DYNAMIC BANDWIDTH RESERVATION IN HIERARCHICAL WIRELESS ATM NETWORKS USING GPS-BASED PREDICTION

Wee-Seng Soh and Hyong S. Kim

Department of Electrical & Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213-3890, USA

Abstract - In this paper, we propose a distributed admission control scheme based on dynamic bandwidth reservation, with the aid of GPS to track the position of the mobile terminal (MT) and predict its trajectory. We consider both wireless and wired links as potential bottlenecks and perform reservation in both types of links. By predicting a MT's target handoff cell and its remaining time to handoff, bandwidth reservation can be performed dynamically and efficiently. We perform simulations using an arbitrary hierarchical wireless ATM network with 256 cells, and show that the proposed scheme is able to prioritize handoff requests with moderate tradeoff of the blocking probability for new calls. We also show that the proposed scheme does not affect the amount of bandwidth utilization significantly, and the amount of prediction errors is approximately 1% of the total number of handoffs.

I. INTRODUCTION

In wireless networks, mobility and handoff [1, 2] place stringent requirement on network resources. From the point of view of a mobile user, forced termination of an ongoing call is less desirable than blocking a new call. Many works have proposed handoff-prioritizing techniques to reduce handoff-dropping probability (HDP) at the expense of a reduction in the total admitted traffic and an increase in new call blocking probabilities (CBP). In early work, Hong and Rappaport [5] propose to reserve a fixed number of channels in each cell exclusively for handoffs. The use of fixed number of guard channels, however, is effective only under stationary traffic conditions, and risks under-utilizing the spectrum. In order to solve the problem, a logical step would be to dynamically predict the amount of bandwidth that needs to be reserved, using information such as position, speed, direction, and bandwidth demand of each MT.

Currently, little attention has been given to the underlying wired backbone infrastructure in conjunction with handoff-prioritization and its corresponding admission control, especially when heterogeneous wireless bandwidths are required. In future broadband wireless networks such as wireless ATM (WATM), peak user bit rate could be at least 25Mbps [3, 4]. Both wireless and the backbone wired links could potentially become bottlenecks. When wired backbone is considered in conjunction with handoff prioritization, there are some practical issues that must be taken into consideration. Previous schemes in the literature that propose queuing of handoff requests for radio bandwidth availability cannot be applied in a similar way for wired backbone. It is difficult to determine which handoff request currently queued would have sufficient end-to-end bandwidth for path rerouting, as a result of bandwidth being released by other handoff rerouting or connection completion procedures throughout the entire network. Similarly, conventional techniques that set aside a common pool of bandwidth for handoff prioritization are not easily extendable to the wired backbone. It is difficult to determine how much bandwidth should be reserved in each link when different connections have diverse bandwidth requirements and different end-to-end routes. A handoff attempt would fail when there is insufficient reservation in any link along the new route. Such reservations do not improve HDP, yet they increase the CBP for new calls.

In this work, we propose a distributed admission control scheme based on dynamic bandwidth reservation, with the aid of Global Positioning System (GPS) to track the MT's position and predict its trajectory. With GPS tracking, a MT's target handoff cell and remaining time to handoff can be predicted, so that bandwidth is reserved only when they are deemed to be necessary and timely, and only along the computed rerouting path.

The remaining of this paper is organized as follows. Section II describes the proposed GPS-based bandwidth reservation technique in detail. In Section III, we present the simulation model and assumptions made. In Section IV, simulation results are presented to evaluate the performance of the proposed scheme. Finally, we give our concluding remarks in Section V.

II. BANDWIDTH RESERVATION USING GPS-BASED PREDICTION

Reserving bandwidth in anticipation of possible rerouting when a handoff occurs can reduce the HDP of handoff requests. Depending on the reservation techniques used, the level of bandwidth utilization could vary widely for the same HDP. A logical approach would be to minimize bandwidth wastage due to wrong or unnecessary prolonged reservations. This can be achieved if we can accurately predict a MT's target handoff cell and estimate its remaining time to handoff initiation. Some works in this area propose the use of Received Signal Strength (RSS) from neighboring base stations to estimate a MT's current position. However, such methods tend to suffer from both slow and fast fading. When system performance is directly coupled with the precision of position estimation, it is attractive to look into other alternatives with higher precision. We propose the use of GPS in each MT for this purpose.

Whenever a MT joins a new cell, the BS informs the MT about the topology of the cell and its neighbors. On the other hand, a MT obtains its own position information at regular time intervals ΔT using GPS, and keeps track of its previous positions over the last N intervals. We propose the use of simple linear regression to fit the N points so as to predict the direction of travel from the MT's current position. Predictions are performed by individual MTs to avoid overloading the network. In order to minimize the prediction computation frequencies, a MT shall only perform the prediction algorithm when its current position is farther than a threshold distance D_T away from the current BS. The use of linear regression is a plausible approach for predicting the direction of travel; it is simple, yet moderately accurate for short-term vehicular movement. In a short time interval, a vehicle tends to move in a straight-line manner with occasional changes in its direction of travel. If the BS's coverage area is known, we could approximate the handoff point to be along its boundary, and estimate its coordinates using simple geometry.

The remaining time to handoff (T_{remain}) can now be estimated using a MT's current speed information and its current distance from the estimated handoff point. Upon predicting T_{remain}, if it is shorter than a threshold time known as Remaining Time to Handoff Threshold (*RTHT*), the MT shall proceed to predict its target handoff cell C_{target} by noting which neighboring cell does the estimated handoff point fall within. Upon determining C_{target}, the MT sends a reservation message to its BS, which will then initiate bandwidth reservation between an anchor switch known as crossover switch (COS) [3, 4] and C_{target}, inclusive of radio bandwidth at C_{target}. Note that by restricting the reservation attempts to within *RTHT* from the estimated handoff time, we can minimize the reservation duration.

The reservation procedure attempts to reserve bandwidth in every link within the handoff segment of the rerouted path. If any link does not have sufficient bandwidth for reservation, bandwidth is not reserved in the entire handoff segment, and the reservation fails. Thus, bandwidth is only reserved when it can guarantee a handoff rerouting, else it could lead to unwieldy wastage if the reserved bandwidth could neither serve handoffs nor new calls. In order to increase the chance of a successful reservation, when a reservation attempt fails, it will be attempted again when T_{remain} computed at any time interval is lesser than half the previous T_{remain} during which the reservation attempt failed. Therefore, reservation attempts are not repeated in every time interval so that reservation signaling messages would not overload the network. By making several attempts to reserve bandwidth in anticipation of handoff rerouting well in advance, the MT increases its chance of successful handoff. Note that priority has been given to MTs that have shorter T_{remain}, because the interval between successive reservation reattempts decrease exponentially with T_{remain}. The reserved bandwidth will be released prematurely if: (i) the MT ends its call before reaching the target cell, or (ii) subsequent predictions do not indicate the need to reserve bandwidth over a consecutive number of intervals, known as the Reservation Release Threshold Time (*RRTT*). In this way, reserved bandwidth is not held longer than necessary.

When a new call request arrives at a BS, the network estimates its equivalent bandwidth requirement based on the declared traffic parameters and the required QoS. Based on the available bandwidth in each link along the computed path, the call request is accepted if sufficient end-to-end bandwidth is found, and rejected otherwise. Although granting multiple attempts to handoff reservations might improve HDP significantly, CBP for new call requests could be degraded to unbearable levels. By granting a new call request a maximum of M_{new} attempts at ΔT_{new} interval apart, we could adjust the tradeoff between these two performance measures.

During handoff initiation, if previous reservation attempts have failed or if the handoff is unforeseen due to inaccuracy in the prediction algorithm, the MT makes one final attempt to get hold of the required bandwidth for rerouting between the COS and its target handoff cell. In this case, the handoff is successful if the attempt succeeds, otherwise the call is dropped. Note that the final attempt could still make use of any bandwidth that has been reserved by other wrongly predicted C_{target} belonging to the same MT, if that link also lies within the current handoff segment. After a successful handoff, the bandwidth between the COS and the MT's old BS is released. Any reserved bandwidth due to other wrongly predicted C_{target}, if any, is also released if they are unused after the handoff. Similarly, when a call is dropped, any bandwidth currently held by the MT, including any previously reserved bandwidth, must be released from each link.

III. SIMULATION MODEL

Fig. 1 shows the 256-cell network topology used in the simulation. We assume that the underlying wired backbone network has a three-level hierarchical structure (only the top-level hierarchy is shown in Fig. 1). The lowest level in the hierarchy consists of omni-directional BS transceivers located at the center of each cell. In addition, we choose a *cluster size* of four throughout the entire wireless network. The BSs from the four cells within the same cluster are connected to a Cluster Switch (CS) using OC-3 links. The CS's from different clusters form the middle-level of the hierarchy. The network topology is divided into 11 regions indicated by different shadings. The CS's within the same region are connected to a Regional Switch (RS), also by OC-3 links. The RS's from different regions form the top-level of the hierarchical backbone, and are in turn interconnected by OC-12 links. Fig. 2 illustrates the hierarchical structure of the backbone using region 10 as an example.

In the simulations, we use three traffic classes with different characteristics as shown in Table I. We only consider unicast calls between a MT and a fixed terminal (FT), and we assume that both terminals are within the same metropolitan area. In addition, we do not differentiate between calls initiated by MTs or FTs, because the end-to-end bandwidth requirements are essentially the same for both scenarios. The interarrival times of MTs making or receiving a call in each cell are assumed to be exponentially distributed with parameter λ calls/cell/s, and the call has equal probability of belonging to any of the three traffic classes. We also assume that the average call arrival rates are uniform for the entire network. The calls' duration, on the other hand, are exponentially distributed. For class III, its bandwidth requirement is geometrically distributed. For each call, we assume that it is equally likely for the FT to be within any of the 256 cells in the entire network. It is also assumed that the FTs in each cell share the OC-3 BS-CS link with the BS located in that same cell.

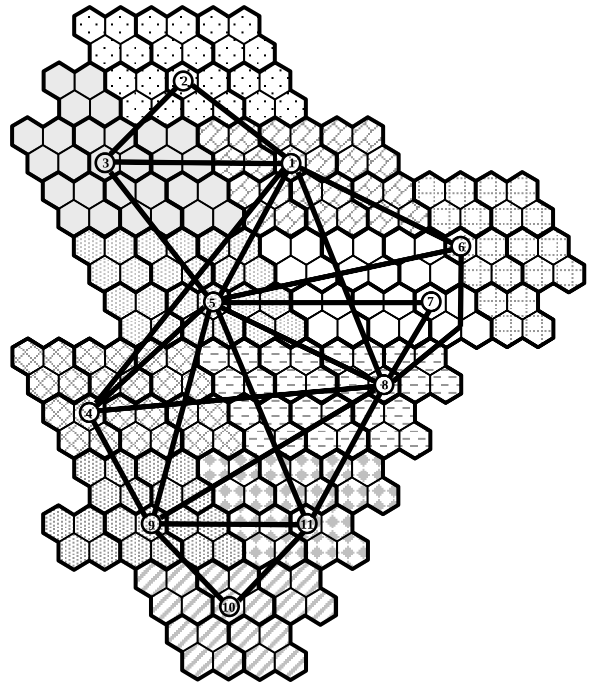

Fig. 1. Network topology showing top-level hierarchical backbone.

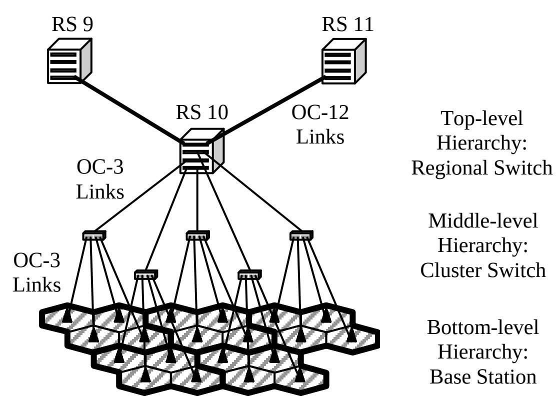

Fig. 2. Hierarchical view of backbone infrastructure.

Table I
Traffic classes used in simulations

Class	Bandwidth(B_i)	Duration	Average Duration(E[τ_i])
1	30 Kbps	1-10 mins	3 mins
2	256 Kbps	1-30 mins	5 mins
3	1-6 Mbps 3 Mbps (mean)	5 mins - 5hrs	10 mins

We assume that a MT's initial position within the cell during call establishment is uniformly distributed. Its initial speed is chosen from a normal distribution $N(S_{\text{mean}}, \sigma_{\text{speed}})$ but limited to $[0, S_{\max}]$. We also assume that each MT will travel towards a random virtual destination cell (VDC) within the network with some degree of randomness. A MT's initial direction of travel is assumed to be the bearing towards the center of the VDC. Its speed and direction may change according to the Bernoulli Process with probability P, at time intervals equal to multiples of its position update time ΔT. The change in speed is chosen from a normal distribution $N(0, \sigma_{\text{speed}\Delta})$. However, the new speed is always confined to $[0, S_{\max}]$. The new heading of each MT, on the other hand, is chosen from a normal distribution $N(0, \sigma_{\text{heading}\Delta})$, from the current bearing towards the center of the VDC. For simulation purpose, it is assumed that a handoff request is initiated whenever a MT crosses its current cell's border.

Table II
Simulation Parameters

Parameter	Value	Description
R	500 m	Cell radius
ΔT	1 s	Location update time interval
C	50 Mbps	Radio capacity per cell
λ	*Variable*	Avg. calls per second per cell
M_{new}	*Variable*	No. of attempts per new call
ΔT_{new}	*Variable*	Time between new call retries
N	5	No. of past location data used
D_T	400 m	Threshold distance from BS
RTHT	15 s	Remaining time to handoff threshold
RRTT	5 s	Reservation release threshold time
P	0.1	Bernoulli prob. for speed/direction Δ
$S_{\max}$	120 km/h	Maximum MT speed
S_{mean}	60 km/h	Mean initial MT speed
σ_{speed}	20 km/h	Standard deviation of initial speed
$\sigma_{\text{speed}\Delta}$	10 km/h	Standard deviation for speed Δ
$\sigma_{\text{heading}\Delta}$	30°	Standard deviation for direction Δ

Table II summarizes the parameter values used in the simulations. Through the simulations, we evaluate the following performance measures with respect to the system's normalized offered load: (i) CBP of new call requests, (ii) HDP of handoff requests, and (iii) bandwidth utilization of hierarchical backbone and radio links. For an average call arrival rate of λ calls/sec/cell that is uniform over the entire network, we define the normalized offered load with respect to the total radio bandwidth capacity as follows:

$$L = \frac{\sum_{i=1}^{3} 256 \times \lambda/3 \times E[\tau_i] \times B_i}{256 \times C} = \frac{\lambda \sum_{i=1}^{3} E[\tau_i] \times B_i}{3 \times C}, \text{ where}$$

$E[\tau_i]$ = average class i call duration in seconds,
B_i = average bandwidth for class i in Mbps, and
C = radio bandwidth capacity per cell in Mbps.

In this work, we perform three sets of simulations:
(i) no reservation, no backbone, $M_{\text{new}} = 1$ (**NB**),
(ii) no reservation, $M_{\text{new}} = 1$ (**NR**), and
(iii) reservation, $M_{\text{new}} = 2$, $\Delta T_{\text{new}} = 10$ secs (**R2**).

IV. SIMULATION RESULTS

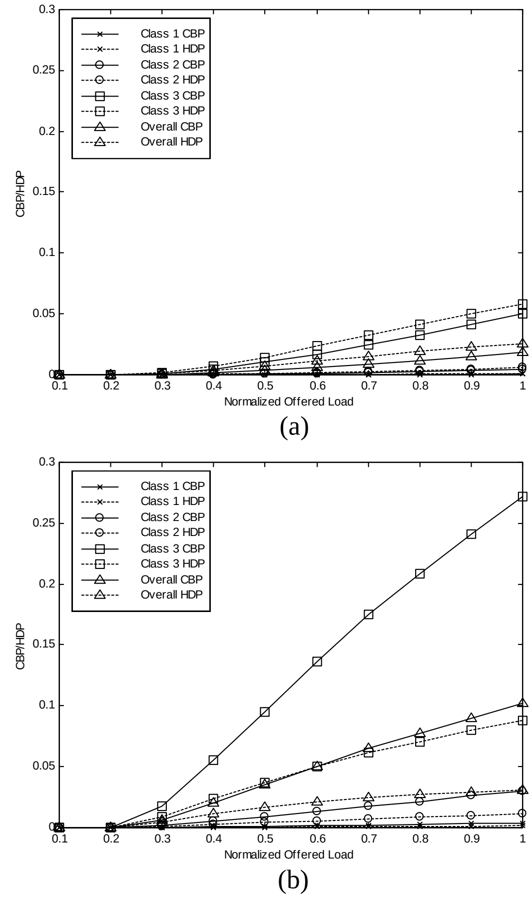

Fig. 3. CBP/HDP vs. normalized load L
(a) simulation NB, (b) simulation NR.

Fig. 3(a) and (b) show how the CBP/HDP values vary with normalized offered load L for NB and NR respectively. Their CBP/HDP values increase with L, as expected. Also, the values increase with the bandwidth requirement of the traffic class, with class 1 having the lowest CBP/HDP while class 3 having the highest. However, there are two important differences between the two sets of simulation results:
(i) In simulation NB (see Fig. 3(a)), the CBP values are very close to the HDP values for each traffic class. When there is no reservation and no consideration for the backbone, both new call and handoff requests are treated the same. Note however, that the CBP and HDP values are not exactly equal because different cells experience different handoff request rates due to non-uniformity in the chosen mobility model, whereas new call arrival rates are assumed to be uniform across the entire network. In simulation NR where backbone network is considered (see Fig. 3(b)), the CBP values are much larger than the HDP values for all traffic classes. A handoff request only needs to find sufficient bandwidth for rerouting between

the COS and the new BS, whereas a new call request needs to find sufficient end-to-end bandwidth. Thus, handoff requests have higher chances of success than new call requests inherently, even without bandwidth reservations.

(ii) In simulation NR, the CBP and HDP values are much higher than those from simulation NB. When backbone network is considered, as is done in simulation NR, both radio and wired links could potentially become bottlenecks for new call and handoff requests. Thus, it is natural for the CBP and HDP values to be higher in simulation NR than the corresponding values in simulation NB.

The above shows that results could be quite misleading if we only consider radio bandwidth limitation, and ignore the wired backbone's bandwidth requirement.

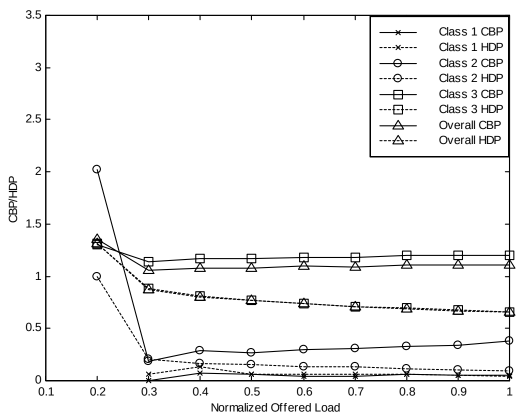

Fig. 4. Normalized CBP/HDP vs. load L for R2.

In order to track how the CBP and HDP values are affected by our reservation scheme, we normalize the values obtained from R2 by those obtained from NR. A normalized value of 1.0 would indicate that the performance neither improve nor degrade. Fig. 4 shows how the normalized CBP/HDP vary with normalized load L for R2. For L = 0.1, normalization is impossible because all CBP/HDP values are zero. For L = 0.2, the normalized CBP/HDP appear to be large because the corresponding values in NR were very small. Note however, that the actual values obtained from R2 are still quite small ($<10^{-2}$). Interestingly, for $L \geq 0.3$, both normalized CBP and HDP of classes 1 and 2 fall below 1.0, meaning that both have improved over the non-reservation case in simulation NR. For class 3, we also see that the normalized CBP values are only approximately 1.2 for $L \geq 0.3$. The corresponding normalized HDP values, on the other hand, are well below 1.0.

We have also investigated the amount of bandwidth utilization and reservation in different types of links. It is found that the amount of bandwidth utilization decreases by at most 1.5% when reservation is enabled. The average amount of reservation in RS-RS links is approximately 2% of their corresponding average utilization. For the RS-CS, CS-BS and radio links, the values are 4%, 7% and 15% respectively. The RS-RS links have the smallest proportion because inter-regional handoffs occur least frequently, while radio links and CS-BS links have high proportions because every predicted handoff would attempt to reserve bandwidth in these two types of links.

In our scheme, two types of prediction errors could potentially occur: (i) unforeseen handoff - no handoff has been predicted by the MT but a handoff occurs, and (ii) wrong target predicted - the target handoff cell is different from the one predicted. From the simulations, it is observed that the amount of unforeseen handoff is approximately (0.25±0.05)%, while the amount of wrong target prediction is approximately (0.7±0.05)%. Therefore, the amount of prediction errors is quite low for the assumed mobility model and the chosen prediction algorithm.

V. CONCLUSION

In this paper, we have proposed a distributed admission control scheme based on dynamic bandwidth reservation, with the aid of GPS to track the MT's position and predict its trajectory. Bandwidth reservation is performed on both wired and radio links. We have shown through simulations that the proposed scheme is able to reduce the HDP of handoff requests significantly (up to a maximum of 3%), and that the amount of bandwidth utilization is not affected significantly (by at most 1.5%). Finally, we have shown that the amount of prediction errors is approximately 1% of the total number of handoffs.

REFERENCES

(1) Acharya, A. *et al.*, "Mobility management in wireless ATM networks," *IEEE Commun. Mag.*, Nov. 1997, pp. 100–109.

(2) Akyol, B.A. and Cox, D.C., "Rerouting for handoff in a wireless ATM network," IEEE Pers. Commun., Oct. 1996, pp. 26–33.

(3) ATM Forum, "Wireless ATM requirements specification – Draft," July 1998.

(4) ATM Forum, "Draft wireless ATM capability set 1 specification," July 1998.

(5) Hong, D. and Rappaport, S.S., "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoff procedures," *IEEE Trans. Veh. Technol.*, vol. VT-35, no. 3, Aug. 1986, pp. 77–92.