Cleveland State University
EngagedScholarship@CSU

Electrical Engineering & Computer Science Faculty
Publications

Electrical Engineering & Computer Science
Department

4-2011

Behavior-Based Mobility Prediction for Seamless Handoffs in Mobile Wireless Networks

Weetit Wanalertlak
Oregon State University, wanalewe@eecs.oregonstate.edu

Ben Lee
Oregon State University, benl@eecs.oregonstate.edu

Chansu Yu
Cleveland State University, c.yu91@csuohio.edu

Myungchul Kim
Information and Communications University, mckim@icu.ac.kr

Seung-Min Park
Electronic and Telecommunications Research Institute, minpar@etri.re.kr

See next page for additional authors

Follow this and additional works at: https://engagedscholarship.csuohio.edu/enece_facpub

Part of the Digital Communications and Networking Commons

How does access to this work benefit you? Let us know!

Publisher's Statement

The final publication is available at Springer via http://dx.doi.org/10.1007/s11276-010-0303-x

Authors

Weetit Wanalertlak, Ben Lee, Chansu Yu, Myungchul Kim, Seung-Min Park, and Won-Tae Kim

Behavior-based mobility prediction for seamless handoffs in mobile wireless networks

Weetit Wanalertlak · Ben Lee · Chansu Yu ·
Myungchul Kim · Seung-Min Park ·
Won-Tae Kim

Abstract The field of wireless networking has received unprecedented attention from the research community during the last decade due to its great potential to create new horizons for communicating beyond the Internet. *Wireless LANs* (WLANs) based on the IEEE 802.11 standard have become prevalent in public as well as residential areas, and their importance as an enabling technology will continue to grow for future pervasive computing applications. However, as their scale and complexity continue to grow, reducing handoff latency is particularly important. This paper presents the Behavior-based Mobility Prediction scheme to eliminate the scanning overhead incurred in IEEE 802.11 networks. This is achieved by considering not only location information but also group, time-of-day, and duration characteristics of mobile users. This captures short-term and periodic behavior of mobile users to provide accurate next-cell predictions. Our simulation study of a campus network and a municipal wireless network shows that the proposed method improves the next-cell prediction accuracy by $23 \sim 43\%$ compared to location-only based schemes and reduces the average handoff delay down to $24 \sim 25$ ms.

Keywords Mobility prediction · Fast handoffs · WLANs · WMNs

W. Wanalertlak · B. Lee (✉)
School of Electrical Engineering and Computer Science,
Oregon State University, Corvallis, OR 97331, USA
e-mail: benl@eecs.orst.edu

W. Wanalertlak
e-mail: wanalewe@eecs.orst.edu

C. Yu
Department of Electrical and Computer Engineering,
Cleveland State University, Cleveland, OH 44115, USA
e-mail: c.yu91@csuohio.edu

M. Kim
Korea Advanced Institute of Science and Technology,
335 Gwahangno, Yuseong-gu, Daejeon 305-701, Korea
e-mail: mckim@cs.kaist.ac.kr

S.-M. Park · W.-T. Kim
Embedded Software Research Division, Electronics and
Telecommunications Research Institute, 138 Gajeongno,
Yuseong-gu, Daejeon 305-700, Korea
e-mail: minpar@etri.re.kr

W.-T. Kim
e-mail: kwtkim@etri.re.kr

1 Introduction

Wireless LANs (WLANs) based on the IEEE 802.11 standard have become pervasive in our society. WLANs offer high data transfer rate and are low-cost and easily deployable. These characteristics allow mobile users with portable devices to not only connect to the Internet but also use various services such as Voice over IP (VoIP) and real-time multimedia data transmission, *e.g.*, streaming audio and video. Until recently, deployment and coverage area of WLANs have been limited to variety of isolated hotspots, such as coffee shops, buildings, airport terminals, etc. But now *Municipal Wireless Networks* [1–4] based on *Wireless Mesh Networks* (WMNs) [5] allow the coverage area to be extended across a larger geographical area, *e.g.*, a city. Moreover, WLANs will play a crucial role as backbone networks for emerging pervasive computing technologies, such as *wearable computers*, *Wireless Body Area Networks* (WBANs) [6], *Wireless Personal Area Networks* (WPANs), and *Vehicular Ad hoc Networks* (VANETs) [7]. Nevertheless, achieving seamless, mobile access is a major

challenge because of large handoff delay incurred scanning for available *access points* (APs) when a *mobile station* (MS) switches connection from one AP, or cell, to another. Recent studies found that passively scanning for APs during a handoff can be as much as a second [8] and actively scanning for APs requires $350 \sim 500$ ms [8]. This becomes a major concern for mobile multimedia applications such as VoIP where the end-to-end delay is recommended to be not greater than 50 ms [9].

The key to reducing the scanning overhead and thus the handoff delay is to predict the next point-of-attachment of a MS, *i.e.*, *mobility prediction*. This way a MS can directly authenticate/re-associate with the target AP thereby eliminating the scanning overhead. Numerous efforts have already been made to reduce the scanning overhead in WLANs. These include methods to employ extra hardware, either in the form of additional radios [10] or an overlay sensor network [11], to detect APs, selectively scan channels based on the topological placement of APs [12], and predict the next point-of-attachment based on signal strength [13].

Unfortunately, these techniques neither provide next cell predictions that can eliminate the need to scan for APs nor consider mobility patterns of MSs, which are dictated by the structure of a building or a city block and the past behaviors of MSs. In our previous work [14], we proposed a location-based technique that considers the long-term history of handoff patterns. Although this technique is effective in performing next cell predictions for majority of MSs, it does not consider short-term and periodic behavior of mobile users that cause some MSs to behave differently from the norm and thus suffer mispredictions.

This paper presents a solution called the *Behavior-based Mobility Prediction* (*BMP*) technique that provides accurate mobility prediction by considering multifaceted user behavior based the following four factors: *location*, *group*, *time-of-day*, and *duration*. The location-based mobility prediction is achieved by maintaining the handoff history of all the MSs in the network, and then monitoring direction of movements of MSs relative to the topological placement of cells to predict their next point-of-attachment. In addition, next cell predictions are based on the frequencies of occurrences rather than signal strength. Therefore, it takes into consideration that mobility patterns are dictated by the structure of a building or a city block and the past behaviors of MSs. Moreover, the handoff frequencies are treated as time-series data, thus when next cell predictions fail future predictions are recalibrated based on different groups, where each group of MSs has similar mobility patterns, time-of-day, and duration characteristics.

The paper is organized as follows: Section 2 overviews the 802.11 scanning process. Section 3 presents the proposed BMP method and Sect. 4 evaluates its performance. Section 5 discussed the related work. Finally, Sect. 6 concludes the paper and discusses future work.

2 Background: 802.11 scanning process

In a WLAN, when a MS moves from one cell to another, its network interface senses the degradation of signal quality in the current channel. The signal quality continues to degrade as MS moves further away from the current AP, and a handoff to a new cell is initiated when the signal quality reaches a preset threshold [15]. This process starts with probing for new cells using either passive or active scanning. In *passive scanning*, a MS switches its transceiver to a new channel and waits for a beacon to be sent by the new AP, typically every 100 ms, or until the waiting time reaches a predefined maximum duration, which is longer than the beacon interval. The time a MS has to wait varies since beacons sent by APs are not synchronized. For these reasons, a recent study has shown that an MS can spend up to 1 s to search all possible channels [8], which results in unacceptable handoff delay.

In *active scanning* shown in Fig. 1, a MS broadcasts a probe request and waits for a response. If the MS receives a response from an AP, it assumes there may be other APs in the channel and waits for *MaxChannelTime*. Otherwise, the MS only waits for *MinChannelTime*. *MinChannelTime* is shorter than *MaxChannelTime* to keep the overall handoff delay low, but it should be long enough for MS to receive a possible response. A typical duration for scanning each channel is around 25 ms and $350 \sim 500$ ms for all 11 channels [8].
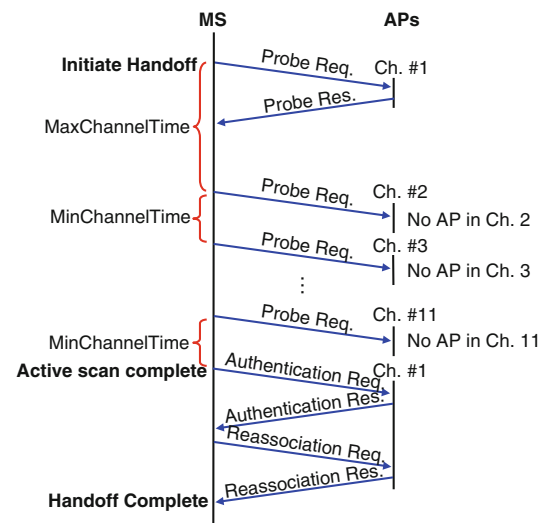

Fig. 1 The 802.11 active scanning process

After scanning, MS joins the cell with the strongest signal strength, which is done by performing authentication and (re)association. Authentication is the process that a MS uses to announce its identity to a new AP. In the IEEE 802.11 standard, authentication is performed using either open system or shared key. Open system authentication is the default method, and involves the MS sending authentication request frame to the AP containing its source address in the frame header and the type of authentication in the frame body. Then, the AP sends the authentication response frame back to the MS. This frame has the authentication result and the information that indicates the type of authentication.

The next step is (re)association, which allows the distribution system to keep track of the location of each MS so that frames destined for a MS can be forwarded to the correct AP. How (re)association requests are processed is implementation specific, but typically involves allocation of frame buffers and, in the case of re-association, communicating with the old AP so that any frames buffered at the old AP are transferred to the new AP. Finally, the last step involves rerouting the network traffic by updating the Forwarding Table in the switch that connects both the old and new APs.

3 The behavior-based mobility prediction scheme

The basic idea behind BMP is to improve the prediction accuracy by exploiting the behavior of mobile users. This is achieved by identifying temporal segments of handoff history that characterize certain user behaviors, and performing handoff predictions that pertain to those periods.

3.1 User behavior

Behavior of mobile users can be characterized in many different ways. In the proposed method, the four characteristics that define user behavior are location, group, time-of-day, and duration. The following discusses the motivation for using these characteristics.

The *location* factor, as discussed in Sect. 1, represents the history of mobility patterns that can be either static or dynamic. Static mobility patterns are dictated by fixed structures, such as roads, building structures, and city blocks. On the other hand, dynamic mobility patterns are caused by user behavior as well as frequent and drastic changes in the operating environment of wireless networks due to multipath effects and electromagnetic interference.

The *group* factor reflects the fact that MSs often behave as groups. For example, MSs in an academic setting can be categorized as students, graduate students, and faculty/ staff, and the mobility behaviors of these groups are very different. Moreover, MSs can be associated with specific events that are derived from user habits, *e.g.*, departmental staff spends most of the day near the administrative offices while students congregate in the atrium, classrooms, and computer labs. The group factor can be statically applied during the network registration phase. For example, users in a typical campus network are registered with Unix accounts that are grouped based on their status. Users in large community networks are also registered based on different types of memberships, such as residential, business, free subscriber, etc. In addition to these pre-assigned groups, other groups can be dynamically formed from a set of MSs that suffer from high misprediction rate.

The *time-of-day* factor indicates that user behaviors change as function of time. For example, mobility patterns observed in an academic setting will change during the course of a day depending on the schedule of classes. There will be bursts of repetitive mobility patterns when students move between classes, and mobility behaviors during the evening will be different from the daytime. Most of MSs in an academic network in the evening are graduate students and, for most part, they tend to move only within limited areas, *i.e.*, graduate student offices, laboratories, and hallways. Similarly, most users in a community network in the evening tend to stay within residential areas. In addition, both environments typically exhibit periodic behaviors such as students attending classes and workers commuting.

The *duration* factor, which can be categorized as short, medium, and long, directly represents how long a MS is connected to a cell and indirectly represents the speed at which it moves through a cell. Figure 2(a) shows an example of a *medium* duration, which represents the default case when a MS transits through a cell. For example, the cells c_x and c_z in Fig. 2(a) do not overlap, thus the MS has to transit through c_y before it can reach the destination c_z. This is also the case when c_x and c_z overlap but a direct pathway does not exist between the two cells. In WLANs, most handoffs occur with medium duration because the traveling distance of a MS is typically a lot further than coverage area of a cell. Figure 2(a) also illustrates a *long* duration that represents a MS performing some activity at the destination cell c_y. For example, in an academic setting, MS performs a handoff to a cell that covers a classroom and then spends a long period of time listening to a lecture. Similarly, in a community network, a MS performs a handoff in the morning to a cell that covers an office building and stays connected to this cell for many hours.

A *short* duration often represents an unnecessary or false handoff. For instance, consider a MS moving across three adjacent cells c_x, c_y, and c_z in Fig. 2(b). As the MS moves from c_x to c_y, and then to c_z, if the connection duration for c_y is very short then it indicates the three adjacent cells are highly overlapped. This is important because if the degree
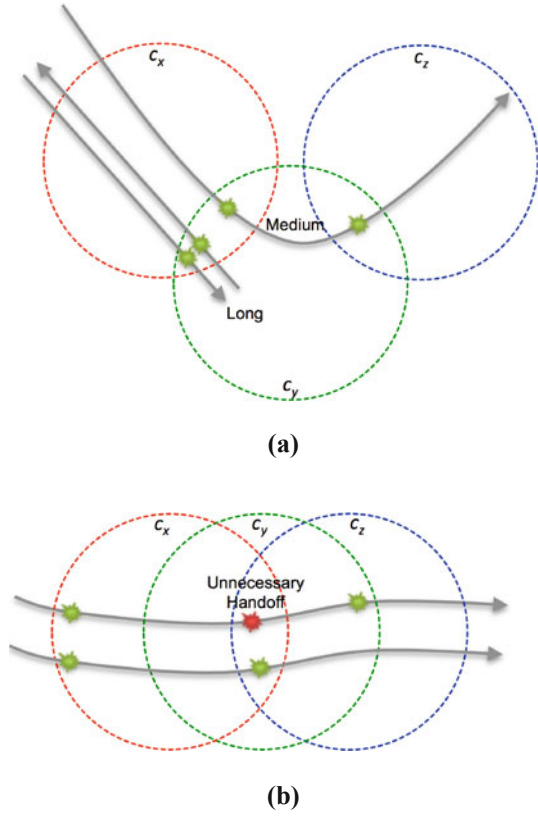
Fig. 2 Long, medium, and short durations. **a** Medium and long durations. **b** Short duration

Fig. 3 A possible BMP configuration in a WLAN environment

Fig. 4 The BMP scheme

of overlapping is sufficiently high then the prediction for the handoff sequence $c_x \rightarrow c_y \rightarrow c_z$ can be replaced by $c_x \rightarrow c_z$, which allows MS to move from c_x directly to c_z, eliminating an extra handoff. Therefore, short duration together with other behavior factors can be used to identify and eliminate these unnecessary handoffs.

Based on our analysis of the simulated networks (see Sect. 4), a typical short duration is few seconds while a long duration lasts more than few minutes. All others are considered as medium duration.

3.2 The proposed method

Figure 3 illustrates a possible configuration for the proposed method, where the BMP scheme is assumed to be implemented by a server and is collocated with the authentication server. Moreover, the next-cell prediction c_w for the current handoff from c_x to c_w is assumed to be available from the last handoff from c_y to c_x. Therefore, a MS directly authenticates/re-associates with c_w eliminating the need to perform scanning. A full scan is performed when an MS joins the network for the first time.

When the MS requires a handoff from c_x to c_w, it sends $pred\_req\langle ID, HS, D\rangle$, where ID is the ID of the MS, HS is the handoff sequence for the MS, and D is the duration of
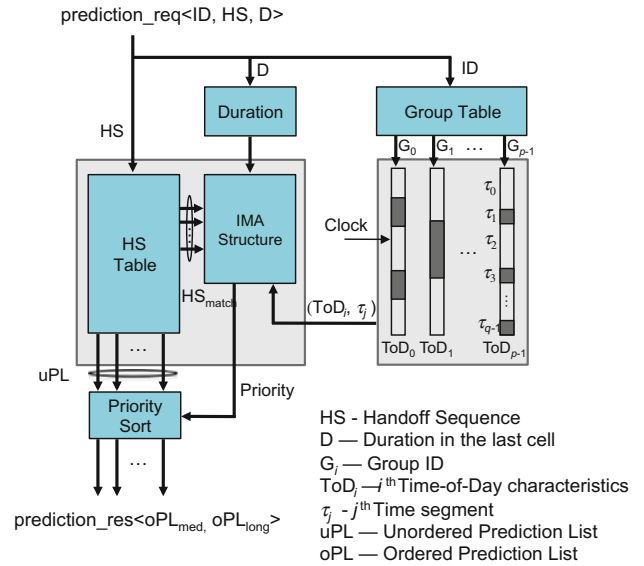
time spent by the MS in the last cell c_x, to the server as a part of the authentication request via c_w to obtain predictions for the next handoff (Step 1).

The server then performs *Next-cell Prediction* (Step 2), which involves the following operations illustrated in Fig. 4. First, the last k handoffs in the handoff sequence, i.e., $c_x \rightarrow c_w$, are used to search the *HS Table* for a set of matching entries representing an *unordered* next-cell *Prediction List* (*uPL*). At the same time, *ID* is used to index the *Group Table* to obtain the group ID G_i in $\{G_0, G_1, \ldots, G_{p-1}\}$, where p represents the number of groups. Second, G_i is used to select a particular group's Time-of-Day characteristic ToD_i. As the name suggests, *ToD*s model time-of-day characteristics of different groups

of mobile users. The default period for a ToD_i is T, e.g., $T = 1$ day, which means MSs that belong to this group exhibit similar handoff behavior during the entire period T. Moreover, a group of MSs that exhibit short-term and periodic handoff behaviors can be separately modeled as $ToD_i = \{\tau_0, \tau_1, \ldots, \tau_{q-1}\}$, where τ_j represents a time segment, q represents the total number of time segments, and $\sum_j \tau_j = T$. Thus, the current time (i.e., clock) determines τ_j for a particular group G_i, i.e., (ToD_i, τ_j). The importance of (ToD_i, τ_j) is that it uniquely defines a set *Integrated Moving Averages* (IMAs) to represent the frequencies of handoff sequences for a particular group of MSs with specific time-of-day and duration characteristics. Therefore, (ToD_i, τ_j) together with HS_{match} determine the proper subset of IMAs to be used in applying the priority order for the next-cell predictions, i.e., *ordered Prediction Lists*, for both *medium* (oPL_{med}) and *long* (oPL_{long}) durations.

Once the priority is determined, the BMP server returns $pred\_resp\langle oPL_{\text{med}}, oPL_{\text{long}}\rangle$ as a part of the authentication response (Step 3). After receiving the authentication response, the MS reassociates with the AP in cell c_w (Step 4). This is followed by *Ethernet Switch Update* (Step 5) and *Prediction Update* operations (step 6), which complete the handoff.

During the next handoff, the 1st prediction in either oPL_{med} or oPL_{long} is used as the next cell prediction based on whether the duration of the MS in the current cell c_w is medium or long, respectively. The next-cell prediction defaults to oPL_{med} if the MS experiences a short duration in the current cell. If the 1st prediction fails, i.e., authentication/re-association fails, the 2nd prediction is used, and so on. A full scan is performed when all the predictions in $oPL_{\text{med}}/oPL_{\text{long}}$ fails,

The following discusses the detailed operations of BMP.

3.2.1 HS Table

The mobility history of a MS is represented by a handoff sequence of length l denoted as $(c_{n-l-1}, \ldots, c_n)$, where c_i

HS Table

Past	Current	Next
c_x	c_w	c_x
c_x	c_w	c_y
c_x	c_w	c_z
...	...	...
c_y	c_x	c_w

HS_{match}

Unorder Prediction List (uPL)

Fig. 5 Handoff Sequence Table. This table shows a possible entries for the example shown in Fig. 3, where the last $k = l - 1$ visited cells (i.e., $c_x \rightarrow c_w$) are used to search for the next cell predictions

indicates the cell ID of ith visited cell. Figure 5 shows the HS Table, which is a collection of unique handoff sequences that represents the global history of mobility patterns in the network. The mobility prediction based on the HS Table can be represented by an order-k Markov process

$$\hat{P}(X_{n+1} = c_{n+1}|X(n-k+1, n) = (c_{n-k+1}, \ldots, c_n))$$
$$= \frac{N((c_{n-k+1}, \ldots, c_n, c_{n+1}), L)}{N((c_{n-k+1}, \ldots, c_n), L)}, \quad (1)$$

where $k = l - 1$, L is the overall observed history of mobility patterns, and N is the average frequency of mobility patterns. Our simulation study in Sect. 4 shows that $l = 3$ provides accurate next-cell predictions and yet the amount of mobility history needed is minimal.

3.2.2 IMA

IMAs are typically used in forecasting time-series data and can be derived from the more general *AutoRegressive Integrated Moving Average* (ARIMA) model. ARIMA is defined as ARIMA (p, d, q), where p, d, and q refer to the order of the autoregressive, the differencing, and the moving average parts of the model, respectively. *Exponential Weighted Moving Average* (EWMA) is equivalent to ARIMA $(0, 1, 1)$ [16, 17] and is much simpler to formulate than the general ARIMA model. The predicted frequency of each HS for the next period $t + 1$, $\overline{z}_{t+1}$, can be defined by the following EWMA:

$$\overline{z}_{t+1} = (1 - \lambda)\overline{z}_t + \lambda z_t, \quad (2)$$

where z_t is the frequency of the handoff sequence during the current period t, $\overline{z}_t$ is the predicted frequency of the handoff sequence in the current period t, and λ is the smoothing factor $0 < \lambda < 1$. t represents the minimum time interval for the time series data (i.e., $t = 1$ min). The parameter λ determines the characteristic of the EWMA model and is typically chosen experimentally. Based on our analysis of the simulated networks (see Sect. 4), λ for the time-series data representing the frequency of handoff sequences is chosen to be 0.1. Although EWMA does not rely on a full statistical analysis, this simple model gives results that are relatively close to ones from ARIMA.

The IMA structure is shown in Fig. 6, which is essentially a three-dimensional structure. HS_{match} and ToD_i are used to index the first two dimensions of the IMA structure. This results in a set of EWMAs, i.e., $\overline{z}_{t+1}$, which corresponds to a particular group's (G_i) time-of-day characteristics (ToD_i). Each element in the last dimension is indexed by τ_j corresponding to a particular time segment containing a set EWMAs $\langle \overline{z}_{t+1}^{\text{med}}, \overline{z}_{t+1}^{\text{long}}\rangle_c$, where $\overline{z}_{t+1}^{\text{med}}$ and $\overline{z}_{t+1}^{\text{long}}$ represent EWMAs for medium and long durations, respectively, and c represents the predicted cell. Then, uPL from the HS
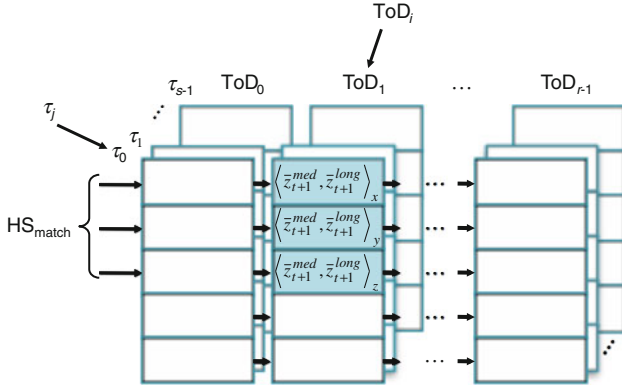
Fig. 6 The IMA structure

Table is prioritized based on these EWMAs to generate oPL_{med} and oPL_{long}.

3.2.3 Dynamic group formation

Separate IMAs are maintained for different user groups. Groups that are formed statically are referred to as *parent* groups. *Child* groups can be dynamically formed from a parent group by applying heuristics to a set of MSs that suffer from a high misprediction rate, which can be cause by differences or changes in mobility behaviors within the group. In order to gather the necessary information, each MS keeps track of and submits its average 1st prediction accuracy, μ, to the BMP server at the end of each period T. If the server detects a set of MSs in a group G_i with their average 1st prediction accuracies lower than the threshold $\rho_i = \mu_i - 0.5\sigma_i$, where μ_i and σ_i represent the average and standard deviation of the 1st prediction accuracy for the group G_i during the last period T, a new child group G_i' is formed with its own ToD_i'. This will improve the prediction accuracy for those set of MSs. However, if a set of MSs within this newly formed child group still suffer from high misprediction rate, *i.e.*, their average 1st prediction accuracies are lower than ρ_i of the parent group, then these MSs rejoin their parent group. After rejoining their parent group, if these MSs as well as other MSs continue to suffer from high misprediction rate, they will form their own child group. This process repeats every period T. In order to reduce the processing load, the number of MSs in a group must be more than some threshold M based on the traffic characteristics of the network and capability of the server.

3.2.4 Time-of-day characteristic

The time-of-day factor is applied when the prediction accuracy of a group is lower than some threshold for a period of time. This is done by associating a separate time segment τ_j with its own IMA. The motivation for generating time-segments is to isolate periods when mobility patterns can be better identified. The BMP server does this by keeping track of the prediction accuracy of each group during a period of τ_{min} (*e.g.*, 1 h), which is the minimum time-segment length. If the prediction accuracy of a group G_i is lower than ρ_i during the last τ_{min}, the server assigns a new IMA to the time-segment. Again, the value of τ_{min} is based on the network traffic and server capability.

3.2.5 Update operations

Different update operations are performed depending on different time periods.

The following operation is performed after each handoff:

- *HS Table Update*—A new HS Table entry is allocated and updated when a new handoff sequence is encountered.
- The following operation is performed after each period t:
- *IMA Structure Update*—z_t is determined based on the number of times handoff sequences of MSs match with an entry in the HS Table during the current period t. If there are no matches, then z_t is 0. Then, $\bar{z}_{t+1}$ is calculated for the next period $t + 1$ using Eq. 2. Note that z_t is not maintained for a handoff sequence with short durations. This way, any false or unnecessary handoffs are eventually removed from the behavior model.
- The following operation is performed after each period T:
- *Group Table Update*—A MS is always registered to a parent (*i.e.*, static) group. A MS can also be registered to a child group, and unregistered from the child group when they rejoin the parent group. Thus, a MS registered to a child group will also be registered to a parent group. Once a child group is formed, the priority is given to the child group for determining their time-of-day characteristics.
- The following operation is performed after each period τ_{min}:
- *ToD Update*—When IMA values of two consecutive time segments are close together (*e.g.*, closer than 10%) they are combined.

3.3 A distributed implementation

Note that the discussion of BMP thus far has been based on a centralized scheme. However, BMP can also be implemented using a distributed scheme where each AP maintains its own portion of the global mobility history shown in Fig. 4. This can be achieved by relaying next-cell

prediction requests using re-association requests. For example, consider again the handoff between c_x and c_w shown in Fig. 3. After authentication, MS sends reassociation request containing $pred\_req\langle ID, HS, D\rangle$ to c_w. Then, c_w performs next-cell prediction and returns $pred\_resp\langle oPL_{med}, oPL_{long}\rangle$ as a part of the reassociation response. Finally, c_w sends the handoff sequence as a part of either *Inter-Access Point Protocol* (IAPP) [18] or a vendor specific protocol to c_x. After AP_x receives the handoff sequence, update operations are performed. This allows MSs handing off from c_x to obtain the next-cell prediction list.

4 Performance evaluation

This section presents the performance evaluation of the BMP scheme and compares its accuracy and delay results against using only location information, such as our prior work on *Global Path Cache* (GPC) [14], which has been shown to be superior to *Neighbor Graph* [12] and *Selective Scanning with Caching* [13].

4.1 Simulation environment

The two network topologies used in the simulation study are shown in Figure 7, which are the first floor of the four-story, 153,000-ft^2 Kelley Engineering Center (KEC) at Oregon State University, and a public WLAN service that covers 2.5-mile^2 area of Portland, Oregon, called MetroFi® [1]. The APs in KEC are connected by an Ethernet switch, while APs in the MetroFi® network are interconnected by a WMN [5].

The simulated coverage area for KEC contains 6 APs and 450 MSs, and the paths taken by MSs are limited to hallways and the atrium. There are three groups of users, i.e., Students, Graduate Students, and Faculty/Staff, with each having a different type of mobility behavior. For example, Students mostly move between the atrium, the cafe, and the computer lab. In addition, Students move in and out of the classrooms during the last ten minutes of each class hour between 8 AM and 6 PM. In contrast, Graduate Students mainly move between their offices, the atrium, and the computer lab. Finally, Faculty/Staff moves mainly between their offices and the atrium. All MSs move at a pedestrian speed of 1.5 m/s.

The coverage area for Portland (indicated by the dotted line) contains 40 APs and 4,500 MSs and the paths taken by MSs are limited to sidewalks. The results for Portland were generated based on nine different groups of users. Nomadic represents a group of MSs that can move anywhere, any time within the simulated area. The next four groups represent Commuters who work in each of the four
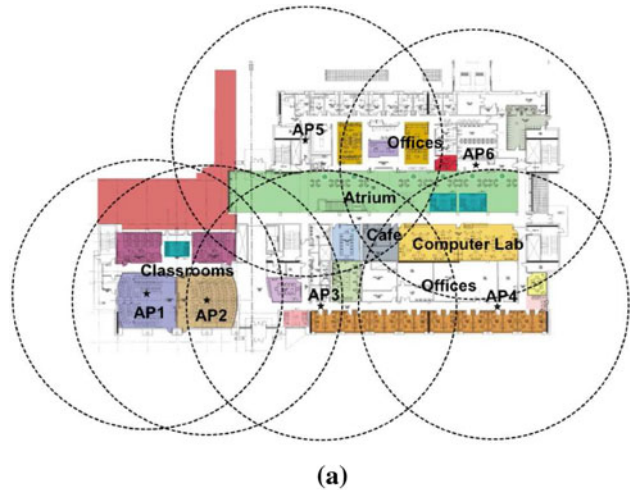

(a)

(b)

Fig. 7 Simulated WLAN coverage areas. **a** Kelley Engineering Center building. **b** Public WLAN in Portland, Oregon (MetroFi®)

quadrants or regions, i.e., C-I, C-II, C-III, and C-IV in Fig. 7(b), and are likely to travel long distances (i.e., 15–20 blocks) to work. Moreover, these groups of MSs only move between 6 AM to 10 AM and 6 PM to 10 PM. The last four groups represent Residents who live in each of the four regions, i.e., R-I, R-II, R-III, and R-IV in Fig. 7(b). These groups of MSs tend to stay at their homes most of time (several hours) but when they move they are likely to only move within few blocks (5–10 blocks) from their homes.

In order to accurately simulate mobility patterns and handoffs in large networks, we developed a simulator that implements a WLAN radio model, generates realistic mobility patterns based on building and city layouts, and supports management frames needed to implement scanning, authentication, and re-association. The structure of the simulator is shown in Fig. 8, which consists of Path
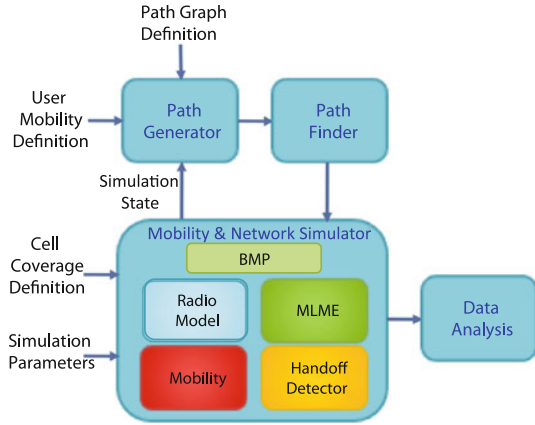
Fig. 8 Simulation model

Fig. 9 Path graphs for Fig. 7. **a** KEC. **b** Portland

Generator, Path Finder, Mobility & Network Simulator, and Data Analysis modules.

The *Path Generator* generates a new destination way-point based on the Path Graph definition, User Mobility definition, and the current Simulation State for each MS that has completed its trip between the original and destination waypoints. The *Path Graph* definition is a graphical representation of all the possible paths MSs can traverse within a simulated area, which is similar to the ones proposed in [19, 20]. Figure 9 shows the path graphs for the two network topologies used in the simulation study—KEC and Portland in Fig. 7. They consist of vertices representing waypoints and segments representing paths between adjacent waypoints. The *User Mobility* definition defines the number of MSs and APs as well how MSs move, including when and where a MS moves to and how long it stays at a waypoint. The *Simulation State* defines the current time and the locations of all the MSs in the simulated network area. The Path Generator uses these two definitions together with the current state of the simulator to randomly select destination waypoints based on a *modified* Random Waypoint model [21, 22]. Our modified Random Waypoint model allows a probability distribution to be assigned to sub-areas or regions within a path graph based on different groups of MSs at different times. The *Path Finder* module then uses the path-finder algorithm [23] to generate the shortest path between the source and destination waypoints. The resulting path consists of multiple segments, which are then fed to the *Mobility & Network Simulator*.

The Mobility & Network Simulator consists of Mobility module, Radio Model, Handoff Detector, and MAC sub-Layer Management Entity (MLME). The *Mobility* module simulates the movements of MSs on the segments at one meter resolution. The *Handoff Detector* monitors each MS's movement, and based on the *Cell Coverage* defi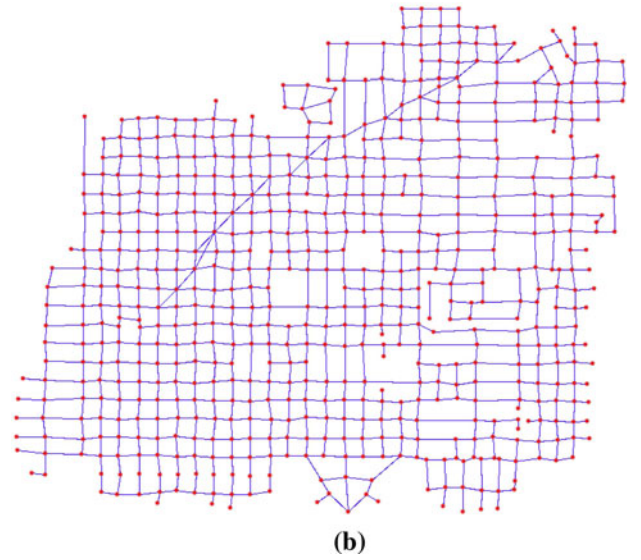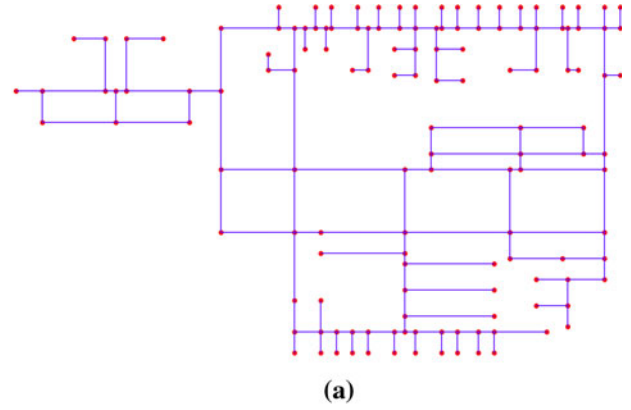nition and the *Radio Model*, which is based on log-distance path loss model [24], performs a handoff when the distance between a MS and its associated AP reaches the maximum radius of the coverage area. Handoffs are performed using the *MLME* module, which supports beacons, probing, authentication, and re-association. Finally, the *Data Analysis* module records the number of channel switches, the number of times MS has to wait for MaxChannelTime, MinChannelTime, Authentication delay/timeout, and Re-association delay (see Sect. 4.2).

Table 1 Delay parameters used in the simulation

Parameters	Delay
Channel switching time (t_{switch})	11.4 ms
MinChannelTime (t_{min})	20 ms
MaxChannelTime (t_{max})	200 ms
Authentication delay (t_{auth})	6 ms
Re-association delay ($t_{reassoc}$)	4 ms

4.2 Simulation delay parameters

The delay parameters used in the simulation are shown in Table 1: *Channel Switching Time* (t_{switch}) is the time required to switch from one channel to another; *Min-ChannelTime* (t_{min}) is the minimum amount of time a MS has to wait on an empty channel; *MaxChannelTime* (t_{max}) is the maximum amount of time a MS has to wait to collect all the probe responses, which is used when a response is received within *MinChannelTime*; *Authentication delay/timeout* (t_{auth}) is the time required to perform authentication based on MAC addresses; and *Re-association delay* (t_{assoc}) is the time required to perform re-association.

These delay values represents the current off-the-shelf NICs and obtained with an experimental setup that consisted of two laptops with 802.11a/b/g NICs based on Atheros AR 5002X chipset [25] running Linux 2.6 with MadWiFi driver [26], a Sun SPARC Server with Ethernet LAN NIC, and an HP ProCurve Wireless Access Point 420. Measurements were gathered by having one laptop serve as a MS that performs a channel switch and continuously sends UDP traffic to a daemon running on the server while the second laptop sniffs the traffic using tcpdump (see [14]).

4.3 Simulation results

All the results in this section were obtained based on the assumption that the HS Table contains a complete history of handoff patterns. This is achieved by first running the simulations for 10^4 handoffs for KEC and 10^6 handoffs for Portland to allow the BMP system to "learn" all the handoff sequences [14], and then gathering statistics for up to 10^7 handoffs.

Figures 10 and 11 show the number of required HS Table entries and the average number of next-cell predictions returned per handoff, respectively, as function of l (KEC on the left, Portland on the right). As can be seen
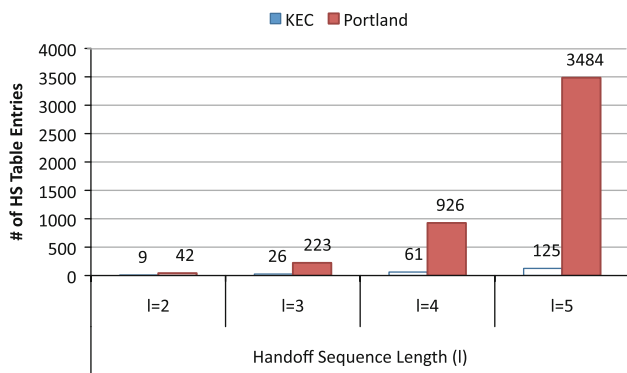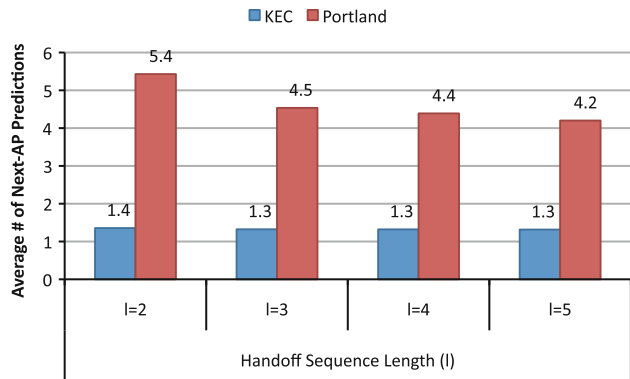

Fig. 11 Average number of next-cell predictions returned per handoff

from Fig. 10, more entries are required to keep track of MSs' more complex moving paths as l increases. Moreover, the number of entries needed for Portland is higher and grows significantly faster than KEC due to larger number of APs. Similarly, Portland has, on average, a larger set of next-cell predictions due to a larger number of highly overlapped regions. Note that, for implementation purposes, each entry in the HS Table contains multiple next-cell predictions instead of one prediction entry as illustrated in Fig. 5. Therefore, Figure 10 represents the number of unique handoff sequences of length $l-1$ observed in the two networks. The product of the number of entries in Fig. 10 and the average number of predictions returned per handoff in Fig. 11 represents the total number of unique handoff sequences of length l.

Figure 12 shows the accuracies of individual next-cell predictions based only on location information (starting with the 1st prediction result on the bottom), which are prioritized based only on their frequencies of occurrence. The overall accuracy for both KEC and Portland is 100%, which means one of the next-cell predictions is guaranteed to be correct. However, the accuracy of each prediction is more important because a misprediction incrementally increases the average handoff delay by $t_{switch}+t_{auth}$. For KEC, the 1st next-cell prediction accuracy is 68% and increase slightly as function of l. The 1st next-cell predictions that fail are satisfied by the 2nd next-cell predictions, which represent accuracy rate of $28.5 \sim 25.1\%$ relative to all the next-cell predictions. Similarly, the 3rd next-cell predictions that succeed make up $3.5 \sim 3.1\%$ relative to all the predictions. For Portland, the 1st next-cell predictions accuracy start at 43% and increases as function of l. The 2nd, 3rd, and 4th next-cell prediction accuracies are $28.7 \sim 22.6\%$, $17.1 \sim 9\%$, and $7.9 \sim 3\%$, respectively, relative to all the next-cell predictions. The combined prediction accuracy beyond the 4th next-cell prediction is $3 \sim 1.43\%$.
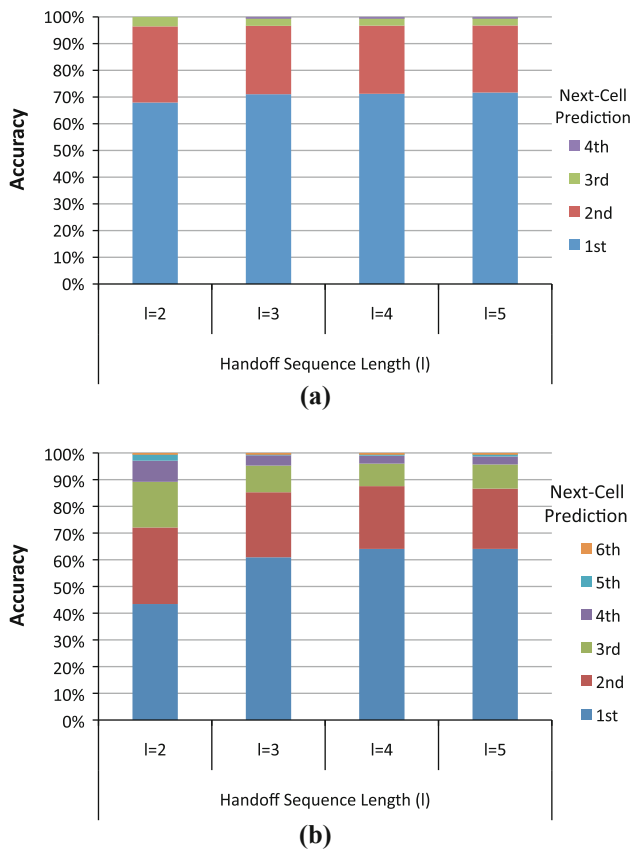

Fig. 10 Number of entries in HS Table

Fig. 12 Overall and individual prediction accuracies based only on location information. **a** KEC. **b** Portland

Fig. 13 Improvement in 1st prediction accuracy for BMP. **a** KEC. **b** Portland

Based on Figs. 10–12, it is clear that suggested size of l is 3, which provides good 1st prediction accuracy and yet the number of entries required for the HS Table is minimal.

Figure 13 shows the 1st prediction accuracy of BMP, which is importance because each misprediction from the oPL incrementally increases the average handoff delay by $t_{switch} + t_{auth}$.

The average improvement for BMP compared to using only location information for KEC is 23%. The largest improvement comes from Static Group at 12.7%, followed by Dynamic Group, ToD, and Duration with 7, 1.7, and 1.7%, respectively. These improvements also vary for different user groups. For example, BMP improves the 1st prediction accuracy for all three groups in KEC. The largest improvement of 77% comes from Students because their behaviors are dictated by schedule of classes, which causes their handoffs to be bursty and periodic and also results in longer durations. This is followed by Graduate Students with an overall improvement of 61%, which is the result of being able to better capture the fact that they work late at night and on weekends. In contrast, improvement for Faculty/Staff was less at 9.5% because the number of MSs in this group is much less than the other groups and the
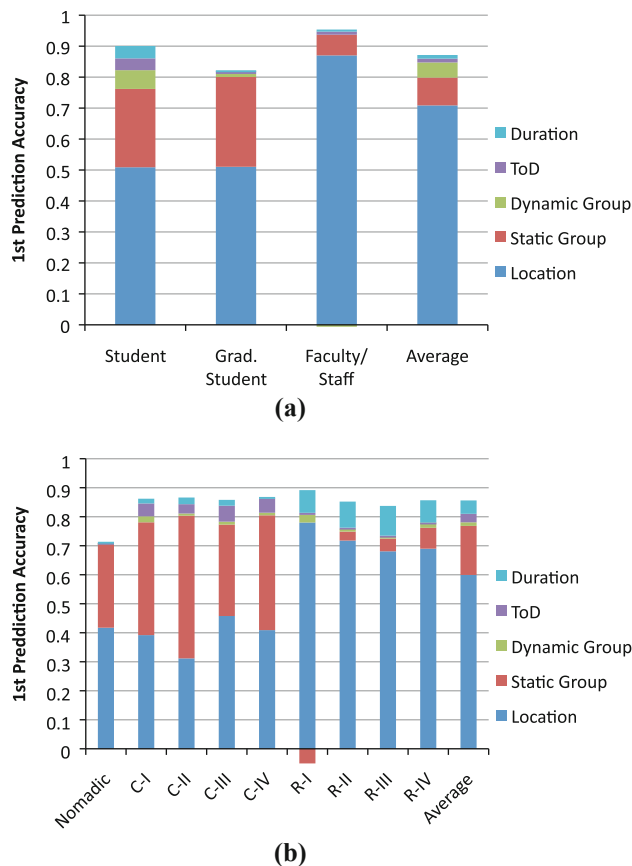
coverage area affected by their mobility is much more limited.

The average improvement for Portland is much more significant at 43%. The largest improvement came from Static Group at 28%, and this is followed by Duration at 7.7%, ToD at 4.9%, and Dynamic Group at 2%. Again, all of the user groups in Portland resulted in improvements. Among them, Nomadic and Commuter groups (C-I ~ IV) exhibit large improvements due to short-term surges in handoffs caused by groups of users commuting during rush hour. The Nomadic group is hardly affected by Dynamic Group, ToD, and Duration factors because their mobility behavior is uniformly random across the entire network and independent of time. In contrast, the largest improvement for Commuter groups came from the ToD characteristic. For Resident groups (R-I ~ IV), the Duration factor provided the most improvement.

Figure 14 compares the average handoff delays based on the parameter set defined in Table 1 (Location-based scheme on the left, and BMP on the right). These results show that BMP provides 10 and 21% improvement for KEC and Portland, respectively. This may appear to be only a small improvement compared to the location-only
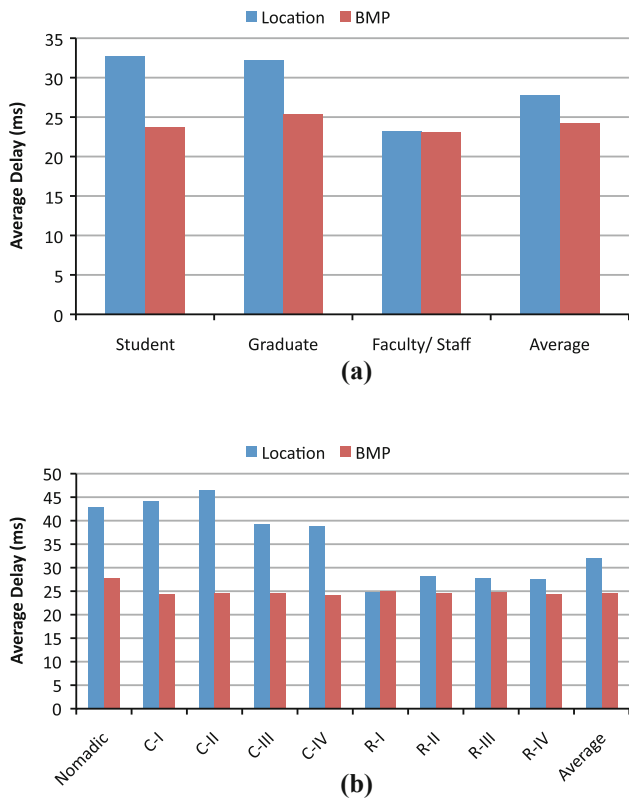
Fig. 14 Average handoff delay. a KEC. b Portland

based scheme, but the resulting delay for BMP is very close to the lower bound delay, which is 21.4 ms $= t_{switch} + t_{auth} + t_{assoc}$. More importantly, when individual handoff delays are considered, they resulted in significant improvements for some user groups. For example, the Student group in KEC resulted in 21.7% improvement, while Grad. Students had 16.5% improvement. This was also the case for Portland, where the Nomadic group resulted in 107.2% improvement. In addition, all groups resulted in similar average delay.

5 Related work

Many different mobility prediction techniques have been proposed for a variety of wireless networks, such as cellular [27–32], WLANs [10–13, 33, 34], ad hoc networks [35, 36], and mesh networks [5], and applied to reduce handoff latency [8, 12, 13, 37], provide efficient resource reservation [27–33], improve routing protocols [35], and conserve power [36]. However, these methods tend to be general and thus do not consider the special characteristics of WLANs, such as highly overlapped cell coverage, MAC contention, and variations in link quality.

Location-based schemes provide predictions using the current and past locations (*i.e.*, cells). These are all based on the order-*k* Markov predictor [34], and include *Global Path Cache* [14], *Selective Scan with Caching* [13], *Movement Model* [38], *Two-Tier prediction* [39, 40], and a technique based on the use of directional vectors [28]. However, they do not consider other characteristics of mobile users, such as group, time-of-day, and duration. Therefore, these techniques will not properly capture mobility patterns that deviated from the norm, such as behaviors exhibited by a small group of users that tend repeat in certain periods of time. There is a technique that applies different predictors to different time segments [38]. There are also techniques that reduce the number of location update operations in cellular networks by associating locations of individual users to different periods of time [41–43]. However, time is not the only factor that affects mobility patterns. Therefore, these techniques are unable to properly capture the behaviors exhibited by different groups.

Topology-based schemes define directed graphs that represent topological placements of APs and mobility patterns of MSs. These techniques are typically applied to WLANs and include *Neighbor Graph* [8, 12] and *Pre-Authentication path* [37]. Although these techniques reduce the number of channels to scan, they do not provide next-cell predictions.

Activity-based schemes provide next-cell predictions by relating locations to users' interests, such as schedules and activities. These include *Activity-based Mobility Prediction* [44], *ComMotion* [45], and GPS-based techniques [46, 47]. The mobility prediction is made based on the shortest path between the current and the predicted locations. However, these techniques may not be accurate since individual users may take a path based on their point-of-interests, such as ATM to withdraw money or simple idiosyncrasy of passing by a park, rather than the shortest path.

Data Mining-based schemes [27, 31, 32] reduce the signaling overhead during handoff and provide resource reservation to MSs in cellular networks by logging users' visited cells and time in a database. In addition, some techniques record geographic coordinates and directional movements of MSs from either GPS or triangulation of signal strengths [29, 30]. The basic idea is to provide predictions by searching the database using user contexts stored in MSs. If a match or multiple matches are found, predictions are provided based on location stored in the context. However, data-mining techniques require a large storage and fast processors to properly analyze long-term mobility behavior of users. In addition, most techniques typically require a GPS device to obtain information about locations and directions of MSs. For systems that rely on signal triangulation, their effectiveness may be limited due to the fact that mobile devices are mainly used for indoors and crowded outdoor areas where the signal strength is highly affected by noise rather than distance [24].

6 Conclusion

This paper presented the BMP technique to improve mobility prediction in WLANs. BMP models mobility patterns not only based on the location information of MSs but also captures their group, time-of-day, and duration characteristics. Therefore, it monitors the next-cell prediction accuracies and readjusts its prediction model based on the dynamic behavior of MSs. Our simulation study shows that BMP provides better predictions than location-only based schemes, and results in much lower handoff delay for all MSs.

For future work, we plan to investigate couple of issues. First, we plan to study the effectiveness of BMP for other wireless networks, such as cellular networks, WiMAX, and WMNs, not only to improve handoff latency but also to improve resource reservation, routing protocols, and energy efficiency. Second, we would like to investigate how BMP can be utilized to speed up vertical handoffs in heterogeneous networks.

References

1. MetroFi Portland Free Wi-Fi. Online. Available: http://www.metrofiportland.com.
2. SeattleWireless. Online. Available: http://SeattleWireless.net.
3. NYCwireless. Online. Available: http://NYCwireless.net.
4. Rooftop@Media. Online. Available: http://rooftops.media.mit.edu.
5. *Draft Standard for Information Technology— Telecommunications and Information Exchange Between Systems—LAN/MAN Specific Requirement - Part 11:Wireless LAN Medium Access Control and Physical Layer Specifications: Amendment: ESS Mesh Networking*, IEEE Unapproved draft Std. P802.11s/D1.02, Mar 2007.
6. Jovanov, E., Milenkovic, A., Otto, C., & de Groen, P. C. (2005). A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of Neuroengineering Rehabiltation (Online)*. Mar 2005.
7. Ashbrook, D., & Starner, T. (2006). Poster: Slope: A system for rapid deployment of vanet communication protocols. In *The international conference on mobile systems, applications, and services (MOBISYS)*. Jun 2006.
8. Ramani, I., & Savage, S. (2005). Syncscan: Practical fast handoff for 802.11 infrastructure networks. In *IEEE INFOCOM* (pp. 675–684). Mar 2005.
9. ITU-T recommendation G.114, (1993). International Telecommunication Union, Tech. Rep.
10. Brik, V., Mishra, A., & Banerjee, S. (2005). Eliminating handoff latencies in 802.11 WLANs using multiple radios: Applications, experience, and evaluation. In *Internet measurement conference (IMC)* (pp. 27–27). Oct 2005.
11. Waharte, S., Ritzenthaler, K., & Boutaba, R. (2004). Selective active scanning for fast handoff in WLAN using sensor networks. In *Mobile and wireless communications networks (MWCN)* (pp. 59–70). Oct 2004.
12. Shin, M., Mishra, A., & Arbaugh, W. A. (2004). Improving the latency of 802.11 hand-offs using neighbor graphs. In *The international conference on mobile systems, applications, and services (MOBISYS)* (pp. 70–83). Jun 2004.
13. Shin, S., Forte, A. G., Rawat, A. S., & Schulzrinne, H. (2004). Reducing mac layer handoff latency in IEEE 802.11 wireless LANs. In *ACM international workshop on mobility management and wireless access (MOBIWAC)* (pp. 19–26). Sep 2004.
14. Wanalertlak, W., & Lee, B. (2007). Global path-cache technique for fast handoffs in WLANs. In *International conference on computer communications and networks (ICCCN)* (pp. 45–50). Aug 2007.
15. *Local and Metropolitan Area Network, Part 11: Wireless LAN Medium Access Control and Physical Layer Specifications*, IEEE Std. 802.11, (2007).
16. Box, G. E. P., & Jenkins, G. (1994). *Time series analysis, forecasting and control*, 3rd ed. NJ: Prentice Hall.
17. Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications: With R examples*, 2nd ed. New York: Springer.
18. Ieee 802.11f standard: Recommend practice for multi-vendor access point interpretability via an inter-access point protocol. Online. Available: http://grouper.ieee.org/groups/802/11/private/Draft_Standards/11f/802.11f-D3.1.pdf.
19. Yoon, J., Noble, B. D., Liu, M., & Kim, M. (2006). Building realistic mobility models from coarse-grained traces. In *The international conference on mobile systems, applications, and services (MOBISYS)* (pp. 177–190). Jun 2006.
20. Umedu, T., Urabe, H., Tsukamoto, J., Sato, K., & Higashino, T. H. T. (2006). A manet protocol for information gathering from disaster victims. In *Fourth annual IEEE international conference on pervasive computing and communications workshops* (pp. 442–447). Mar 2006.
21. Broch, J., Maltz, D. A., Johnson, D. B., Chun Hu, Y., & Jetcheva, J. (1998). A performance comparison of multi-hop wireless ad hoc network routing protocols. In *ACM international conference on mobile computing and networking (MobiCom)* (pp. 85–97). Oct. 1998.
22. Network simulator (ns2). Online. Available: http://www.isi.edu/nsnam/ns.
23. Amit's thoughts on path-finding and A*. Online. Available: http://theory.stanford.edu/amitp/GameProgramming.
24. Rappaport, T. S. (2002). *Wireless communications: Principles and practice*, 2nd ed. NJ: Prentice Hall.
25. Atheros ar5002x 802.11a/b/g universal WLAN solution. Online. Available: http://www.atheros.com/pt/AR5002XBulletin.htm.
26. MadWIFI_0.9.2. Online. Available: http://www.madwifi.org.
27. Katsaros, D., Nanopoulos, A., Karakaya, M., Yavas, G., Ulusoy, U., & Manolopoulos, Y. (2003). *Clustering mobile trajectories for resource allocation in mobile environments*, ser. Lecture notes in computer science, vol. 2779/2003. New York: Springer. Sep. 2003.
28. Aljadhai, A., & Znati, T. F. (2001). Predictive mobility support for QoS provisioning in mobile wireless environments. *IEEE Journal on Selected Areas in Communications, 19*(10), 1915–1930.
29. Kim, T.-H., Yang, Q., Lee, J.-H., Park, S.-G., & Shin, Y.-S. (2007). A mobility management technique with simple handover prediction for 3G LTE systems. In *Vehicular technology conference (VTC)* (pp. 259–263). Jun 2007.
30. Soh, W.-S., & Kim, H. S. (2004). Dynamic bandwidth reservation in cellular networks using road topology based mobility predictions. *IEEE INFOCOM, 4*, 2766–2777.

31. Wu, H.-K., Jin, M.-H., Horng, J.-T., & Ke, C.-Y. (2001). Personal paging area design based on mobile's moving behaviors. *IEEE INFOCOM, 1*, 21–23.

32. Yavas, G., Katsaros, D., Ulusoy, O., & Manolopoulos, Y. (2005). A data mining approach for location prediction in mobile environments. *Data and Knowledge Engineering, 54*(2), 121–146.

33. Song, L., Deshpande, U., Kozat, U. C., Kotz, D., & Jain, R. (2006). Predictability of WLAN mobility and its effects on bandwidth provisioning. In *IEEE INFOCOM* (pp. 1–13). Apr 2006.

34. Franois, J.-M. (2007). Performing and making use of mobility prediction. Ph.D. dissertation, University of Lige.

35. Su, W., Lee, S.-J., & Gerla, M. (2001). Mobility prediction and routing in ad hoc wireless networks. *International Journal of Network Management, 11*(1), 3–30.

36. You, C.-W., Chen, Y.-C., Chiang, J.-R., Huang, P., Chu, H.-H., & Lau, S.-Y. (2006). Sensor-enhanced mobility prediction for energy-efficient localization. *Sensor and Ad Hoc Communications and Networks (SECON), 2*, 565–574.

37. Pack, S., & Choi, Y. (2004). Fast handoff scheme based on mobility prediction in public wireless LAN systems. *IEE Proceedings Communications, 151*(5) 489–495.

38. Liu, T., Bahl, P., & Chlamtac, I. (1998). Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE Journal on Selected Areas in Communications, 16*(6), 922–936.

39. Chan, J., Zhou, S., & Seneviratne, A. (1997). A hybrid handoff scheme with prediction enhancement for wireless ATM network. In *IEEE Asia Pacific Conference on Communications (APCC)*, Dec. 1997, pp. 494–498.

40. Chan, J., Zhou, S., & Seneviratne, A. (1998). A QoS adaptive mobility prediction scheme for wireless networks. In *IEEE conference and exhibition global telecommunications conference (GLOBECOM)* (pp. 1414–1419). Nov 1998.

41. Ma, W., & Fang, Y. (2002). A new location management strategy based on user mobility pattern for wireless networks. In *IEEE conference on local computer networks (LCN)* (pp. 451–457). Nov 2002.

42. Tabbane, S. (1995). An alternative strategy for location tracking. *IEEE Journal on Selected Areas in Communications, 13*(5), 880–892.

43. Cayirci, E., & Akyildiz, I. F. (2002). User mobility pattern scheme for location update and paging in wireless systems. *IEEE Transactions on Mobile Computing, 1*(3), 236–247.

44. Sricharan, M., Vaidehi, V., & Arun, P. (2006). An activity based mobility prediction strategy for next generation wireless networks. In *Wireless and optical communications networks: The next generation of internet (WOCN)*. Apr 2006.

45. Marmasse, N., & Schmandt, C. (2002). A user-centered location model. *Personal Ubiquitous Computer, 6*(5–6), 318–321.

46. Ashbrook, D., & Starner, T. (2002). Learning significant locations and predicting user movement with GPS. In *IEEE international symposium on wearable computers (ISWC)* (pp. 101–108). Oct 2002.

47. Liao, L., Fox, D., & Kautz, H. (2007). Extracting places and activities from GPS traces using hierarchical conditional random fields. *International Journal of Robotics Research, 26*(1), 119–134.