Title

Access and Mobility of Wireless PDA Users

Authors
McNett, Marvin
Voelker, Geoffrey M

Publication Date
2004-02-09

Peer reviewed

Access and Mobility of Wireless PDA Users

Marvin McNett and Geoffrey M. Voelker
Department of Computer Science and Engineering
University of California, San Diego

Abstract

In this paper, we analyze the mobility patterns of users of wireless handheld PDAs in a campus wireless network using an 11 week trace of wireless network activity. Our study has three goals. First, we characterize the high-level mobility and access patterns of handheld PDA users and compare these characteristics to previous workload mobility studies focused on laptop users. Second, we develop two wireless network topology models for use in wireless mobility studies: an *evolutionary topology model* based on user proximity and a *campus waypoint model* that serves as a trace-based complement to the random waypoint model. Finally, we use our wireless network topology models as a case study to evaluate ad-hoc routing algorithms on the network topologies created by the access and mobility patterns of users of modern wireless PDAs.

1 Introduction

As wireless access proliferates, understanding user behavior and wireless network performance has become crucial as a basis for developing and evaluating new applications, such as context-aware applications, new network infrastructure, such as middleware support for public-area networks, and new wireless communication architectures, such as ad-hoc networking.

Over the past few years there have been a number of wireless workload studies characterizing user behavior and network performance in a variety of settings, including metropolitan networks [24], university campuses [17], conferences [5], and most recently corporate networks [6]. The goals of these studies have ranged from developing low-level radio signal and error models [11], network installation and maintenance issues [5], and characterizing user workload models, network performance, and mobility of laptop users [17].

In this paper, we extend previous wireless studies by characterizing the mobility patterns of users of wireless handheld PDAs in a campus wireless network, and evaluating the implications of these mobility patterns on new wireless communication architectures like ad-hoc networks. We use a trace of wireless network access by 275 freshmen with HP Jornada PDAs over the course of the Fall, 2002, term at our University. A key aspect of our trace is the focus on handheld PDA users. Even more so than laptop users, we expect handheld PDA users to exhibit high degrees of both casual and extended wireless access and mobility.

Our study has three goals. First, we characterize the high-level mobility and access patterns of handheld PDA users, and compare these characteristics to previous workload mobility studies focused on laptop users.

Second, we develop two wireless network topology models for use in wireless mobility studies. Previous work frequently uses synthetic models of user mobility patterns, such as the popular "random waypoint" model [9], to derive wireless network topologies that change due to user mobility. To complement these synthetic models, we derive two new models of network topologies that incorporate user mobility patterns from our traces. Our *evolutionary topology model* represents connectivity among users solely based on observed network proximity: an edge connects two nodes if two users can reasonably "hear" each other. In this model, the network topology evolves over time as nodes and edges appear and disappear based upon user connections, disconnections, and movements observed in our trace. Our *campus waypoint model* serves as a trace-based analog of the random waypoint model. In this model, we associate users with geographic locations on campus, and model their mobility vectors and potential interactions as they access the wireless network over time. However, rather than choosing user locations, speeds, and directions using random distributions, we instead use the access and mobility patterns of users in our trace.

Finally, we use the evolutionary network topology model as a case study to evaluate ad-hoc routing algorithms in a realistic setting. Ad-hoc routing has been a popular research topic for a while, and many protocols have been proposed and evaluated in the literature [21, 16, 20, 22]. Surprisingly, however, very few have been evaluated in realistic user settings [19]. As a result, we know little about the trade-offs and applicability of these algorithms to common, expected situations. Since groups of users with handheld PDAs have often been used as a motivating setting for ad-hoc networking, we use our trace to evaluate popular ad-hoc routing algorithms on the network topologies created by the access and mobility patterns of users of modern wireless PDAs.

The remainder of this paper is organized as follows. Section 2 discusses related work. Section 3 describes our user population and trace methodology. Section 4 characterizes the access and mobility patterns of the users in our trace. Section 5.1 describes the two mobility models we create based upon our trace, and Section 6 evaluates ad-hoc routing algorithms using one of those models. Finally, Section 7 summarizes our results and concludes.

2 Related Work

In this section, we first discuss previous trace based studies of wireless network user behavior. We then discuss various mobility models that have been proposed and studied in the literature.

2.1 Trace Studies

Three early studies on wireless network usage were performed at Stanford University [18, 24, 25]. In [18], the authors study a small set of 8 users for an 8-day period. In their trace, laptop users switch between wired and wireless networks. The wide area wireless network used in their study covered an entire metropolitan area and is intrinsically different from the 802.11b wireless network used in our study. Their focus was on how often users switched between the wired and wireless networks, how often users changed position geographically, latency in the wireless network, and optimizations for improving wireless network performance. The granularity of geographic location was limited by the large, half mile radius of the wireless radios used.

In [24], the same wireless network infrastructure was used as in [18]. However, a set of 24,773 users were observed in [24] over a 7 week period to study user mobility and network access properties. These goals are similar to our own, but were for a much wider area with a larger audience, using a very different network setup. Due to the dates of the trace period, we also assume that most of the users in the trace were laptop users.

The last Stanford study [25] focused on local area wireless network activity of 74 users in the Stanford Computer Science Department over a 12-week period. In that study, they analyzed data polled from access points using SNMP, subnet router tcpdump data, and an authentication log to characterize network access, network loads, and network traffic types.

Three subsequent studies followed the Stanford research. Balachandran et al. [5] supplement previous studies with a workload from a wireless LAN at an ACM conference. Their goal was to characterize user behavior in this setting to facilitate the planning and deployment of wireless networks. Since the network was confined to a large auditorium for the conference, the mobility characteristics of the users in the trace were very limited.

Kotz and Essien [17] recently extended the study in [25] to an entire campus wide wireless network at Dartmouth College. They collected data from 1706 users across 476 access points over an 11 week period. Their study focused on a more general population of wireless users. However, their trace data did not allow for them to distinguish among device types (PDA, laptop, etc.), nor did they focus on modeling mobility.

Most recently, Balazinska and Castro [6] analyzed a corporate wireless LAN workload of 1366 users across 177 access points. Again, data was collected by using SNMP to periodically poll all access points involved. In their study, Balazinska and Castro compare their results with results from the previous studies to distinguish user behavior among the various user groups and identify factors which contribute to these differences.

The recent studies of 802.11 wireless networks use AP associations to track user mobility. In our study, we refine this technique to use signal strength information for *all* access points detected by the device at a particular time.

2.2 Mobility Models

Numerous mobility models have been proposed for use in simulation and evaluation of ad-hoc routing protocols. A recent survey presents a number of these models which are divided into entity and group mobility models [10]. The survey also present simulation results from these models to emphasize the importance of properly choosing a mobility model for research simulations. They show that the choice of mobility model greatly impacts the performance of various ad-hoc routing protocols.

Perhaps the most widely used mobility model has been the random waypoint model [9]; this is in part due to its implementation in the popular Network Simulator, ns2 [1]. In this model, each node begins a simulation stationary for *pause time* seconds, then randomly chooses a destination in the simulation space and moves to that location with a speed between $speed_{min}$ and $speed_{max}$ (chosen from a uniform distribution). Once a node reaches its destination, the process repeats. A few studies have been performed on this mobility model to overcome shortcomings. Aside from being far from realistic user behavior in most settings [14], random waypoint creates non-uniform [8] and fluctuating [23] node densities within the simulation area, as well as decrease the average nodal speed over time [26, 27].

Recently, there have been a number of mobility models proposed as alternatives to the standard random waypoint model (e.g., [4, 7, 10, 12, 13, 14, 15, 27]). In general, the goal of these models is to increase the realism of random waypoint in particular settings. For example, [12] proposes a group mobility model called Reference Point Group Mobility (RPGM). In this model, node velocity vectors are calculated as the sum of a group center velocity vector and a random motion vector for each node. Such a model could be used for group movement in disaster recovery where teams work and move together throughout a disaster area, or to model movement of convention attendees moving from room to room between project demonstrations. In [15], three movement models were proposed: conference, event coverage, and disaster area[1]. New to these models was the incorporation of obstacles through which neither users nor radio signals could pass. [7] proposes a mobility model with smooth, rather than sudden, changes in both velocity and direction. A "Mobility Vector Model" is proposed in [13] which is similar to the model in [12]. However, [13] introduces an "acceleration factor" which, when properly chosen, also generates smoother node trajectories. "Freeway" and "Manhattan" mobility models are proposed in [4]. These models restrict movement to paths defined prior to simulation. In addition, node movement depends upon the node's previous velocity and the velocity of nearby nodes considered to be in the same lane.

Most recently, [27] shows that any mobility model in which speed and destination are chosen independently suffers from average speed decay. They propose a framework

[1] A common theme, but with many differing models.

in which any given mobility model can be transformed to eliminate variations in average node speed as simulation time progresses. Finally, [14] proposes a mobility model that incorporates both obstacles and paths. Node movement is not only restricted to paths, but follows a shortest path route to the node's randomly chosen destination. In addition, node transmissions do not pass through obstacles. This framework was used to model student movement between buildings on a college campus.

Each of the above mentioned works test various ad-hoc routing protocols using their mobility models. Although protocol performance results vary from model to model, results indicate that a few key characteristics of the mobility models play major roles in the effectiveness of the ad-hoc protocol under study. For example, average link lifetime and link change frequency which result from the model are key characteristics which affect the performance of the protocol.

Until now, these synthetic models have been the only means of specifying user movement for ad hoc routing protocol testing. Our study provides much needed data to compliment, as well as, validate the synthetic models in use today.

3 Trace Methodology

In this section, we describe the procedure used to collect the trace data as well as definitions and data analysis methods.

3.1 Data Collection

We collected trace data from approximately 275 freshmen PDA users[2] for an 11 week period between September 22, 2002 and December 8, 2002. The freshmen were the initial students in a new college (anonymized as "New College") on our university campus. Our colleges have a particular academic theme and define graduation requirements for students in the college, but are independent of major; colleges have students representing all majors on campus. Each PDA was equipped with a Symbol Wireless Networker 802.11b Compact Flash card. The trace PDAs consisted of 97 Jornada 548s and 185 Jornada 568s running the Windows Pocket PC 2000 and 2002 operating systems, respectively[3]. We identify users according to their registered wireless card MAC address, and assume that there is a fixed one-to-one mapping between users and wireless cards. The mapping is anonymous; we have no mapping of MAC address to user names.

The University campus has extensive 802.11b coverage in which students can roam. The students in our trace resided at the "Roosevelt" College housing facility – an approximately 130m x 130m square area near the southeast corner of the University campus with complete 802.11b coverage. Overall wireless activity was extensive: students associated with over 400 unique APs in our trace.

We developed a background data collection tool called WTD (Wireless Topology Discovery) and installed the tool on each PDA prior to distribution. For our trace, WTD periodically recorded the following information:

- Access point (AP) signal strength (for each AP detected)
- AP MAC address (for each AP detected)
- Current AP association
- WTD program version number
- Device type (Jornada 548 or 568)
- Power state (on AC or battery power)

Note that WTD recorded the AP signal strength and MAC address for *all* APs that it could sense across all frequencies for each time interval, not just the AP the wireless card was associated with at the time. Recording all APs provides much richer topology information than just the associated AP.

As a trade-off between the granularity of samples and the resource and power overhead of collecting the data, we used a sampling period of 20 seconds. Once either the local data file reached a critical size or a maximum data file age was reached, WTD contacted our server to upload its sample collection in bulk.

A practical feature of the software is its ability to periodically check our database for new releases and bug fixes. Once detected, a new version is automatically downloaded to the device and re-launched. This feature makes it easy to add functionality to the data collection software or adapt to unexpected problems.

3.2 Preliminary Data Analysis

A key aspect of our data analysis involved determining when and where a sample was taken, in addition to defining a user session and its duration. This information enabled us to reconstruct the necessary topology information for later experiments, as well as describe device usage patterns.

Since all PDAs maintained their own clocks, we needed a method for resolving the time in which a sample was taken. Our approach compared the sample upload time according to the PDA with the sample upload time according to our database server. The clock skew between the PDA and our server was then calculated and added to the sample timestamp recorded when the PDA took the sample.

We determined user geographic locations using the locator software from another project [3]. This software uses trilateration based on recorded AP signal strengths (for all APs detected in the sample) along with empirical corrections to calculate user locations on campus.

One difficulty encountered during initial data analysis was to determine when a session began and ended. We observed that many sessions occurred on the edge of AP

[2] There were 3 development PDAs which uploaded data during the trace period.

[3] The slight discrepancy in the number of total users and number of total trace PDAs is due to a few network cards – likely developers' cards – moving between a 548 and 568.

detection. This caused a single session to appear to be several very short sessions as the AP signal fluctuated between detected and undetected. As a result, we used a simple heuristic requiring sessions to have a minimum length of one minute and 1.5 minutes between sessions.

Finally, we observed a steady decline in user population over the trace period. Though not unforeseen, the dropout rate was higher than expected. Although it would be impossible to say what the exact cause was for the decline, we attribute the decline to two contributing factors. First, rather than purchasing the PDAs for an explicit need (e.g., as with laptops), students were simply given the devices for being a New College freshman independent of whether students wanted the device. Because these users are from a variety of different majors and interests, it is likely that some opted to stop using the device after an initial trial. We have also heard of this device abandonment trend in other studies in which devices were simply given to the user population [2]. This trend raises the interesting question of the extent to which students, usually considered early adopters, actually have a strong demand for current handheld PDA technology.

A second contributing factor was the fact that a complete depletion of battery resulting in a device hard reset occurred more often than initially expected. Since our tracing module resided in soft state on the PDAs, a hard reset permanently removed all pre-installed software (including WTD). This situation may have also contributed to students abandoning their devices. After losing all their personal data and settings, some users may have opted to cease using their PDA out of frustration. Since the Jornada 568s possessed a small, replaceable backup battery, this problem was more prevalent in the Jornada 548s, where only 23 devices (22 percent of 548s) recorded data the final week of the trace, as opposed to 64 Jornada 568s (33 percent of 568s). Due to complicated issues involved with the PDA distribution, there was no means for recalling and/or restoring devices once they were in the student's possession. These issues were, unfortunately, out of our control.

(a)

(b)

Figure 1: User locations across the University campus at noon (a) and 1PM (b) on September 24, 2002 (dark circled 'x'). The dense areas at the lower right are Roosevelt College student housing buildings (where the New College freshmen were temporarily housed).

4 User Behavior

In this section, we study user behavior based upon our trace data. First, we characterize overall activity of our PDA users, focusing on daily usage patterns. Next, we discuss user mobility focusing on movement among access points on campus. Finally, we discuss user access to the wireless network, analyzing user and AP session counts and length. Where appropriate, we compare and contrast our results with previous trace studies. For most measurements, we provide cumulative distribution functions and explicitly quantify the median, 80th percentile, and 90th percentile statistics.

4.1 Activity

We start by illustrating the location of wireless activity on campus, and then quantify overall user PDA activity during our trace period. Figure 1 shows snapshots of user locations taken at noon and 1PM on September 24. The majority of activity takes place at the student housing facility in the lower right hand corner of the map. This was the case throughout the duration of the trace. Other locations of moderate activity were lecture halls where a large number of undergraduate courses are taught.

In terms of overall user PDA activity, Figure 2 shows the number of active users per hour across the 11 weeks of the trace, and Figure 3 shows the number of active users per hour for just the first full week of the trace (the most active week). These graphs show a number of high-level characteristics of the user population in our trace. First, device usage and network access follow regular diurnal patterns, with peak usage typically between 1pm–2pm and minimum usage between 5am–6am. Activity during the week is significantly higher than the weekend: an average

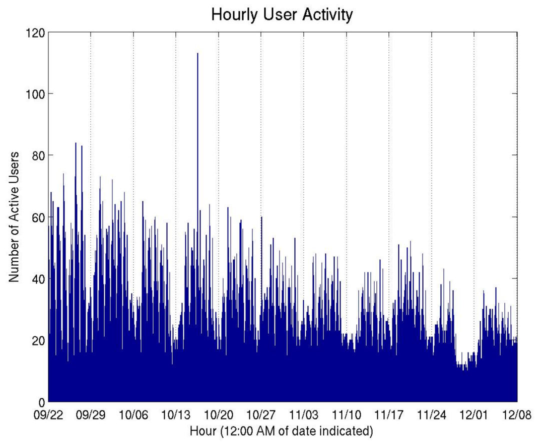

Figure 2: Number of active users each hour over study duration. The shallow decline indicates user dropout, a result of device abandonment or loss of trace module.

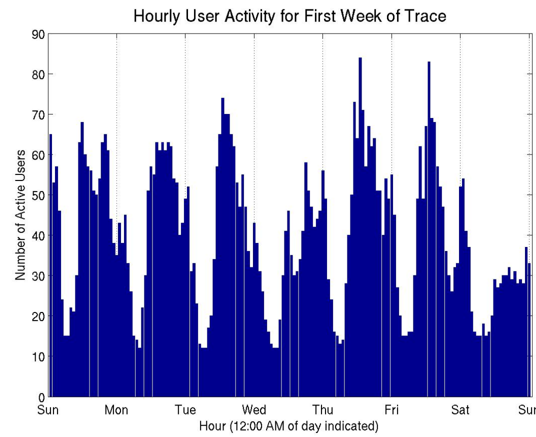

Figure 3: Number of active users each hour for first week of trace.

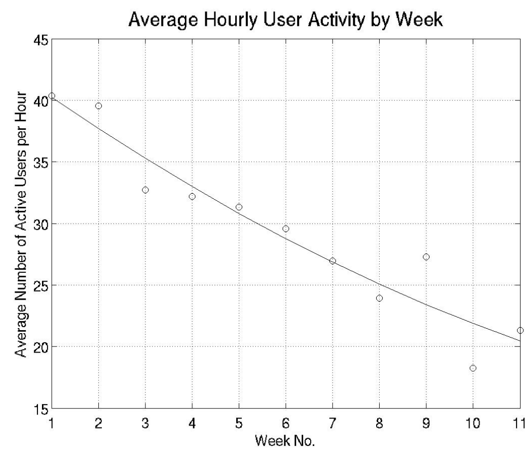

Figure 4: Average number of active users per hour for each week of trace. Week 10 was Thanksgiving break.

of 33 percent more users per hour on the weekdays than during the weekends. We also find that student usage of PDAs is relatively bursty — likely reflecting the ease with which users can carry and activate their PDAs as well as the constraints of using PDAs for extended periods of time.

The extended time span of low usage around 11/30 corresponds to the Thanksgiving weekend. The peak on 10/16 corresponds to a day in which many users participated in a psychological experiment for a New College class.

Second, as mentioned in Section 3.2, the graph shows a clear decline in recorded user activity over the trace period. Again, it it not clear what the contributions of user dropout and loss of trace module are to this decline. Figure 4 shows the average number of users per hour for each week of the trace. A second degree polynomial has been fitted to the data in a least squares sense to highlight the user activity decline trend. The trace starts with an average of 40 users per hour for the first week and ends with 21 – a loss of nearly 2 active users per hour each week.

Third, there are typically between 10 and 15 users active even in the early hours of the morning. For the most part, we can assume that these are PDAs left on over night and sitting in their cradles. This assumption is based on the typically long session durations and device immobility observed among this set of PDAs.

The Dartmouth [17] and IBM [6] traces also show clear diurnal patterns of usage. The decline in user activity only appears in our trace due to the reasons discussed above.

4.2 Access

Next we study user access to the wireless network in terms of session counts and session durations.

4.2.1 Session counts

Figure 5 shows the CDF of the number of user sessions. Note that 50 percent of the users initiated more than 77 sessions over the trace period. This means that the median user initiated an average of one session per day over the trace period. Clearly, many of the students were not relying upon their device in their daily activities. Still, 20 percent of the users did initiate over 231 sessions (roughly three times per day), and 10 percent initiated over 335 (roughly four times per day).

To understand user activity on a day to day basis, Figure 6 presents a CDF of the number of days the users actually turned on their PDAs. From Figure 6 we can see that half the users turned on their devices less than 21 days during the trace. This is lower than the median number of days from the Dartmouth study [17] in which laptop users connected a median of 28 days over their 77 day trace. Furthermore, the distribution of number of days in which a user used their device was nearly uniform in the Dartmouth study. Clearly this is not the case from the inset graph of Figure 6. There, we see that there were 20 users who only used their PDAs one day during our trace period. The number of users for each number of active days tends to drop from there. This is further indication of user dropout during our trace period.

Figures 5 and 6 indicate that some users did find their PDA useful. Twenty percent used their PDAs more than 46 days (60 percent of the 77 days), and 10 percent used them more than 58 days (75 percent of the 77 days). And there were a few 'die hard' users who used their PDAs

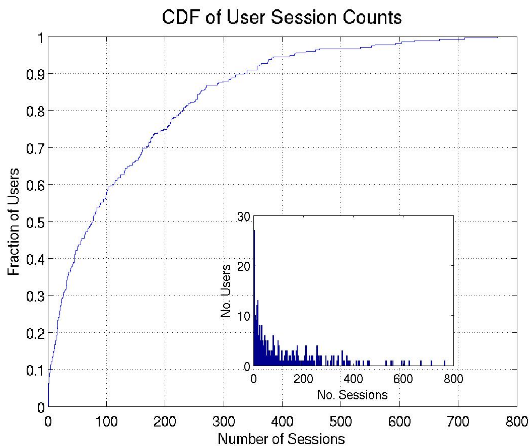

Figure 5: Cumulative distribution function of session counts. Inset is histogram of the same using 200 bins. Half the users initiated more than 77 sessions over the trace. A few extreme users initiated over 700 sessions!

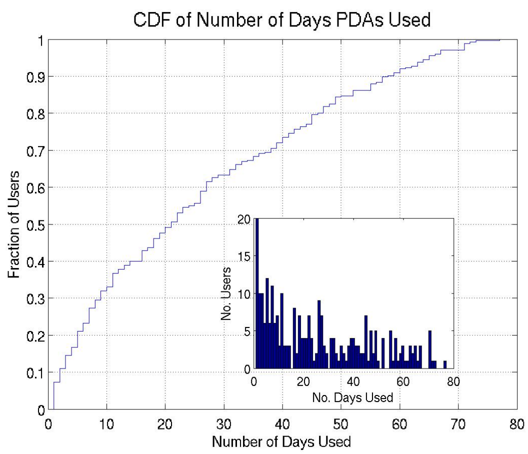

Figure 6: Cumulative distribution function of the number of days the users powered on their devices during the trace. Inset is histogram of the same using 77 bins. Half the users powered up their PDA more than 21 days during the trace. A few users used their PDA nearly every day.

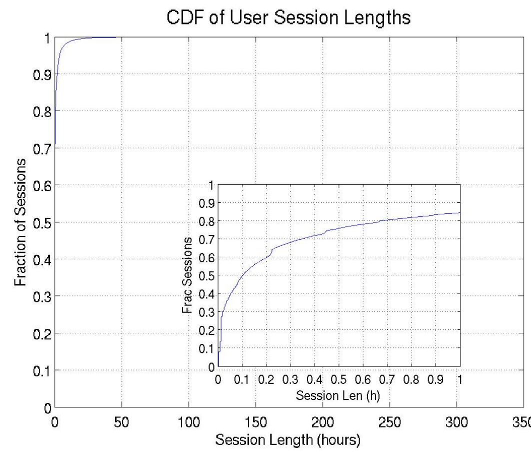

Figure 7: Cumulative distribution function of session lengths. Inset shows data for sessions less than one hour in duration. Nearly half of all sessions were less than 6 minutes. Extremely long sessions are likely PDAs left running in their cradles.

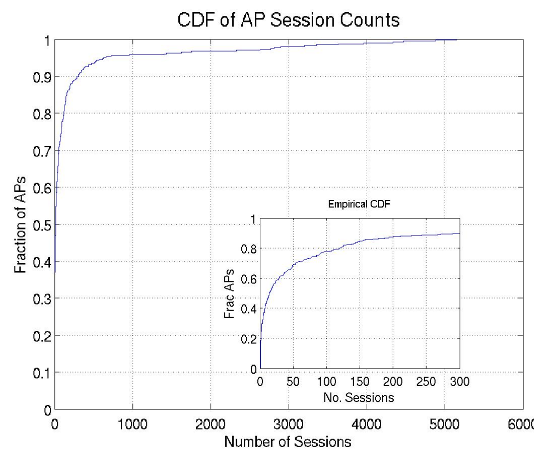

Figure 8: Cumulative distribution function of AP session counts. Inset shows data for APs with fewer than 300 sessions.

nearly every day.

4.2.2 Session durations

Figure 7 shows the heavy tailed distribution of user session lengths. A user session length is the time duration in which a user PDA is powered on and able to detect nearby access points. The extremely long sessions are likely PDAs left running in their cradles. One session actually lasted 333 hours – nearly two weeks! The median session duration, or the time a PDA remains connected to the wireless network, was only 6.25 minutes compared to 16.6 minutes for laptop users [17]. Further, only 16 percent of all sessions are at least one hour, compared to 29 percent for laptop users. However, for both PDA and laptop users, 27 percent of sessions were under one minute. Still, a substantial number of sessions were long: 20 percent of user sessions were 41 minutes or longer, and 10 percent were 121 minutes or longer.

From the perspective of the access point, Figure 8 contains the distribution of number of AP sessions for the set of APs recorded in our trace. An AP session is the time duration in which a user PDA associates with an access point. This heavy tailed distribution makes it difficult to see that there was a median of 14 sessions at an AP. This heavy tail indicates that a few APs were used very frequently and held a large portion of the total number of AP sessions, while many of the APs from our trace held only a few, infrequent sessions. Around 20 percent of the APs held 120 or more sessions, 10 percent held 304 or more sessions, and just 10 different APs held 50 percent of all AP sessions in our trace. Not surprisingly, these AP were located at the student housing facility, the hotspot of PDA activity in our trace.

Figure 9 shows the CDF of AP session lengths. Note again that this is again a very heavy tailed distribution. Interestingly, AP sessions were a median of only 1.85 minutes

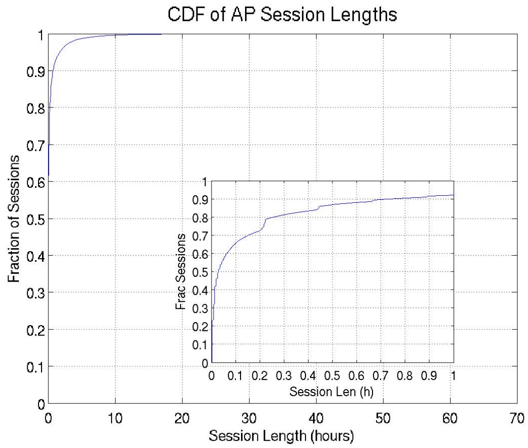

Figure 9: Cumulative distribution function of AP session durations. Inset shows data sessions less than one hour.

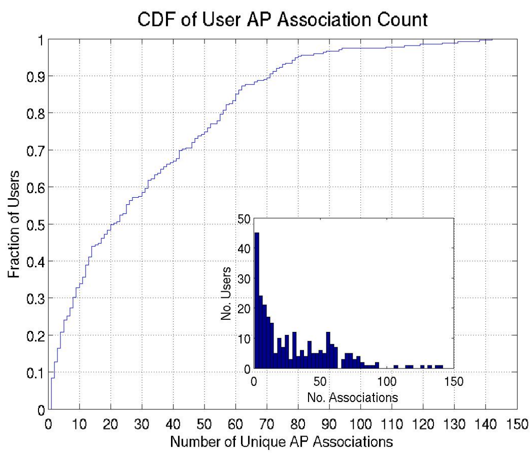

Figure 10: Number of access points with which users associate over the entire trace. Inset is histogram of the same with 50 bins.

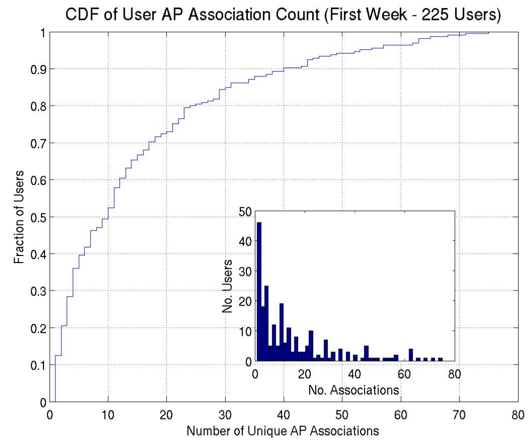

Figure 11: Number of access points with which users associate during the first week. Inset is histogram of the same with 50 bins.

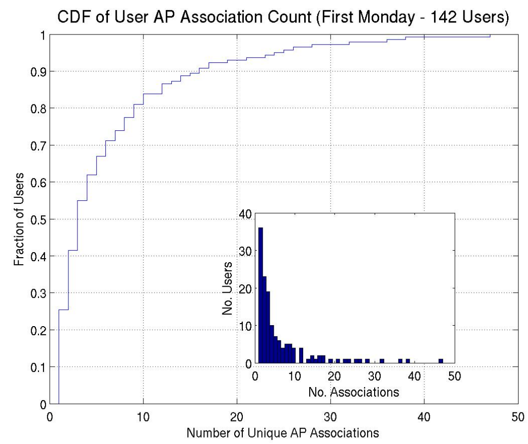

Figure 12: Number of access points with which users associate on first Monday of the trace (9/23/2002). Inset is histogram of the same with 50 bins.

in length. This indicates that there was significant roaming during user sessions, where the median was 6.25 minutes in length. This is not surprising considering that multiple APs were detected in most samples, and our cards periodically scan for the AP with the highest signal strength. If an AP with a higher signal strength is detected, a reassociation will occur unless the card is explicitly instructed not to do so. Since signal strengths can vary significantly from one moment to the next at the same location, even stationary users experience reassociation when their devices detect two or more APs with similar signal strengths. As further evidence, the maximum AP session length was only 62 hours as compared to a 333 hour user session length.

4.3 Mobility

In this section, we characterize overall user mobility from two perspectives: (1) the distribution of the number of access points with which users associate and the number which they detect, and (2) the distribution of number of users which associate with particular access points. Intuitively, the first captures how widely users in the trace roam across campus while using their PDAs. The second captures how concentrated this roaming is.

We find that students are relatively mobile and use their PDAs in many locations. Figure 10 shows the cumulative distribution of the number of unique access points that users associate with throughout the trace (the inset graph shows the raw distribution histogram with 50 bins). From the graph, we see that 50 percent of the users visit 21 APs or more, 20 percent visit 56 APs or more, and 10 percent visit 71 APs or more during the entire trace. In extreme cases, some students associated with over 130 APs. Compared to the laptop users in the Dartmouth study [17], we find that the typical wireless PDA user is over twice as mobile as the typical laptop user in terms of associated access points. This indicates that PDA users tend to operate in a larger number of locations than their laptop counterparts.

Figures 11 and 12 show CDF plots of the user AP association counts for just the first week and first Monday (9/23/2002) of the trace, respectively. These graphs provide a finer grained view of user activity, and demonstrate that users are quite active. Within the first week, 225 users

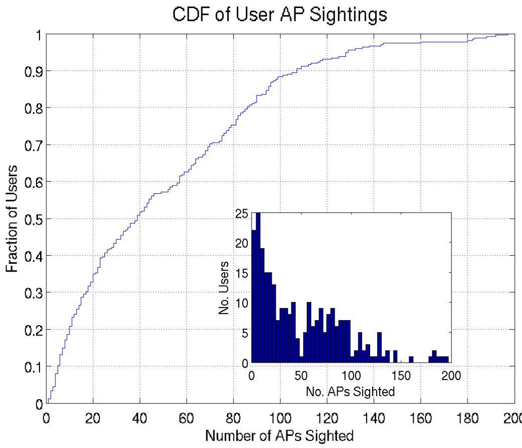

Figure 13: Number of access points which a users detects over entire trace. Inset is histogram of the same with 50 bins.

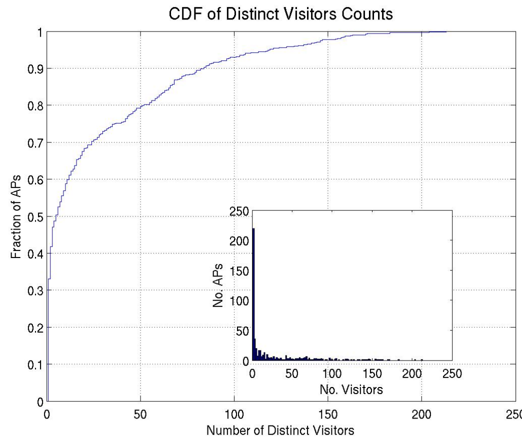

Figure 15: Distribution of number of distinct visitors to an AP. Inset is histogram of the same with 125 bins.

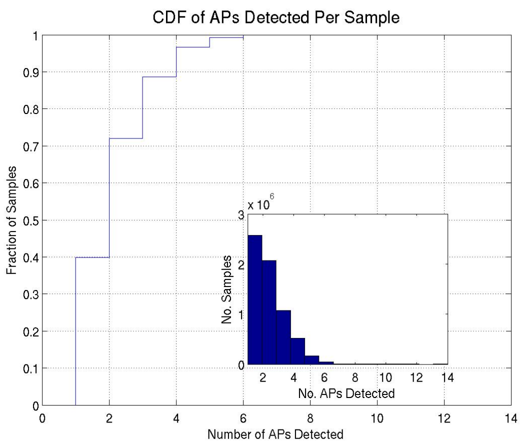

Figure 14: CDF of APs detected in a sample. Most samples detected multiple nearby access points. Inset is histogram of the same.

had associated with an access point, half of which had already associated with 10 or more different APs. Note from Figure 12 that 142 users were active on the first Monday. A few extreme users associated with over 30 APs on that day alone.

Figure 13 is similar to Figure 10, except that it shows the cumulative distribution of the number of access points that users *could have* associated with throughout the trace, i.e., the set of APs detected, but not necessarily associated with during the trace. Recall that our wireless monitor records all access points sensed by the 802.11 wireless card, not just the AP with which the card is associated. Comparing the graphs in Figures 10 and 13, we find that users see many more access points than they associate with: 50 percent of the users see 39 APs or more, 20 percent see 86 APs or more, and 10 percent see 107 APs or more.

Overall, users see an average of 2 access points at a location. Figure 14 shows the distribution of number of APs detected per sample. Notice that 60 percent of all samples detected multiple access points and 10 percent of all samples saw 4 or more APs. Such overlapping wireless network deployments provides useful opportunities for location determination and load balancing.

Looking at user mobility from the opposite perspective, Figure 15 shows the cumulative distribution of distinct users seen across the access points for the entire trace (the inset graph shows the raw distribution histogram using 125 bins). Many access points only see a few users: 50 percent of the APs see 5 users or less. However, a significant fraction of access points see a large number of users: 20 percent see 51 or more users, and 10 percent see 84 or more.

5 Trace-Based Mobility Models

Evaluations of mobile systems frequently use synthetic models of user mobility patterns, such as the popular random waypoint model [9] and its numerous refinements (e.g., [10, 14, 26]), to derive wireless network topologies that change due to user mobility. Synthetic mobility models have a number of advantages, including the ability to arbitrarily scale the various parameters of the model to exercise the system across a wide range of parameter scenarios.

However, synthetic mobility models have the obvious limitation that they are synthetic. Without basing its mobility patterns on realistic, observed object mobility, the extent to which a synthetic model is representative of how a mobile system would behave and perform in the real world remains unclear. As a result, we argue that evaluations of mobile systems need to also include the use of mobility models derived from real, measured mobility patterns.

To complement synthetic models, we propose two new models of network topologies that incorporate user mobility patterns from our traces. These models represent the indoor and outdoor mobility of people with handheld PDAs. Consequently, they only represent a particular mobility scenario, but one that is common in applications of mobile systems. Since these models are derived from traces, we are in the process of making these models available for download and use by other researchers.

5.1 Evolutionary topology model

The *evolutionary topology model* is a constructive model based upon the mobility of the users in our trace as well as the wireless connectivity of the 802.11b devices they used. We call it "evolutionary" because we derive the network topology from the network proximity of the users, and the topology naturally evolves over time as users move about as well as when they join and leave the network.

A compelling feature of the model is that it incorporates the wireless connectivity and propagation characteristics of 802.11. Consequently, it naturally captures and models the range, interference, and obstruction properties of 802.11 that are challenging to realistically model using analytic approaches. For example, rather than determining whether two nodes can communicate with each other based on a range parameter, the evolutionary topology model bases connectivity on the ability of users' 802.11 devices to communicate in a particular location at a particular time. Our experience is that this feature is critical for improving the realism of the network topology (see Section 5.1 below). As a result, we consider this model to be particularly realistic for wireless mobility scenarios using 802.11 devices, which is the most common scenario evaluated using wireless simulators [1, 28]. Note that this model intimately ties together user mobility and network topology, trading off realism with generality. Our second model below removes the dependency on the use of 802.11 wireless, albeit at some expense of realism.

The evolutionary topology model represents connectivity among users solely based on observed network proximity. For each time slot, we create a node in the topology for each active user in our trace, and create edges to connect nodes if users' wireless devices could reasonably communicate with each other at that location and time in the trace. Ideally, the topology would have an edge between two nodes if the users' wireless devices sensed each other in the trace. Recall, though, that in our trace we recorded all of the access points that each user could sense (not just the AP the user was associated with); we had planned to detect *all* 802.11b devices, both user and AP, but device limitations prevented user-user detection.

As a result, we approximate connectivity between two users by creating an edge between two nodes if the intersection of the set of APs sensed by their PDAs is nonempty, and remove the edge if the intersection of APs becomes empty again. In other words, we create an edge between two nodes representing users in the topology if those users could sense at least one AP in common during that time slot. This approximation will have errors, since two users sensing the same AP does not necessarily mean that they can potentially communicate directly with each other. However, in Section 5.2 below we argue that this error is acceptably small compared to the use of current radio models, and that the resulting topology is considerably more realistic as a result. In effect, we are generating a topology very near to the actual topology. We call this the "nearby" topology.

Logically, we recreate this nearby topology for each time slot in the trace. As a result, the network topology naturally evolves to model the mobility patterns of users and radio propagation characteristics over time. Nodes and edges appear and disappear based upon PDA on/off events and user movements observed in our trace. In practice, we only explicitly record the node and edge changes in the topology over time for efficient representation.

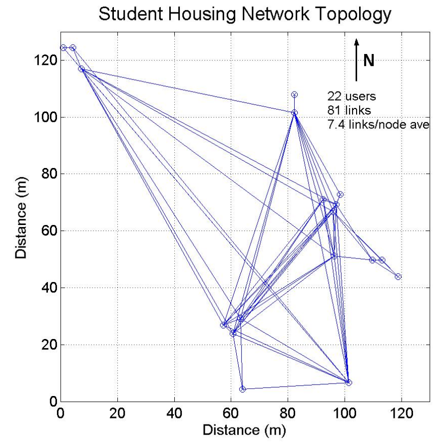

Figure 16: Snapshot of the evolutionary topology model at noon on Monday, September 23, 2002.

Metric	Min	Max	Avg	Median
Nodes	6	48	18	18
Links	2	532	41	35
Degree	1	22	4	4

Table 1: High-level characteristics of the evolutionary topology model for the first week of the trace.

Figure 16 shows an example timeslice of our evolutionary topology model at noon on Monday, September 23, 2002, in the Roosevelt College part of campus. The nodes in the nearby topology represent 22 users, with each node having an average degree of 7.4 for a total of 81 edges. The evolutionary model also emphasizes the limitation of the popular two-ray ground reflection model for radio propagation in an environment containing obstacles. This radio propagation model is often used in wireless networking simulations to determine when nodes are within communication range by calculating the strength of the received signal(s). Figure 16 shows that the evolutionary model accounts for obstacles to generate a more realistic connectivity graph than the two-ray ground model. For this geographic setting, the two-ray ground model would create a completely connected graph under typical radio settings.

Overall, from our trace we found that the evolutionary topology model results in interesting topologies only within the Roosevelt College location of campus. This part of campus is home to users with the deployed PDAs. Once these users move throughout campus, they create numerous islands of disconnected graphs; the campus waypoint model described next better captures and takes advantage of this extended, outdoor mobility. Focusing on the mobility patterns and topology of users within the Roosevelt College, Table 1 shows the high-level characteristics of the resulting evolutionary topology for the first week of our trace. In Section 6, we use this evolutionary topology model to evaluate the performance of popular ad-hoc routing algorithms.

5.2 Campus waypoint model

The *campus waypoint model* serves as a trace-based analog to the random waypoint model. In this model, we associate users with geographic locations on campus and model their mobility vectors and potential interactions as they access the wireless network over time. However, rather than choosing user locations, speeds, and directions using random or other synthetic distributions and models, we instead use the access and mobility patterns of users in our trace. In this section we describe the model and compare its characteristics to common synthetic models.

In the campus waypoint model, the campus geography with 802.11 wireless coverage serves as the geographic region in which nodes roam. This region is roughly a 1400m by 1700m rectangle. For each time step in the model, we determine user location based upon the set of APs that each user senses in that time slot and the known geographic locations of those APs relative to the campus geography. For each user, we estimate user location via trilateration among the locations of sensed APs.

We model user mobility over time based upon (1) the evolving set of sensed APs, and (2) the disappearance and reappearance of users at different AP locations on campus assuming reasonable velocities. A user always appears in the model when connected to the wireless network. For each time slot, we update user location whenever the set of sensed APs changes. The location of a continuously connected user over time determines the mobility rate, direction, and pause time of that user. The mobility patterns from these kinds of users generally represent roaming within a building or within building clusters.

A user *may* appear in the model depending upon its wandering status. A user *wanders* when it disassociates from the network and reassociates at a different location; these mobility patterns generally represent outdoor roaming among buildings across campus. We model wandering direction simply as the vector between the locations of disassociation and reassociation. We model wandering speed by computing the geographic distance between locations and dividing by the time between associations. To ignore situations where users wander far off a straight line, we impose a minimum speed of 0.5 mph on wandering speeds along a straight line. Wandering users can optionally appear in the model depending upon the scenario simulated. Although in our trace users disassociated from the network, ad-hoc routing simulations, for example, can still take advantage of wandering users to model scenarios that assume users keep their devices active while wandering.

When comparing the mobility characteristics in our campus waypoint model to those used in typical synthetic models, we make three significant observations. First, unlike node mobility in typical parameterizations of synthetic simulations, we find that only a small percentage of users – 11% on average for the first, most active week of trace – are actually in motion at any one time. In contrast, default parameterizations for synthetic scenarios, independent of the complexity of the mobility model, result in most nodes being mobile. Second, given that our users are walking, users move in the campus waypoint model at an average speed of 2.2 mph, or roughly a meter per second on average. For comparison, the default node speed for ad-hoc routing in ns2 wireless scenarios draws from a uniform distribution between 0–20 meters per second. Lastly, users appear and disappear from the network. This behavior, absent in most documented simulations, can and does have drastic effects on network topology and connectivity. Perhaps the reason for the absence of this behavior in most simulations is in part due to the fact that node on/off events are not currently implemented in the ns2. We have, however, extended ns2 to model these events for the protocol evaluations of Section 6.

From our observations, we conclude that default parameterizations of mobility models commonly used to evaluate mobile systems represent very aggressive mobility scenarios. Such scenarios indeed stress the ability of systems to deal with mobility, but at the expense of realism. For scenarios where user movement is limited to walking, at least, our results suggest that commonly parameterized synthetic mobility models are too aggressive. We explore this issue in more detail in the next section.

6 Ad-Hoc Routing Evaluation

Groups of users with handheld PDAs have often been used as a motivating setting for ad-hoc networking. In this section, we use the evolutionary topology model to study popular ad-hoc routing algorithms on the network topologies formed among users of modern wireless PDAs. Our goal is to be indicative of the performance of these protocols in a realistic mobility scenario.

We used ns2 version 2.1b8 to simulate and compare the performance of the DSR, DSDV, and AODV ad-hoc routing protocols. Our study focused on the Roosevelt College student housing area from noon to 1pm on Monday, September 23, 2002. Recall that this area consists of eight buildings laid out in approximately a 130m x 130m square area (lower right corner of Figures 1(a) and 1(b)). This area and time corresponds to a scenario where users establish an ad-hoc network among a cluster of nearby buildings, with mobility both inside individual buildings as well as among buildings. Though node numbers fluctuated, there were at least 30 PDAs active during the hour of simulated communication.

Using the default ns2 wireless constant bit rate (CBR) traffic of 4 packets/sec, 512 byte packets, we ran simulations with a random 10, 25, and 50 percent of the nodes communicating at any one time. The simulation results indicate that all three protocols were roughly equivalent in performance; the packet delivery ratios for all protocols were within 2–4% of each other across the various workload scenarios. Even with nodes appearing and disappearing, all three protocols were able to quickly adapt to find new routes between senders and receivers, and user mobility was not a key factor in the performance of the protocols.

These results reflect the low rate of change of the network topology: for the topology we simulated, the topology change at the rate of only 11 link changes per minute. Students moved around in the network during the hour, but the rate of movement was low compared to the ability of the routing protocols to quickly adapt. Although dependent on our trace scenario, these results again suggest that many ad-hoc routing simulations are overly aggressive in their parameter choices for speed (up to 20 m/sec) and mobility. Such choices emphasize topological change and, consequently, perhaps place too much emphasis on an *uncommon* case. The basic routing protocols perform well in this scenario and do not appear to require substantial additional optimization – a topic that has been the source of considerable effort in ad-hoc routing research.

We do not claim that these results are representative of all interesting user mobility scenarios. However, we argue (1) that the results are more indicative of protocol behavior in a compelling deployment scenario than synthetic models with aggressive mobility, and (2) they underscore the importance of linking models to realistic scenarios and parameterizations.

7 Conclusion

In this paper, we study the access and mobility characteristics of an 11-week trace of wireless PDA users on a university campus. Our study has three goals.

First, we characterized the high-level mobility and access patterns of handheld PDA users. Compared to previous studies focused on laptop users, we found a much wider variation in wireless network usage among PDA users. Furthermore, we found that the PDA users were about twice as mobile as laptop users in terms of the number of access points they associated with in the same time period.

Second, we develop two wireless network topology models for use in wireless mobility studies: an *evolutionary topology model* based upon the mobility of the users in our trace as well as the wireless connectivity of the 802.11b devices they used, and a *campus waypoint model* that serves as a trace-based analog to the random waypoint model. We compared the characteristics of these trace-based models based upon realistic user mobility patterns and wireless connectivity with default parameterizations of synthetic models. The evolutionary topology model overcomes the limitations of the popular two-ray ground reflection model for radio propagation in an environment containing obstacles. For our geographic setting, the two-ray ground model would create a completely connected graph under typical radio settings where our model reflects a much more realistic network topology. The campus waypoint model contrasts sharply with typical synthetic models in terms of percentage of nodes mobile at a time as well as average node speed in a campus scenario. The typical parameterizations of synthetic models are significantly more aggressive than what we found in practice.

Finally, we use the evolutionary network topology model as a case study to evaluate ad-hoc routing algorithms in a realistic setting. The simulation results indicate that all three protocols were roughly equivalent in performance, and that user mobility was not a key factor in their performance. The rate of topology change due to user mobility from the trace was relatively low compared with the ability of the protocols to adapt to those changes. This result underscores the importance of linking models to realistic scenarios and parameterizations.

References

[1] The network simulator - ns-2. http://www.isi.edu/nsnam/ns/, 2003.

[2] Anonymous. Personal Communication.

[3] Anonymous. Employing user feedback for fast, accurate, low-maintenance geolocationing. Technical report, Anonymous for review, 2003.

[4] F. Bai, N. Sadagopan, and A. Helmy. Important: a framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks. In *IEEE INFOCOM*, 2003.

[5] A. Balachandran, G. Voelker, P. Bahl, and P. Rangan. Characterizing user behavior and network performance in a public wireless lan. In *ACM Sigmetrics*, 2002.

[6] M. Balazinska and P. Castro. Characterizing mobility and network usage in a corporate wireless local-area network. In *ACM MobiSys*, 2003.

[7] C. Bettstetter. Smooth is better than sharp: A random mobility model for simulation of wireless networks. In *ACM MSWiM*, 2001.

[8] C. Bettstetter, H. Hartenstein, and X. Prez-Costa. Stochastic properties of the random waypoint mobility model: epoch length, direction distribution, and cell change rate. In *ACM MSwiM*, 2002.

[9] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *ACM MobiCom*, pages 85–97, 1998.

[10] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, 2(5):483–502, 2002.

[11] D. Eckardt and P. Steenkiste. Measurement and analysis of the error characteristics of an in-building wireless network. In *ACM SIGCOMM*, 1996.

[12] X. Hong, M. Gerla, G. Pei, and C. Chiang. A group mobility model for ad hoc wireless networks. In *ACM/IEEE MSWiM*, pages 53–60, 1999.

[13] X. Hong, T. J. Kwon, M. Gerla, D. L. Gu, and G. Pei. A mobility framework for ad hoc wireless networks. *Lecture Notes in Computer Science*, 1987, 2001.

[14] A. Jardosh, E. Belding-Royer, K. Almeroth, and S. Suri. Towards realistic mobility models for mobile ad hoc networks. In *ACM MobiCom*, pages 217–229, 2003.

[15] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark. Scenario-based performance analysis of routing protocols for mobile ad-hoc networks. In *ACM MobiCom*, pages 195–206, 1999.

[16] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In Imielinski and Korth, editors, *ACM MobiCom*, volume 353. Kluwer Academic Publishers, 1996.

[17] D. Kotz and K. Essien. Analysis of a campus-wide wireless network. In *ACM MobiCom*, 2002.

[18] K. Lai, M. Roussopoulos, D. Tang, X. Zhao, and M. Baker. Experiences with a mobile testbed. In *Proceedings of The Second International Conference on Worldwide Computing and its Applications (WWCA'98)*, March 1998.

[19] D. A. Maltz, J. Broch, and D. B. Johnson. Experiences designing and building a multi-hop wireless ad hoc network testbed. Technical Report CMU-CS-99-116, School of Computer Science, Carnegie Mellon University, March 1999.

[20] V. D. Park and M. S. Corson. A highly adaptive distributed routing algorithm for mobile wireless networks. In *IEEE INFOCOM (3)*, pages 1405–1413, 1997.

[21] C. Perkins and P. Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *ACM SIGCOMM*, pages 234–244, 1994.

[22] C. E. Perkins and E. M. Royer. Ad hoc on-demand distance vector routing. In *IEEE WMCSA*, pages 90–100, 1999.

[23] E. Royer, P. Melliar-Smith, and L. Moser. An analysis of the optimum node density for ad hoc mobile networks. In *IEEE International Conference on Communication*, pages 857–861, 2001.

[24] D. Tang and M. Baker. Analysis of a metropolitan-area wireless network. In *ACM MobiCom*, pages 13–23, 1999.

[25] D. Tang and M. Baker. Analysis of a local-area wireless network. In *ACM MobiCom*, pages 1–10, 2000.

[26] J. Yoon, M. Liu, and B. Noble. Random waypoint considered harmful. In *IEEE INFOCOM*, pages 1312–1321, 2003.

[27] J. Yoon, M. Liu, and B. Noble. Sound mobility models. In *ACM MobiCom*, pages 205–216, 2003.

[28] X. Zeng, R. Bagrodia, and M. Gerla. Glomosim: A library for parallel simulation of large-scale wireless networks. In *Workshop on Parallel and Distributed Simulation*, pages 154–161, 1998.