When is it worth involving several cells in the Call Admission Control process for Multimedia Cellular Networks?

Youssef Iraqi and Raouf Boutaba
University of Waterloo, DCS, Waterloo, Canada
{iraqi, rboutaba}@bbcr.uwaterloo.ca

Abstract—**In this paper, we enhance our distributed call admission control scheme developed for cellular mobile networks. The new scheme can dynamically adapt to changes in the network load to maintain a target call dropping probability. We investigate the impact of the number of neighboring cells involved in a call admission decision in addition to the cell where the call originated. The neighboring cells provide significant information about their ability to support the new mobile user in the future. This distributed process allows the original cell to make a more clear-sighted admission decision for the new user. Simulations are presented with a detailed analysis of a comparison between two schemes involving different number of cells.**

Keywords—**Multimedia Cellular Networks, Distributed Call Admission Control, Multimedia Traffic.**

I. INTRODUCTION

Cellular mobile networks have to continue supporting their mobile users after they leave their original cells. This rises a new challenge to Call Admission Control (CAC) algorithms. A call admission process should not only take into consideration the available resources in the original cell but also in neighboring cells as well.

Mobile users are in a growing demand for multimedia applications, and the next generation wireless networks are designed to support such bandwidth greedy applications. The (wireless) bandwidth allocated to a user will not be fixed for the lifetime of the connection as in traditional cellular networks, rather it will be allocated dynamically to users. Many evolving standards for Wireless Broadband Systems, UMTS and IMT2000 have proposed solutions to support such capability [1] [2] [3].

Several call admission algorithms have been proposed for wireless networks to support multimedia users with dynamic bandwidth requirements (e.g. [4]). These algorithms take only local information in the admission decision process, which result in a high call dropping probability. Call Dropping Probability (CDP) is an important connection level QoS parameter in wireless mobile networks. To reduce the call dropping probability, few other CAC algorithms which take into consideration information from neighboring cells have been proposed [5][6][7]. However, those algorithms only support users with fixed bandwidth requirements.

In [8] we have proposed a Distributed CAC scheme designed for wireless mobile multimedia networks with dynamic bandwidth allocation. The call admission process involves the cell that receives the call admission request and a cluster of neighboring cells so the user will not be dropped due to handoffs. Consequently, the network will provide a low call dropping probability while maintaining a high resource utilization.

In this paper, we propose an enhancement of the CAC scheme and propose a mechanism for dynamic adaptation to obtain a target CDP. We investigate the impact of the number of involved cells in the CAC process on the achieved performance, in terms of average bandwidth utilization and call dropping probability.

The paper is organized as follows. In section 2, we describe the model of the system considered in this paper. Section 3 defines the dynamic mobile probabilities used by our distributed call admission control algorithm. In section 4 we present the call admission control performed locally by the cells in our system. Section 5 introduces the overall call admission control scheme involving a cluster of neighboring cells. Section 6 describes the algorithm used to dynamically achieve a target call dropping probability. Section 7 discusses the conducted simulation parameters and presents a detailed analysis of the obtained results. Finally, section 8 concludes the paper.

II. SYSTEM MODEL

We consider a wireless/mobile network with a cellular infrastructure that can support mobile terminals running applications which demand a wide range of resources. Users can freely roam the network and experience a large number of handoffs during a typical connection. We assume that users have a dynamic bandwidth requirement. The wireless network must provide the requested level of service even if the user moves to an adjacent cell. A handoff could fail due to insufficient bandwidth in the new cell, and in such case, the connection is dropped.

To reduce the call dropping probability, we have proposed in [8] to make neighboring cells participate in the admission decision of a new user. Each involved cell will give its local decision and then the cell where the request was issued will finally decide if the new request is accepted or not. By doing so, the new admitted connection will have more chances to survive after experiencing handoffs.

We use the notion of a cluster similar to the shadow cluster concept [9]. The idea is that every connection exerts an influence upon neighboring base stations. As the mobile terminal travels to other cells, the region of influence also moves. The cells influenced by a connection are said to constitute a cluster (see figure 1). Each user in the network, with an active connection has a cluster associated to it. The cells in the cluster are chosen by the cell where the user resides. The shape and the number of cells of a user's cluster depend on factors such as user's QoS, terminal trajectory and velocity.

0-7803-7097-1/01/$10.00 ©2001 IEEE

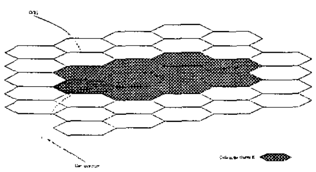

Fig. 1. Cell j and the cluster for a user

III. DYNAMIC MOBILE PROBABILITIES

We consider a wireless network where the time is divided in equal intervals at $t = t_0, t_1, ..., t_m$. Let j denote a base station in the network, and x a mobile terminal with an active wireless connection. Let $K(x)$ denote the set of cells that form the cluster for the active mobile terminal x. We write $P_{x,j,k}(t) = [P_{x,j,k}(t_0), P_{x,j,k}(t_1), ..., P_{x,j,k}(t_{m_x})]$ the probability that mobile terminal x, currently in cell j, to be active in cell k, and therefore under the control of base station k, at times $t_0, t_1, t_2, ..., t_{m_x}$. $P_{x,j,k}(t)$ represents the projected probabilities that mobile terminal x will remain active in the future and at a particular location. It is referred to as the Dynamic Mobile Probability (DMP) in the following. The parameter m_x represents how far in the future the predicted probabilities are computed. It is not fixed for all users and can depend of the user QoS or the actual connection elapsed time.

Those probabilities may be function of several parameters such as: handoff probability, cell size, user mobility profile, etc. The more information we have, the more accurate are the probabilities, however the more complex is their computation. Several schemes to compute these probabilities can be found in [5] [6] [7]. For each user x in the network, the cell responsible for this user decides the size of the cluster $K(x)$. The cells in $K(x)$ are those involved in the CAC process. The cell, then, sends the DMPs to all members in $K(x)$.

IV. LOCAL CALL ADMISSION CONTROL

At each time t_0 each cell, in a cluster $K(x)$ involved in our CAC process for user x, makes a local CAC decision for different times in the future $(t_0, t_1, ..., t_{m_x})$. Based on these CAC decisions, we call Elementary Responses, the cell makes a final decision which represents its local response to the admission of user x in the network. Elementary responses are time dependent. The computation of these responses is different depending on the user location and type. The user can be either a local new user or a new user that has a non null probability to be in this cell in the near future.

The network tries first to continue supporting old users and uses the DMPs to check if a cell can accommodate a new user who will possibly come to the cell in the future. The cell can apply any local call admission algorithm to compute the elementary responses. We write $r_k(x,t)$ the elementary response of cell k for user x for time t. We assume that $r_k(x,t)$ can take one of two values: -1 meaning that cell k can not accommodate user x at time t; and 1 otherwise. A detail description of how to compute the elementary responses is presented in [8].

Since the elementary responses for future foreign users are computed according to local information about the future, they should not be assigned the same confidence degree. Indeed, responses corresponding to the near future are more likely to be more accurate than those of the far future.

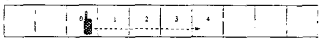

Fig. 2. An example of a highway covered by 10 cells

We write $C_k(x,t)$ the confidence of cell k in its elementary response $r_k(x,t)$. Cell k has to compute (or simply choose) the confidence degree $C_k(x,t)$, typically between 0% and 100%.

If for user x, cell k has an elementary response $r_k(x,t)$ for each t from t_0 to t_{m_x}, those elementary responses are weighted with the corresponding DMPs $P_{x,j,k}(t_0)$ to $P_{x,j,k}(t_{m_x})$, to compute the final response. The final response from cell k to cell j concerning user x is then :

$$R_k(x) = \frac{\sum_{t=t_0}^{t=t_{m_x}} r_k(x,t) \times P_{x,j,k}(t) \times C_k(x,t)}{\sum_{t=t_0}^{t=t_{m_x}} P_{x,j,k}(t)} \tag{1}$$

To normalize the final response, each elementary response is also divided by the sum over time t of the DMPs in cell k. Finally, cell k sends the response $R_k(x)$ to the corresponding cell j. Note that $R_k(x) \in \{-1, 1\}$.

V. DISTRIBUTED CALL ADMISSION CONTROL

Here the decision takes into consideration the responses from all the cells in the user's cluster. The admission process concerns only new users seeking admission to the network, not already accepted ones. We assume that cell j has already decided the cluster $K(x)$ and that cell j has already assigned to each cell k in the cluster $K(x)$ a weight $W_k(x)$. Each weight represents the importance of the contribution of the associated cell to the global decision process. We propose the following formula to compute the weights $W_k(x)$:

$$W_k(x) = \frac{\sum_{t=t_0}^{t=t_{m_x}} P_{x,j,k}(t)}{\sum_{k' \in K(x)} \sum_{t=t_0}^{t=t_{m_x}} P_{x,j,k'}(t)} \tag{2}$$

A. Relevance

In this paper, we introduce a new parameter that we call spatial relevance or simply relevance of a cell. To explain the idea of relevance, let's take the following example: consider a linear highway covered by 10 square cells as in figure 2. Assume that a new user, following the trajectory shown in figure 2, is requesting admission in cell number 0 and that the CAC process involves 5 cells. Responses from cells number 1, 2, 3 and 4 are relevant only if cell number 0 can accommodate the user. Similarly, responses from cells 2, 3 and 4 are relevant only if cell 1 can accommodate the new user when it hands off from cell 0. And the same principle applies to the other cells. This is because a response from a cell is irrelevant if the user can not be supported until that cell. We write $\Phi_k(x)$ the relevance of cell k for user x.

The relevance depends only on the topology of the considered cellular network. For the linear highway example of figure 2,

0-7803-7097-1/01/$10.00 ©2001 IEEE

we propose the following relevance formula:

$$\Phi_0(x) = 1 \; and \; \Phi_k(x) = \prod_{l=1}^{k} \frac{(1 + R_{l-1}(x))}{2} \qquad (3)$$

Note that for each $k \in K(x)$ we have $0 \leq \Phi_k(x) \leq 1$. Note also that in eq. 3, cell j (the cell receiving the admission request) has the index 0 and that the other cells are indexed in an increasing order according to the user direction as in figure 2.

The cell computes the sum of $R_k(x) \times W_k(x) \times \Phi_k(x)$ over k. The final decision of the call admission process for user x is based on:

$$D(x) = \frac{\sum_{k \in K(x)} R_k(x) \times W_k(x) \times \Phi_k(x)}{\sum_{k' \in K(x)} W_{k'}(x) \times \Phi_{k'}(x)} \qquad (4)$$

Note that $-1 \leq D(x) \leq 1$ and that $\sum_{k' \in K(x)} W_{k'}(x) \times \Phi_{k'}(x)$ is never null, since the relevance, $\Phi_0(x)$, of cell j is always equal to 1, its weight $W_j(x)$ is strictly superior to 0, and all other $\Phi_{k'}(x)$ and $W_{k'}(x)$ are positive or null.

If $D(x)$ is higher than a certain threshold, we call acceptance threshold T_{acc}, the user x is accepted; the user is rejected otherwise. The more higher is $D(x)$ the more likely the user connection will survive in the event of a handoff.

A detailed description of the algorithm is presented in [8].

VI. MAINTAINING A TARGET CALL DROPPING PROBABILITY

In this section we explain how our algorithm vary the value of T_{acc} to maintain a target CDP value. We assume that each Mobile Switch Center (MSC), controlling a set of cells in the network, modifies the acceptance threshold of the cells it controls in order to maintain a target CDP. The following is the pseudocode of the algorithm for adjusting the acceptance threshold, we will refer to as algorithm 1 in the remaining of the paper.

Algorithm 1 (Adjusting the acceptance threshold T_{acc})

```
$w = \lceil \frac{1}{target\ CDP} \rceil$; $wobs = w$; $nA = 0$; $nD = 0$
  if a user is accepted
    { $nA++$;
    if ($nA \geq wobs$) {
      if ($nD == wobs/w$) {$wobs = w$; $nA = 0$; $nD = 0$;}
        else {$wobs += w$; if ($T_{acc} > -1.0$) $T_{acc} -= 0.01$;}}}
  if a user is dropped $nD++$; if ($nD > wobs/w$)
    {$wobs += w$; if ($T_{acc} < 0.95$) $T_{acc} += 0.01$;}
```

The MSC begins by selecting a reference observation window w according to the target CDP as follows: $w = \lceil \frac{1}{target\ CDP} \rceil$. Note that we do not include the case where the target CDP is equal to zero, since this one is almost impossible to achieve and not practical from the provider point of view.

The variable representing the observation window $wobs$ is set to w, and the number of accepted users nA as well as the number of dropped users nD are set to zero.

If a new user is accepted in the system then nA is incremented by one. If we have observed at least $wobs$ accepted users ($nA \geq wobs$) then, if the number of users dropped is equal to the maximum allowed dropping value, we set $wobs$ to w and set nA and nD to zero and restart from the beginning. If the number of users dropped is less than the maximum allowed, then we increase $wobs$ and decrease the acceptance threshold.

In case a user is dropped then nD is incremented by one. If the number of dropped users exceeds the maximum allowed value, then we increase $wobs$ and increase the acceptance threshold. This means that we increase our observation window and will allow less users to be admitted in the system.

Note that the proposed algorithm aims to achieve exactly the target CDP. This can easily be modified to let the actual CDP lay between a maximum and a minimum allowed values.

VII. PERFORMANCE EVALUATION

We evaluate here our Distributed CAC scheme with different numbers of cells involved in the CAC process. We compare the performance of the scheme in the two following scenarios:

1. two cells are involved in the CAC process (called **SC1**).
2. five cells are involved in the CAC process (called **SC2**).

A. Simulation parameters

For simplicity, we evaluate the performance of our Distributed CAC for mobile terminals which are traveling along a highway as in figure 2. This is a simplest environment representing a one-dimensional cellular system. In our simulation study we have the following simulation parameters and assumptions:

1. The time is quantized in intervals $\tau = 10s$
2. The whole cellular system is composed of 10 linearly-arranged cells (numbered from 1 to 10), laid at 1-km intervals.
3. Cell 1 and 10 are connected so that the whole cellular system forms a ring architecture as assumed in [10]. This will avoid the uneven traffic load that will be experienced otherwise.
4. Each cell has a fixed capacity of 100 bandwidth units (BUs) except cells 3, 4 and 5 which have 50, 30 and 50 bandwidth units respectively. This is to create a local congestion that will remain for a long period. An example of such case is a temporary increase in the interference level.
5. Connection requests are generated in each cell according to Poisson process. A newly generated mobile terminal can appear anywhere in the cell with equal probability.
6. Mobile terminals can have speeds of: 70, 90, or 105 km/h. The probability of each speed is 1/3, and mobile terminals can travel in either of two directions with equal probability.
7. We consider three possible types of traffic: voice, data, and video. The probabilities associated with these types are 0.3, 0.4 and 0.3 respectively. The number BUs required by each connection type is: voice = 1, data = 5, video = 10.
8. Connection lifetimes are exponentially-distributed with a mean value equal to 180 seconds.
9. We simulate a total of 10 hours of real-time highway traffic, with a constant cell load equal to 720 new calls/h/cell.
10. The DMPs are computed as in [5].
11. The relevance is computed using eq. 3.
12. All users with a specific type of service have the same acceptance threshold. Algorithm 1 is used to adjust T_{acc} of all 10 cells and the target CDP is 10%. We assume that all 10 cells are under the control of one MSC. The accepted thresholds for voice, data and video users are set to $1.7 * T_{acc}$, $1.2 * T_{acc}$ and T_{acc} respectively. This is to achieve fairness between voice, data and video users.

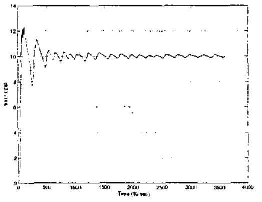

Fig. 3. Call dropping percentage

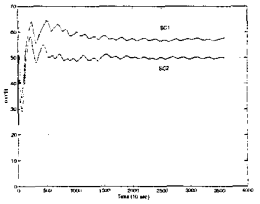

Fig. 4. Percentage of refused calls

The following additional simulation parameters are used for the **SC1** scheme:

- $m_x = 18$ for all users. This means that the DMPs are computed for 18 steps in the future.
- The size of the cluster $K(x) = 2$ for all users. This means that one cell in the direction of the user and the cell where the user resides form the cluster.

For **SC2** scheme, the following additional simulation parameters are assumed:

- $m_x = 25$ for all users.
- The size of the cluster $K(x) = 5$ for all users. This means that four cells in the direction of the user and the cell where the user resides form the cluster.

B. Simulation results

In our simulations, a user x requesting a new connection is accepted into a cell only if the final decision $D(x)$ is above the acceptance threshold corresponding to the user class of service (voice, data or video). Figure 3 depicts the call dropping percentage achieved when using scheme **SC2**. The call dropping percentage represents the ratio of dropped users to the number of admitted users in the system. This is the aggregate call dropping percentage including all types of service. We can notice that algorithm 1 allows the actual CDP to approach the target CDP by varying the value of the acceptance threshold T_{acc}.

In figure 4, we compare the percentage of refused calls, given the offered load, when using scheme **SC1** and **SC2**. We can notice that **SC2** refuses less users than **SC1**. Indeed, **SC2** accepts about 8% more users than **SC1**. At a first sight, this result may seem abnormal. Indeed, scheme **SC2** involves five cells in the CAC decision process (3 more cells than **SC1**), and thus it is more difficult for a new user to be admitted by **SC2** than **SC1**. However, as we will see later in this section, **SC2** has the ability to avoid admitting those users who are most likely to be dropped and can use the saved bandwidth to accept more users who can most likely be supported.

Figure 5 shows that **SC2** not only accepts more users than **SC1** but also allows for a better resource utilization. In fact, **SC2** uses almost 10 bandwidth units more than **SC1**. It is worth noting that the low resource utilization experienced by the two schemes is due to the number of video users in the system and to the assumption that the whole system is controlled by one MSC. The latter assumption means that when a part of the network experience a congestion, the whole network is affected by refusing more users (since the MSC increases the acceptance threshold for all the cells in the network). Although the simulated one MSC configuration is not likely to happen in the real-world, simulation results show the potential benefit of using scheme **SC2** compared to scheme **SC1**.

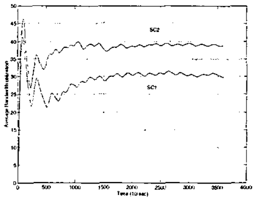

Fig. 5. Average bandwidth utilization

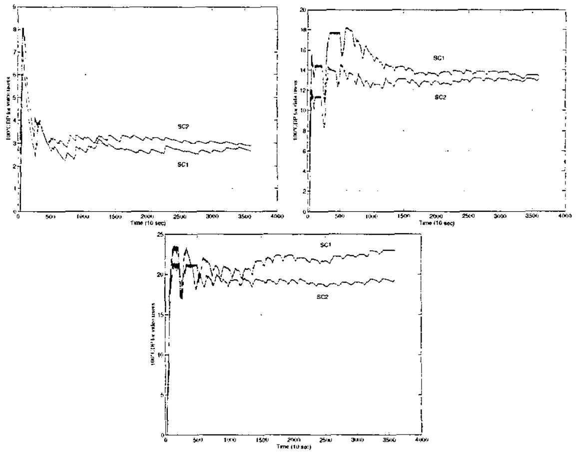

Fig. 6. Percentage of dropped voice, data and video users

To further compare the two considered schemes, we compute the individual dropping percentage among the three considered classes of service, namely voice, data and video. The simulation results are shown in figure 6. In this figure, we can observe that the two schemes, **SC1** and **SC2**, achieve almost the same dropping percentage for voice and data users respectively, with a slightly better performance of **SC2** in case of data users. However, **SC2** drops almost 4% video users less than **SC1**.

As the percentage of dropped users depicted in figure 6 is computed according to the number of accepted users in each class of service, the comparison will not be fair if we do not observe the number of admitted users within each class of service for the two schemes. Figure 7 shows the percentage of refused calls within each class of service, and figure 8 plotted the number of accepted users within each class of service when using the two schemes.

According to figure 7, **SC2** refuses less users than **SC1** irrespective of users classes of service. This means that **SC2** accepts

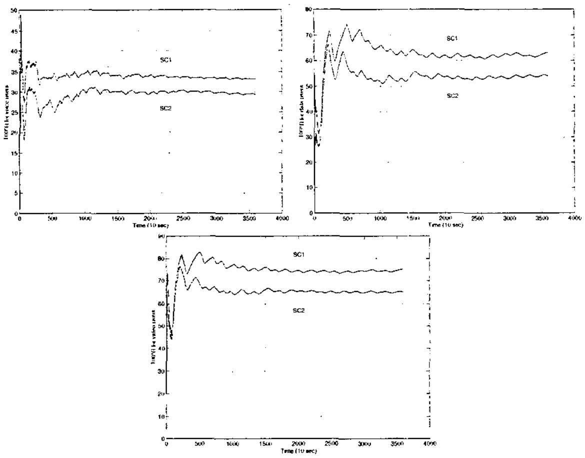

Fig. 7. Percentage of refused voice, data and video calls

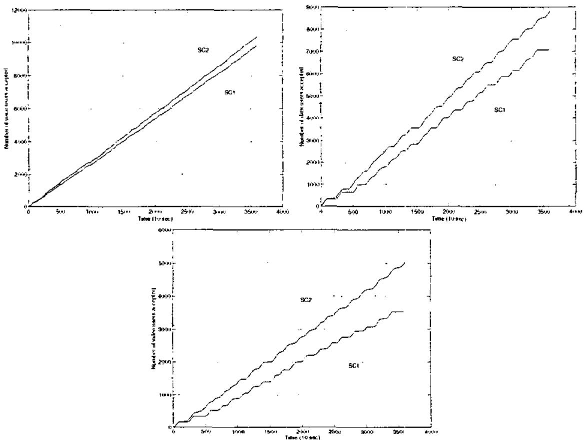

Fig. 8. Number of accepted voice, data and video calls

more users while achieving the same CDP in case of voice and data users, and that it allows more video users to be admitted to the network while achieving a lower CDP compared to **SC1**. According to figure 8, **SC2** accepts about 1500 video users more than **SC1** for the 10 real-time hours considered.

The bad performance achieved by **SC1** is explained by the fact that this scheme can not differentiate between those users who can be supported and those who can not. Its short sight prevents the scheme from being informed about a far congestion. Thus, the only way for **SC1** to reduce the CDP to the target value is to accept less users in the network, which results in a poor resource utilization.

On the other hand, since **SC2** involves more cells in the CAC process than **SC1**, the scheme is able to distinguish between those users who can be supported and those who are most likely to be dropped due to some congestion. This has the two following benefits: (1) the scheme can accept more users without sacrificing the CDP; (2) the bandwidth saved from not allowing some "bad" users to be admitted in the network, can be used to admit more "good" users.

We have conducted several other simulations with different offered loads and different simulation parameters (such as different mean holding time). The main observation worth highlighting here is that the two schemes **SC1** and **SC2** achieve almost the same performance in case of no congestion or in case of a uniformly distributed congestion. The latter case is less important since it can be solved off-line by increasing the network capacity. We have observed in the simulations presented in this paper, **SC2** achieves a better performance in case of a local congestion. The fact that the two schemes achieve the same results in case of a non congested network or in case of a uniformly distributed congestion is foreseeable. This is mainly because the responses from the three additional cells in **SC2** (cells 2, 3 and 4 in figure 2) only confirm what the two involved cells in **SC1** (cells 0, 1 in figure 2) have decided.

Of course, **SC2** does not have only advantages. As **SC2** involves more cells in the CAC decision process, it induces more communications between base stations and also requires more processing power than **SC1**. These resources are less critical compared to the wireless network bandwidth. A good compromise is to use **SC1** when the network is not congested and use **SC2** when a congestion is detected. The process of selecting the good scheme is out of the scope of this paper and is subject to future work.

VIII. CONCLUSION

We have described a call admission control scheme suitable for wireless multimedia networks. The proposed scheme operates in a distributed fashion by involving, in a call admission decision, not only the cell where the call originated, but also a determined number of neighboring cells. We also presented an analysis of the comparison between two call admission control schemes involving different number of cells in the decision process. We have observed that it is worth involving more cells in the CAC decision in case of local congestion. This allows the scheme to take a more clear-sighted admission decision for new users, hence, achieving better resource management and quality of service. The choice of the number of cells to involve and when this should happen is an important issue that will be addressed in the future.

REFERENCES

[1] UMTS, http://www.etsi.org/umts/
[2] BRAN, http://www.etsi.org/bran/
[3] IMT2000, http://www.itu.int/imt/
[4] J. Evans and D. Everitt, 'Effective bandwidth based admission control for multiservice CDMA cellular networks,' IEEE Trans. Vehicular Tech., vol. 48, no. 1, pp 36-46, Jan. 1999.
[5] D. A. Levine, I. F. Akyildz and M. Naghshineh, 'A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks Using the Shadow Cluster Concept,' IEEE/ACM Trans. Net., vol. 5, no. 1, 1997.
[6] Sunghyun Choi and Kang G. Shin, 'Predictive and Adaptive Bandwidth Reservation for Hand-Offs in QoS-Sensitive Cellular Networks,' in Proc. ACM SIGCOMM'98, pp. 155-166, Sept. 1998.
[7] Songwu Lu and Vaduvur Bharghavan, 'Adaptive Resource Management Algorithms for Indoor Mobile Computing Environments,' in Proc. ACM SIGCOMM'96, pp. 231-242, Aug. 1996.
[8] Y. Iraqi and R. Boutaba, 'A Novel Distributed Call Admission Control for Wireless Mobile Multimedia Networks,' in Proc. ACM WoWMoM'2000, pp. 21-27, 2000.
[9] D. A. Levine, I. F. Akyildz and M. Naghshineh, 'The shadow cluster concept for resource allocation and call admission in ATM-based wireless networks,' in Proc. ACM MOBICOM'95, pp. 142-150, Nov. 1995.
[10] M. Naghshineh and M. Schwartz, 'Distributed Call Admission Control in Mobile/Wireless Networks.' IEEE JSAC, vol. 14, no. 4, pp. 711-717, 1996.