An Adaptive Distributed Call Admission Control for QoS-Sensitive Wireless Mobile Networks

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

Abstract—**This paper introduces an adaptive distributed call admission control framework developed for cellular mobile networks. The main feature of the proposed framework is a more efficient support for mobile multimedia users having dynamic bandwidth requirements. This is achieved by imposing an upper bound on the experienced call dropping probability, regardless of network load changes, while maintaining a high network resource utilization. The call admission control algorithm presented in this paper involves not only the original cell (handling the new admission request) but also a cluster of neighboring cells. 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. The cell changes the acceptance threshold dynamically to maintain a target call dropping probability. Simulations are provided to show the improvements obtained using our framework.**

Keywords—**Wireless mobile networks, multimedia traffic, adaptive call admission control.**

I. INTRODUCTION

Cellular mobile networks have to continue supporting their mobile users after they leave their original cells. This poses 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 wireless networks, rather the base station will allocate bandwidth dynamically to users. The Wireless ATM and the UMTS standards have proposed solutions to support such capability.

Several call admission algorithms have been proposed for wireless networks (e.g. [1]) to support multimedia users with dynamic bandwidth requirements. These algorithms take only local information in the admission decision process, and therefore will have 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 neighboring cells information have been proposed [2][3][4]. However, those algorithms only support users with fixed bandwidth requirements. In [5], we have proposed an admission control algorithm for supporting multimedia users with dynamic bandwidth requirements. The call admission process involves not only the cell that receives the call admission request but also a cluster of neighboring cells so the user will not be dropped due to handoffs. However, the proposed scheme can not dynamically adapt to changes in the network load.

In this paper we propose an adaptive CAC scheme for support of multimedia users, which allow a high resource utilization while guarantying an upper bound on the experienced call dropping probability. This is achieved by dynamically changing the acceptance threshold and network load monitoring.

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 propose to make neighboring cells participate in the decision of a new user admission. Each involved cell will give its local decision and finally the cell where the request was issued will 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 [6]. 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 set of cells influenced by a connection are said to constitute a cluster. 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 current call holding time, user's QoS, terminal trajectory and velocity.

III. DYNAMIC MOBILE PROBABILITIES

We consider a wireless network where the time is divided in equal intervals at $t = t_1, t_2, ..., t_m$. 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 denote $P_{x,j,k}(t) = [P_{x,j,k}(t_0), P_{x,j,k}(t_1), ..., P_{x,j,k}(t_{m_x})]$ 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

0-7803-6596-8/00/$10.00

times $t_0, t_1, t_2, ..., t_{m_x}$. $P_{x,j,k}(t)$ represents the projected probabilities that a mobile terminal 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 (e.g. higher QoS users will have a higher value of m) or the actual connection elapsed time (e.g. new connection will have a high value of m).

Those probabilities may be function of several parameters such as:

- residence time of mobile x in cell j,
- handoff probability,
- the distribution of call length for a mobile terminal x when using a given service class,
- cell size,
- user mobility profile.

Of course, the more information we have, the more accurate are the probabilities, and hence the more complex is their computation.

For each user x in the network, the cell that is responsible for, decides the size of the cluster $K(x)$, those are the cells involved in the CAC process, and sends the DMPs to all members in $K(x)$. The cell must specify if the user is a new one (in which case the cell is waiting for responses from the members of $K(x)$) or not.

DMPs could range from simple probabilities to complex ones. Simple probabilities can be obtained by assuming, for example, that call length is exponentially distributed, that the call arrival process follows a Poisson distribution, that the handoff probabilities are equal in any direction and so on (Manhattan cellular network, small cluster).

DMPs can also be complex for example by including information about user mobility profiles. A method for computing dynamic mobile probabilities taking into consideration mobile terminal direction, velocity and statistical mobility data, is presented in [2]. Other schemes to compute these probabilities are presented in [3] [4]. To compute these probabilities, one can also use mobiles' path/direction information readily available from certain applications, such as the route guidance system of the Intelligent Transportation Systems with the Global Positioning System (GPS).

IV. LOCAL CALL ADMISSION CONTROL

User's traffic can be either voice, data or video. Voice users are usually characterized by a fixed bandwidth demand. Data and video users have a dynamic bandwidth requirement due to the burstiness of the carried traffic. Without loss of generality, we assume that all users are characterized by a bandwidth demand distribution $f_x(E_x(c), \sigma_c)$. Where $E_x(c)$ and σ_c are the mean and the standard deviation of the distribution f_x respectively, and c is user's x type of traffic. $E_x(c)$ depends of user x traffic type c (audio, data or video).

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. The cell can apply any local call admission algorithm to compute the elementary responses.

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

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

where $C_k(x,t)$ is the confidence that has cell k about the elementary response $r_k(x,t)$. To normalize the final response each elementary response is also divided by the sum over time t of the DMPs in cell k. Of course, the sum $\sum_{t=t_0}^{t=t_{m_x}} P_{x,j,k}(t)$ should not be null (which otherwise means that all the DMPs for cell k are null!). Cell k, then, sends the response $R_k(x)$ to the corresponding cell j.

V. DISTRIBUTED CALL ADMISSION CONTROL

Here the decision takes into consideration the responses from all the cells in the user cluster. The admission process concerns only new users seeking admission to the network and not already accepted users.

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. Usually a cell that is involved more in supporting the user has a high weight value. Weights $W_k(x)$ depend on the DMPs and the time t.

The final decision of the call admission process for user x is based on

$$D(x) = \sum_{k \in K} R_k(x) \times W_k(x) \qquad (2)$$

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

According to the load experienced, the network should change its acceptance threshold to guaranty that the call dropping probability is always bellow a fixed target value. This allows the scheme to adapt to changes in the traffic intensity or distribution.

VI. MAINTAINING A TARGET CALL DROPPING PROBABILITY

In this section we explain how our algorithm vary the value of the acceptance threshold to maintain a target CDP value. We assume that each Mobile Switch Center, 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 pseudo-code of the algorithm for adjusting the acceptance threshold, we will refer to as algorithm 1 in the remaining of the paper.

Algorithm 1: Pseudo-code for adjusting the acceptance threshold:

```
$refW = \lceil \frac{1}{target\ CDP} \rceil$; $obsW = refW$; $nA = 0$; $nD = 0$
do while the system is running {
  if a user is accepted
    { $nA++$;
    if ($nA \geq obsW$)
      {
        if ($nD == obsW/refW$)
          { $obsW = refW$; $nA = 0$; $nD = 0$; }
        else
          { $obsW += refW$;
          if ($threshold > -1.0$) $threshold -= 0.01$; }
      }
    }
  if a user is dropped
    $nD++$;
    if ($nD > obsW/refW$)
      { $obsW += refW$;
      if ($threshold < 0.95$) $threshold += 0.01$; }
}
```

The MSC begins by selecting a reference observation window $refW$ according to the target CDP as follows: $refW = \lceil \frac{1}{target\ CDP} \rceil$. 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 $obsW$ is set to $refW$, 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 $obsW$ accepted users ($nA \geq obsW$) then, if the number of users dropped is equal to the maximum allowed dropping value, we set $obsW$ to $refW$ 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 $obsW$ and decrease the acceptance threshold. This means that we will allow more users to be admitted in the system.

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 $obsW$ 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. Note also that the maximum allowed acceptance threshold is set to 0.95 in algorithm 1 in order for the network to accept a minimum number of users even if a congestion occurs.

VII. PERFORMANCE EVALUATION

A. Simulation parameters

For simplicity, we evaluate the performance of our Adaptive Call Admission Control for mobile terminals which are traveling along a highway. 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, laid at 1-km intervals. Cells are numbered from 1 to 10.
3. Cell 1 and 10 are connected so that the whole cellular system forms a ring architecture. This allows to avoid the uneven traffic load that will be experienced by cell 1 and 10 otherwise.
4. Each cell has a fixed capacity of 100 bandwidth units.
5. Connection requests are generated in each cell according to a Poisson process with rate λ (connections/second). 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 $R_{voice} = 0.3$, $R_{data} = 0.4$ and $R_{video} = 0.3$ respectively. The number of bandwidth units (BUs) required by each connection type is: $B_{voice} = 1$, $B_{data} = 5$, $B_{video} = 10$. Note that fixed bandwidth amounts are allocated to users for the sake of simplicity.
8. Connection lifetimes are exponentially-distributed with a mean value equal to 180 seconds.
9. m_x is fixed for all users and for the duration of the connection and is equal to 25. This means that the DMPs are computed for 25 steps in the future.
10. The size of the cluster $K(x)$ is fixed for all users and is equal to 5. This means that four cells in the direction of the user along with the cell where the user resides form the cluster.
11. The DMPs are computed as in [2].
12. All users with a specific type of service have the same acceptance threshold. Algorithm 1 is used to adjust the acceptance threshold of all 10 cells and the target CDP is 1%. We assume that all 10 cells are under the control of the same Mobile Switching Center.

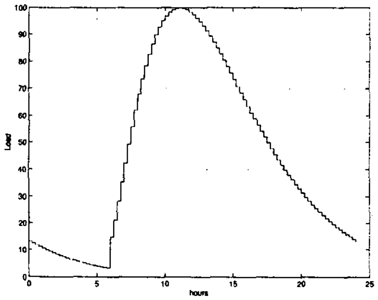

Fig. 1. Offered cell load variation for one day

The offered load per cell, L, is defined as connection generation rate × connections' bandwidth × average connection lifetime, i.e.,

$L = (R_{voice}.B_{voice} + R_{data}.B_{data} + R_{video}.B_{video}).\lambda.180.$

Knowing the average connection lifetime and bandwidth, we changed the connection generation rate to have a variable cell load as depicted in figure 1. We simulate a total of 5 × 24 hours of real-time highway traffic, with a variable cell load as shown in figure 1 for each day.

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. Figure 2 depicts the call dropping percentage achieved when using our scheme. 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.

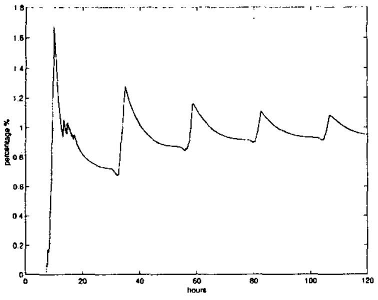

Fig. 2. Call dropping percentage

Figure 3 shows the variation of the acceptance threshold according to the hour of the day. We can notice that when the offered load is high (around 11am each day) the algorithm increases the acceptance threshold to control the number of admitted users in the network. Only those users that can be supported by the corresponding cluster are accepted. When the offered load is low (between 8pm-6am) the acceptance threshold is decreased to allow more users in the network.

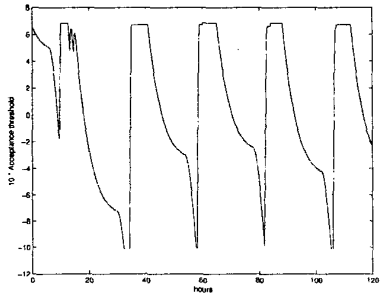

Fig. 3. Acceptance threshold

Figure 4 shows the average bandwidth utilization of the cells in the network. Figure 5 depicts the call blocking percentage achieved by our scheme. The call blocking percentage represents the ratio of blocked users to the number of admission requests in the system. The maximum call blocking percentage is about 10%, which means that about 90% of all admission requests are accepted. It is interesting to note that even when the offered load if high the blocking percentage is still low (below 11%).

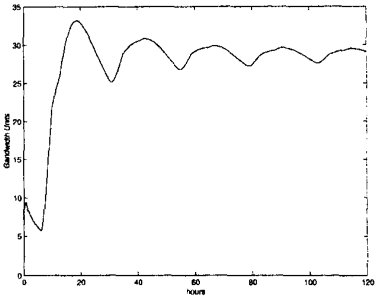

Fig. 4. Average bandwidth utilization

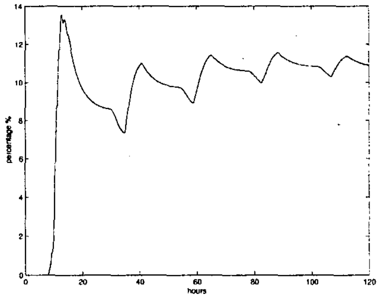

Fig. 5. Call blocking percentage

According to figures 2, 3, 4 and 5, our scheme adapts to changes in the network load and varies its acceptance threshold to achieve the target CDP while maintaining a good resource utilization and a lower blocking probability.

VIII. CONCLUSION

In this paper, we have described an adaptive call admission control algorithm suitable for wireless multimedia networks. The proposed algorithm 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. The goal underlying the design of our algorithm is to achieve a target call dropping probability while maintaining a high network resource utilization and a lower call blocking probability.

Simulations results have shown the adaptive capability of the proposed algorithm. By implementing the proposed algorithm, the system is able to lower the call dropping probability while offering a high average bandwidth utilization. The system is also able to adapt to different network loads, traffic intensity and distributions while maintaining a low blocking probability for new users.

REFERENCES

[1] 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, January 1999.

[2] 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. on Networking*, vol. 5, no. 1, February 1997.

[3] Sunghyun Choi and Kang G. Shin, "Predictive and adaptive bandwidth reservation for hand-offs in qos-sensitive cellular networks," in *ACM SIGCOMM'98 Proceedings*, September 1998, pp. 155–166.

[4] Songwu Lu and Vaduvur Bharghavan, "Adaptive resource management algorithms for indoor mobile computing environments," in *ACM SIGCOMM'96 Proceedings*, August 1996, pp. 231–242.

[5] Youssef Iraqi and Raouf Boutaba, "A novel distributed call admission control for wireless mobile multimedia networks," in *The Third ACM International Workshop on Wireless Mobile Multimedia, WOWMOM'2000 Proceedings*, August 2000.

[6] 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 *ACM MOBICOM'95 Proceedings*, November 1995, pp. 142–150.