An Admission Control Scheme for QoS-Sensitive Cellular Networks

Sunho Lim, Guohong Cao, and Chita R. Das
Department of Computer Science and Engineering
The Pennsylvania State University
University Park, PA 16802

Abstract—We propose an admission control scheme to guarantee a certain level of QoS to on-going connections in cellular networks. This admission control scheme is based on a differential bandwidth reservation policy that uses a sector of cells in making bandwidth reservation for accepting the new call. The sector of cells, which are located along the way to which the MT might move, are further divided into two regions depending on whether they have an immediate impact on the handoff or not. Two different bandwidth reservation policies are applied to cells in the two regions to optimize the connection dropping rate (CDR) while maximizing the connection blocking rate (CBR).

In contrast to most prior policies, the proposed admission control scheme uses varying number of cells in the sector to make admission decision. Depending on the currently measured average CDR of the cells in the sector and the current cell where a new connection is generated, the number of cells involved in admission control can be changed dynamically to satisfy the target QoS (CDR) parameter. Simulation results indicate that our admission control policy guarantees the required CDR over the entire workload, while maintaining a competitive CBR. Comparison of the proposed scheme with two prior schemes shows that our approach is not only capable of providing better QoS guarantees, but also is more flexible in terms of using varying number of cells in satisfying a certain QoS requirement.

I. INTRODUCTION

With the increasing use of wireless networks as a ubiquitous communication infrastructure, Quality-of-Service (QoS) guarantees to on-going and new connections are becoming critical issues in designing wireless systems. Since users are expected to move around during communication sessions, one of the most important QoS factors is related to connection handoff, which is a mechanism of transferring an on-going connection from the current cell to the next cell to which a mobile terminal (MT) moves. A handoff, however, could fail due to unavailability of sufficient bandwidth in the destination cell. As the individual cell size gradually shrinks to accommodate increasing number of MTs within the limited frequency spectrum, the probability of connection handoffs, and hence being dropped due to insufficient bandwidth during the lifetime of a connection could be high.

Connection dropping rate (CDR) can be reduced by reserving some bandwidth solely for handoffs. However, the connection blocking rate (CBR) of new connections may increase due to such bandwidth reservation. Hence, reduction of CBR and CDR are conflicting requirements, and optimization of both is admittedly extremely complex. Two supplementary techniques have been used for controlling the CDR and CBR in cellular networks: bandwidth reservation and admission control. While bandwidth reservation helps in minimizing the CDR of on-going connections, it cannot effectively guarantee a certain level of QoS without an admission control scheme when the traffic load is high. Therefore, admission control is an integral part of the call management mechanism, and has been a current research focus.

The difficulty in admission control in a mobile environment is that it is not adequate to admit a new call only based on the status of the current cell, where the call is generated. This is because when the MT attempts to move from the original cell to a next cell, there may not be sufficient bandwidth in the destination cell for accepting the handoff. This may result in dropping the call, and increasing the CDR. Therefore, an efficient admission control policy should check the bandwidth availability in the adjacent cells to make the handoff smooth.

Recent admission control policies [1] [2] [3], therefore, consider the status of a number of cells, which are located around or along the path to which the MT might move, and then make an admission decision based on the agreement among all cells or a subset of cells, respectively. For example, Naghshineh and Schwartz [3] suggested a distributed admission control scheme where the current cell and all its immediate adjacent cells are considered to decide whether a new connection should be accepted or not. The shadow cluster concept introduced in [2] estimates the future resource requirements and considers a set of cells located around an active MT. However, these schemes are based on strong assumptions such as precise knowledge about handoff and connection termination, as well as the MT’s mobility pattern. Choi and Shin [4] compared several connection level admission control schemes, which can predict the mobility pattern of the MTs based on their aggregated moving history. Aljadhai and Znati [1] defined a most likely cluster (MLC), which is a set of cells to which an MT is likely to move with a higher probability during its lifetime. The shape of the MLC and the number of cells in the MLC are determined based on the moving speed and direction of the MT. Since all cells in the MLC need to reserve bandwidth before admitting a new connection, the approach may waste a large amount of bandwidth.

Recently we have proposed a differential bandwidth reservation scheme [5] that uses a sector-type configuration to reserve/share bandwidth along the path of an MT. The size of the sector and the number of cells in the sector can be dynamically configured for each MT using its mobility information. The sector of cells are further divided into two regions depending on whether they have an immediate effects on the handoff or not. Two different reservation/sharing policies are employed in the two regions.
In this paper, we propose an admission control scheme using the differential bandwidth reservation policy. The admission control algorithm first uses the bandwidth reservation algorithm to check if bandwidth reservation can be done in appropriate cells to admit the new call. In contrast to most prior schemes, not all the cells in the sector are required to reserve or share the bandwidth for satisfying the QoS requirement. The number of cells involved in admission control can be changed dynamically depending on the average CDR of the cells in the sector and that of the current cell where a new connection is generated. The novelty of the proposed admission control mechanism is that it is adaptable to the mobility pattern of the MTs in terms of the sector size, number of cells in the sector, and specific QoS parameters.

We simulate a \((6 \times 6)\) wrap-around network of hexagonal cells to evaluate the effectiveness of the proposed admission control mechanism. We compare the proposed scheme with two related schemes, called STATIC [6] and Predictive Timed QoS guarantees (PT_QoS) that is based on the MLC concept [1]. Both voice and video traffics are used for performance evaluation. Simulation results indicate that our scheme is more adaptable to provide a certain level of CDR guarantee compared to the prior schemes. In particular, it guarantees the specified CDR over the entire workload while maintaining a competitive CBR. In addition, we study the impact of the cell size, the number of cells involved in making admission decisions and other system parameters in satisfying the overall QoS requirements.

The rest of this paper is organized as follows. Section II introduces the system model. The admission control scheme is presented in Section III. Section IV is devoted to performance evaluation and comparisons of the algorithms. Section V concludes the paper.

II. SYSTEM MODEL

In cellular networks, a geographical area is divided into a number of hexagonal regions, called cells. Each cell is supported by a base station (BS) located in the center of the cell. The BSs are connected to each other by a static wired network. Each mobile terminal (MT) in a cell communicates with the BS by wireless links. Each cell is assigned a fixed number of channels (or bandwidth). Since an MT tends to keep its original moving direction, it has a lower probability of making a sudden turn compared to maintaining the current direction. Therefore, let us assume that an MT moves straight, left, right, lower left, lower right, and back with probabilities of \(P_S\), \(P_L\), \(P_R\), \(P_{LL}\), \(P_{LR}\), and \(P_B\), respectively.

We assume that a BS is capable of predicting the direction of a new connection based on its path history, which is recorded at the MT. Typically, a BS periodically broadcasts control messages to the MTs in its area for location and bandwidth management. A BS’s \(id\) consisting of a few bits can be easily added to the control message without much overhead. Usually, a cellular phone/MT stays in the power-saving mode to reduce battery consumption when it is idle. However, it still listens to the control channel. Whenever the MT moves into a new cell, based on the control messages, it identifies the BS \(id\), and caches the \(id\) in its memory. When the MT makes a new call, it sends the list of \(id\)s with the connection request message to the BS. The BS can predict the next cell to which the MT might move from the path \((id)\) list. Similar assumptions have been used in prior research [1] [7]. More accurate moving directions can be obtained by using the Global Positioning System (GPS) [8].

Since an MT has a much higher probability of maintaining its current direction, moving left, or moving right than the other three directions, there exist a sector of cells covered within an angle \(\theta\) to which the MT might move in the near future. This is similar to the MLC concept [1]. This sector of cells can be classified based on their distance from the current cell. Let \(r_{i,d}\) denote the cluster of cells located at a distance \(d\) away and consisted of cells to which the MT might move with a high probability from the current cell \(c_i\). Let \(st(\theta)\) denote the sector of cells located within \(\theta\). Then, we have:

\[
r_{i,d} = \{c_j | \text{distance}(c_i,c_j) = d \land c_j \in st(\theta)\}, \quad d = 1, 2, \ldots, n.
\]

As shown in Figure 1, cluster \(r_{i,1}\) consists of three cells, \(c_1\), \(c_2\), and \(c_3\), while cluster \(r_{1,4}\) has 7 cells, \(c_{12}\) to \(c_{18}\). As the distance \(d\) increases, the number of cells in that cluster increases. Depending on the moving speed, mobility pattern, and the traffic flow of MTs, the number of cells in a cluster and the maximum cluster distance \(n\) should be carefully chosen. Also, the sector angle \(\theta\) is likely to affect the system performance.

![Fig. 1. A classification of sector of cells into different clusters based on their distance from the current cell.](image)

III. A QoS-AWARE ADMISSION CONTROL SCHEME

In this section, we propose a QoS-aware admission control scheme based on our differential bandwidth reservation policy [5]. A BS, as described in the system model, is assumed to have some knowledge of the moving pattern of the new call. Using this information, the BS checks if bandwidth reservation in the appropriate cells can be done to accommodate the new call. The proposed admission control mechanism attempts to keep the CDR below a target QoS value by judiciously using the bandwidth reservation algorithm. We first describe the
bandwidth reservation scheme followed by the admission control algorithm.

A. An Efficient Differential Bandwidth Reservation Policy

Bandwidth reservation is essential for seamless handoff of an MT from one cell to the next cell. When making bandwidth reservations, it is more effective to consider a cluster of cells which includes cells located around or along the way to which the MT might move. Furthermore, there should be different bandwidth reservation policies for different cells depending on their distance from the current cell. For example, suppose an MT in cell \( c_i \) is likely to perform a handoff. \( BS_i \) can predict the direction to which the MT might move based on the past history of the MT, and then construct a sector of cells for bandwidth reservation. It is easy to see that the MT will most likely move to the cells closer to \( c_i \), such as cells in \( r_{i,1} \) and \( r_{i,2} \) in Figure 1. Since the MT may change its direction at any time, it has a relatively lower probability of moving to cells located far away from cell \( c_i \), such as the cells in \( r_{i,4} \). Based on this observation, for each cell \( c_i \), cells that are required for bandwidth reservations are classified into two different regions: \( R_I(i) \) and \( R_{II}(i) \).

- \( R_I(i) \): It contains the cells located closer to \( c_i \). For example, cells in \( r_{i,1} \) and \( r_{i,2} \) are close to \( c_i \) and they have a direct impact on CDR because the handoff will fail if the requested bandwidth is not available.
- \( R_{II}(i) \): It contains the cells further away from \( c_i \). For example, cells in \( r_{i,3} \) and \( r_{i,4} \) are 3 and 4 cell distant away from \( c_i \), and they only have indirect effects on the handoff because the MT may not move to those cells.

These regions are referred to as inner and outer regions, respectively. For each handoff request, a BS reserves or shares bandwidth of cells in both regions. Since several BSs could request reservations for the corresponding handoff requests, a cell \( c_m \) may be involved in multiple reservations from different BSs. As an example, consider two handoff requests from two cells \( c_i \) and \( c_j \). \( BS_i \) and \( BS_j \) make bandwidth reservations in their corresponding regions of cells; i.e., \( R_I(i) \) and \( R_{II}(i) \), and \( R_I(j) \) and \( R_{II}(j) \). As a result, \( c_m \) can be included in \( R_I(i) \) or \( R_{II}(i) \), and in \( R_I(j) \) or \( R_{II}(j) \). In this case, \( c_m \) reserves bandwidth for handoff requests from \( c_i \) and \( c_j \). The differential bandwidth reservation policies for the two regions (inner and outer) are described next.

Prior to initiating a handoff, a \( BS_i \) first constructs the sector configuration and broadcasts the region information along with the reservation request to the corresponding cells. When \( BS_m \) receives a handoff request from \( BS_i \), it already knows its reservation region. If \( c_m \) is in \( R_I(i) \), \( BS_m \) makes bandwidth reservation using the condition: \( N_{used}(m) + bw \leq N_{total} - N_{req}(m) \), where \( N_{used}(m) \) is the currently used bandwidth, \( bw \) is the requested bandwidth, \( N_{total} \) is the total bandwidth of the cell, and \( N_{req}(m) \) is the currently reserved bandwidth for handoffs. If the above condition is not satisfied, \( BS_m \) checks if it can share the already reserved bandwidth with other cells.

For bandwidth sharing, the reserved bandwidth is not strictly assigned to any particular MT, and hence can be used by any MT in a first come first serve (FCFS) manner. To implement this idea, \( BS_m \) checks if the sum of \( N_{req}(m) \), \( bw \), and the existing shared bandwidth \( (N_{share,I}(m) + N_{share,II}(m) \times \frac{\epsilon}{\eta}) \) is less than or equal to \( \epsilon \) times of \( N_{req}(m) \). Note that the shared bandwidth could be different in the inner and outer regions, and they are specified by two different terms. We use a weight factor \( \frac{\epsilon}{\eta} \), which is less than 1.0 for \( N_{share,I}(m) \), because the MT has a low probability to move to this region. Here, \( \epsilon \) is a system parameter, which enables \( BS_m \) to reserve more than the actually allocated bandwidth, i.e., \( \epsilon \geq 1 \) based on the QoS requirement of the system. If any of the two conditions (reservation/sharing) is satisfied, \( BS_m \) sends a positive reply (ACK) to \( BS_i \). Otherwise, the reservation request from \( BS_i \) is rejected.

If \( c_m \) belongs to \( R_{II}(i) \), reservation may waste a large amount of bandwidth since the MT has a relatively low probability of moving into \( c_m \). Thus, only bandwidth sharing is applied in the outer region. Specifically, \( BS_m \) checks if the sum of \( N_{req}(m) \), \( bw \), and the existing shared bandwidth, \( (N_{share,I}(m) + N_{share,II}(m) \times \frac{\epsilon}{\eta}) \), is less than or equal to \( \eta \) times of \( N_{req}(m) \), and then decide whether it can share the bandwidth or not. In contrast to \( R_I(i) \), we use a weight factor \( \frac{\epsilon}{\eta} \), which is greater than 1.0 for \( N_{share,I}(m) \) due to its high probability of being used. Similar to \( \epsilon \), \( \eta \) is a system QoS parameter, and \( \eta \geq 1 \). Intuitively, \( \eta \) should be higher than \( \epsilon \) to facilitate more sharing in the outer region. The pseudo code for the bandwidth reservation is given in Fig. 2. For additional details, please refer to [5].

![Fig. 2. The differential bandwidth reservation scheme used in \( c_m \) when it receives a reservation request from \( c_i \).](image-url)

```plaintext
if \( c_m \in R_I(i) \) \{ 
if \( (N_{req}(m) + bw \leq N_{total} - N_{req}(m)) \) 
send agree to allocate bandwidth to \( c_i \); 
\( N_{req}(m) = N_{req}(m) + bw; \) 
* done only after getting confirmation from \( BS_i \) */!
\}
else 
if \( (N_{req}(m) + bw + N_{share,I}(m) + N_{share,II}(m) \times \frac{\epsilon}{\eta} \leq N_{req}(m) \times \eta) \) 
send agree to share bandwidth to \( c_i \); 
\( N_{share,I}(m) = N_{share,I}(m) + bw; \) 
* done only after getting confirmation from \( BS_i \) */!
\}
else send reject to \( c_i \); 
\}
else if \( c_m \in R_{II}(i) \)
if \( (N_{req}(m) + bw + N_{share,I}(m) + N_{share,II}(m) \times \frac{\epsilon}{\eta} \leq N_{req}(m) \times \eta) \) 
send agree to share bandwidth to \( c_i \); 
\( N_{share,I}(m) = N_{share,I}(m) + bw; \) 
* done only after getting confirmation from \( BS_i \) */!
\}
else send reject to \( c_i \); 
\}
```
B. The Admission Control Algorithm

Next we present an admission control policy for new connections. Whenever a new connection request is generated in a current cell $c_i$, $BS_i$ first checks if Eq. 1 can be satisfied. If the condition cannot be satisfied, the new connection is blocked. Otherwise, $BS_i$ constructs $R_f(i)$ and $R_{II}(i)$ based on the current MT’s moving direction, which can be obtained by using the techniques mentioned in Section II. For cells which always follow similar moving pattern, a look-up table about the moving pattern can be saved and used subsequently to reduce the computation overhead [9]. With the information about $R_f(i)$ and $R_{II}(i)$, $BS_i$ sends reservation request messages to the corresponding cells to check the feasibility of the handoff.

$$N_{used}(i) + bw \leq N_{total} - N_{resv}(i).$$

When a cell, say $c_j$, in $R_f(i)$ or $R_{II}(i)$ receives a reservation request message from $c_i$, $BS_j$ executes the algorithm shown in Figure 2, computes its current CDR ($\rho_j$), and replies to $c_i$. After $BS_j$ receives all reply messages from cells in $R_f(i)$ and $R_{II}(i)$, $c_{i,all}$, it computes $\rho_i$, checks average CDR ($\rho_{i,avg}$), and decides whether it can accept the new connection or not. Here, $\psi$ is a system parameter, which can vary dynamically.

Let $|c_{i,agree}|$ and $|c_{i,all}|$ be the number of cells which sent ACK messages, and the total number of cells in $R_f(i)$ and $R_{II}(i)$, respectively. Then $c_{i,all} = \{c_j | c_j \in R_f(i) \lor c_j \in R_{II}(i)\}$, and $c_{i,agree} = \{c_j | c_j \in c_{i,all} \land c_j \text{ has sent ACK to } c_i\}$. The average CDR is given by $\rho_{i,avg} = \frac{1}{|c_{i,all}| + 1} \left( \sum_{c_k \in c_{i,all}} \rho_k + \rho_i \right)$. After receiving all the responses, the new connection is accepted if Eq. 2 is satisfied.

$$|c_{i,all}| \cdot \psi_i \leq |c_{i,agree}| \land \rho_{i,avg} \leq T_{QoS}.$$  

Once $BS_i$ accepts the new connection, it sends reservation confirm messages to the cells, which sent an ACK message for bandwidth reservation or sharing. Otherwise, the new connection is blocked. When $BS_j$ receives a confirm message from $c_i$, it reserves or shares the bandwidth as specified in Figure 2.

After $BS_i$ makes a decision for a new connection, it changes $\psi_i$ depending on the $\rho_{i,avg}$ value. If the $\rho_{i,avg}$ is higher than $T_{QoS}$, then it implies that $c_{i,all}$ currently does not have enough reserved or shared bandwidth for the future handoff to keep the $\rho_{i,avg}$ below $T_{QoS}$. Therefore, accepting a new connection in $c_i$ can degrade $\rho_{i,avg}$. $BS_i$ can block a new connection by increasing $\psi_i$ by an amount of $\delta$. This implies that $BS_i$ has more cells in $c_{i,all}$ to reserve or share the bandwidth, which in turn lowers the probability of accepting a new connection. However, if the $\rho_{i,avg}$ less than $T_{QoS}$, then the system bandwidth in $c_{i,all}$ is underutilized. In this case, $BS_i$ should accept more new connections by decreasing $\psi_i$ by $\delta$. This helps $BS_i$ in accepting a new connection even though less number of cells in $c_{i,all}$ reserve or share the bandwidth. The complete admission control algorithm for a new connection is shown in Fig. 3.

IV. PERFORMANCE EVALUATION

A. The Simulation Testbed

We use a $(6 \times 6)$ wrap-around cellular network with 1 mile cell distance to examine the proposed schemes. We consider two different types of communication sessions: voice and video. The connection time (T) is exponentially distributed with a mean of 180 seconds, and the connection arrival follows Poisson distribution with a rate of $\lambda$. The speed of the MTs is uniformly distributed between 45 to 70 miles per hour. The total bandwidth allocated to each cell is 600 kbps. The voice connection and the video connection require 10 kbps and 100 kbps, respectively. The percentage distribution of voice and video connections is given by $P_{voice} = 0.8$ and $P_{video} = 0.2$. Similar assumptions have been used in most prior studies [1] [4] [10].

We have written an event-driven simulator using CSIM to conduct the performance study. The simulation results are illustrated as a function of the offered load, where offered load is defined as

$$\text{Offered Load} = T \cdot \lambda \cdot (P_{voice} \cdot B_{voice} + (1 - P_{voice}) \cdot B_{video}).$$

Note that for a given offered load, we can find the arrival rate $\lambda$ since all other parameters in Eq. 3 are known. For most of the experiments, we use two clusters and the number of cells in the inner and outer cluster is 3 and 3, respectively. Since the probabilities of moving straight ($P_5$) and moving left or right ($P_L$ or $P_R$) are much higher than the other probability, we choose $P_5 = 0.5$, $P_L = 0.15$, $P_R = 0.15$, $P_{LL} = 0.075$, $P_{LR} = 0.075$, $P_{RL} = 0.075$, $P_{RR} = 0.075$. The Sharing rates ($\epsilon$ or $\eta$) are chosen between 1.0 to 3.0. The $\psi$ in the admission control algorithm is incremented or decremented by $\delta=0.1$ in our experiment.

B. Simulation Results

We analyze and compare the performance of our approach with two previously proposed schemes: the STATIC approach [6] and the PT-QoS approach [1]. In the static reservation (STATIC), a fraction ($g$) of the total bandwidth in a cell is exclusively reserved for handoffs. This scheme relies only on the local information to admit a new connection, and the current
cell does not negotiate with neighboring cells. The PT_{QoS} approach defines an MLC (Mostly Likely Cluster) based on the mobility pattern of the MTs, and reserves bandwidths from a fraction of cell \((s)\) in the MLC based on parameters such as the MT’s expected arrival time, the latest arrival times, and the departure time. Unlike our scheme, the PT_{QoS} algorithm uses a fixed number of cells to admit a new connection, and drops the on-going connection if the MT does not conform the prediction. Both these schemes have no mechanism to support QoS guarantee for handoffs. For the differential bandwidth reservation algorithm, we use one cluster in \(R_{I}^{f}\) and one cluster in \(R_{II}\) for the sector configuration, and two sharing rates \(\epsilon\) and \(\eta\), which are set to 1.5 and 3.0, respectively. As shown in Fig. 4 (a), the STATIC approach has the lowest CBR because it only examines the status of the current cell where the new connection request is initiated. The PT_{QoS} scheme has the maximum CBR since it reserves more than the required bandwidth for a longer time. Our scheme has a lower CBR than the PT_{QoS} approach, but slightly higher CBR than the STATIC approach.

Figure 4 (b) compares the CDRs of three schemes. As can be seen, the proposed policy keeps the CDR below the \(T_{QoS}\) compared to the other two schemes. An acceptable CDR of 0.01 or 0.02 has been used as a target QoS in prior studies [4] [10] [11]. The PT_{QoS} approach has a lower CDR than the STATIC scheme except for \(g = 0.2\), which shows a little lower CDR than the PT_{QoS} approach (with \(s = 0.7\)) when the offered load is high. The PT_{QoS} approach with \(s = 0.7\) has a higher CDR than that with \(s = 1.0\), because the former accepts more connections.

Based on the results of Figures 4, we can conclude that proposed scheme not only provides relatively better performance, but also maintains a stable CDR as specified by \(T_{QoS}\) over the entire workload. It has comparable CBR with respect to the STATIC policy. However, the CDRs with both STATIC and PT_{QoS} policies increase with the workload and are higher than the target value of 0.01.

V. CONCLUSIONS

In this paper, we proposed an admission control scheme for new connections, while guaranteeing a certain level of QoS to on-going connections in cellular networks. Unlike the prior schemes, the proposed admission control algorithm uses a subset of cells in the sector to check if the required CDR can be maintained. The number of cells required for bandwidth reservation varies dynamically based on the average CDR of the cells in the sector. Simulation results verified that the proposed schemes outperform the prior STATIC and PT_{QoS} policies.

REFERENCES