvami, 1999). In this algorithm, if all channels in the ordinary channel group are occupied at the arrival time of a new call and there is at least one free channel in guard channel group, then any free guard channel can temporarily be lent to the ordinary channel group to prevent the new call to be blocked. Such transferring can only be carried out if the base station can predict that there are no handoff attempts from neighboring cell, while the borrowed channel is used for the new call. This prediction is done with the aid of power measurements. AlQahtani & Mahmoud (AlQahtani & Mahmoud, 2008) extended complete partitioning and the queuing priority call admission schemes for operation in 3G WCDMA networks. In their complete partitioning, each class of calls has its own queue and resource partition whereas in queuing priority, each call class has its own queue and all classes share the available resources. Then they develop an analytical model for the queuing priority algorithm to study the behavior of this algorithm.

4.3 Queuing Priority Schemes

These schemes reduce the blocking probability of new calls and the dropping probability of handoff calls by employing a queuing mechanism. In queuing priority schemes, calls of each class are accepted whenever there is a free channel for that class. When there is no free channels for a class, calls may be queued and calls of other classes are blocked and cleared from system. One key point of using queuing in call admission control algorithms is that the service differentiation could be managed by modifying the queuing discipline. For example, instead of FIFO queuing strategy, other prioritized queuing discipline can be used to maintain priority level in each service class. Another key point is the mobility of the users, which results difficulties in management of queue. These schemes consider two traffic classes, new calls and handoff calls. Based on the type of calls that is queued, these schemes are divided in three groups: *new call queuing schemes*, *handoff call queuing scheme* and *all call queuing schemes*. Some of the reported schemes are briefly described below.

4.3.1 New Call Queuing Schemes

In a new call queuing scheme, a certain number of channels is reserved in each cell exclusively for handoff calls. In new call queuing schemes, the new calls and the handoff calls are treated equally on a first-come first-served basis for channel allocation until the number of occupied channels in the cell becomes T_1. When the predetermined channel utilization threshold, T_1, is reached, new calls are queued and only handoff call requests are accepted. In other words, a new call is accepted if $c < C - T_1$, where $T_1 \geq 0$ is the number of channels reserved specifically for handoff (guard channels), that is,

$$u(x,w) = \begin{cases} accept & if\ x < C\ and\ w = handoff\ calls \\ accept & if\ x < T_1 and\ w = newcalls \\ queue & if\ x \geq T_1 and\ w = new\ calls \\ reject & if\ x = C \end{cases} \quad (33)$$

The only reported new call queuing scheme is given in (Guern, 1998). In this scheme, when the number of free channels is less than the number of guard channels, the new calls are queued. It is pointed out that the blocking probability of new calls can be drastically reduced by reserving some channels for handoff calls and using a queuing mechanism for new calls.

4.3.2 Handoff Calls Queuing Schemes

Handoff queuing schemes reserves a number of channels for use of handoff calls. In these schemes, the new calls are serviced as same as handoff calls until the number of free channels becomes less than the number of reserved channels $\left(C - T_1\right)$. When the number of occupied channels is greater than threshold T_1, new calls are blocked and handoff call requests are accepted. When all channels are occupied, the handoff calls are queued, that is

$$u(x,w) = \begin{cases} accept & if\ x < C\ and\ w = handoff\ calls \\ queue & if\ x = C\ and\ w = handoff\ calls \\ accept & if\ x < T_1\ and\ w = new\ calls \\ reject & if\ x \geq T_1\ and\ w = new\ calls \end{cases} \quad (34)$$

Hong and Rappaport analyzed handoff queuing scheme with an infinite buffer for handoff calls and this scheme with finite buffer is analyzed in (Yoon & Kwan, 1993). The extension of handoff queuing scheme with finite buffer size to multi-class of calls is proposed in (Tian & Ji, 2001). Agrawal & Anvekar & Naredran (1996) introduced a handoff call queuing scheme, which reserves no channels for handoff calls. In this scheme, when a new call arrives and all channels are busy, then the call will be blocked; when a handoff call arrives and all channels are busy, the call will be queued. Both types of calls will be accepted if there are any free channels. When a channel becomes free, then a handoff call from the queue, if queue is not empty, will be serviced. Agrawal & Anvekar & Naredran (1996) also proposed some queuing discipline such as first-in first out, most critical first. Cho & Ko & Kwang (1997) proposed a dynamic channel reservation scheme with handoff queuing. In this scheme, the number of channels to be reserved is adjusted based on the handoff traffic and the current number of reserved channels. Zheng & Lam (2002) introduced a dynamic channel reservation scheme with handoff queuing in which the number of channels to be reserved is adjusted based on the occupied channels in the neighboring cells. It must be pointed out that queuing of handoff calls is more sensitive to delay (time between request and the time for allocation of channels) in the service than queuing of new calls, because as mobile users move the signal strength decreases and the call may be dropped. However, this delay depends on the speed of the mobile user.

4.3.3 All Calls Queuing Schemes

These schemes wok as same as guard channel scheme when the number of occupied channels in the cell is less than T_1. When the number of occupied channels is equal or greater than T_1, new calls are queued and only handoff call requests are accepted. When all channels are occupied, the handoff calls are also queued, that is

$$u(x,w) = \begin{cases} accept & if\ x < C\ and\ w = handoff\ calls \\ queue & if\ x = C\ and\ w = new\ calls \\ accept & if\ x < T_1\ and\ w = new\ calls \\ queue & if\ x \geq T_1\ and\ w = new\ calls \end{cases} \quad (35)$$

Yoon & Kwan (1993) proposed a call admission scheme in which the value of T_1 is equal to C. In this scheme, the new calls are put after all handoff calls in the queue and the queue is serviced in the FIFO manner. When the queue is full, then all incoming calls will be blocked. Yoon & Kwan (1993) also used a rearranging mechanism in which when the queue is full, then the last new call is pushed out from the queue and the incoming handoff call will be placed after the last handoff call. Chang & Chang & Lo (1999) introduced a call admission scheme in which all calls are queued with certain rearrangements in the queue.

5. OPTIMAL CALL ADMISSION POLICIES

Let assign a cost to each blocked call, low cost for new calls and high cost for handoff calls, the optimal policy is the one that that finds $u(x,w)$ in such a way that the cost is minimized. In these policies, the call admission is formulated as as Markov decision process and actions of this Markov decision process are used as function $u(x,w)$. Saquib & Yates (1995) used value iteration algorithm of Markov decision process as a technique to search for the optimal policy, that is, the policy which minimizes a weighted blocking criterion. In (Kwon & Choi & Naghshineh (1998); Choi & Kwon & Choi & Naghshineh (2000)), a call admission control algorithm is given which focuses the forced termination probability (call dropping probability) as the main QoS requirement. In this approach the cellular system is modeled using semi-Markov decision process. The linear programming method for solving semi-Markov decision process is employed to find out the optimal call admission control decision in each state. Morley & Grover (2000) formulated the call admission problem in dual-mode cellular networks as a Markov decision process and the linear programming is used for finding the optimal call admission policy.

6. FUTURE RESEARCH DIRECTIONS

Voice telephony and short message services were two first applications that mobilized. Now mobile networks support many other services such as email, web browsing, and push to talk by introduction of packet based networks. Current 3G and 3.5G wireless networks are able to cope with several such applications and offer a sufficient bandwidth. Due to the rapid development and growth of mobile communications, there will be a rapid growth in demand for new wireless services in next-generation wireless networks. The next generation wireless networks such as UMTS long term evolution (LTE) and WiMAX will support a wide variety of multimedia services at higher bandwidths. These services have different traffic characteristics, bandwidth requirements, and quality of service requirements. To support such integrated services, call admission control algorithms become more important. Most of the call admission algorithms reported in the literature support only voice service.

One challenge is how to implement handoff in next generation networks with minimum packet loss and handoff latency. Call admission and handoff management in these networks are more complex, as they must cover both horizontal and vertical handoffs. Therefore, fast and seamless handover is a big challenge for these networks. Since these networks must support real-time multimedia applications that require small delay and high-rate data transmission, the future researches on call admission algorithms and handoff management will focus on algorithms for services that have different bandwidth requirements, different quality of service requirements, and delay and cross-layer call admission and handoff management is one of such area.

Table 1. Features of the proposed classification

Call admission algorithm	Ability to have control on blocking / dropping probabilities	Ability to track traffic variation	Usability for delay sensitive applications
Non-prioritized schemes	No	No	Yes
Prioritized schemes			
• Static equal access sharing with priority	Low	No	Yes
• Dynamic equal access sharing with priority	Low	Yes	Yes
• Static reservation based schemes			
o Fractional number of channels	High	No	Yes
o Integral number of channels	Medium	No	Yes
• Dynamic reservation based schemes			
o Fractional number of channels	High	High	Yes
o Integral number of channels	Medium	Medium	Yes
• Queuing schemes	Yes	No	No, it needs careful management of queues

7. CONCLUSION

In this chapter, we proposed a classification of user based call admission policies in mobile cellular networks. The proposed classification not only provides a coherent framework for comparative studies of existing approaches, but also helps future researches and developments of new call admission policies. Much of research has been done in reservation based call admission policies. One critical issue in all reservation based call admission control policies is how the reservation is made. In traditional guard channel policy, the number of guard channels is determined based on the priori knowledge of the cell traffic and the QoS requirements. Obviously, the performance will degrade if the cell traffic is not conformal to the priori knowledge; thus it will be better to use dynamic reservation schemes: adjusting the number of guard channels with the network traffic. In order to determine an optimal or near optimal value for number of guard channels one first answer the following question: when do reserve channels for incoming handoff calls? If the reservation is made at time when it is needed, the resulting scheme will definitely achieve the best performance. However, such timing will be very difficult, if it is not impossible, to acquire. Since the reservation is a waste of resources if it not used by handoff calls, the shorter the time the reservation is actually used (reservation time), the better performance will be achieved. Table 1 summarizes some features of algorithms in our classification.

REFERENCES

Acampora, A. S., & Naghshineh, M. (1994a). An Architecture and Methodology for Mobile-Executed Handoff in Cellular ATM Networks. *IEEE Journal on Selected Areas in Communications*, *12*, 1365–1374. doi:10.1109/49.329339

Acampora, A. S., & Naghshineh, M. (1994b). Control and Quality-of-Service Provisioning in High Speed Microcellular Networks. *IEEE Personal Communications*, 36-43.

Agrawal, P., Anvekar, D. K., & Naredran, B. (1996). Channel management policies for handovers in cellular networks. *Bell Labs Technical Journals*, *1*(2), 97–110. doi:10.1002/bltj.2018

Ahn, K., & Kim, S. (2003). Optimal bandwidth allocation for bandwidth adaptation in wireless multimedia networks. *Computers & Operations Research, 30*, 1917–1929. doi:10.1016/S0305-0548(02)00115-6

AlQahtani, S. A., & Mahmoud, A. S. (2008). Performance analysis of two throughput-based call admission control schemes for 3G WCDMA wireless networks supporting multi-services. *Computer Communications, 31*, 49–57. doi:10.1016/j.comcom.2007.10.029

Ayyagari, D., & Empremides, A. (1999). Admission control with priorities: approaches for multi-rate wireless systems. *ACM Mobile Networks and Applications, 4*, 209–218. doi:10.1023/A:1019103016426

Beigy, H., & Meybodi, M. R. (2002a). An adaptive limited fractional guard channel policy using continuous action learning automata. In *Proceedings of the 10th International Conference on Software, Telecommunications and Computer Networks (SoftCom 2002)*, Croatia, Italy (pp. 536-540).

Beigy, H., & Meybodi, M. R. (2002b). A limited fractional guard channel scheme using learning automata. In *Proceedings of the First EurAsian Conference on Advances in Information and Communication Technology (EURASIA-ICT 2002)*, Shriaz, Iran (pp. 343-347).

Beigy, H., & Meybodi, M. R. (2003a). Adaptive uniform fractional guard channel algorithm: the steady state analysis. In *Proceedings of the 9th International Symposium on Wireless Systems and Networks (ISWSN'03)*, Dhahran, Saudi Arabia (pp. 177-181).

Beigy, H., & Meybodi, M. R. (2003b). An optimal prioritized channel assignment scheme for using in mobile transaction environments. In *Proceedings of the 8th Annual International Computer Society of Iran Computer Conference (*CISCC-2003), Iran (pp. 68-74).

Beigy, H., & Meybodi, M. R. (2003c). An optimal channel assignment scheme. In *Proceedings of the 11th Iranian Conference on Electrical Engineering (ICEE-2003)*, Shiraz, Iran (Vol. 2, pp. 634-641).

Beigy, H., & Meybodi, M. R. (2003c). User based call admission control policies for cellular mobile systems: A survey. *The CSI Journal on Computer Science and Engineering, 1*(2b), 27–40.

Beigy, H., & Meybodi, M. R. (2004a). A new fractional guard channel policy. *Journal of High Speed Networks, 13*(1), 25–36.

Beigy, H., Meybodi, M. R. (2004b). Adaptive uniform fractional guard channel algorithms. *Iranian journal of Electrical and Computer Engineering, 3*(1), 47-53.

Beigy, H., & Meybodi, M. R. (2005a). A general call admission policy for next generation wireless networks. *Computer Communications, 28*, 1798–1813. doi:10.1016/j.comcom.2004.12.009

Beigy, H., & Meybodi, M. R. (2005b). An adaptive call admission algorithm for cellular networks. *Computers & Electrical Engineering, 31*, 132–151. doi:10.1016/j.compeleceng.2004.12.002

Beigy, H., & Meybodi, M. R. (in press). Learning automata based dynamic guard channel algorithms. *Journal of High Speed Networks.*

Boumerdassi, S., & Beylot, A. (1999). Adaptive channel allocation for wireless PCN. *ACM Mobile Networks and Applications, 4*, 111–146. doi:10.1023/A:1019186412792

Bozinovski, M., Popovski, P., & Gavrilovska, L. (2000). Novel strategy for call admission control in mobile cellular network. In *Proceedings of IEEE Conference on Vehicular Technology (VTCC 2000)* (Vol. 4, pp. 1597-1602).

Bozinovski, M., Popovski, P., & Gavrilovska, L. (2000). QoS-Based policy for call admission control in mobile cellular networks. In *Proceedings of IEEE Conference on Wireless Communications and Networking (WCNC 2000)* (Vol. 2, pp. 502-506).

Chang, C., Chang, C. J., & Lo, K. R. (1999). Analysis of hierarchical cellular systems with reneging and dropping for waiting new calls and handoff calls. *IEEE Transactions on Vehicular Technology, 48*(4), 1080–1091. doi:10.1109/25.775357

Chang, K., & Kim, D. (2001). Optimal prioritized channel allocation in cellular mobile systems. *Computers & Operations Research, 28*, 345–356. doi:10.1016/S0305-0548(99)00121-5

Chen, G. C., & Lee, S. Y. (2001). Modeling of static and dynamic guard channel schemes for mobile transactions. *IEICE Transactions on Information and Systems . E (Norwalk, Conn.), 84-D*, 87–99.

Cho, C., Ko, Y., & Kwang, H. L. (1997). Fuzzy adaptive guard channel assignment strategy for handoff in PCS system. *Proceedings of 6th IEEE International Conference on Fuzzy Systems, (pp.* 1511-1518).

Choi, J., Kwon, T., Choi, Y., & Naghshineh, M. (2000). Call admission control for multimedia services in mobile cellular networks: A Markov decision approach. In *Proceedings of 5th IEEE Symposium on Computers and Communications (ISCC 2000)* (pp. 594-599).

Choi, S., & Shin, K. G. (1998). Predictive and adaptive bandwidth reservation for handoffs in QoS-sensitive cellular networks. *ACM SIGCOMM Computer Communication Review, 28*(4), 155–166. doi:10.1145/285243.285278

Chung, S. P., & Chiu, C. L. (2002). Joint call admission control / congestion for wireless integrated voice/data Networks. *Computer Communications, 25*, 1653–1664. doi:10.1016/S0140-3664(02)00050-6

Cruz-Perez, A., & Ortigoza-Guerrero, L. (2007). Fractional resource reservation in mobile cellular systems. In Y. Zhang, H. Hu & M. Fujise (Ed.), *Resource, Mobility, and Security Management in Wireless Networks and Mobile Communications* (pp. 122-138). New York: Auerbach Publishing.

Das, S. K., Sen, S. K., & Jayaram, R. (1998). A novel of load balancing scheme for the tele-traffic hot spot problem in cellular networks. *ACM/Baltzer Journal on Wireless Networks*, 4(4), 325-340.

Fang, Y. (2003). Thinning schemes for call admission control in wireless networks. *IEEE Transactions on Computers, 52*(5), 686–688. doi:10.1109/TC.2003.1197136

Fang, Y., & Zhang, Y. (2002). Call admission control schemes and performance analysis in wireless mobile networks. *IEEE Transactions on Vehicular Technology, 51*, 371–382. doi:10.1109/25.994812

Ghaderi, M., & Boutaba, R. (2006). Call admission control in mobile cellular networks: A comprehensive survey. *Wireless Communications and Mobile Computing, 6*(1), 69–93. doi:10.1002/wcm.246

Guern, R. (1988). Queuing-blocking system with two arrival streams and guard channels. *IEEE Transactions on Communications, 36*, 153–163. doi:10.1109/26.2745

Han, M., & Nilsson, A. A. (2000). Population-based call admission control in wireless cellular networks. In *Proceedings of the 2000 IEEE International Conference on Communications* (pp. 1519-1523).

Haring, G., Marie, R., Puigjaner, R., & Trivedi, K. (2001). Loss formulas and their application to optimization for cellular networks. *IEEE Transactions on Vehicular Technology, 50*, 664–673. doi:10.1109/25.933303

Ho, C.-J., & Lea, C.-T. (1999). Improving call admission policies in wireless networks. *Wireless Networks, 5*, 257–265. doi:10.1023/A:1019107329380

Hong, D., & Rappaport, S. (1986). Traffic modeling and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoffs procedure. *IEEE Transactions on Vehicular Technology, 35*, 77–92. doi:10.1109/T-VT.1986.24076

Hou, J., & Fang, Y. (2001). Mobility-based call admission control schemes for wireless mobile networks. *Wireless Communications and Mobile Computing, 1*, 269–282. doi:10.1002/wcm.18

Hu, F., & Sharma, N. K. (2003). Multimedia call admission control in mobile networks: A dynamical reservation-pool approach. *Computer Networks, 43*(3), 263–288. doi:10.1016/S1389-1286(03)00271-8

Huang, C., Chuang, Y., & Yang, D. (2008). Implementation of call admission control scheme in next generation mobile communication networks using particle swarm optimization and fuzzy logic systems. *Expert Systems with Applications, 35*, 1246–1251. doi:10.1016/j.eswa.2007.08.023

Ishikawa, Y., & Umeda, N. (1997). Capacity design and performance of call admission control in CDMA systems. *IEEE Journal on Selected Areas in Communications, 15*, 1627–1635. doi:10.1109/49.634800

Katzela, I., & Naghshineh, M. (1996). Channel assignment schemes for cellular mobile telecommunication systems: A comprehensive survey. *IEEE Personal Communications, 3*, 10–31. doi:10.1109/98.511762

Kulavaratharash, M. D., & Aghvami, A. H. (1999). Teletraffic performance evaluation of microcellular personal communication networks (PCN's) with prioritized handoff procedures. *IEEE Transactions on Vehicular Technology, 48*, 137–152. doi:10.1109/25.740076

Kwon, T., Choi, Y., & Naghshineh, M. (1998). Optimal distributed call admission control for multimedia services in mobile cellular networks. In *Proceedings of 5th International Workshop on Mobile Multi-Media Communications(MoMuc98).*

Lee, S., & Park, S. (1998). Handoff with dynamic guard channels in ATM-based mobile networks. In *Proc. of 12th International Conference on Information Networking (ICOIN-12)* (pp. 439-442).

Leong, C. W., & Zhuang, W. (2002). Call admission control for voice and data traffic in wireless communications. *Computer Communications, 25*, 972–979. doi:10.1016/S0140-3664(01)00434-0

Levine, D. A., Akylidiz, I. F., & Naghsineh, M. (1997). A resource estimation and call admission algorithm for wireless multi-media networks using the shadow cluster concept. *IEEE/ACM Transactions on Networking*, 5, 1-12.

Martinez-Bauset, J., Gimenenz-Guzman, J. M., & Pla, V. (2008). Optimal admission in multimedia mobile networks with handover prediction. *IEEE Wireless Communications*, 38-44.

Morley, C. D., & Grover, W. D. (2000). Strategies to Maximize Carried Traffic in Dual-Mode Cellular Systems. *IEEE Transactions on Vehicular Technology, 49*(2), 357–366. doi:10.1109/25.832966

Naghshineh, M., & Schwartz, M. (1996). Distributed Call Admission Control in Mobile/Wireless Networks. *IEEE Journal on Selected Areas in Communications, 12*, 711–717. doi:10.1109/49.490422

Oh, S., & Tcha, D. (1992). Prioritized channel assignment in a cellular radio network. *IEEE Transactions on Communications, 40*, 1259–1269. doi:10.1109/26.153371

Oliveria, C., Kim, J. B., & Suda, T. (1998). An adaptive bandwidth reservation scheme for high-speed multimedia wireless network. *IEEE Journal on Selected Areas in Communications, 16*(6), 858–874. doi:10.1109/49.709449

Peha, J. M., & Sutivong, A. (2001). Admission Control Algorithms for Cellular Systems. *Wireless Networks, 7*, 117–125. doi:10.1023/A:1016629421079

Ramanathan, P., Sivalingam, K. M., Agrawal, P., & Kishore, S. (1999). Dynamic resource allocation schemes during handoff for mobile multimedia wireless networks. *IEEE Journal on Selected Areas in Communications*, *17*, 1270–1283. doi:10.1109/49.778185

Ramjee, R., Towsley, D., & Nagarajan, R. (1997). On optimal call admission control in cellular networks. *Wireless Networks*, *3*, 29–41. doi:10.1023/A:1019172226345

Rappaport, S., & Purzynski, C. (1996). Prioritized resource assignment for mobile cellular communication systems with mixed services and platform types. *IEEE Transactions on Vehicular Technology*, *45*, 443–458. doi:10.1109/25.533760

Salamah, M., & Lababidi, H. (2005). Dynamically adaptive channel reservation scheme for cellular networks. *Computer Networks*, *49*, 787–796. doi:10.1016/j.comnet.2005.02.005

Saquib, M., & Yates, R. (1995). Optimal call admission to a mobile cellular network. *Proceedings of IEEE Vehicular Technology Conference* (pp. 190-194).

Tian, X., & Ji, C. (2001). Bounding the performance of dynamic channel allocation with QoS provisioning for distributed admission control in wireless networks. *IEEE Transactions on Vehicular Technology*, *50*, 388–397. doi:10.1109/25.923051

Tzeng, S. S., & Lu, H. Y. (2008). Dual-threshold admission control for non-real-time traffic in wireless data networks. *Computer Communications*, *31*, 2428–2434. doi:10.1016/j.comcom.2008.03.010

Vazquez-Avila, L. J., Cruz-Perez, F. A., & Ortigoza-Guerrero, L. (2006). Performance Analysis of Fractional Guard Channel Policies in Mobile Cellular Networks. *IEEE Transactions on Wireless Communications*, *5*(2), 301–305. doi:10.1109/TWC.2006.1611053

Wang, X., Fan, P., & Pan, Y. (2008). A more realistic thinning scheme for call admission control in multimedia wireless networks. *IEEE Transactions on Computers*, *57*(8), 1143–1148. doi:10.1109/TC.2008.29

Wu, S., Wong, K. Y. M., & Li, B. (2002). A dynamic call admission policy with precision QoS guarantee using stochastic control for mobile wireless networks. *IEEE Transactions on Networking*, *10*, 257–271. doi:10.1109/90.993306

Yavuz, E. A., & Leung, V. C. M. (2006). Computationally efficient method to evaluate the performance of guard channel based call admission control in cellular networks. *IEEE Transactions on Vehicular Technology*, *55*(4), 1412–1424. doi:10.1109/TVT.2006.877470

Yin, L., Li, B., Zhang, Z., & Lin, Y. (2000). Performance analysis of a dual-threshold reservation (DTR) scheme for voice/data integrated mobile wireless networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC-2000)* (pp. 258-262).

Yoon, C. H., & Kwan, C. (1993). Performance of personal portable radio telephone systems with and without guard channels. *IEEE Journal on Selected Areas in Communications*, *11*, 911–917. doi:10.1109/49.232300

Yu, O., & Leung, V. (1996). Self-tuning prioritized call handling mechanism with dynamic guard channels for mobile cellular systems. In *Proceedings of IEEE Vehicular Technology Conference*, (pp. 1520-1524).

Yu, O. T. W., & Leung, V. C. M. (1997). Adaptive Resource Allocation for Prioritized Call Admission over an ATM-Based Wireless PCN. *IEEE Journal on Selected Areas in Communications*, *15*, 1208–1225. doi:10.1109/49.622906

Zheng, Z. H., & Lam, W. H. (2002). Performance analysis of dynamic channel assignment with queuing and guard channel scheme combined scheme for handoff prioritization. *Electronics Letters*, *38*, 1728–1729. doi:10.1049/el:20021177