Adaptive Resource Reservation for Indoor Wireless LANs

Songwu Lu, Rayadurgam Srikant and Vaduvur Bharghavan

Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801 / USA
{s-lu, r-srikant, b-vaduvur}@csl.uiuc.edu

Abstract

Emerging communication-intensive applications such as multimedia teleconferencing require significant networking resources for efficient operation. In an indoor mobile computing environment, limited wireless resources and user mobility require effective resource reservation algorithms. These requirements demand an efficient advance reservation algorithm in order to minimize handoff call dropping while not compromise network utilization.

This paper proposes a design framework and algorithms for advance resource reservation over the wireless link in indoor cellular wireless LANs. The key insight presented in this paper is that different cells in an indoor mobile computing environment behave differently based on location, and need to manage their resources according to their behavior profiles. We present two results: (a) classification of cells and predictive advance reservation based on location and behavior profiles, (b) state-dependent resource reservation based on statistical information.

1 Introduction

Recent years have witnessed explosive research in the field of mobile computing, resulting in the development of indoor mobile computing environments which seek to provide a mobile user with not only a basic toolchest of applications, but also more advanced applications such as video teleconferencing, WWW browser, etc., which will stress the wireless network severely. The fundamental problem is that wireless resources are scarce and need to be managed efficiently in order to provide an acceptable quality of service (QoS) to these communication-intensive applications. In particular, in order to cope with inter-cellular mobility, advance resource reservation in the neighboring cells of the current cell where a mobile is residing seems to be a necessity. However, in order to minimize handoff call dropping, advance resource reservations should not be overly aggressive such that the new call blocking increases drastically. State-of-the-art solutions advocate either brute-force advance reservation [1] or predictive advance reservation [2] that requires a very detailed knowledge of the environments, therefore are inadequate to meet such resource reservation challenges. This paper proposes a design framework and algorithms for advance resource reservation over the wireless link in indoor wireless LANs that have cellular architectures. The reservation algorithm is adaptive in the sense that the resource reservation is adaptively tailored to the location-dependent user behavior.

Our framework has three key features:

- provision of loose and dynamically renegotiable QoS bounds [5]. In this paper, our focus is the wireless link, and we have a more narrow perspective of QoS bounds: (a) we assume that a higher layer has translated the end-to-end QoS bounds into per-hop requirements, and (b) we only consider bandwidth. Essentially, over the wireless link, we require that an application be provided bandwidth in an adjustable range of (b_{min}, b_{max}), and our goal in this paper is to provide at least the minimum required bandwidth to all applications.
- *classification of cells based on location and behavior profiles*, motivated by the following observation: cells have handoff and resource characteristics that are dependent on their location. Besides, in an office environment, monitoring the past behavior of a cell provides insight into the future behavior in the cell. We therefore classify cells based on location and past behavior profiles, and provide different predictive advance resource reservation algorithms for each class of cells.
- *state-dependent resource reservation*. When profiles do not provide enough information to characterize a cell, we use a current-state-dependent resource reservation algorithm, which is based on the ongoing 'calls' in the cell and its neighbors, call arrival, termination and handoff rates. The goal of the algorithm is to find an operating point which achieves an acceptable tradeoff between new-call blocking probability and handoff blocking probability.

The rest of the paper is organized as follows. Section 2 provides a cell classification and Section 3 presents the predictive advance bandwidth reservation based on portable and cell profiles. Section 4 describes the default state-dependent algorithm and Section 5 reports its performance. Section 6 concludes the paper.

2 Cell Classification and Profiles

The network has a standard cellular indoor wireless LAN (WLAN) architecture, consisting of a wired *backbone* component and a wireless cellular component. *Base stations* are connected to the backbone network and provide networking access to *portable* computers via a single-hop wireless link within a geographical region round them called *cells*. Cells are small: each office is typically a single cell, a corridor is a row of cells, and larger meeting rooms and lounges are clusters of cells. Two cells are called *neighbors* if it is possible for a handoff to occur between them.

2.1 Cells and zones

The current cell of a portable and the set of neighboring cells form the *neighborhood* for a portable at any time. The set of cells in the indoor WLAN is called the *universe of cells*. The universe is divided into distinct geographical regions called *zones*. Each zone has a *profile server*. The locational hierarchy for a portable is thus the current cell, neighborhood, zone, and universe.

Cells are divided into three classes based on predictive behavior of occupants: *office*, *corridor* and *lounge*. Lounges are further divided into three sub-classes: *meeting room*, *cafeteria* and *default*. The characteristics of each class are as follows: an office is typically occupied by a few *regular* occupants; users typically move *linearly* along corridors (except at intersections); a meeting room typically has bursts of handoff activity at the *start* and *conclusion* of meetings (the times of which can be pre-specified in a 'booking calendar'); a cafeteria typically has a predictable pattern of activity over the course of a day. A default cell is one which does not fall into any of the above classes or does not have sufficient profile information to classify it. It should be noted that the above classification is based on measurements on handoff activity. An implication of this is that a cell in a real-world cafeteria may end up with being defined as a *default* cell.

2.2 Profiles and Profile Servers

Each portable and cell in the WLAN environment has a *profile*. The profile of a *portable* contains its identification, authentication information, and an aggregated history of its previous handoffs, which is used to predict its next cell given its current cell. The profile of a *cell* contains its identification, authentication information, cell class, the set of its neighbors and their cell class, and an aggregated history of its previous handoffs, which is used to predict the next cell for a portable (in absence of information from the portable's profile). Table 1 specifies the contents of the portable and cell profiles.

Each zone has a profile server. The profile server maintains the cell-profiles for all the cells in its zone, and the portable-profiles for all the portables currently in its zone, and updates the cell/portable-profile upon each handoff. The profile server keeps track of each handoff for each portable in each cell, and computes the next-predicted cell using a brute force method by aggregating the recent history for each portable and each cell. The actions of the profile server are fairly straightforward and described in [5].

3 Predictive Advance Reservation

The overall idea of our resource reservation scheme is as follows: given the profile of a portable and its cell (in descending order of precedence), it will typically be possible to predict the next (neighboring) cell into which the portable will move. In this case, resources are only reserved for the portable in the next cell. The goal is to make accurate predictions based on the individual past history of the portable, and of the aggregate past history for all the portables in the cell. In each cell, a common bufferpool of resources are maintained in order to handle erroneous predictions.

- First-level prediction using the *portable* profile: given the *previous cell id* and the *current cell id*, the base station checks the *next-predicted-cell* field (see Table 1) for the portable. If it is not *empty*, then the prediction is successful; otherwise, prediction will be made based on the *cell* profile.

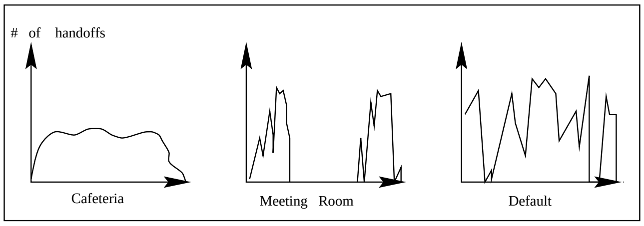

Figure 1: **Handoff Activity In a Lounge**

- Second-level prediction using the *cell* profile: if a neighboring cell is an *office* and the portable is a *regular* occupant of that office, then the next predicted cell is that neighboring office. Otherwise, prediction will be made based on the aggregate history: the next predicted cell is the cell to which most handoffs have occurred in the recent past.
- If the above predictors are not available for a cell, the (default) state-dependent resource reservation algorithm, to be described in next section, is used to conduct resource reservation.
- A separate buffer pool of bandwidth is maintained for handling exception cases, which includes wrong prediction.

We now present the details of cell classification and prediction, based on cell heuristics and aggregate history.

3.1 The office and the corridor case

An *office* is a cell with a small set of *regular* occupants. An *office* cell makes advance bandwidth reservations only when the mobile portable in its neighboring cells is a *regular* occupant of the office.

A *corridor* is a cell such that users typically move in one of several possible directions across the cell. Knowing the previous cell, the next cell can be generally predicted accurately based on recent handoff history.

3.2 The lounge case

A *lounge* is a cell which has many 'non-regular' users. A lounge does not distinguish an individual user's behavior, but aggregates the behavior of all the users in its cell. Based on aggregate behavior, a lounge can be sub-classified into three categories: *cafeteria* with a *slow* time-varying profile; *meeting room* with a profile characterized by bursts of handoffs at the start and conclusion of the meeting, and *default* with a *random* time-varying profile. Figure 1 shows the profile for each class.

3.2.1 The meeting room

In a meeting room, most handoffs occur at and around the start and conclusion of meetings, with few handoffs in between. The profile includes a *booking calendar* of meeting schedules, with each meeting specifying its start time T_s, stop time T_a, and the required resources N_m. The reservation policy for a meeting room is as follows:

(a) Starting from time $T_s - \Delta_s$ (Δ_s is a tunable design parameter set to 10 minutes in our simulations), the base

type		handoff activity	profile contents[1]
office (c)		having limited regular occupants	$\omega(c)$, $\eta(c)$, $\forall i \in \eta(c), < i, \forall j \in \eta(c), \{j, p_j\} >$
corridor (c)		linear movements	$\eta(c)$, $\forall i \in \eta(c), < i, \forall j \in \eta(c), \{j, p_j\} >$
lounge (c)	meeting room	bursts of handoffs at the start and conclusion	$\eta(c)$, booking calendar, $\forall i \in \eta(c), < i, \forall j \in \eta(c), \{j, p_j\} >$
	cafeteria	slow time-varying	$\eta(c)$, $\forall i \in \eta(c), < i, \forall j \in \eta(c), \{j, p_j\} >$
	default	uniformly distributed	$\eta(c)$, $\forall i \in \eta(c), < i, \forall j \in \eta(c), \{j, p_j\} >$
portable			$\forall i, \forall j \in \eta(i), < j, i, \textit{next-predicted-cell} >$

functions	$\omega(c)$	regular occupants in office c	
	$\eta(c)$	neighbors of cell c	

Table 1: Cell and Portable Profiles

station in the meeting room will advance reserve resources of amount N_m. The base station also maintains a counter $N_{arrived}(t)$ to record the amount of resources that have been allocated by time t. At any time t, the base station advance reserves resources of amount $N_m - N_{arrived}(t)$. After time T_s, the base station starts a timer (5 minutes) and releases unused reserved resources for the meeting once the timer expires.

(b) Starting at time $T_a - \Delta_a$ (Δ_a is set to 5 minutes in our simulations), the base station at the meeting room notifies its neighboring cells to reserve bandwidth of total amount $N_{arrived}(T_a - \Delta_a)$ (as defined in (a)) assuming every user is leaving. The base station maintains a counter $N_{left}(t)$ to record the amount of resources that have been released by time t, and notifies its neighbors to reserve resources according to the amount $N_{arrived}(T_a - \Delta_a) - N_{left}(t)$ at time t. At T_a, the base station starts a timer (15 minutes), and asks its neighbors to release reserved resources once the timer expires.

3.2.2 The cafeteria

If the current cell is a cafeteria that has a slow-time varying nature[2], then the reservation policy is as follows: It will predict the amount of reserved resources at next time instant $t+1$, denoted as $N_{handoff}(t+1)$, and ask its neighboring cells to advance reserve bandwidth of total amount $N_{handoff}(t+1)$. Accordingly, if at least one of its neighboring cells is a *default*, it will also predict the amount of resources that need to be reserved in advance (due to handoffs from its neighboring cells), denoted as $N_{arv}(t+1)$. Then it updates its current total reserved resources to amount $N_{arv}(t+1)$ to accommodate expected arriving portables at time $t+1$. The reason for doing this is that since the current cell has a *default* neighbor which provides a poor quality of next-cell prediction, it should not totally "trust" that default cell, therefore, the current cell will also predict by itself the advance reserved resources (due to handoffs from its neighbors) at the next time instant.

The prediction of $N_{handoff}(t+1)$ and N_{arv} is based on a linear model to fit past aggregate handoff data (over a few days at the same time) using least-squares methods.

3.2.3 The default case

The default cell cannot be classified as any of the above. For the default cell, we adopt a one-step-memory policy for the prediction of the advance reserved resources at next time instant, denoted by $N_{handoff}(t+1)$, i.e., $N_{handoff}(t+1) = N_{handoff}(t)$.

In the case when at least one of its neighboring cells is also a *default*, since a default neighbor provides poor quality of next-cell prediction, the current default should not totally "trust" its default neighbor. Therefore, the current cell will also try to predict the total amount of bandwidth to be reserved in it at the next time instant, to accommodate handoff portables from its default neighbor and others. To achieve this goal, the resource reservation algorithm does not use profiles, rather only current state, and recently measured handoff rates. The following section describes the algorithm in detail.

4 State-dependent Reservation Algorithm

In this section, we provide a state-dependent resource reservation algorithm, which is especially useful in the case where the profile system is not available for a cell.

Assuming that the system state information (the number of current ongoing calls in each cell), the arrival rate of new calls in a cell, the average duration of calls in a cell, and the hand-off probability are available, our objective is to keep the hand-off dropping probability below some pre-specified design threshold. As in [3], the idea here is to look ahead at a time window given by $[t, t+T]$, where t is the current time and T is the window size. During this time window, we keep the fraction of dropped handoff calls below some pre-specified level P_{QOS}.

Our reservation decision is based on a two-cell model, depicted in Fig. 2. We denote C_q as the cell in which a resource reservation decision has to be made. We denote C_s as the neighboring cell of C_q, possibly after the aggregation of all

[2] Note: The term "cafeteria" is a misnomer, since cafeteria cells may not depict the above behavior, but tend to behave more like a *default* lounge.

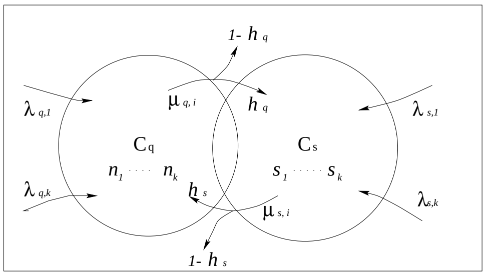

Figure 2: **Model for Reservation Design**

neighboring cells of C_q. It should be noted that, however, our approach can be extended to the multiple neighbor case without any conceptual difficulty.

We consider k call types, labeled $i = 1, 2, \ldots, k$, each having a different bandwidth requirement b_i. Moreover, we assume that the new call (of type i) arrival to cells C_q, C_s are with rates $\lambda_{q,i}, \lambda_{s,i}$, respectively; the mean call duration (of call type i) $1/\mu_{q,i}, 1/\mu_{s,i}$ in cells C_q and C_s is exponentially distributed. When a mobile leaves cell C_q, it is handing off to cell C_s with probability h_q, and it leaves the system entirely (i.e. terminating) with probability $1 - h_q$. Similarly, when a mobile leaves cell C_s, it is handing off to cell C_q with probability h_s, and it leaves the system entirely (i.e. terminating) with probability $1 - h_s$.

Consider any call of type i at time t in cell C_q. We denote the probability that a call remains in the same cell for the next T time units as $p_{s,i}$ and the probability that it hands-off to cell C_s as $p_{m,i}$. Furthermore, we assume that the probability that a mobile hands-off more than once during time T is negligible, i.e. mobiles are not moving rapidly. Besides, we ignore the hand-off effects due to calls newly admitted into cell C_s during $[t, t+T]$. This implies that whenever there is a space conflict between a handoff call and an already existing call in a cell, the call with a later arrival time is dropped. Thus, admission control can be implemented based on the current state only, and future call arrivals can be ignored. Note that this model is slightly different from the one in [3] in the way the handoff probability is modelled. Further, we allow for multiple call types.

Based on the above exponential distribution assumption, $p_{s,i}$ and $p_{m,i}$ can be calculated as follows:

$$p_{s,i} = e^{-\mu_{q,i} T}, \qquad p_{m,i} = (1 - e^{-\mu_{q,i} T}) h_q \ .$$

In our model we consider a homogeneous system, where each cell has maximum bandwidth $B_{\max}$, and denote P_{QOS} as the highest tolerable handoff dropping probability in a wireless system where all calls of different types require the same P_{QOS} throughout their connection. Therefore, the design objective is to reserve a minimum quantity of bandwidth, such that the desired quality of service (in terms of P_{QOS}) of the *ongoing* calls in cell C_q at (future) time $t + T$ is maintained.

Given the above assumptions, and denoting the number of (type-i) calls in cell C_q at current time t as n_i, the probability that j_i calls (out of $n_i + 1$ calls by supposedly admitting a new type-i call at time t) are in the same cell C_q at time $t + T$ has a binomial distribution, denoted $B(j_i; n_i + 1, p_{s,i})$, which is specified as

$$B(j_i; n_i, p_{s,i}) = \binom{n_i + 1}{j_i} p_{s,i}^{j_i} (1 - p_{s,i})^{n_i + 1 - j_i} \ . \qquad (1)$$

Let us assume that there are s_i (type-i) calls in cell C_s at time t. Then the probability that l_i calls of type-i (out of s_i calls at time t) in cell C_s hand-off to cell C_q by time $t + T$ is given by

$$B(l_i; s_i, p_m) = \binom{s_i}{l_i} p_m^{l_i} (1 - p_m)^{s_i - l_i} \ . \qquad (2)$$

At time t, suppose that a new call request of type i is arriving, it will be admitted if the following holds:

$$1 - Prob(\sum_{i=1}^{k} b_i (l_i + j_i) \leq B_{\max}) \leq P_{QOS}, \qquad (3)$$

with $0 \leq j_i \leq n_i, 0 \leq l_i \leq s_i$.

A solution approach to the inequality (3) is through the characteristic function technique via FFT, which is quite standard in literature [4]. We feel that the Gaussian approximation for binomial distributions, as suggested in [3], will not be appropriate for an indoor setting where the capacity of each cell (in terms of the number of calls that they can carry) is quite small.

Note that the reservation given above in current cell C_q is implicit, because the reserved bandwidth is only effective for handoff calls and new calls are not eligible for using it. For a new request, the network is admitting it only when the network resources can accommodate all the existing calls plus the new call in future time $t + T$.

The implementation of this algorithm is as follows: the bandwidth requirements of calls are partitioned into k regions (not necessarily partitioned evenly). A call is characterized as *type* k if its bandwidth requirement falls within the k-th region of the partition. We then choose a prediction window T, and measure the existing calls in the current cell and those in its neighbors (after aggregation). Then, based on above calculation, whenever a new call requesting is coming, we can check whether the inequality (3) is satisfied numerically. It should be noted that in the case when no information regarding the statistical data, i.e. the new call mean arrival rate, call mean departure rate, and call holding time in a cell, is available for a cell, the data can be roughly estimated by the aggregate history of another default cell in the zone by implicitly assuming most default cells in a zone behave similarly.

5 Performance of the State-dependent Reservation Algorithm

In this section, practical design issues of the state-dependent reservation algorithm provided in last section are discussed via a simulation example.

It can be noted that the choice of two parameters for the reservation algorithm is important for design, i.e., the time window T and the target dropping probability P_{QOS}. The choice of these two parameters will balance the tradeoff between actual handoff dropping probability (P_d) and the overall blocking probability (P_b). Ideally, a good choice of P_{QOS} will keep the overall blocking probability as small as possible while achieving the desired P_d, and a good choice of T will make the prediction accurate.

We show how the design is performed using the following example. We consider a situation where the two cells have identical characteristics. The capacity of each cell is 20, arrival rate is 10, holding time is 1 and the handoff probability is 0.2. The performance of the state-dependent resource reservation algorithm is shown in Figure 3.

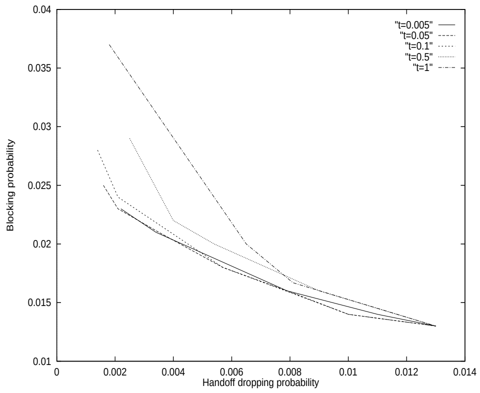

Figure 3: Performance of the state-dependent reservation algorithm for different values of the time window T

A family of curves have been obtained for different values of the time window T. Each curve is a plot of P_d versus P_b. As is to be expected, P_b decreases with increasing P_d. All curves lie on top of each other for large P_d. (Note that, large P_d denotes that all calls are admitted irrespective of whether the handoff calls will be dropped or not, i.e., a new call is admitted if there is sufficient bandwidth independent of its effect on handoff dropping.) Thus, value T_1 is better than another value T_2 if the curve for T_1 lies below the curve for T_2. For this example, it seems beneficial to choose T small. However, there is very little difference for $T < 0.1$. Once a particular value of T is chosen, the network designer must choose an operating point. However, once an operating point is chosen, then the value of P_{QOS} (the target handoff dropping probability) should be available. This is provided in Table 2 for $T = 0.05$ as an example. Suppose we choose $P_d = 0.22\%$ and $P_b = 2.27\%$ as the operating point, then, from Table 2, P_{QOS} will be chosen as 1.6%.

We also present results in Table 3 for a static reservation scheme, i.e., one where a certain capacity r is reserved in each cell for handoff calls. Notice that, for a small size, there is not enough flexibility in choosing an operating point with a static scheme. For example P_d jumps from 0.26% to 0.035%, when r changes from 1 to 2. Thus, if wishes to choose a P_d (which also affects P_b) between 0.26% and 0.035%, the static scheme cannot be used.

P_{QOS} (%)	Actual P_d (%)	Overall P_b (%)
0.1	0.038	3.42
0.4	0.16	2.5
1.6	0.22	2.27
3.2	0.35	2.13
4.2	0.58	1.84
4.8	0.79	1.66
6.0	1.03	1.39
6.4	1.29	1.26

Table 2: Simulation results for a specific time window ($T = 0.05$): State-dependent reservation

Reserved bandwidth	Actual P_d (%)	Total P_b (%)
0	1.29	1.26
1	0.26	2.37
2	0.035	3.77
3	0.008	5.43
4	0.000	7.68
5	0.000	10.51
6	0.000	13.65

Table 3: Simulation results for a static reservation algorithm

6 Conclusion

In this paper, we propose a framework and algorithms for predictive advance resource reservation in WLAN environments. The key insight is that different cells for WLANs behave differently based on location, and need to manage their resources according to their behavior profiles. Some preliminary measurements in an in-door WLAN environment have partially validated our approach and our algorithms are being evaluated in an experimental testbed.

References

[1] K. Lee, "Adaptive Network Support for Multimedia," *Proceedings of the ACM Conference on Mobile Computing and Networking*, Beerkeley, California, November 1995.

[2] D. A. Levine, I. F. Akyildiz and M. Naghshineh, "The Shadow Cluster Concept for Resource Allocation and Call Admission in ATM-based Wireless Networks," *Proceedings of the ACM Conference on Mobile Computing and Networking*, Beerkeley, California, November 1995.

[3] M. Naghshineh and M. Schwartz, "Distributed Call Admission Control in Mobile/Wireless Networks," *IEEE Journal on Selected Areas in Communications*, pp.711-717, Vol. 14, No.5, May 1996..

[4] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[5] S. Lu and V. Bharghavan, "Adaptive Resource Management Algorithms for Indoor Mobile Computing Environments," *Proceedings of ACM SIGCOMM'96*, Stanford, California, August 1996.