AN ADAPTIVE HIERARCHICAL SCHEME FOR BANDWIDTH ALLOCATION IN CELLULAR NETWORKS

Sridhar Machiraju[1] **and C. Siva Ram Murthy**[2]

[1]University of California at Berkeley, machi@cs.berkeley.edu
[2]Indian Institute of Technology Madras, Chennai, murthy@iitm.ac.in

Abstract – The proliferation of wireless networks has led to a demand for providing bandwidth guarantees to mobile users that are comparable to the ones received by users in fixed networks. Error-prone communication links and user mobility make this a challenging task. In this paper, we propose a bandwidth allocation scheme for cellular networks that copes with user mobility. Our proposed scheme constructs, for each user, *Hierarchical Clusters (HCs)* which are tree-like structures having sub-clusters of cells as nodes. The HC of any user represents the cells that he is expected to visit during the lifetime of his call. Bandwidth is reserved in all cells of the HC thereby providing an estimation of future resource availability. Such estimates are used to determine the feasibility of admitting new calls. In constructing HCs, we use the velocity of the user and available data on user movement. Our scheme is adaptive since it works with any amount of available data. Results of simulations conducted are provided to demonstrate the performance obtained using our proposed scheme.

Keywords - Cellular Networks, Call Admission Algorithm, Blocking Probability, Dropping Probability.

I. INTRODUCTION

Rapid advances in wireless technology have resulted in an explosion in the use of wireless networks. Apart from providing ubiquitous connectivity, wireless networks of the future must also possess mechanisms to satisfy the diverse requirements of user applications, especially real time multimedia streams, by providing bandwidth guarantees.

Bandwidth allocation schemes use a *call admission algorithm* to decide if a call can be supported by the system. If the algorithm determines that sufficient resources are available, then the call is admitted. Calls that are not admitted are referred to as *blocked calls*. The admission may not be foolproof and may result in the *dropping* of admitted calls. However, the aim of any network is to drop as few calls as possible even if this is at the expense of blocking other calls. This is because a user would rather be denied service than have an ongoing communication abruptly terminated.

In this paper, we deal with mechanisms to incorporate user mobility into bandwidth allocation schemes of cellular networks. Several such schemes have been proposed in [1], [3], [5] and [6]. The algorithm proposed in [5] uses the number of active calls in neighboring cells along with handoff and call arrival rates to decide on call admission. It does not consider the actual mobility characteristics of users such as speed and direction of movement. It also assumes that each call request is identical. This scheme is, therefore, not suitable in networks supporting multimedia applications.

The scheme proposed in [6] dynamically estimates the resource requirements for calls in neighboring cells that may be handed off to the current cell. Such estimates are used to reserve resources for handoffs. However, this scheme fails to consider calls in non-adjacent cells that may be handed off to the current cell. This is a serious drawback in networks with small cells (and a large number of handoffs).

The schemes proposed in [1] and [3] consider users in non-adjacent cells. But they take different approaches to do this. The scheme proposed in [3] builds *Shadow Clusters (SCs)* for each user. The cells of a user's SC collectively take the call admission decision. Also, the SC of any user consists of all the cells that will consider it in their resource estimations. However, this scheme generates a lot of information that needs to be exchanged between base stations to determine future resource availability. The SC scheme also assumes the availability of data regarding user movement in *every* cell of the network. The division of the geographical region into cells being independent of traffic patterns, user movement in each cell may not be well defined at all. Even in networks with well-defined traffic patterns, collection of extensive data on user movement in each cell is extremely complex.

The scheme proposed in [1] assumes that no data on user movement is available. It uses the velocity of users to build *Most Likely Clusters (MLCs)* of cells. Explicit reservations are made in each cell of the MLC. A drawback of this scheme is that it does not contain any provision to consider known traffic behavior. Also, this scheme requires knowledge about the velocity of users periodically. This makes its implementation quite complex.

In reality, most networks would possess some information on user movement. This may not be as comprehensive and well defined as assumed in [3] or non-existent as assumed in [1]. To generalize the information available on user movement, we introduce the concept of *Special Cells*. Available information provides the expected direction(s) of users out of special cells. We also propose a simple velocity based scheme to predict movement out of non-special cells. We use the expected direction out of special cells and user velocity to build *Hierarchical Clusters (HCs)* for each user. HCs are trees with *sub-clusters* of cells as nodes. The system determines future bandwidth requirements by reserving bandwidth in each cell of the HC of admitted calls. Call admission is also based upon the availability of bandwidth in the HC of the new calls.

The rest of the paper is organized as follows: In Section II, we present the motivation behind our scheme. In Section III, we describe our proposed scheme. In Section IV, we present the results of the simulations that we conducted. In Section V, we present the conclusions drawn from our work.

II. MOTIVATION

As mentioned in the previous section, several QoS provisioning schemes incorporating user mobility have been proposed. The motivation behind our proposed scheme is largely drawn from the Shadow Cluster (SC) scheme proposed in [3] and the Most Likely Cluster (MLC) scheme proposed in [1]. The SC scheme assumes that user movement from every cell of the system to its neighbors is available. Since collection of such comprehensive data is not an easy task, our aim is to propose an adaptive scheme that can work with any amount of data on user movement.

In this paper, we assume data on user movement to be in the form of expected direction out of certain *special cells* of the system. Such data may depend on the user, time of day, or the day of week. Data independent of users could be a result of the layout of roads in cells. Data dependent on time and users would reflect observed traffic patterns.

The scheme proposed in [1] motivates us to use the velocity of users to predict movement out of cells that are not special. But, the actual scheme proposed in [1] is too complicated for implementation because it requires periodic measurement of user's velocity. Also, it uses techniques to reduce the amount of reservation based on the deviation of the user's movement from the predicted movement. Since such reduction amounts to an arbitrary action on behalf of the system, we use a pre-defined or user-specified parameter to determine the amount of reservation in our proposed scheme.

III. HIERARCHICAL CLUSTER (HC) BASED SCHEME

A. Special Cells

Available data on user movement is in the form of *Information Sets (ISs)* in special cells. ISs may be indexed using various characteristics such as user ID, time of day, day of week and initial cell of residence. The IS of any special cell m (of a user x) consists of 2-tuples $<n_i , D_i>$ ($i=1$ to M). Here, n_i is a neighbor of m and D_i is the *expected direction* of x when x enters n_i through m. M is the number of entries in the IS of m. Since the expected direction of the user need not be unique, a neighbor of m may appear more than once in this information set. If x were known not to use certain cells, then these would be special cells with empty ISs.

B. Hierarchical Cluster Construction

Our proposed scheme uses ISs of special cells and user velocity to build, for each user, a Hierarchical Cluster (HC) of cells together with the expected times of residence in each cell of the HC. The HC consists of cells that are likely to be visited by each user.

Let the current cell of residence of a user x be i. Let j be the most recent special cell visited by x. The initial cell of residence is assumed to be j, if no special cell has been visited yet. The *estimated direction* of x, D, is the direction of the line joining the centers of cell i and j. If i and j are the same cell, then D is the current direction of movement. The HC is constructed in the following manner.

1. The angle, ϕ_k, between the line joining each cell k to i and D is determined. The distance, r_k, of k (in terms of number of cells) from i is also determined. A *Level-1 sub-cluster (L-1 sub-cluster)* is constructed. It consists of all cells k such that: $\phi_k \leq \Delta$ and $r_k \leq Win_{sz}$ where Δ and Win_{sz} are system parameters.
2. The L-1 sub-cluster may contain special cells. Each special cell m of the L-1 sub-cluster will provide each n_i ($i=1$ to M) in its IS the expected direction of x on entering n_i from m. Another parameter that is provided to n_i's is R. This is explained later.
3. Further construction of the HC uses W_{rem} which is given as $W_{rem} = T_{rem}/T_{av}$. Here, T_{rem} is the time remaining for the call to finish and T_{av} is the average time of residence in each cell. Basically, W_{rem} is the average number of cells that x is expected to traverse before the call terminates. Using the expected directions and user characteristics, *Level 2 sub-clusters (L-2 sub-clusters)* are constructed by the various n_i's (present in the IS's of special cells in L-1 sub-clusters). For each such n_i, the distance of every cell k from n_i and ϕ_k, the angle that the line joining k and n_i makes with the expected direction D_i is calculated. Using these, L-2 sub-clusters are formed. They consist of cells such that: $\phi_k \leq \Delta$ and $r_k \leq S$ where S is given by: $S = Min(W_{rem} - R, Win_{sz})$. R is a parameter that is explained later.
4. The special cells present in the L-2 sub-clusters send expected direction to the n_i's in their IS's which construct L-3 sub-clusters in a manner similar to the L-2 sub-cluster construction. This construction of sub-clusters at various levels is continued till reservations are made for the duration of the call. A tree-like HC is thus obtained with the L-1 sub-cluster as the root. The

n_i's at each step are the *roots* of their respective sub-clusters.

The parameter R is used in calculating the expected times of residence of x in the cells of the HC. For the cells in the L-1 sub-cluster, R is taken to be 0. Let s be a special cell of a sub-cluster Q, which is used to continue the HC construction. Also, let n be the root of Q and let r_s be the distance (in terms of number of cells) of s from n. s passes r_s+R_n as the value of R to its n_i's, R_n being the value of R passed to n. Thus, the parameter R signifies the shortest distance of a particular sub-cluster's root from the present cell of residence i.

The expected time of residence of x in a cell h of the HC is calculated using R (to reserve bandwidth during that period). Let n be the root of the sub-cluster of h and R_n be the value of R passed to n. Also, let r_h be the distance of h from n. As in the MLC scheme [1], bandwidth in h is reserved from the expected time of earliest arrival, T_{EA}, to the expected time of latest departure, T_{LD}. These reservations cease to be effective if x has not arrived in h by the expected time of latest arrival, T_{LA}. Let t be the current time, T_{av} be the average time of residence in a cell and P_{num} denote the maximum number of extra cells (over the shortest path) that x may be allowed to traverse in order to reach any cell. The expected times of residence are calculated as

$$\mathbf{T}_{EA} = t + (r_h + R_n) \cdot T_{av}$$

$$\mathbf{T}_{LA} = \mathbf{T}_{EA} + P_{num} \cdot T_{av}$$

$$\mathbf{T}_{LD} = \mathbf{T}_{LA} + \mathbf{T}_{av}$$

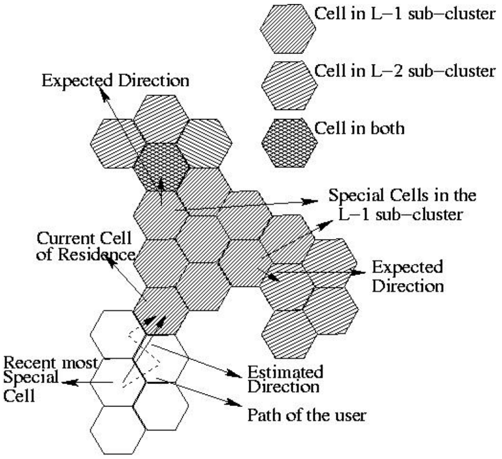

Figure 1 An example of a Hierarchical Cluster with 2 Levels

We now illustrate the construction of a HC with 2 levels through an example in Figure 1. The estimated direction is shown as the direction of movement from the recent most special cell to the current cell of residence. The L-1 sub-cluster contains 2 special cells both of which give rise to 2 L-2 sub-clusters. One cell in the L-1 sub-cluster is also present in an L-2 sub-cluster. It should be noted that in such cases, reservations are made in a cell only the first time it is encountered (in the L-1 sub-cluster in this case).

The HC needs to be reconstructed after every handoff. This is to keep the expected resource requirements up to date. Bandwidth in the cells of the HC is reserved for the time interval $[T_{EA}, T_{LD}]$. Reservations are freed in a cell of the HC if the user has not arrived there by T_{LA}.

C. Call Admission Algorithm

The admission parameters used are the time guarantee T_G and reservation fraction τ. T_G is the time for which the desired bandwidth is to be guaranteed. A call in our scheme is admitted only if 2 conditions are met. The first is that at least τ fraction of the cells in the L-1 sub-cluster possess bandwidth required to support the call in the time interval $[T_{EA}, T_{LD}]$. The second condition requires that at least τ fraction of the cells of the whole HC possess bandwidth during their respective $[T_{EA}, T_{LD}]$ time intervals. In loose terms, the first condition is to prevent short term congestion whereas the second condition is to prevent long term congestion.

IV. PERFORMANCE EVALUATION

We present the results of simulation experiments that we conducted. Performance is measured in terms of *percentage utilization, dropping probability* and *blocking probability*. Percentage utilization is the average percentage of bandwidth used by a cell in the network. Dropping probability is the fraction of accepted calls that are dropped because of lack of bandwidth. The blocking probability is the fraction of calls that are denied admission by the call admission algorithm.

Calls are taken to be of 3 possible types - audio, video and voice. The probabilities of occurrence of each type P_A, P_B, P_C (respectively) are taken to be 0.2, 0.1, 0.7. They require a constant bandwidth of B_A, B_B, B_C which are 10, 20, 2 kbps respectively. Bandwidth available in each cell is 60 kbps. The duration of calls is exponentially distributed with a mean of T_H equal to 180s. The time between two calls in a cell is exponentially distributed with a mean of λ. λ is varied and the various measures of performance are observed. The load given to the system is $T_H.(B_AP_A + B_BP_B + B_CP_C)/\lambda$.

The cellular network is taken to consist of 7 rings around a central cell. Hence, the network has 169 cells. The speed of users is a uniformly random value between 36kmph and 72kmph (i.e., between 10 m/s and 20 m/s). The distance to be covered in each cell is taken to be a constant 200m. For simplicity, no distinction is made between the initial cell of residence and other cells in calculating the time of residence. Also, P_{num} is taken to be 1.

The user's initial handoff is into a randomly chosen neighbor. Subsequent handoffs are *predictable* with a probability, p which is initially p_{init} and decreases by p_{del}

after every predictable handoff. A handoff that is not predictable is *random*. After a random handoff, p is reset to p_{init}. As the name indicates, a random handoff is into a randomly chosen neighboring cell. Predictable movement follows a "0.2-0.6-0.2" rule. This is explained with the help of Figure 2. The previous handoff is from C1 to C2. If the handoff from C2 is predictable, then the next handoff will be into C4 with a probability of 0.6. A predictable handoff may be into one of C3 and C5 with a probability of 0.2 each. A random handoff from a cell makes it a special cell provided that the number of special cells for that user is not more than 10.

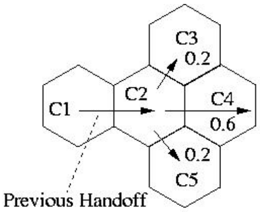

Figure 2 Predictable Handoff

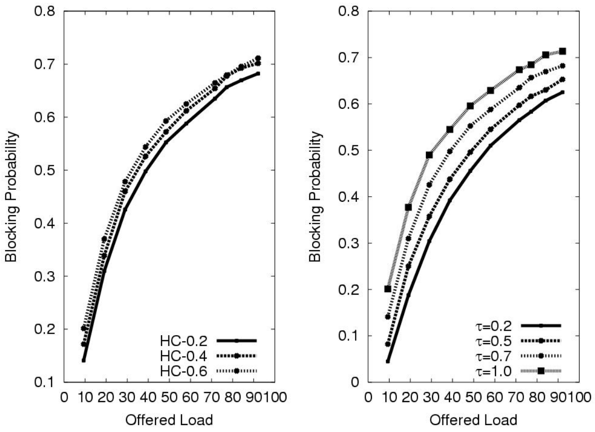

Figure 3 Blocking Probability vs. Offered Load

The admission parameters of our scheme are T_G and τ. T_G is taken to be the duration of the call. The scheme is simulated on the network for 3600 seconds. This corresponds to an addition of about 6000 calls for an extra load of 10. For instance, a load of 50 causes the generation of about 30000 calls. Also, Δ is taken to be 20° and Win_{sz} is 3. The performance of the HC scheme for different values of p_{init} (τ=0.7) and τ are plotted in Figure 3, Figure 4 and Figure 5.

The HC-0.2, HC-0.4, HC-0.6 schemes refer to the HC scheme in systems with user movement following the predictable/random movement described earlier taking the p_{init} parameter to be 0.2, 0.4 and 0.6 respectively. It can be seen from Figure 3 that the blocking probabilities of the HC-0.4 and the HC-0.2 scheme do not differ by much. The utilization plots in Figure 4 show that the HC-0.4 scheme utilizes more than 18% of the network at a load of 60. In practice, better utilization may be achieved by tolerating higher drop rates. This is a subject of future work. An interesting observation can be made from Figure 5. Dropping is actually the least in the HC-0.6 scheme. This is because, in the HC-0.6 scheme, there are a number of special cells near the initial cell of residence. This causes a lot of reservation around the initial cell of residence. Blocking is increased and dropping is reduced because of the resulting over-reservation. With the HC-0.2 scheme, such over-reservation is prevented because special cells are more uniformly distributed throughout the system.

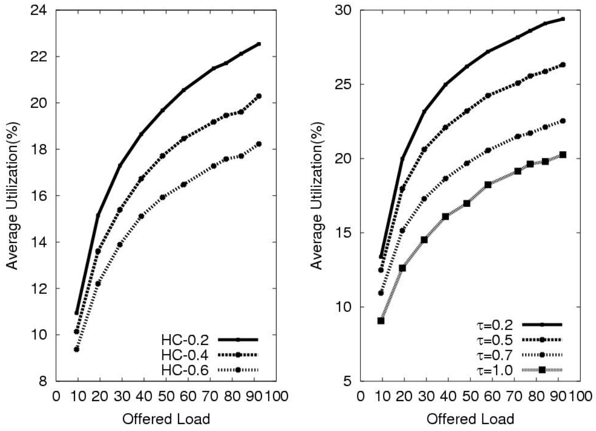

Figure 4 Percentage Utilization vs. Offered Load

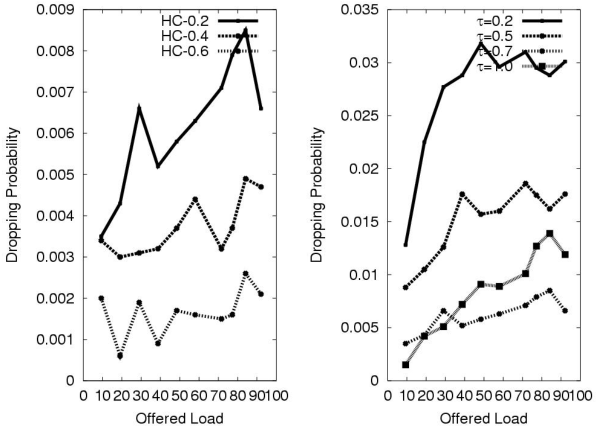

Figure 5 Dropping Probability vs. Offered Load

The plot of blocking probabilities shows the intuitive result that schemes with bigger τ values block more number of calls. This is because, for greater values of τ, admission is more constrained and lots of resources need to be available for admission. It can be seen from Figure 4 that a HC scheme with a τ value of 0.2 achieves more than 25% utilization, with a dropping probability of less than 0.035 under high loads.

One drawback of our scheme is that the size of the HC increases with the duration of calls. Hence, shorter calls are more easily accepted than longer calls. However, this is inevitable to some extent because the admission of longer calls is more constrained as they require resource availability in more number of cells. In fact, such bias towards shorter calls is more acute with larger values of τ.

This bias is responsible for the unusually large dropping of the HC scheme with a τ value of 1.0 because of the following explanation. In such a biased scheme, reservations (for shorter calls) are made in a small HC where the user is expected to move. Longer calls with less than average special cells that stray even a bit from their HC will face a bandwidth shortage. With lesser τ values, more number of longer calls are accepted and their bigger HCs prevent the occurrence of such global congestion due to erratic user movement. Detailed comparisons of our scheme and the MLC scheme are present in the longer version of this paper [13].

V. CONCLUSIONS AND FUTURE WORK

We have, in this paper, provided an adaptive hierarchical clustering scheme for bandwidth allocation in cellular networks. Our scheme is adaptive because it can be used in any network system irrespective of the amount of available information on user movement. Our scheme is quite simple and can be implemented without any need for complex hardware/software. Results of simulation experiments conducted show that our scheme provides up to 25% utilization without dropping more than 4% of the calls.

Presence of too many special cells in the system may lead to large amounts of reservation in the system. This is because the system would expect the user to take a lot of paths. Future work may be aimed at evolving mechanisms by which this problem may be reduced. Schemes to perform QoS provisioning in *ad hoc* wireless networks have been proposed in [2], [11]. Future work may be directed towards using (any amount of) available data on user movement to provide guarantees in ad hoc networks by using such data to determine the existence of routing paths for the entire time guarantee period.

VI. ACKNOWLEDGEMENT

This work was supported by the Department of Science and Technology, New Delhi, India.

REFERENCES

[1] AbdulRahman Aljadhai and Taieb F. Znati, "A Framework for Call Admission Control and QoS Support in Wireless Environments", *Proc. of the INFOCOMM '99* , pp 1019-1026.

[2] S. Chen and K. Nahrstedt, "Distributed Quality-of-Service Routing in Ad Hoc Networks", *IEEE Journal on Selected Areas in Communications*, Vol.17, No.8, pp 1488-1505, Aug 1999.

[3] David A.Levine et. al. "A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks Using the Shadow Cluster Concept", *IEEE/ACM TON*, Vol. 5, No.1, pp 1-12, Feb 1997.

[4] I. Katzela and M. Naghshineh, "Channel Assignment Schemes for Cellular Mobile Telecommunication Systems", *Personal Communications Magazine*, June 1996.

[5] M. Naghshineh and Mischa Schwartz, "Distributed Call Admission Control in Mobile/Wireless Networks", *IEEE JSAC*, Vol.14, No.4, pp 711-717, May 1996.

[6] P. Ramanathan, K.M. Sivalingam, P. Agrawal and S. Kishore, "Dynamic Resource Allocation Schemes During Handoff for Mobile Multimedia Wireless Networks", *IEEE JSAC*, Vol.17, No.7, pp 1270-1283, July 1999.

[7] R. Jayaram, Naveen K. Kakani and S.K. Das, "Call Admission and Control for Quality-of-Service (QoS) Provisioning in Next Generation Wireless Networks", *Proc. of the 5th Intl. Workshop on Mobile Multimedia Communications (MoMuc) '98* .

[8] B. RamaSubramaniam and C. Siva Ram Murthy, "Application Semantics and Seamlessness Based Admission Policy for Multimedia Mobile Networks", *Journal of Multimedia Tools and Applications* , Vol.9, No. 1, pp 7-28, 1999.

[9] S.K. Das, S.K. Sen and R. Jayaram, "Call Admission and Control for Quality-of-Service Provisioning in Cellular Networks", *Proc. of the IEEE International Conference on Universal Personal Communications (ICUPC) '97*, pp. 109-113.

[10] Sirin Tekinay and Bijan Jabbari, "Handover and Channel Assignment in Mobile Cellular Networks", *IEEE Communications Magazine*, Vol. 29, pp. 42-46, Nov. 1991.

[11] J.L. Sobrinbo and A.S. Krishnakumar, "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks", *IEEE JSAC*, Vol.17, No.8, pp. 1353-1368, Aug. 1999.

[12] T.S.Rappaport, T.H.Reed and B.D.Woerner, "Position Location Using Wireless Communications on Highways of the Future", *IEEE Communications Magazine*, Vol.34, No.10, pp 33-41, Oct 1996.

[13] S. Machiraju, “Hierarchical Cluster and Path Suggestion Based QoS Provisioning Schemes for Cellular Networks”, B.Tech Project Report, IIT-Madras, May 2000.